

Mémoire de titres et travaux en vue de l'obtention de  
l'Habilitation à Diriger des Recherches

*par*  
Caroline Smet-Nocca

**Régulation structurale et fonctionnelle de la protéine neuronale Tau  
par les modifications post-traductionnelles**

*Implications physiopathologiques de la phosphorylation, l'acétylation et  
l'O-β-N-acétylglucosaminylation de Tau*

Jury

Pr Fred Van Leuven  
Pr José Martins  
Dr Françoise Jacob-Dubuisson  
Dr Florence Pinet  
Dr Marie-Laure Caillet-Boudin  
Dr Guy Lippens  
Pr Tony Lefebvre

## Sommaire

---

<b>CURRICULUM VITAE</b>	<b>5</b>
<b>ABREVIATIONS</b>	<b>11</b>
<b>INTRODUCTION GENERALE : PARCOURS SCIENTIFIQUE ET THEMATIQUES DE RECHERCHE</b>	<b>12</b>
<b>CHAPITRE 1 : LA PROTEINE NEURONALE TAU ET SON ETUDE PAR RESONANCE MAGNETIQUE NUCLEAIRE</b>	<b>15</b>
1.1 LA PROTEINE TAU, UNE PROTEINE ASSOCIEE AUX MICROTUBULES IMPLIQUEE DANS LA MALADIE D'ALZHEIMER	15
1.1.1. <i>Isoformes et domaines de la protéine Tau dans la régulation fonctionnelle de la liaison aux microtubules</i>	16
1.1.2. <i>Pathologie de la protéine Tau dans la maladie d'Alzheimer</i>	18
1.1.3. <i>Phosphorylation de la protéine Tau : équilibre entre effet physiologique et pathologique</i>	20
1.1.4. <i>Autres modifications post-traductionnelles dans le processus d'agrégation de la protéine Tau</i>	23
1.1.5. <i>Tau comme cible thérapeutique dans la maladie d'Alzheimer</i>	25
1.2 ETUDE DE LA PROTEINE TAU PAR SPECTROSCOPIE RMN	28
1.2.1. <i>Généralités sur les protéines non structurées et leur étude par RMN</i>	28
1.2.2. <i>Etudes structurales et fonctionnelles de Tau par RMN : structures résiduelles, liaison aux microtubules, formation des PHF et impact de la phosphorylation</i>	30
<b>CHAPITRE 2 : PHOSPHORYLATION DE LA PROTEINE TAU ET REGULATION DES INTERACTIONS AVEC LA PEPTIDYL-PROLYL CIS/TRANS ISOMERASE PIN1</b>	<b>34</b>
2.1 INTRODUCTION	34
2.1.1. <i>De l'hétérogénéité conformationnelle des motifs Ser/Thr-Pro phosphorylés à une régulation fonctionnelle par Pin1 ?</i>	34
2.1.2. <i>Méthodes de caractérisation de l'échange conformationnel</i>	34
2.1.3. <i>Interactions fonctionnelles entre Pin1 et Tau phosphorylée, et implications physiopathologiques</i>	36
2.2 RESULTATS	37
2.2.1. <i>Le motif phospho-Thr212-Pro213, un nouveau site de la protéine Tau permettant une interaction fonctionnelle avec Pin1 (Smet, Sambo et al. 2004)</i>	37
2.2.2. <i>Rôle du domaine WW et des phosphorylations multiples dans la régulation de l'activité cis/trans isomérase de Pin1 (Smet, Wieruszeski et al. 2005)</i>	40
2.2.3. <i>Vers l'étude de la protéine Tau entière et des interactions fonctionnelles avec Pin1 par RMN (Lippens, Wieruszeski et al. 2004; Smet, Leroy et al. 2004; Lippens, Sillen et al. 2006)</i>	41
2.2.4. <i>Rôle de Pin1 dans la stimulation de la déphosphorylation de Tau par la phosphatase PP2A (Landrieu, Smet-Nocca et al. 2011)</i>	44
2.2.5. <i>Inhibition de l'interaction de Pin1 avec Tau phosphorylée</i>	45
2.3 DISCUSSION.	50
2.3.1. <i>Rôle moléculaire de Pin1 dans la reconnaissance de substrats phosphorylés</i>	50
2.3.2. <i>Rôle de l'interaction Pin1/Tau phosphorylée dans la maladie d'Alzheimer</i>	52
<b>CHAPITRE 3 : LA RESONANCE MAGNETIQUE NUCLEAIRE DANS LA CARACTERISATION DES MODIFICATIONS POST-TRADUCTIONNELLES</b>	<b>54</b>
3.1 INTRODUCTION	54
3.2 RESULTATS	55
3.2.1. <i>La phosphorylation</i>	55
3.2.2. <i>L'acétylation</i>	61
3.2.3. <i>L'O-<math>\beta</math>-N-acétylglucosaminylation</i>	63

3.3 DISCUSSION : APPORT DE LA RMN A L'ETUDE DES MODIFICATIONS POST-TRADUCTIONNELLES (THEILLET, SMET-NOCCA ET AL. 2012)	64
3.3.1. <i>Le challenge analytique</i>	64
3.3.2. <i>PTM multiples</i>	65
3.3.3. <i>PTM et régions désordonnées</i>	66
3.3.4. <i>PTM et conséquences structurales/fonctionnelles</i>	66
<b>CHAPITRE 4. REGULATION STRUCTURALE ET FONCTIONNELLE DE TAU PAR LES MODIFICATIONS POST-TRADUCTIONNELLES</b>	<b>68</b>
4.1 ETUDE DE MODIFICATIONS POST-TRADUCTIONNELLES SPECIFIQUES AVEC DES PROTEINES TAU SEMI-SYNTHETIQUES	68
4.1.1. <i>Principe de la synthèse de protéines Tau semi-synthétiques</i>	68
4.1.2. <i>Etudes fonctionnelles de phospho-épitopes spécifiques par l'intermédiaire de protéines Tau semi-synthétiques</i>	72
4.2 O-GLCNAcylation : UN MECANISME ANTAGONISTE DE LA PHOSPHORYLATION DE TAU ET PROTECTEUR DE L'AGREGATION ?	74
4.2.1. <i>Introduction</i>	74
4.2.2. <i>O-GlcNAcylation de Tau : étude par RMN</i>	79
4.3 ACÉTYLATION DE TAU : UN MECANISME FAVORISANT L'ACCUMULATION DES FORMES PHOSPHORYLEES ET DES PHFS ?	86
4.3.1. <i>Introduction</i>	86
4.3.2. <i>Etude de l'acétylation de Tau : caractérisation des sites d'acétylation par RMN et impact fonctionnel</i>	88
REFERENCES	95
SELECTION D'ARTICLES	115

## Curriculum Vitae

---

### Caroline SMET-NOCCA

03/11/1977, 35 ans  
mariée, 2 enfants

✉ UGSF – UMR CNRS 8576  
Cité Scientifique - Bâtiment C9  
59650 Villeneuve d'Ascq  
☎ 03 20 43 49 97  
@ [caroline.smet@univ-lille1.fr](mailto:caroline.smet@univ-lille1.fr)

### MAITRE DE CONFERENCES EN BIOCHIMIE

Université des Sciences et Technologies de Lille – UFR de Biologie, Unité de Glycobiologie Structurale et Fonctionnelle (UMR CNRS 8576, Lille).

### Diplômes

---

- 2004** Doctorat de Chimie Organique (Directeur de Thèse : Dr. Guy Lippens ; UMR CNRS 8525 Institut Pasteur/Institut de Biologie de Lille, Université de Lille II, U.F.R. de Médecine) – Mention Très Honorable avec les Félicitations du Jury.  
*Thèse financée par une Bourse Docteur-Ingénieur (BDI) du CNRS et de la Région Nord-Pas de Calais.*
- 2001** Diplôme d'ingénieur de l'Ecole Nationale Supérieure de Chimie de Lille  
DEA (Master) en Chimie Organique et Macromoléculaire – Mention Bien (Lille I)
- 1999** Maîtrise de Chimie – Mention Bien (Lille I)  
Certificat d'études de Chimie Pharmaceutique (Lille II – Faculté de Pharmacie)
- 1998** Licence de Chimie – Mention Bien (Lille I)
- 1997** Diplôme d'Etudes Universitaires Générales "Sciences de la Matière" - Mention Bien (Lille I)
- 1995** Baccalauréat série Scientifique – Mention Bien

### Expériences professionnelles

---

#### Post-doctorats

**Oct. 2007–août 2009** Unité de Glycobiologie Structurale et Fonctionnelle (UMR CNRS 8576), groupe "RMN des Biomolécules", Lille (France). *Financement AIRMA (Association Internationale pour la Recherche sur la Maladie d'Alzheimer).*

*Régulation de la fonction de la peptidyl-prolyl cis/trans isomérase Pin1 et implication fonctionnelle de l'interaction de Pin1 avec Tau phosphorylée.*

**Sept. 2005–oct. 2007** Institut de Recherche Interdisciplinaire (USR CNRS 3078), groupe "Systems Epigenomics", Lille (France). *Financement de l'Agence Nationale de la Recherche.*

*Etude structurale et fonctionnelle du complexe TDG/CBP dans la régulation épigénétique de la transcription.*

**Janv.–juil. 2005** Intervet Innovation GmbH (actuellement Intervet–Schering Plough Animal Health), service "Drug Discovery/Lead Optimization", Francfort (Allemagne).

*Synthèse parallèle de dérivés macrolides à visée antibiotique.*



## Stages

**Avril-sept. 2001** Stage Ingénieur/DEA (6 mois) au laboratoire « Synthèse, Structure et Fonction des Biomolécules » (UMR CNRS 8525), Institut Pasteur de Lille / Institut de Biologie de Lille (France).

*Caractérisation d'interactions protéine-ligand par RMN. Simulation de l'échange chimique.*

**Juill.-août 2000** Stage Technicien (2 mois) au laboratoire « Synthèse, Structure et Fonction des Biomolécules » (UMR CNRS 8525), Institut Pasteur de Lille / Institut de Biologie de Lille/Université de Lille II (France).

*Synthèse peptidique et caractérisation de peptides par Résonance Magnétique Nucléaire.*

**Juill.-sept. 1998** Stage Technicien (3 mois) société LEVER (Groupe Unilever), Haubourdin (France).

*Contrôles des eaux de rejet et de la qualité microbiologique des lessives liquides.*

## Compétences

---

**Biochimie.** Clonage de protéines dans des vecteurs d'expression. Mutagenèse dirigée. Expression et purification de protéines recombinantes. Marquage isotopique de protéines ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ,  $^2\text{H}$ ) pour des études par RMN. Western blot. GST-pull down. Manipulation de radioisotopes ( $^{35}\text{S}$ ).

**RMN.** Acquisition et traitement de spectres de routine de RMN mono- et bi-dimensionnelle (Brüker) sur de petites molécules organiques (1H, 13C, COSY, TOCSY, ROESY, 1H-13C HSQC, 1H-13C HMQC). Acquisition et traitement de spectres en RMN des peptides et protéines (TOCSY, NOESY, 1H-15N HSQC, EXSY). Analyse conformationnelle de peptides par RMN. Attribution séquentielle de protéines par exploitation des données de RMN tri-dimensionnelle. Criblage de ligands par RMN.

**Fluorescence.** Spectres d'émission de fluorescence de protéines, FRET, polarisation de fluorescence. Criblage de composés par FRET.

**Chimie.** Synthèse en phase solide de peptides et de peptides modifiés (modifications post-traductionnelles, sondes fluorescentes, biotine); synthèse supportée de molécules de type peptido-mimétiques dans la recherche de hits pharmacologiques; Synthèse organique en milieu homogène de composés à visée thérapeutique (optimisation de leads pharmacologiques).

## Publications

---

### Articles publiés (dans des revues à comité de lecture).

1. [Smet-Nocca, C.](#), Launay, H. Wieruszkeski, J.M., Lippens, G. and Landrieu, I. (2013) Unraveling a phosphorylation event in a folded protein by NMR spectroscopy: phosphorylation of the Pin1 WW domain by PKA. *J Biomol NMR* 55(4), 323-337.
2. Delaune, A., Cabin-Flaman, A., Legent, G., Gibouin, D., [Smet-Nocca, C.](#), Lefebvre, F., Benecke, A., Vasse, M., and Ripoll, C. (2013) 50nm-Scale Localization of Single Unmodified, Isotopically Enriched, Proteins in Cells. *PLoS ONE* 8(2), e56559.
3. Ando, K., Dourlen, P., Sambo, A.V., Bretteville, A., Bélarbi, K., Vingtdeux, V., Eddarkaoui, S., Drobecq, H., Ghestem, A., Bégard, S., Demey-Thomas, E., Melnyk, P., [Smet, C.](#), Lippens, G., Maurage, C.A., Caillet-Boudin, M.L., Verdier, Y., Vinh, J., Landrieu, I., Galas, M.C., Blum, D., Hamdane, M., Sergeant, N., Buée, L. (2013) Tau pathology modulates Pin1 post-translational modifications and may be relevant as biomarker. *Neurobiol. Aging*, 34(3), 757-769.

4. Theillet, F.X., [Smet-Nocca, C.](#), Liokatis, S., Thongwichian, R., Kosten, J., Yoon, M.K., Kriwacki, R.W., Landrieu, I., Lippens, G., and Selenko, P. (2012) Cell signaling, post-translational protein modifications and NMR spectroscopy. *J. Biomol. NMR* 54(3), 217-236.
5. Landrieu, I., [Smet-Nocca, C.](#), Amniai, L., Louis, J.V., Wieruszeski, J.-M., Goris, J., Janssens, V., and Lippens, G. (2011) Molecular implication of PP2A and Pin1 in the Alzheimer's disease specific hyperphosphorylation of Tau. *PLoS One* 6(6), e21521.
6. Sami, F., [Smet-Nocca, C.](#), Khan, M., Landrieu, I., Lippens, G., and Brautigam, D.L. (2011) Molecular Basis for an ancient partnership between prolyl isomerase Pin1 and phosphatase inhibitor-2. *Biochemistry* 50(30), 6567-6578.
7. [Smet-Nocca, C.](#), Broncel, M., Wieruszeski, J.-M., Tokarski, C., Hanouille, X., Leroy, A., Landrieu, I., Rolando, C., Lippens, G., and Hackenberger, C. P. (2011) Identification of O-GlcNAc sites within peptides of the Tau protein and their impact on phosphorylation. *Mol. Biosyst.* 7(5), 1420-1429.
8. [Smet-Nocca, C.](#), Wieruszeski, J.-M., Leger, H., Eilebrecht, S., and Benecke, A. (2011) SUMO-1 regulates the conformational dynamics of Thymine-DNA Glycosylase regulatory domain and competes with its DNA binding activity. *BMC Biochemistry* 12, 4.
9. [Smet-Nocca, C.](#), Wieruszeski, J.-M., Melnyk, O., and Benecke, A. (2010) NMR-based detection of acetylation sites in peptides. *J. Pept. Sci.* 16, 414-423.
10. Landrieu, I., Leroy, A., [Smet-Nocca, C.](#), Huvent, I., Amniai, L., Hamdane, M., Sibile, N., Buée, L., Wieruszeski, J.-M., and Lippens, G. (2010) NMR spectroscopy of the neuronal Tau protein: normal function and implication in Alzheimer's disease. *Biochem. Soc. Trans.* 38(4), 1006-1011.
11. Eilebrecht, S.\*, [Smet-Nocca, C.\\*](#), Wieruszeski, J.-M., and Benecke, A. (2010) SUMO-1 possesses DNA binding activity. *BMC Res Notes* 3(1), 146-152. \*equal contribution
12. Clantin, B., Leyrat, C., Wohlkönig, A., Hodak, H., Ribeiro, E.D. Jr, Martinez, N., Baud, C., [Smet-Nocca, C.](#), Villeret, V., Jacob-Dubuisson, F. and Jamin, M. (2010) Structure and plasticity of the peptidyl-prolyl isomerase Par27 of *Bordetella pertussis* revealed by X-ray diffraction and small-angle X-ray scattering. *J. Struct. Biol.*, 169(3), 253-265.
13. [Smet-Nocca, C.](#), Wieruszeski, J.-M., Leroy, A., Chaar, V. and Benecke, A. (2008) The Thymine-DNA Glycosylase regulatory domain: residual structure and DNA binding. *Biochemistry* 47(25), 6519-6530.
14. Hodak, H., Wohlkönig, A., [Smet-Nocca, C.](#), Drobecq, H., Wieruszeski, J.-M., Sénéchal, M., Landrieu, I., Loch, C., Jamin, M. and Jacob-Dubuisson, F. (2008) The peptidyl-prolyl isomerase and chaperone Par27 of *Bordetella pertussis* as the prototype for a new group of parvulins. *J. Mol. Biol.* 376(2), 414-426.
15. Lippens, G., Landrieu, I. and [Smet, C.](#) (2007) Molecular mechanisms of the phospho-dependent prolyl *cis/trans* isomerase Pin1. *FEBS J.* 274(20), 5211-5222.
16. Landrieu, I., [Smet, C.](#), Wieruszeski, J. M., Sambo, A. V., Wintjens, R., Buee, L., and Lippens, G. (2006) Exploring the molecular function of Pin1 by Nuclear Magnetic Resonance. *Cur. Prot. Pept. Sci.* 7(3), 179-194.
17. Lippens, G., Sillen, A., [Smet, C.](#), Wieruszeski, J. M., Leroy, A., Buee, L., and Landrieu, I. (2006) Studying the natively unfolded neuronal Tau protein by solution NMR spectroscopy. *Protein Pept. Lett.* 13(3), 235-246.
18. [Smet, C.](#), Wieruszeski, J.-M., Landrieu, I., Buée, L., and Lippens, G. (2005) Regulation of Pin1 peptidyl prolyl *cis/trans* isomerase activity by its WW binding module on a multi-phosphorylated peptide of Tau protein. *FEBS Lett.* 579(19), 4159-4164.
19. [Smet, C.](#), Duckert, J.-F., Wieruszeski, J.-M., Landrieu, I., Buée, L., Lippens, G., and Déprez, B. (2005) Control of protein-protein interactions: structure-based discovery of low molecular weight inhibitors of the interactions between Pin1 WW domain and phosphopeptides. *J. Med. Chem.* 48(15), 4815-4823.
20. [Smet, C.](#), Leroy, A., Sillen, A., Wieruszeski, J.-M., Landrieu, I., and Lippens, G. (2004) Accepting its random coil nature allows a partial NMR assignment of the neuronal Tau protein. *ChemBiochem* 5(12), 1639-1646.

21. [Smet, C.](#), Sambo, A. V., Wieruszeski, J. M., Leroy, A., Landrieu, I., Buée, L., and Lippens, G. (2004) The peptidyl prolyl *cis/trans*-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau. *Biochemistry* 43(7), 2032-2040.
22. Lippens, G., Caron, L. and [Smet, C.](#) (2004) A microscopic view of chemical exchange: Monte Carlo simulation of molecular association. *Concepts Magn. Res. Part A* 21A(1), 1-9.
23. Lippens, G., Wieruszeski, J. M., Leroy, A., [Smet, C.](#), Sillen, A., Buée, L., and Landrieu, I. (2004) Proline-directed random-coil chemical shift values as a tool for the NMR assignment of the Tau phosphorylation sites. *Chembiochem* 5(1), 73-78.
24. Hamdane, M., Delobel, P., Sambo, A. V., [Smet, C.](#), Bégard, S., Violleau, A., Landrieu, I., Delacourte, A., Lippens, G., Flament, S., and Buée, L. (2003) Neurofibrillary degeneration of the Alzheimer-type: an alternate pathway to neuronal apoptosis? *Biochem. Pharmacol.* 66(8), 1619-1625.
25. Hamdane, M., [Smet, C.](#), Sambo, A. V., Leroy, A., Wieruszeski, J. M., Delobel, P., Maurage, C. A., Ghestem, A., Wintjens, R., Bégard, S., Sergeant, N., Delacourte, A., Horvath, D., Landrieu, I., Lippens, G., and Buée, L. (2002) Pin1: a therapeutic target in Alzheimer neurodegeneration. *J. Mol. Neurosc.* 19(3), 275-287.

### Chapitres de livre.

[Smet-Nocca, C.](#), Paldi, A., and Benecke, A. (2006) De l'épigénomique à l'émergence morphogénétique, dans *Morphogenèse, l'origine des formes*. Edtrs.: P Bourguine, A Lesne. Editions Belin.

[Smet-Nocca, C.](#), Paldi, A., and Benecke, A. (2006) From Epigenomic to Morphogenetic Emergence in Morphogenesis, *Origins Of Patterns And Shapes*. Editors : P Bourguine, A Lesne ; Springer : Complexity.

### Résumé de congrès.

Lippens, G., Landrieu, I., Amniai, L., Si, H., Leroy, A., and [Smet-Nocca, C.](#) (2012) The phosphorylation code of the neuronal Tau protein. *FEBS J* 279(S1), 51.

### Thèse.

*Etude des interactions entre la peptidyl-prolyl cis/trans isomérase Pin1 et la protéine microtubulaire Tau. Recherche d'inhibiteurs ciblant la liaison de Pin1 à ses substrats phosphorylés.*

Soutenue le 18 octobre 2004, Université de Lille II, Lille (France). Document téléchargeable à l'adresse: <http://tel.archives-ouvertes.fr/tel-00354986/fr/>

**Rapporteurs** : Dr Doris JACOBS, Dr Georges MASSIOT ; **Examineurs** : Dr Ralf WARRASS, Pr Benoît DEPPEZ, Dr Guy LIPPENS (Directeur de Thèse) ; **Président** : Dr Luc BUEE.

### Communications orales et affichées

---

- [Smet-Nocca C.](#), Broncel M., Hackenberger C. P. and Lippens G. O- $\beta$ -linked N-acetylglucosamylation of the microtubule-associated protein Tau (2010) Leibniz Institute of Molecular Pharmacology (FMP Berlin) conférence invitée, 14<sup>th</sup> June 2010, Berlin (Allemagne). *Communication orale*
- [Smet-Nocca C.](#), Wieruszeski J.-M., et Benecke A. Dynamique conformationnelle du domaine de régulation de la Thymine DNA Glycosylase : structure résiduelle, liaison à l'ADN et interactions avec SUMO-1 (2009) 16<sup>e</sup> congrès du Groupe Français des Peptides et des Protéines, 10-15 mai 2009, Albé (France). *Communication orale*
- Landrieu I, [Smet C.](#), Wieruszeski J-M and Lippens G. Exploring the molecular function of the PIN1 prolyl *cis/trans* isomerase by Nuclear magnetic (2008) Resonance Summer School NEURAD Neurodegeneration in Alzheimer's disease—mechanisms consequences and therapy, 17-20 Sep.2008 Leuven (Belgique). *Communication orale*

- [Smet-Nocca C.](#), Chaar V., Léger H. and Benecke A. (2007) NMR of the human Thymine-DNA Glycosylase: conformational dynamics of the N-terminal regulatory domain. Séminaire Institut de Recherche Interdisciplinaire, 28 mars 2007 (France). *Communication orale*
- Landrieu I, Sillen A., Wieruszeski J-M, Leroy A., [Smet C.](#), Sibille, N. and Lippens G. NMR spectroscopy of the neuronal Tau protein: Normal function and implication in Alzheimer Disease (2007) Advance and management of NMR in Life Sciences : a joint annual meeting of (I3) EU-NMR and (CA) NMR-LIFE, 18-20 Jan., 2007 Florence (Italie). *Communication orale*
- Landrieu, I., [Smet, C.](#), Wieruszeski and Lippens, G. Exploring the molecular function of PIN1 prolyl cis/trans isomerisation by Nuclear Magnetic Resonance (2006), Protein posttranslational modifications: analysis and function, MPSA2006, 29 août-1er septembre 2006, Lille (France), International. *Communication orale*
- Lippens, G., [Smet, C.](#), Sillen, A., Leroy, A., Wieruszeski, J.-M., Buée, L., and Landrieu, I. (2005) Studying the neuronal Tau protein by solution NMR spectroscopy. Congrès SFC-EuroChem Nancy, 28 Août-1er Septembre 2005 (France). *Communication orale*
- Landrieu, I., [Smet, C.](#), Leroy, A., Sillens, A., Wieruszeski, J.-M., Buée, L. and Lippens, G. (2005) Exploring the molecular function of PIN1 with Tau model substrate, NMR in Molecular Biology, ESF Research Conferences, 20-25 Août 2005, Höör (Suède). *Affiche*
- [Smet, C.](#), Landrieu, I., Wieruszeski, J.-M., Leroy, A., Duckert, J.-F., Sambo, A.-V., Déprez, B., Buée, L. and Lippens, G. (2004) Targeting protein-protein interaction : Inhibition of Pin1-phospho-epitope interactions with small-molecule ligands of the Pin1 WW binding module. 40e Rencontres Internationales de Chimie Thérapeutique, Société de Chimie Thérapeutique ; 30 juin-2 juillet 2004, Bordeaux (France). *Affiche*
- [Smet, C.](#), Landrieu, I., Wieruszeski, J.-M., Leroy, A., Duckert, J.-F., Sambo, A.-V., Déprez, B., Buée, L. and Lippens, G. (2004) Peptide substrates for phospho-epitope screening and characterization of catalytical activity of the human prolyl cis/trans isomerase Pin1 by NMR. Design, synthesis and NMR screening of WW potential inhibitors. Workshop "Molecules as modulators: systems biology challenges chemistry", Aventis / Université de Francfort ; 29-31 janvier 2004, Wiesbaden (Allemagne). *Affiche*
- [Smet, C.](#), Wieruszeski, J. M., Landrieu, I., Buée, L., and Lippens, G. (2003) La RMN, un outil pour la détection des interactions de faible affinité. 3e Colloque annuel des doctorants ; 2 septembre 2003, Lille (France). *Communication orale*
- Landrieu, I., Leroy, A., [Smet, C.](#), Wieruszeski, J.-M., Wintjens, R., Sambo, A.-V., Inzé, D., Buée, L. and Lippens, G. (2003) Proline conformation as a regulatory element of molecular function FEBS meeting on Signal Transduction, 3-8 juillet 2003, Bruxelles (Belgique). *Affiche*
- [Smet, C.](#), Wieruszeski, J. M., Leroy, A., Landrieu, I., Buée, L., and Lippens, G. (2003) Etude des interactions entre la peptidyl-prolyl cis/trans isomerase Pin1 et la protéine microtubulaire Tau. Congrès « RMN biologique et les enjeux du vivant », Société Française de Biochimie et de Biologie Moléculaire ; 12-14 mars 2003, Clermont-Ferrand (France). *Communication orale*
- Lippens, G., [Smet, C.](#), Wieruszeski, J. M., Leroy, A., Sillen, A., Sambo, A.-V., Bégard, S., Buée, L., and Landrieu, I. (2003) NMR interaction study of the prolyl cis/trans isomerase Pin1 and the microtubule-associated protein Tau; Obernai (France). *Communication orale*
- Lippens, G., Landrieu, I., Odaert, B., Wintjens, R., [Smet, C.](#), Wieruszeski, J. M., Buée, L. (2002) NMR study of CKS and Pin1, two proteins involved in the eukaryotic cell cycle. Congrès « La RMN, un outil pour la biologie IV », Institut Pasteur de Paris ; 4-6 février 2002, Paris (France). *Communication orale*

---

### **Collaborations**

- Depuis 2009 : Pr Dr C.P. Hackenberger, Chemical Biology Research Group, Leibniz-Institut fuer molekulare Pharmakologie, Berlin (Allemagne)

- Depuis 2011 : Dr Gal Bitan, Brain Research Institute, University of California, Los Angeles (Etats-Unis) et Dr Thomas Schrader, Department of Chemistry, University Duisburg-Essen (Allemagne)
- Depuis 2012 : Dr Florence Pinet, Equipe Recherche des Déterminants Moléculaires des Maladies Cardiovasculaires, Institut Pasteur de Lille/Inserm U744 (France)

### **Enseignements**

---

- **Licence SVTE parcours Biochimie (semestres 4 et 5)** : UE "Caractérisation structurale des biomolécules" (20h TD, 60h TP), UE "Structure des protéines et métabolisme des acides aminés" (8h TD), UE "Technique d'analyse des molécules du vivant" (12h TD), UE "Structure et métabolisme des glucides" (20h TD)
- **Master Biologie et Biotechnologie (semestres 1 et 2)** : UE "Analyse structurale des biomolécules et imagerie" Résonance Magnétique Nucléaire appliquée à l'étude des protéines (6h CM, 4h TD)
- **Master Chimie Biologie parcours Chimie Bioanalytique (semestre 3)** : UE "Biomolécules : structure et interactions moléculaires" Résonance Magnétique Nucléaire appliquée à l'étude des protéines (8h CM, 4h TD)
- **Master Pro Génomique et Protéomique (semestre 2)** : UE Ateliers Biotechnologiques, atelier "Modélisation moléculaire" (30h TP)
- **Licence SVTE (semestre 2)** : UE "Molécules du vivant" (12h TP)

### **Activités d'encadrement**

---

- Sophie Delhalle – Master 1 Chimie Biologie, Lille 1 (stage de 2 mois, 2013) : optimisation de la production de protéine OGT recombinante et tests d'activité *in vitro*.
- Hélène Léger – Thèse (taux d'encadrement 20%), UGSF (2009-2012), Lille 1 : Rôle du complexe CBP/TDG/RAR dans la régulation épigénétique de la transcription (Smet-Nocca, Wieruszkeski et al. 2011; Léger, Smet-Nocca et al. en préparation).
- Sven Potelle – Licence de Biologie, parcours Biochimie, (stage de 1 mois, 2012), Lille 1 : formation à l'expression et à la purification de protéines recombinantes pour les études par RMN, à la préparation d'échantillons pour la RMN.
- Meerakhan Pathan – Master Erasmus Mundus ASC (M2) Advanced Spectroscopy in Chemistry (2009-2010), Lille 1 : interactions de l'inhibiteur-2 de la phosphatase PP1 avec la peptidyl-prolyl *cis/trans* isomérase Pin1 (Sami, Smet-Nocca et al. 2011).
- Julien Michel – Master 2 Chimie Organique et Macromoléculaire - 3<sup>e</sup> année Ecole Nationale Supérieure de Chimie de Lille (2002-2003), Lille 1 : formation à la synthèse peptidique en phase solide, à la purification de peptides et à l'attribution de peptides par RMN homonucléaire.
- Amandine Menet – BTS (stage de 2 mois, 2003), Douai : formation à la synthèse peptidique en phase solide et à la purification de peptides.
- Sylviane Piola – Master 2 Pro (stage de 2 mois, 2002), Lille 1 : formation à la synthèse peptidique en phase solide, à la purification de peptides et à l'attribution de peptides par RMN homonucléaire.

### **Responsabilités administratives et collectives**

---

- Responsable de l'UE "Caractérisation structurale des biomolécules" en Licence SVTE, Biochimie et de l'UE "Biomolécules : structure et interactions moléculaires" dans le Master 2 Chimie Biologie
- Organisation des groupes de TD/TP en Licence SVTE, parcours Biochimie (semestre 5).
- Entre septembre 2005 et octobre 2007 : mise en place d'un laboratoire de biologie structurale à l'Institut de Recherche Interdisciplinaire (USR CNRS 3078) pour la production (clonage-expression-purification) de protéines recombinantes satisfaisant aux exigences des études structurales (quantité, pureté, marquage isotopique).

## Abréviations

---

APP : *Amyloid Precursor Protein* ou protéine précurseur du peptide A $\beta$

CBP : *CREB-binding protein*

CDK : *cyclin-dependent kinases*

CREB : *cAMP-responsive element (CRE)-binding protein*

EPL : *Expressed protein ligation*

ERK2/MAPK1 : *Extracellular signal-regulated kinase 2/Mitogen-activated protein kinase 1*

FKBP : *FK506-binding protein*

FTDP-17 : *Frontotemporal dementia with parkinsonism linked to chromosome 17* ou démences frontotemporales avec syndrome parkinsonien liée au chromosome 17

GSK3 : *Glycogen Synthase Kinase 3*

HAT : Histone acétyltransférase

HDAC : Histone deacétylase

HPLC : *High performance liquid chromatography* ou chromatographie liquide à haute performance

HSQC : *Heteronuclear Single Quantum Correlation*

IDP : *Intrinsically Disordered Proteins* ou protéines intrinsèquement désordonnées

K<sub>D</sub> : constante de dissociation

MA : maladie d'Alzheimer

MALDI-TOF : *Matrix-assisted Laser Desorption Ionization – Time-Of-Flight*

NFT : *Neurofibrillary Tangles* ou enchevêtrements neurofibrillaires

NIMA : *Never In Mitosis A*

NOE : *Nuclear Overhauser Effect*

OGA : *O-GlcNAc hydrolase*

O-GlcNAc : *O- $\beta$ -N-acétylglucosamine*

OGT : *O-GlcNAc transférase (uridine diphosphate-N-acétyl-D-glucosamine: polypeptidyl-transferase)*

PHF : *Paired Helical Filaments* ou paire de filaments appariés en hélice

Pin1 : *Protein-Interacting with NIMA-1*

PKA : *cAMP-dependent Protein Kinase A*

PP2A : *Protein Phosphatase 2A*

PPIase : *peptidyl-prolyl cis/trans isomérase*

PTM : *posttranslational modifications* ou modifications post-traductionnelles

RMN : Résonance Magnétique Nucléaire

TPR : *tetratricopeptide repeat*

UDP-GlcNAc : *uridine diphosphate N-acétyl-D-glucosamine*

WW : domaine Trp-Trp

WW<sup>PKA</sup> : domaine WW phosphorylé par PKA



## Introduction générale : parcours scientifique et thématiques de recherche

---

Mon parcours universitaire et post-doctoral m'a conduit à étudier, après une formation en chimie générale, les protéines au niveau structural, en me focalisant sur leurs interactions moléculaires et leur fonction. Les principaux outils que j'ai utilisés sont des techniques spectroscopiques comme la résonance magnétique nucléaire (RMN) et la fluorescence, mais aussi les techniques de biologie moléculaire et structurale pour le clonage, l'expression et la purification de protéines recombinantes dans *E. coli*. De plus, ma formation en chimie organique m'a permis de synthétiser des petites molécules ou des peptides modifiés ciblant la fonction des protéines ou les interfaces protéine-protéine. Ces molécules ont été utilisées comme modèles dans la compréhension des processus biologiques, des mécanismes associés aux dérégulations pathologiques, ainsi que dans la découverte de nouvelles cibles pharmacologiques. Mon objectif scientifique est donc de caractériser, à l'aide de méthodes structurales comme la RMN, les surfaces d'interaction protéine-protéine pour des complexes impliqués dans des processus physiopathologiques. A ce titre, les modifications post-traductionnelles des protéines qui entrent en jeu dans l'établissement de complexes macromoléculaires spécifiques et dans la régulation de leur fonction, sont d'un intérêt évident. Récemment, les avancées technologiques et méthodologiques ont permis d'étendre le champ d'investigation de la RMN au-delà de la caractérisation structurale de protéines ou de complexes. La vision dynamique d'une protéine en solution, le criblage de composés à visée thérapeutique ou l'utilisation de la RMN comme technique analytique dans la caractérisation des modifications post-traductionnelles en sont quelques exemples qui seront illustrés dans ce manuscrit. Le travail présenté ici concerne l'étude, au niveau moléculaire, de la régulation de la protéine neuronale Tau par diverses modifications post-traductionnelles. Cette protéine est impliquée dans le maintien de l'architecture microtubulaire du neurone et dans le processus de neurodégénérescence de la maladie d'Alzheimer. Avant de commencer, je vais décrire brièvement mon parcours scientifique et professionnel.

Après l'obtention d'un diplôme d'Ingénieur de l'Ecole Nationale Supérieure de Chimie de Lille et d'un DEA de Chimie Organique et Macromoléculaire, j'ai effectué ma thèse (2001-2004, BDI CNRS/Région) sous la direction du Dr. Guy Lippens (UMR CNRS 8525, Institut de Biologie/Institut Pasteur de Lille) avec pour thématique l'étude, au niveau moléculaire, du complexe Pin1/Tau phosphorylée et son implication dans la maladie d'Alzheimer. J'ai pu ainsi acquérir les méthodes associées à l'étude des protéines par RMN, comme l'expression et la purification de protéines recombinantes enrichies avec des isotopes stables, l'acquisition et l'interprétation de spectres RMN bi- et tridimensionnels pour la caractérisation de complexes protéine-protéine, et les stratégies de criblage par RMN. Parallèlement, l'utilisation de peptides modifiés, synthétisés par chimie en phase solide, m'a permis d'obtenir un grand nombre d'informations sur ces interfaces protéine-protéine, difficilement accessibles autrement. La même stratégie de synthèse a été employée pour l'obtention de mimes peptidiques rompant des interactions protéine-protéine spécifiques. Ces travaux ont permis d'ouvrir le chemin vers la compréhension des mécanismes de reconnaissance par Pin1 de la protéine Tau phosphorylée et de régulation de la fonction de Pin1.

Une première expérience post-doctorale en Allemagne (2005), dans l'industrie pharmaceutique (Intervet Innovation GmbH), m'a permis de renforcer mon expérience en synthèse organique avec l'optimisation de la synthèse en phase homogène d'un composé (macrolide) à visée antibiotique, à partir d'un composé naturel, et la synthèse parallèle d'une série de composés dérivés.

J'ai rejoint ensuite le groupe " Systems Epigenomics " du Dr Arndt Benecke (2006-2008, financement ANR), au sein du nouvel Institut de Recherche Interdisciplinaire (USR CNRS 3078) à Lille. Cette équipe a pour objectif la compréhension, selon une approche multi- et interdisciplinaire, des mécanismes à la base de la régulation de l'expression des gènes pour définir de nouvelles stratégies d'interférence thérapeutique, en parfait accord avec mes objectifs scientifiques. J'ai eu l'occasion, dans le contexte de cet institut émergent, de mettre en place un laboratoire de biologie structurale pour permettre l'exécution de mon projet de recherche : l'étude structurale et fonctionnelle des interactions entre la CREB-binding protein (CBP) et la Thymine DNA Glycosylase humaine (TDG), le premier lien entre la réparation d'ADN par excision de base, la transcription et l'épigénèse. Le but de ce projet était de définir les mécanismes moléculaires impliqués dans la réparation de l'ADN au niveau des sites de régulation épigénétique que sont les sites de méthylation de l'ADN au sein des îlots CpG, ainsi que les interactions entre TDG et ses partenaires : le coactivateur général de la transcription CBP et la protéine SUMO-1. Nos données structurales, obtenues par RMN, ont permis de définir les bases moléculaires pour permettre, à partir d'analyses transcriptomiques, de définir le rôle du complexe CBP/TDG dans la régulation épigénétique de la transcription, et d'identifier les réseaux de signalisation associés dans les leucémies.

En 2008, j'ai souhaité rejoindre à nouveau le groupe du Dr. Guy Lippens (financement de l'Association Internationale pour la Recherche sur la Maladie d'Alzheimer), qui fait désormais partie de l'Unité de Glycobiologie Structurale et Fonctionnelle (UMR CNRS 8576, Université des Sciences et Technologies de Lille). J'ai commencé ce post-doctorat avec différents projets autour de la peptidyl-prolyl *cis/trans* isomérase Pin1, dans la continuité de mes travaux de thèse, pour comprendre le rôle de l'isomérisation *cis/trans* des prolines comme modification post-traductionnelle : son impact la fonction des protéines et les interactions avec leurs partenaires. Les différents projets menés autour de Pin1 concernent (i) la mise en évidence du rôle de Pin1, au niveau moléculaire, dans la déphosphorylation de Tau par la phosphatase PP2A (en collaboration avec le Pr Janssens, Faculty of Medicine of Leuven University, Belgique), (ii) l'étude d'un mécanisme allostérique de régulation de la fonction de Pin1 par interaction avec l'inhibiteur-2 de la phosphatase PP1 (IPP-2) dans sa forme non phosphorylée (en collaboration avec le Pr D. Brautigan, Center for Cell Signaling, University of Virginia School of Medicine, Charlottesville, USA) et (iii) l'étude structurale et dynamique de la forme phosphorylée du domaine WW de Pin1 comme mécanisme inhibiteur de la liaison des protéines phosphorylées. Parallèlement, dans le cadre d'une collaboration avec le Dr Françoise Jacob-Dubuisson (Inserm U1019/ CNRS UMR8204/ Institut Pasteur de Lille, Lille), j'ai été impliquée dans la mise en évidence de l'activité PPIase de Par27, une protéine chaperonne de *Bordetella pertussis*.

Depuis 2009, j'ai orienté mes activités de recherche vers l'étude des modifications post-traductionnelles de la protéine neuronale Tau, que je mène en parallèle à mes activités d'enseignement. Ce projet comporte plusieurs approches : l'étude de la phosphorylation, de l'O- $\beta$ -N-acétylglucosaminylation (O-GlcNAcylation) et de l'acétylation de la protéine Tau au niveau structural



et fonctionnel, et de l'effet réciproque de ces différentes modifications. La caractérisation par RMN de la phosphorylation de Tau par différentes kinases purifiées ou dans des extraits cellulaires est une thématique qui a été développée dans notre groupe par le Dr I. Landrieu. Je vais m'intéresser à la phosphorylation de Tau, avec une approche complémentaire, au travers de la production d'une protéine Tau semi-synthétique permettant l'étude de l'impact de sites de phosphorylation spécifiques sur la conformation et la fonction de Tau (en collaboration avec le Pr C. Hackenberger, Leibniz-Institut fuer molekulare Pharmakologie, Berlin, Allemagne). Cette stratégie pourra également être utilisée pour l'étude d'autres modifications post-traductionnelles. L'O-GlcNAcylation et l'acétylation de Tau sont deux modifications post-traductionnelles qui ont été décrites plus récemment et qui offrent de nouvelles perspectives pour interférer dans le processus pathologique d'agrégation de Tau dans la maladie d'Alzheimer. Une étude des sites de modification et des mécanismes au niveau moléculaire est un prérequis pour pouvoir interférer dans ces processus au niveau cellulaire.

## Chapitre 1 : La protéine neuronale Tau et son étude par Résonance Magnétique Nucléaire

---

### 1.1 LA PROTEINE TAU, UNE PROTEINE ASSOCIEE AUX MICROTUBULES IMPLIQUEE DANS LA MALADIE D'ALZHEIMER

Les neurones ont un cytosquelette constitué de plusieurs types de filaments : les microtubules, les filaments d'actine et les neurofilaments. Les microtubules sont particulièrement abondants dans l'axone des neurones où ils jouent un rôle essentiel dans leur organisation tridimensionnelle et dans le transport de vésicules du corps cellulaire vers les terminaisons axonales. La dynamique du cytosquelette des neurones joue un rôle prépondérant dans leur différenciation, le maintien de leur forme et des connexions entre les neurones qui en résultent. Les microtubules sont des objets macromoléculaires de 20-30 nm de diamètre, constitués d'une protéine dimérique de 110 kDa, la tubuline. Ces structures, très dynamiques, sont régulées par la polymérisation/dépolymérisation de la tubuline via l'hydrolyse du GTP en GDP. D'autre part, des protéines microtubulaires associées, les MAP (*microtubule-associated proteins*), stabilisent les microtubules par liaison, donc en intervenant plutôt dans des quantités stœchiométriques que catalytiques. La protéine neuronale Tau est une des protéines associées aux microtubules et joue un rôle prépondérant dans un certain nombre de pathologies neurodégénératives, appelées tauopathies, dont la plus connue est la maladie d'Alzheimer (MA). La protéine Tau est, dans sa forme libre, dissociée des microtubules, une protéine intrinsèquement désordonnée (IDP pour *Intrinsically Disordered Proteins*). Du fait de la nature instable des microtubules et de la nature désordonnée de la protéine Tau, le mode de liaison de Tau aux microtubules et son mode d'action sont difficiles à établir, et restent encore à élucider.

Décrite pour la première fois en 1975, la protéine Tau a été isolée de cerveaux de porcs après dépolymérisation des microtubules et dissociation de la tubuline (Weingarten, Lockwood et al. 1975; Cleveland, Hwo et al. 1977; Cleveland, Hwo et al. 1977). Elle y est montrée comme un facteur thermiquement stable, nécessaire à la polymérisation de la tubuline *in vitro*. Tau est principalement une protéine neuronale, localisée dans l'axone, mais on la retrouve également dans le compartiment somato-dendritique du neurone ou dans le noyau (Loomis, Howard et al. 1990; Brady, Zinkowski et al. 1995). Une des premières caractérisations structurales de Tau, réalisée par RMN  $^1\text{H}$  à 270 MHz en 1983, montre une protéine très flexible qui perd en partie sa flexibilité en présence de dimères de tubuline (Woody, Clark et al. 1983). Par la suite, de nombreux travaux se sont intéressés à la caractérisation structurale de la protéine Tau, notamment aux différentes isoformes et à leur expression spatio-temporelle ainsi qu'à la modulation de la liaison de Tau aux microtubules en fonction des isoformes ou de modifications post-traductionnelles (PTM), en particulier, la phosphorylation. Ainsi, des régions fonctionnelles ont été identifiées correspondant à des domaines directement impliqués dans la liaison des microtubules ou à des domaines régulateurs de cette interaction (voir paragraphe 1.1.1).

D'autre part, la protéine Tau a été identifiée comme un des marqueurs pathologiques de la MA. Plus généralement, différentes maladies neurodégénératives impliquant la protéine Tau ont été définies sous le terme de tauopathies, où la forme pathologique de Tau présente des caractéristiques

communes (maladie de Pick, démences frontotemporales avec syndrome parkinsonien liée au chromosome 17 ou FTDP-17, paralysie supranucléaire progressive ou PSP...). Deux types de dépôts ont été observés dans la MA : les dépôts amyloïdes, encore appelés plaques séniles, constitués d'agrégats extracellulaires de peptides A $\beta$  provenant de la coupure de la protéine transmembranaire APP (*Amyloid Precursor Protein*) par les sécrétases, et la dégénérescence neurofibrillaire (NFT pour *neurofibrillary tangles*), correspondant à des dépôts intraneuronaux de PHF (*Paired Helical Filaments* ou paires de filaments hélicoïdaux) constitués principalement de protéines Tau hyperphosphorylées (voir paragraphe 1.1.2). Les caractéristiques des lésions de type NFT (leur composition en protéines Tau et leur morphologie) ainsi que leur propagation dans les différentes régions du cerveau varient selon les tauopathies. Ces lésions, déjà observées par Aloïs Alzheimer en 1907, n'ont commencé à être caractérisées qu'à partir des années 1960.

Les lésions NFT se propagent dans le cerveau avec le développement de la maladie, suggérant un transfert interneuronal des protéines Tau pathologiques à la manière du prion (Braak and Braak 1991; Su, Deng et al. 1997; Armstrong, Cairns et al. 2001). La dégénérescence neurofibrillaire va suivre un chemin précis dans les différentes régions du cerveau de manière progressive et commence bien avant (plusieurs années) l'apparition des signes cliniques (pertes de mémoires, aphasia, apraxie, agnosie). Ces signes cliniques évoluent avec le chemin de propagation de la pathologie Tau, selon les aires du cerveau touchées. En revanche, la distribution des plaques amyloïdes est plus hétérogène et plus diffuse. Les deux types de lésions apparaissent généralement simultanément mais n'occupent pas les mêmes régions du cerveau. Alors que la pathologie amyloïde a longtemps été considérée comme prépondérante dans la MA, la pathologie Tau a pris récemment toute son importance depuis qu'il a été établi que l'étendue de lésions NFT est corrélée avec la progression de la maladie, contrairement aux plaques séniles. De plus, des mutations du gène *tau*, connues pour accélérer la formation des fibres pathologiques, sont un des facteurs étiologiques dans d'autres tauopathies. Enfin, la protéine Tau peut être utilisée comme biomarqueur de la MA car elle est détectée spécifiquement et précocement dans le fluide cérébrospinal de patients atteints de la MA et non dans ceux des contrôles correspondant au vieillissement normal ou à d'autres tauopathies (Arai, Terajima et al. 1995; Johnson, Seubert et al. 1997; Sjogren, Rosengren et al. 2000; Saman, Kim et al. 2012). Cependant, les pathologies Tau et A $\beta$  sont étroitement liées. En particulier, il a été montré qu'une dérégulation dans la transformation de Tau (phosphorylation et clivage protéolytique) est une conséquence de l'accumulation anormale de peptide A $\beta$ , et inversement, Tau augmente significativement la cytotoxicité de A $\beta$  dans la MA (Rapoport, Dawson et al. 2002; Roberson, Scarce-Levie et al. 2007).

### ***1.1.1. Isoformes et domaines de la protéine Tau dans la régulation fonctionnelle de la liaison aux microtubules***

Dans le cerveau humain, la protéine Tau est produite à partir d'un gène unique (localisé sur le chromosome 17) par épissage alternatif de l'ARNm. Les six isoformes de Tau, constituées de 352 à 441 acides aminés, diffèrent par l'insertion des régions peptidiques codées par les exons 2, 3 et 10. Les exons 2 et 3 codent pour un insert de 29 (exon 2) ou 58 (exons 2 et 3) acides aminés dans la partie N-terminale de la protéine, tandis que l'exon 10 code pour un fragment répété de 31 acides aminés, le R2, dans la région C-terminale. La protéine Tau contient au minimum trois segments

répétés, appelés R1, R3 et R4. Trois des six isoformes contiennent également le segment R2 (Tau4R) et trois n'en contiennent pas (Tau3R). Ces trois ou quatre fragments répétés, avec les régions avoisinantes, notamment le C-terminus, constituent le domaine de liaison aux microtubules (Gustke, Trinczek et al. 1994; Trinczek, Biernat et al. 1995) qui présente des homologies de séquences avec d'autres MAP ; le reste de la protéine est appelé domaine de projection. Le segment R2 constitue un point d'ancrage supplémentaire aux microtubules. Il a en effet été montré que les isoformes contenant quatre segments répétés ont une affinité plus importante pour la tubuline que les isoformes n'en possédant que trois (Lu and Kosik 2001), contribuant davantage à la stabilité des microtubules. Les protéines Tau sont donc des protéines essentielles au transport intra-neuronal en contrôlant la stabilité des microtubules, et interviennent ainsi indirectement dans la plasticité neuronale. Les segments répétés du domaine de liaison aux microtubules sont constitués de séquences similaires, mais non identiques, de 31 ou 32 résidus. Chaque segment est composé de 18 résidus appelés *repeat* correspondant à la région minimale de liaison à la tubuline, espacés du *repeat* voisin par une séquence moins conservée de 13 ou 14 résidus, appelée *inter-repeat*. Les deux premiers *repeats* ainsi que l'*inter-repeat* R1-R2 présentent la plus grande capacité à lier les microtubules (Goode, Denis et al. 1997). La liaison de Tau à la tubuline implique des interactions ioniques entre le domaine C-terminal acide de la tubuline et la région basique de Tau. En particulier, deux résidus lysine, K274 et K281, localisés dans le premier *inter-repeat* R1-R2, sont essentiels à la liaison de la tubuline et à la promotion de l'assemblage microtubulaire (Goode and Feinstein 1994; Trinczek, Biernat et al. 1995). Il a été proposé que Tau se lie à la fois à la queue C-terminale acide, non structurée, des sous-unités  $\alpha$  et  $\beta$  de la tubuline ainsi qu'aux régions exposées à la surface externe des microtubules (Nogales, Wolf et al. 1998). Deux modes de liaison ont été proposés : Tau pourrait se lier à un seul monomère de tubuline ou à deux monomères adjacents (Chau, Radeke et al. 1998). D'autre part, il a été montré que la polyglutamylation de la tubuline, au niveau de la région C-terminale des sous-unités  $\alpha$  et  $\beta$ , régule de manière indirecte l'interaction avec Tau, via un changement conformationnel (dé)masquant le site de liaison à la tubuline (Boucher, Larcher et al. 1994).

D'un point de vue physico-chimique, la protéine Tau est globalement basique (avec un pI supérieur à 9) et présente notamment une forte concentration de résidus lysine (44 dans l'isoforme la plus longue, soit 10% de la séquence). Cependant, la région N-terminale du domaine de projection contient une forte proportion de résidus acides comparée au reste de la protéine. La deuxième partie du domaine de projection contient une forte proportion de résidus proline, d'où son nom de domaine riche en prolines. Différents rôles ont été attribués au domaine de projection. Il joue notamment un rôle dans l'espacement des microtubules en se projetant vers l'extérieur de leur surface et en établissant un pont avec les microtubules voisins, d'où sa dénomination (Hirokawa, Shiomura et al. 1988; Chen, Kanai et al. 1992). Il peut également jouer un rôle de médiateur d'interactions avec d'autres protéines, plus particulièrement au niveau de la région riche en prolines. La présence de motifs PPXXP ou PXXP dans cette région pourrait jouer un rôle dans l'interaction de Tau avec les domaines SH3, modules d'interaction protéine-protéine spécifiques des régions polyprolines. Cette région pourrait par conséquent être le médiateur d'interactions avec des protéines associées à la membrane cytoplasmique (Brandt, Leger et al. 1995). De plus, la région riche en prolines a été décrite dans la régulation de la liaison de Tau aux microtubules. Plus précisément, la séquence peptidique <sup>224</sup>KKVAVVR<sup>230</sup>, localisée en amont du domaine de liaison aux microtubules,

favorise fortement la liaison des microtubules et leur assemblage (Goode, Denis et al. 1997). De manière significative, la région riche en prolines est le siège de nombreux sites de phosphorylation, des PTM décrites dans la modulation de la liaison de Tau aux microtubules (voir paragraphe 1.1.3).

### **1.1.2. Pathologie de la protéine Tau dans la maladie d'Alzheimer**

Observés par microscopie électronique pour la première fois en 1963 (Kidd 1963; Kidd 1964), les fibres insolubles appelées PHF sont des filaments appariés en hélice de 10 nm de diamètre. Les PHF sont une signature morphologique de la dégénérescence neurofibrillaire de type Alzheimer et semblent spécifiques à l'espèce humaine. Ce n'est qu'en 1986 qu'a été identifié le constituant majeur des PHF, la protéine Tau. Elle y apparaît sous une forme hyperphosphorylée et toutes les isoformes y sont représentées (Grundke-Iqbal, Iqbal et al. 1986; Grundke-Iqbal, Iqbal et al. 1986; Kosik, Joachim et al. 1986; Wood, Mirra et al. 1986). Assez rapidement, des formes tronquées de Tau dans les PHF ont été détectées (Wischik, Novak et al. 1988; Wischik, Novak et al. 1988). Le mécanisme de la formation des PHF n'est pas encore totalement élucidé et toujours sujet à controverse. Autre question d'intérêt : est-ce que les PHF sont la cause primaire de la MA ou jouent-ils un rôle secondaire, notamment parce qu'ils seraient une conséquence de la formation des plaques amyloïdes ? En effet, la pathologie A $\beta$  a longtemps été considérée comme primordiale du fait de l'existence de mutations dans les gènes de l'APP, de la préséniline-1 ou -2 qui engendrent toujours la MA, en accélérant notamment la formation de peptide A $\beta$  (formes familiales, 1% des cas) alors qu'aucune mutation de Tau n'a été décrite dans la MA. En revanche, des mutations du gène *tau* ont été rapportées dans la FTDP-17 et constituent un facteur étiologique de la maladie.

Le mécanisme par lequel Tau induit une toxicité neuronale reste également non élucidé. La formation des PHF dans la MA implique à la fois une perte de fonction de Tau, qui ne peut plus assurer sa fonction physiologique sous cette forme insoluble, mais également un gain de fonction toxique avec la formation d'espèces oligomériques solubles, cytotoxiques. Il a été montré que ces formes oligomériques sont présentes dans les neurones aux stades précoces de la pathologie contrairement aux PHF et perturbent le transport axonal antérograde. Au niveau moléculaire, ces oligomères sont riches en structures en feuillet  $\beta$ . De manière intéressante, il a été proposé que ces oligomères servent à la nucléation et à la propagation des fibres pathologiques selon un mécanisme similaire au prion (Lasagna-Reeves, Castillo-Carranza et al. 2012; Lasagna-Reeves, Castillo-Carranza et al. 2012; Ward, Himmelstein et al. 2012). En effet, des oligomères solubles, purifiés de cerveaux de patients atteints de la MA, propagent la pathologie Tau *in vivo*, dans des souris saines, non transgéniques. *In vitro*, ces oligomères sont bien plus toxiques que des oligomères générés *in vitro* à partir de protéine Tau recombinante car ils sont capables d'induire la formation d'agrégats à partir de monomères de protéine Tau recombinante de manière plus rapide et robuste. Ces données suggèrent que ces oligomères *in vivo* issus de cerveaux humains MA, sont porteurs de PTM ou de co-facteurs, ou encore présentent une conformation pathologique stable qu'on ne retrouve pas dans les oligomères *in vitro* (Lasagna-Reeves, Castillo-Carranza et al. 2012). Sur base de ces résultats, il a été suggéré que la formation des PHF pourrait avoir un rôle protecteur en débarrassant la cellule des formes oligomériques plus toxiques mais cette théorie fait encore débat (Bretteville and Planel 2008; Congdon and Duff 2008).

Au niveau moléculaire, la région de Tau impliquée dans la formation des PHF *in vitro* est le domaine à segments répétés (Wille, Drewes et al. 1992) qui constitue le cœur des PHF (Wischik, Novak et al. 1988; von Bergen, Barghorn et al. 2006). De manière intéressante, cette région coïncide avec celle impliquée dans la liaison des microtubules suggérant un recouvrement des fonctions physiologiques et pathologiques de Tau. Deux courtes séquences peptidiques constituées de six acides aminés chacune, l'une au début du deuxième *repeat* appelée PHF6\* (<sup>275</sup>VQIINK<sup>280</sup>), l'autre au début du troisième *repeat* appelée PHF6 (<sup>306</sup>VQIVYK<sup>311</sup>) sont étroitement liées au processus d'agrégation (von Bergen, Friedhoff et al. 2000; von Bergen, Barghorn et al. 2001). Ces deux peptides PHF6\* et PHF6 sont capables de s'assembler de façon autonome *in vitro*.

La distribution des différentes isoformes de Tau est un facteur qui a également été examiné dans la formation des PHF, en particulier le ratio Tau3R : Tau4R. Dans le cerveau adulte sain, ce ratio est de 1 :1. Puisque les isoformes 4R sont prédominantes dans les NFT, il est possible qu'une dérégulation de la balance entre les deux groupes d'isoformes en soit l'une des causes. Par exemple, le ratio Tau 3R :4R est modifié dans la FTDP-17 sous l'effet de différentes mutations du gène *tau* (Hong, Zhukareva et al. 1998; Hutton, Lendon et al. 1998). Dans les formes sporadiques de la MA où aucune mutation de *tau* n'a été détectée, l'étiologie de la pathologie reste encore indéterminée et d'autres facteurs interviennent. Une dérégulation de la machinerie impliquée dans l'épissage alternatif de Tau pourrait modifier le ratio 3R :4R dans la MA (Glatz, Rujescu et al. 2006). D'autres mutations FTDP-17, localisées dans le domaine à segments répétés (P301L, P301S, R406W, V337M) altèrent plutôt la liaison de Tau aux microtubules et augmentent significativement la vitesse de formation des PHF *in vitro* (Nacharaju, Lewis et al. 1999). Sur base de ces observations, un modèle de souris MA transgénique exprimant la mutation P301L a été construit et constitue, à l'heure actuelle, l'un des meilleurs modèles d'étude (Lewis, McGowan et al. 2000). Ces souris développent, avec l'âge, des NFT et d'autres lésions caractéristiques de la MA, et présentent également des déficiences cognitives et motrices, associées aux pertes neuronales. De plus, ces événements ne sont pas associés à un niveau d'expression anormalement élevé de Tau.

Par ailleurs, les PTM jouent un rôle central dans la physiopathologie de la protéine Tau. De nombreuses PTM, de différentes natures, sont représentées au sein de la protéine, en raison probablement de la structure désordonnée de Tau qui la rend flexible et facilement accessible (Pevalova, Filipcik et al. 2006). Elles sont presque toujours liées à la pathologie Tau comme la phosphorylation, la *N*-glycosylation, l'ubiquitination (Morishima-Kawashima, Hasegawa et al. 1993), la sumoylation (Hatakeyama, Matsumoto et al. 2004; Dorval and Fraser 2006), l'oxydation (Schweers, Mandelkow et al. 1995), l'acétylation (Min, Cho et al. 2010; Cohen, Guo et al. 2011), l'*O*-GlcNAcylation (Arnold, Johnson et al. 1996). La plus étudiée est, de loin, la phosphorylation car elle joue également un rôle dans la régulation de la liaison de Tau aux microtubules (paragraphe 1.1.3). Elle tient également une place prépondérante dans la pathologie car, historiquement, c'est la première PTM qui a été associée aux PHF (Grundke-Iqbal, Iqbal et al. 1986; Wood, Mirra et al. 1986) et on pense aujourd'hui encore, que c'est l'augmentation de la phosphorylation qui favorise, de manière prépondérante, l'agrégation pathologique selon un mécanisme encore mal défini. Il faut noter que la proportion des PTM est plus importante dans le domaine de projection et la région C-terminale que dans le domaine de liaison aux microtubules, ce qui peut s'expliquer par une accessibilité moindre de ce dernier lorsqu'il est lié aux microtubules (dans des conditions

physiologiques, étant données les concentrations relatives de Tau et de tubuline, et la constante de dissociation estimée, toute la protéine Tau serait sous forme liée dans le neurone). La phosphorylation, par exemple, est plus importante dans la région riche en prolines, car celle-ci contient de nombreux motifs Ser/Thr-Pro ciblés par des kinases spécifiques. Cependant, certaines PTM, comme l'acétylation, ciblent davantage le domaine de liaison aux microtubules, en lien avec le caractère basique de cette région.

### ***1.1.3. Phosphorylation de la protéine Tau : équilibre entre effet physiologique et pathologique***

La phosphorylation joue un rôle crucial dans la régulation de la liaison de Tau aux microtubules et dans sa capacité à stimuler la polymérisation de la tubuline. De plus, dans le contexte pathologique, la protéine Tau se présente sous forme de fibres insolubles, anormalement phosphorylées, appelées NFT, qui constituent un des marqueurs moléculaires de la maladie d'Alzheimer. La phosphorylation est donc un processus central dans la régulation physiopathologique de Tau. Dans le cerveau normal, la protéine Tau contient 2 à 3 moles de phosphate par mole de protéine alors que dans la MA, la stœchiométrie est au moins trois fois plus importante : elle est dite hyperphosphorylée (Morishima-Kawashima, Hasegawa et al. 1995; Morishima-Kawashima, Hasegawa et al. 1995). L'hyperphosphorylation se rapporte à la fois au niveau de phosphorylation des différents sites, dont le rôle est notamment de moduler l'affinité de Tau pour les microtubules, et également à l'apparition de phosphorylations dites pathologiques, sites qu'on ne trouve qu'associés aux PHF. L'augmentation du niveau de phosphorylation, en diminuant la capacité d'interaction avec les microtubules, favorise la déstabilisation du réseau microtubulaire (Alonso, Zaidi et al. 1994). L'autre effet est d'augmenter le réservoir de protéines Tau libres, susceptible de favoriser l'association des protéines Tau entre elles pour former des oligomères. Ainsi, la phosphorylation semble jouer un rôle décisif dans la pathologie Tau (Gotz, Gladbach et al. 2010). Preuve en est que l'évolution de la maladie peut être suivie dans des coupes de cerveau à l'aide de l'anticorps monoclonal AT8, révélant l'apparition progressive de Tau phosphorylée en Ser202 et Thr205 (Braak, Alafuzoff et al. 2006). Le rôle spécifique de ces phosphorylations multiples (Tau contient 80 sites potentiels de phosphorylation, soit 20% de la séquence), à savoir si elles sont la cause ou la conséquence des dérèglements liés à la tauopathie reste à élucider. Plusieurs kinases et phosphatases sont bien connues pour modifier Tau mais pas la cascade d'événements moléculaires qui conduit à la formation de Tau hyperphosphorylée. Mieux comprendre cette séquence d'événements aura pour but à long terme de trouver des molécules qui pourraient bloquer la formation d'espèce toxiques hyperphosphorylées de Tau dans les neurones, à un stade précoce de la maladie.

Bien que cette phosphorylation soit un processus cellulaire normal, la protéine Tau hyperphosphorylée ne serait plus capable de se lier aux microtubules, ce qui pourrait être à l'origine de l'agrégation de la protéine sous forme de PHF (Alonso, Grundke-Iqbal et al. 1996; Merrick, Trojanowski et al. 1997; Alonso, Zaidi et al. 2001). L'état de phosphorylation de Tau est majoritairement caractérisé en utilisant des anticorps spécifiques des sites phosphorylés. Cependant, ces anticorps ne sont pas disponibles pour tous les sites, et l'aspect quantitatif de l'immunochimie reste à confirmer. La caractérisation des formes hyperphosphorylées de Tau est dès lors compromise par ces limitations. Tau, dans les PHF, est phosphorylée sur 40 sites dont la plupart sont localisés

dans les régions encadrant le domaine à segments répétés (Morishima-Kawashima, Hasegawa et al. 1995; Morishima-Kawashima, Hasegawa et al. 1995; Johnson and Stoothoff 2004). Différents sites de phosphorylation sont impliqués dans la modulation de la liaison de Tau aux microtubules. Certains sites, localisés dans le domaine à segments répétés diminuent fortement la liaison de Tau aux microtubules comme la phospho-Ser262, générée par la kinase MARK, dans le motif KxGS du premier *repeat* (Biernat, Gustke et al. 1993; Augustinack, Schneider et al. 2002). D'autres se trouvent dans le domaine riche en prolines, qui influence indirectement la liaison de Tau aux microtubules, comme la phospho-Ser214, phosphorylée par PKA (*cAMP-dependent protein kinase A*) (Jicha, Weaver et al. 1999). Bien que la phosphorylation de Ser262 et Ser214 détache Tau des microtubules, elle protège également de l'agrégation (Schneider, Biernat et al. 1999), ce qui souligne la difficulté d'établir un lien clair entre phosphorylation, diminution de la liaison aux microtubules et agrégation.

D'autres kinases dites *proline-directed* sont spécifiques des sites Ser/Thr-Pro. Il existe 17 motifs Ser/Thr-Pro au sein de l'isoforme la plus longue de Tau (Tau441) qui sont principalement localisés dans les régions flanquant le domaine de liaison aux microtubules, en particulier le domaine riche en prolines. Les protéines GSK3 $\beta$  et Cdk5 sont des kinases majeures impliquées dans la pathologie Tau car elles phosphorylent Tau sur un grand nombre de sites dont la plupart sont communs aux deux kinases. Les niveaux d'expression de ces deux kinases sont élevés dans le cerveau et elles ont été toutes deux étroitement impliquées dans l'hyperphosphorylation et la formation des NFT dans la MA. Une surexpression de GSK3 $\beta$  conduit à l'hyperphosphorylation de Tau dans des souris transgéniques et dans des cellules en culture. Cdk5 doit être activé par les protéines p35 ou p25, le produit de protéolyse de p35 par la calpaïne. Une surexpression de p25 induit l'hyperphosphorylation de Tau dans des souris transgéniques (Ahlijanian, Barrezueta et al. 2000; Noble, Olm et al. 2003). Comme pour GSK3 $\beta$  et Cdk5, les kinases ERK, de la famille des MAPK (*mitogen-activated protein kinases*), et JNK, de la famille des *stress-activated kinases*, phosphorylent Tau au niveau de sites associés à l'hyperphosphorylation anormale de Tau dans les PHF. Parmi les sites de phosphorylation associés à la pathologie, on trouve entre autres les sites pS202/pT205 reconnus par l'anticorps AT8, pT212/pS214, reconnus par AT100, pT231/pS235 reconnus par AT180, pS396/pS404 reconnus par PHF-1, pS422 reconnus par l'anticorps du même nom. Une corrélation a été observée dans les cerveaux entre l'immunodétection spécifique par ces différents anticorps et l'évolution de la pathologie Alzheimer (Augustinack, Schneider et al. 2002). Certains phospho-épitopes sont obtenus par des phosphorylations séquentielles mettant en jeu différentes kinases, comme AT180 et PHF-1. GSK3 $\beta$  fonctionne sur ce mode : elle requiert un *priming*, c'est-à-dire la phosphorylation préalable par d'autres kinases, avant la phosphorylation séquentielle de motifs S/TXXXpS (Fiol, Mahrenholz et al. 1987; Fiol, Haseman et al. 1988; Fiol, Wang et al. 1990; Liu, Zhang et al. 2004; Leroy, Landrieu et al. 2010). Contrairement aux phosphorylations des résidus Ser262 et Ser214, la phosphorylation d'un unique site dans les motifs Ser/Thr-Pro a peu d'effet sur la liaison de Tau aux microtubules. Néanmoins, la multiplication des phosphorylations, en particulier dans la région riche en prolines, réduit fortement la capacité de liaison de Tau (Sergeant, Bretteville et al. 2008). Il faut aussi souligner que la phosphorylation physiologique ou pathologique peut être modulée par l'interaction de Tau avec d'autres protéines comme les microtubules ou par l'association des protéines Tau entre elles lors de la formation des PHF, rendant certains sites inaccessibles à la phosphorylation ou à la déphosphorylation. Parallèlement à l'activation des kinases, une diminution de l'expression et de l'activité de PP2A, la phosphatase majeure de Tau, a été décrite dans la MA, contribuant ainsi à



l'hyperphosphorylation anormale de Tau (Gong, Shaikh et al. 1995; Sontag, Hladik et al. 2004; Liu, Grundke-Iqbal et al. 2005; Rudrabhatla and Pant 2011). De plus, PP2A est une phosphatase majeure dans le cerveau (l'activité de PP2A compte pour 70% de l'activité phosphatase). Elle est responsable de l'activation de plusieurs kinases dans le cerveau. Ainsi PP2A stimule indirectement la phosphorylation de Tau en favorisant l'activité de PKA, de MAPKK ou de ERK.

D'autre part, il a été montré que la phosphorylation mitotique de Tau s'apparente au profil de phosphorylation observé dans la MA (Preuss and Mandelkow 1998). En effet, le niveau de phosphorylation de Tau varie au cours du cycle cellulaire en réponse au besoin de modulation du fuseau mitotique au cours des différentes étapes, passant par des phases de polymérisation et de dépolymérisation des microtubules. De même, le niveau de phosphorylation de Tau varie au cours du développement (la forme fœtale de Tau est hautement phosphorylée par rapport à la forme adulte normale) ou en fonction des conditions (stress thermique, stress oxydant, anesthésie). De plus, une expression anormale de marqueurs du cycle cellulaire, comme la cycline D1 ou le complexe Cdc2/cyclineB1, a été observée dans les neurones de cerveaux atteints de la MA ou dans un modèle de neuroblastomes SH-SY5Y exprimant p25 de façon inductible. Une réactivité forte avec l'anticorps mitotique MPM-2 (Davis, Tsao et al. 1983) des NFT et des neurones malades suggère une réactivation des kinases mitotiques dans la MA. Il a été proposé qu'en réponse à des signaux inappropriés, une réactivation anormale du cycle cellulaire dans les neurones matures pourrait être à l'origine de la neurodégénérescence conduisant, entre autres, à l'hyperphosphorylation de Tau observée dans la MA (Vincent, Rosado et al. 1996; Vincent, Zheng et al. 1998; Delobel, Flament et al. 2002; Hamdane, Delobel et al. 2003; Hamdane, Sambo et al. 2003; Schindowski, Belarbi et al. 2008).

La phosphorylation est médiateur d'interactions phospho-dépendantes aux niveaux physiologique et pathologique. On citera par exemple, les interactions avec les peptidyl-prolyl cis/trans isomérases (PPIases) Pin1 (Lu, Wulf et al. 1999) et FKBP52 (Chambraud, Sardin et al. 2010) ou la protéine 14-3-3 (Hashiguchi, Sobue et al. 2000), qui jouent un rôle sur la localisation, la stabilité, la phosphorylation de Tau et régulent ses interactions avec d'autres partenaires. L'interaction de Tau phosphorylée avec la protéine régulatrice Pin1 permettrait de restaurer la fonction de Tau hyperphosphorylée, piégée dans des agrégats neurofibrillaires dans la pathologie Alzheimer, et serait donc un événement permettant à la cellule neuronale de récupérer une protéine Tau fonctionnelle. Ainsi, il semble que Pin1, par son affinité pour Tau hyperphosphorylée, soit séquestrée dans les PHF et ne pourrait plus exercer sa fonction cellulaire normale (Lu, Wulf et al. 1999). Des souris Pin1 *knockout* (Pin1<sup>-/-</sup>) développent une tauopathie liée à l'âge, avec la formation de NFT composés de protéines Tau hyperphosphorylées, accompagnée de déficits cognitifs et moteurs. De plus, Pin1 favorise la déphosphorylation de Tau par la phosphatase PP2A (Zhou, Kops et al. 2000). Il semble donc que Pin1 ait un rôle neuroprotecteur qui serait compromis par la ségrégation de Pin1 dans les PHF (Liou, Sun et al. 2003). En revanche, Pin1 aurait un rôle opposé dans des souris transgéniques exprimant la protéine Tau mutée P301L, typiquement utilisées comme modèle de la MA, où il a été montré que Pin1 exacerbe la pathologie Tau et la dégénérescence neuronale (Lim, Balastik et al. 2008). Plus récemment, l'interaction de Tau hyperphosphorylée avec FKBP52, membre d'une autre famille de PPIases, a été montrée avec un rôle potentiel dans la MA. Mais, au contraire de Pin1, FKBP52 semble exercer une activité "anti-Tau". FKBP52 inhibe l'assemblage des microtubules par la protéine Tau *in vitro*. La surexpression de FKBP52 dans des cellules PC12 en culture empêche l'accumulation de Tau

et réduit la longueur des neurites (Chambraud, Sardin et al. 2010). Bien que présente à un niveau d'expression anormalement bas dans les cerveaux de patients atteints de la maladie d'Alzheimer ou de FTDP-17, FKBP52 n'est pas séquestrée dans les fibres pathologiques (Giustiniani, Sineus et al. 2012). L'association de FKBP52 avec la chaperonne Hsp90, dont le rôle dans la dégradation des espèces phosphorylées de Tau via le protéasome a été décrit (Dickey, Kamal et al. 2007), pourrait être impliquée dans l'inhibition de l'accumulation et de la fonction de Tau dans les neurones.

#### **1.1.4. Autres modifications post-traductionnelles dans le processus d'agrégation de la protéine Tau**

De nombreuses hypothèses ont été proposées pour expliquer le rôle de Tau dans la pathogenèse de la MA, notamment le lien entre l'accumulation de formes hyperphosphorylées et l'agrégation. Nous venons de décrire comment l'activation anormale des kinases et/ou le défaut d'activité phosphatase pourrait mener aux formes hyperphosphorylées. Parmi les différents mécanismes potentiellement impliqués dans l'agrégation ou dans l'accumulation de protéines Tau hyperphosphorylées, une altération de la conformation de Tau soluble a été suggérée. Des anticorps conformationnels ont été développés pour détecter ces conformations anormales qui précèdent la formation des PHFs et qui sont aussi retrouvées dans les PHF. Ces anticorps, générés par repliement du domaine N-terminal sur les *repeats*, reconnaissent les premiers stades de la pathologie Alzheimer (Carmel, Mager et al. 1996; Jicha, Bowser et al. 1997; Jicha, Berenfeld et al. 1999). Ces épitopes ne sont pas trouvés dans les cerveaux sains. Les anticorps Alz50 et MC1 reconnaissent un épitope discontinu localisé à la fois dans la région N-terminale de Tau et dans la région des *repeats*. D'abord identifiée par FRET puis vérifiée par les expériences RMN de relaxation paramagnétique (PRE), cette conformation anormale, dite *paperclip-like* existe déjà dans la protéine Tau soluble de manière transitoire, et serait exacerbée dans les formes oligomériques ou dans les PHF. Dans cette conformation, les régions N- et C-terminales se replient à proximité du domaine à segments répétés (Jeganathan, von Bergen et al. 2006; Bibow, Mukrasch et al. 2011). De même, des études structurales sur des formes phosphorylées ou mimant la phosphorylation de Tau ont montré des modifications conformationnelles, avec des résultats contradictoires selon les sites et le nombre de phosphorylations (Bielska and Zondlo 2006; Jeganathan, Hascher et al. 2008; Bibow, Ozenne et al. 2011). De manière générale, il semble qu'une modification des réseaux d'interactions transitoires à longue distance est une propriété commune à un certain nombre de protéines désordonnées impliquées dans des processus d'agrégation.

Parmi les causes possibles de l'agrégation de Tau sous forme de PHF, il a été proposé une dérégulation de son ubiquitination et un défaut de dégradation par les chaperonnes moléculaires. Tau et, plus particulièrement, sa forme phosphorylée sont dégradées par le protéasome (Petrucci, Dickson et al. 2004; Tan, Wong et al. 2008). Des protéines chaperonnes, comme Hsp70 ou Hsp90, et leur co-chaperonnes, comme le domaine C-terminal de l'*Hsc70-interacting protein* (CHIP), interagissent avec Tau et régulent le niveau de Tau soluble dans les neurones. Elles jouent ainsi un rôle dans la prévention de l'agrégation de manière phospho-dépendante. De manière intéressante, il a été mis en évidence que la région de Tau engagée dans l'interaction avec les protéines de choc thermique correspond aux peptides formant des structures en feuillets  $\beta$ , impliqués dans l'agrégation (Sarkar, Kuret et al. 2008; Tortosa, Santa-Maria et al. 2009). La protéine chaperonne Hsp90 joue un rôle décisionnaire dans le tri des protéines, en reconnaissant les protéines mal conformées et en les

orientant soit vers la voie du *refolding*, soit vers la dégradation en s'associant avec la co-chaperonne CHIP (Connell, Ballinger et al. 2001). Il a été montré que CHIP est la E3 ubiquitine-ligase spécifique de Tau, qui contrôle la dégradation de Tau *in vitro*. CHIP peut dégrader Tau dans des conditions physiologiques indépendamment de son état de phosphorylation. Dans des conditions pathologiques, CHIP assure la dégradation de la protéine Tau hyperphosphorylée induite par suractivation de GSK3 ou inhibition de PP2A, dans des cerveaux de rats ou des cellules en culture. Ainsi, une surexpression de CHIP pourrait contrebalancer l'accumulation de Tau dans les cerveaux MA (Zhang, Xu et al. 2008). D'autre part, l'inhibition d'Hsp90 favorise une diminution sélective des espèces phosphorylées de Tau dans des souris transgéniques présentant une tauopathie, suggérant un rôle comme cible thérapeutique potentielle dans le traitement des tauopathies (Dickey, Kamal et al. 2007).

Il a été observé la présence de protéines Tau tronquées dans les NFTs. Ce clivage peut résulter de la protéolyse par les caspases, caspase-3 ou caspase-6, aux acides aspartiques 421 et 402 (Guo, Albrecht et al. 2004; Rissman, Poon et al. 2004; Albrecht, Bourdeau et al. 2007) localisés du côté C-terminal du domaine de liaison aux microtubules. Il a été démontré que la protéine Tau clivée au niveau de son extrémité C-terminale agrège plus rapidement que la protéine Tau entière *in vitro* (Gamblin, Chen et al. 2003). Caspase-6 est une cystéinyl protéase qui fait partie des protéases exécutrices de l'apoptose cellulaire. Caspase-6 induit une mort neuronale différée dans des cultures neuronales primaires où elle est activée. L'activation de caspase-6 ne conduit pas à la mort neuronale mais à un processus de neurodégénération, compatible avec une implication de la caspase-6 dans l'apparition des altérations neuropathologiques caractéristiques de la MA (Klaiman, Petzke et al. 2008; Klaiman, Champagne et al. 2009; Nikolaev, McLaughlin et al. 2009). Une étude d'imagerie *in vivo* de cerveaux de souris indique que l'activation transitoire de caspases dans un neurone est corrélée avec la formation de NFTs. Une activation des caspases-3 et 6, ainsi que leur colocalisation avec la protéine Tau tronquée permet de les impliquer dans la formation de NFTs. Bien que les caspases soient normalement des enzymes impliquées dans la mort cellulaire, leur activation suivie de la formation de NFTs ne semble pas induire la mort neuronale. L'hypothèse selon laquelle le clivage de Tau précède la formation des NFTs a été vérifiée en surexprimant Tau tronquée en D421 chez la souris. Les neurones deviennent réactifs aux anticorps Alz50, PHF-1 et AT8, spécifiques des PHFs. Ainsi, la présence de Tau tronquée induit des modifications de la phosphorylation et de la conformation de Tau caractéristiques des NFTs (de Calignon, Fox et al. 2010). Tau contient encore d'autres sites de coupure accessibles à différentes protéases, comme la calpaïne par exemple, qui génèrent des courts fragments qui exercent potentiellement leur action cytotoxique via la voie A $\beta$  (Canu and Calissano 2003; Park and Ferreira 2005). De manière générale, les formes tronquées de Tau au niveau des N- et C-termini, par les caspases ou d'autres protéases, constituent des points de nucléation qui peuvent recruter diverses formes de Tau, dont des formes intactes, et promouvoir la formation des fibres *in vivo* (Binder, Guillozet-Bongaarts et al. 2005).

Il existe encore bien d'autres mécanismes impliqués dans le processus d'agrégation et parmi ceux-ci, différentes PTM sont impliquées dans la pathologie Tau. Par exemple, il a été montré que la glycosylation est une PTM anormale de la protéine Tau et qu'elle est essentiellement associée aux processus pathologiques dans la MA. Les N-glycannes sont notamment impliqués dans le maintien de l'hélicité des PHF (Wang, Grundke-Iqbal et al. 1996). De plus, il a été montré que la glycosylation de

Tau précède et favorise l'hyperphosphorylation (Sato, Naito et al. 2001; Liu, Iqbal et al. 2002; Liu, Zaidi et al. 2002; Liu, Zaidi et al. 2002). L'oxydation des résidus cystéines (au nombre de deux, en positions 291 et 322, dans l'isoforme la plus longue) est également une modification qui apparaît comme essentielle dans le processus d'agrégation car elle serait l'étape préliminaire à la formation des espèces oligomériques (Schweers, Mandelkow et al. 1995). Il faut souligner que l'épissage de l'ARNm module la formation et la nature des espèces oxydées puisque, des deux cystéines présentes dans les motifs KCGS, l'une (C322) se trouve dans le *repeat* R3 tandis que l'autre (C291) se trouve dans le R2, qui n'est présent que dans trois isoformes sur six.

L'acétylation et l'O-GlcNAcylation sont des PTM décrites récemment, qui jouent également un rôle dans le processus d'agrégation de manière directe ou indirecte. L'acétylation inhibe la liaison de Tau à la tubuline, favorise l'agrégation et prévient la dégradation de Tau par le protéasome, contribuant à l'accumulation de formes anormalement phosphorylées. A l'inverse, l'O-GlcNAcylation jouerait plutôt un rôle protecteur de l'agrégation en diminuant la phosphorylation de Tau et en inhibant la formation des PHF de manière indépendante de la phosphorylation. Ces deux PTM seront discutées plus en détails dans le Chapitre 4.

#### **1.1.5. Tau comme cible thérapeutique dans la maladie d'Alzheimer**

Il n'existe pour le moment pas de traitement curatif de la MA, mais uniquement des traitements qui ralentissent sa progression et améliorent, pour un temps, les symptômes. Des tests menés jusqu'aux essais cliniques de phase III n'ont pas abouti. La cible était la formation des plaques amyloïdes avec, notamment la mise au point d'un vaccin. L'un des problèmes actuels est de diagnostiquer plus précocement la maladie avant que les pertes neuronales ne soient trop importantes, avec comme avantage supplémentaire, la possibilité de mieux évaluer les nouvelles approches thérapeutiques. Dans la stratégie d'élimination des plaques amyloïdes par immunothérapie dans les cerveaux atteints de la MA, il n'y a pas eu de gain significatif de fonction cognitive et la progression de la maladie n'a pas été ralentie malgré une nette diminution des plaques (Citron 2010). La protéine Tau est à présent considérée comme une cible thérapeutique d'intérêt (Brunden, Trojanowski et al. 2009). L'hypothèse que Tau pourrait jouer un rôle central dans le mécanisme physiopathologique des démences séniles est en effet de plus en plus supportée. Reconnue à présent comme un composant majeur responsable des fonctions synaptiques altérées et de la destruction irréversible des neurones (Dawson, Cantillana et al. 2010), Tau, ses modifications anormales et leurs conséquences pathologiques doivent être étudiées si on souhaite parvenir à contrôler les mécanismes de pathogenèse (Ballatore, Lee et al. 2007; Brunden, Trojanowski et al. 2008).

Un rôle exécutif de la protéine Tau hyperphosphorylée est reconnu dans la neurodégénération, basé sur la corrélation entre l'envahissement par des agrégats de Tau dans les neurones et l'aggravation des symptômes de la maladie. Les régions du cerveau dont la fonction est altérée sont ainsi envahies par ces agrégats. Il existe en outre des formes familiales de tauopathies qui sont liées à des mutations dans le gène codant la protéine Tau, comme la FTDP-17. C'est une preuve directe qu'une altération de Tau peut conduire à la neurodégénération. D'autre part, la protéine Tau hyperphosphorylée pourrait participer à la progression de la maladie dans les différentes zones du cerveau en étant transmise aux neurones voisins (Clavaguera, Bolmont et al. 2009). L'effet toxique de Tau pourrait provenir non seulement de son agrégation mais aussi de son accumulation (Brunden,

Trojanowski et al. 2008). Cette hypothèse est étayée par plusieurs études qui démontrent qu'en réduisant la quantité cellulaire de Tau, on peut diminuer la toxicité des fibres amyloïdes A $\beta$  (Roberson, Scarce-Levie et al. 2007; Ittner and Gotz 2011). Les travaux les plus récents ont en effet démontré que la fonction de Tau ne se limite pas à la stabilisation des microtubules mais que Tau peut jouer un rôle de protéine plate-forme ("*scaffold*") intervenant ainsi dans une cascade de signalisation qui relaie la toxicité des fibres amyloïdes (Ittner, Ke et al. 2010). De plus, Tau nucléaire protège l'intégrité de l'ADN neuronal en condition de stress (Sultan, Nesslany et al. 2011). Une stratégie immunothérapeutique ciblant la protéine Tau est désormais envisagée avec l'utilisation d'oligomères solubles de Tau ou pré-filaments (Kayed and Jackson 2009) puisque de nombreuses études ont montré la prévalence du rôle pathogénique des formes oligomériques par rapport aux protéines Tau solubles et aux fibres (Santacruz, Lewis et al. 2005; Roberson, Scarce-Levie et al. 2007).

Différentes stratégies thérapeutiques ont été envisagées avec, pour cible, la protéine Tau. En général, la protéine Tau est visée de manière indirecte, sauf dans le cas des inhibiteurs de l'agrégation. Pour cause, l'absence de structure qui empêche la conception de molécules bioactives basées sur une complémentarité de forme et une liaison optimale dans la structure tridimensionnelle d'un site actif ou de tout autre site allostérique. Dans les cas des IDP, d'autres stratégies doivent être employées. Cependant, un aperçu de la structure moléculaire des fibres pourrait mener à des stratégies plus rationnelles dans la conception des molécules antiagrégantes (Landau, Sawaya et al. 2011). Des inhibiteurs de l'agrégation ont été initialement criblés avec des méthodes à haut débit. Des molécules de type colorants ont été ainsi identifiées avec, comme chef de file, le bleu de méthylène, qui présente des effets thérapeutiques positifs et est aujourd'hui en tests cliniques de phase III. D'autres composés de la famille des anthraquinones et des rhodanines montrent non seulement des effets inhibiteurs de l'agrégation mais sont également capables de dissoudre de fibres pré-formées. D'autres familles de composés comme les polyphénols ont été identifiées plus récemment. Tous ces composés ont été évalués dans des tests d'agrégation *in vitro* basés sur l'amorçage de l'agrégation par des polyanions comme l'héparine, et le suivi de l'émission de fluorescence de la Thioflavine S ou T pour détecter la formation de structures en feuillet  $\beta$  auxquelles se lie spécifiquement la Thioflavine. Cependant, il a été mis en évidence que certains composés empêchent la formation des fibres via la formation de peroxydes avec le dithiothréitol (DTT) utilisé dans les tests *in vitro* et sont inactifs en absence de DTT, ce qui soulève la question de leur pertinence dans un contexte *in vivo*.

D'autres stratégies ont pour cibles des partenaires de Tau impliqués dans la modification de la protéine comme les inhibiteurs l'hyperphosphorylation : inhibiteurs de kinases et de l'O-GlcNAc hydrolase (OGA) ou activateurs de phosphatases. Il a été montré qu'il faut la contribution de plusieurs combinaisons de kinases pour conduire à l'hyperphosphorylation de Tau alors que la phosphatase PP2A à elle seule peut réguler la phosphorylation (Iqbal, Liu et al. 2009). Les kinases GSK3 $\beta$ , Cdk5 et MARK ont été envisagées comme des cibles thérapeutiques d'intérêt car ces kinases sont étroitement liées à la phosphorylation anormale de Tau dans la MA. L'inhibition spécifique des kinases n'est pas simple en termes de conception de molécules capables d'inhiber sélectivement un type de kinases. En effet, la plupart des inhibiteurs actuels ciblent le site de liaison à l'ATP partagé par l'ensemble des kinases, posant des problèmes de sélectivité. De plus, les conséquences de cette

inhibition restent également incertaines. Par exemple, il existe un lien entre les activités de Cdk5 et GSK3 $\beta$  (Wen, Planel et al. 2008). Cdk5 inhibe GSK3 $\beta$  par phosphorylation de la Ser9 et il a été observé, dans des modèles souris transgéniques surexprimant p25, qu'une inhibition de Cdk5 entraîne une élévation de la phosphorylation de Tau par GSK3 $\beta$ . Chez des souris traitées par le chlorure de lithium, un inhibiteur de GSK3 $\beta$ , on observe une réduction de la phosphorylation de Tau et des agrégats insolubles, et moins de dégénérescence axonale (Noble, Planel et al. 2005). Cependant, l'inhibition de ces kinases peut avoir des effets néfastes sur les autres substrats. Cdk5 est essentielle chez la souris et joue un rôle primordial dans le développement neuronal. GSK3 $\beta$  intervient dans le métabolisme du glycogène et est également une cible thérapeutique dans les désordres métaboliques associés. L'utilisation d'inhibiteurs de kinases a permis de montrer, dans des souris transgéniques modèles de tauopathies, l'implication de ces kinases dans la pathologie Tau. De manière générale, l'administration de ces inhibiteurs permet de réduire la phosphorylation de Tau et le niveau de protéines insolubles. Leur utilisation précoce permet de retarder le développement de la maladie et prévient la formation des NFT mais une administration plus tardive, lorsque la pathologie Tau est déjà installée, accompagnée de déficits moteurs et cognitifs, diminue la phosphorylation de Tau sans diminuer la charge en NFT, soulignant l'intérêt d'établir un diagnostic précoce. Une autre approche pour réduire la phosphorylation de Tau met à profit la relation antagoniste entre phosphorylation et O-GlcNAcylation. Un inhibiteur de l'O-GlcNAc hydrolase (OGA), le Thiamet-G, permet de réduire la phosphorylation spécifique des résidus Thr231, Ser262 et Ser396 chez des souris traitées *per os*.

La phosphorylation anormale qui diminue fortement la capacité de liaison de Tau aux microtubules, augmentant ainsi le réservoir de Tau libre, contribue pour une part importante à la formation des agrégats insolubles. Pour preuves, des mutations de Tau dans certaines tauopathies affectent directement la liaison aux microtubules suggérant qu'une déficience dans la fonction des microtubules et du transport axonal contribue à la neurodégénération. Ainsi, des stratégies visant à compenser la perte de fonction de Tau par stabilisation des microtubules constituent une approche alternative. Des souris transgéniques Tau *knock-out* présentent des déficits moteurs et une réduction de la densité microtubulaire. Traitées avec le paclitaxel, une molécule stabilisant les microtubules, ces souris montrent une amélioration du transport axonal rapide et de la fonction motrice, accompagnée d'une augmentation de la densité des microtubules, indiquant qu'une compensation de la perte de fonction de Tau est possible par les agents stabilisant les microtubules. De telles molécules (paclitaxel, épothilones), avec une faible capacité à passer la barrière hémato-encéphalique, sont utilisées comme agents anti-mitotiques avec un potentiel de molécules anti-cancéreuses, d'où l'intérêt de nouvelles générations de molécules présentant une pénétration plus importante pour traiter la MA, permettant de limiter les effets secondaires liés à une concentration élevée au niveau périphérique.

Une autre approche consiste à favoriser la clairance des agrégats de Tau via le protéasome avec des inhibiteurs de la chaperonne Hsp90. La destruction des protéines se fait soit par dégradation protéolytique, par le système lysosomal ou par le système ubiquitine-protéasome. Il a été proposé que les voies du catabolisme soient affectées dans le cerveau de patients atteints de la MA et il semble que ces trois voies soient affectées concernant la dégradation de Tau. Les protéases ne sont plus capables de cliver Tau à l'intérieur des PHF. Le système ubiquitine-protéasome nécessite la

liaison covalente d'une ou plusieurs molécules d'ubiquitine qui permet la dégradation de la protéine-cible par le complexe protéasome. Mais son ouverture trop étroite pour les protéines multimériques empêche vraisemblablement l'utilisation de cette voie pour la dégradation des PHF, les orientant préférentiellement vers la macro-autophagie. Il a été montré que les protéines Tau phosphorylées ou présentant une conformation anormale sont reconnues par la protéine de choc thermique Hsp90. Ces chaperonnes moléculaires reconnaissent les protéines mal conformées et les stabilisent pour aider à leur *refolding* selon un mécanisme ATP-dépendant. Il a été montré qu'une inhibition de la fonction ATPase de Hsp90 les détourne de la voie du *refolding* pour les orienter vers la dégradation par le protéasome. Dans des cellules surexprimant une protéine Tau mutée, les inhibiteurs de Hsp90 réduisent la phosphorylation spécifique de plusieurs motifs Ser/Thr-Pro (Ser202, Thr205, Ser296, Ser404) et diminuent le niveau de protéines présentant une conformation anormale. De manière intéressante, certains inhibiteurs de Hsp90 présentent une affinité plus importante pour les complexes impliquant des protéines mal conformées permettant une utilisation à des doses qui laisseraient les autres protéines-clients de Hsp90 intactes.

Enfin des stratégies utilisant des inhibiteurs de HAT ou de peptidyl-prolyl *cis/trans* isomérases (PPlases) sont encore peu explorées. La modulation des interactions protéine-protéine phospho-spécifiques pourrait s'avérer une voie thérapeutique originale, jusqu'ici peu explorée. Il existe, par exemple, des inhibiteurs de la PPlase phospho-spécifique Pin1 (Zhang, Fussel et al. 2002; Uchida, Takamiya et al. 2003; Wildemann, Erdmann et al. 2006; Zhang, Daum et al. 2007; Guo, Hou et al. 2009; Dong, Marakovits et al. 2010; Namanja, Wang et al. 2010; Duncan, Dempsey et al. 2011; Urusova, Shim et al. 2011) conçus dans la recherche de molécules anti-cancéreuses, en lien avec le rôle de Pin1 dans la régulation du cycle cellulaire, et sa surexpression dans de nombreux cancers (Lu 2003; Lu, Suizu et al. 2006). De tels composés n'ont pas encore été évalués dans des modèles de la MA mais pourraient être utilisés pour moduler le rôle de Pin1 dans ce contexte (Hamdane, Smet et al. 2002; Lim and Lu 2005).

## 1.2 ETUDE DE LA PROTEINE TAU PAR SPECTROSCOPIE RMN

### 1.2.1. Généralités sur les protéines non structurées et leur étude par RMN

La protéine Tau est une protéine intrinsèquement non structurée dans sa forme native, dissociée des microtubules, qui adopte une conformation étendue, relativement flexible. Une des raisons de cette absence de structure tridimensionnelle est liée à la faible proportion d'acides aminés hydrophobes. En effet, le repliement des protéines en une structure globulaire est gouverné principalement par la formation d'un cœur hydrophobe. Ces caractéristiques structurales confèrent à la protéine Tau des propriétés particulières qui rendent possible son étude par RMN malgré sa taille (441 acides aminés pour l'isoforme la plus longue), notamment une vitesse de relaxation comparable à celle de petits peptides. En conséquence, la largeur des signaux est relativement faible pour une protéine de cette taille et le gain en résolution résultant compense, en partie seulement, la faible dispersion des déplacements chimiques liés à l'absence de structures secondaires et tridimensionnelles. Un autre facteur est défavorable à l'étude de la protéine Tau par RMN, il s'agit de la mauvaise représentation

des vingt acides aminés usuels dans la séquence, avec cinq acides aminés (Ser, Thr, Lys, Gly, Pro) représentant environ 50% de la séquence.

Les protéines non structurées ou partiellement structurées ont longtemps été ignorées du fait des méthodes utilisées pour l'extraction et la purification des protéines à partir des cellules ou des tissus, où seules les protéines les plus robustes, donc présentant une structure globulaire très stable, survivent. Leur étude au niveau structural reste encore un challenge à l'heure actuelle. Avec les méthodes modernes, il est désormais possible d'isoler une protéine non structurée, la purifier, et d'étudier sa fonction *in vitro* et *in vivo*. La caractérisation directe des protéines partiellement ou non structurées est possible, à l'échelle du résidu acide aminé, par RMN à haut champ avec l'utilisation de protéines marquées uniformément ou sélectivement, combinée aux séquences impulsionsnelles tridimensionnelles dites "isotope-éditées". Ces techniques ont permis de contourner, au moins en partie, le problème de la faible dispersion des signaux protons, qui représente le problème majeur de l'étude des protéines non structurées par RMN. En revanche, la dispersion des résonances des noyaux  $^{15}\text{N}$  reste comparable à celle des protéines structurées. L'attribution des signaux utilise par conséquent la dispersion des résonances  $^{15}\text{N}$ , qui sont plus sensibles à la séquence primaire locale que les résonances  $^1\text{H}$ , et les résonances  $^{13}\text{C}$ , qui sont plus sensibles à l'environnement, notamment la conformation locale ou la structure tridimensionnelle.

La protéine neuronale Tau est totalement non structurée et son étude au niveau moléculaire a donc bénéficié des techniques de RMN et de biologie structurale modernes. D'autres techniques que la RMN sont possibles pour l'étude des protéines non structurées, comme la spectroscopie infrarouge, Raman ou de fluorescence, ou le dichroïsme circulaire, mais seule la RMN offre une résolution à l'échelle du résidu. Ainsi, on peut observer finement des changements conformationnels locaux sous l'effet d'interactions avec d'autres protéines ou ligands, de changements de conditions (agents dénaturants, pH, salinité...) ou encore, sous l'effet de changements covalents comme les PTM. Les stratégies qui peuvent être utilisées en RMN sont nombreuses pour caractériser la conformation d'une protéine intrinsèquement désordonnée, le but étant de déterminer la présence ou non de structures résiduelles locales. Ces structures résiduelles, transitoires, sont indicatives de régions impliquées dans la liaison à un partenaire et peuvent parfois adopter une structure secondaire stable lors de la liaison. Il peut également s'agir de régions ayant une propension à former des oligomères par interactions inter- ou intramoléculaires. Différents indicateurs provenant de différentes expériences RMN existent (Dyson and Wright 2004) : (1) les déplacements chimiques  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^1\text{H}_\alpha$  et  $^1\text{H}_\text{N}$  (Wishart 1992; Wishart, Bigam et al. 1995; Wishart, Bigam et al. 1995) qui permettent d'identifier des régions présentant une propension à former des structures résiduelles avec le calcul du score SSP (*Secondary Structure Propensity*) (Marsh, Singh et al. 2006); (2) les paramètres de relaxation  $^{15}\text{N}$ , T1, T2 et NOEs hétéronucléaires (Golovanov, Blankley et al. 2007) qui permettent de discriminer les régions totalement flexibles et désordonnées des régions plus ordonnées (Kay 1997; Kay 1998; Kay 2005) ; (3) les contraintes de distances proches avec les NOE  $^1\text{H}$ - $^1\text{H}$  (moins de 5Å) ou les constantes de couplage  $^3\text{J H}_\text{N}$ - $\text{H}_\alpha$ , ou à plus longue distance par le marquage de spin avec des espèces paramagnétiques dans les expériences PRE (*Paramagnetic Relaxation Enhancement*), la mesure des couplages dipolaires résiduels (RDC ou *residual dipolar couplings*)... Différentes expériences doivent être combinées pour l'étude des IDP car les déviations par rapport aux valeurs standards ou aux valeurs théoriques calculées sont faibles (Marsh, Singh et al. 2006).



La RMN offre donc de nombreuses approches pour l'étude de protéines désordonnées tant au niveau structural que fonctionnel, notamment au travers des interactions avec leurs partenaires. L'intérêt dans le cas de la protéine Tau est de comprendre son rôle aussi bien physiologique que pathologique, et dans ce contexte, l'importance n'est pas tant d'étudier la protéine Tau isolée que de comprendre les phénomènes qui se produisent dans un contexte plus physiologique, notamment quand celle-ci est liée aux microtubules ou à la tubuline. Que ce soit pour les investigations de la liaison à la tubuline ou aux microtubules ou concernant la structure des PHF et le processus de leur formation, la compréhension de ces phénomènes au niveau moléculaire se heurte aux contraintes liées à l'étude de gros objets, qui est une limite majeure de la RMN. A l'inverse, la cristallographie est plus adaptée pour ces complexes de grande taille mais elle sera confrontée à la dynamique importante que conservent certaines régions de Tau dans les deux contextes.

### **1.2.2. Etudes structurales et fonctionnelles de Tau par RMN : structures résiduelles, liaison aux microtubules, formation des PHF et impact de la phosphorylation**

Le rôle primordial de Tau dans la régulation de la dynamique des microtubules et dans la pathogenèse de la MA a conduit à des efforts considérables en matière de biologie structurale pour déterminer la structure d'une protéine nativement non structurée, ce qui peut sembler paradoxal. Cette nature intrinsèquement désordonnée empêche la cristallisation *per se* et les études consécutives par cristallographie aux rayons X. *A priori*, la taille de 441 résidus devrait décourager également les études par RMN puisqu'elle se situe au-delà de ce qui est accessible pour des études structurales (on considère généralement comme limite une taille d'environ 25 kDa). Les premières données structurales ont été obtenues grâce aux techniques classiques de biologie moléculaire et de biochimie (reconnaissance par des anticorps ciblant des épitopes discontinus, mutagenèse, délétions, *pull-down*, fluorescence/FRET, tests fonctionnels...). Le développement de la RMN appliquée aux protéines, puis, plus récemment, aux protéines désordonnées a permis d'envisager l'étude de la protéine Tau malgré sa taille puisque ces protéines désordonnées ont un comportement très dynamique en solution. Elle apparaît donc comme un objet de poids moléculaire plus faible en termes de largeurs de raie. Ainsi, la RMN permet l'étude structurale des protéines désordonnées, comme Tau ou  $\alpha$ -synucléine impliquée dans la maladie de Parkinson, avec une résolution à l'échelle du résidu (Eliezer 2009). Les spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC constituent une sorte de signature pour une protéine donnée et sont à la base de la plupart des expériences RMN. L'attribution des  $^1\text{H}$ - $^{15}\text{N}$  HSQC de différentes isoformes de la protéine Tau a été initiée dans l'équipe de G. Lippens (Lippens, Wieruszkeski et al. 2004; Smet, Leroy et al. 2004), il y a une dizaine d'année, et sont décrits dans le Chapitre 2 plus en détails car ils ont été l'objet d'une partie de mes travaux de thèse. La possibilité de détecter des structures résiduelles à l'aide des résonances  $^{13}\text{C}$ , obtenues des expériences tridimensionnelles avec la protéine doublement enrichie  $^{15}\text{N}/^{13}\text{C}$ , était la perspective première de ce travail d'attribution. Le problème auquel nous avons été confrontés vient, d'une part, de la similitude de séquence entre les *repeats* qui rend difficile l'attribution complète de la protéine. En effet, la comparaison de ces séquences répétitives montre que 11 acides aminés sont identiques dans ces quatre séquences et chacune de ces séquences contient un motif particulier de quatre acides aminés, PGGG. D'autre part, l'élargissement plus important de certaines résonances du spectre semble indiquer une hétérogénéité dynamique de la protéine ou une dynamique d'échange des fonctions amides avec l'eau différente selon les régions protéiques. Les attributions sont donc

partielles et réalisées manuellement, dans un premier temps. Les problèmes d'attribution ont été contournés avec les marquages sélectifs qui permettent de réduire significativement le recouvrement des résonances ou la comparaison de spectres de différentes isoformes (Lippens, Sillen et al. 2006), jusqu'à l'utilisation de petits peptides (Smet, Leroy et al. 2004). Le clonage de fragments plus courts, ne contenant par exemple que la région des *repeats* (fragments K18 ou K19) (Mukrasch, Biernat et al. 2005), ou contenant également les régions voisines : le domaine C-terminal et la région riche en prolines (fragment K32 [198-394]) (Mukrasch, von Bergen et al. 2007) a permis de réduire la taille des protéines et donc le recouvrement, tout en conservant les régions fonctionnelles. Ces études structurales fines ont permis de détecter la présence de courtes séquences présentant une propension à former des feuillettes  $\beta$  au début des *repeats* R2, R3 et R4, coïncidant avec les hexapeptides PH6\* et PHF6 précédemment identifiés par dichroïsme circulaire et spectroscopie infra-rouge à transformée de Fourier (von Bergen, Friedhoff et al. 2000). La région C-terminale des *repeats* ne présente pas de structure particulière. En revanche, un motif <sup>224</sup>KKVAVVR<sup>230</sup> est plus rigide que le reste de la région riche en prolines probablement sous l'effet d'interactions transitoires à longue distance. De plus, la phosphorylation des sites T231, S235 et S237 induit une conformation en hélice de type poly-proline II dans la région 229-238 (Bielska and Zondlo 2006). Il a également été montré que de courtes régions dans chaque *repeat* présentaient une tendance à former des coudes  $\beta$ , résistante aux conditions dénaturantes, et viendraient perturber la tendance des *repeats* à former des structures en feuillet  $\beta$  étendues (Mukrasch, Markwick et al. 2007).

L'utilisation de séquences impulsives multidimensionnelles combinées à des stratégies d'attribution automatique a mené à l'attribution presque complète de l'isoforme la plus longue de Tau mais requiert des temps d'acquisition très longs (Narayanan, Durr et al. 2010; Harbison, Bhattacharya et al. 2012). Des séquences tridimensionnelles classiques, nécessitant des temps d'acquisition plus courts, ont été combinées à une stratégie d'attribution semi-automatique (Verdegem, Dijkstra et al. 2008) dans notre laboratoire aboutissant au même résultat. A partir des attributions, il est possible d'étudier par RMN différents aspects impliqués dans la physiopathologie de la protéine Tau comme (i) la liaison de Tau aux microtubules (Fischer, Mukrasch et al. 2007; Sillen, Barbier et al. 2007; Fauquant, Redeker et al. 2011), (ii) l'effet de la phosphorylation sur la structure de la protéine Tau soluble (Bibow, Ozenne et al. 2011; Sibille, Huvent et al. 2011), (iii) l'interaction avec des polyanions comme l'héparine induisant l'agrégation de la protéine (Mukrasch, Biernat et al. 2005; Sibille, Sillen et al. 2006; Mukrasch, von Bergen et al. 2007), (iv) le mécanisme antiagrégant de petites molécules (Akoury, Gajda et al. 2013), (v) la dynamique de Tau dans les PHFs par RMN liquide et à l'angle magique (Sillen, Leroy et al. 2005; Sillen, Wieruszkeski et al. 2005; Bibow, Mukrasch et al. 2011) ou (vi) la structure du cœur des PHF par RMN du solide (Andronesi, von Bergen et al. 2008; Daebel, Chinnathambi et al. 2012).

En particulier, la RMN a permis d'étudier l'effet de la phosphorylation sur la structure de Tau. Dans la protéine Tau entière, une rupture des interactions transitoires à longue distance a été observée sous l'effet de pseudo-phosphorylations au niveau des épitopes AT8, AT100 et PHF-1, obtenues par des mutations glutamates mimant la forme phosphorylée des résidus sérine ou thréonine impliquées dans ces phospho-épitopes (Bibow, Ozenne et al. 2011). Dans un fragment de la protéine Tau (fragment F4 [208-324]) phosphorylé par le complexe CDK2/Cycline A3 au niveau des résidus Thr231, Ser235, Thr212 et Thr217 de la région riche en prolines, a été détectée la formation d'une courte

hélice  $\alpha$  transitoire dans le segment 235-245. La double phosphorylation Thr231/Ser235 forme un N-cap qui stabilise l'hélice. D'autre part, ce profil de phosphorylation diminue significativement la capacité du fragment F4 à polymériser la tubuline (Sibille, Huvent et al. 2011). Il a été montré également que la présence d'au moins trois phosphorylations parmi les épitopes AT8 (Ser202/Thr205) et AT180 (Thr231/Ser235) est nécessaire pour empêcher l'assemblage des microtubules sans pour autant modifier significativement l'affinité de Tau pour les microtubules (Amniai, Barbier et al. 2009). Tau phosphorylée pourrait participer à une interaction non productive avec la tubuline à la manière de la Stathmine (Jourdain, Curmi et al. 1997). D'autre part, une cartographie de l'interaction de Tau avec les microtubules stabilisés par le paclitaxel a montré une disparition des résonances du domaine à segments répétés ainsi que des résidus de la région riche en prolines à partir du résidu Ser214, révélant leur immobilisation à la surface des microtubules. La phosphorylation en Ser214, bien que diminuant significativement l'affinité de Tau, ne permet pas de retrouver l'intensité des résonances observées en l'absence des microtubules, indiquant que la région riche en prolines ne récupère pas une mobilité suffisante. Au contraire des deux domaines précédemment cités, les résonances du domaine N-terminal restent inchangées, indiquant que ces résidus conservent leur flexibilité et n'adoptent pas de structure stable (Sillen, Barbier et al. 2007). Un fragment de Tau a pu être identifié permettant une liaison optimale au complexe constitué de deux dimères de tubuline liant le domaine *stathmin-like* de la protéine RB3 (complexe T2R ; voir la référence (Gigant, Curmi et al. 2000)). Il inclut une partie de la région riche en prolines et va jusqu'au troisième *repeat* (Tau-F4 [208-324]) (Fauquant, Redeker et al. 2011). Ce fragment présente une affinité plus importante que Tau entière indiquant la présence, dans Tau, de régions inhibitrices de l'interaction avec les microtubules. Chacun de ces domaines, le domaine riche en prolines d'une part et le domaine de liaison aux microtubules d'autre part, lie T2R avec une affinité modérée mais ils fonctionnent en synergie pour lier les microtubules (Fauquant, Redeker et al. 2011). Les détails moléculaires qui permettront de comprendre le mode d'action de Tau dans la liaison et la polymérisation de la tubuline n'ont pas encore été élucidés. Des efforts importants ont été faits dans notre laboratoire pour avancer dans cette voie.

La formation des PHF a pu être étudiée par RMN liquide ou l'angle magique pour observer les régions de Tau qui conservent un certain degré de flexibilité, même dans le contexte des fibres. A nouveau, il s'agit principalement de la région N-terminale (Sillen, Leroy et al. 2005; Sillen, Wieruszkeski et al. 2005). D'autre part, la disparition d'un grand nombre de résonances, dans des conditions d'agrégation, permet d'identifier les résidus impliqués dans la formation du cœur des fibres et répond à des phénomènes dynamiques complexes (Sillen, Leroy et al. 2005). Le fait que la protéine Tau n'agrège pas spontanément *in vitro* indique la présence de facteurs inhibiteurs au sein de la protéine. Les régions flanquant les points de nucléation de l'agrégation, incluant la région riche en prolines chargée positivement en amont, pourraient jouer ce rôle d'inhibiteur. L'interaction entre Tau et de petits fragments d'héparine a été étudiée par RMN dans des conditions où l'agrégation ne se produit pas. L'initiation de l'agrégation et la formation des PHF par les polyanions pourraient être similaire à l'effet de l'hyperphosphorylation. Il faut noter toutefois l'incapacité de la protéine Tau phosphorylée seule à former des PHF *in vitro*. Il a été montré que l'héparine interagit, entre autre, avec une zone positivement chargée dans la région riche en prolines. L'interaction avec l'héparine induit la formation de brins  $\beta$  dans plusieurs régions de Tau qui pourraient servir de sites de nucléation de l'agrégation. L'héparine est également intégrée dans le cœur des fibres où elle pourrait

neutraliser des fragments riches en lysines afin de permettre un empilement des feuillets  $\beta$  pour former les fibres (Sibille, Sillen et al. 2006).

Enfin, la RMN a été utilisée comme outil analytique pour la caractérisation de phosphorylations *in vitro* (sites de modification, niveaux relatifs et cinétique de phosphorylation) par des kinases purifiées ou des mélanges plus complexes, comme des extraits cellulaires, à l'échelle du résidu (Landrieu, Lacosse et al. 2006; Amniai, Barbier et al. 2009; Landrieu, Leroy et al. 2010; Leroy, Landrieu et al. 2010). De plus, la RMN dans la cellule (*in-cell NMR*), une des perspectives les plus prometteuses de la RMN biologique à l'heure actuelle, a permis d'observer la protéine Tau injectée dans des ovocytes de *Xenopus laevis* et de détecter les régions associées aux microtubules ainsi que des modifications *in situ* par phosphorylation (Bodart, Wieruszeski et al. 2008).

## Chapitre 2 : Phosphorylation de la protéine Tau et régulation des interactions avec la peptidyl-prolyl *cis/trans* isomérase Pin1

---

### 2.1 INTRODUCTION

#### **2.1.1. De l'hétérogénéité conformationnelle des motifs Ser/Thr-Pro phosphorylés à une régulation fonctionnelle par Pin1 ?**

Le cycle cellulaire est régulé par des protéines kinases cycline-dépendantes ou CDK dont les substrats sont majoritairement composés par le motif Ser/Thr-Pro. Récemment, un nouveau mécanisme de régulation a été proposé impliquant la conformation des prolines dans ces motifs Ser/Thr-Pro phosphorylés (Figure 2.1) (Lu, Hanes et al. 1996; Lu 2000; Lu, Liou et al. 2002; Lu, Liou et al. 2003), dont la plupart sont reconnus par l'anticorps monoclonal mitotique phospho-dépendant, MPM-2 (Davis, Tsao et al. 1983). Ces épitopes sont reconnus par des protéines de la famille des peptidyl-prolyl *cis/trans* isomérases (PPIases) comme la protéine Pin1 (*Protein interacting with NIMA-1*). La régulation de la conformation de la proline dans ces motifs par les PPIases a été proposée comme un nouveau mécanisme de régulation du cycle cellulaire, consécutif au processus de phosphorylation. Sur le plan fonctionnel, Pin1 et ses homologues eucaryotes jouent un rôle essentiel dans la progression du cycle cellulaire. La surexpression de Pin1 dans les cellules HeLa induit l'arrêt en mitose. Inversement, la déplétion de Pin1 provoque l'apparition de phénotypes mitotiques précoces, mais empêche également les cellules de sortir de la mitose et les conduit ainsi tout droit vers l'apoptose (Hani, Stumpf et al. 1995; Lu, Hanes et al. 1996; Crenshaw, Yang et al. 1998; Shen, Stukenberg et al. 1998). Sur le plan moléculaire, hormis pour les modèles peptidiques, peu de données structurales sur des protéines permettent de comprendre le rôle des PPIases et l'impact de leur action sur la structure et la fonction biologique des protéines substrats.

#### **2.1.2. Méthodes de caractérisation de l'échange conformationnel**

Contrairement aux PTM résultant de l'addition d'un ou plusieurs groupements chimiques, les modifications de nature purement conformationnelle, comme l'échange entre les conformères *cis* et *trans* des prolines, avec conservation stricte de la composition chimique de la protéine, sont difficilement détectables et quantifiables par les techniques biochimiques classiques, ou seulement de manière indirecte. La spectrométrie de masse se révèle donc inefficace dans ce cas, puisque l'échange conformationnel n'induit pas de modification de la masse de la protéine. On peut utiliser, par exemple, la reconnaissance d'un seul des deux conformères par une protéase qui n'hydrolysera alors qu'une fraction du substrat peptidique- celle qui est en conformation *trans*- ce qui aboutira à un déplacement de l'équilibre conformationnel en faveur de ce conformère (Fischer, Bang et al. 1984). Malheureusement, toutes les protéines "conformère-dépendantes" décrites à ce jour sont *trans*-spécifiques ce qui rend cette méthode peu sensible pour la détection d'un équilibre conformationnel ou d'une modification de celui-ci sous l'effet d'une PPIase car la conformation *trans* est très souvent majoritaire. A moins qu'il ne soit stabilisé par la structure globulaire de la protéine, le conformère *cis* correspond à une conformation défavorable, associée état énergétique élevé

comparé à celui de la conformation *trans*, où les fragments de la chaîne polypeptidique de part et d'autre de la liaison peptidique en *cis* sont trop proches d'un point de vue stérique. On estime que, dans de petits peptides non structurés, la proportion de formes *cis* ne dépasse pas 10%, et 30% dans le cas particulier d'une liaison Tyr-Pro. Dans ce contexte, la RMN est un outil particulièrement intéressant puisqu'elle permet de détecter directement l'interconversion *cis/trans* des prolines qui se déroule sur une échelle de temps relativement lente (de l'ordre de la minute). Il est donc possible d'observer distinctement les résonances correspondant à ces deux conformères (Figure 2.1). La spectroscopie d'échange utilisant la séquence  $^1\text{H}$ - $^1\text{H}$  EXSY permet d'observer un échange entre deux états. Dans le cas de l'interconversion *cis/trans* des prolines, cet échange est trop lent pour être détecté par spectroscopie d'échange EXSY. En revanche, elle permet de détecter de manière directe l'accélération de l'isomérisation *cis/trans* des prolines en présence d'une PPIase (voir Figure 2.3). On peut alors observer un pic d'échange reliant les résonances des conformères *cis* et *trans* de la proline (par exemple, entre le proton  $\text{H}_\alpha$  de la forme *cis* et le proton  $\text{H}_\alpha$  de la forme *trans*). L'expérience hétéronucléaire *z-exchange* permet d'observer, sur une expérience de type  $^1\text{H}$ - $^{15}\text{N}$  HSQC, un pic d'échange pour les résonances amide des résidus au voisinage d'une proline, qui ressentent l'échange conformationnelle de celle-ci.

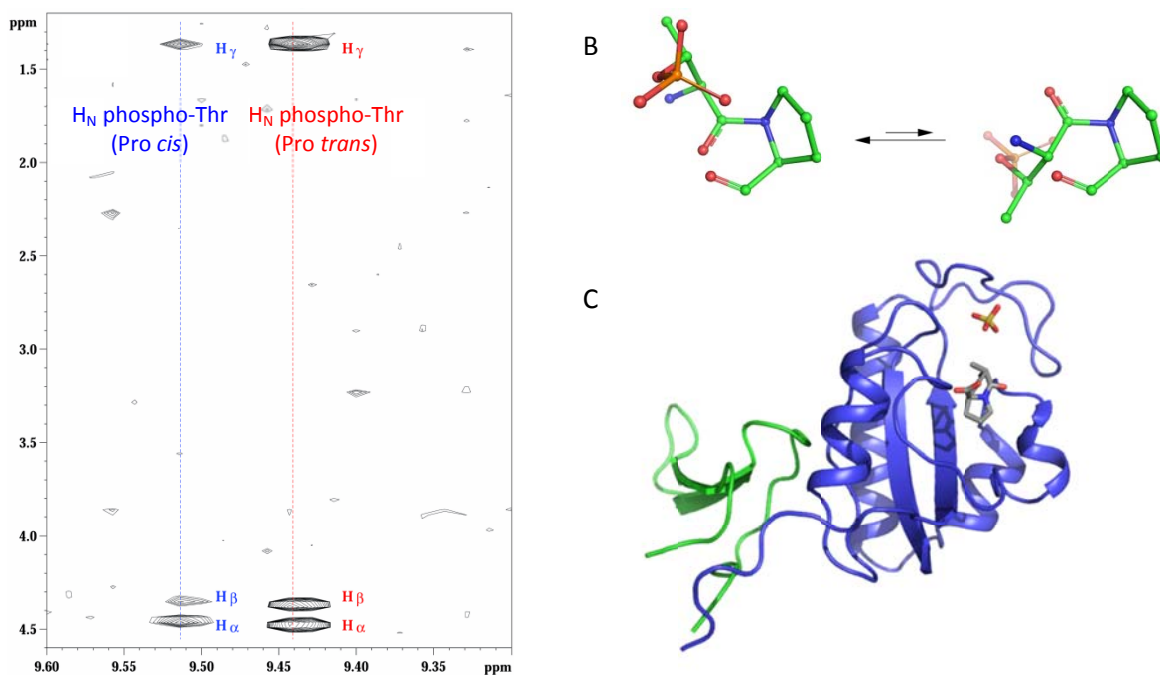


Figure 2.1 : (A,B) Observation d'un équilibre conformationnel relatif aux formes *cis* et *trans* de la proline dans un motif phospho-Thr-Pro. (A) Représentation des contacts scalaires  $^1\text{H}$ - $^1\text{H}$  de la phospho-Thr (corrélations entre les protons  $\text{H}_\text{N}$  et les protons  $\text{H}_\alpha$ ,  $\text{H}_\beta$  et  $\text{H}_\gamma$ ) sur un spectre TOCSY. Les signaux les plus intenses, annotés en rouge, correspondent à la forme majoritaire liée à la Pro en *trans*. Les signaux annotés en bleu correspondent à la forme minoritaire liée à la Pro en *cis*. (B) Equilibre conformationnel *cis/trans* de la liaison peptidique précédant une proline dans un dipeptide phospho-Thr-Pro. (C) Structure tridimensionnelle de la peptidyl-prolyl *cis/trans* isomérase humaine Pin1 qui catalyse l'interconversion conformationnelle des motifs phospho-Ser/Thr-Pro dans les protéines substrats. Le domaine WW de Pin1 est représenté en vert et le domaine PPIase en bleu, en complexe avec un ion sulfate et le dipeptide Ala-Pro (PDB ID : 1PIN) (Ranganathan, Lu et al. 1997).

### 2.1.3. Interactions fonctionnelles entre Pin1 et Tau phosphorylée, et implications physiopathologiques

La fonction de la protéine Tau dans l'association aux microtubules et la polymérisation de la tubuline est régulée par la phosphorylation. De nombreuses kinases sont impliquées, notamment les kinases dites *proline-directed* qui phosphorylent spécifiquement les motifs Ser/Thr-Pro. CDK5/p25, Erk2/MAPK1 et GSK3 $\beta$  sont des kinases majeures impliquées dans la phosphorylation de Tau, et également dans l'hyperphosphorylation pathologique. Il a été proposé qu'une réactivation anormale du cycle cellulaire dans les neurones matures, avec l'expression concomitante de ces kinases, puisse être à l'origine de l'hyperphosphorylation de Tau dans la maladie d'Alzheimer (Hamdane, Smet et al. 2002; Hamdane, Delobel et al. 2003). De plus, la formation d'épitopes MPM-2 est une caractéristique de la maladie d'Alzheimer et d'autres maladies neurodégénératives, et le profil de phosphorylation de Tau observé dans la maladie d'Alzheimer est similaire à celui de cellules mitotiques (Vincent, Zheng et al. 1998; Lu, Liou et al. 2003).

Il a été montré que la peptidyl-prolyl *cis/trans* isomérase humaine Pin1 est un acteur important dans le processus de neurodégénérescence impliquant la protéine Tau (Lu, Wulf et al. 1999; Liou, Zhou et al. 2011). L'interaction Pin1/Tau se fait via les motifs phospho-Ser/Thr-Pro, phosphorylés spécifiquement par les kinases *proline-directed*, notamment des kinases mitotiques telles que les CDK. La régulation par Pin1 de l'isomérisation conformationnelle du lien peptidique précédant les prolines, de manière phospho-dépendante, pourrait constituer un deuxième niveau de régulation de la fonction physiopathologique de la protéine Tau, consécutive à la phosphorylation (Lu 2000; Lu, Liou et al. 2003). L'interaction de Pin1 avec Tau phosphorylée permet de restaurer la fonction de Tau dans la liaison aux microtubules et la promotion de leur assemblage *in vitro* selon un mécanisme qui reste à définir. En revanche, dans le contexte pathologique, la séquestration de Pin1 dans les NFT résulterait en une diminution de Pin1 soluble, induisant l'arrêt en mitose et conduisant les neurones vers l'apoptose (Lu, Wulf et al. 1999). Il a été montré dans des souris Pin1<sup>-/-</sup> que Pin1 aurait un effet protecteur sur la dégénérescence neuronale puisqu'en l'absence de Pin1, les souris développent avec l'âge des déficits moteurs et comportementaux, une hyperphosphorylation de Tau, des PHF et une dégénérescence neuronale (Liou, Sun et al. 2003).

Au niveau moléculaire, Pin1 est constituée d'un petit domaine N-terminal de 40 résidus, le domaine WW, qui se lie spécifiquement aux motifs phospho-Ser/Thr-Pro et, dans sa partie C-terminale, d'un domaine catalytique responsable de la catalyse de l'isomérisation *cis/trans* de la liaison peptidique dans ces mêmes motifs (Figure 2.1) (Ranganathan, Lu et al. 1997; Verdecia, Bowman et al. 2000). De manière surprenante, il a été montré que Pin1 se lie, avec une haute affinité, à la protéine Tau au niveau d'un unique phospho-épitope, le motif phospho-Thr231-Pro232 (Lu, Wulf et al. 1999), qui est quantitativement phosphorylé dans le contexte de la MA (Morishima-Kawashima, Hasegawa et al. 1995). Pin1 se trouve ainsi séquestrée dans les agrégats de Tau formés dans le cytosol des neurones en dégénérescence, conduisant à une diminution de Pin1 soluble. Une diminution de la concentration de Pin1 au niveau nucléaire, par interactions avec de nombreux phospho-épitopes générés par l'activation des kinases mitotiques dans les neurones en dégénérescence telles que Cdk5/p25, pourrait être à l'origine d'une réactivation du cycle cellulaire dans les neurones matures (Hamdane, Smet et al. 2002; Hamdane, Delobel et al. 2003).

## 2.2 RESULTATS

### **2.2.1. Le motif phospho-Thr212-Pro213, un nouveau site de la protéine Tau permettant une interaction fonctionnelle avec Pin1 (Smet, Sambo et al. 2004)**

De manière surprenante, l'interaction Pin1/Tau qui a été décrite par Lu et collaborateurs, est spécifiquement régulée par le phospho-épitope phospho-Thr231-Pro232 parmi les 17 motifs phospho-Ser/Thr-Pro que contient la protéine Tau (Lu, Wulf et al. 1999). La liaison de Pin1 à ce phospho-épitope serait nécessaire et suffisante pour permettre à Pin1 de restaurer l'activité de Tau phosphorylée sur la polymérisation de la tubuline (Lu, Wulf et al. 1999). Bien que cet épitope soit décrit comme étant hautement phosphorylé dans la protéine Tau fœtale et dans les PHFs (Morishima-Kawashima, Hasegawa et al. 1995), la spécificité de l'interaction avec Pin1 via cet épitope est difficile à comprendre d'après les données structurales, qui montrent que la reconnaissance des protéines phosphorylées par Pin1 se limite au seul motif phospho-Ser/Thr-Pro, (Ranganathan, Lu et al. 1997; Verdecia, Bowman et al. 2000; Wintjens, Wieruszeski et al. 2001). De plus, l'impact structural d'une activité PPIase sur une protéine non structurée comme Tau n'est pas évident. Il pourrait néanmoins exister des motifs structuraux particuliers impliquant une proline qui nécessiteraient l'intervention d'une PPIase pour faciliter les interactions avec et/ou les modifications par un troisième partenaire. Par exemple, Pin1 a été décrite comme médiateur dans le recrutement de partenaires de la protéine Tau, comme la phosphatase PP2A. Il a été montré que Pin1 stimule la déphosphorylation de Tau par la phosphatase PP2A (Zhou, Kops et al. 2000). A nouveau, le site phospho-Thr231-Pro232 a été décrit comme étant la cible privilégiée de PP2A. Par ailleurs, Pin1 pourrait également réguler des interactions entre Tau et ses partenaires via la liaison des mêmes sites. Une compétition entre le domaine WW de Pin1 et la sous-unité CKS (*cyclin-dependent kinase subunit*) pour la liaison de motifs phospho-Ser/Thr-Pro a été mise en évidence (Landrieu, Odaert et al. 2001).

Nous avons mesuré les constantes de dissociation des complexes formés par Pin1, ou son domaine WW isolé, avec différents peptides issus de Tau, centrés autour de divers phospho-épitopes décrits dans la forme physiologique ou plus spécifiques de la forme pathologique observées dans la maladie d'Alzheimer. La RMN a la capacité de détecter des interactions de faible affinité (comprise entre 10  $\mu$ M et 1 mM), indétectables par bon nombre de techniques de mesure d'interactions moléculaires, grâce à l'acquisition de spectres bidimensionnels  $^1\text{H}$ - $^{15}\text{N}$  HSQC qui nécessite de travailler avec une protéine enrichie en  $^{15}\text{N}$  (Figure 2.2 A et B). La même expérience  $^1\text{H}$ - $^{15}\text{N}$  HSQC peut également permettre de déterminer le site de liaison d'un ligand (Figure 2.2 C et D). De plus, les constantes de vitesse d'association ( $k_{\text{on}}$ ) et de dissociation ( $k_{\text{off}}$ ) des complexes peuvent être extraites à partir du  $K_D$  et des largeurs de raies (Lippens, Caron et al. 2004) (Figure 2.2 E).



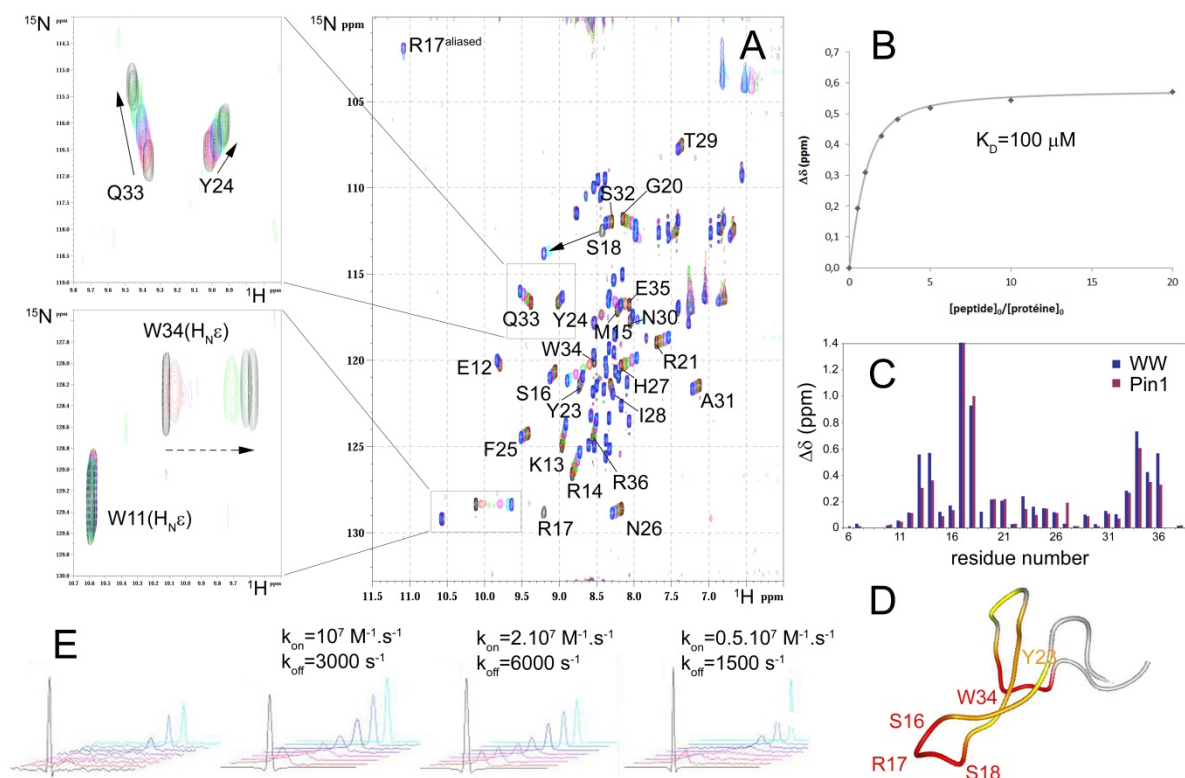


Figure 2.2 : (A) Titration du domaine WW isolé de Pin1, enrichi uniformément en  $^{15}\text{N}$ , par un peptide de Tau phosphorylé en Thr212. Des agrandissements de régions du spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC montrent différents comportements des résonances lors de l'ajout de ligand en fonction de la variation de fréquences entre l'état libre et l'état lié ( $\Delta\omega_{\text{max}}$ ). (B) La variation de déplacement chimique en fonction des concentrations initiales de protéine et de ligand permet d'extraire la constante de dissociation du complexe. (C et D) La cartographie des déplacements chimiques permet d'identifier le site de liaison du ligand. (E) Simulation de Monte Carlo de l'échange chimique permettant la modélisation de l'élargissement des raies lors de l'expérience de titration et l'extraction des constantes de vitesse d'association ( $k_{\text{on}}$ ) et de dissociation ( $k_{\text{off}}$ ) en fonction du  $K_D$ .

Nous avons observé une gamme restreinte d'affinités pour des peptides de Tau provenant de différentes régions comme la région riche en prolines et la région carboxy-terminale, allant de  $100 \mu\text{M}$  à  $1 \text{mM}$  pour les formes mono-phosphorylées (Smet, Sambo et al. 2004). Ces affinités ont également été mesurées par des techniques de fluorescence faisant intervenir la fluorescence intrinsèque du domaine WW ou le FRET (*Fluorescence Resonance Energy Transfer*) avec un peptide marqué par le groupement 5-(diméthylamino)naphthalène-1-sulfonyle (ou dansyle) en position N-terminale, comme sonde fluorescente (Figure 2.3 A et B). Seuls les complexes de plus haute affinité ont pu être détectés par ces méthodes, notamment ceux impliquant le peptide mono-phosphorylé en Thr212 ou des peptides multi-phosphorylés, avec des affinités similaires à celles déterminées par RMN.

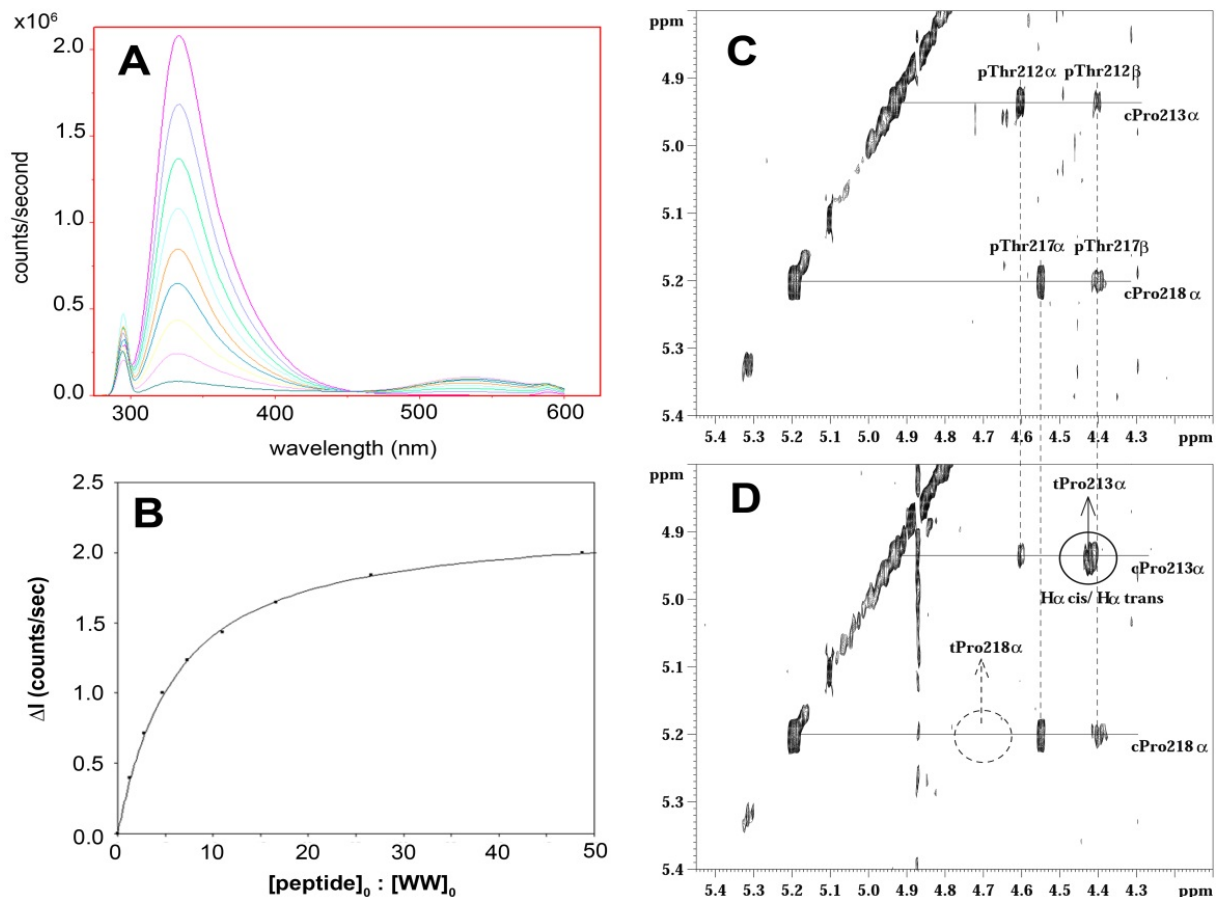


Figure 2.3 : (A,B) Titration de Pin1 par un peptide contenant le phospho-épitope Thr212-Pro213 réalisée par FRET. (A) Spectres d'émission de Pin1 après excitation à 295nm pour différentes concentrations de peptide comportant deux maxima d'émission à 334 et 530 nm, correspondant respectivement aux Trp et au dansyle. (B) Courbe de titration représentant la variation de l'intensité d'émission à 334 nm en fonction du ratio peptide sur protéine. (C,D) Activité *cis/trans* isomérase de Pin1 mesurée par RMN sur le peptide contenant les motifs phospho-Thr212-Pro213 et phospho-Thr217-Pro218. Spectres  $^1\text{H}$ - $^1\text{H}$  EXSY sur le peptide seul à 2 mM (C) et en présence de 50  $\mu\text{M}$  Pin1 (D). Le pic encerclé en trait plein correspond au pic d'échange entre les protons  $\text{H}\alpha$  *cis* et  $\text{H}\alpha$  *trans* du résidu Pro213. Le cercle en pointillés indique l'absence d'un pic similaire pour le résidu Pro218.

Cette absence de spécificité de liaison de Pin1 ou du domaine WW isolé concorde avec les données structurales qui indiquent que la reconnaissance par le domaine WW ou le domaine catalytique de Pin1 se fait principalement au niveau du motif phospho-Ser/Thr-Pro et que la séquence environnante ne module que très peu l'affinité. D'ailleurs, les dipeptides Ac-phospho-Ser/Thr-Pro-NH<sub>2</sub> sont reconnus dans la même gamme d'affinité (Smet, Sambo et al. 2004; Smet, Duckert et al. 2005). Nous avons constaté que la discrimination se fait principalement entre la phospho-Thr et la phospho-Ser avec une affinité environ trois fois plus importante pour la première (respectivement, 150 et 420  $\mu\text{M}$ ). De plus, les affinités de ces peptides pour le domaine WW isolé diffèrent peu des affinités pour Pin1 entière, indiquant une participation négligeable du domaine catalytique à la liaison des substrats. Ces résultats ne permettent donc pas d'expliquer la spécificité pour l'épitope phospho-Thr231 au niveau de la liaison et ont mis en évidence un site de plus haute affinité, le motif phospho-

Thr212-Pro213. Alors que nous n'avons pas observé de sélectivité pour la liaison des phospho-épitopes de Tau, nous avons montré une sélectivité de l'activité *cis/trans* isomérase de Pin1 pour le site phospho-Thr212-Pro213 (Figure 2.3 C et D) (Smet, Sambo et al. 2004). Ce site, en combinaison avec la phospho-Ser214, constitue l'épitope AT100, un phospho-épitope caractéristique des filaments insolubles de la protéine Tau associés à la MA (Hoffmann, Lee et al. 1997; Zheng-Fischhofer, Biernat et al. 1998). Il pourrait permettre une interaction fonctionnelle entre Pin1 et Tau, dont les conséquences biologiques restent à définir.

### **2.2.2. Rôle du domaine WW et des phosphorylations multiples dans la régulation de l'activité *cis/trans* isomérase de Pin1 (Smet, Wieruszeski et al. 2005)**

Dans les modèles peptidiques issus de la région riche en prolines de Tau, le ratio *trans*:*cis* de la liaison peptidique des motifs Ser/Thr-Pro est largement en faveur de la forme *trans*, mais peut aller toutefois jusqu'à 10% de forme *cis*. Ce ratio varie peu avec l'état de phosphorylation des résidus sérine ou thréonine impliqués dans ces motifs (Smet, Sambo et al. 2004). L'activité enzymatique de Pin1 via son domaine PPlase est plus sélective que la liaison des motifs phospho-Thr/Ser-Pro par le domaine WW, et n'a été observée que sur un seul site de phosphorylation parmi les différents sites criblés de la région riche en prolines, l'épitope phospho-Thr212-Pro213 (Smet, Sambo et al. 2004; Lippens, Landrieu et al. 2007).

Nous avons étudié l'activité catalytique de Pin1 entière et de son domaine catalytique isolé (PPlase), et montré que le domaine WW joue un rôle positif sur l'activité isomérase au niveau d'un substrat peptidique simple, ne contenant qu'un seul site de phosphorylation centré sur le motif phospho-Thr212-Pro213. L'enzyme est plus efficace en présence du domaine WW et l'échange *cis/trans* de la Pro213 (de l'ordre de la seconde) se caractérise par une vitesse 10 fois supérieure avec Pin1 qu'avec son domaine catalytique isolé (Figure 2.4) (Smet, Wieruszeski et al. 2005). Ces données suggèrent que le domaine WW fonctionne probablement en permettant l'ancrage de l'enzyme au substrat. En revanche, l'effet est plus subtil au niveau d'un substrat plus complexe présentant un deuxième site de phosphorylation, l'épitope phospho-Thr231-Pro232.

Pour cette étude, nous avons synthétisé en phase solide un peptide de 40 acides aminés comportant trois sites de phosphorylation aux positions Thr212, Ser214 et Thr231. Pour ce substrat, nous avons observé une diminution de l'activité isomérase de Pin1 d'un facteur 3 par rapport au peptide mono-phosphorylé en Thr212 indiquant que, bien que le domaine PPlase n'ait pas d'activité isomérase sur l'épitope phospho-Thr231-Pro232, une liaison non productive du domaine PPlase est néanmoins possible sur ce site et pourrait diminuer l'activité sur l'épitope phospho-Thr212-Pro213 (Figure 2.4). De plus, le domaine WW, en permettant la liaison de Pin1 alternativement sur l'un ou l'autre des motifs phospho-Thr-Pro, pourrait jouer un rôle régulateur sur la fonction isomérase de Pin1 au niveau d'un site particulier dans le contexte d'un substrat multi-phosphorylé (Smet, Wieruszeski et al. 2005).

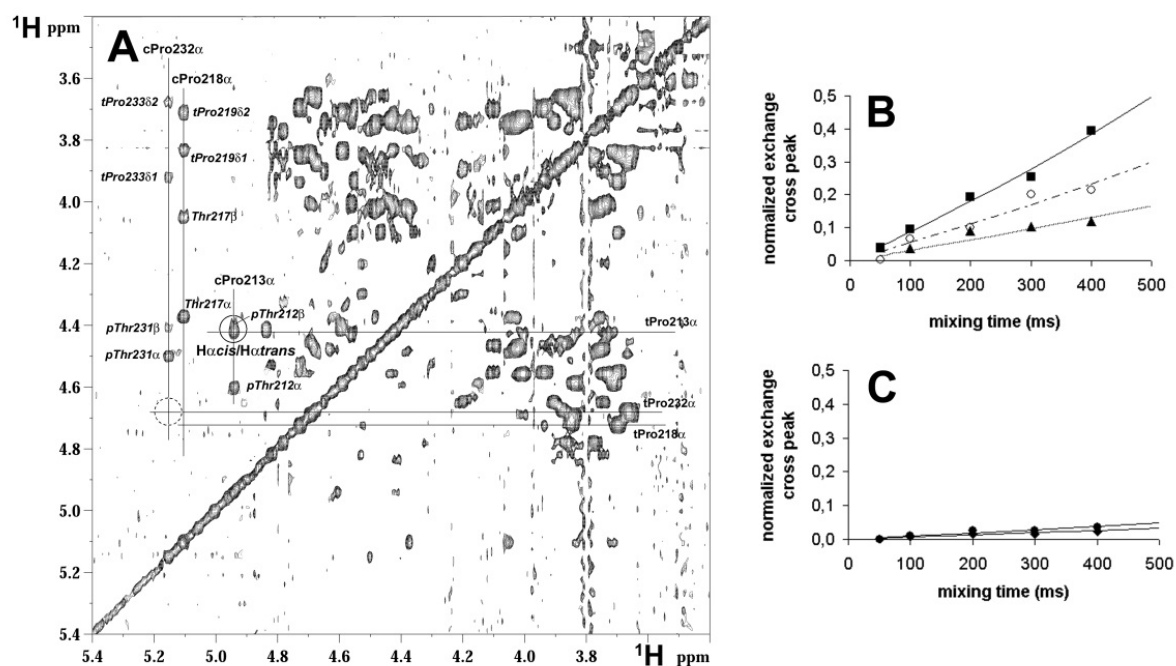


Figure 2.4 : (A) Activité *cis/trans* isomérase de Pin1 mesurée par RMN sur le peptide contenant les motifs phospho-Thr212-Pro213 et phospho-Thr231-Pro232. Spectres  $^1\text{H}$ - $^1\text{H}$  EXSY sur le peptide à 2mM en présence de Pin1 à 50  $\mu\text{M}$ . Le pic encerclé en trait plein correspond au pic d'échange entre les protons  $\text{H}\alpha$  *cis* et  $\text{H}\alpha$  *trans* du résidu Pro213. Le cercle en pointillés indique l'absence d'un pic similaire pour le résidu Pro232. Activité *cis/trans* isomérase de Pin1 (B) ou du domaine catalytique isolé Pin1<sup>CAT</sup> (C) sur différents peptides. (B) Activité de Pin1 sur le peptide court [208-221] comportant un seul motif en phospho-Thr212-Pro213 (■,  $k_{\text{exch}}=0.8 \text{ s}^{-1}$ ), sur le peptide long [201-240] comportant les deux motifs phospho-Thr212-Pro213 et phospho-Thr231-Pro232 (▲,  $k_{\text{exch}}=0.3 \text{ s}^{-1}$ ), et sur le mélange des deux peptides mono-phosphorylés [208-221] et [224-240] (○,  $k_{\text{exch}}=0.5 \text{ s}^{-1}$ ). (C) Activité de Pin1<sup>CAT</sup> sur le peptide court [208-221] comportant un seul motif en phospho-Thr212-Pro213 (●,  $k_{\text{exch}}=0.1 \text{ s}^{-1}$ ), sur le peptide long [201-240] comportant deux motifs en phospho-Thr212-Pro213 et phospho-Thr231-Pro232 (◆,  $k_{\text{exch}}=0.06 \text{ s}^{-1}$ ).

### 2.2.3. Vers l'étude de la protéine Tau entière et des interactions fonctionnelles avec Pin1 par RMN (Lippens, Wieruszeski et al. 2004; Smet, Leroy et al. 2004; Lippens, Sillen et al. 2006)

Bien que la protéine Tau soit intrinsèquement non structurée, une structuration particulière autour des sites de phosphorylation est envisageable, notamment dans le contexte de la protéine hyperphosphorylée ou en interaction avec la tubuline, et pourrait être à l'origine de la sélectivité observée dans l'interaction avec Pin1, au niveau du phospho-épitope Thr231-Pro232. Une étude structurale des peptides issus de la protéine Tau a permis de mettre en évidence une absence de structure secondaire ainsi qu'une hétérogénéité conformationnelle des prolines, impliquées dans les motifs phospho-Thr/Ser-Pro reconnus par Pin1, qui reste toutefois limitée (populations de formes *cis* inférieures à 10%). Mais le modèle peptidique a ses limites et ne permet pas de prendre en compte un effet de structuration globale ou de structures résiduelles locales qui pourraient exister dans le contexte de la protéine entière. Cependant, la caractérisation conformationnelle des peptides a

permis d'utiliser ces derniers comme modèles *random coil* dans la recherche de structures résiduelles ou de prolines dont l'équilibre *cis/trans* serait modifié au sein de la protéine Tau entière.

La comparaison des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de peptides d'environ vingt résidus et de la protéine Tau entière n'a pas permis de détecter de structures résiduelles ou de proline majoritairement en conformation *cis* autour de ces motifs Ser/Thr-Pro (Lippens, Wieruszeski et al. 2004; Smet, Leroy et al. 2004; Lippens, Sillen et al. 2006) qui puissent expliquer une sélectivité du complexe Pin1/Tau autour d'un site de phosphorylation particulier et le rôle fonctionnel de Pin1 sur la conformation de ses substrats. Dans le contexte de la protéine non phosphorylée, des déplacements chimiques strictement identiques ont été observés pour les résonances amides dans ces petits peptides non structurés ou dans le contexte de la protéine Tau entière (Figure 2.5), indiquant, une absence de structure, même résiduelle, dans les régions observées (Smet, Leroy et al. 2004). Des travaux publiés récemment par Zweckstetter et collaborateurs ont montré la présence de structures résiduelles dans la protéine Tau grâce à l'optimisation de la stratégie d'attribution du spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC de différentes isoformes de Tau et l'analyse des résonances  $^{13}\text{C}$  (Mukrasch, Bibow et al. 2009; Narayanan, Durr et al. 2010; Bibow, Mukrasch et al. 2011). Néanmoins, ces premières expériences de RMN sur la protéine Tau entière réalisées dans notre groupe ont été précurseurs de l'étude de Tau par RMN à haute résolution (Sillen, Leroy et al. 2005; Sillen, Wieruszeski et al. 2005; Landrieu, Lacosse et al. 2006; Sibille, Sillen et al. 2006; Sillen, Barbier et al. 2007; Amniai, Barbier et al. 2009).

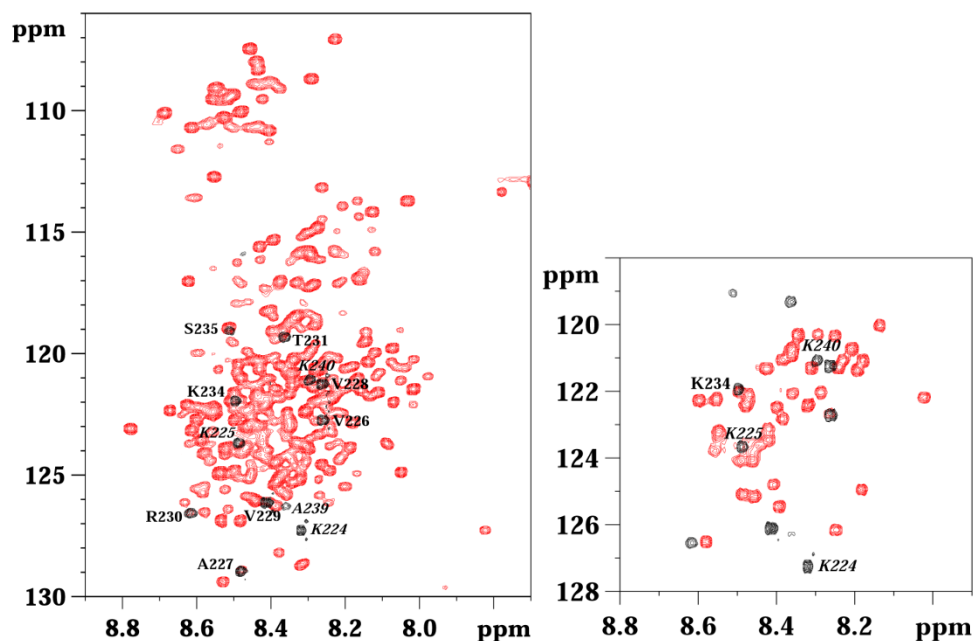


Figure 2.5 : Comparaison des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de la protéine Tau entière (rouge) uniformément enrichie en  $^{15}\text{N}$  (à gauche) ou sélectivement enrichie en  $^{15}\text{N}$ -Lys (à droite) avec le spectre du peptide de Tau [224-240] en abondance naturelle  $^{15}\text{N}$  (noir).

Dans le contexte de la protéine Tau phosphorylée, la situation est différente. La protéine Tau phosphorylée *in vitro* par le complexe CDK2/cyclineA3 purifié présente de nombreux sites de

phosphorylation Ser/Thr-Pro dont la plupart sont localisés dans la région riche en prolines (Figure 2.6 A). La région centrée sur l'épitope AT8 (pSer202/pThr205), dans la protéine Tau entière ou dans le fragment Tau-F5 [165-245] comprenant la région riche en prolines, présente un déplacement chimique inhabituel pour le résidu Gly207, dont le proton amide établit probablement un pont hydrogène avec le phosphate des résidus phosphorylés de l'épitope AT8, qui pourrait témoigner d'une structure locale particulière qui toutefois n'implique pas une modification de l'équilibre *cis/trans* des prolines (aux positions 203 et 206) (Figure 2.6 B). L'interaction de la protéine Tau phosphorylée avec Pin1 est difficile à étudier dans la mesure où la liaison de Pin1 conduit à un élargissement des résonances des résidus phosphorylés et de leurs voisins immédiats dans la séquence (Figure 2.6 A). De plus, on observe une précipitation importante des protéines de l'échantillon aux concentrations utilisées pour l'acquisition des spectres RMN. Cette précipitation pourrait être spécifique de certains complexes car elle apparaît pour des ratios Pin1 : Tau au-delà des concentrations équimolaires. Par ailleurs, nous avons pu observer une activité PPIase de Pin1 dans le contexte de la protéine Tau entière, via l'expérience hétéronucléaire *z-exchange*, au niveau de résidus voisins des prolines qui présentent une hétérogénéité conformationnelle *cis/trans* (Figure 2.6 B) (données non publiées).

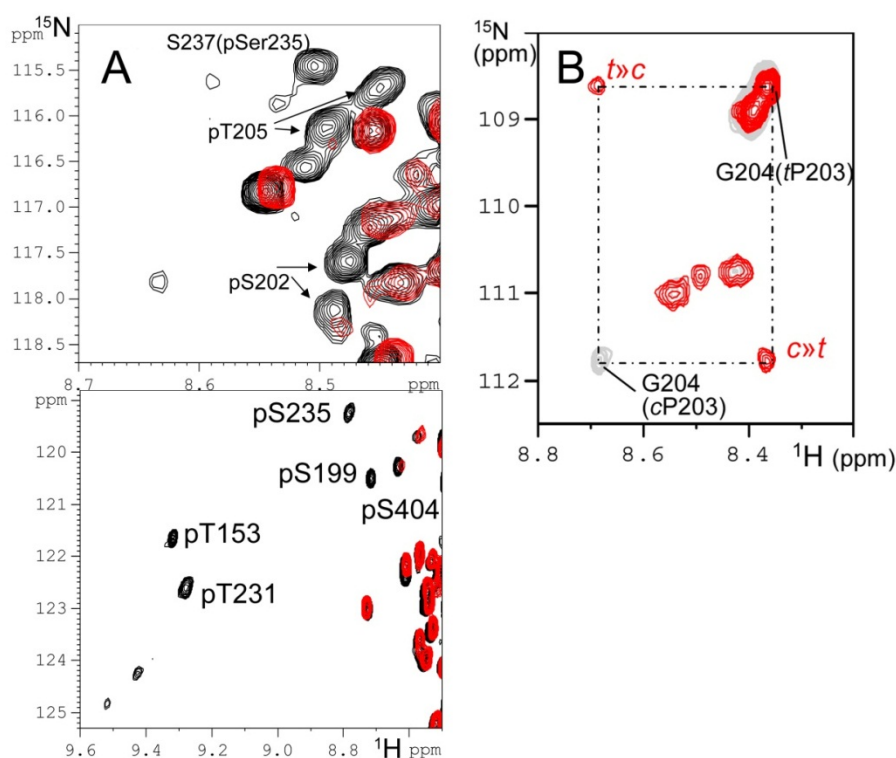


Figure 2.6 : (A) Interactions entre Pin1 et  $^{15}\text{N}$ -Tau phosphorylée par le complexe CDK2/cyclineA3. Spectre de Tau phosphorylée seule (noir) ou en présence de Pin1 (rouge). (B) Activité PPIase de Pin1 sur la protéine Tau phosphorylée. Spectre de contrôle de la protéine Tau phosphorylée par le complexe CDK2/cyclineA3 à 400  $\mu\text{M}$  (gris) et en présence d'une quantité catalytique de Pin1 (40  $\mu\text{M}$ ) (rouge). Les pics d'échange, dans l'expérience de *z-exchange*, indique une accélération de l'interconversion *cis* vers *trans* (noté *c*»*t*) et *trans* vers *cis* (noté *t*»*c*), observable pour le résidu Gly204 lié à la Pro203 en conformation *cis* (cP203) ou *trans* (tP203), et lié au résidu phospho-Thr205.

#### **2.2.4. Rôle de Pin1 dans la stimulation de la déphosphorylation de Tau par la phosphatase PP2A (Landrieu, Smet-Nocca et al. 2011)**

La régulation de la fonction physiologique de Tau dans la stabilisation des microtubules par la phosphorylation et la déphosphorylation suit un processus qui reste encore mal compris, tout comme la formation des fibres dans la pathologie Alzheimer, composées de protéines Tau hyperphosphorylées. Un profil de phosphorylation spécifique est donc le résultat d'une balance entre les activités kinases et phosphatases, et l'hyperphosphorylation résulterait donc d'une dérégulation de cette balance. La progression de la MA peut être caractérisée *post-mortem* avec des anticorps monoclonaux spécifiques, qui peuvent définir des lésions neurofibrillaires à différentes étapes de la pathologie. Par exemple, le site T231 est précocement phosphorylé (phospho-épitope AT180) dans la progression de la maladie et précède la phosphorylation du site S202/T205 (phospho-épitope AT8) (Goedert, Jakes et al. 1994; Goedert, Jakes et al. 1995; Braak, Alafuzoff et al. 2006). La protéine hétérotrimérique PP2A (*Protein Phosphatase 2A*) est la phosphatase majeure de Tau *in vivo*, qui contribue à son état de phosphorylation final (Drewes, Mandelkow et al. 1993; Gong, Grundke-Iqbal et al. 1994; Goedert, Jakes et al. 1995). Elle comprend un corps dimérique (PP2A<sub>D</sub>), constitué d'une unité catalytique (sous-unité C) et d'une unité "plate-forme" (sous-unité A), ainsi qu'une sous-unité régulatrice (sous-unité B) qui régule de manière spatio-temporelle l'activité de PP2A. La sous-unité régulatrice la plus importante dans les neurones est la sous-unité PR55/B $\alpha$  (PP2A<sub>T55 $\alpha$</sub> ) (Janssens and Goris 2001; Schmidt, Kins et al. 2002; Martens, Stevens et al. 2004; Janssens, Longin et al. 2008; Xu, Chen et al. 2008; Virshup and Shenolikar 2009). Il a été montré, dans un modèle de souris transgénique, qu'une réduction de l'activité de PP2A induit une hyperphosphorylation de Tau, incluant la phosphorylation de l'épitope AT8 (Kins, Crameri et al. 2001). Dans le cerveau de patients atteint de la MA, une réduction de l'expression de l'ARNm de la sous-unité C et de diverses sous-unités B de PP2A a été observée dans l'hippocampe (Vogelsberg-Ragaglia, Schuck et al. 2001). Au niveau protéique, une diminution du niveau de PR55/B $\alpha$  due à une augmentation du *turnover* a été associée avec la pathologie Alzheimer (Sontag, Luangpirom et al. 2004).

La spectroscopie RMN a été utilisée ici pour mesurer la vitesse de déphosphorylation de Tau phosphorylée par PP2A au niveau de sites spécifiques incluant les phospho-épitopes AT8 (pS202/pT205) et AT180 (pT231) (Figure 2.7 A). La phosphorylation de Tau a été générée *in vitro* par la kinase CDK2/cycline A3 recombinante (Jeffrey, Russo et al. 1995; Brown, Noble et al. 1999; Welburn and Endicott 2005). Nous avons montré que la sous-unité régulatrice PR55/B $\alpha$  est importante pour l'activité de déphosphorylation de phospho-Tau et que la phosphorylation du site pT231 interfère négativement dans la déphosphorylation du site pS202/pT205 (Figure 2.7 B). D'autre part, il a été montré que la peptidyl-prolyl *cis/trans* isomérase phospho-dépendante Pin1 stimule la déphosphorylation de Tau par PP2A (Zhou, Kops et al. 2000) (Figure 2.7 B). Nous avons donc utilisé notre approche par RMN de la déphosphorylation de Tau par PP2A pour étudier le mode d'action de Pin1 au niveau moléculaire. Nous avons observé que l'effet inhibiteur de la phosphorylation en T231 sur l'activité de PP2A est levé par Pin1. Puisque l'effet stimulant de Pin1 sur la déphosphorylation de Tau est perdu en utilisant le corps dimérique de l'enzyme (PP2A<sub>D</sub>) ou avec un mutant de Tau T231A qui ne peut plus être phosphorylé au niveau du site AT180, nous proposons que Pin1 régule l'interaction entre la sous-unité PR55/B $\alpha$  et le phospho-épitope AT180 de Tau.



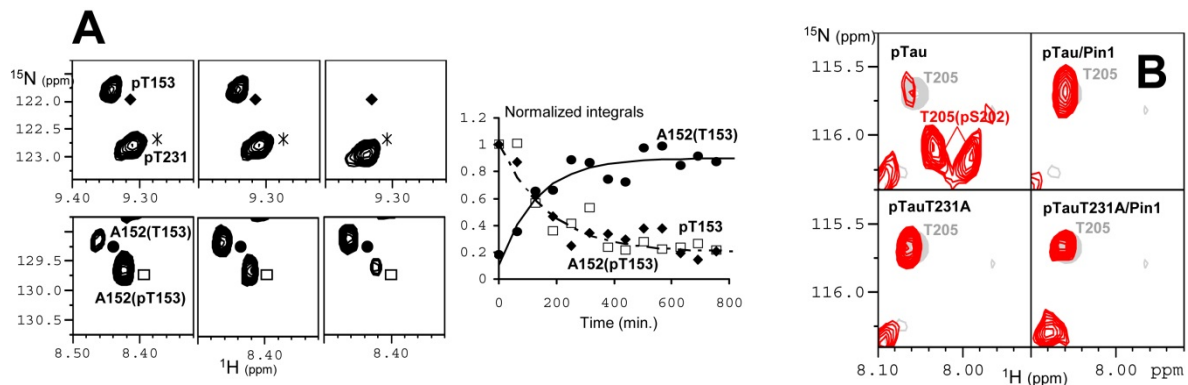


Figure 2.7 : (A) Cinétique de déphosphorylation de Tau phosphorylée par PP2A observée directement sur les résidus phosphorylés, pT153 (◆) et pT231 (\*) (en haut) ou sur un résidu voisin d'un site de phosphorylation, le résidu A152 (● pour le résidu A152 lié à la forme non phosphorylée du résidu T153, □ pour le résidu A152 lié à la forme phosphorylée du résidu T153) (en bas). (B) Déphosphorylation par PP2A de Tau phosphorylée. La phosphorylation de la protéine Tau a été obtenue *in vitro* avec le complexe CDK2/cycline A3 (spectres gris). La déphosphorylation par PP2A (spectres rouges) a été mesurée avec PP2A seule et la protéine Tau sauvage (pTau) ou avec le mutant Tau T231A (pTauT231A), en présence de Pin1 avec la protéine Tau sauvage (pTau/Pin1) ou le mutant Tau T231A (pTauT231/Pin1).

En conséquence, l'amorçage par une déphosphorylation en AT180 est requis pour une déphosphorylation efficace de l'épitope AT8 par PP2A<sub>T55α</sub>. Ainsi, une interaction complexe entre les mécanismes d'amorçage rapportés pour certaines kinases de Tau et celui décrit ici pour la phosphatase PP2A<sub>T55α</sub> pourrait contribuer à l'hyperphosphorylation de Tau observée dans les neurones dans la pathologie Alzheimer.

### 2.2.5. Inhibition de l'interaction de Pin1 avec Tau phosphorylée

Dans l'étude des interactions Pin1/Tau, nous avons envisagé d'inhiber l'activité de Pin1 pour éclaircir son rôle au niveau moléculaire dans des modèles *in vitro* et *in vivo* (Hamdane, Smet et al. 2002; Smet, Duckert et al. 2005). Trois voies d'inhibition de l'activité de Pin1 ont été envisagées :

- La conception rationnelle de petites molécules inhibitrices ciblant la liaison de Pin1 via son domaine WW aux substrats phosphorylés (Hamdane, Smet et al. 2002; Smet, Duckert et al. 2005)
- L'impact structural et dynamique de la phosphorylation du domaine WW de Pin1 dans la régulation de la liaison des substrats phosphorylés (Smet-Nocca, Launay et al. 2013), la phosphorylation de Pin1 ayant été impliquée dans la modification de la localisation subcellulaire et dans l'inactivation de Pin1 (Lu, Zhou et al. 2002).
- L'interaction de Pin1 avec l'inhibiteur 2 de la phosphatase PP1 (IPP-2) dans sa forme non phosphorylée (Sami, Smet-Nocca et al. 2011). Il a été montré, en effet, que l'interaction Pin1/IPP-2 modifie de manière allostérique la spécificité de Pin1 pour les phospho-protéines mitotiques (Li, Stukenberg et al. 2008; Sami, Smet-Nocca et al. 2011).



**Petites molécules interférant dans les interactions protéine-protéine entre le domaine WW de Pin1 et Tau phosphorylée**

Les interactions impliquant les motifs phospho-Ser/Thr-Pro dans les protéines jouent un rôle-clé dans un grand nombre de processus de régulation dans la cellule. Nous avons synthétisé et évalué des ligands potentiels du domaine de liaison de Pin1 dans le but d'inhiber les interactions protéine-protéine entre Pin1 et des peptides phosphorylés. Notre stratégie de conception des molécules est basée sur la structure du domaine WW en complexe avec des substrats phosphorylés et implique la synthèse d'analogues du dipeptide Ac-Thr(PO<sub>3</sub>H<sub>2</sub>)-Pro-NH<sub>2</sub> ainsi que l'évaluation de la liaison de ces molécules au domaine WW par RMN à haute résolution. Ce dipeptide a été défini comme étant le motif d'interaction minimal avec le domaine WW (Hamdane, Smet et al. 2002; Smet, Sambo et al. 2004). La RMN a permis, d'une part, de mesurer la constante de dissociation des complexes (K<sub>D</sub>) avec l'avantage de pouvoir mesurer des K<sub>D</sub> dans une gamme d'affinité relativement faible (de l'ordre de 1 mM). Nous avons pu ainsi discriminer des ligands dans une gamme de K<sub>D</sub> allant de 5 µM à quelques mM. De plus, la capacité d'identifier le site d'interaction des ligands a permis d'augmenter la complexité structurale des ligands de manière rationnelle et de relier le gain d'affinité de ces molécules à l'établissement d'interactions additionnelles au niveau du domaine WW (Figure 2.8) (Smet, Duckert et al. 2005).

Dans un premier temps, des composés peptido-mimétiques ont été synthétisés en phase solide dans lesquels la proline a été remplacée par un analogue structural ou le phosphate substitué par une tête polaire. Ces composés ont ensuite été criblés par RMN. Tous les composés constitués de mimes de phosphate ont montré une perte d'affinité significative (K<sub>D</sub>>2mM), en dehors de la gamme mesurable par RMN, indiquant une adaptabilité très limitée de la boucle β-hairpin liant le phosphate. En revanche, les mimes de proline ont abouti à des composés dont l'affinité reste mesurable et comprise entre 20 µM (pour la 3,4-dehydroproline) et 800 µM (pour la 4-*trans*-hydroxyproline). La poche aromatique, constituée des chaînes latérales des résidus Tyr23 et Trp34, responsable de la liaison de la proline, est capable d'adapter des cycles homologues au cycle pyrrolidinique ainsi que l'énantiomère de la proline sans perte d'affinité notable. Cependant, l'introduction d'un isostère ou d'un substituant en position 4 induit généralement une perte d'affinité, sauf dans le cas de la 4-*trans*-fluoroproline (40 µM). Devant les possibilités restreintes de modifications pharmacophoriques autour de ces deux éléments, l'introduction de motifs d'interaction supplémentaires a été envisagée et guidée par la connaissance de la structure du domaine WW en complexe avec des peptides phosphorylés. L'introduction de groupements hydrophobes sur l'amine de la phosphothréonine et/ou de groupements polaires sur la fonction acide de la proline a permis d'augmenter significativement l'affinité des composés (jusqu'à 5 µM). Le *mapping* des déplacements chimiques du domaine WW en présence de ces ligands plus complexes a permis d'identifier un nouveau site d'interaction du domaine WW ciblé par les groupements hydrophobes portés par le ligand (Smet, Duckert et al. 2005).

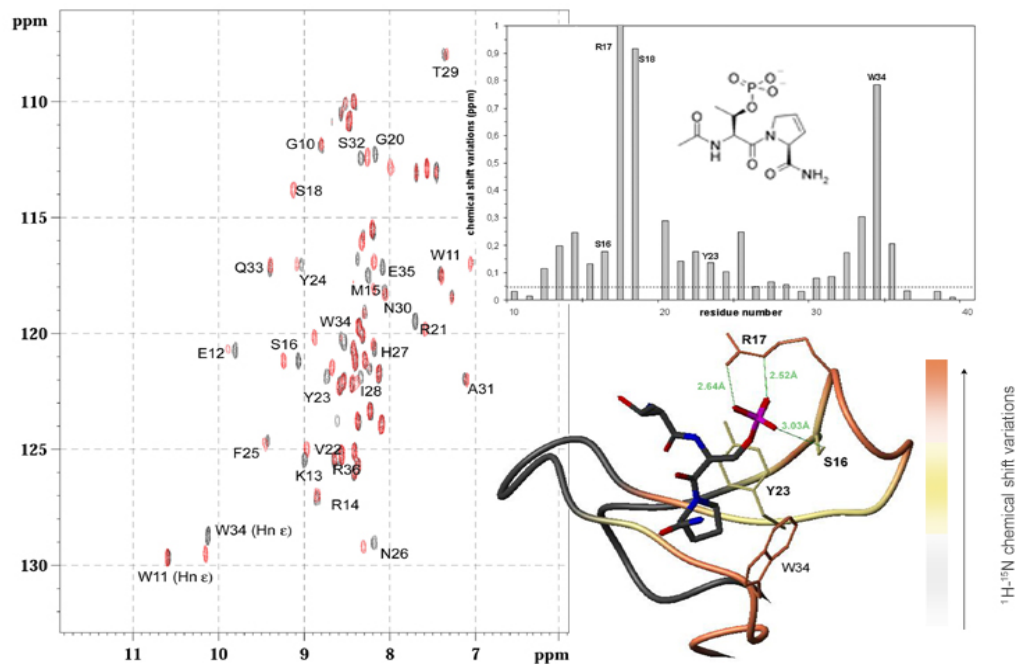


Figure 2.8 : Cartographie de l'interaction des ligands du domaine WW de Pin1 par RMN. Variations des déplacements chimiques des résidus du domaine WW en l'absence de ligand (spectre noir) et à saturation (spectre rouge, ratio [WW] : [ligand] 1 :20) observées sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC du domaine WW enrichi en  $^{15}\text{N}$ . L'amplitude des variations des résonances est reportée sur un graphique en fonction de la séquence primaire du domaine WW et sur la structure tridimensionnelle du domaine WW (représentation en ruban) en complexe avec le dipeptide phospho-Thr-Pro (représentation en bâtonnets).

Dans des expériences de compétition, mesurées par déplacement du FRET entre le domaine WW et un phospho-peptide issu de la protéine Tau marqué en position N-terminale par un fluorophore, nous avons montré que les composés sont capables de déplacer le peptide phosphorylé du domaine WW, avec des  $K_i$  comparables aux  $K_D$  déterminées par RMN. De plus, dans des lignées cellulaires (neuroblastomes SY5Y) transfectées de manière stable avec Tau et surexprimant Pin1 par induction, il a été montré une corrélation entre l'inhibition fonctionnelle de Pin1 par la juglone (un inhibiteur spécifique de Pin1) et la diminution du niveau de cycline D1 dans les neurones (Hamdane, Smet et al. 2002). Ces effets ont été reproduits avec un des composés synthétisés, avec des doses similaires à la juglone (Smet, Duckert et al. 2005). Il faut souligner ici que la présence de phosphate dans les composés synthétisés, en plus d'être instable d'un point de vue métabolique, est un obstacle à l'entrée des composés dans la cellule. L'amélioration de l'affinité des composés comportant un point d'interaction supplémentaire, de nature hydrophobe, a permis de revisiter le remplacement du phosphate. Le groupement squaramide, notamment, a pu être utilisé en remplacement du phosphate avec des  $K_D$  relativement faibles mais mesurables (compris entre 500  $\mu\text{M}$  et 1 mM) (Smet, Duckert et al. 2005), ce qui laisse de nouvelles perspectives de synthèse pour améliorer la stabilité métabolique des composés ainsi que leur perméabilité membranaire.

***Inhibition des interactions protéine-protéine Pin1/Tau phosphorylée par la phosphorylation du domaine WW (Smet-Nocca, Launay et al. 2013)***

Les modules d'interactions protéine-protéine qui lient les résidus sérine ou thréonine phosphorylés sont impliqués dans la transduction de signaux spécifiques, notamment ceux liés à la régulation du cycle cellulaire. Le domaine WW de Pin1 joue un rôle essentiel en ciblant Pin1 à ses substrats de manière phospho-dépendante et est essentiel pour la fonction de Pin1 *in vivo* (Lu, Zhou et al. 1999; Winkler, Swenson et al. 2000). Si la phosphorylation des protéines-substrats est requise pour la liaison à ces domaines, leur propre phosphorylation pourrait également réguler la liaison de ces derniers aux protéines phosphorylées. Il a été montré que le domaine WW de Pin1 est phosphorylé *in vitro* et *in vivo* par PKA sur le résidu S16 (Lu, Zhou et al. 2002). La phosphorylation du domaine WW régule sa capacité à recruter des protéines phosphorylées et la localisation subcellulaire de Pin1. Un mutant Pin1-S16A induit l'arrêt en mitose et l'apoptose. Dans les tissus sains, la phosphorylation du domaine WW sur la sérine 16 inhibe sa fonction dans la liaison des substrats phosphorylés et conduit à l'inactivation de Pin1. En revanche, dans les tissus de cancer du sein, le domaine WW est majoritairement retrouvé sous une forme non phosphorylée, active (Lu, Zhou et al. 2002). De plus, dans un modèle de tauopathie et dans des cerveaux MA, Pin1 est hyperphosphorylée sur la Ser16 (Ando, Dourlen et al. 2013).

Nous avons produit le domaine WW de Pin1 phosphorylé par PKA *in vitro* (WW<sup>PKA</sup>). L'analyse par RMN du domaine WW<sup>PKA</sup> est décrite en détails dans le Chapitre 3, qui décrit l'utilisation de la RMN comme technique analytique pour la caractérisation des PTM. Brièvement, le résidu Ser16 a été identifié comme unique site de phosphorylation par PKA, par comparaison des déplacements chimiques <sup>13</sup>Cβ entre les formes phosphorylée et non phosphorylée, et la phosphorylation est quasi-quantitative. Le spectre <sup>1</sup>H-<sup>15</sup>N HSQC du domaine WW<sup>PKA</sup> montre que la phosphorylation de la Ser16, localisée dans la boucle de liaison du phosphate entre les feuilletts β1 et β2, affecte la largeur des résonances de cette boucle. Une titration du domaine WW<sup>PKA</sup> par un peptide de Tau phosphorylé en position Thr212 montre une diminution significative de l'affinité du domaine WW pour ce substrat (K<sub>D</sub> estimé à 5 mM vs. 120 μM pour le domaine WW non phosphorylé). L'index des déplacements chimiques basé sur les valeurs des <sup>1</sup>H<sub>N</sub>, <sup>15</sup>N<sub>H</sub>, <sup>13</sup>Cα, <sup>13</sup>Cβ et <sup>13</sup>CO et les constantes de couplage <sup>3</sup>J(H<sub>N</sub>-Hα) calculées à partir de l'expérience HNHA sont similaires pour les domaines WW<sup>PKA</sup> et WW indiquant que la phosphorylation perturbe peu la structure tridimensionnelle du domaine WW. En revanche, la dynamique de la boucle et les échanges avec l'eau sont modifiés par la phosphorylation d'après les expériences de relaxation <sup>15</sup>N et de mesure des vitesses d'échanges avec l'eau (CLEANEX). De manière remarquable, les variations des résonances des résidus de la boucle sous l'effet de la phosphorylation sont similaires à celle du domaine WW non phosphorylé lors de la liaison à un phospho-peptide (voir Chapitre 3, Figure 3.3). La modélisation du domaine WW<sup>PKA</sup> par dynamique moléculaire montre que le phosphate porté par la Ser16 entre compétition avec le phosphate du substrat lors de la liaison d'un phospho-peptide (Figure 2.9), contribuant à la perte de l'affinité du domaine WW<sup>PKA</sup> pour les phospho-protéines (Smet-Nocca, Launay et al. 2013). Comme nous avons montré que le domaine WW joue un rôle positif sur l'activité PPIase de Pin1 probablement en ancrant Pin1 à son substrat (Smet, Wieruszeski et al. 2005), la phosphorylation du domaine WW, en régulant négativement la liaison des substrats phosphorylés, pourrait induire une diminution de l'activité PPIase de Pin1.

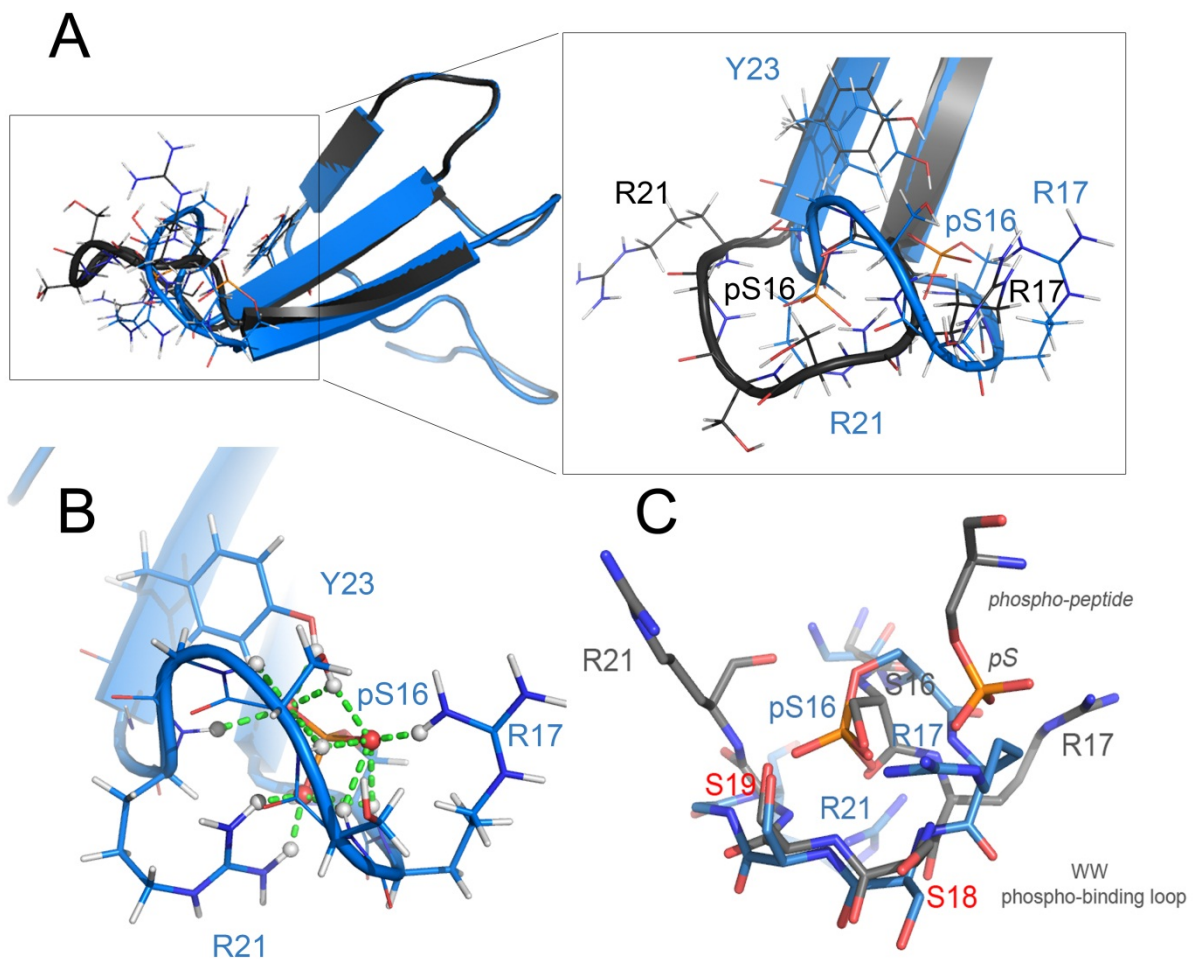


Figure 2.9: Structure du domaine WW phosphorylé (WW<sup>PKA</sup>) obtenue par dynamique moléculaire. (A) Comparaison des structures du domaine WW non phosphorylé (PDB ID 1F8A), représenté en gris, et du domaine WW<sup>PKA</sup>, représenté en bleu, selon un alignement global. (B) Détail de la boucle impliquée dans la liaison du phosphate, dans sa forme phosphorylée sur le résidu Ser16, montrant (en trait pointillés verts) les contacts inter-résidu stabilisant la phospho-Ser16. (C) Comparaison des structures de la boucle du domaine WW lié à un peptide phosphorylé (en gris) et celle du domaine WW<sup>PKA</sup> (en bleu). Ici, l'alignement des structures des deux protéines a été fait sur les résidus de la boucle.

### **Liaison allostérique de l'inhibiteur-2 de la phosphatase PP1 (IPP-2) à Pin1.**

Comme Pin1 reconnaît les protéines phosphorylées au niveau de motifs phospho-Ser/Thr-Pro et que l'inhibiteur-2 de phosphatase PP1 (IPP-2) est phosphorylé pendant la mitose au niveau d'un site PSpTP, ce site pourrait être un site de reconnaissance pour Pin1. De manière surprenante, il a été proposé que la forme non phosphorylée d'IPP-2 est préférentiellement reconnue par Pin1 et qu'IPP-2 serait un ligand allostérique de Pin1, modifiant sa spécificité vis-à-vis des substrats (Li, Stukenberg et al. 2008). Nous avons déterminé l'organisation du complexe Pin1 : IPP-2 par *GST-pull down*, chromatographie d'exclusion et RMN. Bien que l'homologie de séquence entre les protéines IPP-2 humaine, de xénope et de drosophile ne soit que de 50%, elles présentent toutes une même capacité de liaison à GST-Pin1, avec des constantes d'inhibition de 0.3  $\mu$ M. La marquage <sup>15</sup>N de l'un des

partenaires en présence de l'autre non marqué a permis de cartographier l'interaction Pin1/IPP-2. D'une part, les résonances  $^1\text{H}$ - $^{15}\text{N}$  des domaines WW et catalytique de Pin1 varient avec l'addition d'IPP-2 impliquant les deux modules dans l'interaction avec IPP-2. D'autre part, des régions spécifiques d'IPP-2 sont affectées par Pin1. Nous avons ainsi pu identifier le résidu Ile68 comme important pour la liaison à Pin1 via le domaine WW : la mutation I68A diminue de moitié la liaison de Pin1 et empêche la liaison du domaine WW isolé. De même, la troncation d'IPP-2 au niveau de l'acide aminé 152 (sur 205) modifie peu l'interaction avec le domaine WW mais élimine celle avec le domaine PPlase de Pin1. La chromatographie d'exclusion a montré que Pin1 et IPP-2 forme un large complexe macromoléculaire dont la taille dépasse 300 kDa tandis que le mutant I68A d'IPP-2 est impliqué dans un complexe deux fois plus petit. Nous avons proposé que Pin1 et IPP-2 s'associent pour former un hétérotétramère  $(\text{Pin1})_2(\text{IPP-2})_2$  et que le complexe formé par ces deux protéines conservées chez les eucaryotes est un mode de régulation de la spécificité et la fonction de Pin1.

## 2.3 DISCUSSION.

### ***2.3.1. Rôle moléculaire de Pin1 dans la reconnaissance de substrats phosphorylés***

Le rôle moléculaire de Pin1 reste encore largement incompris plus de 15 ans après sa découverte (Lu and Hunter 1995; Lu, Hanes et al. 1996) mais a été l'objet de beaucoup d'attentions. Son rôle est essentiel dans le déroulement du cycle cellulaire et un grand nombre de phospho-protéines, notamment des phospho-protéines mitotiques, ont été décrites comme substrat de Pin1 (Lu 2000; Lu, Liou et al. 2002; Lu, Finn et al. 2007). Son implication a également été mise en cause dans certaines pathologies comme les cancers et la maladie d'Alzheimer (Lu 2003; Lu, Liou et al. 2003; Lu 2004; Lu, Suizu et al. 2006; Lu and Zhou 2007). Ce sont le rôle de Pin1 dans cette dernière pathologie et l'interaction avec la protéine Tau qui nous ont intéressés. Il faut souligner que Pin1 interagit également (au niveau du motif phospho-Thr668-Pro) avec l'autre marqueur pathologique de la maladie d'Alzheimer, la protéine APP, et stimule la production du peptide A $\beta$ , qui est le principal constituant des plaques amyloïdes extra-neuronales (Ramelot and Nicholson 2001; Akiyama, Shin et al. 2005; Pastorino, Sun et al. 2006; Balastik, Lim et al. 2007).

Bien que de nombreuses structures de Pin1 en complexe avec divers substrats existent, les informations apportées par ces structures restent limitées du fait de la nature même de ces substrats, de petits peptides ou petites molécules. De ce fait, l'interaction complexe entre Pin1 et des protéines multi-phosphorylées comme Tau, qui comporte 17 motifs phospho-Ser/Thr-Pro donc autant de sites d'interaction potentiels, n'est pas encore comprise. Nos données indiquent que Pin1 reconnaît plusieurs sites phospho-Ser/Thr-Pro de Tau avec des affinités modestes plutôt qu'un seul site avec une haute affinité comme décrit par Lu et collaborateurs (Lu, Wulf et al. 1999). Cet aspect de régulation par des phosphorylations multiples a déjà été décrit dans la dynamique du cycle cellulaire avec des travaux montrant comme ces événements de phosphorylation multiple sont coordonnés pour se traduire par une activité de type "tout ou rien" (Deshaies and Ferrell 2001). Il faut souligner que de nombreux substrats de Pin1, comme l'ARN polymérase II par exemple (Albert, Lavoie et al. 1999; Xu, Hirose et al. 2003; Xu and Manley 2007), contiennent plusieurs sites de

phosphorylation dans des régions désordonnées (Lu, Liou et al. 2002; Albert, Lavoie et al. 2004; Lu 2004; Lu 2004).

La fonction cellulaire de Pin1 n'est pas encore bien définie bien qu'il soit généralement admis que Pin1 catalyse l'isomérisation *cis/trans* des liaisons peptidiques phospho-Ser/Thr-Pro dans les phospho-protéines, régulant ainsi leur fonction biologique dans le transport intracellulaire, les voies de signalisation, la transcription, la progression du cycle cellulaire et l'apoptose. Cependant, on peut se demander, sur base des données actuelles, si l'activité PPlase de Pin1 est sa fonction principale ou s'il s'agit d'une activité artefactuelle observée *in vitro* qui n'existe pas dans le contexte cellulaire. En cause, la difficulté de mettre en évidence une activité qui n'impose qu'une modification conformationnelle du substrat et non une modification covalente, comme les (autres) PTM. D'autre part, la redondance des PPlases dans la cellule, regroupées en trois familles : les cyclophilines, les *FK506-binding proteins* (FKBP) et les parvulines, avec des similitudes au niveau structural et fonctionnel (Zhang, Daum et al. 2007), complique la détermination du rôle fonctionnel de Pin1 au niveau cellulaire. Ainsi, il faudra déterminer si l'activité primaire de Pin1 est l'activité PPlase ou la liaison des motifs phospho-Ser/Thr-Pro et, dans les deux cas, quelles en seront les conséquences pour le substrat (régulation de l'activité, des interactions protéine-protéine, de la localisation cellulaire, de la dégradation...). Ces deux fonctions, liaison et activité PPlase, sont portées par deux domaines distincts qui ont, néanmoins, les mêmes spécificités structurales, à savoir la reconnaissance des dipeptides phospho-Ser/Thr-Pro avec peu d'influence de la séquence primaire environnante (Ranganathan, Lu et al. 1997; Yaffe, Schutkowski et al. 1997; Verdecia, Bowman et al. 2000; Wintjens, Wieruszeski et al. 2001; Smet, Sambo et al. 2004) : elles sont donc difficiles à dissocier. Le domaine catalytique présente cependant une plus large spécificité de reconnaissance que le domaine WW, incluant la reconnaissance des motifs Glu-Pro, par exemple (Yaffe, Schutkowski et al. 1997; Eidenmuller, Fath et al. 2000; Sami, Smet-Nocca et al. 2011). De plus, il est difficile de dissocier liaison et catalyse dans la mesure où des mutants du site actif perdent également leur capacité de liaison des substrats phosphorylés (Behrsin, Bailey et al. 2007). Le domaine PPlase pourrait simplement être un second site de liaison de plus faible affinité, dont l'occupation nécessiterait une augmentation de la concentration locale en substrat, obtenue par la liaison du domaine WW (sur le même site ou sur un site voisin) (Smet, Wieruszeski et al. 2005).

La fonction PPlase de Pin1 est d'autant plus intrigante que la plupart des motifs phospho-Ser/Thr-Pro reconnus par Pin1 sont localisés dans des régions désordonnées, caractérisées par une minorité de proline en conformation *cis*. Le rôle d'une activité PPlase ciblant les régions désordonnées au sein des protéines-substrats reste une question entière. Il a été proposé que Pin1, en stabilisant l'état de transition entre les états *cis* et *trans*, sans changer la population relative de ces deux conformations, pouvait faciliter l'interaction de l'un ou l'autre des deux conformères lors d'une interaction avec un troisième partenaire conformère-spécifique, menant ainsi au déplacement de l'équilibre conformationnel. Cependant, la plupart des partenaires des substrats de Pin1, comme le complexe kinase CDK2/cyclineA (Brown, Noble et al. 1999) ou la phosphatase PP2A (Zhou, Kops et al. 2000) montrent une spécificité pour le conformère *trans*, qui est décrit comme le conformère majoritaire dans ces fragments protéiques ou protéines totalement non structurés coïncidant avec les sites de phosphorylation (Iakoucheva, Radivojac et al. 2004). Le bénéfice d'une accélération de l'isomérisation conformationnelle catalysée par Pin1, dans ce contexte, ne semble pas donc trivial.

### 2.3.2. Rôle de l'interaction Pin1/Tau phosphorylée dans la maladie d'Alzheimer

Pin1 pourrait jouer un rôle dans la MA à la fois via la régulation de facteurs de transcription (Ryo, Nakamura et al. 2001; Wulf, Ryo et al. 2001; Wulf, Liou et al. 2002; Zacchi, Gostissa et al. 2002; Zheng, You et al. 2002; Barone, Desouza et al. 2008) ou via son rôle de régulation post-traductionnelle. Pin1 est une protéine essentiellement nucléaire, qui régule le cycle cellulaire et prévient l'entrée en mitose (Joseph, Yeh et al. 2003). C'est cette fonction de Pin1 qui pourrait intervenir dans la réactivation anormale du cycle cellulaire dans les neurones différenciés. Il a été observé en effet, que Pin1 augmente le niveau de cycline D1 dans les neuroblastomes, ce qui pourrait avoir pour conséquence de faciliter la transition de la phase G0 à la phase G1, une dédifférenciation neuronale anormale, conduisant vers l'apoptose (Hamdane, Smet et al. 2002; Hamdane, Delobel et al. 2003). D'autre part, Pin1 a été co-localisée dans les PHF des cerveaux de patients atteints de la MA (Lu, Wulf et al. 1999; Ramakrishnan, Dickson et al. 2003). Il a été proposé que Pin1, en se liant à la protéine Tau hyperphosphorylée, soit ainsi séquestrée dans les PHF et maintenue dans le cytoplasme, et ne pourrait plus assurer sa fonction dans la régulation du cycle cellulaire au niveau nucléaire, contribuant ainsi à la mort neuronale (Lu, Wulf et al. 1999; Lu, Liou et al. 2003). De plus, des souris Pin1 *knock-out* développent des caractéristiques de tauopathies liées à l'âge comme des déficits moteurs et comportementaux, une hyperphosphorylation de Tau, la formation de PHF et une dégénérescence neuronale, impliquant Pin1 comme facteur neuro-protecteur (Liou, Sun et al. 2003). Un des rôles les plus convaincants de Pin1 concerne la stimulation de la déphosphorylation des phospho-protéines (Zhou, Lu et al. 1999; Galas, Dourlen et al. 2006; Landrieu, Smet-Nocca et al. 2011; Kimura, Tsutsumi et al. 2013). La perturbation de l'équilibre entre Pin1 et ses substrats phosphorylés conduirait à une boucle d'amplification du phénomène d'hyperphosphorylation, via notamment un défaut de déphosphorylation par PP2A.

Dans le cas de l'interaction Pin1/phospho-Tau, de nombreuses questions concernant le rôle de Pin1 persistent, incluant le(s) mécanisme(s) moléculaire(s) (liaison et/ou activité PPlase). Pin1 pourrait agir indirectement en stimulant la déphosphorylation de Tau par PP2A (Zhou, Kops et al. 2000; Hamdane, Dourlen et al. 2006; Landrieu, Smet-Nocca et al. 2011), favorisant ainsi une forme moins phosphorylée, capable de se lier aux microtubules. Une des hypothèses est que Pin1 pourrait servir de module d'ancrage de PP2A aux protéines cibles à la manière des sous-unités CKS (*cyclin-dependent kinase subunit*) des CDK. Pin1 pourrait également agir directement par interactions avec la protéine Tau phosphorylée en imposant, par exemple, une modification conformationnelle à cette dernière permettant de réguler sa fonction dans la liaison des microtubules et la polymérisation de la tubuline. Pour cela, il faudra déterminer au préalable, l'effet de la phosphorylation sur la structure de la protéine de Tau et comment ces phosphorylations multiples modifient la liaison de Tau aux microtubules et sa capacité à favoriser leur polymérisation, avant d'envisager le rôle de Pin1 dans ces processus.

Des études structurales basées sur de petits peptides de Tau ont montré un faible pourcentage de conformations *cis* (moins de 10%) pour les prolines impliquées dans les motifs phospho-Ser/Thr-Pro (Daly, Hoffmann et al. 2000; Smet, Sambo et al. 2004). Nos études par RMN de la protéine Tau entière ou de fragments plus courts montrent les mêmes populations de conformères *cis* que dans les peptides (Lippens, Wieruszkeski et al. 2004; Smet, Leroy et al. 2004; Lippens, Sillen et al. 2006). Il a été montré que des phosphorylations multiples pourraient créer des contraintes conformationnelles

locales favorables à une conformation *cis* pour une ou plusieurs prolines dans la région riche en prolines de Tau (Bielska and Zondlo 2006). Cependant, aucune proline en conformation *cis* n'a été détectée dans le contexte de la protéine Tau libre mais il pourrait en être autrement dans le cas d'une protéine Tau associée aux microtubules. Cette région est particulièrement intéressante car elle présente une concentration importante en motifs phospho-Ser/Thr-Pro et joue un rôle prépondérant dans la régulation de la liaison de Tau aux microtubules (Goode, Denis et al. 1997). Des premières données structurales sur le domaine riche en prolines phosphorylé par le complexe CDK2/cycline, obtenues dans notre groupe, indiquent, dans un contexte de protéine globalement désordonnée, la formation d'une courte hélice transitoire allant du phospho-résidu pSer235 jusqu'au début du domaine de liaison aux microtubules (Sibille, Huvent et al. 2011). Ces données montrent une stabilisation de l'hélice dépendante de la séquence puisque le phospho-épitope AT180 (pThr231/pSer235) stabilise cette hélice tandis que la phosphorylation des résidus Thr212 et Thr217 n'a pas d'impact sur cette structure secondaire locale transitoire. De plus, des travaux utilisant différents mutants de la protéine Tau et un test de polymérisation de la tubuline, montrent qu'au moins trois phosphorylations parmi les sites Ser202/Thr205 (AT8) et Thr231/Ser235 (AT180) dans cette même région riche en prolines, sont nécessaires pour perdre l'activité sur la polymérisation de la tubuline bien que la protéine phosphorylée sur ces deux phospho-épitopes conserve sa capacité de liaison aux microtubules (Amniai, Barbier et al. 2009). Les différents mutants pourront être testés en présence de Pin1 pour mettre en évidence quels sites de phosphorylation sont importants pour la liaison et l'activité de Pin1.

La complexité liée aux phosphorylations multiples limite l'étude et la compréhension du rôle de phosphorylations spécifiques dans l'agrégation, la liaison à la tubuline ou les interactions avec d'autres protéines comme Pin1. L'utilisation de phospho-peptides a permis de simplifier le modèle à un ou deux sites de phosphorylation mais ne rend évidemment pas compte du contexte qui est celui d'une protéine de 441 résidus, avec ce que cela sous-tend, notamment la présence de structures résiduelles ou d'une structuration globale, la capacité à lier la tubuline ou à former des oligomères. Nous allons envisager l'étude structurale et fonctionnelle de protéines Tau semi-synthétiques obtenues par ligation chimique native. Cette approche va permettre l'introduction quantitative de PTM spécifiques via la synthèse peptidique en phase solide de fragments de Tau, qui seront ensuite reliés, par un lien peptidique natif, au reste de la protéine obtenue par expression recombinante dans *E. coli*. Nous allons ainsi étudier des phospho-épitopes spécifiques mais aussi d'autres PTM, comme l'acétylation ou l'O-GlcNAcylation. La RMN va permettre de caractériser les protéines semi-synthétiques au niveau structural avant d'étudier l'impact fonctionnel de ces modifications. Parallèlement, nous allons étudier ces PTM récemment décrites que sont l'acétylation et l'O-GlcNAcylation, en faisant appel à la RMN comme technique analytique pour la caractérisation des profils d'acétylation ou d'O-GlcNAcylation obtenus après modification *in vitro* par des enzymes recombinantes. Cette stratégie nécessite la production des enzymes intervenant dans l'ajout de ces PTM et la caractérisation préalable des sites de modifications avant une étude fonctionnelle des protéines Tau modifiées. Ces projets sont décrits dans le Chapitre 4. Nous allons d'abord aborder la question de l'utilisation de la RMN comme technique analytique des PTM dans le Chapitre 3.



## Chapitre 3 : La Résonance Magnétique Nucléaire dans la caractérisation des modifications post-traductionnelles

---

### 3.1 INTRODUCTION

Les PTM constituent un mode de régulation de la structure et/ou de la fonction des protéines après leur synthèse en introduisant un niveau de complexité supplémentaire qui n'est pas directement codé dans le gène. Elles permettent une régulation spatio-temporelle de l'activité des protéines dans la cellule en modifiant la surface d'interactions avec leurs partenaires cellulaires, leur localisation sub-cellulaire, leur durée de vie... Elles nécessitent le recrutement, à un moment donné, d'enzymes spécifiques qui participent à la modification dynamique du substrat via l'introduction et la coupure d'un groupement chimique ou d'une protéine sur un résidu spécifique, ou la protéolyse du substrat. En plus des mécanismes de régulation transcriptionnelle et d'épissage alternatif, qui interviennent au niveau du gène et de l'ARN messager, respectivement, les PTM induisent une grande diversité au niveau du protéome et sont donc impliquées dans la régulation de processus biologiques variés tels que le cycle cellulaire et la différenciation, les voies métaboliques, le transport intracellulaire et les voies de transduction de signaux, la communication cellule-cellule, l'apoptose... Les erreurs au niveau de l'incorporation des PTM sont impliquées dans un certain nombre de pathologies comme les cancers, les maladies neurodégénératives, les diabètes et désordres métaboliques, d'où l'intérêt de comprendre leur régulation et comment elles interfèrent dans la fonction de leurs divers substrats. En effet, le problème est complexe puisque (i) l'introduction et l'impact d'une PTM sont directement dépendantes du substrat modifié car codée, en premier lieu, dans la structure primaire, (ii) la modification peut être liée à la présence ou à l'absence d'autres PTM à proximité, (iii) l'identification d'une PTM est dépendante de la nature chimique de celle-ci, du contexte (PTM multiples, présence ou non de sites de clivage protéolytique, présence ou non d'une séquence consensus,...). L'interprétation quantitative aussi bien que qualitative des PTM est une question essentielle et complexe du fait de leur nature dynamique et de leur aspect multiple, séquentiel ou combinatoire.

La RMN se présente comme une approche nouvelle pour l'étude des PTM. Sa nature non destructive permet, à partir d'un même échantillon, de caractériser les PTM à la fois sur le plan analytique, en termes de position et quantification, et sur le plan structural et fonctionnel. La sensibilité du déplacement chimique vis-à-vis de modifications de nature chimique (covalentes) ou conformationnelle (non covalentes) peut être mise à profit pour la détection, l'identification et la quantification des PTM. Si, actuellement, la spectrométrie de masse est la technique de choix pour l'étude des PTM, la RMN peut néanmoins apparaître comme une technique complémentaire (Prabakaran, Everley et al. 2011). Loin de se suffire à elle-même dans une approche analytique, même à l'échelle d'une protéine unique, purifiée, la RMN peut être d'une aide précieuse dans l'identification de PTM, utilisant des approches méthodologiques dépendantes de la modification étudiée. Nous allons présenter ici, sommairement, les caractéristiques spécifiques de trois PTM que sont la phosphorylation, l'acétylation et l'O-GlcNAcylation, qui vont nous intéresser ensuite pour l'étude de la protéine Tau. Une présentation exhaustive de la caractérisation par RMN de PTM impliquant l'addition covalente de groupements chimiques, ce qui exclut les PTM impliquant

l'addition de protéines telles que l'ubiquitine ou SUMO, est donnée dans la revue (Theillet, Smet-Nocca et al. 2012).

L'introduction d'une nouvelle fonction chimique, de même que toute modification conformationnelle ou toute modification de l'environnement chimique quelle qu'elle soit, va entraîner une modification du déplacement chimique des noyaux du résidu concerné, en particulier du proton amide ( $H_N$ ), mais aussi de ses voisins, qu'ils soient voisins dans la séquence primaire de la protéine ou dans la structure tridimensionnelle. Ceci permet de localiser un site de modification ou d'interaction par *mapping* des déplacements chimiques : on observe alors une ou plusieurs régions de la protéine subissant des variations plus ou moins importantes des déplacements chimiques. Ainsi, le *mapping* des déplacements chimiques ne suffit généralement pas pour identifier un site de PTM. Une analyse des déplacements chimiques proton et carbone est nécessaire pour identifier sans ambiguïté le résidu modifié. De plus, il faut distinguer le cas des protéines non structurées et petits fragments peptidiques, des protéines globulaires. Dans le cas des protéines non structurées, la faible déviation des déplacements chimiques des résidus de la protéine par rapport aux valeurs *random coil* permet de détecter facilement des variations attribuables aux PTM. En revanche, dans les protéines structurées, la situation est différente car les écarts par rapport aux valeurs *random coil* sont plus importants du fait de la structure tridimensionnelle. L'effet d'une PTM n'induit pas nécessairement des variations significatives des résonances.

## 3.2 RESULTATS

### 3.2.1. La phosphorylation

La phosphorylation correspond à l'ajout d'un groupement phosphate sur les résidus sérine/thréonine par les sérine/thréonine-kinases ou sur les résidus tyrosine par les tyrosine-kinases. La phosphorylation des protéines se produit principalement dans les régions désordonnées (Iakoucheva, Radivojac et al. 2004). Du fait de la nature chimique même du groupement phosphate, la phosphorylation a un effet important sur le déplacement chimique des résidus sérine et thréonine phosphorylés. Dans le cas des peptides et protéines non structurées, on observe généralement une variation importante de la résonance du proton amide du résidu sérine/thréonine phosphorylé vers les résonances à bas champ, ce qui rend cette modification facilement détectable sur le spectre RMN. Cette résonance s'écarte significativement du "massif" des autres résonances (Figure 3.1), qui sont peu dispersées dans la dimension proton, et se retrouve dans une région vide du spectre  $^1H$ - $^{15}N$  HSQC. Ces résonances sont détectables même dans une expérience  $^1H$  à une dimension, qui peut donc être utilisée pour évaluer rapidement l'état de phosphorylation d'une protéine non structurée. Il a été proposé que l'établissement d'un pont hydrogène intra-résidu entre le phosphate et le proton amide du squelette soit responsable de cette variation prononcée du déplacement chimique du proton amide du résidu phosphorylé (Du, Li et al. 2005). Les résonances amides des résidus voisins du site de phosphorylation sont modifiées également. Cependant, dans les protéines non structurées, les résidus affectés sont ceux qui se trouvent dans le voisinage immédiat du résidu phosphorylé dans la séquence primaire. On estime que seuls les résidus se trouvant à moins de 3 résidus de part et d'autre du site de phosphorylation sont concernés (Figure 3.2).

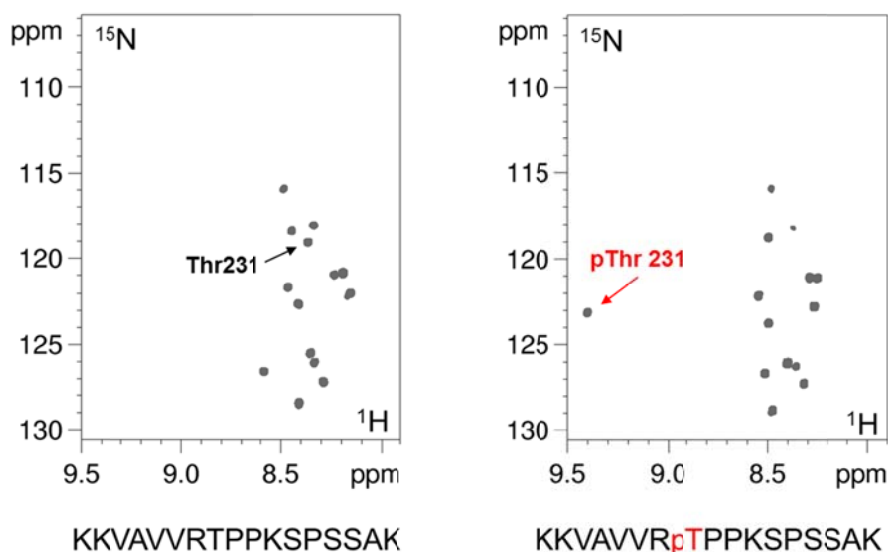


Figure 3.1 : Variation du déplacement chimique du résidu thréonine (T231) entre la forme non phosphorylée (à gauche) et la forme phosphorylée (à droite). Spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC du peptide issu de la protéine Tau (résidu 224 à 240, numérotation de Tau humaine, isoforme de 441 résidus) obtenu en abondance naturelle.

Dans le cas des peptides, l'attribution séquentielle à l'aide des expériences RMN 2D homonucléaires TOCSY et NOESY, ainsi que les expériences hétéronucléaires (en abondance naturelle)  $^1\text{H}$ - $^{13}\text{C}$  HMQC et  $^1\text{H}$ - $^{15}\text{N}$  HSQC permettent d'identifier précisément le(s) site(s) de phosphorylation. En général, on observe des variations importantes du déplacement chimique du proton amide, et plus faibles pour les protons  $\alpha$  et  $\beta$  du résidu phosphorylé. Les résonances des résidus voisins sont également modifiées. L'identification du site de phosphorylation vient des variations des résonances  $^{13}\text{C}$  qui sont significativement plus importantes pour le résidu phosphorylé (Figure 3.2 B).

Dans le cas d'une protéine globulaire, la dispersion des déplacements chimiques de l'ensemble des résonances est plus importante et l'effet de la phosphorylation sur le déplacement chimique du résidu phosphorylé est moins remarquable car la résonance ne s'écarte pas nécessairement des autres. De plus, l'effet de la phosphorylation peut être fortement marqué pour les résidus situés au voisinage dans la structure tridimensionnelle, ce qui rend délicate l'identification du site de phosphorylation. Ce cas de figure est illustré par la phosphorylation du domaine WW de Pin1 par PKA (Smet-Nocca, Launay et al. 2013). Ce domaine possède une boucle  $\beta$ -hairpin, flexible, impliquée dans la liaison des substrats phosphorylés, localisée entre les brins  $\beta 1$  et  $\beta 2$  du triple feuillet  $\beta$ . Il a été montré que cette boucle est phosphorylée au niveau du résidu Ser16 par PKA avec, pour conséquence, un changement de localisation subcellulaire et une inactivation du domaine WW (Lu, Zhou et al. 2002). Cette région du domaine WW contient trois sérines aux positions 16, 18 et 19 dont les résonances  $\text{H}_\text{N}$  sont toutes fortement affectées par la phosphorylation. Les résonances des Ser18 et Ser19 sont même davantage modifiées que celle de la Ser16 sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC (Figure 3.3 A, B, C). De plus, on observe un élargissement des résonances sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC du domaine WW, à des pH supérieurs à 6.4, pour les résidus R17 à S19 qui diminue avec une diminution du pH ou sous l'effet de la phosphorylation.

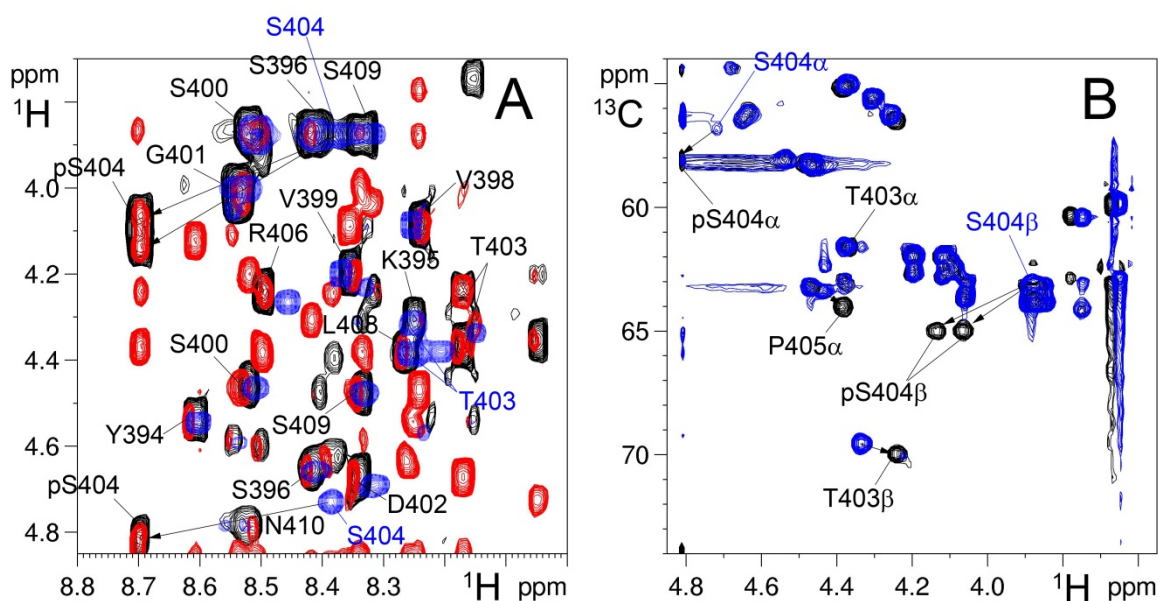


Figure 3.2 : Effet de la phosphorylation sur les résonances  $^1\text{H}$  et  $^{13}\text{C}$  du peptide Tau[392-411] phosphorylé en S404. (A) Attribution séquentielle du peptide Tau[392-411]-pS404 à l'aide des séquences homonucléaires TOCSY (noir) et NOESY (rouge). (B) Spectre  $^1\text{H}$ - $^{13}\text{C}$  HSQC du peptide Tau[392-411]-pS404 (noir). En bleu, spectres TOCSY (A) et  $^1\text{H}$ - $^{13}\text{C}$  HSQC (B) du peptide Tau[392-411] non phosphorylé.

L'analyse par spectrométrie de masse MALDI-TOF indique la présence d'une seule phosphorylation. De même, deux séries de résonances sont observables sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC concernant plusieurs résidus, indiquant un dédoublement de certaines résonances sous l'effet de la phosphorylation : on observe une résonance faible au même déplacement chimique que dans le domaine WW non phosphorylé et une résonance plus intense qui n'apparaît que dans le spectre du domaine WW phosphorylé ( $\text{WW}^{\text{PKA}}$ ). Ces deux signaux correspondent à la résonance d'un résidu du domaine WW dans sa forme non phosphorylée et phosphorylée, respectivement. Ce dédoublement touche également les résonances des résidus au voisinage du site de phosphorylation. Ceci devrait permettre une quantification du niveau de phosphorylation.

Un *mapping* des déplacements chimiques montre que les variations les plus importantes concernent les résidus R17, S18 et S19 de la boucle  $\beta$ -hairpin ainsi que les résidus R21 et Y23 du brin  $\beta_2$ , et E35 (Figure 3.3 A et B). Dans le cas présent, il faut noter que le *mapping* ne permet pas d'identifier sans ambiguïté le site de phosphorylation. De même, alors que dans les protéines non structurées, une comparaison des déplacements chimiques  $^{13}\text{C}$  de la forme phosphorylée avec les valeurs *random coil* suffit généralement pour identifier le site de phosphorylation, dans le cas d'une protéine globulaire, les variations des résonances  $^{13}\text{C}$  sous l'effet de la phosphorylation sont du même ordre que celles imputables aux structures secondaires, et ne permettent donc pas d'identifier facilement le résidu phosphorylé. L'identification de la Ser16 comme site de phosphorylation a pu être réalisée par comparaison des déplacements chimiques  $^{13}\text{C}$  des  $\text{C}\alpha$  et  $\text{C}\beta$  entre les formes non phosphorylée et phosphorylée du domaine WW (Figure 3.4). Ainsi, nous avons observé une forte variation de la résonance du  $\text{C}\beta$  de la Ser16, sous l'effet de la phosphorylation par PKA, par rapport à la même résonance dans la forme non phosphorylée. Ce site a été confirmé par analyse des fragments de digestion tryptique par spectrométrie de masse MALDI-TOF, notamment grâce à la présence d'un

signal correspondant à un incrément de +80 Da pour les peptides  $^{14}\text{RMSR}^{17}$  et  $^{15}\text{MSR}^{17}$ . Il faut souligner que, dans le cas présent, l'identification d'une phosphorylation sur les résidus Ser18 ou Ser19 n'aurait pas été possible du fait de l'absence de site de coupure spécifique de la trypsine entre ces deux résidus.

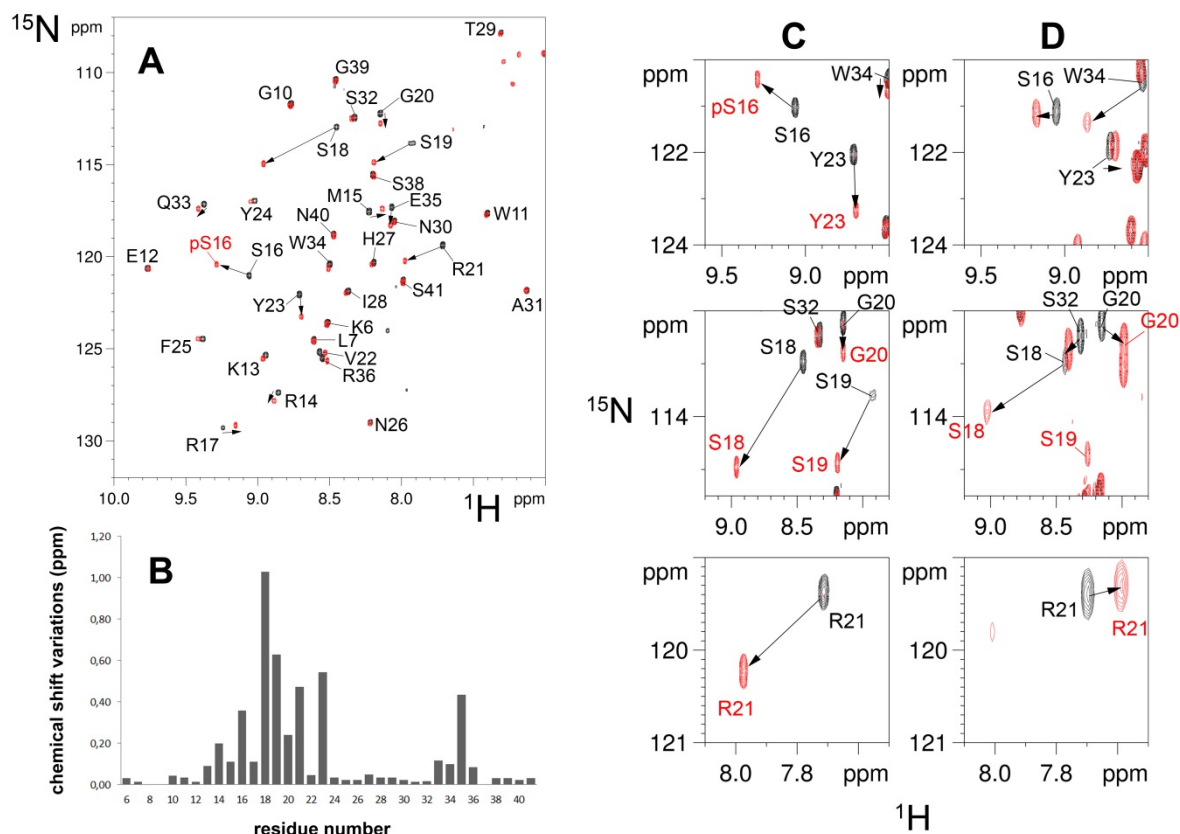


Figure 3.3 : Phosphorylation du domaine WW de Pin1 et liaison d'un substrat phosphorylé. (A,B) Effet de la phosphorylation du domaine WW de Pin1 par PKA sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC. (A) Spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC à pH 5.8 du domaine WW dans sa forme non phosphorylée (noir) et phosphorylée par PKA (rouge). (B) Représentation graphique des variations de déplacements chimiques du proton amide et de l'azote pour chaque résidu (résidus 6 à 41, numérotation Pin1 humaine). (C,D) Comparaison de l'effet de la phosphorylation ou de la liaison d'un substrat phosphorylé sur les résonances du domaine WW. (C) Variations des résonances entre la forme phosphorylée (rouge) et la forme non phosphorylée (noir) pour les résidus de la boucle de liaison du phosphate. (D) Variations de ces mêmes résonances du domaine WW non phosphorylé entre la forme apo (noir) et la forme liée à un peptide phosphorylé comportant un motif phospho-Thr-Pro (rouge).

La quantification d'une PTM est également une question centrale et l'intégration des résonances relatives des formes phosphorylée et non phosphorylée devrait, en théorie, donner le taux de modification. A nouveau, des précautions doivent être prises car des différences significatives ont été observées au niveau des largeurs de raies pour les résonances de la boucle, qui ne donne pas la même précision sur l'intégration des signaux. Dans le cas du domaine WW, à pH 6.8, on n'observe pas les résonances des résidus R17 à S19 du fait de l'élargissement important des résonances qui les rend indétectables. Si on effectue l'intégration des signaux correspondant aux formes phosphorylée et non phosphorylée, on trouvera donc, à partir de ces résonances, 100% de phosphorylation. En

revanche, à des pH plus bas ou par intégration des signaux d'autres résidus, on trouve un niveau de phosphorylation de l'ordre de  $90 \pm 3\%$ . La quantification du niveau de phosphorylation est donc une question délicate qui doit tenir compte d'une modification de la largeur de raie en présence ou en l'absence de phosphorylation sous l'effet d'une modification de la dynamique de la région concernée ou d'une modification de l'échange avec l'eau, car les régions phosphorylées, même dans le cas de protéines globulaires, sont des régions généralement désordonnées et accessibles au solvant.

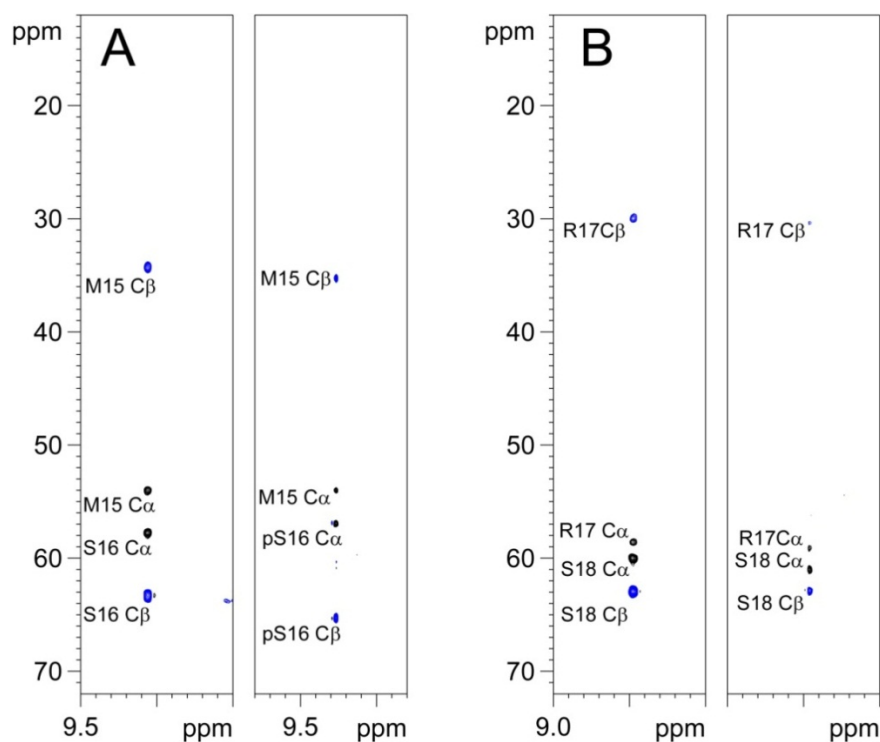


Figure 3.4 : Identification du résidu Ser16 comme site de phosphorylation du domaine WW par PKA. *Strips*  $^{13}\text{C}$  issus de l'expérience HNCACB sur le domaine  $^{13}\text{C}/^{15}\text{N}$ -WW phosphorylé par PKA (A et B, panels gauches) et non phosphorylé (A et B, panels droits) montrant une variation des déplacements chimiques des C $\alpha$  et C $\beta$  sous l'effet de la phosphorylation avec un effet plus marqué pour le C $\beta$  de la S16.

La RMN a été utilisée ici comme outil analytique mais l'un de ses grands avantages est de pouvoir caractériser au niveau structural et fonctionnel l'impact d'une PTM. Dans le cas du domaine WW de Pin1, la caractérisation de la phosphorylation par RMN a été l'étape préliminaire à l'étude de l'effet de cette phosphorylation sur la fonction du domaine WW de Pin1. Nous avons montré que la forme phosphorylée du domaine WW ne peut plus se lier à un peptide phosphorylé de la protéine Tau, au niveau de l'épitope phospho-Thr212 (voir Chapitre 2). Nous n'avons pas détecté de changements structuraux notables sous l'effet de la phosphorylation, notamment la structure en triple feuillet  $\beta$  est conservée. Des études comparatives de la forme non-phosphorylée et phosphorylée ont été réalisées pour expliquer la perte significative de l'affinité du domaine WW pour les substrats phosphorylés. Des études dynamiques, avec la mesure des paramètres de relaxation  $^{15}\text{N}$  (T1, T2 et hétéro-NOE), ainsi que la mesure des vitesses d'échange des protons amides avec l'eau ont permis de mettre en évidence une modification significative de la dynamique de la boucle de liaison des substrats phosphorylés qui contient le site de phosphorylation, de l'ordre de la microseconde à la



milliseconde (relaxation R2) ainsi qu'une diminution de l'échange avec l'eau pour les résidus de la boucle sous l'effet de la phosphorylation de la Ser16. Une minimisation d'énergie et une dynamique moléculaire ont permis de modéliser la conformation de la boucle dans son état phosphorylé. Le modèle a permis d'expliquer les modifications observées par RMN sur les déplacements chimiques, la dynamique de cette boucle et l'échange réduit avec l'eau, et montre que le phosphate porté par la Ser16 pourrait entrer en compétition avec le phosphate du substrat phosphorylé, ce qui explique la perte de fonction observée (Smet-Nocca, Launay et al. 2013).

Ici, la phosphorylation quasiment quantitative (de l'ordre de 90 %) du domaine WW par PKA sur un site unique a facilité l'identification du site de phosphorylation ainsi que la détermination de l'impact de cette phosphorylation spécifique sur la structure et la fonction du domaine WW. Ce cas de figure est rarement atteint et la phosphorylation par les kinases conduit généralement à des phosphorylations multiples et incomplètes et, par conséquent, à une hétérogénéité de l'échantillon qui complique grandement les spectres RMN et leur interprétation (voir Figure 3.5) (Amniai, Barbier et al. 2009; Landrieu, Leroy et al. 2010). Ce qui semble être un obstacle à première vue peut néanmoins apporter beaucoup d'informations sur l'aspect combinatoire des PTM lorsque plusieurs modifications sont localisées sur des sites proches. Alors que la spectrométrie de masse aura des difficultés à identifier des sites proches, liées à l'absence de site de digestion pour des protéases spécifiques ou à la modification de leur activité protéolytique sous l'effet de la PTM même, la RMN pourra détecter deux PTM proches, de moins de 2 ou 3 acides aminés, car chacune d'elles va influencer la résonance de l'autre (Figure 3.6 A, voir les résidus S199 et S202). De même, les résidus voisins d'un site de phosphorylation voient leur résonance dédoublée, avec des intégrations relatives proportionnelles au niveau de phosphorylation de leur voisin (Figure 3.6 A, voir la résonance de A152).

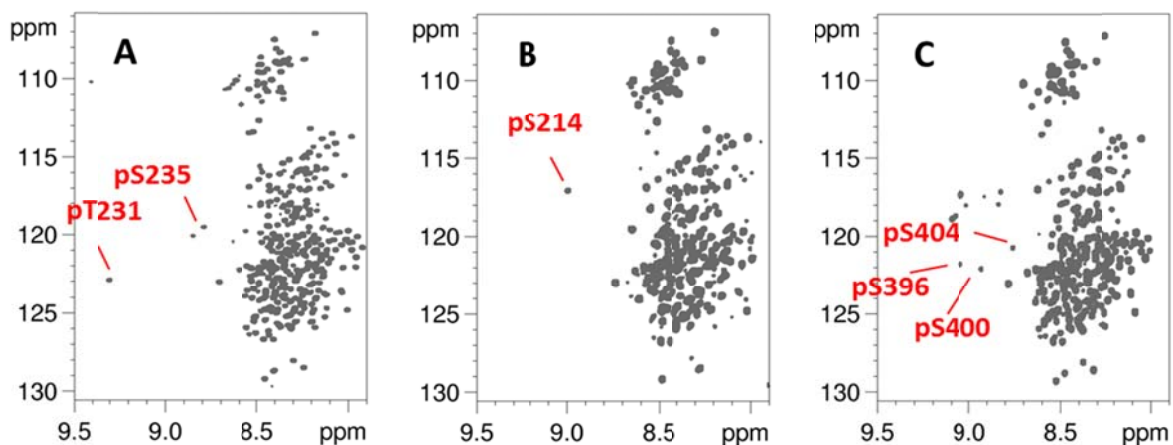


Figure 3.5 : Phosphorylation de la protéine Tau441 par le complexe CDK2/cyclineA3 (A), par PKA (B), par PKA puis GSK3 $\beta$  (C). Quelques sites de phosphorylation sont annotés en rouge. On observe un décalage remarquable vers les bas champs pour les résonances des résidus phosphorylés.

### 3.2.2. L'acétylation

L'acétylation correspond à l'ajout d'un groupement acétyle sur la fonction amine primaire NH $\epsilon$  des résidus lysine, qui se trouve sur le carbone C $\epsilon$  de la chaîne latérale. La fonction amide résultante donne une résonance détectable sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC, alors que la fonction amine primaire, avant acétylation, ne l'était pas du fait de l'échange rapide avec l'eau résultant en l'élargissement de la résonance NH $_3^+$ . De plus, ce groupement NH $_3^+$  résonne à une fréquence  $^{15}\text{N}$  très différente des résonances amides (environ 40-60 ppm pour la résonance NH $_3^+$ , 100-140 ppm pour la résonance H $_N$ ). Cette nouvelle résonance correspondant à la forme NH $\epsilon$  acétylée apparaît dans une région du spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC qui est non peuplée par les résonances des fonctions amide du squelette polypeptidique (à environ 8.01 ppm en proton et 127.4 ppm en azote), dans le cas de peptides ou protéines non structurées. De même, on peut détecter sur le spectre proton, une résonance à 1.98 ppm, sous forme de singulet, correspondant au méthyle du groupement acétyle. Alors que, dans le contexte des protéines non structurées, la phosphorylation induit un *shift* important à bas champ rendant ces PTM facilement détectables, l'acétylation induit, à l'inverse, un *shift* faible à haut champ qui ne permet généralement pas aux résonances de sortir du massif (voir Figures 3.6 B et 3.7).

En revanche, la corrélation scalaire entre le proton amide de la chaîne latérale de la lysine acétylée (NH $\epsilon$ -Ac) et les protons aliphatiques de la lysine permet d'identifier, à partir de l'expérience homonucléaire TOCSY, la(les) lysine(s) modifiée(s) même dans le contexte d'acétylations multiples (Smet-Nocca, Wieruszski et al. 2010). De plus, un *shift* relativement important, d'environ 0.2 ppm à bas champ, du proton H $\epsilon$ , localisé sur le carbone porteur de la fonction N-acétyle permet d'identifier aisément les lysines acétylées. De même, au niveau des résonances  $^{13}\text{C}$ , des variations des déplacements chimiques  $^{13}\text{C}\epsilon$ ,  $^{13}\text{C}\delta$ ,  $^{13}\text{C}\gamma$  permettent de mettre en évidence la présence de résidus lysines acétylés mais restent relativement faibles (de l'ordre de 0.2-0.4 ppm pour  $^{13}\text{C}\epsilon$  et  $^{13}\text{C}\gamma$ , 1.5 ppm pour  $^{13}\text{C}\delta$ ). Elles ne permettent pas de distinguer les différentes lysines, qu'elles soient acétylées ou non, car elles ne sont pas suffisamment résolues dans le contexte de protéines non structurées (Smet-Nocca, Wieruszski et al. 2010).

Dans le cas d'une protéine non structurée, on citera le cas ici de la protéine Tau, l'acétylation par l'histone acétyltransférase CBP (*CREB-binding protein*) conduit à la détection de plusieurs résonances amides peu résolues correspondant à des chaînes latérales de lysines acétylées (NH $\epsilon$ -Ac) indiquant plusieurs sites d'acétylation (Figure 3.6 B, encadré). L'identification des sites d'acétylation dans le contexte d'une protéine entière nécessite une stratégie d'attribution et l'utilisation de séquences impulsionnelles spécifiques qui seront développées dans la partie Projet (voir Chapitre 4). Il faut noter que, les résidus lysines étant la cible de la modification par acétylation, un marquage sélectif des lysines en  $^{15}\text{N}$  présente l'avantage de simplifier considérablement les spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC : on obtiendra un sous-spectre ne montrant que les résidus lysines, acétylés ou non.



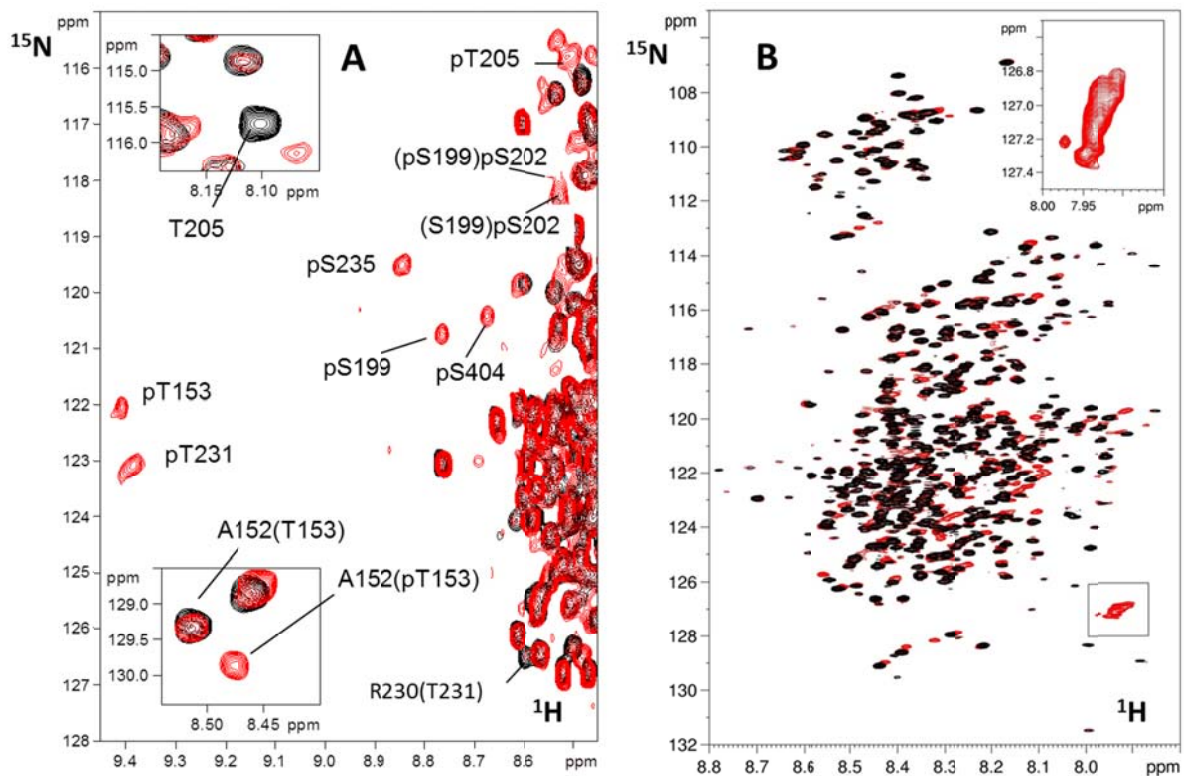


Figure 3.6 : Complexité du spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC de la protéine Tau phosphorylée par le complexe CDK2/cyclineA3 (A) ou acétylée par CBP (B). En noir, spectre de la protéine Tau non modifiée (A,B). (A) La phosphorylation incomplète du résidu T153 induit un dédoublement de la résonance de A152 (en encadré). De même, le dédoublement de la résonance de pS202 est dû à la phosphorylation incomplète de S199. (B) L'acétylation de Tau par le domaine acétyltransférase de CBP (résidus 1202 à 1848, numérotation de CBP de souris) montre une multiplication de certaines résonances avec l'apparition de nouvelles résonances à haut champ. En encadré, les résonances correspondant aux groupements amides (NH $\epsilon$ -Ac) formés au niveau des chaînes latérales des lysines acétylées.

A nouveau, il faut souligner l'intérêt de la RMN pour mettre en évidence l'aspect combinatoire des acétylations multiples via leur impact sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC au niveau du résidu acétylé même et de ses voisins (voir Figure 3.6 B). On illustrera l'effet d'acétylations multiples dans le cas du peptide issu de la Thymine-DNA Glycosylase (TDG) humaine. TDG est acétylée *in vivo* par CBP. L'acétylation réalisée *in vitro* avec CBP recombinante a permis d'identifier, dans des modèles peptidiques d'environ vingt acides aminés, quatre résidus lysine acétylés aux positions 59, 83, 84 et 87 (Tini, Benecke et al. 2002). La synthèse en phase solide de peptides de TDG a permis l'introduction d'acétylations de manière sélective et quantitative, et l'obtention de peptides dans différentes combinaisons d'acétylation. L'acquisition de spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC en abondance naturelle  $^{15}\text{N}$  montre l'influence d'acétylations à différentes positions sur le déplacement chimique du résidu acétylé et des résidus voisins, et d'acétylations multiples (Figure 3.7) (Smet-Nocca, Wieruszkeski et al. 2010).

Dans le cas des trois sites proches K83, K84 et K87, on observe des effets différents de chaque site acétylé sur les résonances  $^1\text{H}$  et  $^{15}\text{N}$ . La résonance amide de la lysine 83 est peu influencée par l'état d'acétylation de la lysine 84 (Figure 3.7 A). En revanche, la résonance amide de la lysine 84 est

différente selon l'état d'acétylation de la lysine 83, que la lysine 84 soit elle-même acétylée ou non. On constate un effet cumulatif de l'acétylation de K83 et K84 sur la résonance de K84 (Figure 3.7 B). De même, la lysine 87 est sensible à l'état d'acétylation des lysines 83 et 84, avec, là aussi, un effet cumulatif sur la résonance de la lysine 87 (Figure 3.7 C). On peut donc extrapoler l'allure du spectre d'une protéine TDG acétylée par la protéine CBP sur ces trois positions où on retrouverait, dans un même spectre, l'ensemble des résonances observées dans ces différents peptides, avec des intensités relatives dépendant de leur taux d'acétylation.

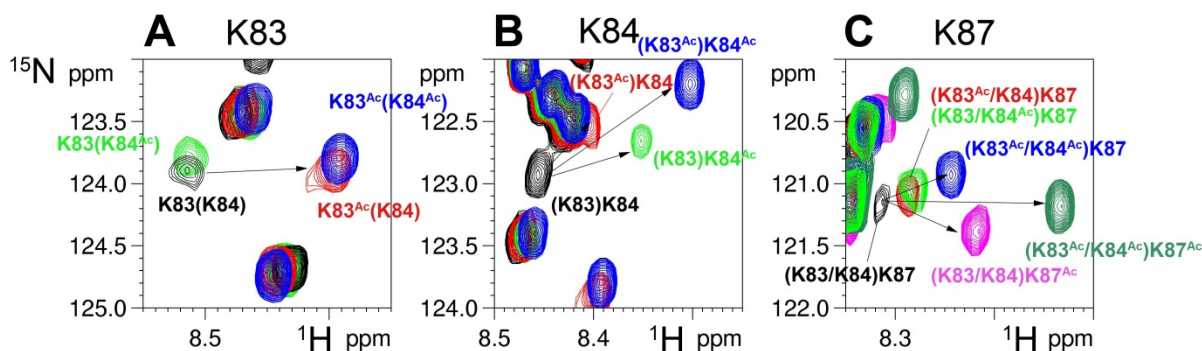


Figure 3.7 : Impact sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC de l'aspect combinatoire de plusieurs acétylations sur un peptide issu de la protéine TDG (résidus 72 à 95, numérotation de la protéine TDG humaine). Effet des différentes acétylation sur la résonance  $\text{H}_\text{N}$  du résidu K83 (A), K84 (B) et K87 (C). Spectres du peptide non acétylé (noir), des peptides mono-acétylés en K83 (rouge), en K84 (vert clair), en K87 (rose), du peptide di-acétylé en K83/K84 (bleu) et tri-acétylé en K83/K84/K87 (vert foncé).

### 3.2.3. L'O- $\beta$ -N-acétylglucosaminylation

L'O- $\beta$ -N-acétylglucosaminylation (ou O-GlcNAcylation) est une PTM ajoutée sur les résidus sérine et thréonine par une enzyme codée, chez l'homme, par un gène unique, l'O-GlcNAc transférase (OGT). Cette PTM est retirée par l'enzyme antagoniste, l'O-GlcNAcase (OGA), codée elle aussi par un seul et unique gène chez l'homme. De manière intéressante, la modification par O-GlcNAc cible les résidus sérine et thréonine comme la phosphorylation et pourrait être un mécanisme compétitif de la phosphorylation, soit pour la modification des mêmes sites, soit dans la régulation réciproque de sites adjacents (Kamemura and Hart 2003; Wang, Gucek et al. 2008). Comme pour l'acétylation, l'O-GlcNAcylation n'aboutit pas à la détection des résidus O-GlcNAc de manière évidente car elle induit des *shifts* modérés avec des déplacements chimiques qui ne s'écartent pas du massif des résonances amide du squelette polypeptidique. Il est parfois possible d'observer des variations plus importantes de la résonance du proton amide du résidu voisin du résidu O-GlcNAcylé, ce qui peut prêter à confusion lorsque celui-ci est un résidu sérine ou thréonine, comme dans le cas du peptide C-terminal issu de la protéine CKII (Dehennaut, Hanouille et al. 2008). L'identification des résidus O-GlcNAc par RMN n'est pas triviale car l'O-GlcNAcylation ne permet pas d'observer de couplages scalaires entre les protons du résidu modifié et ceux du groupement GlcNAc du fait de la liaison osidique (O-glucosidique) qui les relie. Donc, comme pour la phosphorylation, l'identification du résidu O-GlcNAc passe par l'analyse des résonances  $^{13}\text{C}$  (Figure 3.8) (Dehennaut, Hanouille et al. 2008; Smet-NoCCA, Broncel et al. 2011).

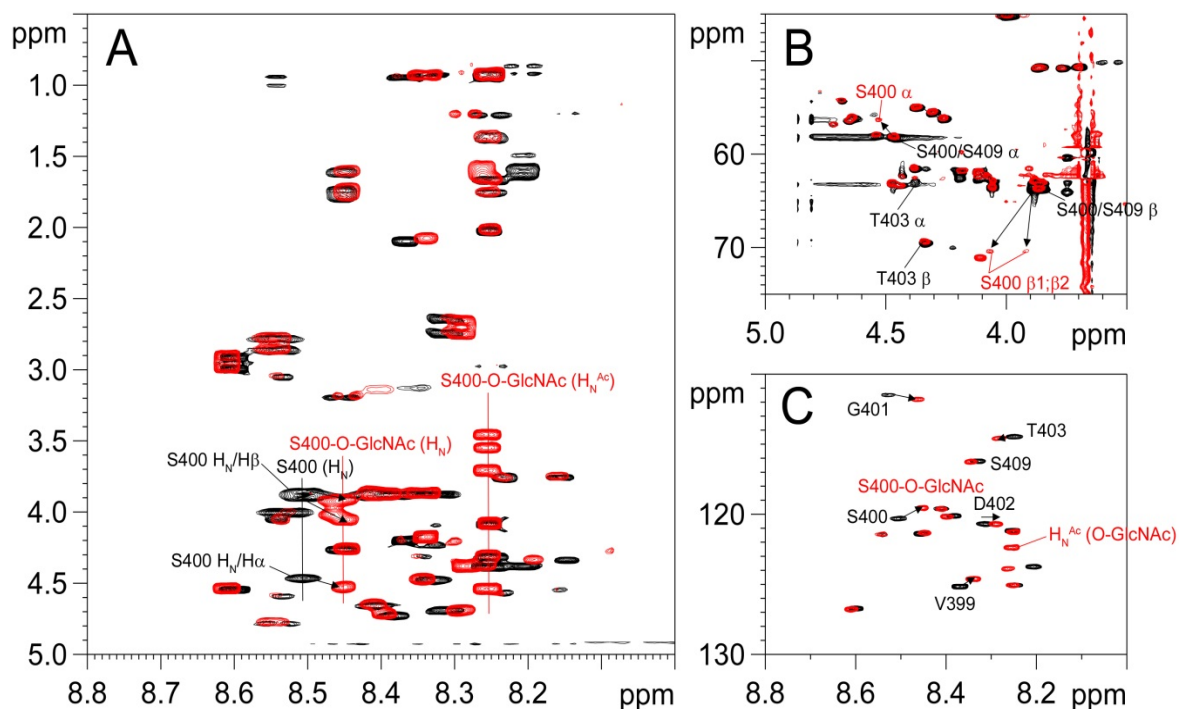


Figure 3.8 : Influence du groupement *O*-GlcNAc sur les résonances  $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}$  du peptide Tau[392-411]. (A)  $^1\text{H}$ - $^1\text{H}$  TOCSY, (B)  $^1\text{H}$ - $^{13}\text{C}$  HSQC et (C)  $^1\text{H}$ - $^{15}\text{N}$  HSQC du peptide Tau[392-411] non modifié (noir) et *O*-GlcNAcylé (rouge).

De plus, comme pour l'acétylation, lors de l'analyse par RMN de peptides non enrichis, une nouvelle résonance amide, correspondant à la fonction glucosamine acétylée, apparaît sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC en abondance naturelle et peut potentiellement renseigner sur la présence d'un groupement *O*-GlcNAc (Figure 3.8 C,  $\text{H}_\text{N}^{\text{Ac}}$  *O*-GlcNAc). Cependant, il faut souligner que, contrairement à la fonction NH-acétylée sur les chaînes latérales des lysines ( $\text{NH}_\epsilon\text{-Ac}$ ), elle ne provient pas de la protéine même mais de l'UDP-GlcNAc, et n'apparaîtra pas dans le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC d'une protéine  $^{15}\text{N}$  enrichie. De plus, elle est localisée dans le massif des résonances amide du squelette polypeptidique et ne peut donc pas être identifiée facilement (Figure 3.8 C).

### 3.3 DISCUSSION : APPORT DE LA RMN A L'ETUDE DES MODIFICATIONS POST-TRADUCTIONNELLES (THEILLET, SMET-NOCCA ET AL. 2012)

#### 3.3.1. Le challenge analytique

La RMN permet de détecter des PTM à l'échelle du résidu. L'introduction d'une PTM va modifier l'environnement local du résidu modifié et de ses voisins permettant ainsi de localiser, grâce à l'expérience  $^1\text{H}$ - $^{15}\text{N}$  HSQC, la région où se trouve(nt) la(les) PTM(s), mais l'identification du site modifié requiert des informations supplémentaires comme la mesure des résonances  $^{13}\text{C}$ . De plus, certaines PTM ont une signature caractéristique ("*indicator*" *signal*) qui permet d'identifier de quel type de modification il s'agit à partir d'une ou quelques résonances facilement identifiable (Theillet, Smet-NoCCA et al. 2012). Ainsi, il n'est pas nécessaire d'attribuer chaque résonance d'une protéine pour savoir, d'un point de vue qualitatif, si elle contient des PTM et d'en identifier la nature à

condition de posséder un spectre de référence de la protéine dans sa forme non modifiée. La nature non destructive de la RMN permet de quantifier le niveau de modification en plus de l'identification du site spécifique, et de mesurer la cinétique de l'addition ou de la suppression d'une PTM dans une réaction reconstituée *in vitro* dans le tube RMN. Par ailleurs, il est aussi possible d'observer des PTM dans un environnement complexe tel qu'un extrait cellulaire ou directement dans la cellule intacte (Bodart, Wieruszkeski et al. 2008; Lippens, Landrieu et al. 2008; Selenko, Frueh et al. 2008; Liokatis, Dose et al. 2010; Bekei, Rose et al. 2012; Bekei, Rose et al. 2012; Bekei, Rose et al. 2012). Dans cet optique, de nombreux efforts sont encore à fournir pour que cette stratégie soit applicable à plus large échelle mais il s'agit probablement de l'aspect le plus intéressant à développer. L'aspect analytique est également couvert par la capacité de détecter plusieurs PTM dans le même échantillon, incluant des PTM de même nature ou de nature différente. La RMN ne nécessite pas de techniques sélectives d'enrichissement ou de purification de l'échantillon, par immuno-précipitation par exemple, qui sont nécessairement directement liées à la nature même de la modification étudiée, et qui empêchent l'analyse quantitative. En revanche, l'intégration des signaux correspondant au résidu dans sa forme modifiée et non modifiée, respectivement, dont les résonances apparaissent dans le même spectre, permet une comparaison directe de l'intensité des signaux, donc une quantification. Cette méthode s'applique assez bien dans le cas des protéines non structurées mais des précautions doivent être prises dans le cas des protéines structurées, où l'introduction ou la suppression d'une PTM peut modifier la dynamique locale et ainsi influencer significativement sur les largeurs de raie, ce qui peut compromettre l'intégration des signaux (Smet-Nocca, Launay et al. 2013).

### **3.3.2. PTM multiples**

Lorsque deux PTM sont suffisamment proches, soit dans la séquence primaire (de 3 à 4 résidus), soit dans la structure globulaire, l'une va directement influencer la résonance du résidu portant l'autre PTM, et *vice versa*. Un des principaux avantages d'étudier les PTM par RMN réside justement dans cette capacité à identifier des PTM multiples lorsqu'elles sont proches dans la séquence de la protéine modifiée, ce qui est un des principaux obstacles de la spectrométrie de masse, qui requiert une fragmentation de la protéine en peptides plus courts par protéolyse (*bottom-up mass spectrometry*) et l'analyse des fragments peptidiques générés par spectrométrie de masse en tandem (MS/MS) pour l'identification et la localisation de PTM. Dans certains cas, les techniques de fragmentation induisent une coupure de la PTM (cas de la phosphorylation ou de l'*O*-GlcNAcylation). La PTM peut également interférer dans la reconnaissance du site de coupure protéolytique ou modifier l'activité de la protéase (cas de l'acétylation ou de la méthylation). L'identification précise du site de modification par spectrométrie de masse nécessite la présence de sites de coupure protéolytique encadrant le résidu modifié mais ne suffit pas toujours. Dans le cas de PTM multiples, la présence d'un site de coupure protéolytique entre les PTM empêche la détermination de la distribution des PTM (Figure 3.9). La RMN, en revanche, permet l'identification des distributions sans ambiguïté puisqu'une PTM va modifier l'environnement chimique des résidus voisins, qu'ils soient eux-mêmes modifiés ou non. Il faut toutefois souligner que la modification des différents résidus, souvent partielle dans les réactions enzymatiques, complique énormément les spectres lorsque ces résidus sont situés à proximité les uns des autres. Cependant, lorsque les PTM sont distantes, la RMN n'apportera pas d'informations quant à leur distribution.

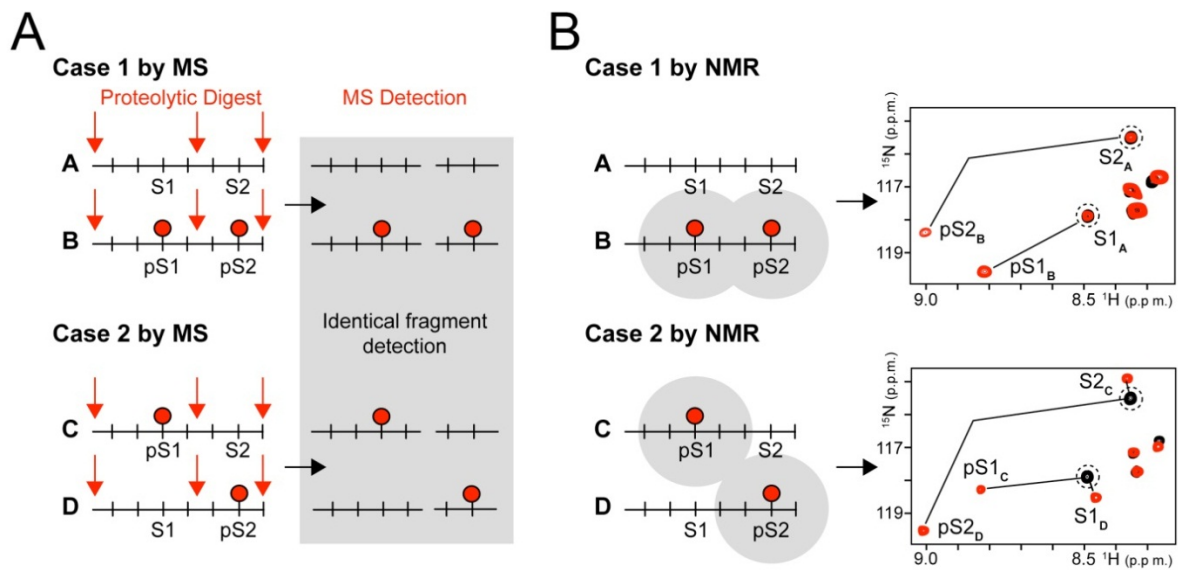


Figure 3.9 : Comparaison de l'analyse de PTM multiples par spectrométrie de masse (A) et RMN (B). Dans le cas de la spectrométrie de masse, les cas 1 et 2 conduisent à la détection d'un même ensemble de fragments et ne permet donc pas de distinguer les deux contextes, alors que ces deux cas de figures induiront sur le spectre <sup>1</sup>H-<sup>15</sup>N HSQC des modifications distinctes des résonances des résidus modifiés et de leurs voisins.

### 3.3.3. PTM et régions désordonnées

Les PTM sont généralement localisées dans des régions désordonnées dans les protéines-substrats, coïncidant avec des régions régulatrices (Iakoucheva, Radivojac et al. 2004). L'accessibilité de ces régions protéiques très exposées au solvant favorise l'interaction dynamique avec les enzymes intervenant dans l'ajout ou la suppression des PTM. Par ailleurs, la RMN est un outil particulièrement bien adapté à l'étude des protéines ou fragments protéiques déstructurés et se prête donc bien à l'étude des PTM dans ces régions désordonnées. L'avantage est que l'impact spectral d'une PTM donnée sera facilement identifiable dans ce contexte, puisque l'environnement chimique de la PTM sera peu influencé par le site de modification. On a montré par exemple, que la phosphorylation induit généralement un *shift* prononcé de la résonance amide du résidu phosphorylé vers une région vide du spectre <sup>1</sup>H-<sup>15</sup>N HSQC, ce qui permet de détecter facilement l'état de phosphorylation d'une protéine. De même, l'acétylation conduit à un signal facilement identifiable correspondant à la nouvelle fonction amide formée sur la chaîne latérale du résidu lysine modifié, qui apparaît également dans une région caractéristique du spectre <sup>1</sup>H-<sup>15</sup>N HSQC, non peuplée dans le cas des protéines non structurées. Le cas des protéines globulaires est complètement différent car les régions spectrales en question ne sont plus vides dans ce contexte, et les *shifts*, rapportés à la dispersion des déplacements chimiques directement due à la structure tridimensionnelle, apparaissent bien moins prononcés

### 3.3.4. PTM et conséquences structurales/fonctionnelles

En plus de l'aspect analytique, la RMN peut apporter des informations structurales et fonctionnelles. La nature non destructive de la RMN permet, en utilisant le même échantillon, à la fois d'identifier une PTM et d'en étudier l'impact sur la structure et la fonction de la protéine modifiée, en particulier

comment celle-ci va moduler l'interaction avec d'autres partenaires. En effet, la plupart des PTM sont "lues" par des modules d'interaction protéine-protéine spécifiques, comme les bromodomains pour les lysines acétylées, les chromodomains pour les lysines méthylées, les domaines 14-3-3 et WW pour les sérines/thréonines phosphorylés, les domaines SH<sub>2</sub> pour les tyrosines phosphorylées... La nature modulaire de nombreuses protéines eucaryotes est en partie liée à l'adaptation des protéines pour la reconnaissance des PTM, en incluant de tels modules pour l'établissement d'interactions protéine-protéine spécifiques régulées par ces PTM. Des modifications structurales résultant directement de l'incorporation spécifique d'une PTM sont également attendue dans le contexte de protéines globulaires. La conséquence est une augmentation de la complexité spectrale avec l'observation de perturbations des résonances à plus longue distance du fait du réarrangement structural, concernant un nombre plus important de résidus de la protéine que ceux immédiatement au voisinage du site de modification dans la séquence primaire.

Dans le chapitre suivant, nous allons proposer un projet d'étude des PTM de la protéine Tau telles que l'acétylation et l'O-GlcNAcylation avec l'identification et la caractérisation de ces PTM par RMN dans le contexte d'une protéine entière. La caractérisation des profils de modifications permettra d'envisager des études structurales et fonctionnelles avec pour objectifs, de clarifier, au niveau moléculaire, les mécanismes de régulation de la fonction physiopathologique de Tau via ces PTM.



## Chapitre 4. Régulation structurale et fonctionnelle de Tau par les modifications post-traductionnelles

---

Deux nouvelles PTM de Tau, l'*O*- $\beta$ -N-acétylglucosaminylation (*O*-GlcNAcylation) et l'acétylation, ont été décrites ces dernières années et ouvrent la voie à de nouvelles approches pour le contrôle (i) de la phosphorylation, (ii) de la fonction de Tau dans la liaison et la polymérisation des microtubules et (iii) de l'agrégation de Tau avec, comme enjeux, non seulement la compréhension des mécanismes moléculaires au niveau physiopathologique, mais également, de nouvelles perspectives thérapeutiques pour traiter la maladie d'Alzheimer. D'après les données actuelles de la littérature, ces deux mécanismes ont *a priori* un effet opposé sur la phosphorylation et la formation des fibres pathologiques. L'*O*-GlcNAcylation, d'une part, décrite par G. Hart en 1996 (Arnold, Johnson et al. 1996), a fait l'objet de nombreuses études cette dernière décennie en raison de son lien compétitif avec la phosphorylation. Récemment, il a été montré que l'*O*-GlcNAcylation de Tau réduit la formation des fibres pathologiques *in vitro* (Yuzwa, Shan et al. 2012). L'acétylation, au contraire, pourrait empêcher la dégradation des formes oligomériques anormalement phosphorylées de Tau par le protéasome et contribuer à l'accumulation de formes hyperphosphorylées pathologiques qui aboutissent à la formation des PHF (Mattson 2010; Min, Cho et al. 2010). De plus, l'acétylation réduit l'interaction de Tau avec les microtubules et stimule l'agrégation de Tau *in vitro* (Cohen, Guo et al. 2011).

Une approche en plein essor pour l'étude des PTM est la voie des protéines de synthèse ou semi-synthétiques. Ces méthodologies sont diverses et l'EPL (*Expressed Protein Ligation*), en particulier, qui permet de combiner les méthodologies d'expression de protéines recombinantes et de synthèse peptidique en phase solide constitue une voie intéressante pour l'analyse par RMN de la protéine ainsi obtenue. Cette approche va être utilisée, dans un premier temps, pour l'étude de phospho-épitopes spécifiques de Tau et pourra être étendue à l'étude d'autres PTM.

### 4.1 ETUDE DE MODIFICATIONS POST-TRADUCTIONNELLES SPECIFIQUES AVEC DES PROTEINES TAU SEMI-SYNTHETIQUES

#### 4.1.1. Principe de la synthèse de protéines Tau semi-synthétiques

Une dérégulation de la phosphorylation peut transformer la protéine Tau en une protéine inactive qui s'accumule dans les neurones sous forme de fibres insolubles hyperphosphorylées. A l'inverse, l'*O*-GlcNAcylation est un mécanisme potentiellement salvateur qui contribue à diminuer le niveau de phosphorylation et prévient l'agrégation de manière indépendante de la phosphorylation. Les profils de phosphorylation sont multiples et complexes par nature étant donné le nombre de kinases et phosphatases impliquées, leur (dé)régulation spatio-temporelle. L'impact de phosphorylations ou d'*O*-GlcNAcylation spécifiques sur la fonction de Tau ou l'agrégation est encore mal connu et demande à être clarifié au niveau moléculaire. Des protéines Tau modifiées spécifiquement seront des modèles simplifiés pour évaluer le rôle de ces PTM dans la régulation de la fonction de Tau et son agrégation. Nous voulons étudier l'effet de phosphorylations ou d'*O*-GlcNAcylation spécifiques sur

la conformation et la fonction de la protéine Tau avec la production d'une protéine semi-synthétique obtenue par ligation chimique native de type EPL (Figure 4.1). Ce projet sera réalisé dans le cadre d'une collaboration avec le Pr Christian Hackenberger (Leibniz Institut fuer Molekulare Pharmakologie, Allemagne), qui possède l'expertise dans le domaine des ligations chimiques (Hackenberger and Schwarzer 2008).

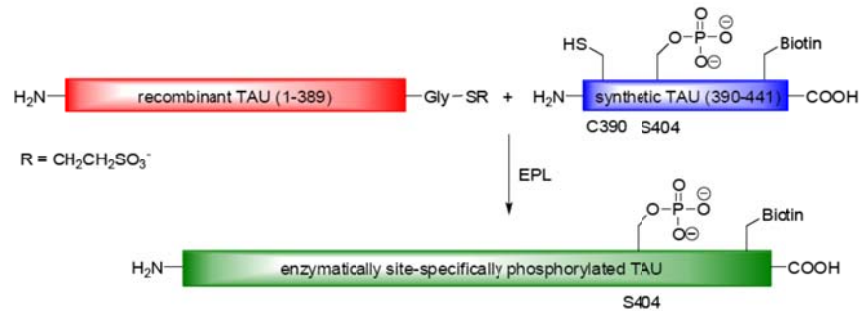


Figure 4.1 : Principe de production de la protéine Tau semi-synthétique (isoforme de 441 acides aminés) sélectivement phosphorylée en S404, par ligation chimique native du segment [1-389] produit par voie recombinante dans *E. coli* avec le segment [390-441] obtenu par synthèse peptidique en phase solide.

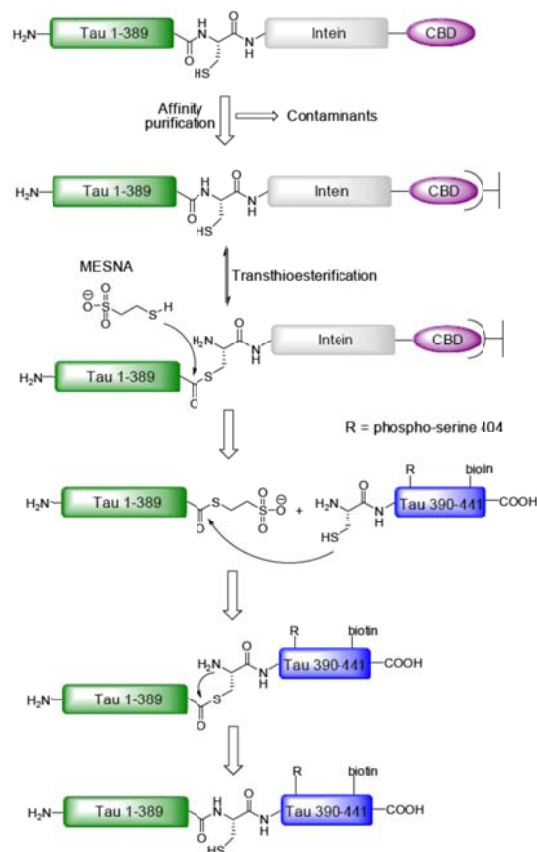


Figure 4.2 : Etapes de l'obtention de la protéine Tau semi-synthétique. Dans un premier temps, la protéine Tau[1-389] fusionnée à l'intéine et au *chitin-binding domain* (CBD) est purifiée par affinité sur colonne de chitine. La liaison peptidique avec l'intéine induit un réarrangement spontané sous forme de N→S shift



conduisant à une fonction thioester. La fonction thioester de l'extrémité carboxy-terminale est obtenue par réaction de transthioestérification avec le 2-mercaptoéthylsulfonate de sodium (MESNa) et conduit à l'élu­tion simultanée de la protéine. La protéine Tau semi-synthétique est ensuite obtenue par condensation de la fonction thioester carboxy-terminale avec la fonction thiol de la cystéine amino-terminale du peptide Tau[390-441] synthétisé en phase solide permettant l'introduction sélective et quantitative de PTM ou autres marqueurs. Le réarrangement S→N *shift* permet de restaurer une liaison peptidique native, en laissant néanmoins une mutation A390C.

Dans un premier temps, nous allons étudier l'impact de l'épito­pe PHF-1 spécifique de la MA (Otvos, Feiner et al. 1994; Augustinack, Schneider et al. 2002), localisé dans la région C-terminale de Tau, phosphorylé sur les sites S396/S400/S404, sur la conformation et la fonction de la protéine Tau (Wang, Udeshi et al. 2010; Smet-Nocca, Broncel et al. 2011; Yuzwa, Yadav et al. 2011). Le fragment N-terminal (résidus 1 à 389) sera produit par voie recombinante chez *E. coli* comme protéine de fusion à l'intéine, permettant le marquage en <sup>15</sup>N et/ou <sup>13</sup>C pour les études par RMN (Figure 4.2). Le fragment C-terminal triplement phosphorylé ou O-GlcNAcylé sera produit par synthèse peptidique en phase solide pour permettre l'introduction sélective et quantitative de PTM ou de marqueurs (biotine, fluorophores) (Broncel, Krause et al. 2012) (Figure 4.2). Des protéines Tau semi-synthétiques fonctionnelles ont déjà été obtenues selon ce schéma, dans des formes non phosphorylée ou phosphorylée en S404 (Broncel, Krause et al. 2012). L'obtention du peptide de 52 acides aminés triplement phosphorylé est plus problématique du fait de la proximité des trois phosphorylations et a nécessité la mise au point d'une stratégie de synthèse optimisée (Figure 4.3).

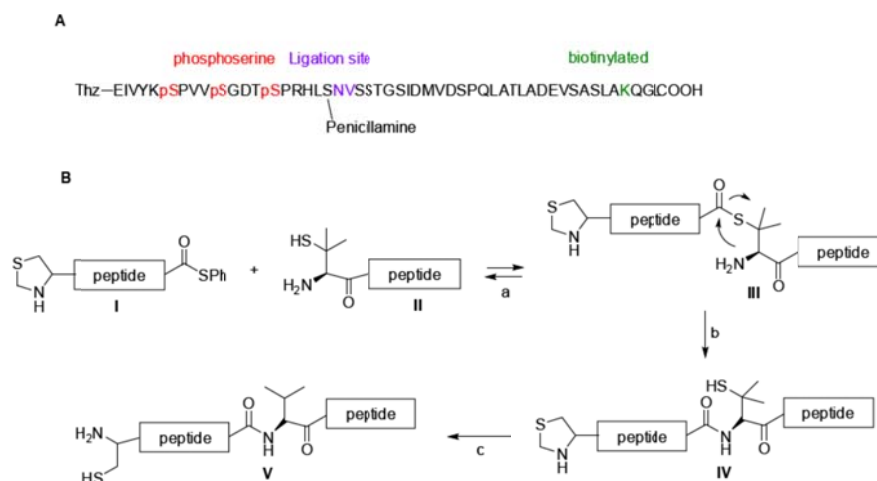


Figure 4.3 : (A) Séquence du peptide C-terminal de Tau de 52 acides aminés comportant trois sérines phosphorylées aux positions 396, 400 et 404 (en rouge). Le site de biotinylation (K438) permettant la purification du produit de ligation chimique est noté en vert. Le site de ligation permettant la liaison du peptide N-terminal de 21 résidus au peptide C-terminal de 31 résidus est noté en violet. (B) Stratégie de synthèse pour l'obtention du peptide de 52 résidus par ligation chimique native du peptide N-terminal de 21 résidus (résidus 390 à 410 ; peptide I), portant les trois sites de phosphorylation, au peptide C-terminal de 31 résidus (résidus 411 à 441 ; peptide II). La jonction se fait au niveau d'un groupement pénicillamine (β-mercaptovaline) du peptide II sur la fonction thioester du peptide II qui, après réaction de désulfuration, restaure une valine native en position 411. Au cours de cette étape, le groupement thiazolidine en position N-terminale du peptide de 52 acides aminés est clivé pour donner le résidu cystéine qui réagira avec le groupement thioester de la fonction carboxy-terminale de la protéine Tau[1-389] produite par voie recombinante, lors de l'étape EPL.

Des premières données obtenues par comparaison des spectres de la protéine Tau tronquée (Tau[1-389]) et de la protéine entière indiquent que la délétion du domaine C-terminal induit des modifications importantes de l'intensité relative des signaux qui affecte plus particulièrement la région riche en prolines ainsi que le segment répété R1 et, dans une moindre mesure, le R2 (Figure 4.4). Pour ces régions, on observe un élargissement des signaux qui s'accompagne, pour certains résidus, d'une variation du déplacement chimique. Au contraire, le R4 et la région voisine (résidus 340 à 389) montrent une nette augmentation de l'intensité des résonances par rapport à l'isoforme entière. Ces données indiquent globalement une modification de la dynamique et/ou des échanges avec l'eau lors de la délétion du domaine C-terminal, qui s'étend sur l'ensemble de la région riche en prolines et du domaine de liaison aux microtubules (le segment répété R3 est moins affecté). Ces observations concordent avec la conformation globale *paperclip-like* qui a été décrite par Mandelkow et collaborateurs (Jeganathan, von Bergen et al. 2006), qui montre notamment un repliement transitoire du domaine C-terminal sur le domaine à segments répétés. En revanche, contrairement à cette conformation *paperclip-like* où il existe une proximité spatiale transitoire entre les domaines N- et C-terminaux, l'absence du domaine C-terminal ne semble pas affecter ici la région N-terminale de manière significative.

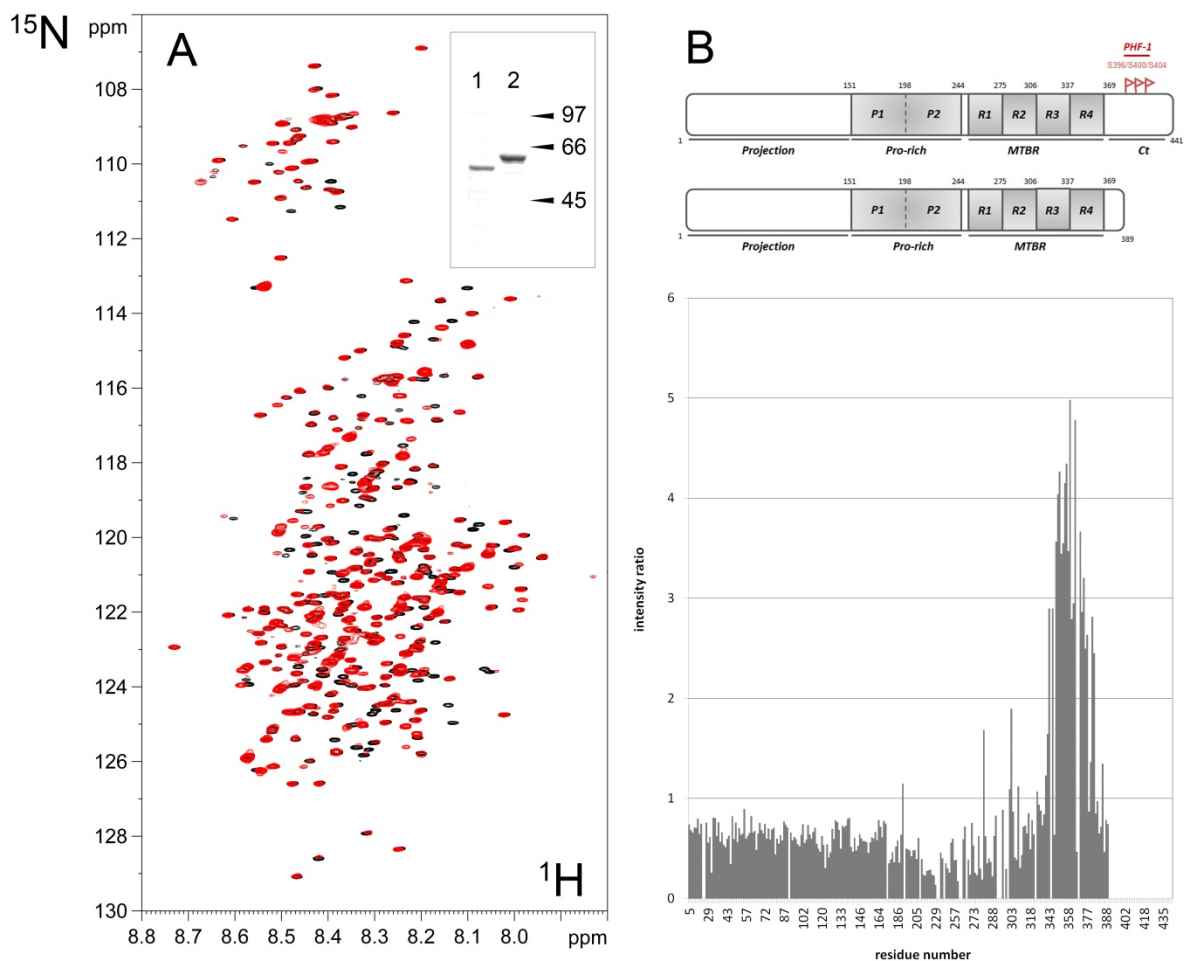


Figure 4.4 : Effet de la délétion du domaine C-terminal de Tau sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC. (A) Superposition des spectres de Tau441 (noir) et Tau[1-389] (rouge) enrichies en  $^{15}\text{N}$ . En encadré, analyse sur gel dénaturant de

polyacrylamide des protéines Tau441 (bande 2) et Tau[1-389] (bande 1). Les marqueurs de poids moléculaires sont indiqués à droite en kDa. (B) Cartographie des intensités relatives des signaux entre les deux protéines. Les intensités ont été normalisées sur l'enveloppe des signaux des spectres  $^1\text{H}$  correspondant. Au-dessus, représentation schématique des domaines de Tau441 (l'épitope PHF-1 est indiqué en rouge) et Tau[1-389]. Projection : domaine de projection ; Pro-rich : domaine riche en prolines ; MTBR : *microtubule-binding repeats* ; Ct : domaine C-terminal.

Il a été montré que la phosphorylation module la conformation globale de Tau selon le nombre et la localisation des sites phosphorylés, et pourrait rompre des interactions intramoléculaires transitoires à longue distance (Jeganathan, Hascher et al. 2008; Bibow, Ozenne et al. 2011). L'étude structurale d'une protéine Tau semi-synthétique phosphorylée de manière homogène permettra d'évaluer l'impact de phosphorylations spécifiques sur la conformation de Tau.

#### **4.1.2. Etudes fonctionnelles de phospho-épitoques spécifiques par l'intermédiaire de protéines Tau semi-synthétiques**

L'étude de l'impact de la phosphorylation des sites S396/S400/S404 ou de l'O-GlcNAcylation en S400 sur la conformation de Tau seule ou liée à la tubuline pourra être étudiée par RMN. Cette méthode permet l'étude, au niveau moléculaire, de protéines non structurées, là où la cristallographie par rayons X est limitée par l'obtention de cristaux pour ces protéines particulièrement flexibles et conformationnellement hétérogènes (Lippens, Wieruszkeski et al. 2004; Smet, Leroy et al. 2004; Lippens, Sillen et al. 2006). Ainsi, l'étude de la conformation des protéines non structurées est possible avec, par exemple, la recherche de structures locales transitoires (appelées également structure résiduelles) coïncidant généralement avec des régions fonctionnelles. On pourra également définir, de manière fine, les régions impliquées dans l'agrégation ou dans la liaison aux microtubules (Sillen, Leroy et al. 2005; Sillen, Wieruszkeski et al. 2005; Sibille, Sillen et al. 2006; Sillen, Barbier et al. 2007). De cette façon, la conformation et la fonction de Tau modifiée par des phosphorylations ou glycosylations spécifiques pourront être étudiées *in vitro*. La liaison des microtubules sera envisagée par RMN comme pour la protéine non modifiée et un test de cinétique de polymérisation de la tubuline permettra de mettre en évidence d'éventuelles différences d'activité sous l'effet de modifications spécifiques (Sillen, Barbier et al. 2007; Amniai, Barbier et al. 2009). De plus, la RMN offre la possibilité de caractériser les PTM, en complément de la spectrométrie de masse, dans le cas de PTM multiples et localisées sur des sites proches (Landrieu, Lacosse et al. 2006; Amniai, Barbier et al. 2009; Landrieu, Leroy et al. 2010; Landrieu, Smet-Nocca et al. 2011; Prabakaran, Everley et al. 2011; Theillet, Smet-Nocca et al. 2012). On pourra ainsi caractériser l'effet de l'O-GlcNAcylation en S400 sur les profils de phosphorylation par différentes kinases MA-spécifiques. Dans un deuxième temps, des PTM localisées dans la région riche en prolines, qui intervient dans la régulation de la liaison aux microtubules, pourront être envisagées mais requièrent l'élaboration de stratégies de ligation plus complexes, faisant intervenir plusieurs fragments polypeptidiques.

Il a été proposé que Tau interagisse avec les groupements polyglutamates de la tubuline via sa région basique. Ces chaînes polyglutamates, qui peuvent contenir jusqu'à six résidus Glu, proches du C-terminus de la plupart des isoformes de tubuline  $\alpha$  et  $\beta$ , ne seraient pas impliquées directement dans la liaison de Tau mais réguleraient sa liaison via un changement conformationnel en (dé)masquant le site de liaison de Tau à la tubuline (Boucher, Larcher et al. 1994). La phosphorylation de Tau, en

augmentant l'acidité de la protéine, pourrait modifier la liaison de Tau à la tubuline via des interactions électrostatiques défavorables avec les queues polyglutamates. Par ailleurs, un repliement transitoire global de la protéine Tau soluble a été mis en évidence montrant une proximité spatiale entre les extrémités N- et C-terminales, et les *repeats* R2/R3, dite conformation *paperclip* (Jeganathan, von Bergen et al. 2006). Cette conformation coïncide avec la capacité d'anticorps dits conformationnellement-dépendants, Alz50 et MC-1, de reconnaître une conformation anormale de Tau au travers d'un épitope discontinu, dans les stades précoces de la MA. La pseudo-phosphorylation des épitopes AT8, AT100 et PHF-1 ensemble augmente la compaction de la conformation *paperclip* et favoriserait l'agrégation sous formes de PHF, tandis que les pseudo-phosphorylations de AT8 ou PHF-1 seules ont l'effet inverse. La réactivité vis-à-vis de MC-1 est augmentée en présence des 3 phospho-épitopes suggérant qu'il pourrait exister une conformation pathologique de Tau dans la MA (Jeganathan, Hascher et al. 2008). Ainsi, la phosphorylation de l'épitope PHF-1 dans la protéine Tau semi-synthétique permettra d'examiner en détails l'impact structural de cette phosphorylation spécifique sur la conformation de Tau libre ou liée à la tubuline. En effet, la formation d'un complexe Tau : tubuline est susceptible d'affecter la conformation de Tau et stabiliser une conformation préexistante qui serait transitoire dans l'état libre. Dans le cadre d'une collaboration avec le Dr Knossow (Centre de Recherche CNRS de Gif-sur-Yvette, Paris), l'étude de la liaison de Tau à deux hétérodimères de tubuline stabilisés par le domaine *stathmin-like* de la protéine RB3 (T2R ; PDB ID : 3HKB) a permis d'identifier précisément les régions de Tau impliquées dans la liaison, et a conduit au clonage du fragment Tau-F4 [208-324] (Fauquant, Redeker et al. 2011). D'autres études réalisées dans notre laboratoire ont montré que Tau phosphorylée au niveau des épitopes pathologiques, AT8 et AT180, ne peut plus promouvoir l'assemblage des microtubules mais conserve sa capacité à lier les microtubules, concluant ainsi à une liaison non productive de Tau phosphorylée aux microtubules (Amniai, Barbier et al. 2009). Le complexe T2R pourra être utilisé pour identifier l'effet d'un phospho-épitope unique sur la structure de Tau en complexe avec la tubuline. Des données obtenues dans notre laboratoire (Guy Lippens, données non publiées) ont permis de cartographier l'interaction de Tau avec T2R qui pourrait expliquer son rôle moléculaire dans la polymérisation de la tubuline. En effet, Tau liée à T2R conserve une dynamique dans certaines régions qui serait nécessaire pour promouvoir l'assemblage des microtubules, en liant un ou deux dimères de tubuline. Dans ce contexte, la phosphorylation de l'épitope PHF-1 dans la région C-terminale pourrait potentiellement stabiliser la conformation de Tau, lui permettant toujours de se lier à T2R mais diminuerait la dynamique requise pour favoriser la polymérisation.

La présence de nombreux sites de phosphorylation dans la protéine Tau phosphorylée par le complexe CDK2/cycline A3 complique l'étude des interactions avec la peptidyl-prolyl *cis/trans* isomérase Pin1 en raison d'une interaction multi-site, sur plusieurs sites de faible affinité. Un élargissement des résonances des résidus phosphorylés en présence de Pin1 est observé pour des quantités de Tau sub-stœchiométriques, et une précipitation des protéines est observée en présence d'un excès de Tau. La présence d'un seul site de phosphorylation dans la protéine Tau semi-synthétique permettra de contourner les problèmes liés aux phosphorylations multiples. Le site PHF-1 comporte deux motifs phospho-Ser-Pro, qui présentent individuellement, dans le contexte de petits peptides, des  $K_D$  de l'ordre du millimolaire (Smet, Sambo et al. 2004). D'autre part, il a été proposé que Pin1 induise, par son activité PPIase, un changement conformationnel au sein de ses

substrats qui régulerait ainsi leur fonction mais qui n'a encore jamais été montré. Dans le contexte neuronal, il a été observé que Pin1 se lie à la protéine Tau phosphorylée et restaure sa capacité à lier les microtubules et à favoriser leur polymérisation (Lu, Wulf et al. 1999). L'interaction Pin1 :phospho-Tau pourra être mise dans le contexte de la liaison de Tau à la tubuline avec le modèle T2R et les protéines Tau semi-synthétiques, ce qui n'était pas possible avec les modèles peptidiques décrits dans le Chapitre 2. Il faut souligner cependant, que nous ne pourrions observer l'impact conformationnel d'une interaction avec Pin1 via l'épitope PHF-1 qu'à longue distance, du fait de l'absence de marquage isotopique dans la région C-terminale comprenant le motif d'interaction. Un tel effet structural à longue distance reste encore très hypothétique, et probablement très difficile à observer. Nous pourrions également déterminer si, via l'épitope PHF-1, cette liaison à Pin1 a un effet sur la liaison de Tau phosphorylée à T2R et sur la polymérisation des microtubules.

## **4.2 O-GLCNAcylation : UN MECANISME ANTAGONISTE DE LA PHOSPHORYLATION DE TAU ET PROTECTEUR DE L'AGREGATION ?**

### **4.2.1. Introduction**

De nombreuses études ont montré que la phosphorylation anormale de Tau joue un rôle crucial dans le processus de neurodégénérescence de la MA et d'autres tauopathies (Iqbal, Alonso Adel et al. 2005). En conséquence, l'inhibition ou le recul de l'hyperphosphorylation de Tau sont des voies thérapeutiques d'intérêt majeur dans la MA. En plus d'être régulée par les kinases et phosphatases, le niveau de phosphorylation de Tau est régulé par O-GlcNAcylation. L'O-GlcNAcylation de Tau dans le cerveau humain a été découverte relativement récemment par rapport à la phosphorylation (Arnold, Johnson et al. 1996; Liu, Iqbal et al. 2004).

La glycosylation par O-GlcNAc est une modification post-traductionnelle réversible résultant de l'addition, à partir de l'UDP-GlcNAc, d'un groupement N-acétylglucosamine sur les chaînes latérales des résidus sérine et thréonine. 2-5% du glucose intracellulaire est utilisé pour synthétiser l'UDP-GlcNAc, un sucre activé, via la voie de signalisation des hexosamines (HBP pour *Hexosamine Signalling Pathway*) (Love and Hanover 2005). En plus du glucose interviennent d'autres métabolites comme la glutamine dont le groupement amide se combine à l'acétyl-coenzyme A (AcCoA) pour former le groupement N-acétyl de l'UDP-GlcNAc. L'O-GlcNAcylation est par conséquent une modification sensible au stress et aux nutriments, qui régule un large panel de fonctions biologiques. Cette modification post-traductionnelle est dynamiquement régulée par l'action antagoniste de deux enzymes, l'O-GlcNAc transférase (OGT) et O-GlcNAc hydrolase (OGA), chacune codée par un gène unique chez l'homme (Iyer and Hart 2003). Dans des conditions normales, la modification par O-GlcNAc concerne plus d'un millier de protéines aux niveaux nucléaires, mitochondriales et cytoplasmiques. L'O-GlcNAcylation a été impliquée dans la transduction des signaux, les interactions protéine-protéine, la transcription, le cycle cellulaire et la réponse au stress cellulaire (Kamemura and Hart 2003; Zachara and Hart 2004; Slawson, Zachara et al. 2005; Zachara and Hart 2006; Zachara, Molina et al. 2011) mais l'implication fonctionnelle de la modification des protéines par O-GlcNAc reste encore largement mal définie. La régulation de l'O-GlcNAcylation est non seulement conditionnée par la régulation des deux enzymes impliquées dans le cyclage du GlcNAc mais elle est

également sensible à la concentration en glucose qui se trouve à la base de la voie de signalisation des hexosamines (Hart, Housley et al. 2007). La dérégulation de la modification par *O*-GlcNAc a été impliquée dans des pathologies comme les diabètes, les cancers et la maladie d'Alzheimer (Akimoto, Hart et al. 2005; Lefebvre, Guinez et al. 2005; Dias and Hart 2007; Butkinaree, Park et al. 2009; Slawson, Copeland et al. 2010).

L'*O*-GlcNAcylation est un mécanisme analogue à la phosphorylation des protéines et la plupart des protéines *O*-GlcNAcylées sont aussi des phospho-protéines (Haltiwanger, Busby et al. 1997). Souvent, ces deux PTM sont décrites comme mutuellement exclusives. Elles peuvent entrer en compétition pour l'occupation d'un même site ou inhiber l'occupation d'un site adjacent (Hart, Greis et al. 1995; Cheng, Cole et al. 2000; Comer and Hart 2000; Cheng and Hart 2001; Wells, Vosseller et al. 2001; Zachara and Hart 2002; Hu, Shimoji et al. 2010). De plus, il a été montré récemment que l'OGT peut former des complexes transitoires avec la sous-unité catalytique de la phosphatase PP1 (PP1c) suggérant que, dans certains cas, de tels complexes enzymatiques puissent être recrutés à la fois pour déphosphoryler et *O*-GlcNAcyler les protéines (Wells, Kreppel et al. 2004). Cependant, des données récentes indiquent que les relations phosphorylation/*O*-GlcNAcylation observées sur une population globale peuvent être très différentes, voire à l'opposé, de ce qui est observé à l'échelle d'une sous-population spécifique, soulignant l'intérêt de développer des méthodologies quantitatives pour étudier la balance entre deux modifications. Il a en effet été montré que la glycosylation par *O*-GlcNAc du répresseur transcriptionnel MeCP2 se produisait uniquement sur la fraction phosphorylée de la protéine, un mécanisme à l'opposé du mécanisme *yin-yang* généralement évoqué où les deux PTM sont mutuellement exclusives (Kohler 2010; Rexach, Rogers et al. 2010).

### ***Balance phosphorylation/O-GlcNAcylation dans la protéine Tau***

On sait de longue date que le métabolisme du glucose est perturbé dans le cerveau de patients atteints de la MA (Meier-Ruge and Bertoni-Freddari 1996) et une diminution du niveau de glycosylation par *O*-GlcNAc a été observé (Liu, Shi et al. 2009). Une corrélation inverse a été observée entre phosphorylation et *O*-GlcNAcylation dans le cas de la protéine Tau. L'*O*-GlcNAcylation régule la phosphorylation de Tau, et *vice versa*, à la fois dans des cellules en culture et dans des coupes de cerveaux de rats métaboliquement actives. De plus, une diminution de l'*O*-GlcNAcylation des protéines, et en particulier de Tau, a été observée dans les cerveaux atteints de la MA.

Différentes approches visant à augmenter le niveau d'*O*-GlcNAcylation, par utilisation d'inhibiteurs de l'OGA (Fischer 2008; Yuzwa, Macauley et al. 2008), ou à le diminuer, par privation en glucose (Liu, Iqbal et al. 2004) ou réduction de l'expression de l'OGT dans des cellules HEK-293 (Liu, Shi et al. 2009), ou encore visant à modifier le niveau de phosphorylation par l'action d'un inhibiteur de la kinase GSK3 $\beta$  (Robertson, Moya et al. 2004) ou de la phosphatase PP2A (Lefebvre, Ferreira et al. 2003), ont permis de mettre en évidence un lien antagoniste entre phosphorylation et *O*-GlcNAcylation, dans la plupart des cas étudiés. Par exemple, il a été montré qu'une augmentation de l'*O*-GlcNAcylation diminue la phosphorylation de Tau sur des sites spécifiques, dans un certain nombre de modèles cellulaires et animaux (Arnold, Johnson et al. 1996; Liu, Iqbal et al. 2004; Dehennaut, Lefebvre et al. 2007; Yuzwa, Macauley et al. 2008). Par conséquent, une diminution de la phosphorylation par augmentation de l'*O*-GlcNAcylation pourrait avoir un effet bénéfique sur

l'agrégation. De plus, il a été montré récemment dans un modèle de souris transgéniques P301L, capables de former des NFT et qui expriment un mutant de Tau ayant une cinétique d'agrégation plus rapide que la protéine sauvage *in vitro*, que l'*O*-GlcNAcylation pourrait être un mécanisme protecteur de l'agrégation en soi, en diminuant la formation des fibres de Tau *in vitro*, de manière indépendante de la phosphorylation (Yuzwa, Shan et al. 2012). Sur bases de ces observations, il a été proposé qu'une dérégulation du métabolisme du glucose dans les cerveaux atteints de la MA pouvait être à l'origine de la dégénérescence neurofibrillaire via une diminution du niveau d'*O*-GlcNAcylation de Tau menant à son hyperphosphorylation (Gong, Liu et al. 2006).

Dans ce contexte, une élévation du niveau d'*O*-GlcNAcylation de Tau pourrait être une nouvelle voie prometteuse dans le traitement de la MA ou d'autres tauopathies, en inhibant ou en renversant l'hyperphosphorylation de Tau (Yuzwa and Vocadlo 2009). La conception d'inhibiteurs de l'*O*-GlcNAcase basée sur la structure de l'enzyme en complexe avec un substrat *O*-GlcNAcylé, a abouti à des molécules mimant l'état de transition du groupement *O*-GlcNAc dans la poche spécifique de l'OGA humaine (Macauley and Vocadlo 2009). Dans cette série, le Thiamet-G présente une activité élevée (avec  $K_i$  de 27 nM), une sélectivité élevée pour l'OGA (de 37 000 fois supérieure par rapport aux  $\beta$ -hexosaminidases lysosomales humaines et n'inhibe d'autres glycosides hydrolases qu'à des concentrations supérieures à 500  $\mu$ M) ainsi qu'une très faible toxicité. Cette molécule est également capable de passer la barrière hémato-encéphalique ce qui en fait un chef de file intéressant pour le développement de candidats médicaments dans le traitement de la MA. Le traitement de cellules PC12 avec le Thiamet-G induit une augmentation de l'*O*-GlcNAcylation des protéines ainsi qu'une diminution de la phosphorylation des sites T231, S396 et S422. Des résultats similaires ont été observés chez des rats traités par voies orales ou intraveineuses (Yuzwa, Macauley et al. 2008). Il faut noter qu'une augmentation du niveau d'*O*-GlcNAcylation d'autres protéines par cette voie pourrait avoir d'autres conséquences, qui restent à définir. Notamment, comme de nombreuses protéines sont modifiées par *O*-GlcNAc, une toxicité potentielle liée à la perturbation des profils de glycosylation d'autres protéines est à prévoir. De plus, il a été observé une augmentation du niveau de phosphorylation de certains sites au sein de la protéine Tau, spécifiques de la MA. Cette élévation de la phosphorylation pourrait être liée à l'activation de GSK3 $\beta$ , une des kinases les plus importantes dans le cerveau. Il a en effet été montré que l'AKT et d'autres protéines de la voie de signalisation de l'AKT sont modifiées par *O*-GlcNAcylation et que l'*O*-GlcNAcylation de l'AKT inhibe son activité. L'AKT est une kinase impliquée dans la phosphorylation de GSK3 $\beta$  au niveau de la S9, résultant en une inhibition de son activité kinase. Une élévation du niveau d'*O*-GlcNAcylation de l'AKT pourrait en conséquence conduire à l'activation de GSK3 $\beta$ .

### ***L'OGT, une enzyme essentielle et unique***

Le clonage du cDNA de l'OGT (UDP-GlcNAc:polypeptidyltransferase) du rat et de l'homme date de 1997 (Kreppel, Blomberg et al. 1997; Lubas, Frank et al. 1997), et de l'OGA issue de cerveaux humains de 2001 (Gao, Wells et al. 2001). Chez l'homme, le gène de l'OGT se trouve sur le chromosome X sous une seule copie (Shafi, Iyer et al. 2000). Face aux 400 Ser/Thr kinases, un seul domaine OGT a été identifié à l'heure actuelle et on dénombre plus de 600 protéines modifiées par *O*-GlcNAc. Ces données suggèrent que, contrairement aux kinases par exemple dont la sélectivité est en partie codée dans leur gène, le mode de reconnaissance des substrats par l'OGT reste encore mal compris à

l'heure actuelle. L'OGT est une protéine ubiquitaire, abondante plus particulièrement dans le pancréas et le cerveau. Elle est majoritairement nucléaire mais un variant d'épissage est également trouvé au niveau mitochondrial. L'activité de l'OGT est sensible au niveau d'UDP-GlcNAc sur une large gamme de concentrations qui affectent le niveau d'O-GlcNAcylation des protéines (Kreppel and Hart 1999).

L'OGT est exprimée sous trois isoformes : l'OGT nucléocytoplasmique (ncOGT) de 116 kDa, l'isoforme courte (sOGT) de 70 kDa - ces deux isoformes sont localisées dans le noyau- et l'OGT mitochondriale (mOGT) de 103 kDa. Les différentes isoformes diffèrent essentiellement par leur nombre de TPR (*tetrapetptide repeats*). Le domaine TPR est localisé dans la partie N-terminale de l'OGT. Il s'agit d'une séquence répétée de peptides de 34 acides aminés. Les TPR sont nécessaires à la spécificité de l'OGT pour des peptides substrats *in vitro* et interviennent dans l'oligomérisation de l'OGT (Kreppel and Hart 1999; Iyer and Hart 2003). La ncOGT humaine contient 11.5 TPR qui forment une super-hélice relativement flexible. La surface interne est riche en résidus asparagines absolument conservés tandis que la surface externe intervient dans la dimérisation (Jinek, Rehwinkel et al. 2004). Cette structure présentant une surface en super-hélice exposant un sillon de résidus asparagine conservés est similaire à celle de la  $\beta$ -caténine (Huber and Weis 2001) ou de l'importine  $\alpha$  qui lie les séquences peptidiques de type signal de localisation nucléaire (NLS pour *nuclear localization signals*) (Conti and Kuriyan 2000; Fontes, Teh et al. 2000). Dans le cas de l'importine  $\alpha$ , il a été proposé que différents sous-groupes d'asparagines interviennent pour reconnaître différents peptides NLS (Fontes, Teh et al. 2000). Le mécanisme de reconnaissance des protéines par l'OGT pourrait être similaire, ce qui expliquerait l'absence de séquence consensus des substrats. De plus, cette organisation tridimensionnelle favorise plutôt la liaison des régions protéiques non structurées. Les TPR joueraient par conséquent le rôle de domaine "plate-forme", médiateur d'un grand nombre d'interactions protéine-protéine.

A ce jour, il n'y a pas encore de structure tridimensionnelle de l'OGT entière. Seules les structures d'une protéine homologue intacte (XcOGT de *Xanthomonas campestris*) (Martinez-Fleites, Macauley et al. 2008) ou de la protéine humaine contenant 4.5 TPR et le domaine catalytique existent (Lazarus, Nam et al. 2011). La structure des TPR en entier a été résolue mais sans le domaine catalytique (Jinek, Rehwinkel et al. 2004). L'ensemble ne permet de proposer aujourd'hui qu'un modèle de la structure de l'OGT humaine (voir Figure 4.6 A). La structure de l'OGT en complexe avec un peptide de CKII et l'UDP ou avec l'UDP seul a permis dans un premier de disséquer le mécanisme enzymatique, notamment de mettre en évidence les résidus du site actif impliqués dans la liaison du peptide et de l'UDP (Lazarus, Nam et al. 2011). Le mécanisme enzymatique de l'OGT implique une inversion de la configuration du carbone anomérique lors du transfert du groupement GlcNAc du donneur, l'UDP-GlcNAc, vers le résidu Ser ou Thr accepteur. Deux publications indépendantes ont ensuite donné un aperçu du mécanisme enzymatique avec les structures cristallographiques détaillant les coordonnées de la réaction (Lazarus, Jiang et al. 2012; Schimpl, Zheng et al. 2012). L'OGT fonctionne selon un mécanisme cinétique bi-bi ordonné : le peptide substrat vient se fixer sur l'OGT après liaison de l'UDP-GlcNAc dans le site actif et ensuite, le glyco-peptide quitte la poche catalytique suivi de l'UDP hydrolysé. Pour l'obtention des structures du complexe ternaire impliquant l'OGT, le peptide substrat et l'UDP-GlcNAc, les auteurs ont utilisé un analogue de l'UDP-GlcNAc, l'UDP-5SGlcNAc, un inhibiteur de l'OGT qui possède une activité 3000 fois plus lente comme donneur, permettant de ralentir la



réaction durant le processus de cristallisation. Mais il faut l'utilisation supplémentaire d'un substrat incompétent pour la réaction de glycosylation, qui diffère entre les deux équipes : il s'agit du peptide CKII où la sérine glycosylée a été mutée en alanine (Lazarus, Jiang et al. 2012) ou du peptide TAB1 modifié où le résidu sérine accepteur est muté en aminoalanine (Schimpl, Zheng et al. 2012). Les mécanismes proposés par les deux groupes sont similaires. Seul diffère le résidu basique impliqué dans la déprotonation de l'hydroxyle accepteur, qui reste encore à déterminer. Il a été proposé que la catalyse basique ne soit pas essentielle dans le mécanisme enzymatique comme cela a déjà été observé pour des glycosidases fonctionnant par inversion de la stéréochimie du carbone anomérique (Withers and Davies 2012).

### **Identification des sites O-GlcNAc dans les protéines**

Des anticorps ont été développés pour détecter des modifications par O-GlcNAcylation dans les protéines, comme RL-2 et CTD110.6, mais leur spécificité a été remise en question. Des réactivités croisées ont été montrées avec d'autres types de structures glycosidiques comme les N-glycannes, par exemple. La spectrométrie de masse a permis de détecter des protéines O-GlcNAcylées dans des cellules en culture ou dans des tissus mais a nécessité des stratégies particulières car les modifications par phosphorylation et O-GlcNAcylation sont labiles dans les conditions d'analyse classique par MS/MS ce qui pose des problèmes pour leur identification. Ainsi, des méthodes sophistiquées combinant modifications chimiques/enzymatiques et spectrométrie de masse ont permis d'identifier les sites O-GlcNAc dans les protéines (Reason, Blench et al. 1991; Zachara, Cole et al. 2001; Wells, Vosseller et al. 2002; Zachara, Hart et al. 2002; Vocadlo, Hang et al. 2003; Whelan and Hart 2003; Zachara, Cheung et al. 2004; Whelan and Hart 2006; Wang, Udeshi et al. 2010; Zachara, Vosseller et al. 2011; Alfaro, Gong et al. 2012; Banerjee, Hart et al. 2012). Ces méthodes associent l'enrichissement des échantillons en peptides O-GlcNAcylés, par des marquages spécifiques avec des enzymes modifiées et purification, à l'utilisation de techniques de fragmentation douce pour la spectrométrie de masse en tandem, spécifiques des liaisons peptidiques comme la dissociation par transfert d'électrons (ETD pour *Electron Transfer Dissociation*) ou la dissociation par capture d'électrons (ECD pour *Electron Capture Dissociation*) (Rexach, Clark et al. 2008; Wang, Udeshi et al. 2010).

La spectrométrie de masse est la technique de choix pour détecter, identifier et localiser des PTM. Plus récemment, des méthodes ont été développées pour en faire une technique structurale permettant d'obtenir des informations sur la structure et la fonction des protéines ou des complexes protéiques. La RMN, à l'inverse, est une technique qui se veut structurale et qui évolue sur le plan analytique pour la caractérisation des PTM, avec la possibilité de caractériser ces PTM sur le plan structural et fonctionnel. Ainsi, il apparaît généralement que ces deux techniques sont complémentaires dans l'étude des PTM. Par exemple, l'analyse par spectrométrie de masse MALDI-TOF du facteur de transcription ER $\beta$  a montré que le résidu sérine 16 est occupé alternativement par phosphorylation et O-GlcNAcylation, ces deux PTM régulant la stabilité et l'activité transcriptionnelle d'ER $\beta$  (Cheng, Cole et al. 2000; Cheng and Hart 2000; Cheng and Hart 2001). Il a été proposé qu'un échantillonnage conformationnel différent du N-terminus non structuré d'ER $\beta$  puisse être à l'origine de la modulation transactivationnelle. L'étude structurale par RMN du peptide d'ER $\beta$  centré autour de ce résidu, dans sa forme non modifiée, phosphorylée ou O-GlcNAcylée a montré une différence

conformationnelle imposée par ces deux PTM, avec une conformation plus étendue sous l'effet de la phosphorylation tandis que l'*O*-GlcNAcylation instaure la formation d'un coude  $\beta$  autour du site de modification (Chen, Du et al. 2006; Hart and Sakabe 2006).

#### **4.2.2. *O*-GlcNAcylation de Tau : étude par RMN**

##### ***Etude des mécanismes d'*O*-GlcNAcylation de Tau in vitro et identification des sites *O*-GlcNAc dans des modèles peptidiques***

Malgré ces nouvelles approches permettant l'identification de sites d'*O*-GlcNAcylation par spectrométrie de masse en tandem (Vocadlo, Hang et al. 2003), l'identification de tels sites dans la protéine Tau reste difficile (Yuzwa, Yadav et al. 2011). Un anticorps dirigé contre le résidu S400 de Tau dans sa forme *O*-GlcNAcylée a été produit (Yuzwa and Vocadlo 2009). La détermination de l'impact de l'*O*-GlcNAcylation sur la phosphorylation de Tau a été réalisée jusqu'à maintenant de manière indirecte, par immunodétection avec des anticorps spécifiques des sites de phosphorylation. La RMN devrait permettre de caractériser, de manière directe, les sites d'*O*-GlcNAcylation, en terme de définition des sites *O*-GlcNAc et quantification relative des niveaux d'*O*-GlcNAcylation. Elle pourra également mettre en évidence l'effet de l'*O*-GlcNAcylation sur des sites de phosphorylation spécifiques, et réciproquement. A cet effet, nous disposons de différentes kinases qui peuvent être produites par voie recombinante (CDK2/cyclineA3, PKA, GSK3 $\beta$ ). La phosphorylation de Tau réalisée *in vitro* avec ces kinases a déjà été caractérisée par RMN (Landrieu, Lacosse et al. 2006; Amniai, Barbier et al. 2009; Landrieu, Leroy et al. 2010; Leroy, Landrieu et al. 2010). L'étude des phosphorylations de Tau par les kinases Erk activées ou des extraits de cerveaux de rats est actuellement en cours dans notre laboratoire. D'autre part, la production d'OGT recombinante dans *E. coli* n'est pas triviale (Lubas and Hanover 2000). L'utilisation d'un plasmide contenant une séquence à codons optimisés pour l'expression dans *E. coli*, obtenu du laboratoire du Dr S. Walker (Harvard Medical School, USA) (Gross, Kraybill et al. 2005), permet d'obtenir de grandes quantités de protéine recombinante (environ 5mg/l), marquée avec une étiquette poly-histidine C-terminale. *In vitro*, cette protéine est active sur un peptide synthétique de la caséine kinase II ( $^{340}$ PGGSTPVSSANMM $^{352}$ ) (Dehennaut, Hanouille et al. 2008).

Nos premiers essais de glycosylation de Tau *in vitro* par l'*O*-GlcNAc transférase (OGT) recombinante n'ont pas permis de détecter de glycosylation. Une sous-unité régulatrice pourrait être nécessaire pour favoriser l'interaction de l'OGT avec un substrat spécifique et ainsi stimuler son activité, comme cela a été décrit pour la phosphatase PP2A, qui recrute une sous-unité régulatrice de type B pour l'interaction spécifique avec le substrat (Janssens and Goris 2001; Xu, Chen et al. 2008). Nous avons identifié, au travers d'un criblage de peptides courts issus de la protéine Tau, deux nouveaux sites d'*O*-GlcNAcylation localisés dans la région riche en prolines, les résidus S208 et S238 (Smet-Nocca, Broncel et al. 2011). Cette stratégie a confirmé le résidu S400 comme site majeur d'*O*-GlcNAcylation (Yuzwa, Yadav et al. 2011). Dans cette région C-terminale de Tau en particulier, le phospho-épitope PHF-1 (S396/S400/S404), centré autour de la sérine 400, constitue un site de phosphorylation pathologique, généré par combinaisons de deux kinases majeures impliquées dans la maladie d'Alzheimer, CDK5/p25 et GSK3 $\beta$ . Nous avons montré, sur des substrats peptidiques, de manière directe, que la phosphorylation des sérines 396 et S404 diminue l'*O*-GlcNAcylation en S400 *in vitro*

par l'*O*-GlcNAc transférase (OGT) recombinante. Réciproquement, l'*O*-GlcNAcylation en S400 diminue la phosphorylation en S404 par le complexe CDK2/cyclineA3 et supprime la phosphorylation séquentielle de l'épitope PHF-1 par GSK3β (Figure 4.5) (Smet-Nocca, Broncel et al. 2011).

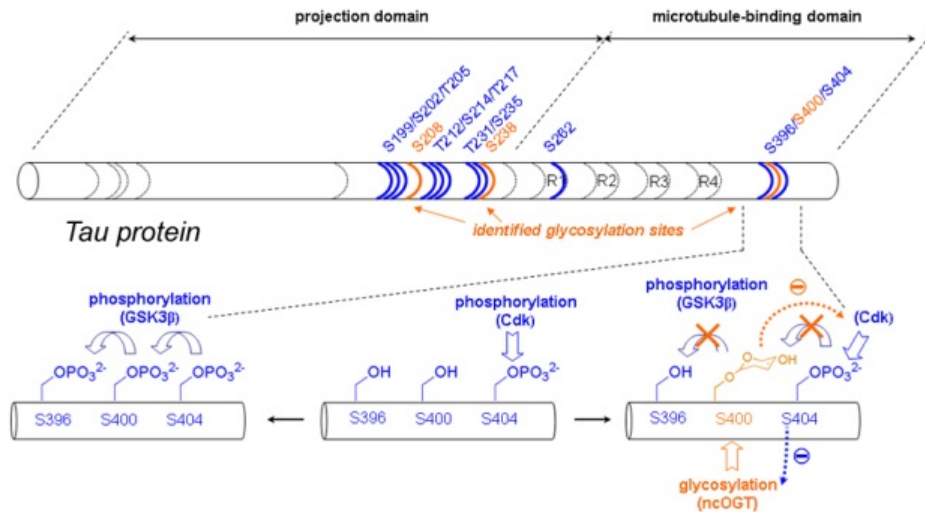


Figure 4.5 : Régulation négative et réciproque de la phosphorylation et de l'*O*-GlcNAcylation dans un peptide de Tau au niveau de l'épitope PHF-1.

### **Hypothèse d'une régulation de l'activité *O*-GlcNAc transférase de l'OGT par une sous-unité régulatrice**

De manière surprenante, la co-expression de la protéine Tau et de l'OGT dans *E. coli* permet de produire une fraction significative de Tau *O*-GlcNAcylée (50%) avec, toutefois, un rendement très faible (1 mg pour 6 litres de culture) (Yuzwa, Yadav et al. 2011; Yuzwa, Shan et al. 2012). Le mécanisme mis en jeu dans ce contexte n'a pas été établi. Une protéine d'*E. coli* pourrait jouer le rôle de partenaire de l'OGT pour faciliter l'*O*-GlcNAcylation de Tau. Bien qu'une telle protéine ne soit pas pertinente, il pourrait exister, chez l'homme, des protéines qui joueraient le même rôle. En effet, comme l'OGT est codée par un gène unique chez l'homme, le recrutement de sous-unités spécifiques permettrait une régulation de l'activité de l'OGT via la formation de complexes macromoléculaires. D'autres protéines fonctionnent sur ce mode, en formant de nombreux complexes protéine-protéine transitoires pour réguler leur spécificité (Cohen 2002; Virshup and Shenolikar 2009). On citera pour exemple la protéine PP2A, qui possède des similarités structurales avec l'OGT, avec un domaine "plate-forme", représenté par les TPR, et un domaine catalytique. PP2A possède en plus de ces deux domaines, une sous-unité régulatrice qui permet de cibler spécifiquement le substrat, comme la protéine B55 dans le cas de Tau (Figure 4.6). Il faudra envisager l'identification d'une éventuelle sous-unité régulatrice spécifique de la protéine Tau qui favoriserait l'interaction fonctionnelle Tau/OGT, et qui permettrait d'obtenir la protéine Tau sous une forme *O*-GlcNAcylée pour les études par RMN.

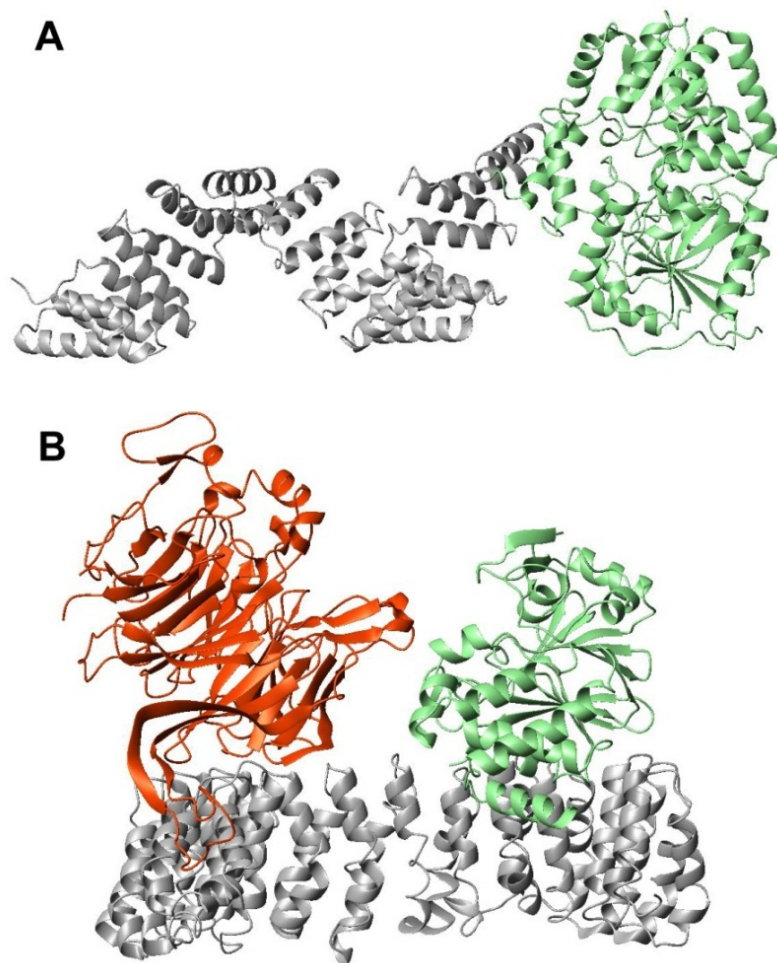


Figure 4.6 : Comparaison des structures tridimensionnelles d'un modèle de l'OGT nucléocytoplasmique (A, PDB ID : 1W3B et 2VSN) et de PP2A (B, PDB ID : 3DW8). Les sous-unités "plate-forme" sont représentées en gris, les sous-unités catalytiques en vert, et la sous-unité régulatrice B55 $\alpha$  de PP2A en rouge.

Des protéines partenaires de l'OGT ont déjà été identifiées dans un criblage en double hybride chez la levure mais aucune sous-unité régulatrice potentielle de Tau n'a été mise en évidence (Iyer, Akimoto et al. 2003; Iyer and Hart 2003; Cheung, Sakabe et al. 2008). Récemment, la protéine chaperonne Hsp90 a été identifiée comme protéine interagissant avec l'OGT (Zhang, Snead et al. 2012). Il a été montré, par ailleurs, que Tau s'associe avec les chaperonnes Hsp70 et Hsp90 (Tortosa, Santa-Maria et al. 2009; Salminen, Ojala et al. 2011). La protéine chaperonne Hsp70 interagit avec Tau hyperphosphorylée et lui permet de conserver sa capacité de liaison aux microtubules. Hsp70 préviendrait également de l'agrégation par la régulation du *turnover* de Tau (Petrucelli, Dickson et al. 2004; Patterson, Ward et al. 2011; Voss, Combs et al. 2012). De plus, Hsp70 présente une activité lectinique via la liaison aux protéines O-GlcNAcylées (Guinez, Losfeld et al. 2006; Guinez, Mir et al. 2007). Ainsi, la formation d'un complexe Tau/Hsp70/OGT pourrait intervenir dans la régulation de l'O-GlcNAcylation de Tau. Dans des souris transgéniques ou dans des cerveaux MA, les niveaux de protéines Tau agrégées et de protéines de choc thermique Hsp70/90 sont inversement corrélés. Une augmentation de l'expression de ces deux protéines favorise la solubilité de Tau, sa liaison aux microtubules et diminue son niveau de phosphorylation (Dou, Netzer et al. 2003). Des données plus

récentes sont contradictoires et indiquent qu'une inhibition de Hsp90 favorise une diminution sélective de la phosphorylation de Tau dans un modèle de souris transgénique MA (Dickey, Kamal et al. 2007). De plus, l'interaction Tau/Hsp90, décrite au niveau du peptide PHF6 (VQIVYK), localisé dans le R3, favoriserait la phosphorylation par GSK3 et la formation des fibres (Tortosa, Santa-Maria et al. 2009). Enfin, des interactions entre le complexe Hsp90/Tau et des co-chaperones partenaires de Tau comme les peptidyl-prolyl *cis/trans* isomérase FKBP51 et FKBP52 ont été décrites (Chambraud, Sardin et al. 2010; Jinwal, Koren et al. 2010). FKBP52 possède un domaine TPR similaire à celui de l'OGT qui lui permet d'établir des interactions avec la région C-terminale acide de l'Hsp90 (voir Figure 4.7) (Wu, Li et al. 2004). De manière similaire, la formation d'un complexe fonctionnel Tau/OGT/Hsp90 pourrait favoriser la liaison et la stabilisation des microtubules et prévenir la formation des formes phosphorylées de Tau et des PHF. L'interaction OGT/Hsp90 pourrait avoir lieu au niveau des TPR de l'OGT. Un alignement des TPR de FKBP52 avec ceux de l'OGT montre que l'alignement des structures se fait au niveau de la poche de liaison du peptide C-terminal MEEVD de l'Hsp90 (Figure 4.7).



---

Figure 4.7 : Alignement des TPR de l'OGT (bleu, PDB ID : 1W3B) et de FKBP52 (vert, PDB ID : 1QZ2) en complexe avec le peptide C-terminale MEEVD de l'Hsp90 lié à FKBP52 (bâtonnets roses). Le RMSD est de 1.96Å sur 167 atomes (1.30Å sur les C $\alpha$ , soit 27 atomes).

Dans un premier temps, nous avons identifié le site d'interaction de Tau pour l'OGT par RMN. En effet, une inhibition de l'activité *O*-GlcNAc transférase de l'OGT par interaction avec une région spécifique de Tau pourrait être à l'origine de l'absence d'activité observée en favorisant un ancrage trop fort au substrat. Deux sites d'interaction ont été détectés avec la ncOGT ou la sOGT en quantités stœchiométriques, en l'absence ou en présence d'UDP-GlcNAc, dans des conditions où la réaction

d'O-GlcNAcylation est défavorisée. De manière intéressante, deux courts fragments d'une dizaine de résidus, [308-319] et [284-289], ont été mis en évidence coïncidant avec les peptides PHF6\* et PHF6, impliqués dans l'amorçage du processus d'agrégation et dans la formation des structures en feuillets  $\beta$  dans les PHF (von Bergen, Friedhoff et al. 2000; Mukrasch, Biernat et al. 2005; von Bergen, Barghorn et al. 2005). Les résultats sont similaires pour les deux protéines et indépendants de l'UDP-GlcNAc. L'interaction se traduit par un élargissement significatif de ces deux régions (Figure 4.8 C), plus prononcé pour la ncOGT que pour la sOGT, et plus prononcé pour le segment PHF6 que pour PHF6\*. Il s'agit de deux régions présentant un caractère hydrophobe ainsi qu'une propension à l'agrégation en feuillets  $\beta$  (von Bergen, Friedhoff et al. 2000; Mukrasch, Bibow et al. 2009; Moore, Huang et al. 2011; Daebel, Chinnathambi et al. 2012). La liaison de l'OGT pourrait avoir un rôle protecteur, indépendant de l'O-GlcNAcylation, vis-à-vis de ces régions probablement impliquées dans le processus de nucléation lors de la formation des PHF. L'activité O-GlcNAc transférase pourra être évaluée en présence d'un peptide PHF6 compétiteur, qui modulerait l'interaction Tau/OGT.

De manière intéressante, le peptide PHF6 de Tau a été également identifié dans l'interaction avec Hsp90. Ainsi Hsp90 et OGT pourraient entrer en compétition pour la liaison à Tau via cette courte séquence PHF6, dont les conséquences fonctionnelles restent à déterminer. D'un point de vue mécanistique, il est possible que la liaison d'OGT au peptide PHF6 permette l'ancrage à Tau via les TPR favorisant ainsi à l'O-GlcNAcylation de la protéine au détriment du *turnover* enzymatique. Un fonctionnement similaire a déjà été décrit pour d'autres enzymes comme la *Thymine-DNA Glycosylase* (TDG), par exemple (Hardeland, Steinacher et al. 2002). Il faut alors l'intervention d'autres protéines pour faciliter la dissociation de l'enzyme de son substrat et augmenter le *turnover*, comme APE1 ou SUMO-1 dans le cas de TDG (Hardeland, Steinacher et al. 2002; Fitzgerald and Drohat 2008; Smet-Nocca, Wieruszkeski et al. 2008; Eilebrecht, Smet-Nocca et al. 2010; Smet-Nocca, Wieruszkeski et al. 2011). Dans le cas de l'OGT, Hsp90 pourrait, en entrant en compétition pour la liaison de PHF6, faciliter le départ de l'OGT et ainsi augmenter le niveau de glycosylation par augmentation du *turnover* enzymatique. Ces hypothèses pourront être vérifiées en réalisant l'O-GlcNAcylation de Tau par l'OGT en présence du peptide C-terminal de Hsp90.

#### **Données préliminaires sur l'O-GlcNAcylation de la protéine Tau in vitro.**

Nous avons pu améliorer significativement l'activité de l'OGT recombinante sur des substrats peptidiques, en montrant notamment qu'une augmentation de la concentration en substrat (entre 100  $\mu$ M et 2 mM) augmente significativement l'activité enzymatique (d'un facteur 3). Nous avons donc réalisé des expériences d'O-GlcNAcylation avec la protéine Tau entière dans des conditions similaires. L'analyse par immunodétection de l'échantillon avec les anticorps CTD110.6 et ab3925 (Yuzwa, Yadav et al. 2011) montre la présence de formes O-GlcNAcylées (Dr Fred Van Leuven, communication personnelle). Les données collectées par RMN n'ont pas montré directement d'O-GlcNAcylation mais le profil de l'interaction Tau/OGT est modifié par rapport à ce qui a été observé dans des conditions où l'O-GlcNAcylation est impossible (voir paragraphe précédent et Figure 4.8 C). D'une part, on ne voit plus d'interactions avec les peptides PH6 et PH6\* mais, dans le cas présent, le ratio Tau :OGT (environ 10 :1) est nettement inférieur à celui qui avait été utilisé pour cartographier l'interaction en absence de réaction de glycosylation (ratio 1 :1). D'autre part, de nombreuses résonances sont fortement élargies sous l'effet de la liaison de l'OGT du fait de la taille du complexe.

La région des sérines et thréonines du spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC est particulièrement affectée, suggérant une modification possible de plusieurs résidus le long de la séquence et une interaction multi-site. L'intensité des résonances est particulièrement diminuée dans la région riche en prolines, qui concentre un grand nombre des résidus Ser/Thr de la protéine (Figure 4.8 A et B).

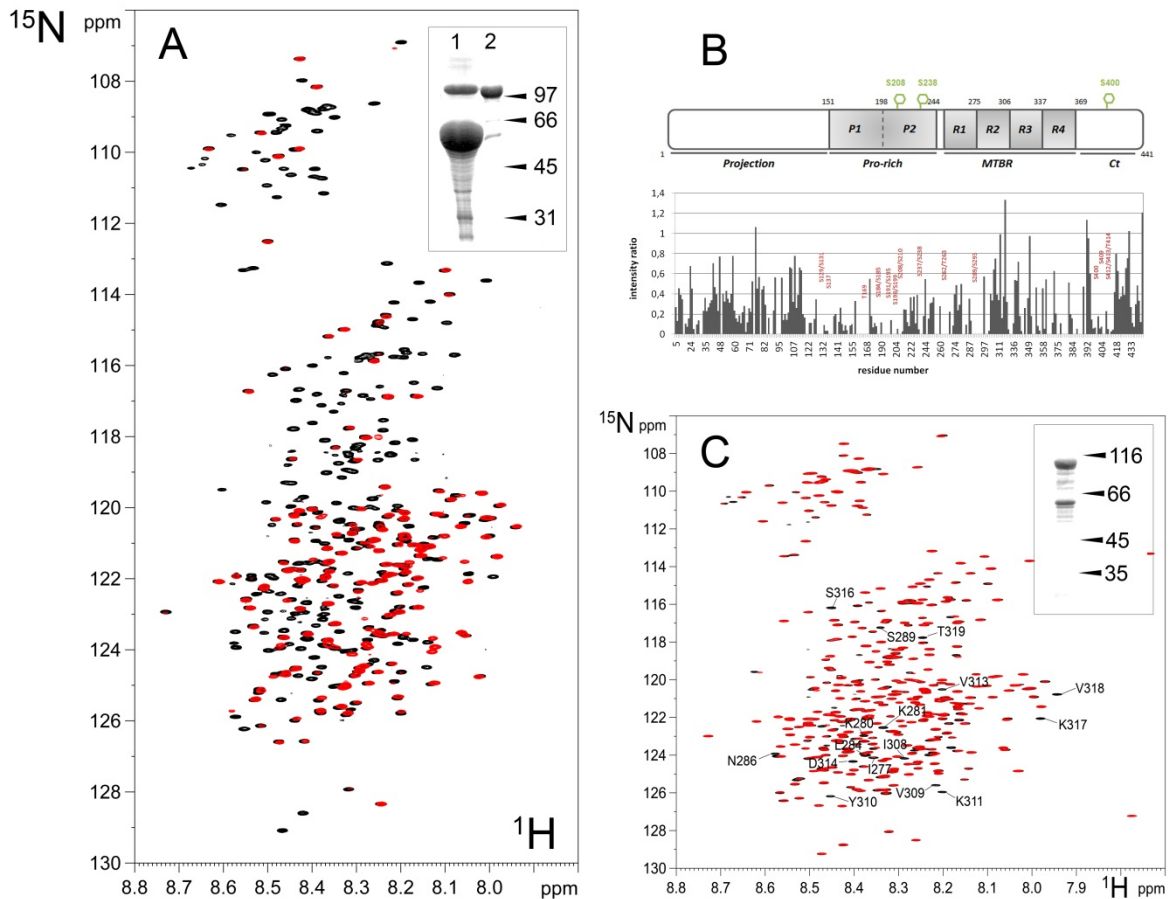


Figure 4.8 : Cartographie de l'interaction Tau/ncOGT avant et après réaction d'*O*-GlcNAcylation. (A) Comparaison des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de la protéine  $^{15}\text{N}$ -Tau441 seule (noir) ou en présence de ncOGT (rouge). En encadré, analyse sur gel dénaturant de polyacrylamide du mélange Tau441/ncOGT (ratio 10 :1) (bande 1) et de l'OGT seule (bande 2). Les marqueurs de poids moléculaires sont indiqués à droite en kDa. (B) Cartographie des intensités relatives des signaux de Tau entre les deux conditions. Les intensités ont été normalisées sur l'enveloppe des signaux des spectres  $^1\text{H}$  correspondant. Au-dessus, représentation schématique des domaines de Tau441 où sont indiqués les sites *O*-GlcNAc identifiés dans les modèles peptidiques. Projection : domaine de projection ; Pro-rich : domaine riche en prolines ; MTBR : *microtubule-binding repeats* ; Ct : domaine C-terminal. (C) Interactions de  $^{15}\text{N}$ -Tau441 avec la ncOGT en absence d'UDP-GlcNAc. En rouge, spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC de la protéine Tau à 50  $\mu\text{M}$  en présence d'une quantité équimolaire de ncOGT. En noir, spectre contrôle de Tau seule. En encadré, analyse sur gel dénaturant de polyacrylamide du mélange Tau441/ncOGT (ratio 1 :1).

Ces données sont préliminaires et l'investigation des sites d'*O*-GlcNAcylation doit être poursuivie. L'interaction avec l'OGT empêche la détection des sites de glycosylation dans la mesure où elle induit un élargissement des résonances, liée à la taille du complexe. Mais il est possible qu'un taux d'*O*-GlcNAcylation trop faible ne permette pas de détecter les résonances correspondant aux résidus



modifiés, autrement que via l'interaction avec l'OGT. Nous allons également réaliser la co-expression de Tau et de l'OGT dans *E. coli* en l'adaptant pour satisfaire aux conditions requises pour les études par RMN (marquages isotopiques, quantité).

### ***Etudes fonctionnelles d'O-GlcNAcylation spécifiques dans des protéines Tau semi-synthétiques***

Le nombre de modifications par phosphorylation ou O-GlcNAcylation dans la protéine Tau rend difficile la mise en évidence de l'impact de modifications spécifiques sur sa fonction physiologique ou sur l'agrégation pathologique. Une approche pour l'obtention de glycoprotéines est de synthétiser des protéines modifiées de manière stœchiométrique, qui contiennent une ou plusieurs modifications sur des sites spécifiques, par incorporation des acides aminés modifiés codés génétiquement. Cette stratégie a été appliquée pour la synthèse de protéines O-GlcNAcylées ou portant des oligosaccharides plus complexes dans *E. coli* et permet des études structurales et fonctionnelles ultérieures (Zhang, Gildersleeve et al. 2004). En effet, il n'est pas possible de distinguer l'effet d'une phosphorylation ou d'une O-GlcNAcylation sur la fonction d'une protéine en utilisant la mutagenèse dirigée, lorsque les deux PTM ciblent le même site, ou éventuellement des sites proches. Une autre approche est celle employant la synthèse peptidique en phase solide qui permet de synthétiser des protéines entières (jusqu'à environ 100 acides aminés) ou des protéines semi-synthétiques avec la stratégie EPL décrite précédemment pour l'obtention de protéines Tau spécifiquement phosphorylées.

L'étude de sites d'O-GlcNAcylation spécifiques, qui ont déjà été identifiés via les modèles peptidiques (Smet-Nocca, Broncel et al. 2011) ou qui seront caractérisés par RMN sur la protéine Tau entière comme décrits précédemment, pourra être réalisée avec des protéines Tau O-GlcNAcylées semi-synthétiques où l'introduction d'un groupement O-GlcNAc peut se faire par synthèse chimique (voir paragraphe 4.1). Par exemple, il est possible de synthétiser le peptide de 52 acides aminés correspondant à la région C-terminale de Tau par synthèse peptidique en phase solide en y incorporant le résidu Ser400-O-GlcNAc et de procéder à la ligation chimique native avec le reste de la protéine produite par voie recombinante, qui peut donc être enrichie en  $^{15}\text{N}$  et/ou  $^{13}\text{C}$  pour des études par RMN. On pourra alors étudier l'impact spécifique de l'O-GlcNAcylation en S400 sur la structure du fragment marqué [1-389] par RMN et sur la fonction de Tau dans la liaison des microtubules ou la polymérisation de la tubuline. Le profil de phosphorylation de la protéine Tau O-GlcNAcylée en S400 pourra être étudié par RMN comme cela a été décrit pour Tau avec l'inconvénient que la région [390-441] ne sera pas visible sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC. Les sites de phosphorylation voisins de la Ser400, qui ont la plus grande probabilité d'être affectés par l'O-GlcNAcylation, ne pourront pas être observés. Seules d'éventuelles modifications des phosphorylations à longue distance pourront être détectées.



### 4.3 ACÉTYLATION DE TAU : UN MECANISME FAVORISANT L'ACCUMULATION DES FORMES PHOSPHORYLEES ET DES PHFS ?

#### 4.3.1. Introduction

L'acétylation des lysines est une PTM importante dans la cellule et presque aussi abondante que la phosphorylation (Allfrey, Faulkner et al. 1964). L'exemple le plus connu est la régulation de la transcription des gènes via des modifications structurales de la chromatine sous l'effet de l'acétylation des queues d'histones (Strahl and Allis 2000). Depuis, différents rôles dans des fonctions cellulaires variées ont été décrits (Kim and Yang 2011). L'acétylation des lysines consiste en un transfert du groupement acétyle à partir de l'acétyl-coenzyme A (AcCoA) sur la fonction amine primaire NH<sub>2</sub> des résidus lysines des protéines. Un grand nombre de protéines, regroupées sous le terme générique d'acétylome, sont modifiées par un ou plusieurs groupements acétyles. Il faut noter que cette PTM se distingue de l'acétylation N-terminale des protéines. Comme pour la phosphorylation, diverses enzymes modifiant les lysines ont été identifiées chez les eucaryotes, et même chez les bactéries, et sont regroupées en deux familles antagonistes : les acétyltransférases (HAT pour histone acétyltransférases ou, plus généralement, KAT pour lysine acétyltransférases) et les deacétylases (HDAC pour histone deacétylases ou, plus généralement, KDAC pour lysine deacétylases) qui introduisent et coupent, respectivement, les groupements acétyles sur les protéines-substrats. Le rôle de l'acétylation dans la régulation des voies de signalisation peut impliquer la perte ou le gain de fonction. D'une part, l'acétylation, en neutralisant les charges positives qui se trouvent au bout des chaînes latérales des résidus lysine, affecte les interactions électrostatiques dans les lesquelles ces résidus sont mis en jeu. D'autre part, l'acétylation des lysines produit une nouvelle interface qui peut potentiellement modifier la structure, la fonction, la localisation ou la dégradation des protéines-cibles, en modifiant des interactions protéine-protéine. En particulier, les lysines acétylées sont reconnues spécifiquement par les bromodomains, ce qui permet, pour les protéines modulaires qui ont incorporé ces domaines, d'intervenir spécifiquement dans les réseaux de signalisation associé à l'acétylation (Bannister and Kouzarides 1996; Bannister and Miska 2000; Bannister, Miska et al. 2000; Plevoda and Sherman 2002; Brunet, Sweeney et al. 2004; Yang 2004; Yang 2004; Faiola, Liu et al. 2005; Glozak, Sengupta et al. 2005; Yuan, Guan et al. 2005). Il faut souligner que, contrairement à la phosphorylation qui cible préférentiellement les régions désordonnées des protéines (Iakoucheva, Radivojac et al. 2004; Gnad, Ren et al. 2007; Malik, Nigg et al. 2008), l'acétylation est plus fréquemment localisée dans les régions ordonnées et une proportion plus importante de lysines acétylées a été observée dans ces régions (Kim, Sprung et al. 2006; Choudhary, Kumar et al. 2009).

Récemment, deux études ont mis en évidence l'acétylation de Tau au travers, d'une part, de l'identification des enzymes impliquées, et d'autre part, l'identification des sites d'acétylation. Les acétyltransférases CBP (Cohen, Guo et al. 2011) et son proche homologue p300 (Min, Cho et al. 2010) ont toutes deux été identifiées dans l'acétylation de Tau. En revanche, les deux deacétylases identifiées, HDAC6 (Cohen, Guo et al. 2011) et SIRT1 (Min, Cho et al. 2010), n'appartiennent pas à la même famille. Les études ont mis en évidence, par incubation de Tau *in vitro* avec CBP ou p300, de nombreux sites d'acétylation dont certains communs aux deux acétyltransférases (Lys163, Lys280, Lys281, Lys369). En particulier, la Lys280 a été décrite comme site majeur d'acétylation (Cohen, Guo

et al. 2011; Irwin, Cohen et al. 2012). Ce résidu est localisé dans l'*inter-repeat* R1-R2, qui a été impliqué dans la liaison des microtubules avec une affinité deux fois plus importante que chacun des *repeats* individuels (Goode and Feinstein 1994). Deux lysines en particulier, aux positions 274 et 281, participent à la liaison des microtubules en contribuant de manière importante à l'affinité de l'*inter-repeat* R1-R2. De plus, la délétion de K280 ( $\Delta$ K280) est associée à des formes familiales de tauopathie FTDP-17. Il a été montré *in vitro* que  $\Delta$ K280 altère l'assemblage des microtubules et augmente l'agrégation sous forme de PHF. L'acétylation de Tau, et en particulier de la Lys280, diminue significativement la liaison de Tau aux microtubules et sa capacité à promouvoir leur assemblage, et stimule l'agrégation pathologique de Tau *in vitro* (Cohen, Guo et al. 2011).

De plus, acétylation et phosphorylation pourrait être liées puisque l'acétylation empêche l'ubiquitination de Tau et sa dégradation par le protéasome, comme cela a déjà été montré pour d'autres protéines (comme p53, Smad7 ou  $\beta$ -caténine) (Li, Luo et al. 2002; Monteleone, Del Vecchio Blanco et al. 2005; Ge, Jin et al. 2009), menant potentiellement à l'accumulation des formes phosphorylées de Tau (Min, Cho et al. 2010). L'augmentation de l'acétylation diminue le *turnover* de Tau dans des neurones primaires traités avec un inhibiteur de SIRT1 : la forme acétylée et la forme phosphorylée de Tau reconnue par AT8 (pS202/pT205) sont dégradées plus lentement que la forme non acétylée. Une augmentation de la forme acétylée de Tau est également détectée dans des neurones primaires traités avec de faible quantité d'oligomères A $\beta$ , de manière dose-dépendante, ou exprimant la mutation pathologique P301L, retrouvée dans la pathologie FTDP-17 (Min, Cho et al. 2010).

De manière intéressante, l'acétylation de Tau pourrait être un phénomène précoce dans la MA qui précéderait la formation des fibres hyperphosphorylées. La forme acétylée, détectée par l'anticorps polyclonal dirigé contre la Lys280 acétylée, est retrouvée dans les PHF isolés de souris transgéniques, modèles de la MA (Yoshiyama, Higuchi et al. 2007; Hurtado, Molina-Porcel et al. 2010), et dans différentes tauopathies humaines. La protéine Tau acétylée est retrouvée uniquement dans les tissus pathologiques, à tous les stades de la MA, mais plus particulièrement dans les stades avancés, associée aux épitopes pathologiques comme Alz50/MC-1, AT8, AT100, TG-3 ou PHF-1 (Irwin, Cohen et al. 2012). L'acétylation de Tau est donc plutôt un événement pathologique qui précéderait la formation d'autres PTM impliquées dans la pathologie Tau, et pourrait donc être un marqueur de la MA et d'autres tauopathies. Ainsi, les enzymes impliquées dans la dynamique de l'acétylation de Tau pourraient être de nouvelles cibles pharmacologiques dans le traitement des tauopathies.

De manière générale, les mécanismes de régulation de l'acétylation de Tau, ses effets sur la phosphorylation, la dégradation de Tau et sa fonction physiopathologique sont encore mal compris. Des données récentes indiquent que la protéine Tau serait indirectement liée à la stabilité des microtubules en modulant le niveau d'acétylation de l' $\alpha$ -tubuline. En effet, l'acétylation réversible de l' $\alpha$ -tubuline est impliquée dans la régulation de la stabilité et de la fonction des microtubules (Piperno, LeDizet et al. 1987). L' $\alpha$ -tubuline acétylée est plus abondante dans les microtubules stables et est absente des structures cellulaires dynamiques (Schulze, Asai et al. 1987; Schulze and Kirschner 1987). HDAC6, responsable de la deacétylation de Tau, est également la deacétylase majeure de l' $\alpha$ -tubuline. Elle est localisée exclusivement dans le cytoplasme où elle s'associe aux microtubules (Hubbert, Guardiola et al. 2002). Il a été montré que Tau se lie à la protéine HDAC6 via son domaine de liaison aux microtubules (Ding, Dolan et al. 2008) et inhibe l'activité deacétylase de HDAC6, ce qui

a pour effet notamment d'augmenter l'acétylation de l' $\alpha$ -tubuline (Perez, Santa-Maria et al. 2009). D'autre part, HDAC6 joue un rôle neuroprotecteur dans la coordination de la réponse cellulaire face à l'agrégation anormale des protéines via les protéines de choc thermique dont Hsp90 (Boyault, Zhang et al. 2007; Cook, Gendron et al. 2012; Cook and Petrucelli 2013). L'inactivation de HDAC6 conduit à une diminution de phosphorylations spécifiques de Tau et son niveau d'expression est significativement augmenté dans les cerveaux atteints de la MA suggérant un rôle de HDAC6 dans la régulation de la phosphorylation de Tau (Ding, Dolan et al. 2008). Cependant, ces données sont contradictoires avec le lien établi entre l'acétylation de Tau et la MA qui semble indiquer une augmentation de l'acétylation dans le contexte pathologique. De plus, l'analyse des microtubules dans les cerveaux de patients atteints de la MA montre une diminution du niveau d' $\alpha$ -tubuline ainsi qu'une augmentation de son niveau d'acétylation, principalement associées avec les neurones présentant des NFT ce qui pourrait signifier la mise en place d'un mécanisme de sauvetage visant à stabiliser les microtubules (Perez, Santa-Maria et al. 2009). Par ailleurs, il a été montré que le 4-phénylbutyrate (PBA), un inhibiteur de HDAC, qui augmente la transcription d'un grand nombre de gènes, exerce également un effet neuroprotecteur. Dans une modèle de souris transgénique MA, l'administration de PBA a permis de restaurer les processus d'apprentissage et les déficits de la mémoire (Ricobaraza, Cuadrado-Tejedor et al. 2009; Cuadrado-Tejedor, Ricobaraza et al. 2013). Bien que n'ayant aucune conséquence sur les plaques amyloïdes, le PBA a montré un effet sur la diminution de la phosphorylation de Tau et pourrait être une nouvelle approche dans le traitement de la MA.

#### **4.3.2. Etude de l'acétylation de Tau : caractérisation des sites d'acétylation par RMN et impact fonctionnel**

Lors de mon post-doctorat à l'Institut de Recherche Interdisciplinaire (2006-2008), j'ai travaillé sur la détermination du rôle du complexe formé par les protéines TDG (*Thymine DNA Glycosylase*) et CBP (*CREB-binding protein*) dans la régulation épigénétique de la transcription, impliquant notamment l'activité acétyltransférase de CBP. Pour ce projet, j'ai cloné différents fragments de la protéine CBP autour du domaine HAT, permettant leur expression recombinante dans *E. coli* et leur purification. J'ai mis au point les conditions permettant (i) l'acétylation *in vitro* de peptides ou de fragments protéiques, (ii) l'analyse des peptides et protéines acétylées par chromatographie en phase inverse et par spectrométrie de masse MALDI-TOF, et (iii) la caractérisation de l'acétylation par RMN sur des modèles peptidiques (Smet-Nocca, Wieruszski et al. 2010) qui a été décrite dans le Chapitre 3. Ces outils m'ont permis de mettre en œuvre un projet visant à déterminer le rôle de l'acétylation de Tau dans sa fonction physiopathologique qui nécessite une caractérisation préalable du profil d'acétylation.

#### **Acétylation de Tau par CBP et identification des lysines acétylées par RMN**

L'acétylation de Tau par CBP/p300 a été décrite *in vitro* (Min, Cho et al. 2010; Cohen, Guo et al. 2011). L'approche utilisant la spectrométrie de masse a permis une identification exhaustive des sites d'acétylation mais ne donne pas un aperçu des niveaux d'acétylation relatifs entre les différents sites. En particulier, le résidu Lys280 a focalisé l'attention en raison de son impact sur la liaison des microtubules et sa localisation dans le peptide PHF6\* qui joue un rôle dans la nucléation des agrégats

pathologiques avec la formation de structures en feuillets  $\beta$ . Nos essais d'acétylation avec le domaine HAT de CBP [1202-1849] (numérotation de CBP de souris) permettent de détecter par RMN la présence de plusieurs sites d'acétylation (Figure 4.9 ; voir Chapitre 3, Figure 3.6). Ici, l'avantage est de pouvoir utiliser le marquage sélectif  $^{15}\text{N}$  des résidus lysines, ce qui facilite la détection des acétylations par rapport au marquage uniforme. L'analyse des variations de l'intensité des signaux correspondant à la forme non acétylée permet dans un premier temps d'identifier les régions où vont se trouver une ou plusieurs acétylations (Figure 4.9). En effet, l'acétylation, si elle est partielle, va conduire à une diminution de l'intensité de la résonance correspondant à la forme non acétylée au profit d'une nouvelle résonance correspondant à la forme acétylée. Ceci est valable pour le résidu acétylé lui-même mais également pour ses proches voisins dans la séquence. La quantification des niveaux d'acétylation par cette méthode est possible mais reste délicate car une modification de la dynamique ou de l'échange avec l'eau, et donc de la largeur de raie, sous l'effet de l'acétylation n'est pas à exclure.

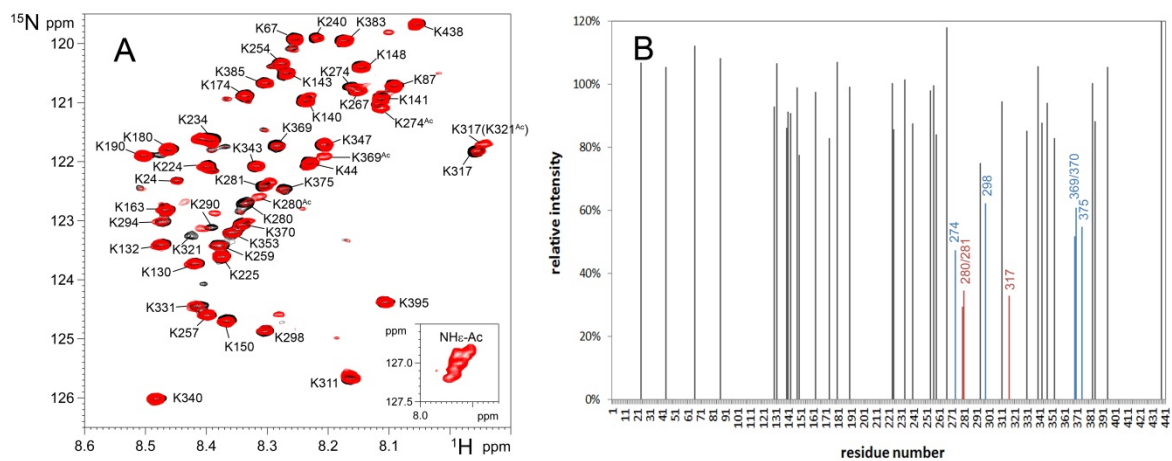


Figure 4.9 : Acétylation par CBP [1202-1849] de la protéine Tau441 (isoforme la plus longue) enrichie uniformément en  $^{13}\text{C}$  et sélectivement en  $^{15}\text{N}$ ,  $\alpha,\epsilon$ -Lys (rouge), contrôle non acétylé (noir).

L'attribution des sites d'acétylation requiert des expériences tridimensionnelles dont la plus sensible est le TOCSY-HSQC, qui permet de corréliser la résonance amide  $\text{H}_\text{N}\alpha$  ou  $\text{H}_\text{N}\epsilon$  avec les résonances protons des chaînes latérales des résidus lysines. Par exemple, des résonances spécifiques des lysines acétylées pour les protons  $\text{H}_\gamma$ ,  $\text{H}_\delta$  et  $\text{H}_\epsilon$  peuvent être utilisées en combinaison pour identifier sur le plan  $^1\text{H}$ - $^{15}\text{N}$  HSQC les résonances appartenant aux formes acétylées, et donc discriminer des signaux correspondant à une lysine acétylée de ceux d'une lysine proche d'un site d'acétylation. Les premières expériences que nous avons réalisées sur Tau441 (Figure 4.9) et Tau-F4 [208-324] (Figure 4.10) indiquent que la plupart des sites d'acétylation sont localisés dans la région basique, au niveau du domaine à segments répétés. On y retrouve les sites précédemment décrits, identifiés par spectrométrie de masse, notamment les deux motifs (N)K(K) aux positions 280 et 369 (Min, Cho et al. 2010; Cohen, Guo et al. 2011). D'autres sites, comme K224, K240, K274, K317 et K321 sont également acétylés de manière significative d'après notre étude par RMN. Jusqu'à présent, les analyses conduites par spectrométrie de masse n'avaient pas permis de graduer les niveaux d'acétylation des différents sites.

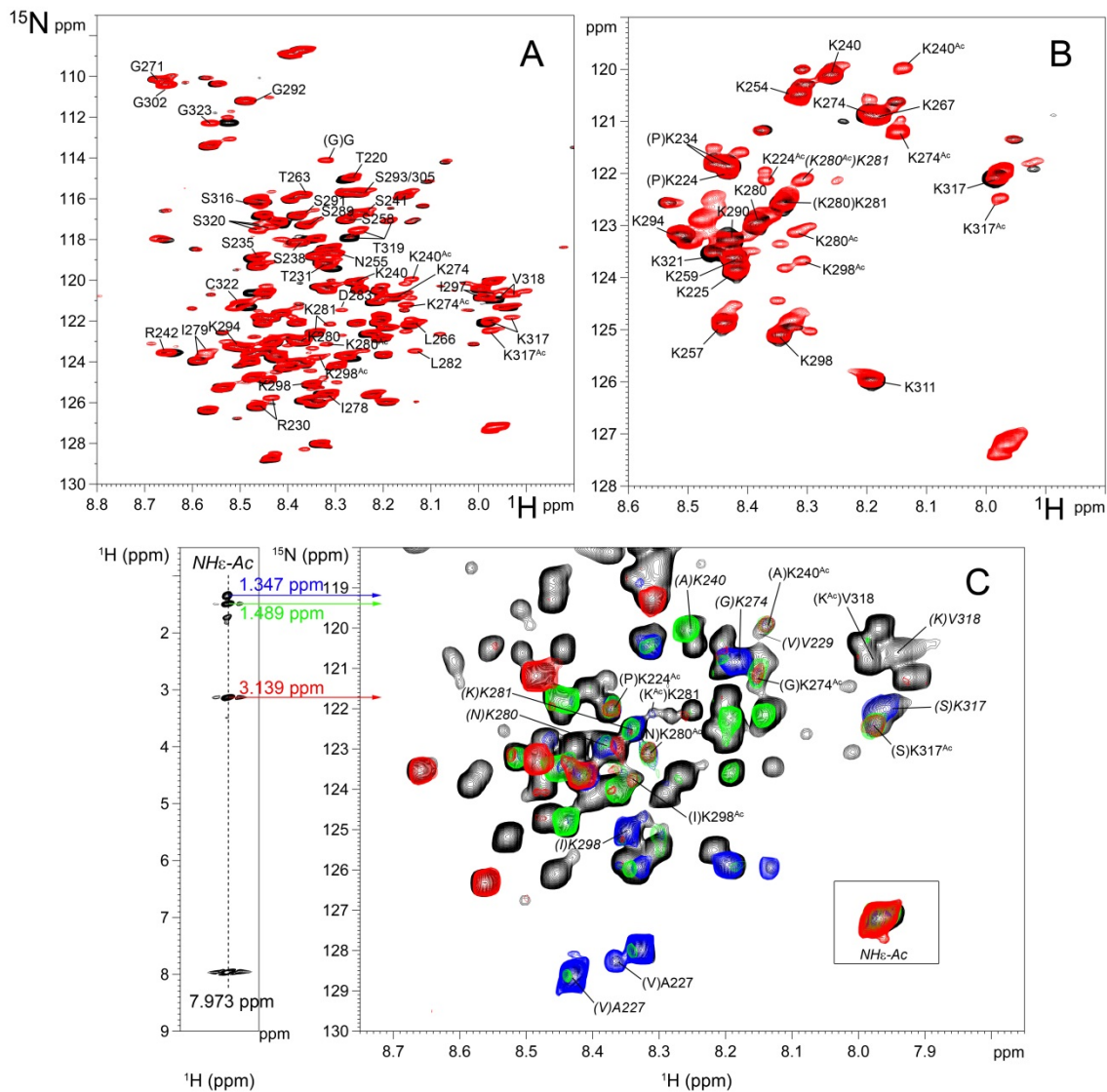


Figure 4.10 : Acétylation du fragment Tau-F4[208-324] par CBP [1202-1849]. (A,B) Spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC du fragment F4 avec marquage  $^{15}\text{N}/^{13}\text{C}$  uniforme (A) ou  $^{15}\text{N}$ -Lys/ $^{13}\text{C}$  (B) dans sa forme non acétylée (noir) et acétylée (rouge). (C) Attribution des sites d'acétylation avec l'expérience TOCSY-HSQC. A gauche, plan  $^1\text{H}$ - $^1\text{H}$  TOCSY du signal  $\text{NH}\epsilon\text{-Ac}$  (en encadré) montrant les corrélations  $^1\text{H}$ - $^1\text{H}$  des lysines acétylés. Les plans correspondant aux résonances proton des  $\text{H}\gamma$  (bleu),  $\text{H}\delta$  (vert),  $\text{H}\epsilon$  (rouge) ont été extrait du TOCSY-HSQC et sont superposés sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC (noir) à droite. Les sites d'acétylation (indiqués par Ac en exposant) correspondant aux résonances présentes dans les trois plans peuvent ainsi être détectés. Les résonances correspondant aux lysines ou aux résidus proches de lysines dans leur forme non acétylée sont indiquées en italique.

### ***Liaison aux microtubules et agrégation de Tau acétylée***

La région dans laquelle nous avons identifié la plupart des sites d'acétylation coïncide avec la région impliquée dans liaison des microtubules et dans l'agrégation. L'effet de l'acétylation sur l'agrégation pourra être étudié *in vitro*. La protéine Tau est très soluble mais son agrégation peut être induite par l'addition de polyanions comme l'héparine. L'initiation de l'agrégation et la formation des fibres par

les polyanions *in vitro* pourrait mimer l'effet de l'hyperphosphorylation. Les molécules de Tau peuvent alors former des filaments morphologiquement semblables aux PHF trouvés dans le cerveau atteint de la MA. La cinétique d'agrégation est suivie par fluorescence de la Thioflavine T, qui donne un signal de fluorescence en se liant aux structures en feuillets  $\beta$ . La formation de PHF *bona fide* peut être confirmée par microscopie électronique. Des études réalisées par RMN dans notre groupe ont également permis de cartographier le cœur des fibres de protéines  $^{15}\text{N}$ -Tau assemblées par l'héparine, en observant les régions qui restent flexibles (Sillen, Leroy et al. 2005; Sillen, Wieruszeski et al. 2005). Ces expériences pourront être reproduites avec la protéine Tau ou le fragment F4 acétylés, pour déterminer si l'acétylation modifie la cinétique d'agrégation ainsi que les régions impliquées dans la formation des fibres. L'acétylation neutralise les charges positives des lysines qui sont les résidus impliqués dans la liaison des polyanions comme l'héparine (Mukrasch, Biernat et al. 2005) et mime ainsi le mode d'action supposé de ces composés comme inducteurs de l'agrégation. Dans un premier temps, l'interaction avec l'héparine, déjà observée par RMN dans des conditions où l'agrégation ne se produit pas (Sibille, Sillen et al. 2006), pourra être reproduite avec la protéine Tau acétylée pour déterminer si l'acétylation modifie le profil d'interaction avec l'héparine auquel cas on peut s'attendre à ce que l'agrégation induite par l'héparine dans des tests *in vitro* soit perturbée par l'acétylation.

Du point de vue des charges, l'acétylation a un effet similaire à la phosphorylation. La première neutralise des charges positives tandis que la seconde apporte des charges négatives. Le bilan des charges revient alors à une augmentation globale de l'acidité de la protéine. Cependant, la phosphorylation s'adresse plutôt aux régions qui flanquent le domaine des *repeats* alors que l'acétylation cible plutôt ce dernier. Donc, localement, l'effet des deux PTM sera différent. Ainsi, on pourra s'intéresser à l'effet de l'acétylation sur la liaison à T2R et sur la polymérisation des microtubules, comme on l'a envisagé pour la phosphorylation. Comme pour l'agrégation, il est possible d'observer par RMN, les régions de Tau qui conservent leur flexibilité/mobilité lors de la liaison aux microtubules ou à la tubuline via le modèle T2R (Dr Benoît Gigant et Dr Marcel Knossow, Paris), constitué d'un dimère de tubuline, stabilisé par le domaine RB3 et incapable de polymériser (Gigant, Curmi et al. 2000; Fauquant, Redeker et al. 2011). Puisque des résidus lysines, notamment K274 et K281, ont été décrits comme essentiels à la liaison des microtubules, l'acétylation de K280 pourrait jouer un rôle important dans la modulation de la liaison de Tau aux microtubules. D'autres sites pourraient avoir un effet similaire et des mutations ponctuelles des sites d'acétylation pourront être envisagées pour évaluer l'effet de sites spécifiques dans cette fonction.

### ***Interactions de Tau et Tau acétylée avec les lysine-tweezers***

La protéine Tau  $^{15}\text{N}$  sélectivement enrichie a été utilisée pour cartographier les interactions avec des pinces moléculaires (*tweezers*) décrites dans la liaison spécifique des résidus lysines, qu'on retrouve fréquemment impliqués dans l'agrégation des protéines amyloïdogéniques (Talbinsky, Bastkowski et al. 2008; Sinha, Lopes et al. 2011; Sinha, Lopes et al. 2012). Parmi ces molécules, le *lysine-tweezer* CLR01 (Talbinsky, Bastkowski et al. 2008) inhibe l'agrégation et la toxicité de nombreuses protéines amyloïdogéniques dont le peptide amyloïde, en interférant dans les interactions intermoléculaires impliquées dans les processus d'agrégation (Sinha, Lopes et al. 2011). Il a été proposé que ces molécules puissent intervenir aux premiers stades de l'assemblage, lors de la formation des

oligomères solubles toxiques. Cette molécule CLR01 inhibe la toxicité du peptide A $\beta$  1-42 *in vivo*. Dans un modèle de souris transgénique de la MA, CLR01 diminue les agrégats de peptide A $\beta$  1-42 dans le cerveau ainsi que l'hyperphosphorylation de la protéine Tau (Attar, Ripoli et al. 2012). Ces composés ne présentent pas de toxicité aux concentrations requises pour détecter l'effet sur l'inhibition de l'agrégation chez la souris et passent la barrière hémato-encéphalique. Ces données indiquent un mode d'action spécifique de ces *tweezers* moléculaires qui pourraient par conséquent être le point de départ d'une nouvelle stratégie thérapeutique dans le traitement de la MA et d'autres maladies apparentées en régulant le processus d'assemblage via la formation d'oligomères ou d'agrégats non toxiques (Liu and Bitan 2012).

Nous avons commencé l'investigation des interactions du *lysine-tweezer* CLR01 avec la protéine Tau dans le cadre d'une collaboration avec les Dr Gal Bitan (University of California, Los Angeles, USA) et Thomas Schrader (University Duisburg-Essen, Allemagne). Dans un premier temps, le but est d'identifier les résidus lysine impliqués dans l'interaction avec le composé. Bien que dans le contexte d'une protéine désordonnée comme Tau, il n'y ait *a priori* pas de discrimination selon l'accessibilité, nous avons observé une sélectivité de la liaison du CLR01 pour un certain nombre de lysines, localisées principalement dans la région des *repeats* et les régions voisines (fragment [224-441]). Nous avons observé une diminution de l'intensité des résonances pour une vingtaine de pics dans l'expérience  $^1\text{H}$ - $^{15}\text{N}$  HSQC sur les 44 lysines de Tau441, due à l'élargissement des signaux. Ce phénomène est dû à l'échange entre les états libre et lié pour chaque résidu participant à l'interaction avec CLR01. Un tel élargissement suggère une différence importante de fréquence de résonance entre les deux états ( $\Delta\omega_{\text{max}}$ ) sous l'effet de la liaison. L'interaction semble donc cibler la région fonctionnelle de Tau impliquée dans la liaison des microtubules, qui coïncident avec une plus grande proportion de résidus basiques (29 lysines sur 44). De manière intéressante, on commence à détecter cet élargissement pour certaines résonances dès le ratio Tau :CLR01 1 :0.5, soit 0.025 molécule CLR01 par résidu lysine, si on considère que 20 lysines participent à l'interaction, ce qui indique une interaction dans une gamme de 1-10  $\mu\text{M}$ . Dans cette gamme de  $K_D$ , on devrait pouvoir atteindre la saturation avec un excès molaire de 40 de CLR01 par rapport à Tau. Une précipitation importante de la protéine en présence d'un large excès de CLR01 a été observée et ne permet pas de mener le titrage jusqu'à saturation et mesurer le  $\Delta\omega_{\text{max}}$ .

Nous allons donc réaliser, sur un fragment plus court, le fragment Tau-F4 [208-324], contenant 17 lysines, des expériences de dispersion de relaxation pour identifier les lysines qui se lient préférentiellement au CLR01. En effet, avec ce type d'expérience, l'élargissement des raies peut être détecté de manière beaucoup plus sensible que dans l'expérience  $^1\text{H}$ - $^{15}\text{N}$  HSQC. On pourra donc descendre à des ratios Tau :CLR01 inférieurs à 1 :0.5. Le but est d'identifier les sites préférentiels de l'interaction et de mettre en évidence le mode d'action de ce composé dans l'inhibition de l'agrégation de Tau. Des expériences d'agrégation *in vitro* seront réalisées en présence d'héparine et la formation des PHF détectée par l'émission de fluorescence de la Thioflavine T (Friedhoff, Schneider et al. 1998; Sillen, Leroy et al. 2005; Sibille, Sillen et al. 2006). Les données que nous avons collectées à ce jour vont dans le sens de ce qui a été décrit pour l'interaction de ce *tweezer* moléculaire avec le peptide A $\beta$  où la molécule se lie aux différents résidus lysines exposés à la surface (Sinha, Lopes et al. 2011; Sinha, Lopes et al. 2012). La cartographie des interactions que nous avons observées avec Tau, montre une sélectivité modérée dans ce contexte de protéines non structurées, ciblant toutefois

préférentiellement la région impliquée dans la formation des fibres pathologiques. Ces interactions multi-site empêcheraient les interactions intermoléculaires pour former des oligomères. La situation est différente dans le cas des protéines globulaires, comme décrit récemment avec la protéine adaptatrice 14-3-3 où une seule lysine exposée au solvant, à l'extrémité d'une hélice  $\alpha$ , est capable de se lier au composé CLR01 dont la liaison inhibe les interactions protéine-protéine avec d'autres partenaires (Bier, Rose et al. 2013).

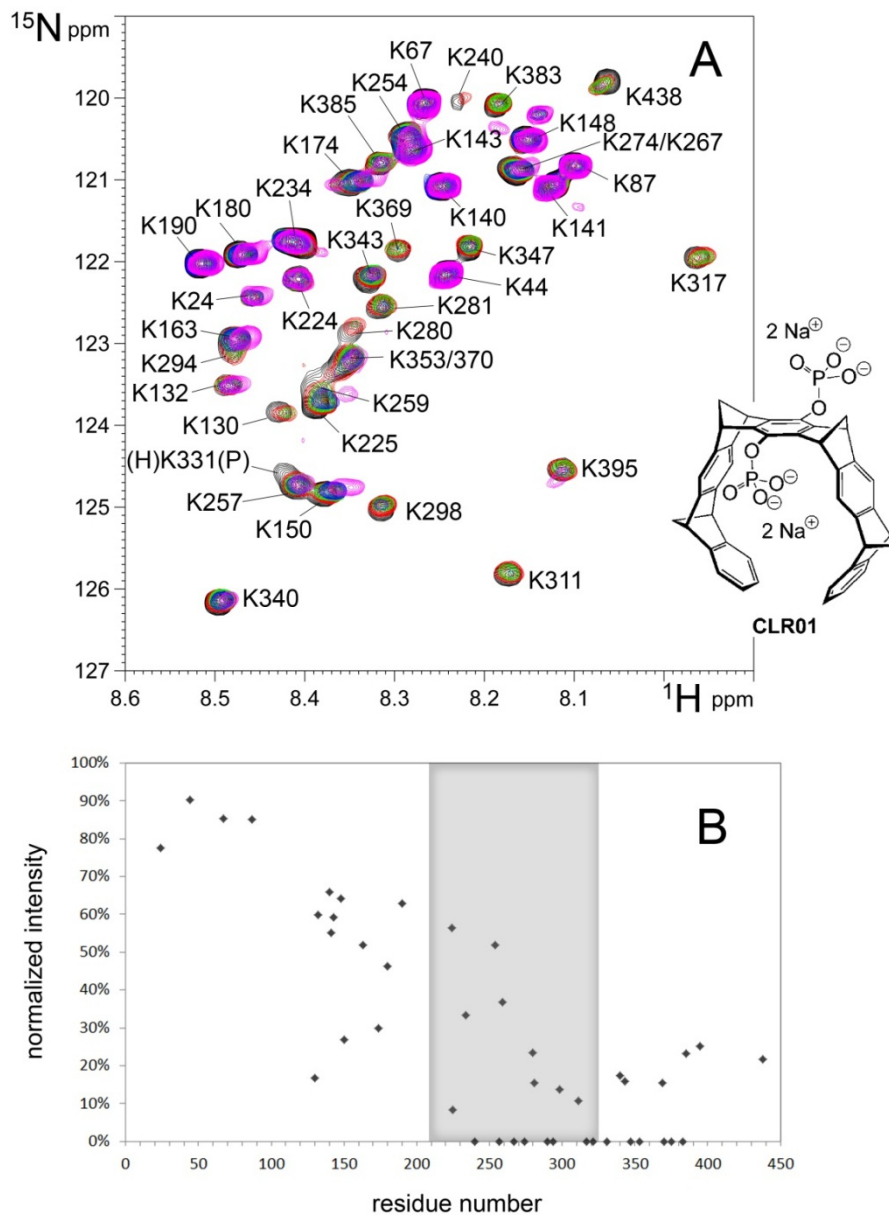


Figure 4.11 : Interaction du composé CLR01 dont la structure est indiquée (A) avec la protéine  $^{15}\text{N}\alpha,\epsilon\text{-Lys-Tau441}$ . (A) Titration de la protéine à  $100\mu\text{M}$  (noir) avec différentes concentrations de CLR01 :  $50\mu\text{M}$  (rouge),  $100\mu\text{M}$  (vert),  $200\mu\text{M}$  (bleu) et  $400\mu\text{M}$  (rose). (B) Cartographie des interactions de CLR01 avec les lysines de Tau à un ratio Tau :CLR01 de 1 :4. La région grisée correspond au fragment Tau-F4 [208-324].



L'autre volet de ce projet est d'envisager l'utilisation du composé CLR01 comme outil de détection des sites d'acétylation. Le but est de déterminer s'il est possible de détecter, par différence, des formes acétylées dans un échantillon hétérogène avec ces *tweezers* moléculaires. En effet, l'identification des lysines acétylées dans un spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC n'est pas triviale du fait de la dégénérescence des signaux qui rend difficile la détection des formes acétylées. L'acétylation devrait empêcher la liaison du *tweezer* moléculaire et, par conséquent, si l'interaction du *tweezer* avec les lysines non acétylées "éteint" les résonances de ces dernières, la détection et l'identification des résonances correspondant aux lysines acétylées deviendrait plus facile.

### ***Interactions de Tau acétylée avec l'ADN***

Bien que Tau ait été décrite comme protéine associée aux microtubules, elle est également trouvée au niveau nucléaire dans les cellules neuronales et non neuronales (Loomis, Howard et al. 1990). Une augmentation de Tau déphosphorylée a été observée dans le noyau des neurones après un choc thermique. Sa capacité d'interactions avec l'ADN est augmentée et elle joue un rôle protecteur de l'ADN génomique en réponse au stress (Sultan, Nessler et al. 2011). Les lysines sont des résidus importants pour l'interaction avec les acides nucléiques. Une étude des interactions Tau :ADN, menée avec le Dr Isabelle Landrieu dans notre laboratoire et le Dr Marie-Christine Gallas (Centre de Recherche Jean Pierre Aubert, Inserm U837, Lille) a récemment débutée. Ces travaux ont déjà permis d'identifier par RMN les régions de Tau impliquées dans la liaison à l'ADN et d'évaluer l'impact de la phosphorylation sur ces interactions. Nous allons également analyser les conséquences de l'acétylation des lysines de Tau sur son interaction avec l'ADN.

---

## Références

---

- Ahlijanian, M. K., N. X. Barrezueta, et al. (2000). "Hyperphosphorylated tau and neurofilament and cytoskeletal disruptions in mice overexpressing human p25, an activator of cdk5." Proceedings of the National Academy of Sciences of the United States of America **97**(6): 2910-2915.
- Akimoto, Y., G. W. Hart, et al. (2005). "O-GlcNAc modification of nucleocytoplasmic proteins and diabetes." Medical molecular morphology **38**(2): 84-91.
- Akiyama, H., R. W. Shin, et al. (2005). "Pin1 promotes production of Alzheimer's amyloid beta from beta-cleaved amyloid precursor protein." Biochem Biophys Res Commun **336**(2): 521-529.
- Akoury, E., M. Gajda, et al. (2013). "Inhibition of tau filament formation by conformational modulation." J Am Chem Soc **135**(7): 2853-2862.
- Albert, A., S. Lavoie, et al. (1999). "A hyperphosphorylated form of RNA polymerase II is the major interphase antigen of the phosphoprotein antibody MPM-2 and interacts with the peptidyl-prolyl isomerase Pin1." J Cell Sci **112 ( Pt 15)**: 2493-2500.
- Albert, A. L., S. B. Lavoie, et al. (2004). "Multisite phosphorylation of Pin1-associated mitotic phosphoproteins revealed by monoclonal antibodies MPM-2 and CC-3." BMC Cell Biol **5**(1): 22.
- Albrecht, S., M. Bourdeau, et al. (2007). "Activation of caspase-6 in aging and mild cognitive impairment." The American journal of pathology **170**(4): 1200-1209.
- Alfaro, J. F., C. X. Gong, et al. (2012). "Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain-specific O-GlcNAc transferase targets." Proceedings of the National Academy of Sciences of the United States of America **109**(19): 7280-7285.
- Allfrey, V. G., R. Faulkner, et al. (1964). "Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis." Proceedings of the National Academy of Sciences of the United States of America **51**: 786-794.
- Alonso, A., T. Zaidi, et al. (2001). "Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments/straight filaments." Proc Natl Acad Sci U S A **98**(12): 6923-6928.
- Alonso, A. C., I. Grundke-Iqbal, et al. (1996). "Alzheimer's disease hyperphosphorylated tau sequesters normal tau into tangles of filaments and disassembles microtubules." Nature medicine **2**(7): 783-787.
- Alonso, A. C., T. Zaidi, et al. (1994). "Role of abnormally phosphorylated tau in the breakdown of microtubules in Alzheimer disease." Proceedings of the National Academy of Sciences of the United States of America **91**(12): 5562-5566.
- Amniai, L., P. Barbier, et al. (2009). "Alzheimer disease specific phosphoepitopes of Tau interfere with assembly of tubulin but not binding to microtubules." Faseb J **23**(4): 1146-1152.
- Ando, K., P. Dourlen, et al. (2013). "Tau pathology modulates Pin1 post-translational modifications and may be relevant as biomarker." Neurobiology of aging **34**(3): 757-769.
- Andronesi, O. C., M. von Bergen, et al. (2008). "Characterization of Alzheimer's-like paired helical filaments from the core domain of tau protein using solid-state NMR spectroscopy." J Am Chem Soc **130**(18): 5922-5928.
- Arai, H., M. Terajima, et al. (1995). "Tau in cerebrospinal fluid: a potential diagnostic marker in Alzheimer's disease." Annals of neurology **38**(4): 649-652.
- Armstrong, R. A., N. J. Cairns, et al. (2001). "What does the study of the spatial patterns of pathological lesions tell us about the pathogenesis of neurodegenerative disorders?" Neuropathology : official journal of the Japanese Society of Neuropathology **21**(1): 1-12.

- Arnold, C. S., G. V. Johnson, et al. (1996). "The microtubule-associated protein tau is extensively modified with O-linked N-acetylglucosamine." *J Biol Chem* **271**(46): 28741-28744.
- Attar, A., C. Ripoli, et al. (2012). "Protection of primary neurons and mouse brain from Alzheimer's pathology by molecular tweezers." *Brain : a journal of neurology* **135**(Pt 12): 3735-3748.
- Augustinack, J. C., A. Schneider, et al. (2002). "Specific tau phosphorylation sites correlate with severity of neuronal cytopathology in Alzheimer's disease." *Acta neuropathologica* **103**(1): 26-35.
- Balastik, M., J. Lim, et al. (2007). "Pin1 in Alzheimer's disease: multiple substrates, one regulatory mechanism?" *Biochim Biophys Acta* **1772**(4): 422-429.
- Ballatore, C., V. M. Lee, et al. (2007). "Tau-mediated neurodegeneration in Alzheimer's disease and related disorders." *Nature reviews. Neuroscience* **8**(9): 663-672.
- Banerjee, P. S., G. W. Hart, et al. (2012). "Chemical approaches to study O-GlcNAcylation." *Chemical Society reviews*.
- Bannister, A. J. and T. Kouzarides (1996). "The CBP co-activator is a histone acetyltransferase." *Nature* **384**(6610): 641-643.
- Bannister, A. J. and E. A. Miska (2000). "Regulation of gene expression by transcription factor acetylation." *Cellular and molecular life sciences : CMLS* **57**(8-9): 1184-1192.
- Bannister, A. J., E. A. Miska, et al. (2000). "Acetylation of importin-alpha nuclear import factors by CBP/p300." *Current biology : CB* **10**(8): 467-470.
- Barone, M. C., L. A. Desouza, et al. (2008). "Pin1 promotes cell death in NGF-dependent neurons through a mechanism requiring c-Jun activity." *J Neurochem* **106**(2): 734-745.
- Behrsin, C. D., M. L. Bailey, et al. (2007). "Functionally important residues in the peptidyl-prolyl isomerase Pin1 revealed by unigenic evolution." *J Mol Biol* **365**(4): 1143-1162.
- Bekei, B., H. M. Rose, et al. (2012). "In-cell NMR in mammalian cells: part 1." *Methods in molecular biology* **895**: 43-54.
- Bekei, B., H. M. Rose, et al. (2012). "In-cell NMR in mammalian cells: part 2." *Methods in molecular biology* **895**: 55-66.
- Bekei, B., H. M. Rose, et al. (2012). "In-cell NMR in mammalian cells: part 3." *Methods in molecular biology* **895**: 67-83.
- Bibow, S., M. D. Mukrasch, et al. (2011). "The dynamic structure of filamentous tau." *Angewandte Chemie* **50**(48): 11520-11524.
- Bibow, S., V. Ozenne, et al. (2011). "Structural impact of proline-directed pseudophosphorylation at AT8, AT100, and PHF1 epitopes on 441-residue tau." *J Am Chem Soc* **133**(40): 15842-15845.
- Bielska, A. A. and N. J. Zondlo (2006). "Hyperphosphorylation of tau induces local polyproline II helix." *Biochemistry* **45**(17): 5527-5537.
- Bier, D., R. Rose, et al. (2013). "Molecular tweezers modulate 14-3-3 protein-protein interactions." *Nature chemistry* **5**(3): 234-239.
- Biernat, J., N. Gustke, et al. (1993). "Phosphorylation of Ser262 strongly reduces binding of tau to microtubules: distinction between PHF-like immunoreactivity and microtubule binding." *Neuron* **11**(1): 153-163.
- Binder, L. I., A. L. Guillozet-Bongaarts, et al. (2005). "Tau, tangles, and Alzheimer's disease." *Biochimica et biophysica acta* **1739**(2-3): 216-223.
- Bodart, J. F., J. M. Wieruszkeski, et al. (2008). "NMR observation of Tau in Xenopus oocytes." *J Magn Reson* **192**(2): 252-257.
- Boucher, D., J. C. Larcher, et al. (1994). "Polyglutamylolation of tubulin as a progressive regulator of in vitro interactions between the microtubule-associated protein Tau and tubulin." *Biochemistry* **33**(41): 12471-12477.
- Boyault, C., Y. Zhang, et al. (2007). "HDAC6 controls major cell response pathways to cytotoxic accumulation of protein aggregates." *Genes & development* **21**(17): 2172-2181.

- Braak, H., I. Alafuzoff, et al. (2006). "Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry." *Acta neuropathologica* **112**(4): 389-404.
- Braak, H. and E. Braak (1991). "Neuropathological staging of Alzheimer-related changes." *Acta neuropathologica* **82**(4): 239-259.
- Brady, R. M., R. P. Zinkowski, et al. (1995). "Presence of tau in isolated nuclei from human brain." *Neurobiology of aging* **16**(3): 479-486.
- Brandt, R., J. Leger, et al. (1995). "Interaction of tau with the neural plasma membrane mediated by tau's amino-terminal projection domain." *The Journal of cell biology* **131**(5): 1327-1340.
- Bretteville, A. and E. Planel (2008). "Tau aggregates: toxic, inert, or protective species?" *Journal of Alzheimer's disease : JAD* **14**(4): 431-436.
- Broncel, M., E. Krause, et al. (2012). "The Alzheimer's disease related tau protein as a new target for chemical protein engineering." *Chemistry* **18**(9): 2488-2492.
- Brown, N. R., M. E. Noble, et al. (1999). "The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases." *Nat Cell Biol* **1**(7): 438-443.
- Brunden, K. R., J. Q. Trojanowski, et al. (2008). "Evidence that non-fibrillar tau causes pathology linked to neurodegeneration and behavioral impairments." *Journal of Alzheimer's disease : JAD* **14**(4): 393-399.
- Brunden, K. R., J. Q. Trojanowski, et al. (2009). "Advances in tau-focused drug discovery for Alzheimer's disease and related tauopathies." *Nature reviews. Drug discovery* **8**(10): 783-793.
- Brunet, A., L. B. Sweeney, et al. (2004). "Stress-dependent regulation of FOXO transcription factors by the SIRT1 deacetylase." *Science* **303**(5666): 2011-2015.
- Butkinaree, C., K. Park, et al. (2009). "O-linked beta-N-acetylglucosamine (O-GlcNAc): Extensive crosstalk with phosphorylation to regulate signaling and transcription in response to nutrients and stress." *Biochim Biophys Acta*.
- Canu, N. and P. Calissano (2003). "In vitro cultured neurons for molecular studies correlating apoptosis with events related to Alzheimer disease." *Cerebellum* **2**(4): 270-278.
- Carmel, G., E. M. Mager, et al. (1996). "The structural basis of monoclonal antibody Alz50's selectivity for Alzheimer's disease pathology." *The Journal of biological chemistry* **271**(51): 32789-32795.
- Chambraud, B., E. Sardin, et al. (2010). "A role for FKBP52 in Tau protein function." *Proceedings of the National Academy of Sciences of the United States of America* **107**(6): 2658-2663.
- Chau, M. F., M. J. Radeke, et al. (1998). "The microtubule-associated protein tau cross-links to two distinct sites on each alpha and beta tubulin monomer via separate domains." *Biochemistry* **37**(51): 17692-17703.
- Chen, J., Y. Kanai, et al. (1992). "Projection domains of MAP2 and tau determine spacings between microtubules in dendrites and axons." *Nature* **360**(6405): 674-677.
- Chen, Y. X., J. T. Du, et al. (2006). "Alternative O-GlcNAcylation/O-phosphorylation of Ser16 induce different conformational disturbances to the N terminus of murine estrogen receptor beta." *Chem Biol* **13**(9): 937-944.
- Cheng, X., R. N. Cole, et al. (2000). "Alternative O-glycosylation/O-phosphorylation of the murine estrogen receptor beta." *Biochemistry* **39**(38): 11609-11620.
- Cheng, X. and G. W. Hart (2000). "Glycosylation of the murine estrogen receptor-alpha." *The Journal of steroid biochemistry and molecular biology* **75**(2-3): 147-158.
- Cheng, X. and G. W. Hart (2001). "Alternative O-glycosylation/O-phosphorylation of serine-16 in murine estrogen receptor beta: post-translational regulation of turnover and transactivation activity." *The Journal of biological chemistry* **276**(13): 10570-10575.
- Cheung, W. D., K. Sakabe, et al. (2008). "O-linked beta-N-acetylglucosaminyltransferase substrate specificity is regulated by myosin phosphatase targeting and other interacting proteins." *J Biol Chem* **283**(49): 33935-33941.

- Choudhary, C., C. Kumar, et al. (2009). "Lysine acetylation targets protein complexes and co-regulates major cellular functions." *Science* **325**(5942): 834-840.
- Citron, M. (2010). "Alzheimer's disease: strategies for disease modification." *Nature reviews. Drug discovery* **9**(5): 387-398.
- Clavaguera, F., T. Bolmont, et al. (2009). "Transmission and spreading of tauopathy in transgenic mouse brain." *Nat Cell Biol* **11**(7): 909-913.
- Cleveland, D. W., S. Y. Hwo, et al. (1977). "Physical and chemical properties of purified tau factor and the role of tau in microtubule assembly." *J Mol Biol* **116**(2): 227-247.
- Cleveland, D. W., S. Y. Hwo, et al. (1977). "Purification of tau, a microtubule-associated protein that induces assembly of microtubules from purified tubulin." *J Mol Biol* **116**(2): 207-225.
- Cohen, P. T. (2002). "Protein phosphatase 1--targeted in many directions." *Journal of cell science* **115**(Pt 2): 241-256.
- Cohen, T. J., J. L. Guo, et al. (2011). "The acetylation of tau inhibits its function and promotes pathological tau aggregation." *Nature communications* **2**: 252.
- Comer, F. I. and G. W. Hart (2000). "O-Glycosylation of nuclear and cytosolic proteins. Dynamic interplay between O-GlcNAc and O-phosphate." *J Biol Chem* **275**(38): 29179-29182.
- Congdon, E. E. and K. E. Duff (2008). "Is tau aggregation toxic or protective?" *Journal of Alzheimer's disease : JAD* **14**(4): 453-457.
- Connell, P., C. A. Ballinger, et al. (2001). "The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins." *Nat Cell Biol* **3**(1): 93-96.
- Conti, E. and J. Kuriyan (2000). "Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha." *Structure* **8**(3): 329-338.
- Cook, C., T. F. Gendron, et al. (2012). "Loss of HDAC6, a novel CHIP substrate, alleviates abnormal tau accumulation." *Hum Mol Genet* **21**(13): 2936-2945.
- Cook, C. and L. Petrucelli (2013). "Tau triage decisions mediated by the chaperone network." *Journal of Alzheimer's disease : JAD* **33** Suppl 1: S145-151.
- Crenshaw, D. G., J. Yang, et al. (1998). "The mitotic peptidyl-prolyl isomerase, Pin1, interacts with Cdc25 and Plx1." *Embo J* **17**(5): 1315-1327.
- Cuadrado-Tejedor, M., A. L. Ricobaraza, et al. (2013). "Phenylbutyrate is a multifaceted drug that exerts neuroprotective effects and reverses the Alzheimer s disease-like phenotype of a commonly used mouse model." *Current pharmaceutical design*.
- Daebel, V., S. Chinnathambi, et al. (2012). "beta-Sheet core of tau paired helical filaments revealed by solid-state NMR." *J Am Chem Soc* **134**(34): 13982-13989.
- Daly, N. L., R. Hoffmann, et al. (2000). "Role of phosphorylation in the conformation of tau peptides implicated in Alzheimer's disease." *Biochemistry* **39**(30): 9039-9046.
- Davis, F. M., T. Y. Tsao, et al. (1983). "Monoclonal antibodies to mitotic cells." *Proc Natl Acad Sci U S A* **80**(10): 2926-2930.
- Dawson, H. N., V. Cantillana, et al. (2010). "Loss of tau elicits axonal degeneration in a mouse model of Alzheimer's disease." *Neuroscience* **169**(1): 516-531.
- de Calignon, A., L. M. Fox, et al. (2010). "Caspase activation precedes and leads to tangles." *Nature* **464**(7292): 1201-1204.
- Dehennaut, V., X. Hanouille, et al. (2008). "Microinjection of recombinant O-GlcNAc transferase potentiates Xenopus oocytes M-phase entry." *Biochem Biophys Res Commun* **369**(2): 539-546.
- Dehennaut, V., T. Lefebvre, et al. (2007). "O-linked N-acetylglucosaminyltransferase inhibition prevents G2/M transition in Xenopus laevis oocytes." *J Biol Chem* **282**(17): 12527-12536.
- Delobel, P., S. Flament, et al. (2002). "Abnormal Tau phosphorylation of the Alzheimer-type also occurs during mitosis." *J Neurochem* **83**(2): 412-420.
- Deshaies, R. J. and J. E. Ferrell, Jr. (2001). "Multisite phosphorylation and the countdown to S phase." *Cell* **107**(7): 819-822.

- Dias, W. B. and G. W. Hart (2007). "O-GlcNAc modification in diabetes and Alzheimer's disease." Molecular bioSystems **3**(11): 766-772.
- Dickey, C. A., A. Kamal, et al. (2007). "The high-affinity HSP90-CHIP complex recognizes and selectively degrades phosphorylated tau client proteins." J Clin Invest **117**(3): 648-658.
- Ding, H., P. J. Dolan, et al. (2008). "Histone deacetylase 6 interacts with the microtubule-associated protein tau." J Neurochem **106**(5): 2119-2130.
- Dong, L., J. Marakovits, et al. (2010). "Structure-based design of novel human Pin1 inhibitors (II)." Bioorg Med Chem Lett **20**(7): 2210-2214.
- Dorval, V. and P. E. Fraser (2006). "Small ubiquitin-like modifier (SUMO) modification of natively unfolded proteins tau and alpha-synuclein." The Journal of biological chemistry **281**(15): 9919-9924.
- Dou, F., W. J. Netzer, et al. (2003). "Chaperones increase association of tau protein with microtubules." Proceedings of the National Academy of Sciences of the United States of America **100**(2): 721-726.
- Drewes, G., E. M. Mandelkow, et al. (1993). "Dephosphorylation of tau protein and Alzheimer paired helical filaments by calcineurin and phosphatase-2A." FEBS Lett **336**(3): 425-432.
- Du, J. T., Y. M. Li, et al. (2005). "Low-barrier hydrogen bond between phosphate and the amide group in phosphopeptide." J Am Chem Soc **127**(47): 16350-16351.
- Duncan, K. E., B. R. Dempsey, et al. (2011). "Discovery and characterization of a nonphosphorylated cyclic peptide inhibitor of the peptidylprolyl isomerase, Pin1." Journal of medicinal chemistry **54**(11): 3854-3865.
- Dyson, H. J. and P. E. Wright (2004). "Unfolded proteins and protein folding studied by NMR." Chemical reviews **104**(8): 3607-3622.
- Eidenmuller, J., T. Fath, et al. (2000). "Structural and functional implications of tau hyperphosphorylation: information from phosphorylation-mimicking mutated tau proteins." Biochemistry **39**(43): 13166-13175.
- Eilebrecht, S., C. Smet-Nocca, et al. (2010). "SUMO-1 possesses DNA binding activity." BMC research notes **3**(1): 146.
- Eliezer, D. (2009). "Biophysical characterization of intrinsically disordered proteins." Curr Opin Struct Biol **19**(1): 23-30.
- Faiola, F., X. Liu, et al. (2005). "Dual regulation of c-Myc by p300 via acetylation-dependent control of Myc protein turnover and coactivation of Myc-induced transcription." Mol Cell Biol **25**(23): 10220-10234.
- Fauquant, C., V. Redeker, et al. (2011). "Systematic identification of tubulin-interacting fragments of the microtubule-associated protein Tau leads to a highly efficient promoter of microtubule assembly." The Journal of biological chemistry **286**(38): 33358-33368.
- Fiol, C. J., J. H. Haseman, et al. (1988). "Phosphoserine as a recognition determinant for glycogen synthase kinase-3: phosphorylation of a synthetic peptide based on the G-component of protein phosphatase-1." Arch Biochem Biophys **267**(2): 797-802.
- Fiol, C. J., A. M. Mahrenholz, et al. (1987). "Formation of protein kinase recognition sites by covalent modification of the substrate. Molecular mechanism for the synergistic action of casein kinase II and glycogen synthase kinase 3." The Journal of biological chemistry **262**(29): 14042-14048.
- Fiol, C. J., A. Wang, et al. (1990). "Ordered multisite protein phosphorylation. Analysis of glycogen synthase kinase 3 action using model peptide substrates." The Journal of biological chemistry **265**(11): 6061-6065.
- Fischer, D., M. D. Mukrasch, et al. (2007). "Structural and microtubule binding properties of tau mutants of frontotemporal dementias." Biochemistry **46**(10): 2574-2582.

- Fischer, G., H. Bang, et al. (1984). "[Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides]." Biomed Biochim Acta **43**(10): 1101-1111.
- Fischer, P. M. (2008). "Turning down tau phosphorylation." Nat Chem Biol **4**(8): 448-449.
- Fitzgerald, M. E. and A. C. Drohat (2008). "Coordinating the initial steps of base excision repair. Apurinic/aprimidinic endonuclease 1 actively stimulates thymine DNA glycosylase by disrupting the product complex." The Journal of biological chemistry **283**(47): 32680-32690.
- Fontes, M. R., T. Teh, et al. (2000). "Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha." J Mol Biol **297**(5): 1183-1194.
- Friedhoff, P., A. Schneider, et al. (1998). "Rapid assembly of Alzheimer-like paired helical filaments from microtubule-associated protein tau monitored by fluorescence in solution." Biochemistry **37**(28): 10223-10230.
- Galas, M. C., P. Dourlen, et al. (2006). "The peptidylprolyl cis/trans-isomerase Pin1 modulates stress-induced dephosphorylation of Tau in neurons. Implication in a pathological mechanism related to Alzheimer disease." The Journal of biological chemistry **281**(28): 19296-19304.
- Gamblin, T. C., F. Chen, et al. (2003). "Caspase cleavage of tau: linking amyloid and neurofibrillary tangles in Alzheimer's disease." Proc Natl Acad Sci U S A **100**(17): 10032-10037.
- Gao, Y., L. Wells, et al. (2001). "Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain." J Biol Chem **276**(13): 9838-9845.
- Ge, X., Q. Jin, et al. (2009). "PCAF acetylates {beta}-catenin and improves its stability." Molecular biology of the cell **20**(1): 419-427.
- Gigant, B., P. A. Curmi, et al. (2000). "The 4 A X-ray structure of a tubulin:stathmin-like domain complex." Cell **102**(6): 809-816.
- Giustiniani, J., M. Sineus, et al. (2012). "Decrease of the immunophilin FKBP52 accumulation in human brains of Alzheimer's disease and FTDP-17." Journal of Alzheimer's disease : JAD **29**(2): 471-483.
- Glatz, D. C., D. Rujescu, et al. (2006). "The alternative splicing of tau exon 10 and its regulatory proteins CLK2 and TRA2-BETA1 changes in sporadic Alzheimer's disease." J Neurochem **96**(3): 635-644.
- Glozak, M. A., N. Sengupta, et al. (2005). "Acetylation and deacetylation of non-histone proteins." Gene **363**: 15-23.
- Gnad, F., S. Ren, et al. (2007). "PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites." Genome biology **8**(11): R250.
- Goedert, M., R. Jakes, et al. (1994). "Epitope mapping of monoclonal antibodies to the paired helical filaments of Alzheimer's disease: identification of phosphorylation sites in tau protein." Biochem J **301 ( Pt 3)**: 871-877.
- Goedert, M., R. Jakes, et al. (1995). "Protein phosphatase 2A is the major enzyme in brain that dephosphorylates tau protein phosphorylated by proline-directed protein kinases or cyclic AMP-dependent protein kinase." J Neurochem **65**(6): 2804-2807.
- Goedert, M., R. Jakes, et al. (1995). "Monoclonal antibody AT8 recognises tau protein phosphorylated at both serine 202 and threonine 205." Neurosci Lett **189**(3): 167-169.
- Golovanov, A. P., R. T. Blankley, et al. (2007). "Isotopically discriminated NMR spectroscopy: a tool for investigating complex protein interactions in vitro." J Am Chem Soc **129**(20): 6528-6535.
- Gong, C. X., I. Grundke-Iqbal, et al. (1994). "Dephosphorylation of Alzheimer's disease abnormally phosphorylated tau by protein phosphatase-2A." Neuroscience **61**(4): 765-772.
- Gong, C. X., F. Liu, et al. (2006). "Impaired brain glucose metabolism leads to Alzheimer neurofibrillary degeneration through a decrease in tau O-GlcNAcylation." J Alzheimers Dis **9**(1): 1-12.

- Gong, C. X., S. Shaikh, et al. (1995). "Phosphatase activity toward abnormally phosphorylated tau: decrease in Alzheimer disease brain." J Neurochem **65**(2): 732-738.
- Goode, B. L., P. E. Denis, et al. (1997). "Functional interactions between the proline-rich and repeat regions of tau enhance microtubule binding and assembly." Molecular biology of the cell **8**(2): 353-365.
- Goode, B. L. and S. C. Feinstein (1994). "Identification of a novel microtubule binding and assembly domain in the developmentally regulated inter-repeat region of tau." The Journal of cell biology **124**(5): 769-782.
- Gotz, J., A. Gladbach, et al. (2010). "Animal models reveal role for tau phosphorylation in human disease." Biochimica et biophysica acta **1802**(10): 860-871.
- Gross, B. J., B. C. Kraybill, et al. (2005). "Discovery of O-GlcNAc transferase inhibitors." J Am Chem Soc **127**(42): 14588-14589.
- Grundke-Iqbal, I., K. Iqbal, et al. (1986). "Microtubule-associated protein tau. A component of Alzheimer paired helical filaments." The Journal of biological chemistry **261**(13): 6084-6089.
- Grundke-Iqbal, I., K. Iqbal, et al. (1986). "Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology." Proceedings of the National Academy of Sciences of the United States of America **83**(13): 4913-4917.
- Guinez, C., M. E. Losfeld, et al. (2006). "Modulation of HSP70 GlcNAc-directed lectin activity by glucose availability and utilization." Glycobiology **16**(1): 22-28.
- Guinez, C., A. M. Mir, et al. (2007). "Hsp70-GlcNAc-binding activity is released by stress, proteasome inhibition, and protein misfolding." Biochem Biophys Res Commun **361**(2): 414-420.
- Guo, C., X. Hou, et al. (2009). "Structure-based design of novel human Pin1 inhibitors (I)." Bioorg Med Chem Lett **19**(19): 5613-5616.
- Guo, H., S. Albrecht, et al. (2004). "Active caspase-6 and caspase-6-cleaved tau in neuropil threads, neuritic plaques, and neurofibrillary tangles of Alzheimer's disease." The American journal of pathology **165**(2): 523-531.
- Gustke, N., B. Trinczek, et al. (1994). "Domains of tau protein and interactions with microtubules." Biochemistry **33**(32): 9511-9522.
- Hackenberger, C. P. and D. Schwarzer (2008). "Chemoselective ligation and modification strategies for peptides and proteins." Angewandte Chemie **47**(52): 10030-10074.
- Haltiwanger, R. S., S. Busby, et al. (1997). "O-glycosylation of nuclear and cytoplasmic proteins: regulation analogous to phosphorylation?" Biochem Biophys Res Commun **231**(2): 237-242.
- Hamdane, M., P. Delobel, et al. (2003). "Neurofibrillary degeneration of the Alzheimer-type: an alternate pathway to neuronal apoptosis?" Biochemical pharmacology **66**(8): 1619-1625.
- Hamdane, M., P. Dourlen, et al. (2006). "Pin1 allows for differential Tau dephosphorylation in neuronal cells." Molecular and cellular neurosciences **32**(1-2): 155-160.
- Hamdane, M., A. V. Sambo, et al. (2003). "Mitotic-like tau phosphorylation by p25-Cdk5 kinase complex." J Biol Chem **278**(36): 34026-34034.
- Hamdane, M., C. Smet, et al. (2002). "Pin1: a therapeutic target in Alzheimer neurodegeneration." Journal of molecular neuroscience : MN **19**(3): 275-287.
- Hani, J., G. Stumpf, et al. (1995). "PTF1 encodes an essential protein in *Saccharomyces cerevisiae*, which shows strong homology with a new putative family of PPLases." FEBS Lett **365**(2-3): 198-202.
- Harbison, N. W., S. Bhattacharya, et al. (2012). "Assigning backbone NMR resonances for full length tau isoforms: efficient compromise between manual assignments and reduced dimensionality." PloS one **7**(4): e34679.
- Hardeland, U., R. Steinacher, et al. (2002). "Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover." The EMBO journal **21**(6): 1456-1464.



- Hart, G. W., K. D. Greis, et al. (1995). "O-linked N-acetylglucosamine: the "yin-yang" of Ser/Thr phosphorylation? Nuclear and cytoplasmic glycosylation." Advances in experimental medicine and biology **376**: 115-123.
- Hart, G. W., M. P. Housley, et al. (2007). "Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins." Nature **446**(7139): 1017-1022.
- Hart, G. W. and K. Sakabe (2006). "Fine-tuning ER-beta structure with PTMs." Chem Biol **13**(9): 923-924.
- Hashiguchi, M., K. Sobue, et al. (2000). "14-3-3zeta is an effector of tau protein phosphorylation." The Journal of biological chemistry **275**(33): 25247-25254.
- Hatakeyama, S., M. Matsumoto, et al. (2004). "U-box protein carboxyl terminus of Hsc70-interacting protein (CHIP) mediates poly-ubiquitylation preferentially on four-repeat Tau and is involved in neurodegeneration of tauopathy." J Neurochem **91**(2): 299-307.
- Hirokawa, N., Y. Shiomura, et al. (1988). "Tau proteins: the molecular structure and mode of binding on microtubules." The Journal of cell biology **107**(4): 1449-1459.
- Hoffmann, R., V. M. Lee, et al. (1997). "Unique Alzheimer's disease paired helical filament specific epitopes involve double phosphorylation at specific sites." Biochemistry **36**(26): 8114-8124.
- Hong, M., V. Zhukareva, et al. (1998). "Mutation-specific functional impairments in distinct tau isoforms of hereditary FTDP-17." Science **282**(5395): 1914-1917.
- Hu, P., S. Shimoji, et al. (2010). "Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation." FEBS letters **584**(12): 2526-2538.
- Hubbert, C., A. Guardiola, et al. (2002). "HDAC6 is a microtubule-associated deacetylase." Nature **417**(6887): 455-458.
- Huber, A. H. and W. I. Weis (2001). "The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin." Cell **105**(3): 391-402.
- Hurtado, D. E., L. Molina-Porcel, et al. (2010). "A{beta} accelerates the spatiotemporal progression of tau pathology and augments tau amyloidosis in an Alzheimer mouse model." The American journal of pathology **177**(4): 1977-1988.
- Hutton, M., C. L. Lendon, et al. (1998). "Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17." Nature **393**(6686): 702-705.
- Iakoucheva, L. M., P. Radivojac, et al. (2004). "The importance of intrinsic disorder for protein phosphorylation." Nucleic Acids Res **32**(3): 1037-1049.
- Iqbal, K., C. Alonso Adel, et al. (2005). "Tau pathology in Alzheimer disease and other tauopathies." Biochimica et biophysica acta **1739**(2-3): 198-210.
- Iqbal, K., F. Liu, et al. (2009). "Mechanisms of tau-induced neurodegeneration." Acta neuropathologica **118**(1): 53-69.
- Irwin, D. J., T. J. Cohen, et al. (2012). "Acetylated tau, a novel pathological signature in Alzheimer's disease and other tauopathies." Brain : a journal of neurology **135**(Pt 3): 807-818.
- Ittner, L. M. and J. Gotz (2011). "Amyloid-beta and tau--a toxic pas de deux in Alzheimer's disease." Nature reviews. Neuroscience **12**(2): 65-72.
- Ittner, L. M., Y. D. Ke, et al. (2010). "Dendritic function of tau mediates amyloid-beta toxicity in Alzheimer's disease mouse models." Cell **142**(3): 387-397.
- Iyer, S. P., Y. Akimoto, et al. (2003). "Identification and cloning of a novel family of coiled-coil domain proteins that interact with O-GlcNAc transferase." J Biol Chem **278**(7): 5399-5409.
- Iyer, S. P. and G. W. Hart (2003). "Dynamic nuclear and cytoplasmic glycosylation: enzymes of O-GlcNAc cycling." Biochemistry **42**(9): 2493-2499.
- Iyer, S. P. and G. W. Hart (2003). "Roles of the tetratricopeptide repeat domain in O-GlcNAc transferase targeting and protein substrate specificity." J Biol Chem **278**(27): 24608-24616.
- Janssens, V. and J. Goris (2001). "Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling." Biochem J **353**(Pt 3): 417-439.

- Janssens, V., S. Longin, et al. (2008). "PP2A holoenzyme assembly: in cauda venenum (the sting is in the tail)." *Trends Biochem Sci* **33**(3): 113-121.
- Jeffrey, P. D., A. A. Russo, et al. (1995). "Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex." *Nature* **376**(6538): 313-320.
- Jeganathan, S., A. Hascher, et al. (2008). "Proline-directed pseudo-phosphorylation at AT8 and PHF1 epitopes induces a compaction of the paperclip folding of Tau and generates a pathological (MC-1) conformation." *The Journal of biological chemistry* **283**(46): 32066-32076.
- Jeganathan, S., M. von Bergen, et al. (2006). "Global hairpin folding of tau in solution." *Biochemistry* **45**(7): 2283-2293.
- Jicha, G. A., B. Berenfeld, et al. (1999). "Sequence requirements for formation of conformational variants of tau similar to those found in Alzheimer's disease." *J Neurosci Res* **55**(6): 713-723.
- Jicha, G. A., R. Bowser, et al. (1997). "Alz-50 and MC-1, a new monoclonal antibody raised to paired helical filaments, recognize conformational epitopes on recombinant tau." *J Neurosci Res* **48**(2): 128-132.
- Jicha, G. A., C. Weaver, et al. (1999). "cAMP-dependent protein kinase phosphorylations on tau in Alzheimer's disease." *J Neurosci* **19**(17): 7486-7494.
- Jinek, M., J. Rehwinkel, et al. (2004). "The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha." *Nat Struct Mol Biol* **11**(10): 1001-1007.
- Jinwal, U. K., J. Koren, 3rd, et al. (2010). "The Hsp90 cochaperone, FKBP51, increases Tau stability and polymerizes microtubules." *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30**(2): 591-599.
- Johnson, G. V., P. Seubert, et al. (1997). "The tau protein in human cerebrospinal fluid in Alzheimer's disease consists of proteolytically derived fragments." *J Neurochem* **68**(1): 430-433.
- Johnson, G. V. and W. H. Stoothoff (2004). "Tau phosphorylation in neuronal cell function and dysfunction." *Journal of cell science* **117**(Pt 24): 5721-5729.
- Joseph, J. D., E. S. Yeh, et al. (2003). "The peptidyl-prolyl isomerase Pin1." *Prog Cell Cycle Res* **5**: 477-487.
- Jourdain, L., P. Curmi, et al. (1997). "Stathmin: a tubulin-sequestering protein which forms a ternary T2S complex with two tubulin molecules." *Biochemistry* **36**(36): 10817-10821.
- Kamemura, K. and G. W. Hart (2003). "Dynamic interplay between O-glycosylation and O-phosphorylation of nucleocytoplasmic proteins: a new paradigm for metabolic control of signal transduction and transcription." *Prog Nucleic Acid Res Mol Biol* **73**: 107-136.
- Kay, L. E. (1997). "NMR methods for the study of protein structure and dynamics." *Biochem Cell Biol* **75**(1): 1-15.
- Kay, L. E. (1998). "Protein dynamics from NMR." *Nature structural biology* **5 Suppl**: 513-517.
- Kay, L. E. (2005). "NMR studies of protein structure and dynamics." *Journal of magnetic resonance* **173**(2): 193-207.
- Kayed, R. and G. R. Jackson (2009). "Prefilament tau species as potential targets for immunotherapy for Alzheimer disease and related disorders." *Current opinion in immunology* **21**(3): 359-363.
- Kidd, M. (1963). "Paired helical filaments in electron microscopy of Alzheimer's disease." *Nature* **197**: 192-193.
- Kidd, M. (1964). "Alzheimer's Disease--an Electron Microscopical Study." *Brain : a journal of neurology* **87**: 307-320.
- Kim, G. W. and X. J. Yang (2011). "Comprehensive lysine acetylomes emerging from bacteria to humans." *Trends in biochemical sciences* **36**(4): 211-220.
- Kim, S. C., R. Sprung, et al. (2006). "Substrate and functional diversity of lysine acetylation revealed by a proteomics survey." *Molecular cell* **23**(4): 607-618.
- Kimura, T., K. Tsutsumi, et al. (2013). "Pin1 Stimulates Dephosphorylation of Tau at Cdk5-Dependent Alzheimer Phosphorylation Sites." *The Journal of biological chemistry*.

- Kins, S., A. Crameri, et al. (2001). "Reduced protein phosphatase 2A activity induces hyperphosphorylation and altered compartmentalization of tau in transgenic mice." J Biol Chem **276**(41): 38193-38200.
- Klaiman, G., N. Champagne, et al. (2009). "Self-activation of Caspase-6 in vitro and in vivo: Caspase-6 activation does not induce cell death in HEK293T cells." Biochimica et biophysica acta **1793**(3): 592-601.
- Klaiman, G., T. L. Petzke, et al. (2008). "Targets of caspase-6 activity in human neurons and Alzheimer disease." Molecular & cellular proteomics : MCP **7**(8): 1541-1555.
- Kohler, J. J. (2010). "Post-translational modifications: A shift for the O-GlcNAc paradigm." Nature chemical biology **6**(9): 634-635.
- Kosik, K. S., C. L. Joachim, et al. (1986). "Microtubule-associated protein tau (tau) is a major antigenic component of paired helical filaments in Alzheimer disease." Proc Natl Acad Sci U S A **83**(11): 4044-4048.
- Kreppel, L. K., M. A. Blomberg, et al. (1997). "Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats." The Journal of biological chemistry **272**(14): 9308-9315.
- Kreppel, L. K. and G. W. Hart (1999). "Regulation of a cytosolic and nuclear O-GlcNAc transferase. Role of the tetratricopeptide repeats." J Biol Chem **274**(45): 32015-32022.
- Landau, M., M. R. Sawaya, et al. (2011). "Towards a pharmacophore for amyloid." PLoS biology **9**(6): e1001080.
- Landrieu, I., L. Lacosse, et al. (2006). "NMR analysis of a Tau phosphorylation pattern." J Am Chem Soc **128**(11): 3575-3583.
- Landrieu, I., A. Leroy, et al. (2010). "NMR spectroscopy of the neuronal tau protein: normal function and implication in Alzheimer's disease." Biochemical Society Transactions **38**(4): 1006-1011.
- Landrieu, I., B. Odaert, et al. (2001). "p13(SUC1) and the WW domain of PIN1 bind to the same phosphothreonine-proline epitope." J Biol Chem **276**(2): 1434-1438.
- Landrieu, I., C. Smet-Nocca, et al. (2011). "Molecular implication of PP2A and Pin1 in the Alzheimer's disease specific hyperphosphorylation of Tau." PLoS one **6**(6): e21521.
- Lasagna-Reeves, C. A., D. L. Castillo-Carranza, et al. (2012). "Alzheimer brain-derived tau oligomers propagate pathology from endogenous tau." Scientific reports **2**: 700.
- Lasagna-Reeves, C. A., D. L. Castillo-Carranza, et al. (2012). "Identification of oligomers at early stages of tau aggregation in Alzheimer's disease." FASEB journal : official publication of the Federation of American Societies for Experimental Biology **26**(5): 1946-1959.
- Lazarus, M. B., J. Jiang, et al. (2012). "Structural snapshots of the reaction coordinate for O-GlcNAc transferase." Nature chemical biology **8**(12): 966-968.
- Lazarus, M. B., Y. Nam, et al. (2011). "Structure of human O-GlcNAc transferase and its complex with a peptide substrate." Nature **469**(7331): 564-567.
- Lefebvre, T., S. Ferreira, et al. (2003). "Evidence of a balance between phosphorylation and O-GlcNAc glycosylation of Tau proteins--a role in nuclear localization." Biochim Biophys Acta **1619**(2): 167-176.
- Lefebvre, T., C. Guinez, et al. (2005). "Does O-GlcNAc play a role in neurodegenerative diseases?" Expert Rev Proteomics **2**(2): 265-275.
- Léger, H., C. Smet-Nocca, et al. (en préparation). "A TDG/CBP/RARalpha tertiary complex mediates the retinoic acid-dependent expression of de novo methylation-sensitive genes." Nucleic Acids Res.
- Leroy, A., I. Landrieu, et al. (2010). "Spectroscopic Studies of GSK3 $\beta$  Phosphorylation of the Neuronal Tau Protein and Its Interaction with the N-terminal Domain of Apolipoprotein E." Journal of Biological Chemistry **285**(43): 33435-33444.
- Lewis, J., E. McGowan, et al. (2000). "Neurofibrillary tangles, amyotrophy and progressive motor disturbance in mice expressing mutant (P301L) tau protein." Nature genetics **25**(4): 402-405.

- Li, M., J. Luo, et al. (2002). "Acetylation of p53 inhibits its ubiquitination by Mdm2." The Journal of biological chemistry **277**(52): 50607-50611.
- Li, M., P. T. Stukenberg, et al. (2008). "Binding of phosphatase inhibitor-2 to prolyl isomerase Pin1 modifies specificity for mitotic phosphoproteins." Biochemistry **47**(1): 292-300.
- Lim, J., M. Balastik, et al. (2008). "Pin1 has opposite effects on wild-type and P301L tau stability and tauopathy." J Clin Invest **118**(5): 1877-1889.
- Lim, J. and K. P. Lu (2005). "Pinning down phosphorylated tau and tauopathies." Biochimica et biophysica acta **1739**(2-3): 311-322.
- Liokatis, S., A. Dose, et al. (2010). "Simultaneous detection of protein phosphorylation and acetylation by high-resolution NMR spectroscopy." J Am Chem Soc **132**(42): 14704-14705.
- Liou, Y. C., A. Sun, et al. (2003). "Role of the prolyl isomerase Pin1 in protecting against age-dependent neurodegeneration." Nature **424**(6948): 556-561.
- Liou, Y. C., X. Z. Zhou, et al. (2011). "Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins." Trends Biochem Sci **36**(10): 501-514.
- Lippens, G., L. Caron, et al. (2004). "A microscopic view of chemical exchange : Monte Carlo simulation of molecular association." Concepts Magn. Res. Part A **21A**(1): 1-9.
- Lippens, G., I. Landrieu, et al. (2008). "Studying posttranslational modifications by in-cell NMR." Chem Biol **15**(4): 311-312.
- Lippens, G., I. Landrieu, et al. (2007). "Molecular mechanisms of the phospho-dependent prolyl cis/trans isomerase Pin1." The FEBS journal **274**(20): 5211-5222.
- Lippens, G., A. Sillen, et al. (2006). "Studying the natively unfolded neuronal Tau protein by solution NMR spectroscopy." Protein and peptide letters **13**(3): 235-246.
- Lippens, G., J. M. Wieruszkeski, et al. (2004). "Proline-directed random-coil chemical shift values as a tool for the NMR assignment of the tau phosphorylation sites." Chembiochem : a European journal of chemical biology **5**(1): 73-78.
- Liu, F., I. Grundke-Iqbal, et al. (2005). "Contributions of protein phosphatases PP1, PP2A, PP2B and PP5 to the regulation of tau phosphorylation." Eur J Neurosci **22**(8): 1942-1950.
- Liu, F., K. Iqbal, et al. (2002). "Involvement of aberrant glycosylation in phosphorylation of tau by cdk5 and GSK-3beta." FEBS letters **530**(1-3): 209-214.
- Liu, F., K. Iqbal, et al. (2004). "O-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease." Proc Natl Acad Sci U S A **101**(29): 10804-10809.
- Liu, F., J. Shi, et al. (2009). "Reduced O-GlcNAcylation links lower brain glucose metabolism and tau pathology in Alzheimer's disease." Brain **132**(Pt 7): 1820-1832.
- Liu, F., T. Zaidi, et al. (2002). "Aberrant glycosylation modulates phosphorylation of tau by protein kinase A and dephosphorylation of tau by protein phosphatase 2A and 5." Neuroscience **115**(3): 829-837.
- Liu, F., T. Zaidi, et al. (2002). "Role of glycosylation in hyperphosphorylation of tau in Alzheimer's disease." FEBS letters **512**(1-3): 101-106.
- Liu, S. J., J. Y. Zhang, et al. (2004). "Tau becomes a more favorable substrate for GSK-3 when it is prephosphorylated by PKA in rat brain." The Journal of biological chemistry **279**(48): 50078-50088.
- Liu, T. and G. Bitan (2012). "Modulating self-assembly of amyloidogenic proteins as a therapeutic approach for neurodegenerative diseases: strategies and mechanisms." ChemMedChem **7**(3): 359-374.
- Loomis, P. A., T. H. Howard, et al. (1990). "Identification of nuclear tau isoforms in human neuroblastoma cells." Proceedings of the National Academy of Sciences of the United States of America **87**(21): 8422-8426.
- Love, D. C. and J. A. Hanover (2005). "The hexosamine signaling pathway: deciphering the "O-GlcNAc code"." Sci STKE **2005**(312): re13.

- Lu, K. P. (2000). "Phosphorylation-dependent prolyl isomerization: a novel cell cycle regulatory mechanism." *Prog Cell Cycle Res* **4**: 83-96.
- Lu, K. P. (2003). "Prolyl isomerase Pin1 as a molecular target for cancer diagnostics and therapeutics." *Cancer Cell* **4**(3): 175-180.
- Lu, K. P. (2004). "Pinning down cell signaling, cancer and Alzheimer's disease." *Trends Biochem Sci* **29**(4): 200-209.
- Lu, K. P. (2004). "Pinning down phosphorylated Tau and tauopathies." *Biochim Biophys Acta*.
- Lu, K. P., G. Finn, et al. (2007). "Prolyl cis-trans isomerization as a molecular timer." *Nat Chem Biol* **3**(10): 619-629.
- Lu, K. P., S. D. Hanes, et al. (1996). "A human peptidyl-prolyl isomerase essential for regulation of mitosis." *Nature* **380**(6574): 544-547.
- Lu, K. P. and T. Hunter (1995). "The NIMA kinase: a mitotic regulator in *Aspergillus nidulans* and vertebrate cells." *Prog Cell Cycle Res* **1**: 187-205.
- Lu, K. P., Y. C. Liou, et al. (2003). "Proline-directed phosphorylation and isomerization in mitotic regulation and in Alzheimer's Disease." *Bioessays* **25**(2): 174-181.
- Lu, K. P., Y. C. Liou, et al. (2002). "Pinning down proline-directed phosphorylation signaling." *Trends Cell Biol* **12**(4): 164-172.
- Lu, K. P., F. Suizu, et al. (2006). "Targeting carcinogenesis: a role for the prolyl isomerase Pin1?" *Mol Carcinog* **45**(6): 397-402.
- Lu, K. P. and X. Z. Zhou (2007). "The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease." *Nat Rev Mol Cell Biol* **8**(11): 904-916.
- Lu, M. and K. S. Kosik (2001). "Competition for microtubule-binding with dual expression of tau missense and splice isoforms." *Molecular biology of the cell* **12**(1): 171-184.
- Lu, P. J., G. Wulf, et al. (1999). "The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein." *Nature* **399**(6738): 784-788.
- Lu, P. J., X. Z. Zhou, et al. (2002). "Critical role of WW domain phosphorylation in regulating phosphoserine binding activity and Pin1 function." *J Biol Chem* **277**(4): 2381-2384.
- Lu, P. J., X. Z. Zhou, et al. (1999). "Function of WW domains as phosphoserine- or phosphothreonine-binding modules." *Science* **283**(5406): 1325-1328.
- Lubas, W. A., D. W. Frank, et al. (1997). "O-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats." *The Journal of biological chemistry* **272**(14): 9316-9324.
- Lubas, W. A. and J. A. Hanover (2000). "Functional expression of O-linked GlcNAc transferase. Domain structure and substrate specificity." *J Biol Chem* **275**(15): 10983-10988.
- Macauley, M. S. and D. J. Vocadlo (2009). "Increasing O-GlcNAc levels: An overview of small-molecule inhibitors of O-GlcNAcase." *Biochim Biophys Acta*.
- Malik, R., E. A. Nigg, et al. (2008). "Comparative conservation analysis of the human mitotic phosphoproteome." *Bioinformatics* **24**(12): 1426-1432.
- Marsh, J. A., V. K. Singh, et al. (2006). "Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation." *Protein science : a publication of the Protein Society* **15**(12): 2795-2804.
- Martens, E., I. Stevens, et al. (2004). "Genomic organisation, chromosomal localisation tissue distribution and developmental regulation of the PR61/B' regulatory subunits of protein phosphatase 2A in mice." *J Mol Biol* **336**(4): 971-986.
- Martinez-Fleites, C., M. S. Macauley, et al. (2008). "Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation." *Nat Struct Mol Biol* **15**(7): 764-765.
- Mattson, M. P. (2010). "Acetylation unleashes protein demons of dementia." *Neuron* **67**(6): 900-902.
- Meier-Ruge, W. and C. Bertoni-Freddari (1996). "The significance of glucose turnover in the brain in the pathogenetic mechanisms of Alzheimer's disease." *Reviews in the neurosciences* **7**(1): 1-19.

- Merrick, S. E., J. Q. Trojanowski, et al. (1997). "Selective destruction of stable microtubules and axons by inhibitors of protein serine/threonine phosphatases in cultured human neurons." The Journal of neuroscience : the official journal of the Society for Neuroscience **17**(15): 5726-5737.
- Min, S. W., S. H. Cho, et al. (2010). "Acetylation of tau inhibits its degradation and contributes to tauopathy." Neuron **67**(6): 953-966.
- Monteleone, G., G. Del Vecchio Blanco, et al. (2005). "Post-transcriptional regulation of Smad7 in the gut of patients with inflammatory bowel disease." Gastroenterology **129**(5): 1420-1429.
- Moore, C. L., M. H. Huang, et al. (2011). "Secondary nucleating sequences affect kinetics and thermodynamics of tau aggregation." Biochemistry **50**(50): 10876-10886.
- Morishima-Kawashima, M., M. Hasegawa, et al. (1993). "Ubiquitin is conjugated with amino-terminally processed tau in paired helical filaments." Neuron **10**(6): 1151-1160.
- Morishima-Kawashima, M., M. Hasegawa, et al. (1995). "Proline-directed and non-proline-directed phosphorylation of PHF-tau." J Biol Chem **270**(2): 823-829.
- Morishima-Kawashima, M., M. Hasegawa, et al. (1995). "Hyperphosphorylation of tau in PHF." Neurobiol Aging **16**(3): 365-371; discussion 371-380.
- Mukrasch, M. D., S. Bibow, et al. (2009). "Structural polymorphism of 441-residue tau at single residue resolution." PLoS biology **7**(2): e34.
- Mukrasch, M. D., J. Biernat, et al. (2005). "Sites of tau important for aggregation populate {beta}-structure and bind to microtubules and polyanions." The Journal of biological chemistry **280**(26): 24978-24986.
- Mukrasch, M. D., P. Markwick, et al. (2007). "Highly populated turn conformations in natively unfolded tau protein identified from residual dipolar couplings and molecular simulation." J Am Chem Soc **129**(16): 5235-5243.
- Mukrasch, M. D., M. von Bergen, et al. (2007). "The "jaws" of the tau-microtubule interaction." The Journal of biological chemistry **282**(16): 12230-12239.
- Nacharaju, P., J. Lewis, et al. (1999). "Accelerated filament formation from tau protein with specific FTDP-17 missense mutations." FEBS letters **447**(2-3): 195-199.
- Namanja, A. T., X. J. Wang, et al. (2010). "Toward flexibility-activity relationships by NMR spectroscopy: dynamics of Pin1 ligands." J Am Chem Soc **132**(16): 5607-5609.
- Narayanan, R. L., U. H. Durr, et al. (2010). "Automatic assignment of the intrinsically disordered protein Tau with 441-residues." J Am Chem Soc **132**(34): 11906-11907.
- Nikolaev, A., T. McLaughlin, et al. (2009). "APP binds DR6 to trigger axon pruning and neuron death via distinct caspases." Nature **457**(7232): 981-989.
- Noble, W., V. Olm, et al. (2003). "Cdk5 is a key factor in tau aggregation and tangle formation in vivo." Neuron **38**(4): 555-565.
- Noble, W., E. Planel, et al. (2005). "Inhibition of glycogen synthase kinase-3 by lithium correlates with reduced tauopathy and degeneration in vivo." Proceedings of the National Academy of Sciences of the United States of America **102**(19): 6990-6995.
- Nogales, E., S. G. Wolf, et al. (1998). "Structure of the alpha beta tubulin dimer by electron crystallography." Nature **391**(6663): 199-203.
- Otvos, L., Jr., L. Feiner, et al. (1994). "Monoclonal antibody PHF-1 recognizes tau protein phosphorylated at serine residues 396 and 404." J Neurosci Res **39**(6): 669-673.
- Park, S. Y. and A. Ferreira (2005). "The generation of a 17 kDa neurotoxic fragment: an alternative mechanism by which tau mediates beta-amyloid-induced neurodegeneration." The Journal of neuroscience : the official journal of the Society for Neuroscience **25**(22): 5365-5375.
- Pastorino, L., A. Sun, et al. (2006). "The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-beta production." Nature **440**(7083): 528-534.

- Patterson, K. R., S. M. Ward, et al. (2011). "Heat shock protein 70 prevents both tau aggregation and the inhibitory effects of preexisting tau aggregates on fast axonal transport." Biochemistry **50**(47): 10300-10310.
- Perez, M., I. Santa-Maria, et al. (2009). "Tau--an inhibitor of deacetylase HDAC6 function." J Neurochem **109**(6): 1756-1766.
- Petrucci, L., D. Dickson, et al. (2004). "CHIP and Hsp70 regulate tau ubiquitination, degradation and aggregation." Hum Mol Genet **13**(7): 703-714.
- Pevalova, M., P. Filipcik, et al. (2006). "Post-translational modifications of tau protein." Bratislavské lekarske listy **107**(9-10): 346-353.
- Piperno, G., M. LeDizet, et al. (1987). "Microtubules containing acetylated alpha-tubulin in mammalian cells in culture." The Journal of cell biology **104**(2): 289-302.
- Polevoda, B. and F. Sherman (2002). "The diversity of acetylated proteins." Genome biology **3**(5): reviews0006.
- Prabakaran, S., R. A. Everley, et al. (2011). "Comparative analysis of Erk phosphorylation suggests a mixed strategy for measuring phospho-form distributions." Molecular systems biology **7**: 482.
- Preuss, U. and E. M. Mandelkow (1998). "Mitotic phosphorylation of tau protein in neuronal cell lines resembles phosphorylation in Alzheimer's disease." Eur J Cell Biol **76**(3): 176-184.
- Ramakrishnan, P., D. W. Dickson, et al. (2003). "Pin1 colocalization with phosphorylated tau in Alzheimer's disease and other tauopathies." Neurobiol Dis **14**(2): 251-264.
- Ramelot, T. A. and L. K. Nicholson (2001). "Phosphorylation-induced structural changes in the amyloid precursor protein cytoplasmic tail detected by NMR." J Mol Biol **307**(3): 871-884.
- Ranganathan, R., K. P. Lu, et al. (1997). "Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent." Cell **89**(6): 875-886.
- Rapoport, M., H. N. Dawson, et al. (2002). "Tau is essential to beta -amyloid-induced neurotoxicity." Proceedings of the National Academy of Sciences of the United States of America **99**(9): 6364-6369.
- Reason, A. J., I. P. Blench, et al. (1991). "High-sensitivity FAB-MS strategies for O-GlcNAc characterization." Glycobiology **1**(6): 585-594.
- Rexach, J. E., P. M. Clark, et al. (2008). "Chemical approaches to understanding O-GlcNAc glycosylation in the brain." Nat Chem Biol **4**(2): 97-106.
- Rexach, J. E., C. J. Rogers, et al. (2010). "Quantification of O-glycosylation stoichiometry and dynamics using resolvable mass tags." Nature chemical biology **6**(9): 645-651.
- Ricobaraza, A., M. Cuadrado-Tejedor, et al. (2009). "Phenylbutyrate ameliorates cognitive deficit and reduces tau pathology in an Alzheimer's disease mouse model." Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology **34**(7): 1721-1732.
- Rissman, R. A., W. W. Poon, et al. (2004). "Caspase-cleavage of tau is an early event in Alzheimer disease tangle pathology." J Clin Invest **114**(1): 121-130.
- Roberson, E. D., K. Scarce-Levie, et al. (2007). "Reducing endogenous tau ameliorates amyloid beta-induced deficits in an Alzheimer's disease mouse model." Science **316**(5825): 750-754.
- Robertson, L. A., K. L. Moya, et al. (2004). "The potential role of tau protein O-glycosylation in Alzheimer's disease." J Alzheimers Dis **6**(5): 489-495.
- Rudrabhatla, P. and H. C. Pant (2011). "Role of protein phosphatase 2A in Alzheimer's disease." Current Alzheimer research **8**(6): 623-632.
- Ryo, A., M. Nakamura, et al. (2001). "Pin1 regulates turnover and subcellular localization of beta-catenin by inhibiting its interaction with APC." Nat Cell Biol **3**(9): 793-801.
- Salminen, A., J. Ojala, et al. (2011). "Hsp90 regulates tau pathology through co-chaperone complexes in Alzheimer's disease." Progress in neurobiology **93**(1): 99-110.

- Saman, S., W. Kim, et al. (2012). "Exosome-associated tau is secreted in tauopathy models and is selectively phosphorylated in cerebrospinal fluid in early Alzheimer disease." The Journal of biological chemistry **287**(6): 3842-3849.
- Sami, F., C. Smet-Nocca, et al. (2011). "Molecular basis for an ancient partnership between prolyl isomerase Pin1 and phosphatase inhibitor-2." Biochemistry **50**(30): 6567-6578.
- Santacruz, K., J. Lewis, et al. (2005). "Tau suppression in a neurodegenerative mouse model improves memory function." Science **309**(5733): 476-481.
- Sarkar, M., J. Kuret, et al. (2008). "Two motifs within the tau microtubule-binding domain mediate its association with the hsc70 molecular chaperone." J Neurosci Res **86**(12): 2763-2773.
- Sato, Y., Y. Naito, et al. (2001). "Analysis of N-glycans of pathological tau: possible occurrence of aberrant processing of tau in Alzheimer's disease." FEBS letters **496**(2-3): 152-160.
- Schimpl, M., X. Zheng, et al. (2012). "O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis." Nature chemical biology **8**(12): 969-974.
- Schindowski, K., K. Belarbi, et al. (2008). "Neurogenesis and cell cycle-reactivated neuronal death during pathogenic tau aggregation." Genes, brain, and behavior **7 Suppl 1**: 92-100.
- Schmidt, K., S. Kins, et al. (2002). "Diversity, developmental regulation and distribution of murine PR55/B subunits of protein phosphatase 2A." Eur J Neurosci **16**(11): 2039-2048.
- Schneider, A., J. Biernat, et al. (1999). "Phosphorylation that detaches tau protein from microtubules (Ser262, Ser214) also protects it against aggregation into Alzheimer paired helical filaments." Biochemistry **38**(12): 3549-3558.
- Schulze, E., D. J. Asai, et al. (1987). "Posttranslational modification and microtubule stability." The Journal of cell biology **105**(5): 2167-2177.
- Schulze, E. and M. Kirschner (1987). "Dynamic and stable populations of microtubules in cells." The Journal of cell biology **104**(2): 277-288.
- Schweers, O., E. M. Mandelkow, et al. (1995). "Oxidation of cysteine-322 in the repeat domain of microtubule-associated protein tau controls the in vitro assembly of paired helical filaments." Proceedings of the National Academy of Sciences of the United States of America **92**(18): 8463-8467.
- Selenko, P., D. P. Frueh, et al. (2008). "In situ observation of protein phosphorylation by high-resolution NMR spectroscopy." Nat Struct Mol Biol **15**(3): 321-329.
- Sergeant, N., A. Bretteville, et al. (2008). "Biochemistry of Tau in Alzheimer's disease and related neurological disorders." Expert Rev Proteomics **5**(2): 207-224.
- Shafi, R., S. P. Iyer, et al. (2000). "The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny." Proceedings of the National Academy of Sciences of the United States of America **97**(11): 5735-5739.
- Shen, M., P. T. Stukenberg, et al. (1998). "The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins." Genes Dev **12**(5): 706-720.
- Sibille, N., I. Huvent, et al. (2011). "Structural characterization by nuclear magnetic resonance of the impact of phosphorylation in the proline-rich region of the disordered Tau protein." Proteins.
- Sibille, N., A. Sillen, et al. (2006). "Structural impact of heparin binding to full-length Tau as studied by NMR spectroscopy." Biochemistry **45**(41): 12560-12572.
- Sillen, A., P. Barbier, et al. (2007). "NMR investigation of the interaction between the neuronal protein tau and the microtubules." Biochemistry **46**(11): 3055-3064.
- Sillen, A., A. Leroy, et al. (2005). "Regions of tau implicated in the paired helical fragment core as defined by NMR." Chembiochem : a European journal of chemical biology **6**(10): 1849-1856.
- Sillen, A., J. M. Wieruszski, et al. (2005). "High-resolution magic angle spinning NMR of the neuronal tau protein integrated in Alzheimer's-like paired helical fragments." J Am Chem Soc **127**(29): 10138-10139.
- Sinha, S., D. H. Lopes, et al. (2012). "A Key Role for Lysine Residues in Amyloid beta-Protein Folding, Assembly, and Toxicity." ACS chemical neuroscience **3**(6): 473-481.



- Sinha, S., D. H. Lopes, et al. (2011). "Lysine-specific molecular tweezers are broad-spectrum inhibitors of assembly and toxicity of amyloid proteins." *J Am Chem Soc* **133**(42): 16958-16969.
- Sjogren, M., L. Rosengren, et al. (2000). "Cytoskeleton proteins in CSF distinguish frontotemporal dementia from AD." *Neurology* **54**(10): 1960-1964.
- Slawson, C., R. J. Copeland, et al. (2010). "O-GlcNAc signaling: a metabolic link between diabetes and cancer?" *Trends in biochemical sciences* **35**(10): 547-555.
- Slawson, C., N. E. Zachara, et al. (2005). "Perturbations in O-linked beta-N-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis." *J Biol Chem* **280**(38): 32944-32956.
- Smet-NoCCA, C., M. Broncel, et al. (2011). "Identification of O-GlcNAc sites within peptides of the Tau protein and their impact on phosphorylation." *Molecular bioSystems* **7**(5): 1420-1429.
- Smet-NoCCA, C., H. Launay, et al. (2013). "Unraveling a phosphorylation event in a folded protein by NMR spectroscopy: phosphorylation of the Pin1 WW domain by PKA." *Journal of biomolecular NMR* **55**(4): 323-337.
- Smet-NoCCA, C., J. M. Wieruszeski, et al. (2008). "The thymine-DNA glycosylase regulatory domain: residual structure and DNA binding." *Biochemistry* **47**(25): 6519-6530.
- Smet-NoCCA, C., J. M. Wieruszeski, et al. (2011). "SUMO-1 regulates the conformational dynamics of thymine-DNA Glycosylase regulatory domain and competes with its DNA binding activity." *BMC biochemistry* **12**: 4.
- Smet-NoCCA, C., J. M. Wieruszeski, et al. (2010). "NMR-based detection of acetylation sites in peptides." *Journal of peptide science* **16**(8): 414-423.
- Smet, C., J. F. Duckert, et al. (2005). "Control of protein-protein interactions: structure-based discovery of low molecular weight inhibitors of the interactions between Pin1 WW domain and phosphopeptides." *Journal of medicinal chemistry* **48**(15): 4815-4823.
- Smet, C., A. Leroy, et al. (2004). "Accepting its random coil nature allows a partial NMR assignment of the neuronal Tau protein." *Chembiochem : a European journal of chemical biology* **5**(12): 1639-1646.
- Smet, C., A. V. Sambo, et al. (2004). "The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau." *Biochemistry* **43**(7): 2032-2040.
- Smet, C., J. M. Wieruszeski, et al. (2005). "Regulation of Pin1 peptidyl-prolyl cis/trans isomerase activity by its WW binding module on a multi-phosphorylated peptide of Tau protein." *FEBS letters* **579**(19): 4159-4164.
- Sontag, E., C. Hladik, et al. (2004). "Downregulation of protein phosphatase 2A carboxyl methylation and methyltransferase may contribute to Alzheimer disease pathogenesis." *J Neuropathol Exp Neurol* **63**(10): 1080-1091.
- Sontag, E., A. Luangpirom, et al. (2004). "Altered expression levels of the protein phosphatase 2A A $\beta$ Alphac enzyme are associated with Alzheimer disease pathology." *J Neuropathol Exp Neurol* **63**(4): 287-301.
- Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." *Nature* **403**(6765): 41-45.
- Su, J. H., G. Deng, et al. (1997). "Transneuronal degeneration in the spread of Alzheimer's disease pathology: immunohistochemical evidence for the transmission of tau hyperphosphorylation." *Neurobiology of disease* **4**(5): 365-375.
- Sultan, A., F. Nessler, et al. (2011). "Nuclear tau, a key player in neuronal DNA protection." *The Journal of biological chemistry* **286**(6): 4566-4575.
- Talbiersky, P., F. Bastkowski, et al. (2008). "Molecular clip and tweezer introduce new mechanisms of enzyme inhibition." *J Am Chem Soc* **130**(30): 9824-9828.
- Tan, J. M., E. S. Wong, et al. (2008). "Lysine 63-linked ubiquitination promotes the formation and autophagic clearance of protein inclusions associated with neurodegenerative diseases." *Hum Mol Genet* **17**(3): 431-439.

- Theillet, F. X., C. Smet-Nocca, et al. (2012). "Cell signaling, post-translational protein modifications and NMR spectroscopy." *Journal of biomolecular NMR* **54**(3): 217-236.
- Tini, M., A. Benecke, et al. (2002). "Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription." *Mol Cell* **9**(2): 265-277.
- Tortosa, E., I. Santa-Maria, et al. (2009). "Binding of Hsp90 to tau promotes a conformational change and aggregation of tau protein." *Journal of Alzheimer's disease : JAD* **17**(2): 319-325.
- Trinczek, B., J. Biernat, et al. (1995). "Domains of tau protein, differential phosphorylation, and dynamic instability of microtubules." *Molecular biology of the cell* **6**(12): 1887-1902.
- Uchida, T., M. Takamiya, et al. (2003). "Pin1 and Par14 peptidyl prolyl isomerase inhibitors block cell proliferation." *Chem Biol* **10**(1): 15-24.
- Urusova, D. V., J. H. Shim, et al. (2011). "Epigallocatechin-gallate suppresses tumorigenesis by directly targeting Pin1." *Cancer prevention research* **4**(9): 1366-1377.
- Verdecia, M. A., M. E. Bowman, et al. (2000). "Structural basis for phosphoserine-proline recognition by group IV WW domains." *Nat Struct Biol* **7**(8): 639-643.
- Verdegem, D., K. Dijkstra, et al. (2008). "Graphical interpretation of Boolean operators for protein NMR assignments." *Journal of biomolecular NMR* **42**(1): 11-21.
- Vincent, I., M. Rosado, et al. (1996). "Mitotic mechanisms in Alzheimer's disease?" *The Journal of cell biology* **132**(3): 413-425.
- Vincent, I., J. H. Zheng, et al. (1998). "Mitotic phosphoepitopes precede paired helical filaments in Alzheimer's disease." *Neurobiol Aging* **19**(4): 287-296.
- Virshup, D. M. and S. Shenolikar (2009). "From promiscuity to precision: protein phosphatases get a makeover." *Mol Cell* **33**(5): 537-545.
- Vocadlo, D. J., H. C. Hang, et al. (2003). "A chemical approach for identifying O-GlcNAc-modified proteins in cells." *Proc Natl Acad Sci U S A* **100**(16): 9116-9121.
- Vogelsberg-Ragaglia, V., T. Schuck, et al. (2001). "PP2A mRNA expression is quantitatively decreased in Alzheimer's disease hippocampus." *Experimental neurology* **168**(2): 402-412.
- von Bergen, M., S. Barghorn, et al. (2005). "Tau aggregation is driven by a transition from random coil to beta sheet structure." *Biochim Biophys Acta* **1739**(2-3): 158-166.
- von Bergen, M., S. Barghorn, et al. (2001). "Mutations of tau protein in frontotemporal dementia promote aggregation of paired helical filaments by enhancing local beta-structure." *J Biol Chem* **276**(51): 48165-48174.
- von Bergen, M., S. Barghorn, et al. (2006). "The core of tau-paired helical filaments studied by scanning transmission electron microscopy and limited proteolysis." *Biochemistry* **45**(20): 6446-6457.
- von Bergen, M., P. Friedhoff, et al. (2000). "Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming beta structure." *Proc Natl Acad Sci U S A* **97**(10): 5129-5134.
- Voss, K., B. Combs, et al. (2012). "Hsp70 alters tau function and aggregation in an isoform specific manner." *Biochemistry* **51**(4): 888-898.
- Wang, J. Z., I. Grundke-Iqbal, et al. (1996). "Glycosylation of microtubule-associated protein tau: an abnormal posttranslational modification in Alzheimer's disease." *Nature medicine* **2**(8): 871-875.
- Wang, Z., M. Gucek, et al. (2008). "Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc." *Proc Natl Acad Sci U S A* **105**(37): 13793-13798.
- Wang, Z., N. D. Udeshi, et al. (2010). "Enrichment and site mapping of O-linked N-acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation mass spectrometry." *Mol Cell Proteomics* **9**(1): 153-160.
- Ward, S. M., D. S. Himmelstein, et al. (2012). "Tau oligomers and tau toxicity in neurodegenerative disease." *Biochemical Society Transactions* **40**(4): 667-671.

- Weingarten, M. D., A. H. Lockwood, et al. (1975). "A protein factor essential for microtubule assembly." Proc Natl Acad Sci U S A **72**(5): 1858-1862.
- Welburn, J. and J. Endicott (2005). "Methods for preparation of proteins and protein complexes that regulate the eukaryotic cell cycle for structural studies." Methods Mol Biol **296**: 219-235.
- Wells, L., L. K. Kreppel, et al. (2004). "O-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits." The Journal of biological chemistry **279**(37): 38466-38470.
- Wells, L., K. Vosseller, et al. (2002). "Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications." Molecular & cellular proteomics : MCP **1**(10): 791-804.
- Wells, L., K. Vosseller, et al. (2001). "Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc." Science **291**(5512): 2376-2378.
- Wen, Y., E. Planel, et al. (2008). "Interplay between cyclin-dependent kinase 5 and glycogen synthase kinase 3 beta mediated by neuregulin signaling leads to differential effects on tau phosphorylation and amyloid precursor protein processing." The Journal of neuroscience : the official journal of the Society for Neuroscience **28**(10): 2624-2632.
- Whelan, S. A. and G. W. Hart (2003). "Proteomic approaches to analyze the dynamic relationships between nucleocytoplasmic protein glycosylation and phosphorylation." Circ Res **93**(11): 1047-1058.
- Whelan, S. A. and G. W. Hart (2006). "Identification of O-GlcNAc sites on proteins." Methods Enzymol **415**: 113-133.
- Wildemann, D., F. Erdmann, et al. (2006). "Nanomolar inhibitors of the peptidyl prolyl cis/trans isomerase Pin1 from combinatorial peptide libraries." J Med Chem **49**(7): 2147-2150.
- Wille, H., G. Drewes, et al. (1992). "Alzheimer-like paired helical filaments and antiparallel dimers formed from microtubule-associated protein tau in vitro." The Journal of cell biology **118**(3): 573-584.
- Winkler, K. E., K. I. Swenson, et al. (2000). "Requirement of the prolyl isomerase Pin1 for the replication checkpoint." Science **287**(5458): 1644-1647.
- Wintjens, R., J. M. Wieruszkeski, et al. (2001). "<sup>1</sup>H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides." J Biol Chem **276**(27): 25150-25156.
- Wischik, C. M., M. Novak, et al. (1988). "Structural characterization of the core of the paired helical filament of Alzheimer disease." Proc Natl Acad Sci U S A **85**(13): 4884-4888.
- Wischik, C. M., M. Novak, et al. (1988). "Isolation of a fragment of tau derived from the core of the paired helical filament of Alzheimer disease." Proc Natl Acad Sci U S A **85**(12): 4506-4510.
- Wishart, D. S., C. G. Bigam, et al. (1995). "<sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects." J Biomol NMR **5**: 67-81.
- Wishart, D. S., C. G. Bigam, et al. (1995). "<sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shift referencing in biomolecular NMR." Journal of biomolecular NMR **6**(2): 135-140.
- Wishart, D. S., Sykes, B. D., Richards, F. M. (1992). "The chemical shift index : a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy." Biochemistry **31**: 1647-1651.
- Withers, S. G. and G. J. Davies (2012). "Glycobiology: The case of the missing base." Nature chemical biology **8**(12): 952-953.
- Wood, J. G., S. S. Mirra, et al. (1986). "Neurofibrillary tangles of Alzheimer disease share antigenic determinants with the axonal microtubule-associated protein tau (tau)." Proceedings of the National Academy of Sciences of the United States of America **83**(11): 4040-4043.
- Woody, R. W., D. C. Clark, et al. (1983). "Molecular flexibility in microtubule proteins: proton nuclear magnetic resonance characterization." Biochemistry **22**(9): 2186-2192.
- Wu, B., P. Li, et al. (2004). "3D structure of human FK506-binding protein 52: implications for the assembly of the glucocorticoid receptor/Hsp90/immunophilin heterocomplex." Proceedings of the National Academy of Sciences of the United States of America **101**(22): 8348-8353.

- Wulf, G. M., Y. C. Liou, et al. (2002). "Role of Pin1 in the regulation of p53 stability and p21 transactivation, and cell cycle checkpoints in response to DNA damage." *J Biol Chem* **277**(50): 47976-47979.
- Wulf, G. M., A. Ryo, et al. (2001). "Pin1 is overexpressed in breast cancer and cooperates with Ras signaling in increasing the transcriptional activity of c-Jun towards cyclin D1." *Embo J* **20**(13): 3459-3472.
- Xu, Y., Y. Chen, et al. (2008). "Structure of a protein phosphatase 2A holoenzyme: insights into B55-mediated Tau dephosphorylation." *Mol Cell* **31**(6): 873-885.
- Xu, Y. X., Y. Hirose, et al. (2003). "Pin1 modulates the structure and function of human RNA polymerase II." *Genes Dev* **17**(22): 2765-2776.
- Xu, Y. X. and J. L. Manley (2007). "Pin1 modulates RNA polymerase II activity during the transcription cycle." *Genes Dev* **21**(22): 2950-2962.
- Yaffe, M. B., M. Schutkowski, et al. (1997). "Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism." *Science* **278**(5345): 1957-1960.
- Yang, X. J. (2004). "The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases." *Nucleic acids research* **32**(3): 959-976.
- Yang, X. J. (2004). "Lysine acetylation and the bromodomain: a new partnership for signaling." *BioEssays : news and reviews in molecular, cellular and developmental biology* **26**(10): 1076-1087.
- Yoshiyama, Y., M. Higuchi, et al. (2007). "Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model." *Neuron* **53**(3): 337-351.
- Yuan, Z. L., Y. J. Guan, et al. (2005). "Stat3 dimerization regulated by reversible acetylation of a single lysine residue." *Science* **307**(5707): 269-273.
- Yuzwa, S. A., M. S. Macauley, et al. (2008). "A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo." *Nat Chem Biol* **4**(8): 483-490.
- Yuzwa, S. A., X. Shan, et al. (2012). "Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation." *Nature chemical biology*.
- Yuzwa, S. A. and D. J. Vocadlo (2009). "O-GlcNAc modification and the tauopathies: insights from chemical biology." *Curr Alzheimer Res* **6**(5): 451-454.
- Yuzwa, S. A., A. K. Yadav, et al. (2011). "Mapping O-GlcNAc modification sites on tau and generation of a site-specific O-GlcNAc tau antibody." *Amino acids* **40**(3): 857-868.
- Zacchi, P., M. Gostissa, et al. (2002). "The prolyl isomerase Pin1 reveals a mechanism to control p53 functions after genotoxic insults." *Nature* **419**(6909): 853-857.
- Zachara, N. E., W. D. Cheung, et al. (2004). "Nucleocytoplasmic glycosylation, O-GlcNAc: identification and site mapping." *Methods in molecular biology* **284**: 175-194.
- Zachara, N. E., R. N. Cole, et al. (2001). "Detection and analysis of proteins modified by O-linked N-acetylglucosamine." *Curr Protoc Protein Sci Chapter 12*: Unit 12 18.
- Zachara, N. E. and G. W. Hart (2002). "The emerging significance of O-GlcNAc in cellular regulation." *Chemical reviews* **102**(2): 431-438.
- Zachara, N. E. and G. W. Hart (2004). "O-GlcNAc a sensor of cellular state: the role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress." *Biochim Biophys Acta* **1673**(1-2): 13-28.
- Zachara, N. E. and G. W. Hart (2006). "Cell signaling, the essential role of O-GlcNAc!" *Biochimica et biophysica acta* **1761**(5-6): 599-617.
- Zachara, N. E., G. W. Hart, et al. (2002). "Detection and analysis of proteins modified by O-linked N-acetylglucosamine." *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.] Chapter 17*: Unit 17 16.

- Zachara, N. E., H. Molina, et al. (2011). "The dynamic stress-induced "O-GlcNAc-ome" highlights functions for O-GlcNAc in regulating DNA damage/repair and other cellular pathways." Amino acids **40**(3): 793-808.
- Zachara, N. E., K. Vosseller, et al. (2011). "Detection and analysis of proteins modified by O-linked N-acetylglucosamine." Current protocols in protein science / editorial board, John E. Coligan ... [et al.] Chapter 12: Unit12 18.
- Zhang, F., C. M. Snead, et al. (2012). "Hsp90 regulates O-linked beta-N-acetylglucosamine transferase: a novel mechanism of modulation of protein O-linked beta-N-acetylglucosamine modification in endothelial cells." American journal of physiology. Cell physiology **302**(12): C1786-1796.
- Zhang, Y., S. Daum, et al. (2007). "Structural basis for high-affinity peptide inhibition of human Pin1." ACS Chem Biol **2**(5): 320-328.
- Zhang, Y., S. Fussel, et al. (2002). "Substrate-based design of reversible Pin1 inhibitors." Biochemistry **41**(39): 11868-11877.
- Zhang, Y. J., Y. F. Xu, et al. (2008). "Carboxyl terminus of heat-shock cognate 70-interacting protein degrades tau regardless its phosphorylation status without affecting the spatial memory of the rats." Journal of neural transmission **115**(3): 483-491.
- Zhang, Z., J. Gildersleeve, et al. (2004). "A new strategy for the synthesis of glycoproteins." Science **303**(5656): 371-373.
- Zheng-Fischhofer, Q., J. Biernat, et al. (1998). "Sequential phosphorylation of Tau by glycogen synthase kinase-3beta and protein kinase A at Thr212 and Ser214 generates the Alzheimer-specific epitope of antibody AT100 and requires a paired-helical-filament-like conformation." Eur J Biochem **252**(3): 542-552.
- Zheng, H., H. You, et al. (2002). "The prolyl isomerase Pin1 is a regulator of p53 in genotoxic response." Nature **419**(6909): 849-853.
- Zhou, X. Z., O. Kops, et al. (2000). "Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins." Mol Cell **6**(4): 873-883.
- Zhou, X. Z., P. J. Lu, et al. (1999). "Phosphorylation-dependent prolyl isomerization: a novel signaling regulatory mechanism." Cell Mol Life Sci **56**(9-10): 788-806.

## Sélection d'articles

---

1. Smet, C., A. V. Sambo, et al. (2004). "The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau." *Biochemistry* 43(7): 2032-2040.
2. Smet, C., A. Leroy, et al. (2004). "Accepting its random coil nature allows a partial NMR assignment of the neuronal Tau protein." *Chembiochem : a European journal of chemical biology* 5(12): 1639-1646.
3. Smet, C., J. M. Wieruszeski, et al. (2005). "Regulation of Pin1 peptidyl-prolyl cis/trans isomerase activity by its WW binding module on a multi-phosphorylated peptide of Tau protein." *FEBS letters* 579(19): 4159-4164.
4. Smet, C., J. F. Duckert, et al. (2005). "Control of protein-protein interactions: structure-based discovery of low molecular weight inhibitors of the interactions between Pin1 WW domain and phosphopeptides." *Journal of medicinal chemistry* 48(15): 4815-4823.
5. Smet-Nocca, C., J. M. Wieruszeski, et al. (2010). "NMR-based detection of acetylation sites in peptides." *Journal of peptide science* 16(8): 414-423.
6. Smet-Nocca, C., M. Broncel, et al. (2011). "Identification of O-GlcNAc sites within peptides of the Tau protein and their impact on phosphorylation." *Molecular bioSystems* 7(5): 1420-1429.
7. Sami, F., C. Smet-Nocca, et al. (2011). "Molecular basis for an ancient partnership between prolyl isomerase Pin1 and phosphatase inhibitor-2." *Biochemistry* 50(30): 6567-6578.
8. Landrieu, I., C. Smet-Nocca, et al. (2011). "Molecular implication of PP2A and Pin1 in the Alzheimer's disease specific hyperphosphorylation of Tau." *PloS one* 6(6): e21521.
9. Theillet, F. X., C. Smet-Nocca, et al. (2012). "Cell signaling, post-translational protein modifications and NMR spectroscopy." *Journal of biomolecular NMR* 54(3): 217-236.
10. Smet-Nocca, C., H. Launay, et al. (2013). "Unraveling a phosphorylation event in a folded protein by NMR spectroscopy: phosphorylation of the Pin1 WW domain by PKA." *Journal of biomolecular NMR* 55(4): 323-337.

## The Peptidyl Prolyl *cis/trans*-Isomerase Pin1 Recognizes the Phospho-Thr212-Pro213 Site on Tau<sup>†</sup>

Caroline Smet,<sup>‡</sup> Anne-Véronique Sambo,<sup>§</sup> Jean-Michel Wieruszeski,<sup>‡</sup> Arnaud Leroy,<sup>‡</sup> Isabelle Landrieu,<sup>‡</sup> Luc Buée,<sup>§</sup> and Guy Lippens<sup>\*,‡</sup>

*Institut de Biologie de Lille, Institut Pasteur de Lille, UMR CNRS 8525, BP 245, F-59019 Lille Cedex, France, and Institut de Médecine prédictive et recherche thérapeutique, INSERM U422, Place de Verdun, F-59045 Lille Cedex, France*

*Received August 19, 2003; Revised Manuscript Received December 12, 2003*

**ABSTRACT:** The interaction between the neuronal Tau protein and the Pin1 prolyl *cis/trans*-isomerase is dependent on the phosphorylation state of the former. The interaction site was mapped to the unique phospho-Thr231-Pro232 motif, despite the presence of many other Thr/Ser-Pro phosphorylation sites in Tau and structural evidence that the interaction site does not significantly extend beyond those very two residues. We demonstrate here by NMR and fluorescence mapping that the Alzheimer's disease specific epitope centered around the phospho-Thr212-Pro213 motif is also an interaction site, and that the sole phospho-Thr-Pro motif is already sufficient for interaction. Because a detectable fraction of the Pro213 amide bond in the peptide centered around the phospho-Thr212-Pro213 motif is in the *cis* conformation, catalysis of the isomerization by the catalytic domain of Pin1 could be investigated via NMR spectroscopy.

Tau is a microtubule-associated protein that occurs mainly in neurons (1). Although derived from a single gene, it can exist through alternative splicing as six isoforms that are developmentally regulated (2). Further variability is introduced by posttranslational phosphorylation, with a fetal isoform of Tau carrying significantly more phosphate groups in embryonic stages than the six Tau isoforms in adult brain (3). A decade of research has led to the discovery of a large number of kinases that can phosphorylate Tau on specific threonine (Thr) or serine (Ser) residues. Mapping of the phosphorylation sites was done with specific antibodies (4–6) or with protein chemistry techniques, involving mass spectroscopy and/or Edman degradation (3, 7, 8). It was found that many kinases belong to the proline (Pro)-directed protein kinase family, with an enzymatic activity for one of the 17 Ser/Thr-Pro motifs in full-length Tau. Other kinases, however, such as PKA<sup>1</sup> were found to phosphorylate Ser or Thr residues that are not followed by a Pro (9), whereas

GSK-3 $\beta$  can modify both Ser and Thr residues that are or are not followed by a Pro (10–12).

The main motivation for this extensive research was and remains twofold. First, phosphorylation was found to modulate the binding of Tau to the microtubules, with a lower affinity for the phosphorylated Tau, and might as such influence the dynamic equilibrium of the tubulin/microtubule system (13). Second, in the paired helical filaments (PHF), one of the hallmarks of Alzheimer's disease of which Tau is the main component, the protein is heavily phosphorylated (1). Hyperphosphorylation in the later life stages might reflect an erroneous reactivation of the cell cycle in the aging neurons, proposed on the basis of the agreement between AD and mitosis-related phosphorylation patterns (14–16). Still, the relationship between phosphorylation and aggregation is not clear, and certain phosphorylation sites (Ser214 and Ser262) were recently described as being protective against aggregation (17).

The phospho-epitope mapping by immunochemistry heavily relies on a number of specific antibodies that recognize (i) the protein in its unphosphorylated soluble form, (ii) certain phospho-epitopes, or (iii) the phosphorylated protein aggregated into PHF tangles as found in the diseased brain. Examples of the latter are the AT100 and PHF-27/TG3 antibodies that recognize the AD specific phosphorylation sites composed of the Thr212–Ser214 and Thr231–Ser235 epitopes, respectively (14, 18). Whereas those antibodies are commonly classified as being “conformation-dependent”, the precise conformational changes that they detect have still not been fully elucidated (19, 20). Because many of the phosphorylation sites are Pro-directed, one possible hypothesis is that the Pro conformation would change upon phosphorylation and/or aggregation. The Pro residue is indeed unique in the sense that the energetic barrier between its *trans* and *cis* conformation is significantly lower than for

<sup>†</sup> Part of this work was funded by a grant from the Génopole of Lille, by the INSERM, the Région Guadeloupe, and by the Institute for the Study of Aging (ISOA). The 600 MHz facility used in this study was funded by the Région Nord-Pas de Calais, the CNRS, and the Institut Pasteur de Lille.

\* To whom correspondence should be addressed. Telephone: +33/(0) 320871229. Fax: +33/(0) 320871233. E-mail: guy.lippens@pasteur-lille.fr.

<sup>‡</sup> UMR CNRS 8525.

<sup>§</sup> INSERM U422.

<sup>1</sup> Abbreviations: PKA, protein kinase A; GSK-3 $\beta$ , glycogen synthase kinase 3 $\beta$ ; AD, Alzheimer's disease; PP2A, protein phosphatase 2A; PHF, paired helical filaments; CDK, cyclin-dependent kinase; MAPK, mitogen-activated protein kinase; DMF, dimethylformamide; RP-HPLC, reverse phase high-pressure liquid chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; IPTG, isopropyl thiogalactoside; PBS, phosphate-buffered saline; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser exchange spectroscopy; EXSY, exchange spectroscopy; FRET, fluorescence resonance energy transfer;  $K_D$ , dissociation constant; TFE, trifluoroethanol.



all other amino acids, leading to a higher population of the *cis* form. In small peptides, this population is typically on the order of a few percent, but in the context of a folded protein, Pro residues can be found that are completely in the *cis* conformation.

Indirect evidence of a structural role of the proline conformation comes from the recent finding that Pin1, a prolyl *cis/trans*-isomerase that is essential for the cell cycle and highly conserved within all eukaryotic species (21, 22), interacts with Tau (23). Pin1 was reported to restore the microtubule binding capacity of Cdc2/cyclin B-phosphorylated Tau (23). This might result from a direct physical interaction between both partners, or alternatively, Pin1 might assist proline-directed phosphatase PP2A in its catalytic activity on phosphorylated Tau. Indeed, at least at the level of short peptide substrates, PP2A only catalyzes the dephosphorylation of a threonine when it is followed by a proline residue in the *trans* conformation (24). Finally, the large reservoir of phosphorylated Tau in the PHF tangles might lead to depletion of soluble Pin1, and as such interfere with Pin1's other functions. Because Pin1 is a key regulator of the mitotic transition, the earlier finding that Alzheimer's disease might be related to a reactivation of the cell cycle (14–16) is in agreement with a functional role for Pin1 in AD.

Pin1 is constituted by two domains (25) both essential for its *in vivo* activity: a WW binding domain (26) which binds specifically pThr/pSer-Pro motifs and a peptidyl-prolyl *cis/trans*-isomerase domain which can act catalytically to promote conformational changes in pThr/pSer-Pro bonds by accelerating the intrinsic slow *cis/trans* isomerization of these bonds at least on small peptidic substrates (27). When the interaction between phosphorylated Tau and Pin1 was reported, the authors identified a unique interaction site at the phosphorylated Thr231 residue. Combined peptide mapping and mutagenesis studies indicated that phosphorylation of this proline-directed site, which is phosphorylated significantly in both fetal and adult rat brains (3) but is also one of the phosphorylation sites detected on PHF Tau, is both necessary and sufficient for the interaction with Pin1. Thr231 can be phosphorylated by GSK-3 $\beta$ , but only after phosphorylation of Ser235 by CDK5 or MAPK (6, 28, 29). It can be directly phosphorylated by CDK5 in the presence of heparin (30), and with a low efficiency by the p35cdc2-cyclin A complex (31).

Structural details of the interaction of Pin1 with its various substrates come from several high-resolution X-ray and NMR structures. Pin1 in complex with peptide substrates derived from the C-terminal domain of the polymerase II, the CDC25 phosphatase, or Tau itself indicates a dominant role for the WW domain in the interaction, with an aromatic clamp formed by one of the Trp residues and a Tyr holding onto the Pro side chain, whereas Arg and Ser in the loop region connecting the two  $\beta$ -strands make hydrogen bonds with the phosphorylated side chain of the Thr (32, 33). The interaction seems, however, mainly limited to this precise dipeptide, raising the question of the structural basis for the reported selectivity. In a recent NMR study, the only conformational selectivity we could detect was at the level of the proline isomerization state, with the pThr-Pro moiety interacting only when the Pro is in the *trans* conformation (33).

On the basis of the discrepancy between this structural information and the reported uniqueness of the Pin1 interaction site on Tau, we decided to further investigate this aspect by NMR spectroscopy. As the Thr212 residue is one of the few proline-directed and PHF specific phosphorylation sites, we centered our peptide on this residue. Another reason for the choice of this precise peptide is that the two reported interaction sites on human CDC25C, Thr47 and Thr68, are separated by exactly the same number of residues as Thr212 and Thr231 of Tau (34). Although the sequences of both proteins between their respective two threonine residues do not align, they both are considered random coil polymers without well-defined structure, and their separation might therefore be the relevant structural factor. Moreover, as other sites in the direct vicinity of the Thr212 site can equally be phosphorylated (Ser214 and Thr217), we have first investigated the influence of phosphorylation on the conformation of the peptide, by performing a detailed NMR analysis of the non-, mono-, di-, and triphosphorylated peptide. Their interaction with Pin1 was studied at the level of binding to the WW domain in the full-length protein, and the conformational changes of the proline conformation induced by the catalytic domain of Pin1 were followed by exchange spectroscopy. It was found that the different peptides do bind the Pin1 protein even better than the initially reported pThr231-Pro232 motif, and that at least the conformational exchange of the pThr212-Pro213 bond from *cis* to *trans* or *vice versa* can be catalyzed by Pin1. Because these results do not agree with the previously reported uniqueness of the pThr231-Pro231 site, we have looked into the possibility of other interaction sites on Tau, and confirmed that a minimal pThr-Pro dipeptide already can bind with a good affinity. These combined results show that the interaction of Pin1 and Tau is not limited to one unique site, and pose the intriguing question of a functional cooperativity between the WW and catalytic domain of Pin1 while it interacts with the hyperphosphorylated Tau.

## MATERIALS AND METHODS

**Peptide Synthesis and Purification.** Peptides were obtained by solid phase synthesis with introduction of selectively phosphorylated residues using appropriate building blocks of side chain-protected phosphothreonine or phosphoserine, as described previously (33). A dansyl fluorophore was added as a last step in the peptide synthesis of the pT212 peptide, by reaction of 5 equiv of dansyl chloride and 10 equiv of diisopropylethylamine in DMF. After TFA cleavage, peptides were purified by RP-HPLC on a C18 Nucleosil column (Machee-Nagel, Duren, Germany) and eluted with an acetonitrile gradient. The homogeneity of the fractions was verified by MALDI-TOF mass spectrometry.

**Expression and Purification of Recombinant Pin1 and the Catalytic Domain (Pin1<sub>CAT</sub>).** Pin1 with an N-terminal histidine tag fusion was produced in *Escherichia coli* strain BL21(DE3) (Novagen), carrying the recombinant PIN1-pET28 plasmid (generous gift from M. Yaffe, Harvard University, Cambridge, MA). The recombinant strain was grown at 37 °C in an LB medium containing kanamycin (20 mg/L) until the OD<sub>600</sub> reached ~0.6. The cells were then harvested by centrifugation, and the pellet was resuspended in half the volume in M9 medium with <sup>15</sup>NH<sub>4</sub>Cl (Cambridge Isotope Laboratories, Cambridge, MA) as the nitrogen source



[6 g/L Na<sub>2</sub>HPO<sub>4</sub>, 3 g/L KH<sub>2</sub>PO<sub>4</sub>, 0.5 g/L NaCl, 4 g/L glucose, 1 g/L <sup>15</sup>NH<sub>4</sub>Cl, 0.12 g/L MgSO<sub>4</sub>, 20 mg/L kanamycin, and MEM vitamin cocktail (Sigma)]. The culture was incubated at 37 °C for 1 h before induction with 0.5 mM IPTG followed by growth for 3 h at 31 °C. The cell pellet was resuspended in a lysis buffer [50 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> (pH 8.0), 300 mM NaCl (buffer A), 10 mM imidazole, 1 mM DTT, 0.1% NP40, and a protease inhibitor cocktail (Roche)], and cell lysis was performed by sonication. The soluble extract was loaded on a nickel affinity column (Chelating Sepharose Fast Flow, Amersham Pharmacia Biotech). Unbound proteins were extensively washed with 20 mM imidazole in buffer A, and the protein of interest was eluted with 250 mM imidazole in buffer A.

The GST–Pin1<sub>CAT</sub> fusion protein was produced in the BL21(DE3) star strain using the pGEX-6P plasmid (Amersham Pharmacia Biotech). The production was performed in LB medium with 0.5 mM IPTG induction followed by incubation for 3 h at 31 °C before the cells were harvested. The pellet was resuspended in phosphate-buffered saline (PBS) [130 mM NaCl, 3 mM KCl, 10 mM NaHPO<sub>4</sub>, and 2 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.4)] containing 1% Triton X-100, 20 mM β-mercaptoethanol, and a protease inhibitor cocktail. GST fusion proteins were purified from the soluble extract by use of a GSTrap Chelating Sepharose column (Amersham) equilibrated in PBS (pH 7.2). Unbound proteins were extensively washed with PBS, and proteins were eluted after incubation with Precision Protease (Amersham) in 50 mM Tris (pH 8.0), 100 mM NaCl, and 4 mM EDTA at 4 °C for 20 h.

**NMR Spectroscopy.** NMR experiments were performed on a Bruker DMX spectrometer (Bruker, Karlsruhe, Germany) operating at 14.1 T and equipped with a cryogenic triple-resonance probe head. Spectral parameters were as specified in ref 33. Standard TOCSY and NOESY experiments, with mixing times of 60 and 400 ms, respectively, were obtained with 1 or 2 mM peptide solutions in 50 mM deuterated Tris (pH 6.3) (Cambridge Isotope Laboratories) and 100 mM NaCl at 293 K. <sup>15</sup>N-labeled Pin1 HSQC spectra were recorded on a 200 μM protein sample in 50 mM deuterated Tris (pH 6.3), 100 mM NaCl, and a 5 mM DTT/EDTA mixture at the same temperature. For the titration experiments, a 200 μM solution of Pin1 was added to the appropriate amounts of the lyophilized peptide. The chemical shift perturbations of individual resonances calculated with the relationship  $\Delta\delta$  (parts per million) =  $[\Delta\delta^2(^1\text{H}) + 0.2\Delta\delta^2(^{15}\text{N})]^{1/2}$  were used to derive the dissociation constants (33). For the observation of catalytic activity, 50 μM full-length Pin1 or its catalytic domain alone, each solubilized in D<sub>2</sub>O buffer, was added to a 1.5 mM solution in D<sub>2</sub>O of a triply phosphorylated peptide, and EXSY spectra with a mixing time of 400 ms were recorded.

**Fluorescence Spectroscopy.** FRET experiments were carried out on a PTI (Lawrenceville, NJ) fluorescence spectrometer by exciting tryptophan residues of Pin1 at 295 nm. A 20 μM solution of Pin1 in 50 mM Tris (pH 6.3) containing 100 mM NaCl and a 5 mM DTT/EDTA mixture was added to lyophilized aliquots of the peptide to obtain a full titration curve. When the dansylated peptides were used, the fluorescence quenching of the tryptophan indole groups by the dansyl moiety was assessed at 334 nm. The fluorescence

Table 1: Peptide Sequences Synthesized during This Study<sup>a</sup>

peptide	sequence
T231	KKVAVVRT <sub>231</sub> PPKSPSSAK
pT231	KKVAVVR <b>p</b> T <sub>231</sub> PPKSPSSAK
pT231-pS235	KKVAVVR <b>p</b> T <sub>231</sub> PPK <b>p</b> S <sub>235</sub> PSSAK
T212	SRSRT <sub>212</sub> PSLPTPPTR
pT212	SRSR <b>p</b> T <sub>212</sub> PSLPTPPTR
pT212-pS214	SRSR <b>p</b> T <sub>212</sub> <b>p</b> pS <sub>214</sub> LPTPPTR
pT212-pS214-pT217	SRSR <b>p</b> T <sub>212</sub> <b>p</b> pS <sub>214</sub> <b>L</b> pT <sub>217</sub> PPTR
pThr-Pro dipeptide	pTP

<sup>a</sup> Phosphorylated residues are in bold, and numbered according to their position in the 441-residue isoform. The peptides are represented in the text with the one-letter code, whereas the individual residues are represented with the three-letter code.

variation was used to calculate the *K<sub>D</sub>* as described previously (33).

## RESULTS

**NMR Characterization of the Peptide Epitopes.** Four different peptides were synthesized, corresponding to a nonphosphorylated sequence, and peptides containing the single pThr212 residue, pThr212/pSer214, and finally pThr212/pSer214/pThr217. As a control, we also synthesized sequences centered around Thr231, with phosphorylation of this latter site alone or of both Thr231 and Ser235 (Table 1).

Whereas assignment of such short peptides can now be routinely carried out according to the sequential assignment procedure based on combined TOCSY and NOESY spectra (38), the main problem for detecting and characterizing the minor *cis* forms proved to be less straightforward because of sensitivity reasons. We used 2 mM samples dissolved in 270 μL of aqueous buffer in H<sub>2</sub>O or D<sub>2</sub>O together with a cryogenic probe on a 600 MHz instrument to obtain good quality spectra of even the minor forms. The resulting spectra on all peptides show that phosphorylation has a large effect on the amide proton of the carrying Thr or Ser residue (39), making its ready detection from a one-dimensional spectrum feasible (Figure 1). The same phosphorylated residues followed by a Pro residue in the *cis* conformation show up as low-intensity peaks close to the H<sub>N</sub> frequency of the major pThr/pSer signals followed by a Pro in the *trans* conformation.

When we first analyzed the peptides centered around the Thr231 residue, we confirmed the previously reported results for this epitope (20). The *cis* isomer of the pThr231-Pro232 peptide bond is populated at a level of <3% in water, and the conformation of the peptide as determined by NOE spectroscopy is essentially random. For the peptides centered around Thr212, confirmation of the Pro conformation came from the NOESY spectrum recorded in D<sub>2</sub>O on both the non- and triphosphorylated peptides, where we observed specific NOE contacts between the H<sub>α</sub> (and H<sub>β</sub>) protons of Thr212 and the H<sub>α</sub> proton of *cis*-Pro213, and the same contacts between the corresponding protons of Thr217 and Pro218 (Figure 2). Other minor forms correspond to Pro216, which is preceded by a Leu, whereas no *cis* conformation could be detected for Pro219, despite the fact that its conformation could even be more constrained due to the preceding Pro residue.

Integration of the different NMR signals in the one-dimensional spectra or on the H<sub>D</sub>–H<sub>D</sub> NOE cross-peaks,

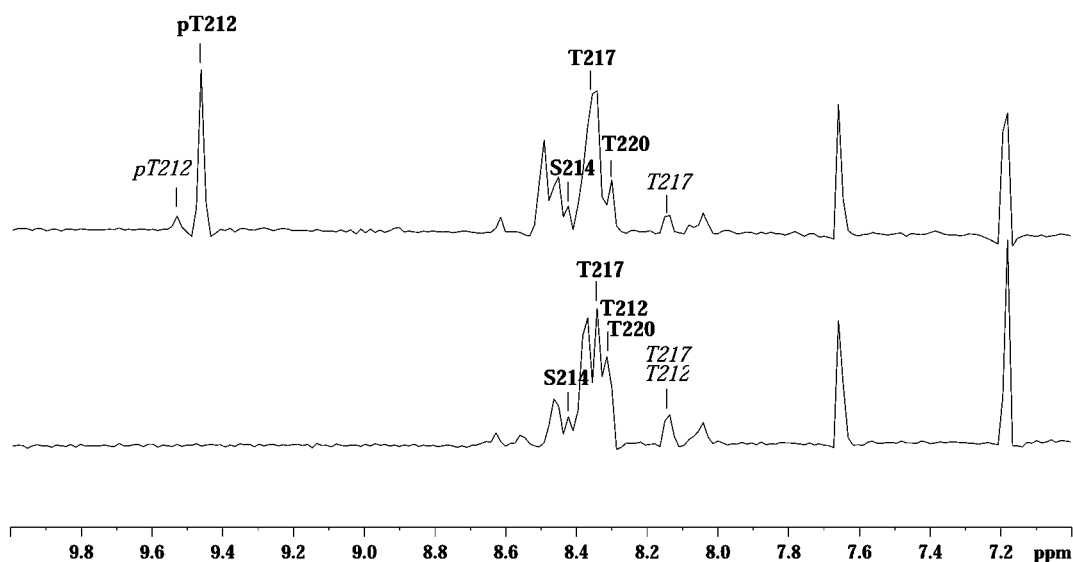


FIGURE 1:  $^1\text{H}$  NMR spectra (zoom of the NH region) of the T212 peptide phosphorylated at the Thr212 position or unphosphorylated. Only Thr212, Ser214, and Thr217 are annotated. Those in bold precede a *trans*-Pro form, and those in italics precede a *cis*-Pro form.

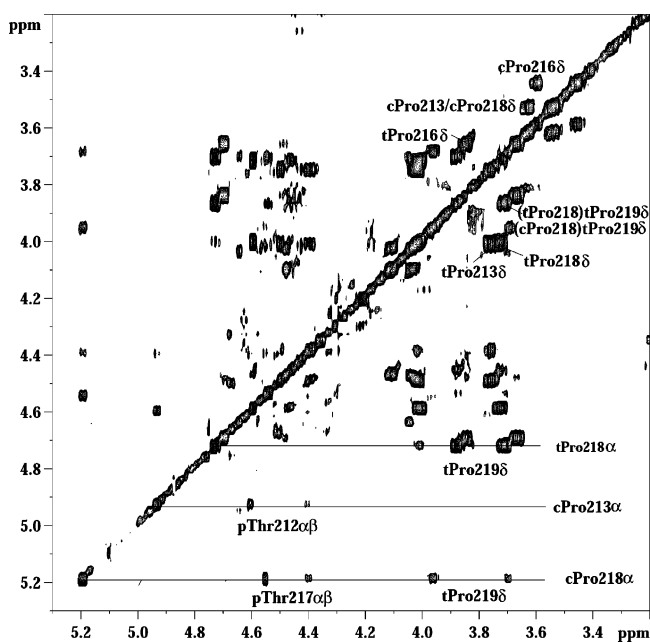


FIGURE 2: Assignment of the triphosphorylated T212 peptide on a 400 ms NOESY spectrum in  $\text{D}_2\text{O}$ .  $\text{H}\delta\text{-H}\alpha$  NOE contacts were assigned for *trans*-Pro isoforms and  $\text{H}\alpha\text{-H}\alpha$  for *cis*-Pro forms of Pro213 and Pro218.  $\text{H}\delta 1\text{-H}\delta 2$  NOE contacts are annotated for all *trans*-Pro forms, for *cis*-Pro213, *cis*-Pro216, and *cis*-Pro218 forms, and for *trans*-Pro219 preceded by *cis*-Pro218.

Table 2: Relative Populations of the *cis/trans* Ratio for the Different Pro Residues in the Non- or Triphosphorylated T212 Peptide

peptide	<i>cis/trans</i> ratio (%)			
	P213	P216	P218	P219
T212	6	8	12	<1
pT212-pS214-pT217	7	8	13	<1

where a reasonable dispersion exists, allowed the estimation of the relative populations of the *trans* and *cis* conformers (Table 2). Importantly, it was found that even a triple phosphorylation did not significantly change the relative populations.

**Interaction Mapping of the Peptides with Pin1.** Rather than using the isolated WW domain as we reported in a previous NMR study (33), we decided to work directly with the full-length  $^{15}\text{N}$ -labeled Pin1 protein. Near-complete assignments for this protein were previously reported (40), enabling a direct identification of the WW resonances in the full-length protein. When we titrated the phosphopeptide centered around the Thr231 epitope against a constant concentration of full-length Pin1, we found a nearly identical value for the dissociation constant. A similar titration experiment with a nonphosphorylated peptide did not yield detectable binding. When we repeated the experiment with the pT212 peptide, carrying a single phospho group on the Thr212 side chain, we detected a 3-fold stronger interaction, confirming that pThr231 is neither a unique nor even the best binding epitope. Resonances on the Pin1 protein that shifted upon addition of the pT212 peptide were mainly assigned to the WW domain, with the strongest shifts for the same residues (Ser16–Gly20 and Gln33–Glu35, with a very pronounced variation for the Trp34  $\text{HN}\epsilon 1$  side chain moiety) that were identified in the titration with the pT231 peptide (33). Small variations were equally detected for some resonances on the catalytic domain. A first category of those was assigned to residues at the interface of the WW and catalytic domain, including, for example, the Phe139 and Leu141 residues. On the basis of the chemical shift variation of their resonances as a function of the increasing peptide concentration, we derived a  $K_D$  constant similar to the one describing the interaction with the WW domain (Table 4). A second group of resonances that shift slightly upon addition of peptides are those of the catalytic site, including Met130, Leu122, Ser114, and Ser115. Because even at the 20-fold peptide excess used in our titration study, saturation of the chemical shift variations was not attained, we could not derive a reliable  $K_D$  value for the catalytic domain. This very weak interaction is in agreement with our findings for the interaction between phosphopeptides and the Pin1 protein of *Arabidopsis thaliana*, which consists only of a catalytic domain (41).

Table 3: Chemical Shift Values of the Triply Phosphorylated pT212-pS214-pT217 Peptide<sup>a</sup>

		H <sub>N</sub>	H <sub>α</sub>	H <sub>β</sub>		H <sub>γ</sub>		H <sub>δ</sub>	
				H <sub>β1</sub>	H <sub>β2</sub>	H <sub>γ1</sub>	H <sub>γ2</sub>	H <sub>δ1</sub>	H <sub>δ2</sub>
208	S	8.38	4.43	3.87					
209	R	8.5	4.44	1.91	1.79	1.67		3.21	
210	S	8.49	4.45	3.89					
211	R	8.48	4.37	1.91	1.79	1.67		3.23	
212 (t)	pT	9.45	4.50	4.40		1.38			
212 (c)	pT	9.52	4.61	4.41		1.27			
213 (t)	P	—	4.43	2.34	2.09	2.04	1.93	4.01	3.78
213 (c)	P	—	4.94	2.41	2.12	1.95	1.84	3.61	3.48
214	pS	9.10	4.44	4.11	3.98				
215 (t)	L	8.40	4.69	1.71	1.63	1.59		0.97	0.94
215 (c)	L	7.80	4.51	1.60		1.43		0.91	
216 (t)	P	—	4.49	2.31	2.03	1.89		3.87	3.68
216 (c)	P	—	4.60	2.41	2.13	1.95	1.85	3.62	3.48
217 (t)	pT	9.03	4.60	4.47		1.37			
217 (c)	pT	9.20	4.56	4.40		1.30			
218 (t)	P	—	4.72	2.4	2.09	2.05	1.9	4.01	3.75
218 (c)	P	—	5.20	2.52	2.17	2.00	1.86	3.62	3.54
219	P	—	4.55	2.36	2.09	2.05	1.93	3.89	3.71
220	T	8.29	4.32	4.22		1.22			

<sup>a</sup> Minor form assignments of residues in the peptide with a Pro in the *cis* conformation are given.

Table 4: Dissociation Constant Values Obtained by NMR for Various Phosphorylated Peptide Substrates against the Full-Length Pin1 Protein<sup>a</sup>

peptide	$K_D^{\text{PIN1}}$ (mM)	peptide	$K_D^{\text{PIN1}}$ (mM)
pT231	0.38 ± 0.1	pT212-pS214	0.16 ± 0.02
pT231-pS235	0.18 ± 0.05	pT212-pS214-pT217	0.10 ± 0.01
pT212	0.10 ± 0.03	pThr-Pro dipeptide	0.15 ± 0.03

<sup>a</sup> NMR dissociation constant values are based on chemical shift perturbations of several residues ( $N > 5$ ). The data for the dipeptide were obtained with the isolated WW domain.

To investigate the influence of multiple phosphorylations on the interaction, we performed similar titration experiments with three different peptides, including doubly phosphorylated Alzheimer-type epitopes pThr231/pSer235 and pThr212/pSer214. Amazingly, the effect on the  $K_D$  for Pin1 of a second phosphorylation is not the same for both peptides; the addition of a phosphate on Ser235 increases the affinity by a factor of 2, whereas phosphorylation of Ser214 decreases the affinity for Pin1. Addition of a third phosphorylation site to the latter peptide (Thr217) was necessary to restore the initial dissociation constant of 100  $\mu\text{M}$ .

The dissociation constant values for the pT212 peptide–Pin1 interaction were confirmed in an independent experiment based on the fluorescence variation of Pin1 upon titration in a dansylated pT212 peptide. Because of energy transfer from the tryptophan emitting at 334 nm to the dansyl group on the N-terminus of the pT212 peptide, the fluorescence of the tryptophans at 334 nm is efficiently quenched (Figure 4a). The resulting decrease in fluorescence intensity can more readily be quantified as the increase in the intrinsic tryptophan fluorescence that we had previously used with a nondansylated peptide as a substrate (42). The  $K_D$  value of 110  $\mu\text{M}$  confirms the data from the NMR titration experiment (Figure 4b), and demonstrates the absence of interference by the dansyl group.

**Catalytic Activity of Pin1.** The absence of a detectable amount of the *cis* form in the pThr231 peptide makes it a

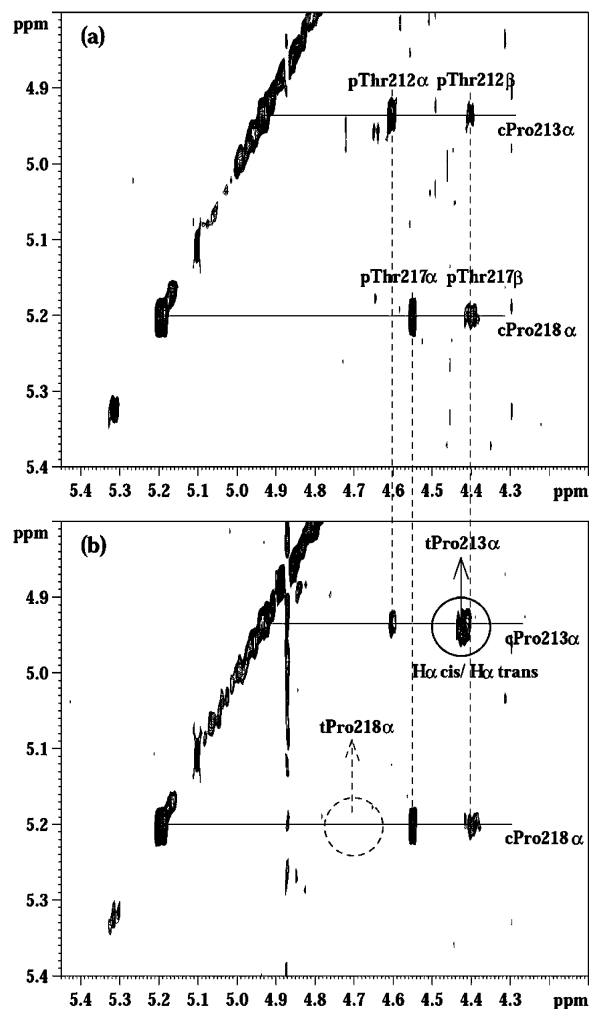


FIGURE 3: (a) Region around H<sub>α</sub> *cis*-Pro213 and Pro218 in the 400 ms EXSY spectrum of the triphosphorylated T212 peptide. Correlations involving pThr212 and pThr217 H<sub>α</sub> and H<sub>β</sub> minority forms are annotated. (b) Same region in the 400 ms EXSY spectrum of the triphosphorylated T212 peptide containing 50  $\mu\text{M}$  Pin1 showing an additional signal that correlated H<sub>α</sub> *cis* and H<sub>α</sub> *trans* of Pro213 (circled). The corresponding signal for Pro218 is not observed as indicated by the dashed circle.

less suitable substrate for the detection of a potential catalytic action of Pin1. Because the situation with the epitope centered around pThr212 is different, and because the use of the cryogenic probe had allowed the full assignment of all minor *cis* forms (Table 3), we used EXSY spectroscopy to probe the catalytic activity of Pin1. The experimental setup was similar to the one we used to demonstrate the catalytic activity of the *A. thaliana* Pin1 protein (35), with 400 ms EXSY spectra of the triply phosphorylated peptide in D<sub>2</sub>O with or without Pin1 added. Without Pin1, the isomerization proved to be too slow to give detectable cross-peaks connecting the signal of the same proton on a peptide in two conformations. When we added a catalytic amount of the protein, however, several additional cross-peaks appeared. The contact between the H<sub>α</sub> protons of Pro213 in the *cis* and *trans* conformations was easily identified (Figure 3), despite the presence of a weak contact between the H<sub>α</sub> proton of Pro213 in the *cis* conformation and the H<sub>β</sub> proton of pThr212. The structural transition of the Thr212-Pro213 amide bond was further confirmed by a contact between the H<sub>δ</sub> protons of Pro213 in the *cis* and *trans* conformations.



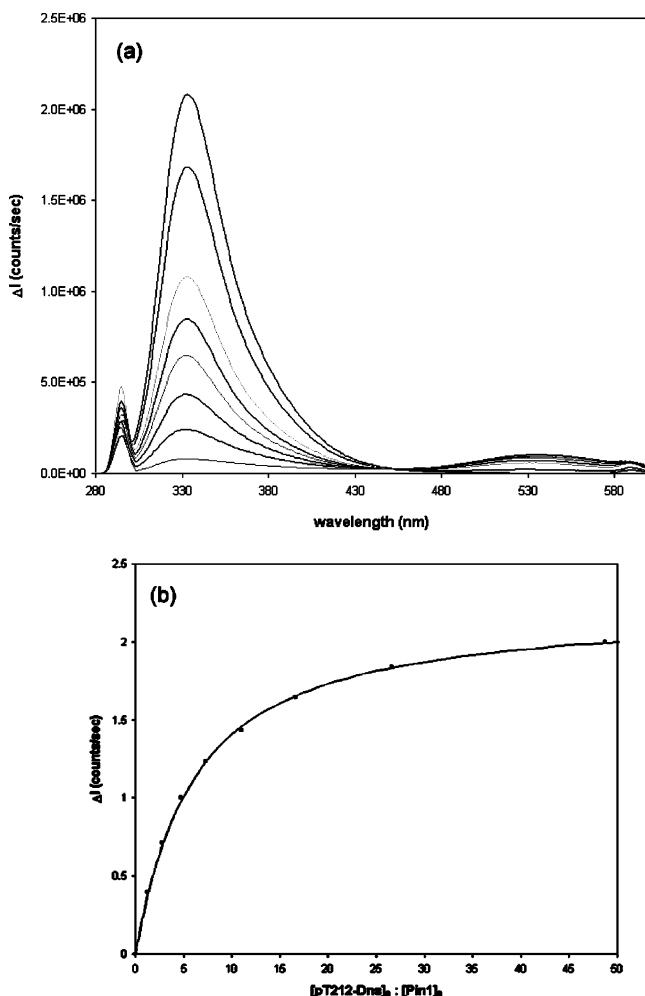


FIGURE 4: (a) Fluorescence titration of Pin1 by the pT212 Tau peptide coupled to a dansyl group at the N-terminus (pT212-Dns) involving FRET. Increasing amounts of the peptide were added to a 20  $\mu$ M solution of Pin1, resulting in a quenching of the fluorescence signal at 334 nm and a slight increase in the dansyl signal at 540 nm. (b) Graphical representation of the  $K_D$  calculation for the pT212 peptide based on the intensity variations of the fluorescence at 334 nm plotted against the molar ratio of the pT212 dansylated peptide to Pin1 concentration.

The isolated character of the H $\alpha$  proton of Pro218 in the *cis* conformation should make the observation of an exchange-mediated contact straightforward, but even with a long mixing time of 400 ms, we could not observe such a contact. This indicates that Pin1 selectively catalyzes the conformational change of the peptide bond between pThr212 and Pro213, but not the one between pThr217 and Pro218. When we performed the same experiment with the isolated catalytic domain of Pin1, very similar results were obtained.

## DISCUSSION

The phosphorylation of Tau has been the subject of intense studies. A first line of research was aimed at the mapping of the *in vivo* phosphorylated sites of fetal or adult soluble Tau (3) as well as the Tau found in the insoluble PHF aggregates (7, 8). Both immunochemical and protein chemistry techniques led to the conclusion that adult Tau carries significantly fewer phosphates than fetal Tau, and that phosphorylation on PHF Tau is even more pronounced than in fetal Tau. The dual character of the hyperphosphorylation stems from the fact that a larger number of phosphorylation

sites has been identified in PHF Tau than in fetal Tau, and second, the occupation level of the individual sites is higher in PHF Tau than in fetal Tau. An example of this latter factor is the Thr231 site, which is only phosphorylated to a level of 30% in fetal Tau, whereas no unphosphorylated Thr231 is found in PHF Tau (8).

Whereas phosphorylation of the Ser262 residue or Ser214 is known to promote detachment of Tau from the microtubules, and might hence lead to microtubule destabilization (13), a clear relationship between phosphorylation and aggregation has not been established. Recent results, for example, have demonstrated that phosphorylation of the same Ser262 site actually prevents aggregation (17), in apparent contradiction with the initial hypothesis that hyperphosphorylation is a causal factor for aggregation. The sole factor of phosphorylation might thus not be sufficient to explain the physicochemical changes that govern microtubule interaction or tangle formation. Additional evidence for this comes from the recent finding that the perturbed interaction between phosphorylated Tau and tubulin might be restored by the parvulin Pin1 (23). This latter enzyme belongs to the class of prolyl *cis/trans*-isomerases that catalyze the conformation of the prolyl bond. However, whereas the cyclophilins or FKBP binding proteins, the two main other classes of prolyl *cis/trans*-isomerases, are at least in yeast not essential for correct cell functioning (43), Pin1 or its yeast analogue Ess1 is crucial for correct progress through the cell cycle. Many of the recognition motifs of Pin1 coincide moreover with the epitopes of the monoclonal MPM2 antibody. This latter recognizes mitosis specific phosphoproteins, confirming further the role of Pin1 in the cell cycle.

Pin1 is characterized by a specificity at the substrate level, as it catalyzes the conformational change only when the proline is preceded by a phosphorylated Thr or Ser residue. The same motif is recognized by the second structural unit, the 40-amino acid WW domain, that seemingly promotes protein-protein interactions and is essential for the correct functioning of the protein (44). The interaction of Pin1 with a large number of phosphoproteins involved in the cell cycle has led to the hypothesis of a second conformational mechanism of posttranslational control after the covalent modification of phosphorylation. A proline *cis/trans* isomerization could be the structural switch, but direct evidence of such a mechanism with true protein targets has only very recently begun to emerge (45). Finally, it should be noted that the described interaction sites on several key mitotic proteins are in the unstructured regulatory parts of the protein. Examples are the Thr48 and Thr67 sites on the regulatory part of CDC25C (46, 47), or the Ser63 and Ser73 epitopes on c-JUN (48).

Because phosphorylated Tau extracted from the brains of people with AD is equally recognized by the mitotic MPM2 antibodies (49), Lu et al. (23) set out to study its interaction with Pin1. They discovered several intriguing aspects. First, Pin1 would be able to restore the microtubule binding capacity of Cdc2/cyclin B-phosphorylated Tau, and second, the interaction with phospho-Thr231 would be necessary and sufficient for this interaction. However, the observed specificity is not easy to explain at the structural level, as several studies have proven that the main interaction motif of the WW domain is the pThr-Pro dipeptide, with a minor contribution of the residue following the Pro (32, 33).

Moreover, although Thr231 is indeed phosphorylated to a higher level in PHF Tau, it can at least partially be phosphorylated in fetal Tau, raising the question of the conformational specificity of the Pin1–Tau interaction. For these reasons, we decided to further investigate the uniqueness of the interaction.

Of those sites that are unique to PHF Tau, two (Thr212 and Ser422) are proline-directed sites. We decided to first investigate the former site, because it is in the vicinity of the Thr231 residue, with the same distance from the latter as the two epitopes described on CDC25C (34). Moreover, because this Thr212 residue has at least two other phosphorylation sites in its immediate vicinity, we synthesized the peptides in different combinations of phosphorylation. NMR spectroscopy was used to investigate first the conformational aspects of those peptides, with specific attention being paid to the *cis/trans* ratios of the different prolines that are involved. Indeed, the absence of interaction between one of those epitopes and Pin1 could be due to the latter containing a Pro residue completely constrained in the *cis* conformation, where interaction with the WW domain is no longer possible (32, 33). A previous study with peptides centered around Thr231 had shown that the phospho-Thr–Pro bond in this peptide is predominantly in the *trans* conformation, with only 3% in the *cis* conformation. Addition of TFE, a solvent that is supposed to mimic the more hydrophobic environment of the inner protein, even further reduced this fraction (20). Therefore, it is not clear how a local structural change in such a small amount of the *cis* conformation containing chains could form the structural switch that explains the functional interaction between Pin1 and Tau. Of course, our results are obtained on a small peptide substrate and hence do not necessarily reflect the situation in the full-length protein. Still, recent NMR results of our group tend to confirm that the random coil chemical shift can be used efficiently to assign at least partially the full-length protein (51), confirming its unstructured nature and validating the peptide approach.

In our T212 peptide, the pool of peptide bonds in the *cis* conformation is somewhat larger, and this occurs not only at the level of the Thr212–Pro213 bond. Indeed, using a 2 mM sample and a measuring time of 12 h on a 600 MHz spectrometer equipped with a cryoprobe, we were able to distinguish the conformational heterogeneity at the level of the Pro213, Pro216, and Pro218 peptide bonds, and assign the different peptides in solution (Figure 2). On the basis of the integration of the one-dimensional proton spectrum and of the proline  $\delta 1$ – $\delta 2$  cross-peaks in the NOESY spectrum, we obtained the relative concentrations for the *trans* and *cis* peptide bonds preceding Pro213, Pro216, and Pro218 (Table 1). Only for Pro219 could we not detect a *cis* conformation, which limits its level to an estimated 1%. However, when comparing the relative populations in the differently phosphorylated peptides, we conclude that there is no clear modification of the *cis/trans* distribution with an increasing level of phosphorylation. As for the secondary structure of the peptides, the only noteworthy NOE contact that we observed was a medium-intensity contact between the amide protons of Ser214 and Leu215. This might indicate a kink in the peptide structure, but it remains to be seen whether this contact is an artifact of the short size of the peptide or whether it is truly present in the full-length protein.

The interaction of the different peptides with Pin1 was first investigated in a series of titration experiments against full-length Pin1. The  $^{15}\text{N}$ -labeled protein was used simultaneously to derive the dissociation constant from a fit of the chemical shift variations as a function of increased peptide concentration and to map those residues that directly interact with the peptide. A first titration experiment with the pT231 peptide resulted in a similar  $K_D$  value as previously determined against the isolated WW domain (33), confirming the primordial role of the WW domain as a protein–protein interaction module. When we performed a second titration series with a peptide phosphorylated at both Thr231 and Ser235, mimicking thereby the recognition site of the Alzheimer specific TG3 antibody, the affinity increased roughly 2-fold. We are at present using calorimetry to study whether this increased affinity results in a loss of entropy in the unbound state, or from a true increase in the enthalpy of binding.

The main result of our study came when we used the pT212 peptide in the same titration study. Indeed, the affinity of this peptide proved to be even better than that of the peptide centered on pThr231. Because this result casts doubt on the proposed uniqueness of this interaction site, we repeated the same measurement by fluorescence spectroscopy with a peptide modified with a dansyl group at its N-terminus. Energy transfer between Trp34 of the WW domain and the dansyl acceptor proved to be very efficient, and the resulting fluorescence decrease proved to be easier to interpret than the intrinsic fluorescence changes that we had previously used to confirm independently the NMR  $K_D$  value for the pT231–WW interaction (42). Despite the addition of a dansyl, the resulting  $K_D$  value confirmed our previous results derived by NMR.

Contrary to the case of the pT231 peptide, a second phosphorylation on the Ser214 site decreases the affinity by a factor of 2. We can rationalize this observation with the crystal structure of Pin1 in complex with a peptide derived from the C-terminal domain of polymerase II, where the hydroxyl Ser residue downstream of the phospho-Thr–Pro recognition motif also makes a hydrogen bond with the WW Trp34 side chain (32). In the pT212 peptide, we also find a Ser214 residue, and from the fact that the affinity of pT212 is higher than that of the pT231 peptide, where the Pro233 side chain does not offer the possibility of a hydrogen bond, we can safely assume that Ser214 indeed contributes to the increased affinity. Phosphorylation, however, would introduce a bulky phosphate group, which might be harder to fit than the simple hydroxyl group. Only when we create a second canonical phospho-Thr217–Pro218 interaction motif in the triply phosphorylated peptide did we find our initial dissociation constant of 100  $\mu\text{M}$ , indicating that both dipeptides are recognized with similar affinities. This latter result motivated us to measure the affinity of the isolated phospho-Thr–Pro dipeptide, devoid of any sequence context. The dissociation constant measured is on the same order of magnitude as that for the pT212 peptide, confirming that the main motif of interaction is indeed the dipeptide, and that the other amino acids only weakly affect this binding.

Because the interaction mapping studies were performed with the full-length Pin1 protein, we could also observe potential interaction sites with the catalytic domain. A first category of residues whose amide correlation shifted slightly

upon addition of the peptide includes residues at the interface between the WW and catalytic domain. Because the  $K_D$  derived from those chemical shift variations is identical to the one obtained for the WW domain, we conclude that it concerns the same binding event, and that binding of the peptide to the WW domain slightly interferes with the catalytic domain. This is in agreement with two recent NMR studies, where chemical shift mapping combined with interdomain NOEs and residual dipolar couplings led to the conclusion of possible cross-talk between the two domains, although the fine details seem to be substrate-dependent (50, 51). An independent interaction site, be it with a very weak affinity, was identified in the long loop that was found to contain the Ala-Pro dipeptide in the first crystal structure (25). However, as was the case for Pin1 from *A. thaliana*, only a lower bound of 10 mM could be determined for the affinity constant.

On the basis of the fact that both *cis* and *trans* conformations can be observed for Pro213 and Pro218 in this triphosphorylated peptide, we have investigated the Pin1 catalytic activity by adding 50  $\mu$ M enzyme to a 1.5 mM solution of the peptide in  $D_2O$ . Unexpectedly, the only Pin1-dependent prolyl *cis/trans* isomerization detected through an *cis-H $\alpha$ -trans-H $\alpha$*  exchange cross-peak was for the Pro213 residue. A similar correlation peak was not observed for the Pro218 residue, although this motif is characterized by twice as much *cis* conformer as the peptide bond preceding Pro213. The presence of a Pro residue before and after the pThr217-Pro218 moiety possibly imposes a structural constraint such that the *cis/trans* interconversion can no longer be catalyzed by the Pin1 enzymatic domain.

Phosphorylation of the different Thr/Ser-Pro sites of the Tau protein is controlled by various well-known kinase-phosphatase complexes. Phosphorylation can be implemented among others *in vitro* with purified activated kinases such as the Cdc2-cyclin B complex used in the initial study identifying the Pin1-Tau interaction (23). However, this latter kinase was previously shown to incorporate only 0.1 phosphate unit per Tau molecule, even with an incubation time of 24 h (29). This latter factor might explain why the interaction with other sites such as the Thr212-Pro213 site described here initially went undetected, as the Cdc2-cyclin B complex probably only slightly catalyzes, if at all, the phosphorylation of this site (41). This is of course in sharp contrast to the peptide studies, where chemical synthesis allows for a 100% incorporation of a phosphate at any given position. As for the important question of whether T212 can interact with Pin1 in the full-length protein, preliminary pull-down results with recombinant Tau phosphorylated by a COS cell extract in the presence of the PP2A phosphatase blocking okadaic acid show that the T231A mutant and the T212A mutant still interact with Pin1 (A.-V. Sambo, M. Hamdane, and L. Buée, unpublished results). The double mutation (T212A/T231A) leads to a further decrease in the level of the retained protein, but still not a complete absence of interaction, suggesting the presence of even other interaction sites.

In conclusion, we have shown through careful peptide mapping that the phospho-Thr231-Pro232 dipeptide is not the only interaction motif between Tau and the cell cycle specific prolyl *cis/trans*-isomerase Pin1, and that the phospho-Thr212-Pro213 motif has an even higher affinity for the

Pin1 WW domain. The enzymatic activity of Pin1 on this latter site was confirmed through exchange NMR spectroscopy. Many questions about the molecular details of the Pin1-Tau interaction remain, however, and one of the most important is the possible cooperativity between both Thr212 and Thr231 sites, which could be a model for the action of Pin1 on other substrates such as Cdc25C. Finally, all our work is related to the interaction between Pin1 and Tau with the latter in its phosphorylated but free state. In the microtubule-bound state, however, we cannot exclude the possibility that one of several phosphorylated Thr/Ser-Pro motifs of Tau is forced into a *cis* conformation, and that those motifs are the true substrates for the prolyl *cis/trans*-isomerase activity of Pin1.

Future efforts will be directed to larger substrates encompassing both sites, in both their free and microtubule-bound state, and should be able to describe in further detail the intriguing functional role of the WW and prolyl isomerase domains of Pin1.

## REFERENCES

- Delacourte, A., and Buée, L. (1997) *Int. Rev. Cytol.* 171, 167–224.
- Goedert, M., Spillantini, M. G., Jakes, R., Rutherford, D., and Crowther, R. A. (1989) *Neuron* 3, 519–526.
- Watanabe, A., Hasegawa, M., Suzuki, M., Takio, K., Morishima-Kawashima, M., Titani, K., Arai, T., Kosik, K. S., and Ihara, K. Y. (1993) *J. Biol. Chem.* 268, 25712–25717.
- Biernat, J., Mandelkow, E.-M., Schröter, C., Lichtenberg-Kraag, B., Steiner, B., Berling, B., Meyer, H., Mercken, M., Vandermeeren, A., Goedert, M., and Mandelkow, E. (1992) *EMBO J.* 11, 1593–1597.
- Otvos, L., Feiner, L., Lang, E., Szendrei, G. I., Goedert, M., and Lee, V. M.-Y. (1994) *J. Neurosci. Res.* 39, 669–673.
- Goedert, M., Jakes, R., Crowther, R. A., Cohen, P., Vanmechelen, E., Vandermeeren, M., and Cras, P. (1994) *Biochem. J.* 301, 871–877.
- Hasegawa, M., Morishima-Kawashima, M., Takio, K., Suzuki, M., Titani, K., and Ihara, Y. (1992) *J. Biol. Chem.* 267, 17047–17054.
- Morishima-Kawashima, M., Hasegawa, M., Takio, K., Suzuki, M., Yoshida, H., Watanabe, A., Titani, K., and Ihara, Y. (1995) *Neurobiol. Aging* 16, 365–371.
- Brandt, R., Lee, G., Teplow, D. B., Shalloway, D., and Abdelghany, M. (1994) *J. Biol. Chem.* 269, 11776–11782.
- Wang, J. Z., Wu, Q., Smith, A., Grundke-Iqbal, I., and Iqbal, K. (1998) *FEBS Lett.* 436, 28–34.
- Godemann, R., Biernat, J., Mandelkow, E., and Mandelkow, E.-M. (1993) *FEBS Lett.* 454, 157–164.
- Illenberger, S., Zheng-Fischhofer, Q., Preuss, U., Stamer, K., Baumann, K., Trinczek, B., Biernat, J., Godemann, R., Mandelkow, E. M., and Mandelkow, E. (1998) *Mol. Biol. Cell* 9, 1495–1512.
- Gustke, N., Steiner, B., Mandelkow, E. M., Biernat, J., Meyer, H. E., Goedert, M., and Mandelkow, E. (1992) *FEBS Lett.* 307, 199–205.
- Vincent, I., Rosado, M., and Davies, P. (1996) *J. Cell Biol.* 132, 413–425.
- Vincent, I., Zheng, J. H., Dickson, D. W., Kress, Y., and Davies, P. (1998) *Neurobiol. Aging* 19, 287–296.
- Delobel, P., Flament, S., Hamdane, M., Mailliot, C., Sambo, A.-V., Bégard, S., Sergeant, N., Delacourte, A., Vilain, J.-P., and Buée, L. (2002) *J. Neurochem.* 83, 1–9.
- Schneider, A., Biernat, J., von Bergen, M., Mandelkow, E., and Mandelkow, E. M. (1999) *Biochemistry* 38, 3549–3558.
- Hoffmann, R., Lee, V. M., Leight, S., Varga, I., and Otvos, L., Jr. (1997) *Biochemistry* 36, 8114–8124.
- Jicha, G. A., Lane, E., Vincent, I., Otvos, L., Jr., Hoffmann, R., and Davies, P. (1997) *J. Neurochem.* 69, 2087–2095.
- Daly, N. L., Hoffmann, R., Otvos, L., Jr., and Craik, D. J. (2000) *Biochemistry* 39, 9039–9046.
- Lu, K., Hanes, S., and Hunter, T. (1996) *Nature* 380, 544–547.



22. Yaffe, M. B., Schutkowski, M., Shen, M., Zhou, X. Z., Stukenberg, P. T., Rahfeld, J. U., Xu, J., Kuang, J., Kirschner, M. W., Fischer, G., Cantley, L. C., and Lu, K. P. (1997) *Science* 278, 1957–1960.
23. Lu, P.-J., Wulf, G., Zhou, X. Z., Davies, P., and Lu, K. P. (1999) *Nature* 399, 784–788.
24. Stukenberg, P. T., and Kirschner, M. W. (2001) *Mol. Cell* 7, 1071–1083.
25. Ranganathan, R., Lu, K., Hunter, T., and Noel, J. (1997) *Cell* 89, 875–886.
26. Sudol, M., Bork, P., Einbond, A., Kastury, K., Druck, T., Negrini, M., Huebner, K., and Lehman, D. (1995) *J. Biol. Chem.* 270, 14733–14741.
27. Fischer, G., and Aumuller, T. (2003) *Rev. Physiol., Biochem. Pharmacol.* 148, 105–150.
28. Ishiguro, K., Shiratsuchi, A., Sato, S., Omori, A., Arioka, M., Kobayashi, S., Uchida, T., and Imahori, K. (1993) *FEBS Lett.* 325, 167–172.
29. Drewes, G., Lichtenberg-Kraag, B., Doring, F., Mandelkow, E. M., Biernat, J., Goris, J., Doree, M., and Mandelkow, E. (1992) *EMBO J.* 11, 2131–2138.
30. Paudel, H. K., and Li, W. (1999) *J. Biol. Chem.* 274, 8029–8038.
31. Vulliet, R., Halloran, S. M., Braun, R. K., Smith, A. J., and Lee, G. (1992) *J. Biol. Chem.* 267, 22570–22574.
32. Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., and Noël, J. P. (2000) *Nat. Struct. Biol.* 7, 639–643.
33. Wintjens, R., Wieruszeski, J.-M., Rousselot-Pailley, P., Drobecq, H., Lippens, G., and Landrieu, I. (2001) *J. Biol. Chem.* 276, 25150–25156.
34. Crenshaw, D., Yang, J., Means, A., and Kornbluth, S. (1998) *EMBO J.* 17, 1315–1327.
35. Landrieu, I., De Veylder, L., Fruchart, J. S., Odaert, B., Casteels, P., Portetelle, D., Van Montagu, M., Inze, D., and Lippens, G. (2000) *J. Biol. Chem.* 275, 10577–10581.
36. Lefebvre, T., Ferreira, S., Dupont-Wallois, L., Bussiere, T., Dupire, M. J., Delacourte, A., Michalski, J.-C., and Caillet-Boudin, M. L. (2003) *Biochim. Biophys. Acta* 1619, 167–176.
37. Mailliot, C., Bussiere, T., Caillet-Boudin, M. L., Delacourte, A., and Buée, L. (1998) *Neurosci. Lett.* 255, 13–16.
38. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
39. Bienkiewicz, E. A. (1999) *J. Biomol. NMR* 15, 203–206.
40. Jacobs, D. M., Saxena, K., Grimme, S., Vogtherr, M., Pescatore, B., Langer, T., Elshorst, B., and Fiebig, K. M. (2002) *J. Biomol. NMR* 23, 163–164.
41. Landrieu, I., Wieruszeski, J.-M., Wintjens, R., Inze, D., and Lippens, G. (2002) *J. Mol. Biol.* 320, 321–332.
42. Hamdane, M., Smet, C., Sambo, A.-V., Leroy, A., Wieruszeski, J. M., Delobel, P., Maurage, C. A., Ghestem, A., Wintjens, R., Bégard, S., Sergeant, N., Delacourte, A., Horvath, D., Landrieu, I., Lippens, G., and Buée, L. (2002) *J. Mol. Neurosci.* 19, 275–287.
43. Dolinski, K., Muir, S., Cardenas, M., and Heitman, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 13093–13098.
44. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) *Science* 283, 1325–1328.
45. Andreotti, A. (2003) *Biochemistry* 42, 9515–9524.
46. Shen, M., Stukenberg, P. T., Kirschner, M. W., and Lu, K. P. (1998) *Genes Dev.* 12, 706–720.
47. Wulf, G. M., Ryo, A., Wulf, G. G., Lee, S. W., Niu, T., Petkova, V., and Lu, K. P. (2001) *EMBO J.* 20, 3459–3472.
48. Kondratick, C. M., and Vandre, D. D. (1996) *J. Neurochem.* 67, 2405–2416.
49. Bayer, E., Goettsch, S., Mueller, J. W., Griewel, B., Guiberman, E., Mayr, L. M., and Bayer, P. (2003) *J. Biol. Chem.* 278, 26183–26193.
50. Jacobs, D. M., Saxena, K., Vogtherr, M., Bernado, P., Pons, M., and Fiebig, K. (2003) *J. Biol. Chem.* 278, 26174–26182.
51. Lippens, G., Wieruszeski, J.-M., Leroy, A., Smet, C., Sillen, A., Buée, L., and Landrieu, I. (2004) *ChemBioChem* 4, 100–105.

BI035479X

# Accepting its Random Coil Nature Allows a Partial NMR Assignment of the Neuronal Tau Protein

Caroline Smet,<sup>[a]</sup> Arnaud Leroy,<sup>[a, b]</sup> Alain Sillen,<sup>[a]</sup> Jean-Michel Wieruszeski,<sup>[a]</sup> Isabelle Landrieu,<sup>[a]</sup> and Guy Lippens<sup>\*[a]</sup>

*A combined strategy to obtain a partial NMR assignment of the neuronal Tau protein is presented. Confronted with the extreme spectral degeneracy that the spectrum of this 441 amino acid long unstructured protein presents, we have introduced a graphical procedure based on residue type-specific product planes. Combining this strategy with the search for pairwise motifs, and*

*combining the spectra of different Tau isoforms and even of peptides derived from the native sequence, we arrive at a partial assignment that is sufficient to map the interactions of Tau with its molecular partners. The obtained assignments equally confirm the absence of regular secondary structure in the isolated protein.*

## Introduction

The physiological function of the neuronal Tau protein involves the dynamic stabilization of tubulin into microtubules.<sup>[1,2]</sup> Hence it plays an important role in neuronal development and plasticity. Moreover, the same microtubule-associated Tau protein is one of the major components of the neurofibrillary paired helical fragments (PHFs) characteristic of Alzheimer's disease (AD).<sup>[3,4]</sup> Primarily composed of Tau, these tangles accumulate intracellularly in the diseased neurons, and their presence correlates rather well with the cognitive impairments that result from the disease.<sup>[5]</sup>

Despite its enormous biological and medical relevance, little is known about the detailed structure of Tau in solution. Small-angle X-ray scattering (SAXS) and other spectroscopic techniques such as circular dichroism (CD) and infrared spectroscopy (IR) have led to a description of the protein as a random coil or Gaussian polymer.<sup>[6,7]</sup> However, as these techniques only give an overall view of the global structure, local elements of secondary and/or tertiary structure might well escape observation. The absence of a stable tertiary structure precludes the use of X-ray crystallography, but, as an alternative, NMR spectroscopy might well give hints about otherwise undetected sequence specific structural elements. Although this latter technique has been successfully applied in the case of several other unfolded proteins,<sup>[8-10]</sup> the literature contains no detailed report about the NMR characterization of Tau.

We have recently shown that proline residues and their distinct influence on the random coil chemical-shift values can be used to assign at least those residues in Tau that precede a proline amino acid.<sup>[11]</sup> In order to extend the assignment to all residues irrespective of their direct neighbours, we demonstrate here the utility of a novel graphical method to handle the huge information content in the triple-resonance spectra of Tau. Pushing the comparison of the spectra of full-length adult Tau and its shorter foetal isoform to the limit of a comparison with natural-abundance spectra of short peptides,

allows us to proceed further in the assignment process. We demonstrate the assignment of Tau to such a level that i) the presence or absence of well-defined structural elements can be ascertained, and ii) the interaction with other molecular partners such as tubulin or aggregating agents can be studied.

## Results

Conventional assignment strategies that use triple resonance experiments, correlate the C $\alpha$ , CO or C $\beta$  of a given residue with that of its immediate neighbour.<sup>[12]</sup> A typical example of this strategy is the use of HNCA and HN(CO)CA experiments, where the C $\alpha$  of residue [i] is seen from its own [i] amide cross peak in the first experiment and from the amide correlation of the next residue [i+1] in the HN(CO)CA. When applied to the full-length Tau protein, spectral overcrowding that results from i) the sheer size of the protein (Figure 1), ii) the limited chemical-shift dispersion that results from the unfolded nature of the protein and iii) the high degeneracy in the primary sequence (Table 1) all made this strategy largely unsuccessful. The only nucleus that gives a good chemical-shift dispersion is the nitrogen, as can be seen from the HSQC spectrum (Figure 2). We have previously observed that all Glycine C $\alpha$  carbons, which represent a total of 49 residues (Table 1), fall within 0.5 ppm of their random-coil value.<sup>[13]</sup> We found the same extreme de-

[a] C. Smet, Dr. A. Leroy, Dr. A. Sillen, Dr. J.-M. Wieruszeski, Dr. I. Landrieu, Dr. G. Lippens  
CNRS—Université de Lille 2, UMR 8525, Institut Pasteur de Lille  
B.P. 245, 59019 Lille Cedex (France)  
Fax: (+33) 320-871-233  
E-mail: guy.lippens@pasteur-lille.fr

[b] Dr. A. Leroy  
Laboratoire de Biochimie Appliquée  
Faculté de Pharmacie à Châtenay-Malabry (Paris XI)  
Tour D4 2<sup>ème</sup> étage, 5 rue Jean-Baptiste Clément  
92296 Châtenay-Malabry Cedex (France)

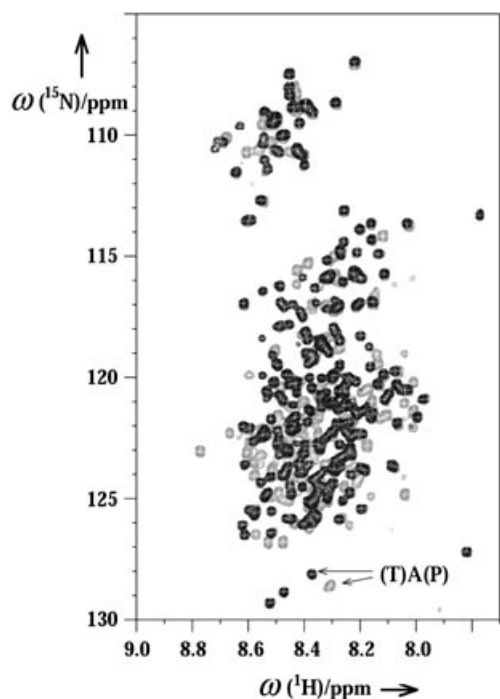


MAEPRQEFEV MEDHAGTYGL GDRKDQGGT MHQDQEGDTD 40  
 AGLKESPLQT PTEDEGSEEPG SETSDAKSTP TADVDTAPLV 80  
 DEGAPGKQAA AQPHTEIPES TTEEAIGD TPLEDEEAA 120  
 HVTQARMVSK SKDTTSSDDK KAKGADGKTK IATPRGAAPP 160  
 GQKGOANAIR IPAKTPPAPK TTPSGEPPK SGDRSGSSP 200  
 GSPGTPGSRG RTPSLTPPT REPKKVAVVR TPKSPSSAK 240  
 SRLQTAPVPM PDLKNYSKI GSTEILKHQP GGGVQIINK 280  
 KLDLNVQSK CGSKDNIKHV PGGGSVQIVY KPVLSVTS 320  
 KCGSLGNHH KPGGGQVEVK SEKLDKFDKRV QSKIGSLDN 360  
 THVPGGGNKK IETHKLTFRE LAKAKDHGA EIVYKSPVVS 400  
 GDTSPRHLSN SSTSIDMV DSPQLATLAD EVSASLAKOGL 441

**Figure 1.** Primary sequence of Tau. The amino acids that are absent from the shortest Tau352 isoform are in bold. The peptides synthesized for the peptide mapping are underlined. Residues that are unambiguously assigned are shown in yellow and those with only a twofold uncertainty in green. Light blue indicates the S/T(P) sites that were previously assigned.<sup>[11]</sup>

**Table 1.** Sequence degeneracy for the full-length Tau441 protein. Only the ten most abundant amino acids are listed.

Amino acid	Number	Percentage
Gly	49	11.1
Ser	45	10.2
Lys	44	10.0
Pro	43	9.8
Thr	35	8.0
Ala	34	7.7
Asp	29	6.6
Glu	27	6.1
Val	27	6.1
Leu	21	4.8



**Figure 2.**  $^1\text{H}, ^{15}\text{N}$  HSQC spectrum of the adult isoform Tau441 (light) and the smaller foetal isoform Tau352 (dark). The two (T)A(P) correlation peaks are indicated by an arrow.

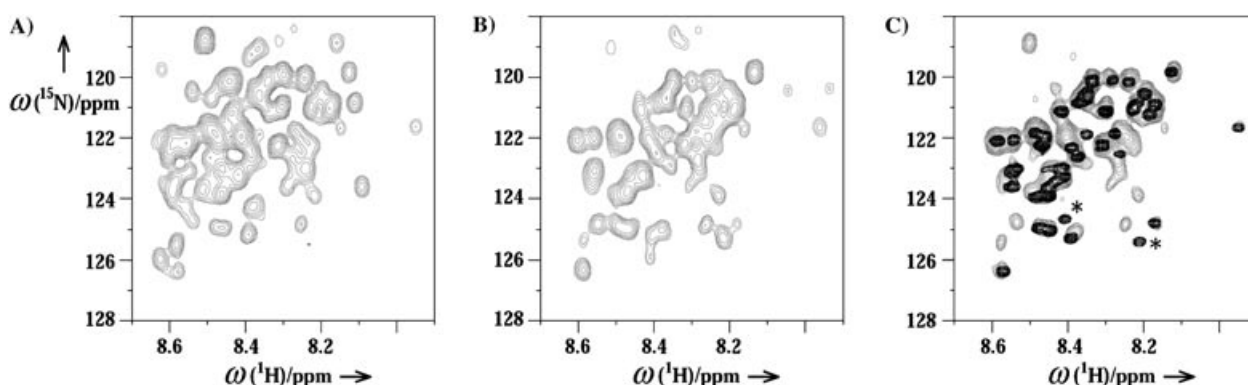
generacy for all other carbons including the carbonyl frequencies.<sup>[11]</sup>

The absence of carbon chemical-shift dispersion can, however, be used constructively once we accept the random-coil chemical-shift values from the onset on to be correct. Indeed, it means that we not only know the carbon chemical-shift values of every residue, but that we can also utilize them to identify those same residues, in a manner similar to amino acid selective labelling. As the random-coil values can be equally used to extract the nature of the residue in the  $[i-1]$  position, we have to deal with a pairwise degeneracy. In the following paragraphs, we will demonstrate that this degeneracy is less stringent than initially thought, but we will first illustrate our graphical strategy based on product planes and pattern searching on the concrete example of the assignment of lysine residues in Tau.

### Lysine assignment

Lys is characterized by a  $\text{C}\alpha$  chemical-shift value of 56.2 ppm and a  $\text{C}\beta$  value of 33.1 ppm.<sup>[13]</sup> Therefore, if we extract the corresponding planes for the  $\text{C}\alpha$  carbon chemical shift from a HNCA spectrum and for the  $\text{C}\beta$  carbons from the CBCANH spectrum, two planes result whose intersection should contain the Lys resonances. To illustrate this observation, we show in Figure 3A the sum of all  $^1\text{H}, ^{15}\text{N}$  HNCA planes with carbon chemical-shift values between  $(56.2-0.5 = 55.7)$  ppm and  $(56.2+0.5 = 56.7)$  ppm, which allow a deviation of 0.5 ppm from the listed random-coil value based on the carbon histograms that we previously derived for Tau.<sup>[11]</sup> A similar procedure for the planes with  $\text{C}\beta$  values comprised between 32.6 and 33.6 ppm (Figure 3B) leads to the two basic planes that define the Lys residues. Rather than superimposing both planes, however, we define the Lys product plane as the  $^1\text{H}, ^{15}\text{N}$  plane resulting from a point-by-point multiplication of both Lys  $\text{C}\alpha$  and  $\text{C}\beta$  defining planes (Figure 3C). The result of this operation is that we only recover intensity where both carbon chemical-shift values fall within the correct interval, thereby simplifying the visual information of both parent planes. For example, the peak at (8.09–123.5) ppm, present in the Lys  $\text{C}\alpha$  plane, is absent from the Lys product plane, because of its different  $\text{C}\beta$  value (30.3 ppm, as extracted from the CBCANH spectrum).

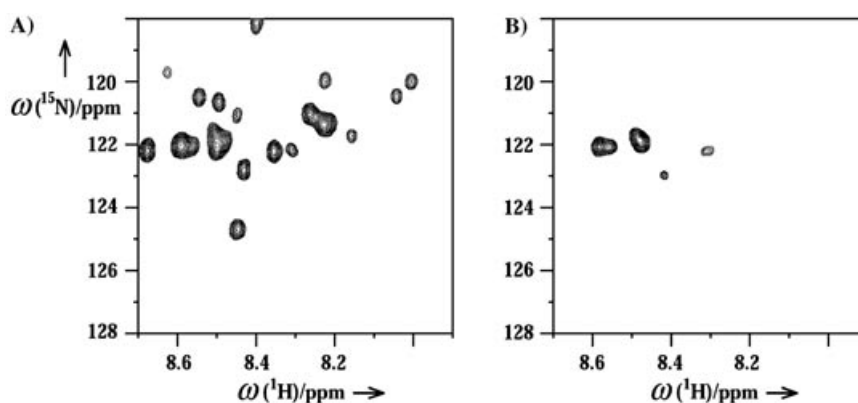
Lysine residues allow an immediate experimental verification of our assignment strategy since a selective Lys labelled sample is relatively simple to produce.<sup>[14]</sup> A comparison of the Lys product plane with the experimental spectrum of the Lys labelled sample is illustrative of the limitations of our method (Figure 3C). First, the majority of the experimental lysine peaks are present in the product plane. The two peaks at (8.21, 125.4) ppm and (8.41, 124.7) ppm that violate this principle, concern Lys residues that precede a Pro and therefore undergo a 2 ppm deviation towards the high field.<sup>[13]</sup> Secondly, the peaks that do coincide between the product plane and the spectrum of the Lys labelled sample are not all of equal intensity. Because the intensity in the product plane directly results from the intensity in the HNCA and CBCANH plane, an obvious



**Figure 3.** A) Plane around the Lys C $\alpha$  random coil chemical shift of 56.2 ppm, extracted from the Tau441 HNCA spectrum and B) around its C $\beta$  value of 33.1 ppm, extracted from the CBCANH spectrum. C) Product plane (light) defined by the point-by-point multiplication of the planes in (A) and (B), with superimposed (dark)  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum of the lysine, selectively labelled in Tau441.

reason for the intensity variations is the less efficient transfer of magnetization in one or both experiments. A deviation of the assumed random coil chemical-shift value might be a second unrelated cause of intensity variations, as the parent planes were constructed as slices around the random-coil values, such that a deviation superior to 0.5 ppm would result in a reduced or absent intensity of the peak. A third observation is that the product plane contains more signals than just the expected Lys correlation peaks. This is the case for the strong correlation found at (8.50, 118.8) ppm, absent from the spectrum of the Lys-labelled sample. Close inspection of the underlying frequencies shows that this residue is a Ser followed by a Pro, leading to a 2 ppm upfield shift of its C $\alpha$  random-coil chemical shift from 58.3 to 56.4 ppm.<sup>[11,13]</sup> Moreover, the residue is preceded by a Lys, and the efficient magnetization transfer from the Ser nitrogen to the preceding Lys C $\alpha$  (and hence its C $\beta$ ) leads to some intensity at the Lys C $\beta$  position in the CBCANH spectrum as well. Whereas specific experiments favouring the intraresidue transfer over the inter-residue one have been developed to simplify the triple-resonance spectra<sup>[15,16]</sup> and would avoid artefacts such as the one described above, the individual subspectra generated by our product-plane approach contain substantially less information than the initial HSQC spectrum; this brings manual inspection within reach. Finally, a close look at the random-coil chemical shift tables reveals that Met residues, equally, might be represented by the intersection of both planes, because the upper limits for both the C $\alpha$  (55.4+0.5=55.9 ppm) and the C $\beta$  (32.9+0.5=33.4 ppm) intersect with the Lys-defining carbon intervals. This overlap further accounts for the observed "leakage" of correlation peaks in the Lys product plane.

These latter observations may appear to cast some doubt on the utility of the product plane approach, but we claim that our method represents a true advantage when one considers that the 3D HNCOCa and CBCACONH experiments can be used in an identical way to extract product planes defining the nature of the residue in the  $[i-1]$  position. In Tau, there are four Pro-Lys patterns. The Pro  $[i-1]$  product plane can readily be constructed from the sum of planes around 63.3 ppm in the HNCOCa and 32.1 ppm in the CBCACONH spectra, and shows all amide signals from those residues that are preceded by a Pro (Figure 4A). If we now form the product plane of the initial Lys product plane of Figure 3C with the Pro  $[i-1]$  plane of Figure 4A, we obtain the  $^1\text{H}$ ,  $^{15}\text{N}$  subspectrum of those Lys residues that are preceded by a Pro (Figure 4B). There are four correlations as expected, demonstrating the validity of the strategy.



**Figure 4.** A) Product plane defining those residues that have a Pro in the  $[i-1]$  position. B) The product of the Lys plane of Figure 3C and the Pro  $[i-1]$  plane of Figure 4A defines the resonance position of the four (Pro) Lys pairs in Tau.

The described procedure does not allow us to go further in the identification of these (Pro)Lys patterns, as we do not connect more than two consecutive residues. However, despite the extreme degeneracy in amino acids for Tau, its sequence does contain 79 unique pairwise patterns, and another 49 residue pairs that occur only twice in the protein. These statistics

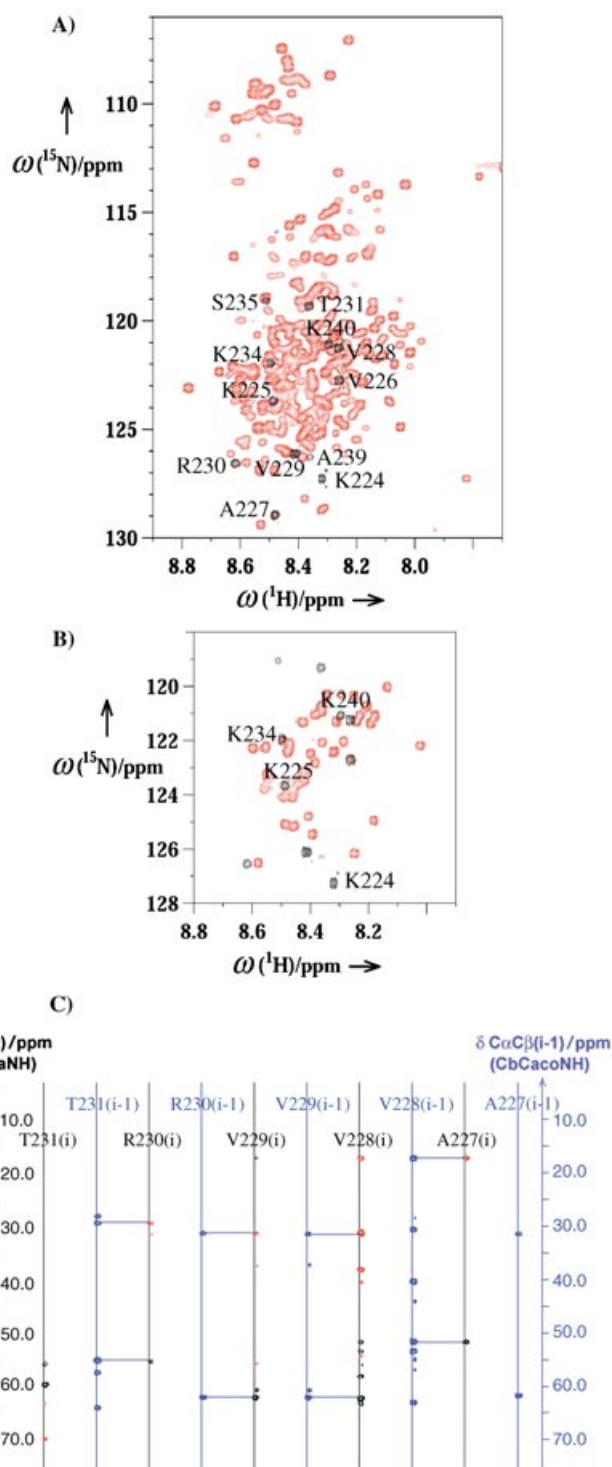
show that the product planes conjugated with a pairwise pattern search, can lead to an acceptable level of assignment without having to go through extensive lists of assignments.

### Alternative splicing

Tau occurs in human neurons in six isoforms. The longest isoform (Tau441) is mainly present in adult brain, whereas the shortest isoform, with 352 amino acids (Tau352), is dominant in foetal brain.<sup>[17]</sup> The two proteins differ in two inserts in the N-terminal region, absent in the shortest isoform, and in the absence of the second microtubule binding repeat for the foetal form (Figure 1). Based on the presumed absence of secondary or tertiary structure, we expected the NMR spectrum of the shorter isoform to be similar to that of the longer Tau441, but with fewer cross peaks and some border effects for those residues that immediately flank the missing regions in Tau352. A comparison of the resulting HSQC spectra of both proteins (Figure 2) confirms this, with the Tau352 HSQC roughly appearing as a subspectrum of that of Tau441. This observation allows us to solve some ambiguities. We have previously reported the assignment of all 5 Ala-Pro motifs in Tau,<sup>[11]</sup> but for the two (Thr)Ala(Pro) residues, the assignment was not unique, especially as both Thr residues have degenerate C $\alpha$  and C $\beta$  chemical shifts. One of the (Thr)Ala(Pro) motifs, however, centred around Ala77 (where we consistently use the Tau441 isoform numbering), is in the N-terminal insert that distinguishes Tau352 from Tau441. In the spectrum of Tau352, the peak at (8.30, 128.5) ppm is absent (Figure 2), allowing us to assign it this time without any ambiguity to Ala77, whereas the other Ala residue at position 246 necessarily has its correlation at (8.37, 128.1) ppm. Other similar examples exist, as of the 49 patterns that occur twice in the Tau sequence, 19 are absent in the Tau352 sequence, and hence become unique. This removes the ambiguity for those 19 pairs, or for approximately another 10% of the protein sequence.

### Splicing to the peptide level

Other isoforms of Tau exist in the brain,<sup>[17]</sup> and could potentially be used in a similar way as Tau352 to assist the assignment. However, we decided to push this strategy to the limit of small peptides, thereby circumventing the need to produce isotope-labelled samples. Several peptides of 13–17 amino acids (Figure 1) were synthesized by solid-phase synthesis. The peptides were solubilised at a concentration of 4 mM in the same buffer as was used for the protein spectra, and standard DQF-COSY, TOCSY and NOESY spectra were recorded for the assignment. The absence of secondary structure elements was verified by the measurement of coupling constants and analysis of the NOE pattern. Coupling constants were consistently found in the range of 6–7 Hz, and except for the sequential H $\alpha$ (i)–H $\beta$ (i+1) contacts, no other NOE cross peaks of significant intensity were detected. Heteronuclear  $^1\text{H}$ ,  $^{15}\text{N}$  correlation spectra at natural abundance were recorded overnight by using a cryoprobe on a 600 MHz instrument and peaks were assigned on the basis of the proton frequencies derived from the homonu-



**Figure 5.** A) Superposition of the uniformly or B)  $^{15}\text{N}$ -Lys labelled Tau441 HSQC spectrum (red) with the same spectrum on the Lys224–Lys240 peptide (black). C) Strips extracted from the CBCANH (black for positive and red for negative contours) and CbCacoNH (blue) spectrum of  $^{15}\text{N}/^{13}\text{C}$  labelled Tau441 confirming the assignment of the Thr231–Ala227 stretch.

clear TOCSY spectra. The HSQC spectra of the Lys224–Lys240 peptide and of full-length Tau show that peptide resonances indeed coincide for the most part with amide signals of the full-length protein (Figure 5 A). Because both Arg230 and

Ala227 are unique (Val)Arg and (Val)Ala motifs had already been unambiguously assigned, we could immediately verify that those two amino acids give the same signal when embedded in the peptide or full-length protein. The correlation peaks of the first two residues of the peptide, Lys224 and Lys225, did not coincide with the equivalent signals in the protein, as verified on a HSQC spectrum of a Lys selectively labelled sample (Figure 5B), leading to the conclusion that border effects extend to at least two residues. Further verification of the other resonances in the crowded region of the full-length protein spectrum was done on the basis of strips extracted from the CBCANH and CBCA(CO)NH spectra on the doubly labelled Tau441 (Figure 5C). The C-terminal part of the peptide following Pro236 was characterized by resonance broadening of the amide signals, hindering the assignment of the two Ser residues.

Whereas the good agreement between peptide and protein spectra was verified for several other peptides (Figure 1), we did experience solubility problems with the Gly303–Lys317 fragment. This peptide contains the PHF<sub>6</sub> motif previously defined to be the nucleus of the aggregation because of its strong interaction with the proteolytic PHF<sub>43</sub> fragment, which also forms fibres in solution.<sup>[18]</sup> The synthesis of the peptide had not given particular problems, but we did not succeed well in dissolving it in our standard buffer. Based on the signal intensity of the NMR signals in the supernatant, we estimate that only 200  $\mu$ M peptide remains in solution, whereas the rest

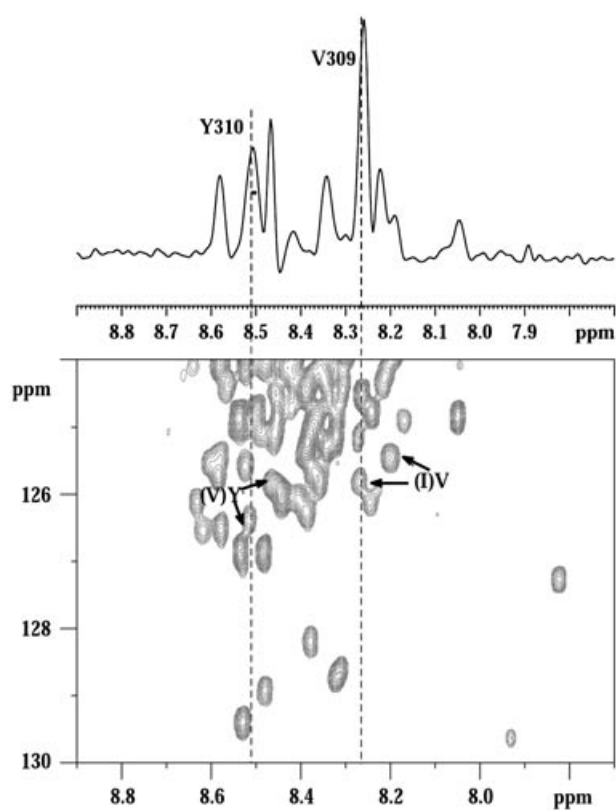
readily precipitates. This concentration, though not sufficient for a <sup>1</sup>H,<sup>15</sup>N HSQC at natural abundance, did allow the recording of the classical homonuclear spectra. Based on the TOCSY and NOESY spectra, we could assign the proton frequencies of the amide protons and in Figure 6, we show the assignment of the Val309 and Tyr310 amide protons on top of the HSQC spectrum of <sup>15</sup>N labelled Tau441. Our pairwise motif search had not succeeded in unambiguously assigning both residues, as exactly the same motif is found later in the sequence (Ile393–ValTyrLys396). However, for both residues, two candidate correlation peaks had been identified in the HSQC spectrum (Figure 6). The H<sub>N</sub> chemical-shift values of Val309 and Tyr310 in the peptide spectrum indicate that the peaks at (8.24, 126.1) ppm and (8.51, 126.4) ppm correspond to the two residues.

## Discussion

The increasing number of natively unfolded proteins<sup>[19]</sup> poses two challenges to structural biology, one fundamental—if the protein is natively unstructured, what is the meaning of a structural biology effort, at least at the level of the isolated biomolecule?—and one methodological—these proteins tend to resist crystallization efforts, and NMR spectroscopy is hampered by a reduced chemical dispersion. Despite these issues, the biological and medical importance of a large number of these proteins has motivated a substantial effort in obtaining information regarding their conformation and dynamics. An important class of natively unfolded proteins is found to be involved with neurological disorders, such as  $\alpha$ -synuclein in Parkinson's disease and Tau in Alzheimer's disease. Whereas an NMR effort on the former  $\alpha$ -synuclein has given a glimpse of the dynamical effect of certain disease-related mutations,<sup>[20]</sup> there is a lack of detailed information on Tau. This protein plays an important role in the dynamic process of tubulin polymerization into microtubules, with four repeat regions of some 30 amino acids, dubbed the MT binding repeats, that are directly involved in this interaction.<sup>[2,17]</sup> Besides its normal physiological role and despite the high solubility of the isolated protein, Tau also aggregates to form the PHF motif, a hallmark of AD that directly correlates to the severity of the cognitive impairments in the patients.<sup>[5]</sup>

Because of the importance of the biological and medical phenomena associated with Tau, a large effort has gone into its structural characterization. However, all macroscopic methods have concluded the absence of regular secondary structure.<sup>[6,7]</sup> Selected studies on isolated microtubule binding domains or short peptide fragments derived from these domains have suggested a tendency for them to adopt an extended  $\beta$ -strand like structure, although this conformation seems only adopted after the initial stages of aggregation.<sup>[18]</sup> Still, the structure of these same peptides within the full protein remains largely unknown.

Since the amide proton dispersion of Tau is limited to less than 1 ppm, it is an unambiguous confirmation that the protein is macroscopically unfolded, but this lack of proton chemical shift dispersion is simultaneously one of the main reasons



**Figure 6.** A) Proton 1D spectrum of the Gly303–Lys317 peptide and B) corresponding zone of the Tau441 HSQC. Candidate peaks for both (Ile)Val and (Val)Tyr are indicated.

that NMR studies of Tau have been devoid of success. A second obstacle for its NMR study is the sequence degeneracy of Tau, where five amino acids make up for almost 50% of its full length (Table 1). Moreover, the  $^{13}\text{C}$  chemical-shift values of these residues are extremely limited, leaving only the nitrogen chemical shift as element of spectral dispersion (Figure 2). Therefore, despite the fact that triple resonance experiments are easily collected (relaxation properties of the protein are very favourable despite its large size, and TROSY<sup>[21,22]</sup> techniques do not lead to any detectable signal enhancement), the traditional triple-resonance NMR strategy of connecting through the combined  $\text{C}\alpha$ ,  $\text{C}\beta$  and CO carbon values a given residue with its next neighbour is of limited value.

Despite these discouraging considerations, the lack of carbon chemical shift dispersion and their coincidence with random-coil values can be used in a constructive way if one adopts them directly as the true values. When we first applied this to those residues that are followed by a proline, where a 2 ppm shift towards the high field is expected<sup>[13]</sup> we were able to assign many of them, although some ambiguity persisted for double of triple pairs.<sup>[11]</sup> Here, we have gone a significant step further in the assignment process of Tau441. Clearly, the basic hypothesis of our previous work—the random coil chemical-shift values apply for all residues over the whole sequence—is maintained. The direct implication is that we know the residue type directly from the  $\text{C}\alpha$  and  $\text{C}\beta$  values. Nevertheless, because the information content of the spectra is so overwhelming, and inspired by the colocalization process commonly employed in confocal microscopy, we introduced a visual procedure in the form of product planes, centred directly around the random-coil values. The advantage of this approach is to directly obtain subspectra of the original  $^1\text{H},^{15}\text{N}$  HSQC map, precluding the need to go through an entire list of peaks to find those that match the required residue type both at the level of their  $\text{C}\alpha$  and  $\text{C}\beta$  chemical-shift values. Potential pitfalls are that residues could be missed if their carbon chemical-shift values fall outside the specified windows, or because magnetization transfer in the  $\text{C}\alpha$  or  $\text{C}\beta$  directed experiment has been hampered by relaxation losses. The experimental approach of residue specific labelling does not suffer the same limitations, but residue specific labelling requires several samples and special efforts to overcome isotopic scrambling. The product plane approach can be performed for all residues starting from a single set of triple resonance experiments recorded on a unique sample, but the equivalent of spectral scrambling cannot be prevented, as we have shown with the example of Lys and Met. A less obvious advantage of the experimental residue specific labelling is that the resulting spectrum can be acquired with any resolution that the operator decides, simply by collecting more increments in the indirect dimension. The product planes, on the other hand, are limited by resolution in the indirect nitrogen dimension of the corresponding 3D experiments, which in most cases comprises substantially fewer than 128 complex points.

The main advantage of the product plane approach is the possibility to derive not only the residue type corresponding to a given correlation peak, but also the residue type of the

preceding residue from the carbonyl relayed triple-resonance spectra. Whereas this has been obtained for the  $[i]$  and  $[i-1]$  residue experimentally by the double labelling of an antibody Fc fragment,<sup>[23]</sup> the approach is clearly not applicable to all amino acid pairs. The assignment of doublets, of course, is equally a major part of the traditional triple-resonance strategy, but there, the main effort is to connect the correlation peaks further to obtain stretches of amino acid residues that are then mapped onto the protein sequence. Here, we did not attempt to go further, but used the fact that Tau441 contains 79 unique pairwise patterns and another 49 residue pairs that occur only twice in the protein; this provides a starting basis for at least a partial assignment of the protein.

Comparing the spectra of the full-length protein and its fragments to divide the assignment problem is clearly limited to the case of natively unstructured proteins. We have used the Tau352 spectrum for the elimination of cross peaks in those cases in which the initial assignment was ambiguous, as shown for the two (T)A(P) cross peaks in Figure 2. Other isoforms of Tau could be used in a similar way, but we have opted to divide the full-length Tau protein into short peptides, and recombine the HSQC spectrum of Figure 2 as a superposition of the HSQC spectra of the short peptides. Central to our approach is the statement that the same amino acid stretch in the isolated peptide or in the full protein should give an identical HSQC spectrum when the individual amino acids sample the same conformational space in both contexts. That this will actually be the case was not obvious at all, as especially the nitrogen chemical shift is extremely sensitive to the time-averaged spatial environment of the amide group. Strikingly, empirical prediction methods that use the peptide or database derived chemical shift parameters still have the most problems with nitrogen chemical shifts as well.<sup>[24,25]</sup> The consequence of this extreme sensitivity is that in the case of a good agreement, we can safely say that the amino acids indeed sample conformational space in a very similar way, be they embedded in a short peptide or in the full protein. Disagreements, on the other hand, indicate structural differences between the isolated peptide and the equivalent amino acid stretch in the full-length protein, and therefore point to regions in the protein that should be examined in more detail by other NMR methods.

Peptides should be sufficiently small allowing for their straightforward chemical synthesis, purification and NMR analysis. An additional reason for a limited length is that they should not adopt a stable structure (stressing our initial hypothesis that the protein is natively unfolded). They should not be too short, however, because we expect some spectral border effects for those residues that are at the N or C terminus of the peptide and, hence, do not have the same chemical environment as their equivalents in the full-length protein. Our present work indicates that these border effects generally do not exceed two residues at both termini. Because we had previously synthesized a number of peptides of 13 to 17 amino acids containing potential phosphorylation sites,<sup>[26]</sup> we first used these and extended them with some additional peptides from the N terminus or third microtubule binding repeat

(Figure 1). NMR analysis led to the conclusion of absence of stable secondary structure elements in all of them. The second stage of the procedure consists in the recording of  $^1\text{H},^{15}\text{N}$  HSQC spectra on these peptides in the same buffer conditions, pH and temperature as for the full-length protein. Present cryogenic probe heads<sup>[27,28]</sup> installed on a 600 MHz spectrometer enable such spectral recordings within one night, if the peptide concentration is in the millimolar range or larger. The graphical superposition of the peptide spectra with the full-length Tau HSQC completes the procedure and the triple-resonance spectra can be used to verify rather than determine the identical assignment of the correlation peaks. The superposition of the Lys224–Lys240 peptide and Tau441 spectra (Figure 5A) indeed shows a perfect agreement for the isolated Ala227 and Arg230 correlation peaks. The strips extracted for the Thr231–Ala227 stretch at the exact positions of the peptide resonances (Figure 5C) show that the sequential walk is indeed possible and thereby confirm that the superposition is valid. However, the CBCA(CO)NH strip extracted for Thr231 shows the difficulty that we would face if we had to use the same information for the *ab initio* assignment of full-length Tau. We indeed observe three  $\text{C}\alpha$  and two  $\text{C}\beta$  peaks, but have no indication how to combine them into meaningful pairs. This situation is even worse for the strip of Val228, the  $^1\text{H},^{15}\text{N}$  correlation peak of which is in the crowded central region of the HSQC spectrum.

How far has the approach described here brought us, and, more importantly, what have we learned from the obtained assignments? First, we have succeeded in the assignment of a significant fraction of the residues occurring in unique patterns. Moreover, some more residues could be assigned on the basis of their unique presence or absence in the Tau352 spectrum or from the peptide mapping. Taken together with the assignments of those residues that are followed by a proline, we presently cover about 30% of the sequence without any ambiguity, and a further 10% with only two choices for the assignment (Figure 1). Overall, these assignments are distributed quite evenly over the sequence, although certain regions such as the quadruply occurring Pro(Gly)<sub>3</sub> quartet clearly have not yet been assigned unambiguously. This extent of the assignment should allow us to probe the interactions of Tau with such molecular partners as tubulin or heparin at the level of the individual residues.

As for the attempt to detect stretches of residual structure, one can question whether our approach, which is based on the validity of the random-coil chemical-shift values, can give any indication of this, as we should miss those assignments corresponding to residues that are in a locally well-defined structure and therefore escape from their respective product plane. The same objection can be made on the second approach, where we assume an identical environment of a given residue in a short unstructured peptide or full-length protein. For this latter approach, though, a discrepancy between the peptide and protein spectrum is easily detected. For the product plane strategy, we actually have one specific example where we did note a clear discrepancy: the isolated peak at (7.82, 127.2) ppm is a Leu, based on the spectrum of a Leu se-

lective labelled sample (data not shown), but its carbon  $\text{C}\alpha$  and  $\text{C}\beta$  frequencies were found at 56.6 and 43.4 ppm. Because both values deviate substantially from the random-coil values of 55.1 and 42.4 ppm, the correlation peak does not show up in the Leu product plane, and we at first thought we had detected a “hotspot” of residual structure. Closer examination, however, revealed that the cross peak represents the C-terminal Leu441. As for the other resonances assigned, next to the Lys labelled sample we have produced two further selectively labelled samples with Leu and Val. The coincidence of the product planes with the HSQC spectra recorded on the respective selectively labelled samples, further demonstrates the validity of our initial assumption. An even more striking example is given by the (Ile)Ile278 doublet. Unique in the sequence and moreover present in the second repeat encoded by the alternatively spliced exon 10, this motif is part of the PHF<sub>6</sub>\* peptide that was previously shown to promote aggregation by the formation of  $\beta$ -structure.<sup>[29]</sup> Despite the compatibility of this peptide with an extended structure in the PHFs, the  $\text{C}\alpha$  and  $\text{C}\beta$  chemical-shift values of the second Ile278 are 60.9 and 38.7 ppm, which is less than 0.2 ppm from their random-coil value. The *a posteriori* justification of our conclusion that Tau lacks well-defined secondary-structure elements is that we have succeeded in the assignment of a substantial fraction of its NMR signals based on two methods that assume this lack of structure from the start. Subtle deviations from the random-coil carbon values, however, that do not exceed our 0.5 ppm limit do exist, and we are currently trying to interpret them in conjunction with the information content in the nitrogen chemical-shift values.<sup>[24,25]</sup> Furthermore, it should be noted that the resonances in the repeat regions are generally weaker in intensity than their counterparts in the N- or C-terminal region of the protein. A full relaxation analysis of this phenomenon is currently in progress, but it is already clear that those residues, despite their sampling of the whole conformational space that defines the “random coil” conformation, do sample this space on a slower time scale. Therefore, it will be interesting to see whether this factor of dynamics will equally contribute to the propensity of fibrilization, on top of the recently determined structural criteria that govern aggregation.<sup>[30]</sup>

In conclusion, we have demonstrated here that the basic assumption of the validity of the random coil chemical-shift values can be successfully exploited to assign, at least partially, this at first sight impossible spectrum. The introduction of product planes allows a graphical view of the different residues and can easily be applied to the preceding residues when one derives them from the carbonyl relayed triple-resonance spectra. The HSQC spectrum of the shortest Tau352 isoform is at least to first order a subspectrum of that of Tau441, although we cannot exclude that subtle conformational differences cause deviations beyond the border effects. Peptide mapping proved possible, and allowed us to ascertain that the same residues embedded in the full-length protein or in a short peptide sequence sample the in same conformational space. Importantly, our present level of coverage and the validity of the random-coil values confirm that, at the level of the individual amino acids, this important protein belongs to the



class of natively unfolded proteins. The results presented form the basis for future structural studies of disease-related mutations of Tau, and of its interactions with its diverse molecular partners.

## Experimental Section

**Expression and purification of the recombinant Tau protein:** Expression and purification of the uniformly labelled samples was as previously described.<sup>[7]</sup> Residue specific labelling was obtained by adding <sup>15</sup>N-labeled Lysine, Leucine or Valine (100 mg) to M9 medium (1 L) supplemented with the other unlabelled amino acids (50 mg) before induction. Induction delays were kept to 2 h, in order to minimize isotopic scrambling.

**Peptide synthesis and purification:** Peptides were synthesized by using a classical Fmoc strategy on a Pioneer™ peptide continuous flow synthesizer with intermediate capping after each coupling step. Cleavage of peptide from the resin and side-chain-protecting groups were performed by using TFA mixtures that contained either 2.5% triisopropylsilane or 1% triisopropylsilane/2.5% ethanedithiol for peptide containing Cys or Met residues. Crude peptides were then purified by RP-HPLC on a C18 Nucleosil column by performing a linear gradient of acetonitrile. Fractions were checked both by RP-HPLC and MALDI-TOF mass spectrometry. NMR samples of peptides were prepared at a concentration of 4 mM in the same buffer that was used for recombinant Tau protein.

**NMR spectroscopy:** All NMR spectra were recorded at 20°C on a 600 MHz Bruker DMX spectrometer (Bruker, Karlsruhe, Germany) equipped with a cryogenic triple resonance probe head with actively shielded z-gradient. Spectral windows were 7800 Hz for proton and 2000 Hz for nitrogen, centred at 4.8 and 115.2 ppm, respectively, and sampled with 1024 and 47 complex points. All spectra were recorded with 8 scans per increment. The 30.7 ppm C $\alpha$  window in the HNCA and HN(CO)Ca spectra was centred at 54.5 ppm and was sampled with 31 complex points. Similarly, the 20 ppm CO window was centred at 175.4 ppm and was sampled with 21 complex points. Finally, the 70.5 ppm combined C $\beta$  and C $\alpha$  window in the CBCANH and CBCAcoNH spectra was centred at 40.2 ppm and was sampled with 72 complex points. All spectra were processed in the SNARF program (F. van Hoesel, Groningen, the Netherlands), with 2048 complex points in the proton dimension and 512 complex points in the indirect nitrogen and carbon dimensions. The final 3D spectra were treated as a three-dimensional matrix, or, alternatively, a collection of <sup>1</sup>H,<sup>15</sup>N planes. The latter can be considered as two dimensional matrices  $A^k[i,j]$ , with  $i$  ranging from 1–2048 and  $j$  from 1–512, containing an index  $k$  corresponding to the carbon frequency of the plane. The summation plane  $S[i,j]$  is defined by  $S[i,j] = \sum A^k[i,j]$ , where the summation is over the indices  $k$  that are within the specified carbon range. Similarly, the point-by-point multiplication plane  $P[i,j]$  of two sum planes  $S^1$  and  $S^2$  is defined by  $P[i,j] = S^1[i,j] \times S^2[i,j]$ . Normalization of the intensity in the latter product planes was consistently done with respect to the C $\alpha$  sum plane.

The homonuclear peptide spectra for peptide assignment and structural verification were recorded with DQFCOSY or Water flip-back NOESY and TOCSY sequences, with mixing times of 500 ms and 69 ms for the latter. <sup>1</sup>H,<sup>15</sup>N HSQC spectra of the peptides at natural abundance were recorded with 256 scans per increment, and 256 points in the nitrogen dimension.

## Acknowledgements

We thank E. Diesis and A. Menet for help with the peptide synthesis, Dr. L. Buée (Lille, France) for insightful discussions about Tau and Alzheimer's disease, and Dr. J. Christodoulou (Cambridge, UK) for careful reading of the manuscript. The 600 MHz facility used in this study was funded by the Région Nord-Pas de Calais (France), the CNRS and the Institut Pasteur de Lille. A.S. is funded by a European Training and Mobility Grant (HPRN-CT-2002-00241).

**Keywords:** neuronal Tau protein • NMR spectroscopy • peptide mapping • protein folding • protein structures

- [1] M. D. Weingarten, A. H. Lockwood, S.-Y. Wy, M. W. Kirschner, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1858–1862.
- [2] M. Goedert, C. M. Wischik, R. A. Crowther, J. E. Walker, A. Klug, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 4051–4055.
- [3] A. Delacourte, A. Defossez, *J. Neurol. Sci.* **1986**, *76*, 173–186.
- [4] S. G. Greenberg, P. Davies, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 5827–5831.
- [5] P. V. Arriagada, J. H. Crowdon, E. T. Hedley-White, B. T. Hyman, *Neurology* **1992**, *42*, 631–639.
- [6] D. W. Cleveland, S. Y. Hwo, M. W. Kirschner, *J. Mol. Biol.* **1977**, *116*, 227–247.
- [7] O. Schweers, E. Schonbrunn-Hanebeck, A. Marx, E. Mandelkow, *J. Biol. Chem.* **1994**, *269*, 24290–24297.
- [8] D. Neri, M. Billeter, G. Wider, K. Wüthrich, *Science* **1992**, *257*, 1559–1563.
- [9] M. Hennig, W. Bermel, A. Spencer, C. M. Dobson, L. J. Smith, H. Schwalbe, *J. Mol. Biol.* **1999**, *288*, 705–723.
- [10] H. J. Dyson, P. Wright, *Adv. Protein Chem.* **2002**, *62*, 311–340.
- [11] G. Lippens, J.-M. Wieruszkeski, A. Leroy, C. Smet, A. Sillen, L. Buée, I. Landrieu, *ChemBioChem* **2004**, *5*, 73–78.
- [12] M. Ikura, L. E. Kay, A. Bax, *Biochemistry* **1990**, *29*, 4659–4667.
- [13] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR*, **1995**, *5*, 67–81.
- [14] L.-Y. Lian, D. A. Middleton, *Progr. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 171–190.
- [15] B. Brutscher, *J. Magn. Reson.* **2002**, *156*, 155–159.
- [16] A. Permi, A. Annala, *Progress Nucl. Magn. Reson. Spectrosc.* **2004**, *44*, 97–137.
- [17] L. Buee, T. Bussiere, V. Buee-Scherrer, A. Delacourte, P. R. Hof, *Brain Res. Rev.* **2000**, *33*, 95–130.
- [18] M. von Bergen, P. Friedhoff, J. Biernat, J. Heberle, E.-M. Mandelkow, E. Mandelkow, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5129–5134.
- [19] J. Dyson, P. Wright, *Curr. Opin. Struct. Biol.* **2002**, *12*, 54–60.
- [20] R. Bussell Jr., D. Eliezer, *J. Biol. Chem.* **2001**, *276*, 45996–46003.
- [21] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12366–12371.
- [22] R. Riek, K. Pervushin, K. Wüthrich, *Trends Biochem. Sci.* **2000**, *25*, 462–468.
- [23] K. Kato, C. Sautès-Fridman, W. Yamada, K. Kobayashi, S. Uchiyama, H. Kim, J. Enokizono, A. Galinha, Y. Kobayashi, W.-H. Fridman, Y. Arata, I. Shimada, *J. Mol. Biol.* **2000**, *295*, 213–224.
- [24] Y. Wang, O. Jardetzky, *J. Am. Chem. Soc.* **2002**, *124*, 14075–14084.
- [25] J. Meiler, *J. Biomol. NMR* **2003**, *26*, 25–37.
- [26] C. Smet, A.-V. Sambo, J.-M. Wieruszkeski, A. Leroy, I. Landrieu, L. Buée, G. Lippens, *Biochemistry* **2004**, *43*, 2032–2040.
- [27] P. Styles, N. F. Soffe, C. A. Scott, D. A. Cragg, F. Row, D. J. White, P. C. J. White, *J. Magn. Reson.* **1984**, *60*, 397–404.
- [28] P. Styles, N. Soffe, C. A. Scott, *J. Magn. Reson.* **1989**, *84*, 376–378.
- [29] M. von Bergen, S. Baghorn, L. Li, A. Marx, J. Biernat, E.-M. Mandelkow, E. Mandelkow, *J. Biol. Chem.* **2001**, *276*, 48165–48174.
- [30] F. Chiti, M. Stefani, N. Taddei, G. Ramponi, C. M. Dobson, *Nature* **2003**, *424*, 805–808.

Received: May 10, 2004

# Regulation of Pin1 peptidyl-prolyl *cis/trans* isomerase activity by its WW binding module on a multi-phosphorylated peptide of Tau protein

Caroline Smet<sup>a</sup>, Jean-Michel Wieruszeski<sup>a</sup>, Luc Buée<sup>b</sup>, Isabelle Landrieu<sup>a,\*</sup>, Guy Lippens<sup>a,\*</sup>

<sup>a</sup> UMR CNRS 8525 – Structure et fonction des Biomolécules, Institut de Biologie de Lille, Institut Pasteur de Lille, BP 245 – F-59019 Lille Cedex, France

<sup>b</sup> INSERM U422, Institut de Médecine prédictive et recherche thérapeutique, Place de Verdun, F-59045 Lille Cedex, France

Received 20 December 2004; revised 25 April 2005; accepted 3 June 2005

Available online 6 July 2005

Edited by Christian Griesinger

**Abstract** The WW module of the peptidyl-prolyl *cis/trans* isomerase Pin1 targets specifically phosphorylated proteins involved in the cell cycle through the recognition of phospho-Thr(Ser)-Pro motifs. When the microtubule-associated Tau protein becomes hyperphosphorylated, it equally becomes a substrate for Pin1, with two recognition sites described around the phosphorylated Thr212 and Thr231. The Pin1 WW domain binds both sites with moderate affinity, but only the Thr212–Pro213 bond is isomerized by the catalytic domain of Pin1. We show here that, in a peptide carrying a single recognition site, the WW module increases significantly the enzymatic isomerase activity of Pin1. However, with addition of a second recognition motif, the affinity of both the WW and catalytic domain for the substrate increases, but the isomerization efficacy decreases. We therefore conclude that the WW domain can act as a negative regulator of enzymatic activity when multiple phosphorylation is present, thereby suggesting a subtle mechanism of its functional regulation.

© 2005 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.

**Keywords:** NMR exchange spectroscopy; Tau protein; Pin1; WW binding module; Peptidyl-prolyl *cis/trans* isomerase

## 1. Introduction

Pin1, a prolyl *cis/trans* isomerase (PPIase) from the parvulin family [1,2], is essential for the correct progression of the cell cycle. Composed of an N-terminal protein–protein interaction module under the form of a short WW domain, and a C-terminal catalytic domain [3], the protein was found to interact via one or more phospho-threonine(serine)-proline dipeptide (phospho-Thr(Ser)-Pro) motifs with a large number of cell cycle regulators such as CDC25 [4,5], Cyclin D1 [6], Wee1 or Myt1 [7], and equally with other proteins such as c-Jun,  $\beta$ -cate-

nin or p53 that indirectly determine the progression of the cell through the cell cycle [8]. Whereas the precise mode of action of Pin1 still remains to be determined, two main mechanisms seem to exist. In the first mode, exemplified by its interaction with c-Jun, p53 or  $\beta$ -catenin, the interaction with Pin1 prevents degradation pathways and hence stabilizes the protein substrate [9–14]. Alternatively, with CDC25 being an example, Pin1 interacts with the regulatory unstructured part of the protein, and isomerization catalysis is thought to regulate the (de)phosphorylation of specific phospho-Thr-Pro motifs through the *trans*-specific PP2A phosphatase [15,16].

The recently described interaction between Pin1 and Tau, a neuronal microtubule associated protein that aggregates into Paired Helical Fragments in Alzheimer's diseased neurons, can be classified as the second type of interaction [15,17]. Aberrant reactivation of mitotic mechanisms leading to Tau phosphorylation and the generation of MPM-2 epitopes might be the link between this interaction with a cytoskeletal protein and the general role of Pin1 in the cell cycle [18,19]. However, because Tau in the tangles is in a hyperphosphorylated state, the initial description of a unique interaction site (pThr231–Pro232) was quite amazing [17]. We recently showed that other motifs, including the pThr212–Pro213 dipeptide that together with pSer214 forms the Alzheimer's disease specific AT100 epitope, equally can be targets for Pin1 [20]. In this latter study, we found that under identical experimental conditions, the WW domain has an even higher affinity for the motif around Thr212 than for the pThr231–Pro232 site. Equally, because the pThr212–Pro213 bond in the peptide can be found in the *cis* form in close to 10% of the peptides, we were able to observe the *cis/trans* isomerase activity of the catalytic domain [20]. Whereas this study challenged the unicity of the interaction, it still did not answer how the two domains of Pin1 interact in the case of the multiply phosphorylated substrate that Pin1 probably encounters in a cellular context. The aim of the present paper is to investigate this interaction for a 40 amino acid peptide of Tau that spans both pThr212 and pThr231 sites. Both binding to the WW domain and catalytic activity have been assessed, using the shorter peptides as independent controls. Whereas the binding event seems mainly independent of the presence of multiply phosphorylated sites, the catalytic activity of Pin1 seems subject to a more subtle regulation. The presence of the WW domain indeed stimulates the catalytic activity in the case of a mono-phosphorylated peptide, but negatively regulates this same activity in the case of a multiply phosphorylated target.

\*Corresponding authors. Fax: +33 320871233.

E-mail addresses: Isabelle.Landrieu@ibl.fr (I. Landrieu), Guy.Lippens@ibl.fr (G. Lippens).

**Abbreviations:** PPIase, peptidyl-prolyl *cis/trans* isomerase; phospho-Thr(Ser)-Pro, phospho-threonine(serine)-proline dipeptide; RP-HPLC, reverse phase-high pressure liquid chromatography; Pin1<sup>CAT</sup>, isolated Pin1 catalytic domain; Bn, benzyl; TFA, trifluoroacetic acid; EXSY, exchange spectroscopy



## 2. Materials and methods

### 2.1. Peptide synthesis and purification

The peptides were synthesized on a continuous flow synthesizer (Pioneer) using the classical Fmoc temporary-protection strategy. Selective phosphorylations were introduced by using Fmoc-Thr/Ser(PO(OBn)OH)-OH building blocks (Novabiochem). Cleavage of the peptidyl-resin and side chain protections was performed with a trifluoroacetic acid (TFA) solution containing 2.5% of triisopropylsilane and 2.5% of water. Purification of the crude peptide was realized by reverse phase-high pressure liquid chromatography (RP-HPLC) on a C18 nucleosil column equilibrated in 0.05% TFA aqueous solution. The peptides were eluted with a linear gradient of acetonitrile (0–40% acetonitrile in 40 min) at 6 mL/min. Homogeneity of fractions was analysed by RP-HPLC and matrix-absorption laser desorption ionization-time of flight mass spectrometry.

### 2.2. Protein expression and purification

Pin1 protein and Pin1<sup>WW</sup> isolated domain with a N-terminal histidine-tag fusion were produced in *Escherichia coli* BL21(DE3) strain, carrying the recombinant Pin1-pET15b plasmid (Novagen). The isolated Pin1 catalytic domain (Pin1<sup>CAT</sup>) with a N-terminal GST fusion protein was produced in the same strain carrying the recombinant Pin1<sup>CAT</sup>-pGEX6P plasmid (GE Healthcare). The recombinant proteins were produced as described [20].

### 2.3. NMR spectroscopy: peptide assignment, binding studies and exchange spectroscopy

NMR experiments were performed at 293 K on a Bruker DMX600 spectrometer (Bruker, Karlsruhe, Germany) equipped with a cryogenic probehead. Standard total correlation spectroscopy and nuclear overhauser effect spectroscopy experiments, with, respectively, 60 and 400 ms mixing times were obtained with 2 mM peptide solutions in 50 mM deuterated Tris, pH 6.4 (Cambridge Isotope Laboratories), 100 mM NaCl buffer at 293 K. For the titration experiments, <sup>15</sup>N-labeled Pin1 <sup>1</sup>H–<sup>15</sup>N HSQC spectra were recorded on a 200 μM protein sample in 50 mM deuterated Tris, pH 6.4, 100 mM NaCl and 5 mM dithiothreitol/EDTA at the same temperature. The solution of Pin1 was added to the appropriate amounts of lyophilized peptide aliquots. Chemical shift variations of Pin1 residues upon increase in Tau peptide concentration were used to determine the dissociation constant value for each peptide. Proton and nitrogen chemical shift variations were averaged to obtain a global Δδ value (Eq. (1)).

$$\Delta\delta(\text{ppm}) = \sqrt{\Delta\delta(^1\text{H})^2 + 0.2\Delta\delta(^{15}\text{N})^2} \quad (1)$$

<sup>1</sup>H–<sup>1</sup>H exchange spectroscopy (EXSY) spectra were recorded at various mixing time (50, 100, 200, 300, 400 ms) on Tau peptides at 3.33 mM concentration in a 25 mM deuterated Tris pH 6.4/50 mM NaCl/D<sub>2</sub>O buffer, in the presence of 0.033 mM (1%) Pin1 or Pin1<sup>CAT</sup>. The exchange peak related to Pro213 H $\alpha$  *cis*/H $\alpha$  *trans* interconversion and NOE peak corresponding to Pro213 H $\alpha$  *cis*/pThr212 H $\beta$  are superimposed (Fig. 2). However, because the latter NOE peak is very weak (as can be seen for the equivalent Pro232 H $\alpha$  *cis*/pThr231 H $\beta$  NOE cross peak in Fig. 2), its contribution was neglected, and both peaks were integrated together. The exchange rate  $k_{\text{exch}}$  (s<sup>-1</sup>) was calculated by fitting the theoretical curve given by Eq. (2) [21] to the experimental data.

$$\%[cis \rightarrow trans] = a \frac{1 - \exp(-(1 + 1/a)k_{\text{exch}} \times MT)}{1 + \exp(-(1 + 1/a)k_{\text{exch}} \times MT)} \quad (2)$$

where % [*cis* → *trans*], expressed as the volume of the exchange cross peak to the diagonal peak, corresponds to the fraction of molecules that undergoes from *cis* to *trans* conformation during the mixing time (MT), and *a* is the excess of *trans* over *cis* forms, determined on the basis of the integrated values of Table 1.

## 3. Results and discussion

### 3.1. Conformational characterization of the Tau [201–240] peptide

This peptide spans both Tau [208–221] and [224–240] peptides (Table 1), investigated in a previous study [20]. The solution structure of the long peptide was verified by high resolution NMR, in order to ascertain that extending the peptide chain to 40 residues did not alter the secondary structure encountered in both individual [208–221] and [224–240] peptides, and also to evaluate the distribution of proline conformational states. The good agreement of chemical shift patterns, both at the level of the amide and aliphatic protons, demonstrates that the large peptide can be considered as the independent union of both small peptides that compose it (see Figure S1 in Supplementary Materials). For the major fully *trans* conformation, only very weak chemical shift changes are observed for H<sub>N</sub> and H $\alpha$  protons, except for the S208, R209, T220, R221, K224 and K225 residues. These resonances are the two N- or C-terminal residues of the short peptides, and are as such submitted to a “border effect” due to modification of their chemical environment. Because the experimental pH (pH 6.4) is close to the pK<sub>a</sub> value of phospho-residues (pK<sub>a</sub><sup>P231</sup> = 5.9) small differences in pH could potentially affect resonances of phospho-residues. The small chemical shift change observed for the phosphorylated T231 amide and aliphatic protons and not for surrounding residues therefore was ascribed to a small pH variation between this sample and the one containing the long peptide. Absence of H<sub>N</sub>–H<sub>N</sub> or other long-range NOE contacts, as well as average H<sub>N</sub>–H $\alpha$  *J* coupling values, further indicated that this long peptide, similarly to the shorter ones, does not adopt a stable structure in solution. Some minority forms corresponding to peptides containing at least one Pro in *cis* conformation can equally be observed in the spectrum of the long peptide, and seem conserved upon peptide elongation. In the aliphatic region of the 2D spectra (Fig. 2), we can distinguish particularly the *cis* forms of P213 and P232, whose integration of H $\alpha$  protons were used to estimate the percentages of both conformers (Table 1).

Table 1

Sequences of synthesized peptides of Tau protein, dissociation constant of Pin1 binding and proportion of *cis* conformations in each peptide for the Pro213 and/or Pro232 following the phosphoThr residue. Numbers in the first column indicates the first and the last residue of each peptide sequence (between brackets) and the position of the phosphorylation site (pT or pS for phospho-Thr and phospho-Ser)

Peptide	Phosphorylation sites	Sequence	K <sub>D</sub> (μM)	% <i>cis</i> conformation	
				Pro213	Pro232
[208–221]	pT212/pS214	S R S R pT P pS L P T P P T R	160	10	–
[224–240]	pT231	K K V A V V R pT P P K S P S S A K	380	–	3
[201–240]	pT212/pS214/pT231	G S P G T P G S R S R pT P pS L P T P P T R E P K K V A V V R pT P P K S P S S A K	70	5	4

### 3.2. Characterizing the complex formed by Pin1 and the Tau [201–240] peptide

Similarly as for the shorter peptides, the majority of the chemical shift perturbations when titrating the long peptide into a solution of full-length Pin1 enzyme were observed for residues of its WW domain, although chemical shift changes were equally detected for residues in the catalytic loop or at the interface of both domains (Fig. 1). In the spectral zoom of Fig. 1, the resonance of the Phe25 residue of the WW domain, the Ser115 residue in the active site of the isomerization domain, and the Ser147 residue that is at the interface between the WW and catalytic domain all are affected by the interaction with the peptide. For the Phe25 residue, the chemical shift variations induced by increasing amounts of peptide go in different directions when the pThr231 or the pThr212 peptide interact. With the long peptide, however, we obtain an intermediate shift, indicating that both phosphorylated residues can interact (Fig. 1, upper panel).

As for the catalytic domain, the short pThr231 peptide only interacts very weakly. Indeed, even at a 10-fold excess of ligand, we only observe a slight broadening of the catalytic residues, without any significant shift (see Ser115 in the upper left panel of Fig. 1). On the contrary, the pThr212/pSer214 peptide does interact with the catalytic domain, although with a lower affinity than with the WW domain, as even a 10-fold excess

does not drive the system into saturation (see Ser115 in the upper middle panel of Fig. 1). From this chemical shift variation, we have derived a dissociation constant of 3.3 mM. For the long peptide, the shifts, although not as pronounced as for the pThr212/pSer214 peptide, go in the same direction, indicating a similar interaction with the pThr212, but the interaction is 4-fold stronger with a derived  $K_D$  of 0.8 mM.

As for the interface between both domains, our results confirm the observation by Jacobs et al. that the nature of the peptide influences the interface. Indeed, whereas we do not see any chemical shift changes for the interfacial Ser147 when titrating in the pThr231 peptide, the same residue shifts in a very similar manner when the pThr212/pSer214 peptide is present, be it on the short or the long peptide.

The pThr212/pSer214 and pThr231 sites when considered on individual short peptides bind the WW domain with dissociation constant values of 160 and 380  $\mu$ M, respectively [20]. In the same way, the dissociation constant of the long peptide was evaluated to 70  $\mu$ M (Table 1). Because the obtained values are still of the same order of magnitude, it suggests that the increased pool of phospho-epitopes carried by the long peptide leads to an apparently increased affinity of the WW domain for the long peptide, but equally allows to conclude to the absence of a truly positive cooperation between both Pin1 domains for substrate binding.

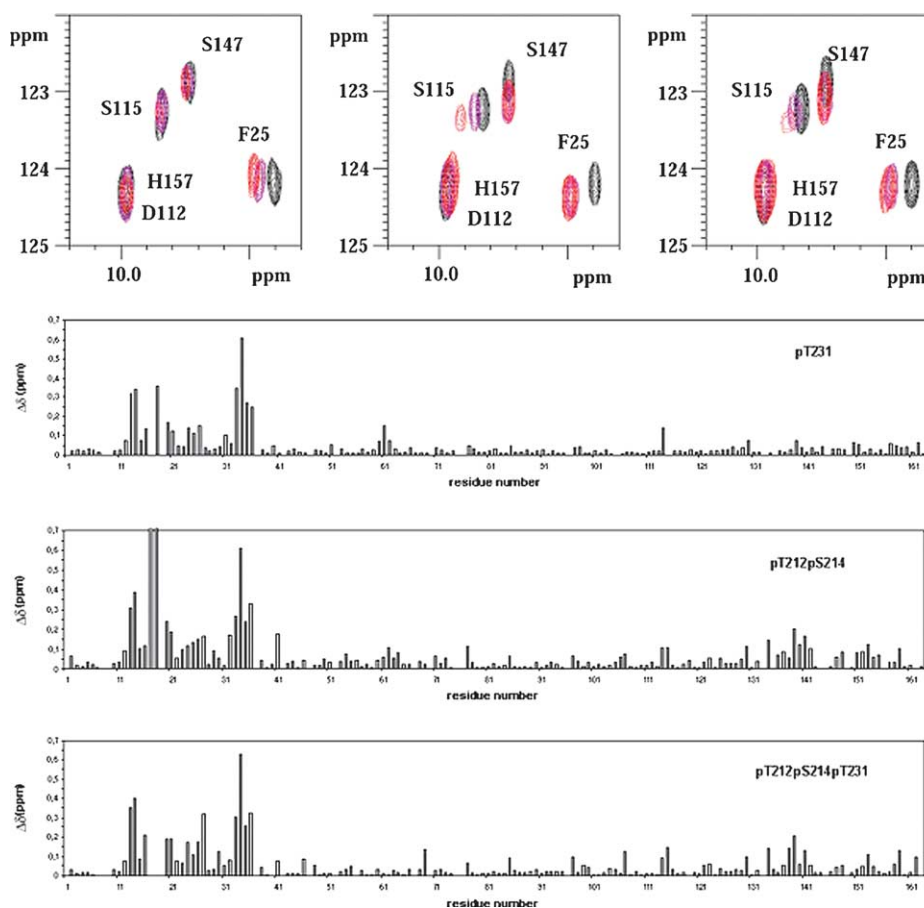


Fig. 1. (Top) Details of the HSQC spectra of  $^{15}$ N-labeled Pin1 without (black), with a 3:1 excess of peptide (violet) and a 10:1 excess of peptide (red), with a constant concentration of 0.2 mM Pin1 enzyme. Peptide substrates are [224–240] centered around pThr231 (left), [208–221] centered around pThr212/pSer214 (middle) or the long [201–240] peptide containing the pThr212/pSer214/pThr231 epitopes (right). (Bottom) Chemical shift maps of the three peptides along the Pin1 primary sequence.

### 3.3. Positive role of the WW domain in Pin1 catalytic activity on a short peptide substrate

Catalytic activities of Pin1 and Pin1<sup>CAT</sup> were evaluated with <sup>1</sup>H-<sup>1</sup>H EXSY NMR experiments by integration of the H $\alpha$  exchange peak related to P213 *cis/trans* isomerization as a function of mixing time (Fig. 2 and Table 2) [22]. No PPIase activity was however measured for phosphoThr231–Pro232 bond, possibly because the smaller population of the *cis* conformer in the small peptide [23] prevents detection by NMR.

In the case of the small [208–221] peptide carrying a single recognition motif, *cis/trans* isomerization catalysis proved to be more efficient in presence of the WW domain since the exchange rate is 8 times higher for full-length Pin1 ( $k_{\text{exch}} = 0.8 \text{ s}^{-1}$ ) than for Pin1<sup>CAT</sup> (Fig. 3). Interestingly, in the case of the first plant Pin1 homologue isolated from *Arabidopsis thaliana* [24] where the WW domain is absent, we had found that the sole catalytic domain of this enzyme, be it towards another peptide substrate, had a catalytic activity that was of the same order ( $k_{\text{exch}} = 1 \text{ s}^{-1}$ ) as found here for the intact human Pin1 protein, and therefore 10-fold higher than for the isolated catalytic domain of human Pin1. At the molecular level, it indicates a positive role of the WW domain on the human Pin1 function, and supports cellular data that have shown the essential biological role of the interaction module in vivo [4,25]. The Pin1 WW module could be implicated in anchoring Pin1 to its substrate to increase the local concentration of enzyme around the substrate molecule, although we cannot exclude that cross talk between both domains leads to an enhanced isomerase activity of the catalytic domain. Indeed, although both Pin1 domains interact only weakly together in solution [26], the relative inter-domain flexibility was found to be substrate-depen-

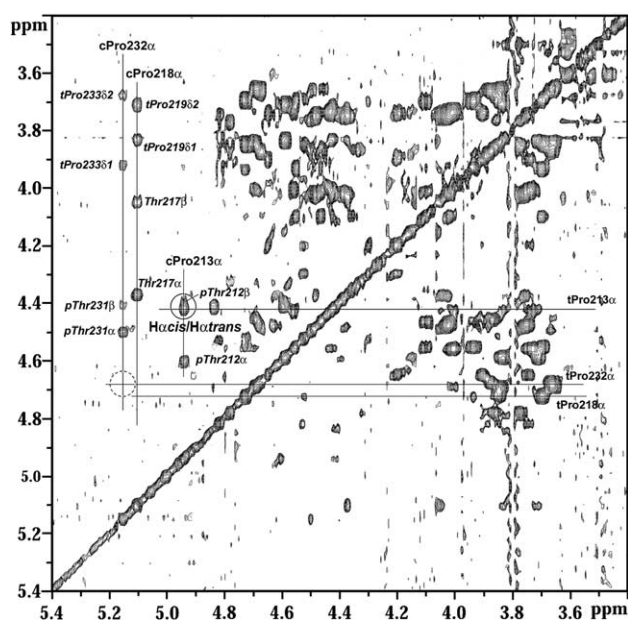


Fig. 2. Detail of the 400 ms <sup>1</sup>H-<sup>1</sup>H EXSY spectrum on the [201–240] large peptide at 3.33 mM in 100% D<sub>2</sub>O in presence of 1% Pin1. The *trans* (annotated by t) and *cis* (annotated by c) H $\alpha$  are assigned for Pro213, Pro218 and Pro232. The NOE correlation peaks between *cis* Pro H $\alpha$  and H $\alpha$ /H $\beta$  of pThr212, Thr217 and pThr232 is indicated in italic as well as *cis* Pro H $\alpha$  and H $\delta$ 1/H $\delta$ 2 of Pro 219 and Pro233 for Pro 218 and Pro232, respectively. Exchange peak between *cis* Pro213 H $\alpha$  and *trans* Pro213 H $\alpha$  is circled. The corresponding exchange cross peak for Pro232 is not observed, as indicated by a dashed circle.

Table 2

Exchange rates corresponding with the *cis* → *trans* isomerization, as measured from the EXSY spectra in the presence of full-length Pin1 or the sole catalytic domain (Pin1<sup>CAT</sup>)

Peptide	$k_{\text{exch}}$	
	Pin1 (s <sup>-1</sup> )	Pin1 <sup>CAT</sup> (s <sup>-1</sup> )
[208–221]	0.8	0.1
[208–221] + [224–240]	0.5	–
[201–240]	0.3	0.06

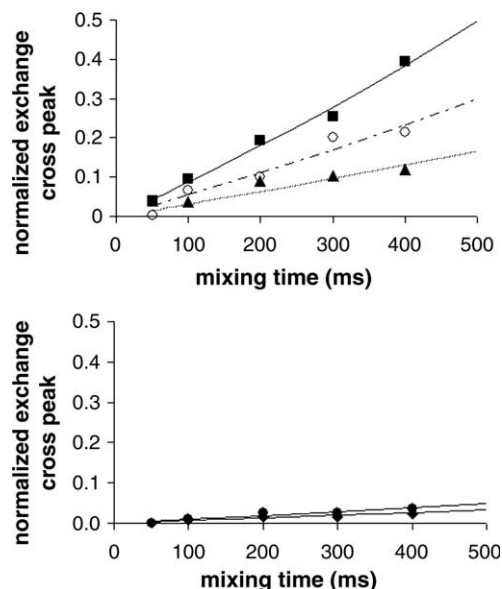


Fig. 3. Build up curves of the normalized exchange cross peaks of the Pro213 H $\alpha$  proton between the *trans* and *cis* forms in the EXSY series. (Top) Measure of catalytic activity of Pin1 on the [208–221] small peptide (■), on the [201–240] large peptide (▲) or on an equimolar mix of both [208–221] and [224–240] small peptides (○). (Bottom) Exchange curves with the sole catalytic domain on the [208–221] small peptide (●), or on the [201–240] large peptide (◆). Theoretical curves were derived from a fit of Eq. (2) to the experimental data points, with the derived exchange constants listed in Table 2.

dant [27]. In agreement with this latter observation, we observe more pronounced chemical shift perturbations at the domain-domain interface with the small peptide di-phosphorylated at Thr212 and Ser214 than for the peptides mono-phosphorylated at Thr231 (Fig. 1, lower panel) or Thr212 (data not shown).

### 3.4. Role of the WW domain in enzymatic regulation on a multi-phosphorylated substrate

Introduction of a second Pin1 binding site at the Thr231 position in the 40-residue peptide induces a remarkable effect on Pin1 catalytic activity. Indeed, the addition of the pThr231 site results in a significant decrease of Pin1 *cis/trans* isomerase activity around the phosphoThr212–Pro213 bond (illustrated by a plain circled peak, Fig. 2), as measured by EXSY spectroscopy with the long peptide as substrate. The deduced exchange rate ( $k_{\text{exch}} = 0.3 \text{ s}^{-1}$ ) is 3-fold lower than that measured for the same motif on the shorter peptide (Fig. 3, upper panel). Despite comparable populations of *cis* conformers (Table 1) for the Pro213 and Pro232 motifs, we did not de-



tect any catalysis around the phosphoThr231–Pro232 bond in the long peptide (that should be present at the location of the dotted circle in Fig. 2). Still, the diminished isomerase activity around the Thr212–Pro213 bond in the long peptide demonstrates that the phosphoThr231 site intrinsically interferes with the isomerase activity of the enzyme on the other site, probably through the enhancement of the binding that we described previously.

As a first control, we measured the Pin1 catalysis when both pThr231 and pThr212/pS214 sites are located on two different peptide chains, by performing the same EXSY experiment on an equimolar mixture of both short peptides. We observed an intermediate isomerase activity ( $k_{\text{exch}} = 0.5 \text{ s}^{-1}$ ) around the pThr212–Pro213 bond. The independent interaction of the pThr231–Pro232 peptide with the WW or catalytic domain might well decrease the available concentration of active Pin1 enzyme, but because isomerase activity is characterized by a further significant decrease when both sites are located on the same peptide chain (Fig. 3, upper panel), we conclude to a negative cooperativity of both sites when they are both phosphorylated on the same peptide chain.

As a second control, we evaluated the isomerase activity of the sole Pin1<sup>CAT</sup> catalytic domain on the long peptide. The catalytic activity, already low on the isolated pThr212–Pro213 bond, further dropped to an exchange rate of  $0.06 \text{ s}^{-1}$  for the *cis/trans* isomerization of the same bond in the long peptide substrate (Fig. 3, lower panel). In absolute terms, however, this change is much less significant than the one observed with the full-length Pin1, supporting a dominant role for the WW domain in mediating the cooperativity between both phosphorylation sites. A third control, where we have measured the isomerization activity of Pin1<sup>CAT</sup> in presence of an equimolar amount of WW domain on the short [208–221] di-phosphorylated peptide, reinforce this hypothesis. The catalysis of Pin1<sup>CAT</sup> in these conditions was as low as for Pin1<sup>CAT</sup> alone (data not shown) suggesting a cooperative role of both Pin1 domains to ensure the catalysis.

#### 4. Conclusion

Our data show that at the level of a single phosphorylation site, the Pin1 WW domain acts by targeting the enzymatic domain to its substrate, thereby increasing the efficiency of isomerization catalysis. However, in the context of a multiple binding sites on the same target as probably is the case in a cellular context, the mode of action is more complex. Adding a second phosphate moiety at the Thr231 position was found to increase the affinity of the ligand for Pin1, but to inhibit the isomerization activity around pThr212/pSer214 bond. This suggests a negative role for the WW binding module on the catalytic activity in a multi-phosphorylated substrate. Further research will be needed to assess its function in the context of an intact protein substrate.

**Acknowledgements:** The authors thank G. Montagne, H. Drobecq and E. Diesis for their technical contribution to this work. The 600 MHz facility used in this study was funded by the Région Nord-Pas de Calais (France), the CNRS, and the Institut Pasteur de Lille. Part of this work was funded by a grant from the Institute of Aging (ISOA, USA).

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.febslet.2005.06.048](https://doi.org/10.1016/j.febslet.2005.06.048).

#### References

- [1] Rahfeld, J.U., Rucknagel, K.P., Schelbert, B., Ludwig, B., Hacker, J., Mann, K. and Fischer, G. (1994) Confirmation of the existence of a third family among peptidyl-prolyl *cis/trans* isomerases. Amino acid sequence and recombinant production of parvulin. FEBS Lett. 352, 180–184.
- [2] Lu, K.P., Hanes, S.D. and Hunter, T. (1996) A human peptidyl-prolyl isomerase essential for regulation of mitosis. Nature 380, 544–547.
- [3] Ranganathan, R., Lu, K.P., Hunter, T. and Noel, J.P. (1997) Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. Cell 89, 875–886.
- [4] Lu, P.J., Zhou, X.Z., Shen, M. and Lu, K.P. (1999) The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. Science 283, 1325–1328.
- [5] Yaffe, M.B., Schutkowski, M., Shen, M., Zhou, X.Z., Stukenberg, P.T., Rahfeld, J.U., Xu, J., Kuang, J., Kirschner, M.W., Fischer, G., Cantley, L.C. and Lu, K.P. (1997) Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. Science 278, 1957–1960.
- [6] Liou, Y.C., Ryo, A., Huang, H.K., Lu, P.J., Bronson, R., Fujimori, F., Uchida, T., Hunter, T. and Lu, K.P. (2002) Loss of Pin1 function in the mouse causes phenotypes resembling cyclin D1-null phenotypes. Proc. Natl. Acad. Sci. USA 99, 1335–1340.
- [7] Shen, M., Stukenberg, P.T., Kirschner, M.W. and Lu, K.P. (1998) The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins. Genes Dev. 12, 706–720.
- [8] Ryo, A., Liou, Y.C., Lu, K.P. and Wulf, G. (2003) Prolyl isomerase Pin1: a catalyst for oncogenesis and a potential therapeutic target in cancer. J. Cell Sci. 116, 773–783.
- [9] Lu, K.P. (2004) Pinning down cell signalling, cancer and Alzheimer's disease. Trends Biochem. Sci. 29, 200–209.
- [10] Zacchi, P., Gostissa, M., Uchida, T., Salvagno, C., Avolio, F., Volinia, S., Ronai, Z., Blandino, G., Schneider, C. and Del Sal, G. (2002) The prolyl isomerase Pin1 reveals a mechanism to control p53 functions after genotoxic insults. Nature 419, 853–857.
- [11] Zheng, H., You, H., Zhou, X.Z., Murray, S.A., Uchida, T., Wulf, G., Gu, L., Tang, X., Lu, K.P. and Xiao, Z.X. (2002) The prolyl isomerase Pin1 is a regulator of p53 in genotoxic response. Nature 419, 849–853.
- [12] Ryo, A., Nakamura, M., Wulf, G., Liou, Y.C. and Lu, K.P. (2001) Pin1 regulates turnover and subcellular localization of beta-catenin by inhibiting its interaction with APC. Nat. Cell Biol. 3, 793–801.
- [13] Wulf, G.M., Ryo, A., Wulf, G.G., Lee, S.W., Niu, T., Petkova, V. and Lu, K.P. (2001) Pin1 is overexpressed in breast cancer and cooperates with Ras signalling in increasing the transcriptional activity of c-Jun towards cyclin D1. Embo J. 20, 3459–3472.
- [14] Wulf, G.M., Liou, Y.C., Ryo, A., Lee, S.W. and Lu, K.P. (2002) Role of Pin1 in the regulation of p53 stability and p21 transactivation, and cell cycle checkpoints in response to DNA damage. J. Biol. Chem. 277, 47976–47979.
- [15] Zhou, X.Z., Kops, O., Werner, A., Lu, P.J., Shen, M., Stoller, G., Kullertz, G., Stark, M., Fischer, G. and Lu, K.P. (2000) Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins. Mol. Cell 6, 873–883.
- [16] Stukenberg, P.T. and Kirschner, M.W. (2001) Pin1 acts catalytically to promote a conformational change in Cdc25. Mol. Cell 7, 1071–1083.
- [17] Lu, P.J., Wulf, G., Zhou, X.Z., Davies, P. and Lu, K.P. (1999) The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. Nature 399, 784–788.

- [18] Vincent, I., Rosado, M. and Davies, P. (1996) Mitotic mechanisms in Alzheimer's disease?. *J. Cell. Biol.* 132, 413–425.
- [19] Vincent, I., Zheng, J.H., Dickson, D.W., Kress, Y. and Davies, P. (1998) Mitotic phosphoepitopes precede paired helical filaments in Alzheimer's disease. *Neurobiol. Aging* 19, 287–296.
- [20] Smet, C., Sambo, A.V., Wieruszkeski, J.M., Leroy, A., Landrieu, I., Buee, L. and Lippens, G. (2004) The peptidyl prolyl *cis/trans*-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau. *Biochemistry* 43, 2032–2040.
- [21] Kaplan, J.L. and Fraenkel, G. (1980) *NMR of Chemically Exchanging Systems*, Academic Press, NY.
- [22] Kern, D., Drakenberg, T., Wikstrom, M., Forsen, S., Bang, H. and Fischer, G. (1993) The *cis/trans* interconversion of the calcium regulating hormone calcitonin is catalyzed by cyclophilin. *FEBS Lett.* 323, 198–202.
- [23] Daly, N.L., Hoffmann, R., Otvos Jr., L. and Craik, D.J. (2000) Role of phosphorylation in the conformation of tau peptides implicated in Alzheimer's disease. *Biochemistry* 39, 9039–9046.
- [24] Landrieu, I., De Veylder, L., Fruchart, J.S., Odaert, B., Casteels, P., Portetelle, D., Van Montagu, M., Inze, D. and Lippens, G. (2000) The Arabidopsis thaliana PIN1At gene encodes a single-domain phosphorylation-dependent peptidyl prolyl *cis/trans* isomerase. *J. Biol. Chem.* 275, 10577–10581.
- [25] Winkler, K.E., Swenson, K.I., Kornbluth, S. and Means, A.R. (2000) Requirement of the prolyl isomerase Pin1 for the replication checkpoint. *Science* 287, 1644–1647.
- [26] Bayer, E., Goettsch, S., Mueller, J.W., Griewel, B., Guiberman, E., Mayr, L.M. and Bayer, P. (2003) Structural analysis of the mitotic regulator hPin1 in solution: insights into domain architecture and substrate binding. *J. Biol. Chem.* 278, 26183–26193.
- [27] Jacobs, D.M., Saxena, K., Vogtherr, M., Bernado, P., Pons, M. and Fiebig, K.M. (2003) Peptide binding induces large scale changes in inter-domain mobility in human Pin1. *J. Biol. Chem.* 278, 26174–26182.

# Control of Protein–Protein Interactions: Structure-Based Discovery of Low Molecular Weight Inhibitors of the Interactions between Pin1 WW Domain and Phosphopeptides

Caroline Smet,<sup>†</sup> Jean-Frédéric Duckert,<sup>†</sup> Jean-Michel Wieruszeski,<sup>†</sup> Isabelle Landrieu,<sup>†</sup> Luc Buée,<sup>‡</sup> Guy Lippens,<sup>†</sup> and Benoît Déprez<sup>\*,†</sup>

CNRS-University of Lille 2 UMR 8525, Institut Pasteur de Lille, Molecular Drug Discovery, 3 rue du Professeur Laguesse, 59000 Lille, France, and Inserm U422, place de Verdun, 59000 Lille, France

Received January 6, 2005

Interactions involving phosphorylated Ser/Thr-Pro motifs in proteins play a key role in numerous regulatory processes in the cell. Here, we investigate potential ligands of the WW binding domain of Pin1 in order to inhibit protein–protein interactions between Pin1 and phosphopeptides. Our structure-based strategy implies the synthesis of analogues of the Ac-Thr(PO<sub>3</sub>H<sub>2</sub>)-Pro-NH<sub>2</sub> dipeptide and relies on high resolution NMR spectroscopy to accurately measure the affinity constants even in the high micromolar range.

## Introduction

Protein–protein interactions have an important role in most biological processes. A number of small domains present in various proteins with different biological functions are recurrent in protein–protein interactions and stand as such in the core of the cellular protein network.<sup>1</sup> Interactions mediated by phosphorylated moieties on their targets are of particular interest, since crucial regulation processes involve a balance between phosphorylation and dephosphorylation pathways allowed by kinase and phosphatase activities.<sup>2</sup>

Pin1 is a peptidyl-prolyl *cis/trans* isomerase constituted by two domains, a catalytic prolyl *cis/trans* isomerase domain and a WW binding module.<sup>3</sup> Both share a specificity for phospho-Thr/Ser-Pro motifs. At the cellular level, Pin1 is shown to be an essential protein involved in cell cycle regulation, since Pin1 inhibition prevents entry into mitosis and cell division.<sup>4,5</sup> Overexpression of Pin1 in numerous human cancers indicates its implication in several oncogenic pathways.<sup>6,7</sup> Therefore, Pin1 is proposed as a therapeutic target for cancer treatment.<sup>8</sup> Pin1 also interacts with the neuronal microtubule-associated tau protein when the latter is phosphorylated at specific motifs, and thereby restores its ability to polymerize the tubulin into microtubules.<sup>9</sup>

Inhibition of Pin1 activity with small molecules targeting the enzymatic domain blocks cell proliferation in cancer cell lines. Juglone was the first of this category of irreversible ligands, and the synthesis and characterization of several derivatives has been described.<sup>10,11</sup> Recently, conformationally constrained isosteres of phosphoSer-*cis/trans*-Pro were described in inhibition of *cis/trans* catalytic activity of Pin1.<sup>12</sup> The WW domain of Pin1, however, is equally essential for its *in vivo* activity.<sup>5,13</sup> Moreover, in healthy tissue, phosphorylation

of the WW Ser16 inhibits its binding role and inactivates the whole protein. In breast cancer cells, the WW domain is dephosphorylated, leading to the active form of the protein.<sup>14</sup>

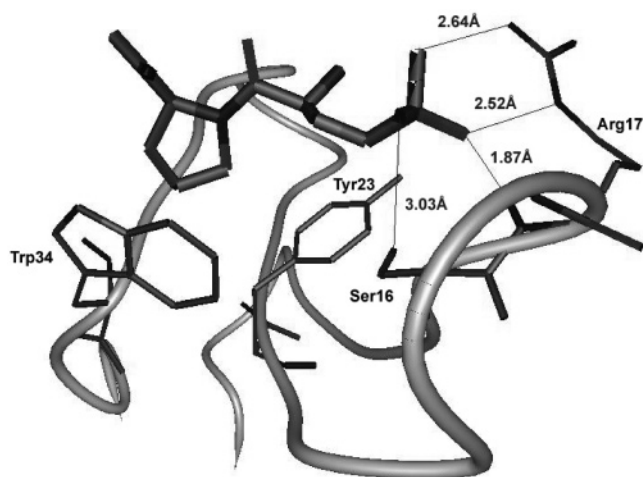
Structural data on Pin1 or isolated WW domain in complexes with phosphopeptides indicate an extremely limited interface between both partners, centered on phosphoThr-Pro dipeptide and a poor contribution of other sequential residues surrounding the dipeptide.<sup>15,16</sup> At the molecular level, the WW domain has a flexible loop encompassing Ser16-Ser19 residues that binds the phosphate moiety on Ser or Thr residues and an aromatic clamp constituted by its Tyr23 and Trp34 side chains that binds the prolyl residue in the +1 position of the phosphoSer/Thr (Figure 1).<sup>15,16</sup> These two points of interaction are of particular interest and are shared by other binding modules that display either the phospho specificity, distinctly for phosphoTyr (SH2,<sup>17</sup> PTB<sup>18</sup>) and phosphoSer/Thr (14–3–3 and FHA<sup>19</sup>), or affinity for proline-rich sequences like SH3 and other classes of WW domains (WW type I, II/III).<sup>20,21</sup> The phosphate moiety brings in strong hydrogen bonds toward the guanidine side chain and backbone amide groups, and is as such a real challenge to the development of potent mimetics.<sup>22</sup> On the other hand, proline displays a unique pattern among the 20 natural amino acids because its side chain is comprised in a five membered ring allowing a critical role in protein structuration (formation of  $\beta$ -turn or PP II helix in proline-rich peptides) or in protein–protein interactions.<sup>21</sup> Although steric and stereoelectronic effects of aryl-, alkyl-, and heteroatom-substituted prolines have been used to study factors that affect prolyl peptide conformation and activity,<sup>23</sup> the impact of such modifications on the recognition of the side chain itself is receptor dependent. Some examples have been described in the case of protease inhibitors.<sup>24,25</sup>

The present study is an exploration of both functional groups in the context of the Pin1 WW domain, using NMR spectroscopy throughout to evaluate the individual compounds.

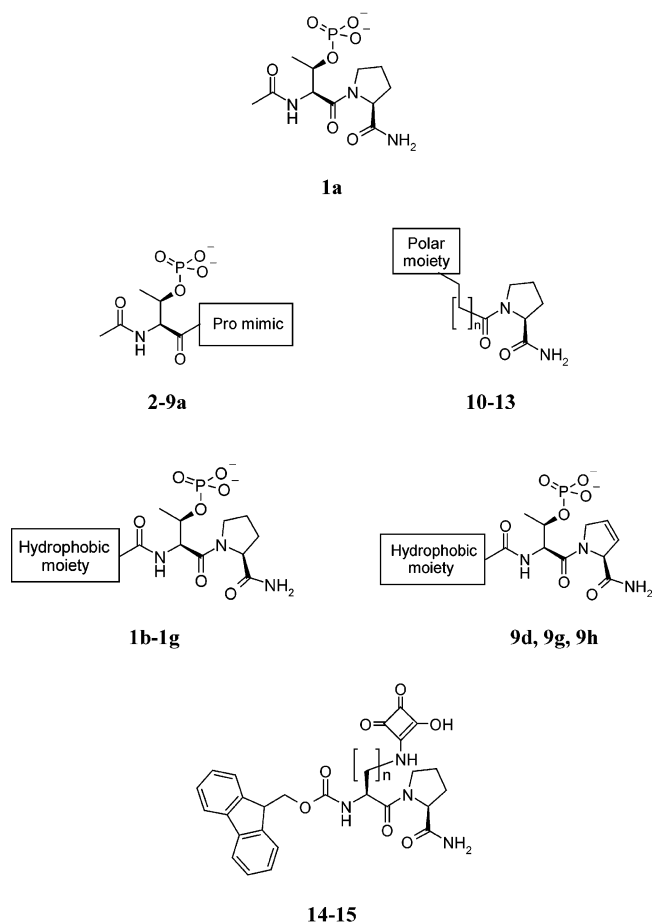
\* To whom correspondence should be addressed. Phone: +33(0) 320964024. Fax: +33(0) 320964709. E-mail: bdeprez@pharma.univ-lille2.fr.

<sup>†</sup> CNRS-University of Lille 2 UMR 8525.

<sup>‡</sup> Inserm U422.



**Figure 1.** Structure of the dipeptide Ac-Thr( $\text{PO}_3\text{H}_2$ )-Pro-NH<sub>2</sub> (**1a**) bound to the Pin1 WW domain. The WW domain is represented as ribbon and the substrate as thick sticks. Side chains of WW residues implicated in dipeptide binding are represented as thin sticks and annotated. Hydrogen bonds between phosphate and Ser16 hydroxyl group and Arg17 guanidinium function are indicated by lines and distances. The proline moiety is encompassed in an aromatic clamp constituted by side chains of Tyr23 and Trp34.



**Figure 2.** Structures of synthetic WW domain-binding analogues. Structure details of proline mimics and polar moieties are given in Table 1 and Table 2, respectively, and structures of hydrophobic moieties in Table 3.

The Ac-Thr( $\text{PO}_3\text{H}_2$ )-Pro-NH<sub>2</sub> dipeptide **1a** (Figure 2) was assessed by NMR as the minimal requirement for binding to the Pin1 WW module exhibiting a relatively

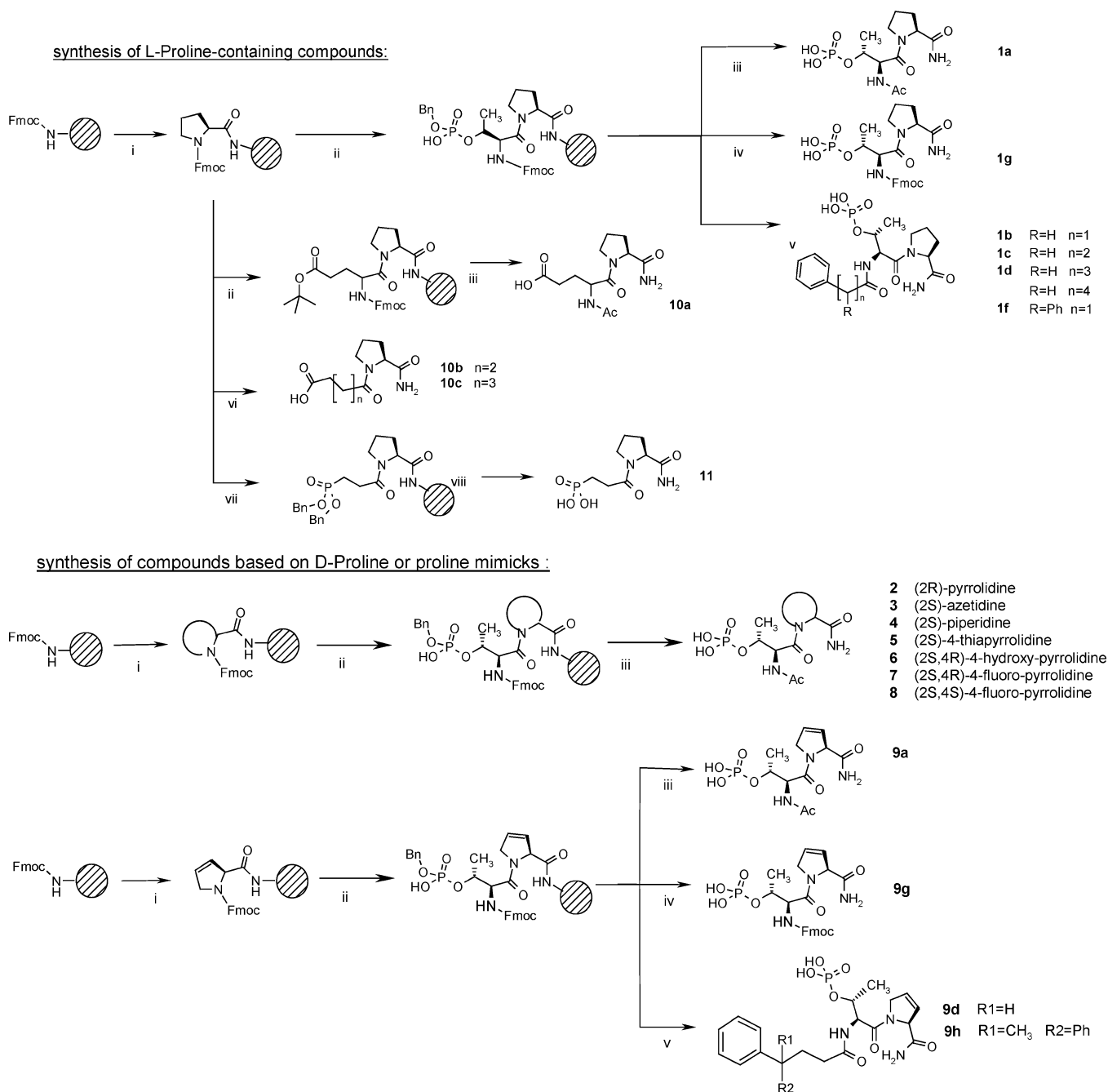
high dissociation constant of 150  $\mu\text{M}$  and a similar mode of binding as larger peptidic substrates that derive from cell cycle-related proteins like RNA Pol II<sup>15</sup> and Cdc25 or the neuronal tau protein.<sup>16,26</sup> On the basis of the molecular structures of the WW domain and the ligand (**1a**), we investigate here potential ligands for the Pin1 WW binding domain, by sequentially considering compounds that contain proline or phosphate mimics, and further extending these compounds with additional hydrophobic moieties (Figure 2). High-resolution NMR proved a sensitive assay even for very weak ligands,<sup>27,28</sup> with cryogenic probe technology limiting appropriately the amount of protein and ligand.<sup>29</sup> Moreover, structural information could be obtained at every step through a simple mapping of chemical shift variations of the WW domain.

## Results and Discussion

**Synthesis.** First, we have investigated separately the replacement of proline and the replacement of phosphothreonine in two series of molecules described in the template Figure 2: (i) proline mimics coupled to phosphothreonine through an amide bond in the compounds **2–9a** and (ii) polar moieties replacing phosphate supported by various backbone coupled to prolinamide in the compounds **10–13**. Then, introduction of hydrophobic moieties in compounds **1b–1g**, **9d**, **9g**, **9h**, **14**, and **15** was investigated in order to improve compound affinities for the WW module. Compounds **1–11** were synthesized by solid phase chemistry on polystyrene solid support functionalized by a rink amide linker (Scheme 1). The chemistry involved is very similar to solid phase peptide synthesis. Amine functions were protected by a Fmoc group and deprotected by suspending the resin in a 20% piperidine solution in DMF. Carboxylic acid activation was performed by a mixture of 1 equiv of HOBt/1 equiv of HBTU and 3 equiv of DIPEA in DMF. Cleavage of compounds from resin beads and deprotection of polar moiety (phosphate, carboxylate) were performed under high concentrations of TFA. Compounds **12–15** were synthesized by classical homogeneous phase synthesis by coupling amino acid derivatives of various chain lengths ( $\beta$ -alanine, ornithine, lysine) on prolinamide and functionalizing the amine group of these compounds with polar moieties (squarate, trifluoromethylsulfonyl) (Scheme 2).

**Pin 1 WW Domain Expression, Purification, and NMR Characterization.** The Pin1 WW binding module [Ala<sub>2</sub>–Asn<sub>40</sub>] was cloned in a pET28 plasmid fused to an N-terminal His<sub>6</sub> tag and overexpressed in *Escherichia coli* that allows uniform isotopic labeling of the protein for NMR studies. The protein was purified from the soluble cell lysate using two chromatography systems: chromatography based on nickel affinity and reverse phase high pressure liquid chromatography. The protein was obtained at a degree of purity superior to 95% (SDS–PAGE) and characterized by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry and high resolution NMR. After purification, the Pin1 WW domain spontaneously folded when dissolved in NMR buffer, and bound correctly to phosphorylated tau peptides as assessed by NMR or FRET experiments.<sup>26</sup> No detectable interaction was observed with nonphosphorylated peptides. Initially, the



Scheme 1. <sup>a</sup> General Scheme of Synthesis for Compounds 1–11

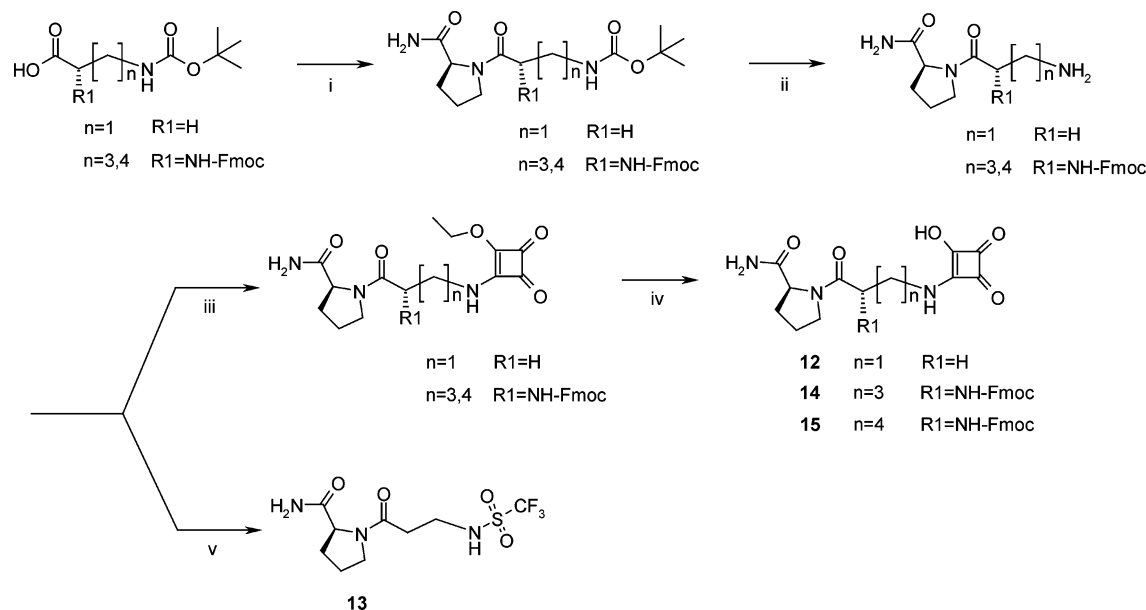
<sup>a</sup> Reagents and conditions: (i) (a) piperidine 20%, DMF, room temperature, 25 min; (b) *N*-Fmoc-(L)-Pro-OH or mimetics, HOBt, HBTU, 3 equiv of DIPEA, DMF, room temperature, 1 h; (ii) (a) piperidine 20%, DMF, room temperature, 25 min; (b) *N*-Fmoc-(L)-Thr(PO(OBn)OH)-OH or *N*-Fmoc-(L)-Glu(tBu)-OH, HOBt, HBTU, 3 equiv of DIPEA, DMF, room temperature, 1 h; (iii) (a) piperidine 20%, DMF, room temperature, 25 min; (b)  $\text{Ac}_2\text{O}$ , 3 equiv of DIPEA, DMF, room temperature, 30 min; (c) TFA 95%, TIS 2.5%, water 2.5%, room temperature, 5 h; (iv) TFA 95%, TIS 2.5%, water 2.5%, room temperature, 5 h; (v) (a) piperidine 20%, DMF, room temperature, 25 min; (b) carboxylic acid precursor, HOBt, HBTU, 3 equiv of DIPEA, DMF, room temperature, 1 h; (c) TFA 95%, TIS 2.5%, water 2.5%, room temperature, 5 h; (vi) (a) piperidine 20%, DMF, room temperature, 25 min; (b) glutaric anhydride or succinic anhydride, 3 equiv of DIPEA, DMF, room temperature, 1 h; (c) TFA 95%, TIS 2.5%, water 2.5%, room temperature, 2 h; (vii) (a) piperidine 20%, DMF, room temperature, 25 min; (b) 3-(dibenzoyloxyphosphoryl)-propionic acid, HOBt, HBTU, 2 equiv of DIPEA, DMF, room temperature, 6 h; (viii) (a) TFA 95%, TIS 2.5%, water 2.5%, room temperature, 5 h; (b)  $\text{H}_2$  at atmospheric pressure, Pd/C 10%, MeOH, room temperature, 48 h.

protein was assigned using conventional tridimensional NMR experiments (HNCA, HNCO, and HN(CO)CA) on a 500  $\mu\text{M}$   $^{15}\text{N}/^{13}\text{C}$ -labeled protein sample on a 600 MHz spectrometer (Bruker, Karlsruhe, Germany).

**Discovery of Novel Reversible Ligands for the Pin1 WW Binding Module.** In an initial phase, we screened compounds by rapid 1D NMR experiments at a 20-fold excess of ligand to WW domain. The WW

domain at 20  $\mu\text{M}$  was  $^{15}\text{N}$ -labeled, allowing the efficient filtering of the ligand resonances. The signal of the HN $\epsilon$  of the Trp34 (at 10.2 ppm) was used as a sensitive probe to detect perturbations of the Trp side chain chemical environment upon binding of the proline ring. When interaction was detected, we performed an additional titration study with a 50  $\mu\text{M}$   $^{15}\text{N}$ -labeled WW sample and increasing amounts of ligands (from 0.5 to 20 equiv).



**Scheme 2.** <sup>a</sup> General Scheme of Synthesis for Compounds **12**–**15**

<sup>a</sup> Reagents and conditions: (i) (a) DCC/DCM; (b) prolinamide/DCM, 0 °C, 2 h; (ii) HCl/dioxane, 0 °C then room temperature, 3 h; (iii) (a) NEt<sub>3</sub>/EtOH; (b) diethylsquarate/EtOH, room temperature, 12 h; (iv) HCl/DCM room temperature, 24 h; (v) (a) NEt<sub>3</sub>/DCM; (b) trifluorosulfonyl chloride/DCM, room temperature, 48 h.

Gradual chemical shift perturbations with increasing ligand concentrations were used to derive the dissociation constant (Figure 3).

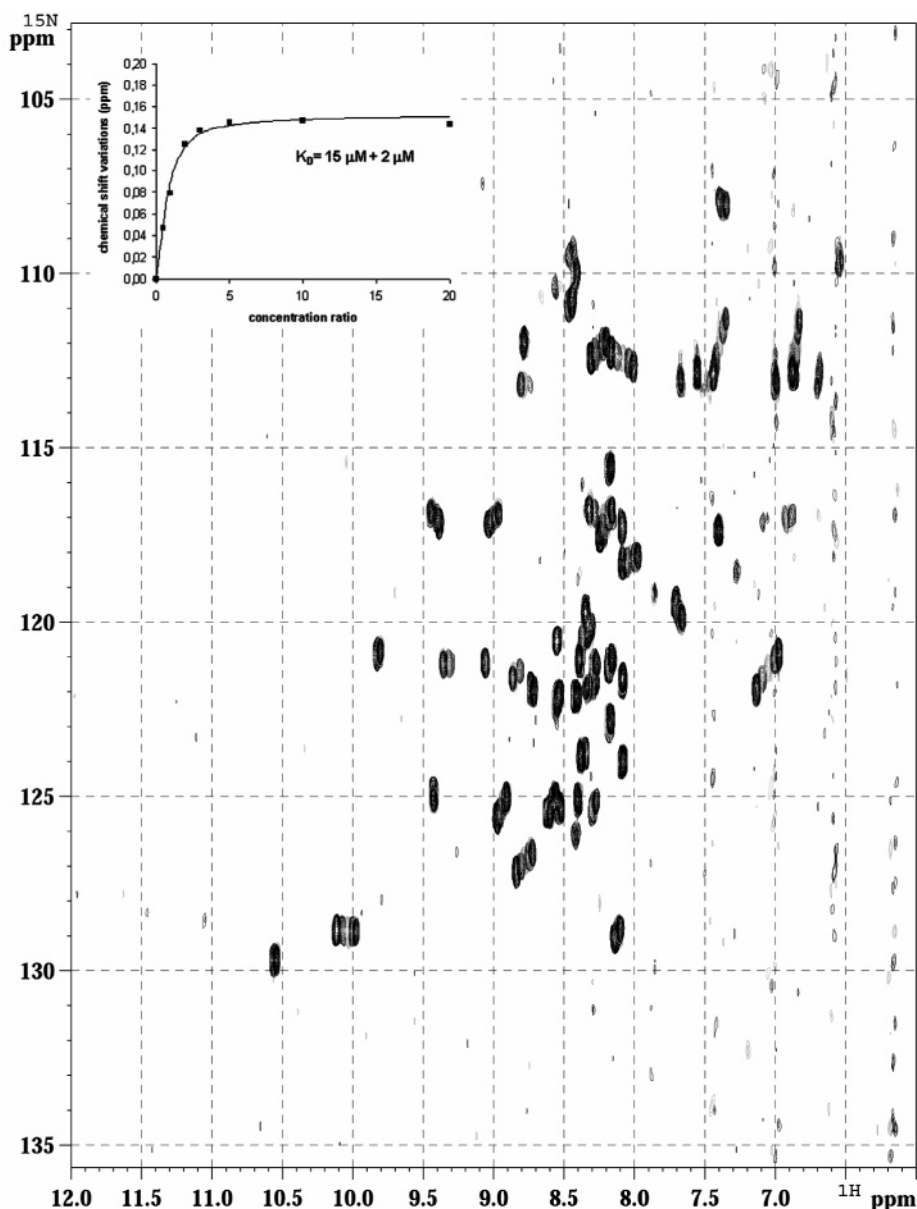
The aromatic clamp was found to accept Pro mimics with structural variations of the pyrrolidine ring. Despite the relatively low affinities, NMR titrations yielded reliable SAR results for analogues of Ac-Thr(PO<sub>3</sub>H<sub>2</sub>)-Pro-NH<sub>2</sub> (**2**–**9a**). Compounds where the pyrrolidine ring is replaced by an azetidine or a piperidine bind the WW domain without significant changes of affinity. The same observation is made for the epimer at the C $\alpha$  of proline. Substitutions have contradictory effects on binding since the replacement of proline by 4-*trans*-hydroxyproline, 4-*cis*-fluoroproline, or thiazolidine-4-carboxylic acid isostere decreases dramatically the affinity for the WW domain. However replacement of proline with a 4-*trans*-fluoroproline improves significantly the binding as well as the introduction of an unsaturation at position 3–4 of the pyrrolidine ring (Table 1). Contrary to the proline mimics, polar groups used to replace the phosphate moiety all exhibited dissociation constants as low as 2 mM, pointing out an extreme specificity of the flexible loop for phosphate binding (Table 2).

Faced with restricted possibilities for structural modifications around the two major points of interaction of dipeptide **1a** with the WW domain, we have examined further WW hydrophobic patches in the near proximity of the main binding site in order to gain in affinity. A potential hydrophobic surface constituted by Phe25 and Ala31 was targeted with a phenyl ring anchored to the Thr(PO<sub>3</sub>H<sub>2</sub>)-Pro dipeptide via aliphatic chains of various lengths in compounds **1b**–**1f** and with the Fmoc group in compound **1g**. A significant increase of binding affinity was observed for these compounds with a maximum for the 4-phenylbutyric acid precursor and for the Fmoc group that confirms the agreement with the 4-carbon-length chain (Table 3). Nevertheless, interactions with the phosphate binding loop and the

prolyl ring binding pocket remain essential and appear as a minimal requirement for ligands to adapt the domain since the replacement of phosphothreonine by a glutamic acid in dipeptides carrying the same hydrophobic moieties did not significantly bind the WW domain (data not shown).

**Consistency of Binding Mode in the Series.** Chemical shift interaction mapping was performed for each compound of significantly improved affinity, to verify the consistency of the binding mode. Very similar modes of binding were observed for the various proline mimics in comparison with the dipeptide **1a**. The 3,4-dehydroproline moiety induces larger chemical shift perturbations for both phosphate and proline binding sites, and especially a stronger effect on the Y23 amide peak (Figure 4A), together with a significant line broadening concerning most of the interacting residues. The binding site of the hydrophobic moiety that replaces the acetyl group in the phospho-Thr-Pro dipeptide is given by the chemical shift mapping of compounds **1b**–**1g**. Larger Ala31 amide shifts (Figure 4B) and a dependence of the Ala31 variations with the chain length carrying a phenyl group (**1b**–**1e**) that is not shared by Trp34 or Arg17 verify our initial structural hypothesis. It cannot be ruled out that observed chemical shift variations could also be due to conformational changes of the domain upon ligand binding. However, since little variations are observed at remote amino acids from the binding pocket, we reasonably make the assumption that, in the series, the conformation of the WW domain in its ligand bound state is not strongly affected.

Combining the data on chain length and nature of the aromatic group with the previously determined optimal proline mimics provides a further perspective for improving the affinity. The combination of the three essential motifs was further investigated in compounds **9d**, **9g**, **9h** on the basis of the 3,4-dehydroproline analogue (Figure 2). In the proline series (**1b**–**1g**), the



**Figure 3.** NMR titration of the  $^{15}\text{N}$ -labeled WW domain with the ligand **1g**. Overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra acquired on a  $50\ \mu\text{M}$  WW domain  $^{15}\text{N}$ -labeled sample with increasing amounts of compound **1g**. In the left top corner, a titration curve plotting the chemical shift variations against the concentration ratio **1g**:WW.

replacement of the acetyl group by  $\omega$ -arylalkanoyl chains or a Fmoc group increased affinity as expected. Surprisingly, except for 4,4-bis(4-hydroxyphenyl)pentanoyl (**9h**), most of the hydrophobic moieties induce a decrease of affinity (Table 3) in the dehydroproline series, suggesting different structural constraints in the free ligand of both series.

**Revisiting Phosphate Mimics on Improved Ligands.** The ligands containing an additional point of interaction with the WW domain through the hydrophobic moiety display affinities in the low micromolar range. This level of affinity allows us to revisit possible phosphate isosteres. Squaramide is singly charged at physiological pH and thanks to delocalization of this charge can accept strong hydrogen bonds at each of its oxygen atoms. Compounds carrying this group instead of the phosphate were synthesized, and their binding was studied by NMR. In the Fmoc series, compounds **14** and **15** displayed measurable affinities (respectively 910 and

540  $\mu\text{M}$ ), thereby providing new entry points in series where the phosphate moiety could be replaced by an isosteric group. This provides new perspectives for improvement of metabolic stability and cell membrane permeation.

**Perspectives in Disrupting Protein-Protein Interactions and Compound Activity in a Cell-Based Assay.** To confirm that our small ligands can actually disrupt interactions between the WW binding module and its proline-containing substrates, we have conducted competition experiments using fluorescence energy transfer. Our ligands were tested for their ability to displace a tetradecapeptide centered on the Thr<sup>212</sup>(PO<sub>3</sub>H<sub>2</sub>)-Pro<sup>213</sup> of Tau protein<sup>26</sup> from the WW domain. In this setup, our best ligands were shown to displace the peptide with  $K_i$  consistent with their  $K_d$  (data not shown).

We have tested some of our lead compounds in a cellular model of SY5Y neuroblastoma cells with induc-

**Table 1.** Structures of Proline Mimics and Dissociation Constants of Resulting Ligands (2–9a)

compound	Pro mimic structures	K <sub>D</sub> (μM)
2		100
3		160
4		140
5		400
6		800
7		36
8		420
9a		18

ible overexpression of Pin1. This cellular assay is based on the monitoring of cyclin D1 level as described.<sup>30</sup> In that model, juglone is used as a reference compound for the inhibition of Pin1 function and induces a marked diminution of the intracellular concentration of cyclin D1. Our results with WW ligands are at this moment very preliminary. Compound **1g** has produced the same effect as juglone at similar concentrations and poses the intriguing question whether it acts directly on Pin1 or also on other target(s). Further assays will be performed in the future to identify the cellular target(s) of our compounds.

## Conclusion

Our NMR-based strategy has led to the conception, synthesis, and testing of new ligands for the Pin1 WW domain. Whereas the proline in the initial phospho-Thr-Pro dipeptide could be successfully replaced by a 3,4-dehydro- or 4-*trans*-fluoroproline, the specificity for the phosphate moiety proved more challenging. On the basis of the structure of the WW domain, we extended these ligands with an N-terminal aromatic group, resulting in a new series of ligands with low micromolar affinity. NMR proved crucial throughout this development, as a technique to bridge the gap between very low affinity hit compounds and conventional tools to measure affinity of small molecules to protein, such as

**Table 2.** Structures of Polar Moieties and Dissociation Constants of Resulting Ligands (10–13)

compound	Polar moiety structures	K <sub>D</sub> (μM)
10a		> 2000
10b		> 2000
10c		> 2000
11		> 2000
12		> 2000
13		> 2000

fluorescence polarization or energy transfer. It allowed initial screening, and generated reliable SAR at low levels of affinity, as well as structural information on the binding mode of our ligands. The exact mode of action of our ligands is currently being studied.

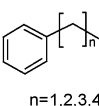
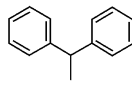
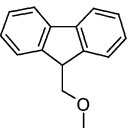
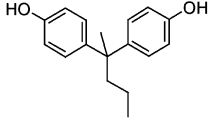
## Experimental Section

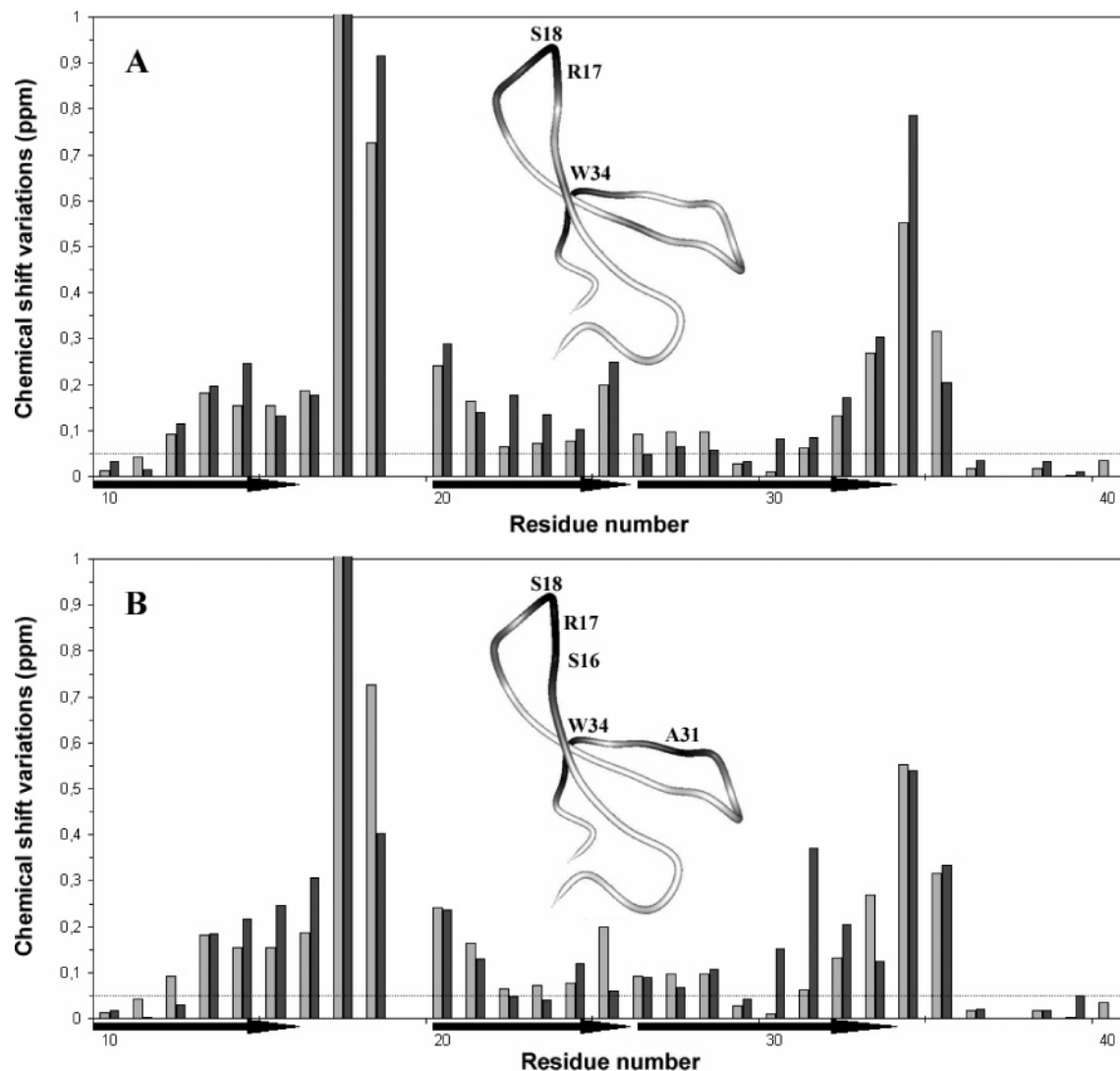
Protected natural amino acid, coupling reagents (HOBt and HBTU), and resin (polystyrene cross-linked with 2% divinylbenzene carrying a rink amide MBHA linker) were purchased from Novabiochem. Purity was checked by HPLC and LC–MS systems. LC–MS data were recorded on a Varian 1200 MS system using a TSKgel OSH supelco column. Water (A) and acetonitrile (B) containing 0.1% and 0.12% formic acid, respectively, were used as eluents. After sample injection, the column is equilibrated in solution A for 2 min and a 0–100% B 7.30 min-gradient is performed at a flow rate of 1 mL/min to elute compounds that are detected at 215 and 254 nm. HPLC data are collected on a C18 nucleosil column equilibrated in an aqueous solution containing 0.05% TFA (solution A). The eluent (solution B) contains 80% acetonitrile and 0.05% TFA. After sample injection, the column is washed with solution A for 5 min and then a 0–100% B 30 min gradient is performed at a flow rate of 1 mL/min. Compounds are detected at 215 nm. Characterization of compounds dissolved in DMSO-*d*<sub>6</sub> was also done with <sup>1</sup>H and <sup>13</sup>C NMR spectra recorded at 300 K on a 300 MHz Bruker spectrometer equipped with a BBI probehead.

**General Procedure (A) for Synthesis of Compounds 1 to 11. Removal of Fmoc Protecting Group.** Cleavage of the Fmoc group was performed with 20% piperidine in DMF (v/v). The resin beads were mixed with the previous solution (3 mL for 300 mg resin), and the reaction was performed at room temperature for 25 min. The solution was removed by filtration under vacuum, and the resin beads were extensively washed with DMF.

**Coupling of Amino Acids.** A mixture of 3 equiv of Fmoc protected amino acid, 3 equiv of HOBt, 3 equiv of HBTU dissolved in DMF, and 9 equiv of DIPEA were added to the resin beads. The reaction was performed at room temperature for 1 h. The solution was removed by filtration under vacuum, and the resin beads were extensively washed with DMF.

**Table 3.** Structures of Hydrophobic Moieties and Dissociation Constants of Resulting Ligands Based on Proline (**1b–1g**) or on 3,4-Dehydroproline (**9d, 9g, 9h**) Derivatives

compound	<b>1b</b>	<b>1c</b>	<b>1d</b>	<b>9d</b>	<b>1e</b>	<b>1f</b>	<b>1g</b>	<b>9g</b>	<b>9h</b>
Hydrophobic moiety structures									
n =	1	2	3	3	4				
K <sub>D</sub> (μM)	35	32	30	28	48	38	15	66	5

**Figure 4.** NMR chemical shift mapping of the <sup>15</sup>N-labeled WW domain. Comparison of the mapping of the WW domain in the presence of 20 equiv of compound **1a** in light gray and (A) compound **9a** or (B) compound **1g** in dark gray. Arrows along the x-axis indicate  $\beta$ -sheet structural elements. In the structure, the ribbon is darkened as a function of chemical shift variations of compound **9a** (A) and compound **1g** (B). Amino acids that display a variation superior to 0.3 ppm are labeled.

**Amine Acetylation.** Five equivalents of acetic anhydride was dissolved in DMF, and 10 equiv of DIPEA was added. The solution was mixed with the resin beads, and the mixture was agitated at room temperature for 30 min. The solution was removed by filtration under vacuum, and the resin beads were extensively washed with DMF.

**Coupling of Hydrophobic Moieties.** A mixture containing the carboxylic acid derivative (5 equiv), HOBt (5 equiv), and HBTU (5 equiv) and DIPEA (10 equiv) in DMF was added to the resin beads, and the coupling reaction was performed

at room temperature for 1 h. The solution was removed by filtration under vacuum, and the resin beads were extensively washed with DMF.

**TFA Cleavage of Resin-Bound Compounds and Benzyl Protecting Group.** Resin beads were extensively washed with DMF, DCM, and 10% TFA in DCM and then mixed with a TFA solution containing 2.5% triisopropylsilane (v/v) and 2.5% water (v/v) (5 mL for 300 mg of resin). The heterogeneous mixture was placed under agitation for 5 h at room temperature. The solution was then filtered under vacuum, and the



resin was washed twice with TFA (2.5 mL each). The combined filtrates were collected in 200 mL of a cold mixture of ether/pentane (v/v). A white precipitate was formed, and the precipitation was achieved at  $-20\text{ }^{\circ}\text{C}$  overnight. The precipitate was isolated by filtration, washed with ether, and then dissolved in water and lyophilized.

**General Procedure (B) for Synthesis of Compounds 12 to 15. Coupling of Carboxylic Acid Derivatives on Prolinamide.** Three equivalents of carboxylic acid derivative (*N*- $\alpha$ -Fmoc-ornithine/lysine(Boc) or *N*-Boc- $\beta$ -alanine) was dissolved in DCM (15.8 mmol in 8 mL) and added to a solution of 1.5 equiv of DCC in DCM at  $0\text{ }^{\circ}\text{C}$  (7.9 mmol in 2 mL). The precipitate was filtered off, and the filtrate was added to a solution of 1 equiv of prolinamide in DCM (5.25 mmol in 3 mL). The reaction mixture was stirred for 2 h at  $0\text{ }^{\circ}\text{C}$ . The crude mixture was evaporated and dissolved in DCM. The precipitate was removed by filtration, and the organic layer was washed twice with a saturated  $\text{NaHCO}_3$  aqueous solution and then dried over  $\text{MgSO}_4$  and evaporated. Products were isolated by purification on silica column in DCM/EtOH 95/5.

**Removal of Boc Protecting Group.** One equivalent of the Boc protected compound was dissolved in DCM (1.37 mmol in 14 mL), and the mixture was cooled in an ice bath. A solution of 4 N HCl in dioxane was added dropwise to the previous solution at  $0\text{ }^{\circ}\text{C}$ . The mixture was stirred at room temperature for 3 h, and then the crude mixture was evaporated. The deprotected product is obtained in quantitative yield.

**Coupling of Diethylsquarate.** A solution of 1.2 equiv of diethylsquarate in ethanol (1.64 mmol in 5 mL) was prepared. One equivalent of  $\beta$ -alanine or *N*- $\alpha$ -Fmoc-ornithine/lysine coupled to prolinamide was dissolved in ethanol (1.37 mmol in 8 mL), and 1.5 equiv of triethylamine was added. The mixture was added dropwise to the diethylsquarate solution, and the reaction was performed at room temperature for 12 h. The crude mixture was evaporated, dissolved in DCM, and purified on a silica column in DCM/MeOH 9/1.

**Removal of Ethyl Protecting Group.** The ethylsquarate derivative was dissolved in THF (1.084 mmol in 6 mL) and stirred with a solution of 4 N HCl in THF (12 mmol in 3 mL) at room temperature for 20 h. The crude mixture was evaporated. The powder was dissolved in DCM and extracted four times with a 0.3 N solution of ammonium carbonate. The aqueous solution was lyophilized.

**Coupling of Trifluoromethylsulfochloride.** Triethylamine (1.2 equiv, 1.044 mmol, 146  $\mu\text{L}$ ) was added to 1 equiv of  $\beta$ -alanine coupled to prolinamide dissolved in DCM. Trifluorosulfonyl chloride (1.2 equiv, 1.044 mmol, 110  $\mu\text{L}$ ) was dissolved in 10 mL of DCM and added to the previous solution dropwise. The mixture was stirred at room temperature for 48 h. The crude mixture was evaporated, dissolved in DCM, and purified on a silica column in DCM/EtOH 8/2.

**Peptide Synthesis (for  $\text{IC}_{50}$  Determination).** The peptide sequence was isolated from the Tau protein around the Thr212-Pro213 epitope as described.<sup>26</sup> The phosphorylated peptide was obtained by solid phase synthesis with introduction of selectively phosphorylated threonine using appropriate building block (Novabiochem). The dansyl fluorescent probe (Sigma) was introduced selectively at the N-terminus by reaction of dansyl chloride (5 equiv) and DIPEA (10 equiv) in DMF and followed by a TFA cleavage. Peptide was isolated by precipitation of the TFA solution in 200 mL of a cold mixture of ether/heptane (v/v) and centrifugation. The peptide was dissolved in water and purified by RP-HPLC on a C18 nucleosil column equilibrated in a 0.05% TFA aqueous solution (solution A). Separation of peptides in the crude mixture was carried out at  $50\text{ }^{\circ}\text{C}$  at a flow rate of 2 mL/min using an acetonitrile linear 60 min gradient from 0 to 50% solution B (80% acetonitrile:20% water–0.05% TFA). Homogeneous fractions as checked by RP-HPLC on a C18 nucleosil column (0 to 100% B linear 30 min gradient) and MALDI-TOF MS were pooled and lyophilized.

**Protein Expression and Purification.** The WW domain was produced in *E. coli* BL21(DE3) strain, carrying the

pET28a plasmid under control of T7 promoter with a sequence encoding a N-terminus His<sub>6</sub> tag. Cells were grown at  $37\text{ }^{\circ}\text{C}$  in a LB medium containing kanamycin (20 mg/L) until O.D. at 600 nm reached about 0.6, allowing a rapid growth to high cell density. Then cells were harvested by centrifugation. The pellet was resuspended in a M9 medium containing 4 g/L glucose, 1 g/L  $^{15}\text{NH}_4\text{Cl}$ , 1 mM  $\text{MgSO}_4$ , 20 mg/L kanamycin, and MEM vitamin cocktail (Sigma) and incubated at  $37\text{ }^{\circ}\text{C}$  for 1 h. Induction was performed with 0.5 mM IPTG (at final concentration) at  $31\text{ }^{\circ}\text{C}$  for 3 h. Cells were harvested by centrifugation, and the pellet was resuspended in 50 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$  pH 8.0, 300 mM NaCl buffer (buffer A) containing 10 mM imidazole, 1 mM DTT, 0.1% NP40 and complemented with a protease inhibitor cocktail (Roche). The soluble lysate was obtained by sonication and centrifugation. The soluble extract was charged on a nickel-nitrilotriacetic affinity column (Chelating Sepharose Fast Flow, Amersham). Unbound proteins were washed with 20 mM imidazole in buffer A, and the WW domain was eluted with 250 mM imidazole in buffer A (pH 7.8). The protein was then purified by RP-HPLC on a Source 15RPC column (Amersham) equilibrated in a 0.05% TFA aqueous solution (solution A). Separation of proteins was carried out at room temperature at a flow rate of 1 mL/min using an acetonitrile linear 20 min gradient from 0 to 60% solution B (85% acetonitrile:15% water–0.05% TFA). Homogeneous fractions as checked by SDS–polyacrylamide gel electrophoresis were pooled and lyophilized.

**NMR Screening.** NMR experiments were performed on a Bruker DMX 600 MHz spectrometer at 293 K equipped with a cryogenic probehead.  $^{15}\text{N}$ -labeled WW was dissolved in 25 mM deuterated Tris-HCl (pH = 6.8) aqueous buffer containing 50 mM NaCl, 5 mM DTT, and 2.5 mM EDTA and 10%  $\text{D}_2\text{O}$ . The spectra were referenced to the signal of sodium 3-trimethylsilyl-d(3,3',2,2')-propionate (TSPA- $d_4$ ). A solution of  $^{15}\text{N}$ -labeled WW at 200  $\mu\text{M}$  (60  $\mu\text{L}$ ) was added to a 2 mM solution of ligand (120  $\mu\text{L}$ ) to reach final concentrations (in 600  $\mu\text{L}$ ) of 20  $\mu\text{M}$  WW and 400  $\mu\text{M}$  ligand in the NMR sample buffer.

**NMR Dissociation Constant Calculation.** Lyophilized aliquots of ligand were prepared from a concentrated solution (2 mM) to get a concentration range of 0.025, 0.05, 0.1, 0.15, 0.25, 0.5, and 1 mM (at final concentration in 600  $\mu\text{L}$  of protein solution). A solution of 50  $\mu\text{M}$   $^{15}\text{N}$ -labeled WW was added to lyophilized aliquots of ligands. Quantification of chemical shift perturbations was calculated according to eq 1.

$$\Delta\delta(\text{ppm}) = \sqrt{\Delta\delta(^1\text{H})^2 + 0.2\Delta\delta(^{15}\text{N})^2} \quad (1)$$

$\Delta\delta(^1\text{H})$  and  $\Delta\delta(^{15}\text{N})$  are the chemical shift variations of amide proton ( $\text{H}_\text{N}$ ) and amide nitrogen ( $\text{N}_\text{H}$ ), respectively, of WW domain residues the upon ligand addition.

Dissociation constants were calculated by fitting on a graphical representation the theoretical curve obtained from eq 2 with experimental points.

$$\Delta\delta(\text{ppm}) = 0.5\Delta\delta_{\text{max}} \left( 1 + X + \frac{K_D}{[\text{protein}]_0} - \sqrt{\left( 1 + X + \frac{K_D}{[\text{protein}]_0} \right)^2 - 4X} \right) \quad (2)$$

$X$  is the molar ratio of ligand on protein,  $[\text{protein}]_0$  the initial concentration of protein, and  $\Delta\delta$  the chemical shift variation calculated using eq 1.  $\Delta\delta_{\text{max}}$  was set at the  $\Delta\delta$  value at the protein saturation level.

**FRET Competition Experiments for  $\text{IC}_{50}$  Determination.** FRET experiments were achieved on a PTI fluorescence spectrometer by exciting tryptophan residues of WW domain at 295 nm. As for NMR titrations, a 20  $\mu\text{M}$  solution of WW in protonated Tris 50mM pH 6.8 buffer containing 50 mM NaCl and 5 mM DTT/EDTA and various concentrations of dansylated peptide were added on lyophilized aliquots of ligand. The dissociation constant of the WW–peptide complex was evalu-

ated to 100  $\mu\text{M}$  according to eq 2 by replacing  $\Delta\delta$  with  $\Delta I^{334}$ , the intensity variation of the fluorescence emission signal measured at 334 nm.<sup>26</sup>  $\text{IC}_{50}$  determination was calculated as described for dissociation constants determined by NMR with eq 2 by replacing  $\Delta\delta$  with the inhibition percentage given by eq 3 and  $K_D$  by  $\text{IC}_{50}$ .

$$\% \text{ inhibition} = \frac{I_0^{334} - I^{334}}{I_0^{334} - I_{100}^{334}} \quad (3)$$

$I^{334}$  is the fluorescence intensity at 334 nm upon ligand addition,  $I_0^{334}$ , the fluorescence intensity without inhibition, and  $I_{100}^{334}$ , the fluorescence intensity at maximal inhibition.

**Acknowledgment.** The authors thank the Institute for the Study Of Aging (ISOA, the United States) for financial support. We thank Professor André Tartar for fruitful discussions, Gérard Montagne, Hervé Drobecq, and Eric Diesis for their technical contribution to this work and Dr. Dragos Horvath for molecular modeling. The 600 MHz facility used in this study was funded by the Région Nord-Pas de Calais (France), the CNRS, and the Institut Pasteur de Lille.

## Appendix

**Abbreviations.** MBHA, 4-methylbenzhydramine; HOBt, 1-hydroxybenzotriazole; HBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; Fmoc, 9-fluorenylmethoxy-carbonyl;  $K_D$ , dissociation constant;  $\text{IC}_{50}$ , inhibition constant at 50%; Boc, *tert*-butoxycarbonyl; Bzl, benzyl; DCM, dichloromethane; DMF, dimethylformamide; EtOH, ethanol; DMSO, dimethyl sulfoxide; DIPEA, diisopropylethylamine; TFA, trifluoroacetic acid; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HSQC, heteronuclear simple quantum correlation; Tris, 2-amino-2-hydroxy-methyl-1,3-propanediol; BBI, inverse broad band (probe); MALDI-TOF, matrix assisted laser desorption ionization time-of-flight; RP-HPLC, reverse phase high pressure liquid chromatography; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

**Supporting Information Available:** Synthetic procedures and spectral characterization of compounds **1a** to **15**. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- Pawson, T.; Raina, M.; Nash, P. Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett.* **2002**, *513*, 2–10.
- Yaffe, M. B.; Smerdon, S. J. The Use of in Vitro Peptide-Library Screens in the Analysis of Phosphoserine/Threonine-Binding Domain Structure and Function. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 225–244.
- Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P. Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* **1997**, *89*, 875–886.
- Lu, K. P.; Hanes, S. D.; Hunter, T. A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **1996**, *380*, 544–547.
- Winkler, K. E.; Swenson, K. I.; Kornbluth, S.; Means, A. R. Requirement of the prolyl isomerase Pin1 for the replication checkpoint. *Science* **2000**, *287*, 1644–1647.
- Wulf, G.; Ryo, A.; Liou, Y. C.; Lu, K. P. The prolyl isomerase Pin1 in breast development and cancer. *Breast Cancer Res.* **2003**, *5*, 76–82.
- Wulf, G. M.; Ryo, A.; Wulf, G. G.; Lee, S. W.; Niu, T.; et al. Pin1 is overexpressed in breast cancer and cooperates with Ras signaling in increasing the transcriptional activity of c-Jun towards cyclin D1. *Embo J.* **2001**, *20*, 3459–3472.
- Lu, K. P. Prolyl isomerase Pin1 as a molecular target for cancer diagnostics and therapeutics. *Cancer Cell* **2003**, *4*, 175–180.
- Lu, P. J.; Wulf, G.; Zhou, X. Z.; Davies, P.; Lu, K. P. The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* **1999**, *399*, 784–788.
- Uchida, T.; Takamiya, M.; Takahashi, M.; Miyashita, H.; Ikeda, H. et al. Pin1 and Par14 peptidyl prolyl isomerase inhibitors block cell proliferation. *Chem. Biol.* **2003**, *10*, 15–24.
- Hennig, L.; Christner, C.; Kipping, M.; Schelbert, B.; Rucknagel, K. P.; et al. Selective inactivation of parvulin-like peptidyl-prolyl cis/trans isomerases by juglone. *Biochemistry* **1998**, *37*, 5953–5960.
- Wang, X. J.; Xu, B.; Mullins, A. B.; Neiler, F. K.; Etkorn, F. A. Conformationally locked isostere of phosphoSer-cis-Pro inhibits Pin1 23-fold better than phosphoSer-trans-Pro isostere. *J. Am. Chem. Soc.* **2004**, *126*, 15533–15542.
- Lu, P. J.; Zhou, X. Z.; Shen, M.; Lu, K. P. Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science* **1999**, *283*, 1325–1328.
- Lu, P. J.; Zhou, X. Z.; Liou, Y. C.; Noel, J. P.; Lu, K. P. Critical role of WW domain phosphorylation in regulating phosphoserine binding activity and Pin1 function. *J. Biol. Chem.* **2002**, *277*, 2381–2384.
- Verdecia, M. A.; Bowman, M. E.; Lu, K. P.; Hunter, T.; Noel, J. P. Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat. Struct. Biol.* **2000**, *7*, 639–643.
- Wintjens, R.; Wieruszkeski, J. M.; Drobecq, H.; Rousselot-Pailley, P.; Buee, L. et al. 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. *J. Biol. Chem.* **2001**, *276*, 25150–25156.
- Songyang, Z.; Shoelson, S. E.; Chaudhuri, M.; Gish, G.; Pawson, T. et al. SH2 domains recognize specific phosphopeptide sequences. *Cell* **1993**, *72*, 767–778.
- Yan, K. S.; Kuti, M.; Zhou, M.-M. PTB or not PTB- that is the question. *FEBS Lett.* **2002**, *73*, 67–70.
- Yaffe, M. B.; Elia, A. E. H. Phosphoserine/threonine-binding domains. *Curr. Opin. Cell Biol.* **2001**, *13*, 131–138.
- Macias, M. J.; Wiesner, S.; Sudol, M. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett.* **2002**, *513*, 30–37.
- Kay, B. K.; Williamson, M. P.; Sudol, M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **2000**, *14*, 231–241.
- Burke, T. R.; Lee, K. Phosphotyrosyl mimetics in the development of signal transduction inhibitors. *Acc. Chem. Res.* **2003**, *36*, 426–433.
- Jeannotte, G.; Lubell, W. D. Large Structural Modification with Conserved Conformation: Analysis of Delta(3)-Fused Aryl Prolines in Model beta-Turns. *J. Am. Chem. Soc.* **2004**, *126*, 14334–14335.
- Grellier, P.; Vendeville, S.; Joyeau, R.; Bastos, I. M.; Drobecq, H. et al. Trypanosoma cruzi prolyl oligopeptidase Tc80 is involved in nonphagocytic mammalian cell invasion by trypanomastigotes. *J. Biol. Chem.* **2001**, *276*, 47078–47086.
- Willand, N.; Joossens, J.; Gesquière, J.-C.; Tartar, A. L.; Evans, M. D. et al. Solid and solution phase synthesis of the 2-cyanopyrrolidide DPP-IV inhibitor NVP-DPP728. *Tetrahedron* **2002**, *58*, 5741–5746.
- Smet, C.; Sambo, A. V.; Wieruszkeski, J. M.; Leroy, A.; Landrieu, I.; et al. The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau. *Biochemistry* **2004**, *43*, 2032–2040.
- Widmer, H.; Jahnke, W. Protein NMR in biomedical research. *Cell. Mol. Life Sci.* **2004**, *61*, 580–599.
- Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins. *Science* **1997**, *278*, 497, 499.
- Hajduk, P. J.; Gerfin, T.; Boehlen, J. M.; Haberli, M.; Marek, D.; et al. High-throughput nuclear magnetic resonance-based screening. *J. Med. Chem.* **1999**, *42*, 2315–2317.
- Hamdane, M.; Smet, C.; Sambo, A. V.; Leroy, A.; Wieruszkeski, J. M.; et al. Pin1: a therapeutic target in Alzheimer neurodegeneration. *J. Mol. Neurosci.* **2002**, *19*, 275–287.

JM0500119



# NMR-based detection of acetylation sites in peptides

Caroline Smet-Nocca,<sup>a\*‡</sup> Jean-Michel Wieruszkeski,<sup>b</sup> Oleg Melnyk<sup>c</sup>  
and Arndt Benecke<sup>a,d</sup>

Acetylation of histone tails as well as non-histone proteins was found to be a major component of the 'chromatin code' that regulates transcription through the recruitment of transcription factors, co-regulators and DNA-binding proteins. Acetylation can have several effects modifying protein-protein interactions, protein activity, localization and stability. Using NMR spectroscopy, we provide a simple way to detect acetyl moieties at the  $\epsilon$ -amino function of lysine residues based on peptides derived from Histone H4 and TDG amino-terminal domains. Significant changes of acetyl-lysine resonances as compared to non-acetylated residues allow a direct identification of specific acetylated lysine. We also show that, in unfolded peptides, acetylation of lysine side chains leads to characteristic NMR signals that vary only weakly depending on the primary sequence or the total number of acetylated sites, indicating that the acetamide group does not establish any interactions with other residues. Furthermore, resonance changes upon acetylation are restricted to residues nearby the acetylation site, indicating that acetylation does not modify the overall peptide conformation. Copyright © 2010 European Peptide Society and John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article

**Keywords:** acetylation; histone; thymine-DNA glycosylase; protein structure; NMR spectroscopy

## Introduction

Posttranslational modifications of proteins are crucial events in the regulation of protein functions or protein-protein interactions. Acetylation of lysine residues is an abundant posttranslational modification within the eukaryotic cells. Acetylation ranks in the same order as phosphorylation [1] in terms of abundance and exhibits the same features as phospho-dependent signaling pathways because addition and removal of acetyl moieties are highly dynamic and are supported by several (de)acetylase enzymes [2]. A large number of non-histone acetylated proteins, the so-called acetylome [1], has been identified in eukaryotic cells. Reversible protein acetylation is involved in the regulation of signaling pathways, in the modification of DNA-binding properties or protein stability and localization, and rivals with other posttranslational modifications targeting lysine residues, such as methylation, ubiquitination and sumoylation. These effects result from neutralizing the positively charged lysine residues and generating new binding sites. Bromodomains were described as specialized domains involved in specific binding of acetylation sites and are found in a number of transcription factors or co-activators like HATs [3,4]. In malignant cells, aberrant acetylation patterns were found as well as modifications of HAT and HDAC activities [5]. The role of HATs and HDACs in development and their deregulation in human diseases like leukemia, breast and colorectal cancers or human developmental disorders emphasizes their functional prevalence. Furthermore, inhibition of HDAC activities has been found to be of promising therapeutic use in cancer treatment because HDAC inhibitors are able to modify the acetylation patterns of histone and non-histone proteins as

well, implying a change in their protein-protein interactions and in their subsequent cellular localization and lifetime [6–9].

Many transcription-related factors and DNA-binding proteins were found to be acetylated. A common feature of these proteins is that they often contain unfolded regions which (i) are specifically involved in protein-protein or protein-DNA interactions, (ii) display structural changes upon substrate binding and

\* Correspondence to: Caroline Smet-Nocca, Unité de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, Université des Sciences et Technologies de Lille, Bâtiment C9, 59655 Villeneuve d'Ascq Cedex, France.  
E-mail: caroline.smet@univ-lille1.fr

a Institut de Recherche Interdisciplinaire, CNRS USR3078, Université de Lille1, Parc de la Haute Borne, 50 Avenue de Halley, 59658 Villeneuve d'Ascq Cedex, France

b Unité de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, Université de Lille1, 59655 Villeneuve d'Ascq Cedex, France

c Institut de Biologie de Lille, CNRS UMR 8161, Université de Lille 1 et 2, 1 rue Professeur Calmette, 59021 Lille Cedex, France

d Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440 Bures-sur-Yvette, France

‡ Present address: Université des Sciences et Technologies de Lille, Unité de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, Bâtiment C9, 59655 Villeneuve d'Ascq Cedex.

**Abbreviations used:** CBP, Creb (cAMP-responsive element binding protein)-binding protein; HAT, histone acetyltransferase; HDAC, histone deacetylase; HSQC, heteronuclear single quantum coherence spectroscopy; SSP, secondary structure propensity; TDG, Thymine-DNA Glycosylase; TOCSY, total correlation spectroscopy.



(iii) are often preferential targets for posttranslational modifications [10–17]. Histones, the protein units at the basis of the nucleosome particle and chromatin architecture, are among the best studied examples of unfolded targets of the posttranslational modification machinery [18–20]. Multiple and combinatorial posttranslational modifications were characterized for histone proteins and more precisely the histone amino-terminal tails [21]. Acetylation is a key modification of histones as their acetylation levels are well correlated with transcriptionally active chromatin domains [22]. Histone acetylation, as it has been well described for the histone H4-K16 site [23,24], is involved in chromatin relaxation [25–27] helping to increase DNA accessibility for transcription-related factors. Acetylation also provides novel binding interfaces for proteins targeting acetylated residues and therefore is an integral component of the chromatin code hypothesis [28]. However, a lack of structural details at the protein level hampers the establishment of structure–activity relationship upon acetylation.

Identification of acetylation sites and their binding partners is crucial for a better description of the acetylome and acetylation-dependent signaling pathways. A major issue consists in the determination of acetylation sites when multiple acetylations occur as this is the case of the histone amino-terminal tails. The identification of the acetylated proteins within cells and the determination of the acetylation sites have been made possible through the use of sensitive radiolabeled  $^{14}\text{C}$ - or  $^3\text{H}$ -acetyl-CoA substrates [29–31]. Determination of specific acetylated lysine residues is typically realized by Edman degradation of peptide substrates [32] or acetylation of chemically synthesized peptides *in vitro* [30], which are either destructive or indirect methods. Finally, mass spectrometry is the method of choice to detect acetylation within purified proteins isolated from cells. However, multiple acetylations render difficult the identification of acetylation sites and necessitate a tandem mass spectrometry strategy [33,34]. Quantification of acetylation levels is also a major issue that requires the isotopic labeling of acetyl moieties to ensure a normalization of signal intensities [33,34].

We describe here the impact of lysine side chain acetylation in peptide substrates using NMR spectroscopy. We have studied the structural effects of acetylation on unfolded peptides derived from two CBP/p300 substrates: the histone H4 amino-terminal tail [31,35] and a non-histone protein, the human TDG [32], which is involved in both transcriptional regulation and DNA repair processes [36–41]. They both contain at the molecular level an extended amino-terminal region that is targeted by the CBP/p300 acetyltransferase activity.

## Materials and Methods

### Peptide synthesis and acetylation

Peptides used in this study are those of the H4 histone *N*-terminal tail and human TDG, whose sequences are Ac-SGRGKGGK<sub>8</sub>GLGKGGAK<sub>16</sub>RHRKVLN-NH<sub>2</sub> and Ac-QPVEPKPVESK<sub>83</sub>K<sub>84</sub>SGK<sub>87</sub>SAKSKEKQ-NH<sub>2</sub>, respectively (Ac and NH<sub>2</sub> indicate that the *N*- and the *C*-terminus are acetylated and amidated, respectively, and the acetylated lysine residues are numbered). Peptides were synthesized automatically by solid-phase synthesis on a continuous flow synthesizer (Perseptive Biosystems, Framingham, MA, USA) using the Fmoc group as temporary protection for the alpha-amine function of each amino acid. Peptides were assembled at a 0.1-mmol scale on a Rink amide resin (purchased from Novabiochem, Merck Biosciences

AG, Läufelfingen, Switzerland), leading after the final TFA cleavage step to an amide function at the carboxy-terminus. Synthesis was carried out using ten equivalents of *N*- $\alpha$ -Fmoc-protected amino acids and coupling was performed with TBTU/HOBt and DIPEA in DMF. Acetylated lysine residues were introduced by coupling of the commercially available *N*- $\alpha$ -Fmoc-*N*- $\epsilon$ -acetyl-lysine synthon (Novabiochem). Fmoc cleavage was performed using 20% piperidine in DMF. Capping steps were performed after each coupling step using acetic anhydride (3%) and DIPEA (0.3%) in DMF. Amino-terminal functions of peptides were acetylated by a final acetylation step that immediately succeeds to the amino-terminal Fmoc deprotection. Peptides were released from the solid support by resin incubation with 5 ml of a TFA solution containing 2.5% triisopropylsilane and 2.5% water for 3 h at room temperature. Crude peptides were precipitated in 250 ml of a cold ether:heptane mixture (v/v) and recovered by centrifugation. Crude peptides were purified by reverse-phase chromatography in 0.05% TFA using a linear gradient of 0–40% acetonitrile in 60 min. Homogeneous fractions, as checked by analytical reverse-phase HPLC and MALDI-TOF mass spectrometry, were pooled and lyophilized.

### NMR spectroscopy

NMR experiments were performed at 293 K on a Bruker DMX 600-MHz spectrometer (Bruker, Karlsruhe, Germany) equipped with a cryogenic triple-resonance probe head. For NMR experiments, peptides were dissolved at 2 mM in a buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>, pH 6.8, 100 mM NaCl and 5% D<sub>2</sub>O. All  $^1\text{H}$  spectra were calibrated with 1 mM sodium 3-trimethylsilyl-3,3',2,2'-d<sub>4</sub>-propionate as a reference. For peptide assignment, standard NOESY and TOCSY experiments were recorded with 400 and 69 ms mixing times, respectively, with 2048 and 512 points and using a DIPSI2 sequence for mixing. Both are recorded with a Water suppression by Gradient-Tailored Excitation (WATERGATE) pulse sequence for water suppression [42].  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded at nitrogen-15 natural abundance with 64 scans per increment, with 2048 and 256 points in the proton and nitrogen dimensions, respectively, and with a window of 24.1802 ppm centered on 118.006 ppm for the nitrogen-15 dimension.  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectra were recorded with 128 scans per increment, with 1024 and 512 points in the proton and carbon dimensions, respectively, and with a window of 70.4950 ppm centered on 37.362 ppm for the carbon-13 dimension. Heteronuclear experiments were recorded with a WATERGATE sequence for water suppression, a Carr–Purcell–Meiboom–Gill sequence and a double-INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) for sensitivity improvement. All experiments were acquired with a recycle delay of 1 s.

The chemical shift perturbations of individual resonances in peptides were calculated with Eqn (1) taking into account the relative dispersion of the proton and nitrogen chemical shifts (1 and 20 ppm, respectively).

$$\Delta\delta(\text{ppm}) = \sqrt{[\Delta\delta(^1\text{H})]^2 + 0.05[\Delta\delta(^{15}\text{N})]^2} \quad (1)$$

Integration of lysine NH $\epsilon$ -acetyl signals was performed either on the proton spectrum for the methyl signals or on the  $^1\text{H}$  traces extracted from the  $^{15}\text{N}$ - $^1\text{H}$  HSQC experiment for the NH $\epsilon$  signal.

### Reverse-phase chromatography and MALDI-TOF mass spectrometry analyses of peptide mixtures

Concentrated solutions of purified peptides at 10 mM in water were mixed up to equimolar ratio at final concentrations of



0.4 mM. The peptide mixture was analyzed by reverse-phase chromatography on a C2/C18 column (GE Healthcare Life Sciences, Uppsala, Sweden) equilibrated in 0.1% TFA buffer containing 2% acetonitrile; 20 nmol of each peptide was separated by a linear gradient of acetonitrile (from 2 to 40% acetonitrile in 16.67 ml, i.e. ten-column volume) at room temperature. The same peptide mixture was analyzed by MALDI-TOF mass spectrometry (Voyager DE-PRO 2000, Applied Biosystems, Foster City, California, USA) on  $\alpha$ -cyano-4-hydroxycinnamic acid matrix.

## Results

### Detection of lysine side chain acetylation in peptides

Histone H4 and TDG peptides analyzed in this study were chosen among the well-characterized acetylation sites in the literature. Peptides were synthesized by Fmoc solid-phase chemistry and acetyl moieties were introduced by the use of the *N*- $\alpha$ -Fmoc-(L)-Lys(*N*- $\epsilon$ -acetyl)-OH derivative. The NMR  $^{15}\text{N}$ - $^1\text{H}$  HSQC experiments were acquired at  $^{15}\text{N}$  natural abundance in peptides with the use of a cryoprobe head at 600 MHz. Resonance assignments were performed with the acquisition for each peptide of homonuclear TOCSY and NOESY experiments. A remarkable consequence of lysine acetylation on NMR resonances is that it converts an amine function which is in fast exchange with water leading to undetectable NMR signals into an acetamide group which adopts the same behavior as any backbone amide. In consequence, the acetamide moiety at the  $\text{NH}\epsilon$  function of lysine side chain exhibits two characteristic signals: the  $\text{NH}\epsilon$  amide at 8.01 ppm (Figures 1 and 2) and the methyl group at 1.98 ppm (Figure 1). Moreover, the relative integration of the acetamide methyl signal of lysine side chain at 1.98 ppm on one hand and those of the *N*-terminal acetyl moiety at 2.09 ppm on the other indicates the number of acetylation sites (Figure 1). Here, the integration values obtained with a recycle delay of 1 s and the number of acetylations introduced by selective  $\text{NH}\epsilon$ -acetyl-lysine incorporation during solid-phase synthesis are in good agreement and illustrates the applicability of the integration of  $\text{NH}\epsilon$ -acetyl signals to quantify the acetylation level. However, one should take care about the recycle delays in the case of unknown acetylation pattern for which a fine measurement is required. Longer recycle delays are recommended to get accurate integration values at the expense of the acquisition time. The  $\text{NH}\epsilon$  amide group of acetyl-lysine exhibits multiple scalar contacts with all side chain protons and a weak contact with  $\text{H}\alpha$  in the  $^1\text{H}$ - $^1\text{H}$  TOCSY experiment allowing the identification of the corresponding acetylated lysine residue in the absence of any NOE contacts between the  $\text{NH}\epsilon$  and other residues (Figure 2). Identification of specific acetylated lysine in unfolded peptides is hampered by the fact that  $^1\text{H}$  chemical shifts differ only weakly from random coil values, hence giving a lot of resonance redundancies. However, as observed in our peptide models from amino-terminal tail of histone H4 and human TDG, a significant change in resonances comprised between +0.13 and +0.20 ppm for the  $\text{H}\epsilon$  signals and between -0.19 and -0.22 ppm for the  $\text{H}\delta$  signals as compared to non-acetylated peptides (Tables S1 and S2, Supporting Information) facilitates the identification of acetylation sites (Figure 2). In contrast to the  $\text{H}\epsilon$  resonances that shift in the lower field, the  $\text{H}\delta$  resonances shift in the higher field as do the  $\text{H}\gamma$  signals, however, with a lower amplitude. The backbone amide signal of the acetylated residue is also significantly shifted as compared to the corresponding non-acetylated lysine falling within a

range of -0.08 to -0.14 ppm (Tables S1 and S2). In contrast, the acetyl effect is less pronounced on  $\text{H}\alpha$  and  $\text{H}\beta$  resonances.

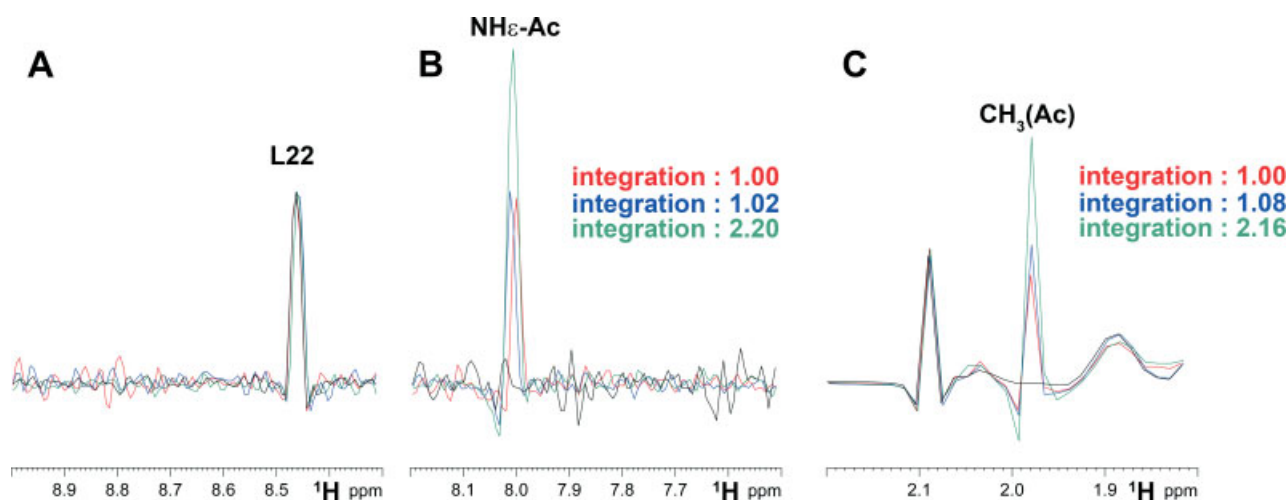
Additionally, shifts of peptide resonances have been mapped on a  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum that can detect variations of  $^{13}\text{C}$  resonances upon lysine side chain acetylation. Resonances of the modified lysines are the most affected with a marked effect for the  $\text{C}\delta$  (Table S3), whereas the remaining resonances are weakly or not perturbed (Figure 3(A)). However, given their high degree of degeneracy,  $^{13}\text{C}$  values of either non-acetylated or acetylated lysine residues are not useful for the identification of the acetylation sites. As  $^{13}\text{C}$  resonances are sensitive to changes in backbone structure, perturbations of  $^{13}\text{C}$  resonances could be used to detect a propensity to adopt  $\alpha$ -helical or  $\beta$ -sheet structures upon lysine acetylation. In the TDG [72–95] peptide,  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  resonances are within a range of 1 ppm with respect to a purely random coil structure indicating that the peptide adopts mainly an extended structure [43,44]. We have measured, by calculating the SSP with  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  values [45], an overall content of 10.7%  $\beta$ -structure localized within the *N*- and *C*-termini (residues 72–79 and 93–95, Figure 3(B)). Small variations of  $^{13}\text{C}\alpha$  and/or  $^{13}\text{C}\beta$  resonances of residues flanking the acetylation sites (V80, E81, S82, S85, G86, S88 and A89) are detected upon multiple lysine acetylation (Figure 3(A)) inducing only slight modifications of the overall  $\beta$ -structure content (Figure 3(B)).

Acetylation of lysine side chains has also an impact on the local peptide conformation and affects only the resonances of neighboring residues. We have examined, using  $^{15}\text{N}$ - $^1\text{H}$  HSQC at natural abundance combined with TOCSY/NOESY experiments, the effect of lysine acetylation (i) in the case of a single acetylation at different positions within the same peptide sequence and (ii) in the case of multiple close or distant acetylation sites.

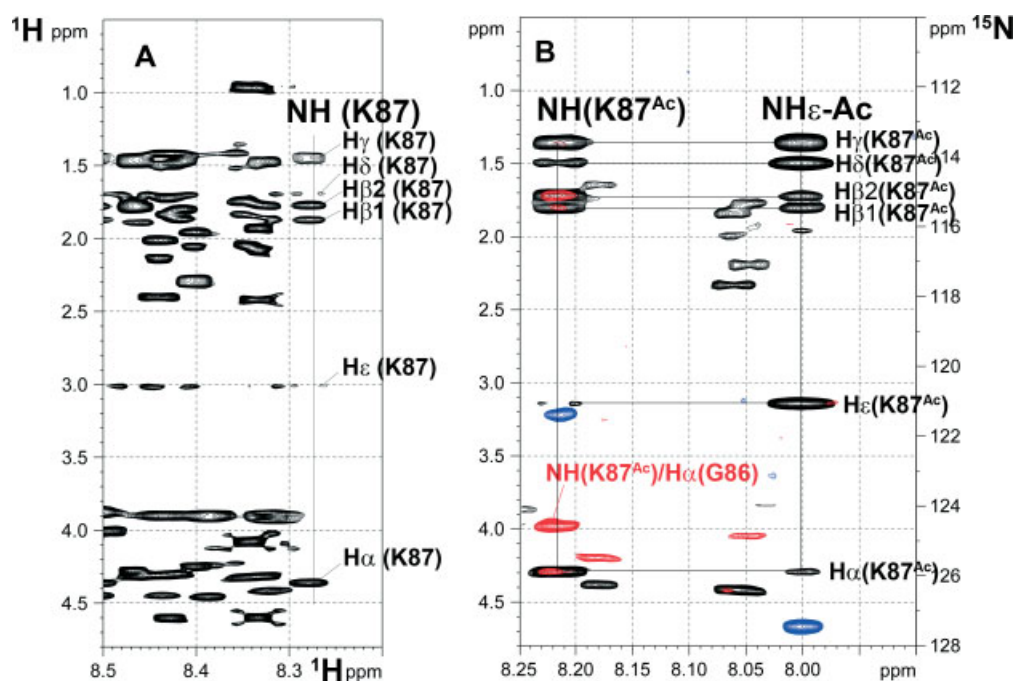
### Effect of different single acetylations on TDG peptide resonances

Multiple acetylation sites has been identified *in vitro* and found to be in close vicinity within the amino-terminal domain of TDG [32], which is involved in the regulation of its DNA repair activity [37,46]. This functional region encompassing the acetylation sites, the so-called regulatory domain, interacts in an intramolecular manner with the catalytic domain and undergoes a conformational equilibrium between this 'closed' state and an 'open' state that adopts an extended conformation with little residual structure [46]. We have synthesized peptides encompassing the three identified acetylation sites at K83, K84 and K87 positions (human TDG numbering) located within the TDG regulatory domain [32]. In the non-acetylated peptide, the detection of the sole sequential contacts between  $\text{NH}$  and  $\text{H}\alpha$  of residues *i* and *i* - 1, respectively, is indicative of a random coil structure. This observation is supported by  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  resonances coinciding with random coil values. We have found that a single-lysine acetylation at each position leads to local perturbations spanning over the two or three neighboring residues in addition to the acetylation site (Figure 4). The resonance of the  $\text{NH}\epsilon$  acetamide moiety is not sensitive to the chemical environment in these unfolded peptides as it does not vary with the acetylated lysine residue indicating no interaction between the acetamide group and other residues.

Furthermore, no NOE contacts were detected between the  $\text{NH}\epsilon$  acetamide and other peptide resonances, indicating an absence of interactions mediated by the acetamide moiety. We have not observed a significant difference of dipolar couplings between residues upon acetylation suggesting that lysine acetylation does



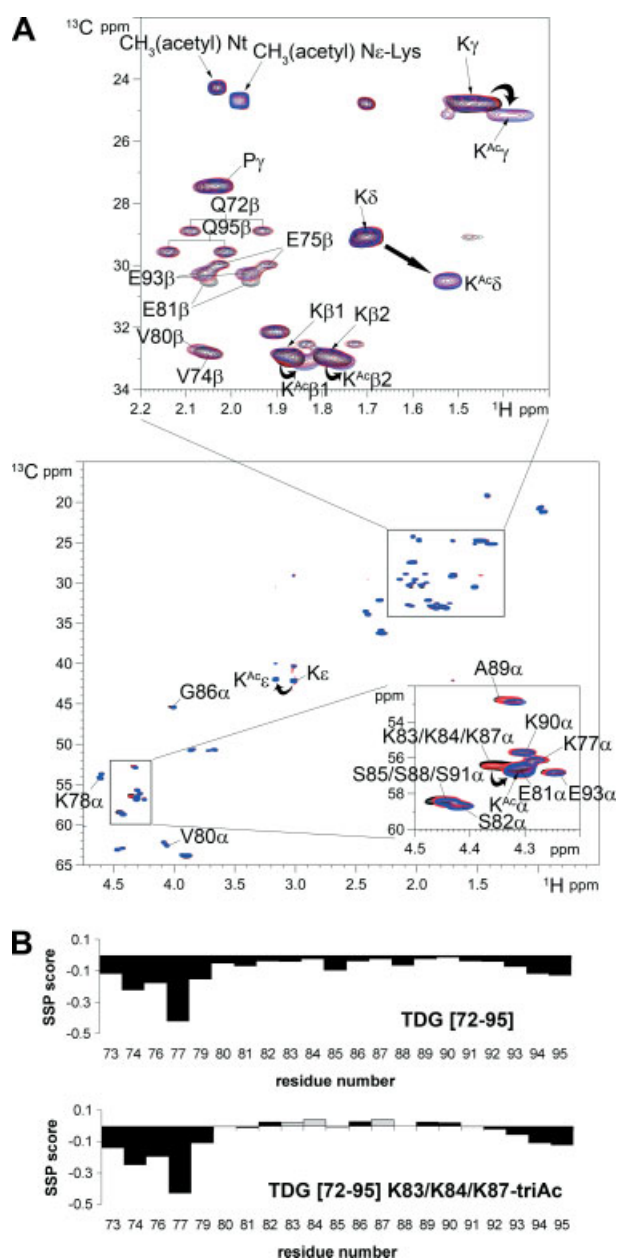
**Figure 1.** One-dimensional projections extracted from  $^{15}\text{N}$ - $^1\text{H}$  spectra in the  $^1\text{H}$  dimension (A and B) for the H4 [1–23] peptides non-acetylated (black), mono-acetylated on K8 (red) or on K16 (blue), di-acetylated at both K8 and K16 (green) showing the row corresponding to L22 residue at 127.2531 ppm (A) as a reference and the row extracted from  $\text{NH}\epsilon$  acetamide signal at 127.3816 ppm (B).  $^1\text{H}$  wet spectra of the same peptides centered on the methyl signal of the acetamide moiety (C). Relative peak intensities are given (B and C): the NH acetamide signals in (B) are normalized on the L22 signal (A) and the methyl signals at 1.98 ppm in (C) are normalized on the signal at 2.09 ppm corresponding to the methyl group of the amino-terminal acetamide moiety.



**Figure 2.** Scalar couplings of the K87 residue in its non-acetylated (A) or acetylated (B) forms. Detail of the  $^1\text{H}$ - $^1\text{H}$  TOCSY spectra (black) of the TDG [72–95] K84-Ac (A) or K87-Ac (B) peptides showing the  $^1\text{H}$ - $^1\text{H}$  scalar couplings between  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{H}\gamma$ ,  $\text{H}\delta$ ,  $\text{H}\epsilon$  and the backbone NH for the K87 (A) and K87-Ac (B) residues or the  $\text{NH}\epsilon$  acetamide of the K87-Ac residue (B).  $^1\text{H}$ - $^1\text{H}$  NOESY spectrum is depicted in red indicating the  $\text{NH}/\text{H}\alpha$  dipolar coupling between K87-Ac and G86, and the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum is shown in blue (B). In the panel (B), the x-axis is common for the three TOCSY, NOESY and  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra, whereas the y-axis corresponds on the left to the  $^1\text{H}$  dimension of the TOCSY and NOESY spectra and on the right to the  $^{15}\text{N}$  dimension of the  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum.

not change the overall extended structure of the peptides. The region encompassed in the TDG [72–95] peptide has a marked propensity to adopt a residual structure within the TDG protein [46]. An analysis of the primary sequence reveals a high content of charged residues that could be responsible for transient long-range interactions (hydrogen bonds, salt bridges, etc.) in the polypeptide chain. In this context, neutralization of positively charged lysine residues by acetylation could contribute to modify

these non-covalent interactions. One can further notice that acetylation of K83 induces larger averaged variations of chemical shifts than acetylation of K84 or K87 suggesting more pronounced conformational changes in the TDG [72–95] peptide when K83 is acetylated (Figure 4). In its non-acetylated form, the  $\text{NH}_3$  group of lysine is hardly detected due to fast proton exchange with water that causes severe line broadening. However, at low pH and depending on the surrounding environment, exchange rates with



**Figure 3.** (A)  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectra of TDG [72–95] non-acetylated (black), K83-Ac/K84-Ac di-acetylated (red) and K83-Ac/K84-Ac/K87-Ac tri-acetylated (blue) peptides (QPVEPKKPVE $\text{SK}_{83}\text{K}_{84}\text{SGK}_{87}\text{SAKSKEKQ}$ ). Relevant resonances are annotated. Acetylated residues are indicated by an Ac superscript and the N-terminus by an Nt. (B) SSP score along the TDG [72–95] primary sequence for the non-acetylated (upper panel) and the K83-Ac/K84-Ac/K87-Ac tri-acetylated (lower panel) peptides. The SSP scores of acetylated lysines are indicated by grey bars.

water are slow enough to allow the observation of  $\text{NH}_3$  resonances. In the TDG [72–95] peptide, the signals of lysine  $\text{NH}_3$  groups were detected only at lower pH values (pH 4.01 and 2.99, Figure S1(B)). The corresponding  $^{15}\text{N}$  resonances are found typically around 30–50 ppm [47]. Here, given the spectral width in the nitrogen dimension, the signal was four-fold aliased and observed at 129.5 ppm. Furthermore, decoupling of nitrogen was poor at this frequency, leading to the observation of a  $^1\text{J}_{\text{HN}}$  coupling of 70.44 Hz (Figure S1(B)), which is significantly smaller than  $^1\text{J}_{\text{HN}}$  couplings observed for backbone amide (around 90 Hz) [47]. In the TOCSY

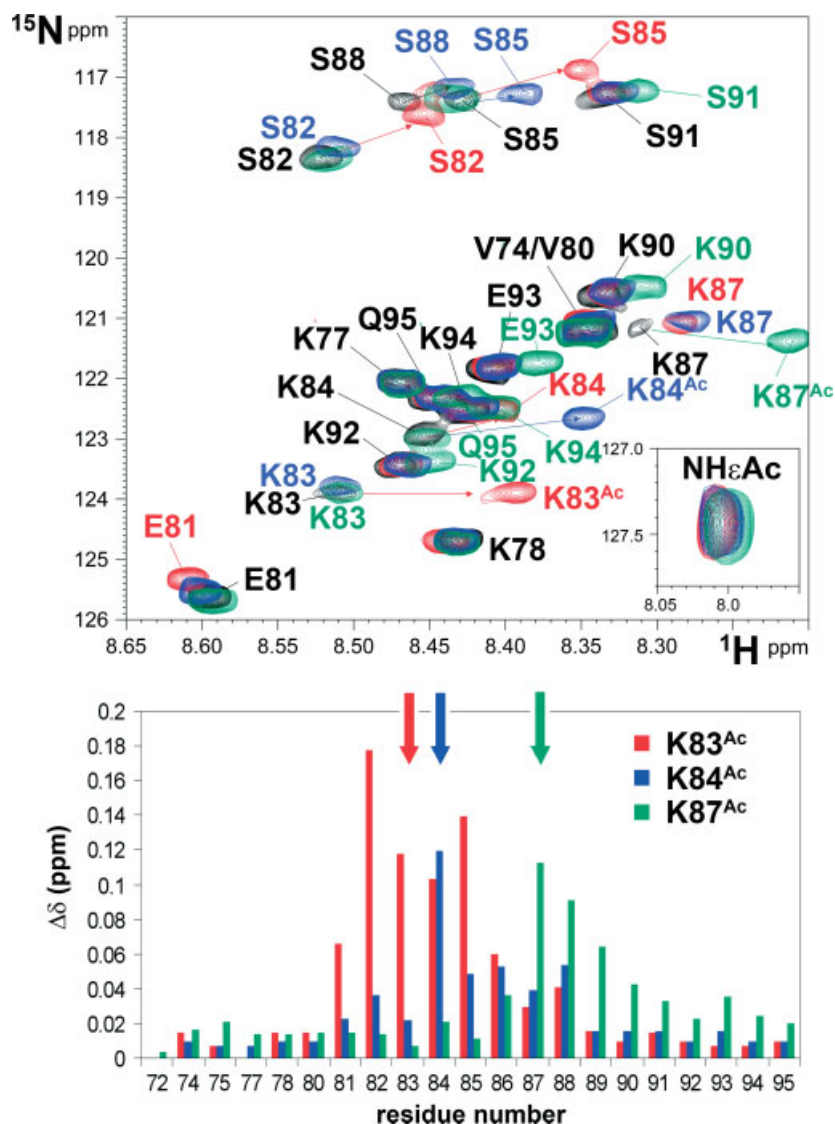
experiment acquired at pH 2.99, scalar couplings between  $\text{NH}_3$  and  $\text{H}\alpha$  or side chain protons do not allow for an identification of the corresponding lysines, all of them being potentially detectable. The important degeneracy of lysine signals (except for K78  $\text{H}\alpha$  at 4.585 ppm due to the presence of a proline at the position 79) hampers the assignment of the  $\text{NH}_3$  lysine resonances (Figure S1(C)). Salt bridges between lysine and nearby glutamate residues (E75, E81 and/or E93) are quite unfavorable at this low pH and the low content of hydrophobic residues argue rather for an extended structure, so that such slow water-exchange rates can merely be explained by a pH effect. Furthermore, a pH titration has been performed to evaluate the impact of a modification of the glutamate charge on the resonances of lysine backbone amides. Because the observation of backbone amides is also related to water-exchange rates, the pH range that can be experienced is restricted to pH values lower than 7. We have acquired TOCSY and  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra on the TDG [72–95] peptide at pH 2.99, 4.01, 5.10 and 6.65. Marked perturbations of chemical shifts upon pH variations were detected for all glutamate residues with a larger shift between pH 4 and 5 (Figure S1(A)) which is in good agreement with the standard  $\text{pK}_R$  value of glutamate (around 4.25). The same transition is found to a lesser extent for surrounding residues in the primary sequence (for residues 72–82 in the N-terminus and 90–95 in the C-terminus). In contrast, in the central region of the peptide (for residues 83–89) that encompasses the three acetylation sites, a transition between pH 5.1 and 6.65 was observed together with smaller shift amplitudes (Figure S1(A)). These results indicate that a modification of the glutamate charge has little impact on the environment of backbone amide of K83, K84 or K87 residues in this unfolded peptide. Hence, the chemical shift perturbations detected upon lysine acetylation stem from the acetyl group itself rather than from disruption of long-range interactions.

### Multiple acetylations: the case of distant acetylation sites in the histone H4 amino-terminus

We have evaluated the reciprocal effect of two distant acetylation sites in the histone H4 peptide (residues 1–23) located within the unfolded amino-terminal tail. The mono-acetylated peptides at K8 and K16 positions, and the di-acetylated peptide at both positions were analyzed by NMR spectroscopy (Figures 1 and 5). Single acetylation of either K8 or K16 residue leads to a minor modification of the peptide conformation restricted to the neighboring residues in the primary sequence (Figure 5) as seen for TDG peptides. Furthermore, no intramolecular interaction involving the acetamide moiety was observed because the  $\text{NH}\epsilon$  signal has nearly the same  $^1\text{H}$  and  $^{15}\text{N}$  values in every peptide (Figure 5). Remarkably, the pattern of chemical shift modifications in the di-acetylated H4 peptide corresponds to a simple additional effect of both K8 and K16 acetyl moieties indicating that remote acetylation sites have no cross-effect on peptide resonances. A detailed analysis of this latter peptide also allows the quantification of its acetylation level (i) through the comparison of the methyl signal integration of N- $\epsilon$ -acetyl-lysine with those of the amino-terminal acetamide (Figure 1(C)) or (ii) through the comparison of  $\text{NH}\epsilon$  signal integration of the acetyl-lysine residues with those of other  $\text{NH}$  signals (Figure 1(A) and (B)).

In contrast to mass spectrometry for which equimolar amounts of peptides at different acetylation states lead to unequal signal intensities (Figures S2 and S3), NMR analyses allow the relative quantification of acetylation levels at each position. For the histone





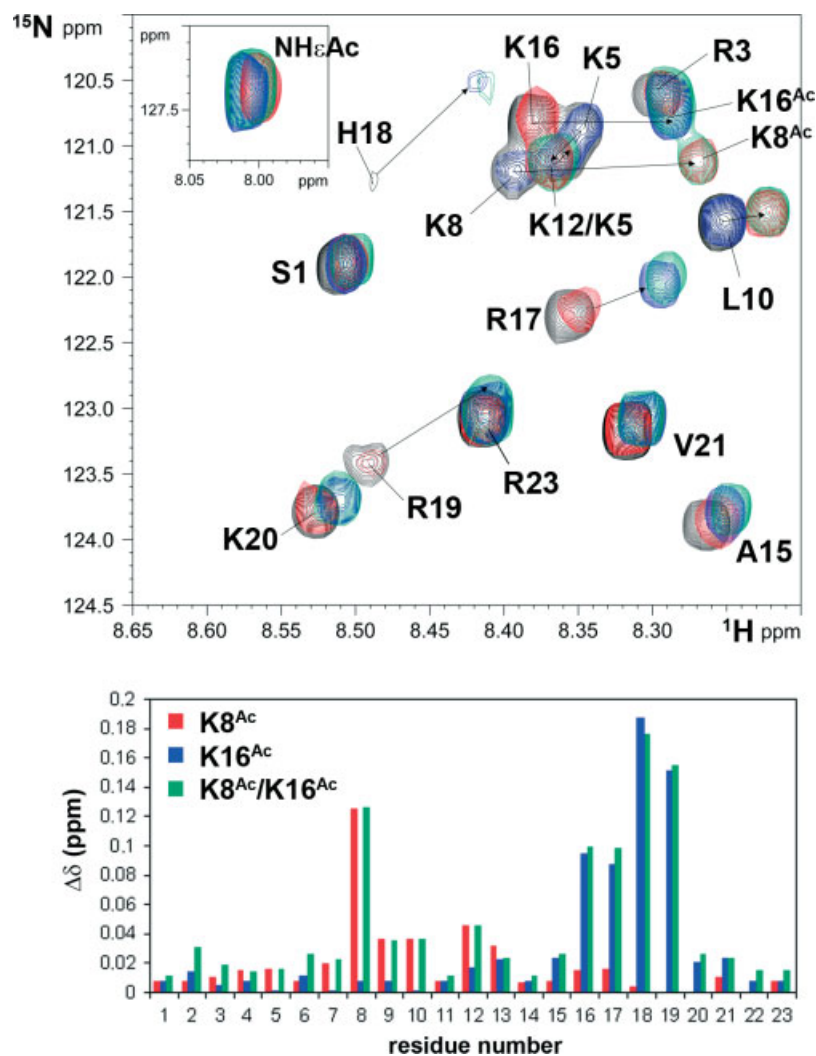
**Figure 4.** Effect of individual acetylations on the amide backbone resonances.  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra of TDG [72–95] peptides (QPVEPKKPVESK<sub>83</sub>K<sub>84</sub>SGK<sub>87</sub>SAKSKEKQ) at natural abundance (upper panel). The non-acetylated peptide is depicted in black, the mono-acetylated peptide on K83 residue in red, on K84 residue in blue and on K87 residue in green. Residues are annotated and acetylated residues are indicated by an Ac superscript. The chemical shift variations between each of the three mono-acetylated peptides and the non-acetylated one along the peptide primary sequence using the same color code as for the NMR spectra (lower panel) are represented graphically.  $\Delta\delta$  values are the averaged  $^1\text{H}$  and  $^{15}\text{N}$  chemical shift variations calculated according to Eqn (1). Acetylation sites are indicated by arrows.

H4 di-acetylated peptide, we have found identical acetylation levels at both K8 and K16 positions with the integration of  $\text{NH}\epsilon/\text{H}\alpha$  and  $\text{NH}/\text{H}\alpha$  contacts in the  $^1\text{H}$ - $^1\text{H}$  TOCSY experiment. We obtained for these couple of signals the same values of peak intensities indicating comparable acetylation levels at both positions (Figure 6).

#### Multiple acetylations: the case of close acetylation sites in the TDG amino-terminus

Again, we have used the sensitivity of the amide backbone resonance to probe modifications of the peptide conformation upon multiple nearby acetylations in the TDG amino-terminal region (residues 72–95) at positions K83, K84 and K87. We have compared the effect of either two vicinal acetylations at positions K83 and K84 or three acetylations (Figure S4). NOE contacts restricted to  $\text{H}_\text{N}/\text{H}\alpha$  between residues  $i$  and  $i - 1$ , respectively,

as well as  $^{13}\text{C}$  resonances that fall within a range of 1 ppm from random coil values still indicate that both peptides adopt an extended conformation with a slight propensity (8.5%) to form  $\beta$ -sheet structures in their  $N$ - and  $C$ -terminal parts as indicated by the SSP score [45] (Figure 3(B)). This propensity ranges, however, in the same order than for the non-acetylated peptide suggesting no marked modification of the peptide structure upon acetylation. In the case of K83/K84 acetylations, a simple additional effect of lysine acetylation was observed (Figure S4(A) and (B)) and no significant shift was detected out of the 81–88 region. In contrast, an additional acetylation on K87 partially counteracts the effect of K83/K84 acetylation on K84 and S85 resonances and induces larger shifts spanning over the entire  $C$ -terminus with a more pronounced effect than the sum of K83<sup>Ac</sup>/K84<sup>Ac</sup> and K87<sup>Ac</sup> individual contributions (Figure S4(C) and (D)). Such long-range perturbations of amide backbone resonances indicate



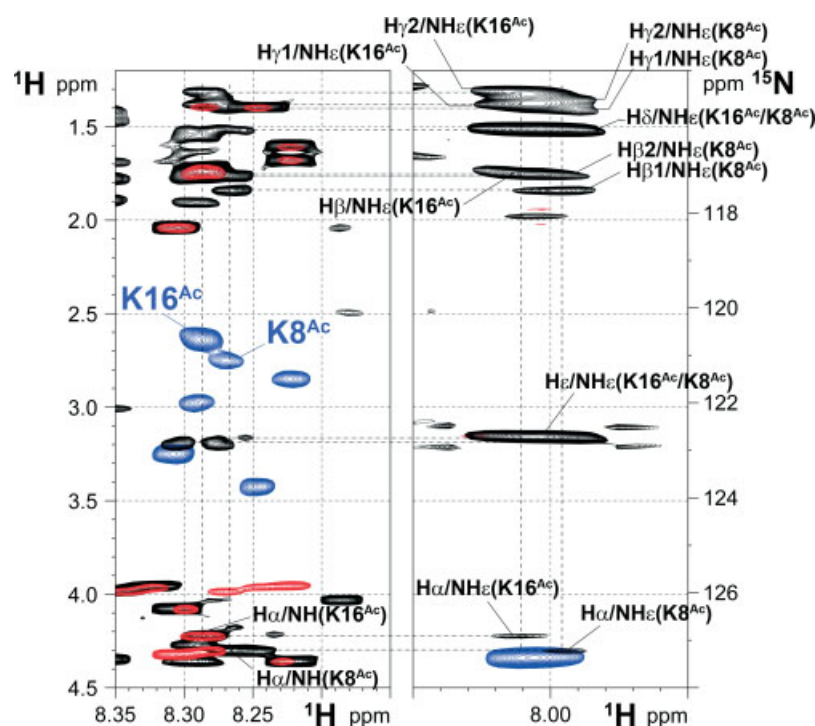
**Figure 5.** Effect of two remote acetylations on the amide backbone resonances.  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra of the histone H4 [1–23] peptides (SGRGKGGK<sub>8</sub>GLGKGGAK<sub>16</sub>RHRKVL<sub>R</sub>) at natural abundance (upper panel). The non-acetylated peptide is depicted in black, the mono-acetylated peptide on K8 residue in red or on K16 residue in blue, and the di-acetylated peptide at both K8 and K16 residues in green. Residues are annotated and acetylated residues are indicated by an Ac superscript. The chemical shift variations between each of the acetylated peptides and the non-acetylated one along the peptide primary sequence with the same color code as for the NMR spectra (lower panel) are represented graphically.  $\Delta\delta$  values are the averaged  $^1\text{H}$  and  $^{15}\text{N}$  chemical shift variations calculated according to Eqn (1).

that K87 acetylation could have a broader effect on the peptide conformation.

## Discussion

We have investigated by NMR spectroscopy the effect of lysine side chain acetylation in peptide substrates using two peptide models, one from the histone H4 and the other from a non-histone protein, the human TDG, which are both substrates of the acetyltransferase CBP/p300. We have shown that acetylation at distinct positions within both substrates has a similar impact on peptide resonances. This methodology allows a direct identification of acetylation sites based on a classical strategy of peptide assignment using TOCSY and NOESY experiments. Scalar contacts between lysine side chain protons and the characteristic  $\text{NH}\epsilon$  acetamide signal at 8.01 ppm enable the identification/quantification of acetylation sites even in the context of a multiple acetylated substrate. The acquisition of  $^{15}\text{N}$ - $^1\text{H}$  and  $^{13}\text{C}$ - $^1\text{H}$  HSQC at natural abundance has allowed

for the investigation of the effect of an acetyl moiety on the  $^{15}\text{N}$  and  $^{13}\text{C}$  chemical shifts and on the peptide structure. These experiments can be used to detect lysine acetylation through characteristic resonances and allow for an overall quantification of acetylation levels. However, they cannot be used for the identification/quantification of each acetylation site due to a high degree of resonance degeneracy concerning the (non)acetylated lysines in the case of random coil peptides. Given the high sensitivity of  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  to backbone structure, these chemical shifts can be used to map a perturbation of the peptide structure upon acetylation. Based on the sensitivity of backbone amide to their chemical environment, a mapping of NH variations upon acetylation of one or more lysine residues can be achieved using the  $^{15}\text{N}$ -HSQC experiment. If used in the context of a folded protein, long-range perturbations can be detected for residues near the acetylation site. Here, we have shown in the context of unfolded peptides that acetylation of individual lysines has only a restricted effect on peptide conformation. No overall changes in the extended structure was observed but only significant shifts



**Figure 6.** Detail of the  $^1\text{H}$ - $^1\text{H}$  TOCSY spectrum (black) of the histone H4 [1–23] di-acetylated peptide (SGRGKGGK $_8^{\text{Ac}}$ GLGKGGAK $_{16}^{\text{Ac}}$ RHRKVLRL) showing the  $^1\text{H}$ - $^1\text{H}$  scalar couplings between  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{H}\gamma$ ,  $\text{H}\delta$ ,  $\text{H}\epsilon$  and the backbone NH (left panel) or the  $\text{NH}\epsilon$  acetamide (right panel) for both the K8-Ac and K16-Ac residues. The  $^1\text{H}$ - $^1\text{H}$  NOESY spectrum is depicted in red and the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum in blue. Note that the x-axis is common for the three TOCSY, NOESY and  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra, whereas the y-axis corresponds on the left to the  $^1\text{H}$  dimension of the TOCSY and NOESY spectra and on the right to the  $^{15}\text{N}$  dimension of the  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum.

of the backbone NH signals of the two or three residues in the immediate vicinity of the acetylation site. In these unfolded peptides, multiple acetylations being close or far from each other have also little influence on the peptide structure. No cross-effect was detected in the case of two remote acetylation sites as studied using a histone H4 K8/K16 di-acetylated peptide. The effect of two or three close acetylation sites exemplified in the case of the TDG peptide has been analyzed and it shows an additional contribution of each acetylation on the overall variations of peptide resonances with a more pronounced effect of K87 acetylation in the triply acetylated peptide. A detailed study of the latter peptide in its non-acetylated state indicates that despite a high charge density spanning over the entire sequence, no long-range interaction is observed between glutamate and lysine residues that could be perturbed by charge neutralization upon lysine acetylation. Moreover, the  $\text{NH}\epsilon$  or methyl resonances of the acetamide moiety are almost the same for all the acetylation sites, be they in the histone H4 or TDG amino-terminal domains. These features suggest that the acetyl group does not establish any interactions with other residues and always experienced the same chemical environment.

A number of works has highlighted the use of NMR spectroscopy to detect residual structures in intrinsically unfolded proteins or the presence of disordered domains in globular proteins [46,48–51]. The use of glutamine as acetylated lysine mimics has shown that a single-charge neutralization can modify the chromatin organization suggesting that acetylation of various histone tails has distinct roles in nucleosome self-assembly and in the regulation of the higher order chromatin structure [52]. Our methodology, when combined with a native peptide ligation strategy [53,54], could determine the structural impact of lysine acetylation itself

in a full protein context. This strategy can take advantage from chemical peptide synthesis to generate peptide fragments of up to 50-amino acid length while selectively and quantitatively introducing one or several posttranslational modifications. The other part of the protein of interest can be synthesized as a recombinant protein with the advantage to be isotopically labeled (with  $^{15}\text{N}$  and/or  $^{13}\text{C}$  and/or  $^2\text{H}$ ) when expressed in *Escherichia coli* (for review, see Refs 55–57). The production of high amounts of natively folded proteins can be achieved at relatively low costs as compared to the introduction of isotopically labeled amino acids during solid-phase peptide synthesis. Then, the native chemical ligation provides segmental isotopically labeled proteins with a fully controlled incorporation of posttranslational modifications. Their structural effect can be investigated by NMR spectroscopy through the assignment of a sub-spectrum corresponding to the  $^{15}\text{N}$  or  $^{13}\text{C}$ -labeled protein fragment. In other words, for histone proteins, the structural effect of homogeneous covalent, and even combinatorial, histone tail acetylations introduced within the unlabeled, and therefore invisible, part of the protein can be studied through the spectrum analysis of the labeled part (the histone core) to evaluate their impact on histone and nucleosome structures. Isotopic labels can also be introduced at strategic points during the chemical synthesis of acetylated peptides to investigate the direct effect of lysine acetylation on the *N*-terminal tail structure and function such as in the establishment of intramolecular interactions or in the binding of protein or DNA molecules. The detection and characterization of *in vitro* acetylation after treatment of recombinant proteins or protein fragments with purified acetyltransferases or cell extracts could take advantage from the per-residue resolution of NMR spectroscopy as similarly described for phosphorylation [58,59].



Such a methodology applied to the detailed investigation of chromatin structure and function would have a significant impact on the comprehension of the chromatin code hypothesis [28], a central dogma in systems biology.

### Acknowledgements

This work has been funded by the Centre National de la Recherche Scientifique (CNRS), the Région Nord-Pas de Calais, the Institut de Recherche Interdisciplinaire, the Institut des Hautes Études Scientifiques and the Agence Nationale de la Recherche. The 600-MHz facility used in this study was funded by the Région Nord-Pas de Calais (France), the CNRS, the University of Lille 1 and the Institut Pasteur de Lille. The authors thank Drs Guy Lippens and Isabelle Landrieu for insightful discussions.

### Supporting information

Supporting information may be found in the online version of this article.

### References

- Spange S, Wagner T, Heinzel T, Kramer OH. Acetylation of non-histone proteins modulates cellular signalling at multiple levels. *Int. J. Biochem. Cell Biol.* 2009; **41**: 185–198.
- Kouzarides T. Acetylation: a regulatory modification to rival phosphorylation? *EMBO J.* 2000; **19**: 1176–1179.
- Mujtaba S, Zeng L, Zhou MM. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* 2007; **26**: 5521–5527.
- Zeng L, Zhou MM. Bromodomain: an acetyl-lysine binding domain. *FEBS Lett.* 2002; **513**: 124–128.
- Timmermann S, Lehrmann H, Polesskaya A, Harel-Bellan A. Histone acetylation and disease. *Cell. Mol. Life Sci.* 2001; **58**: 728–736.
- Mottet D, Castronovo V. Histone deacetylases: target enzymes for cancer therapy. *Clin. Exp. Metastasis* 2008; **25**: 183–189.
- Selvi RB, Kundu TK. Reversible acetylation of chromatin: implication in regulation of gene expression, disease and therapeutics. *Biotechnol. J.* 2009; **4**: 375–390.
- Ocker M, Schneider-Stock R. Histone deacetylase inhibitors: signalling towards p21cip1/waf1. *Int. J. Biochem. Cell Biol.* 2007; **39**: 1367–1374.
- Carew JS, Giles FJ, Nawrocki ST. Histone deacetylase inhibitors: mechanisms of cell death and promise in combination cancer therapy. *Cancer Lett.* 2008; **269**: 7–17.
- Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 2004; **32**: 1037–1049.
- Obradovic Z, Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK. Predicting intrinsic disorder from amino acid sequence. *Proteins* 2003; **53**(Suppl 6): 566–572.
- Tomba P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* 2002; **27**: 527–533.
- Radivojac P, Obradovic Z, Smith DK, Zhu G, Vucetic S, Brown CJ, Lawson JD, Dunker AK. Protein flexibility and intrinsic disorder. *Protein Sci.* 2004; **13**: 71–80.
- Fink AL. Natively unfolded proteins. *Curr. Opin. Struct. Biol.* 2005; **15**: 35–41.
- Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, Iakoucheva LM. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol.* 2006; **2**: e100.
- Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK. Intrinsic disorder and functional proteomics. *Biophys. J.* 2007; **92**: 1439–1456.
- Uversky VN, Radivojac P, Iakoucheva LM, Obradovic Z, Dunker AK. Prediction of intrinsic disorder and its use in functional proteomics. *Methods Mol. Biol.* 2007; **408**: 69–92.
- Hansen JC, Lu X, Ross ED, Woody RW. Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J. Biol. Chem.* 2006; **281**: 1853–1856.
- Bang E, Lee CH, Yoon JB, Lee DW, Lee W. Solution structures of the N-terminal domain of histone H4. *J. Pept. Res.* 2001; **58**: 389–398.
- Cary PD, Crane-Robinson C, Bradbury EM, Dixon GH. Effect of acetylation on the binding of N-terminal peptides of histone H4 to DNA. *Eur. J. Biochem.* 1982; **127**: 137–143.
- Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000; **403**: 41–45.
- Eberharter A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep.* 2002; **3**: 224–229.
- Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 2006; **311**: 844–847.
- Shogren-Knaak M, Peterson CL. Switching on chromatin: mechanistic role of histone H4-K16 acetylation. *Cell Cycle* 2006; **5**: 1361–1365.
- Kurdistani SK, Tavazoie S, Grunstein M. Mapping global histone acetylation patterns to gene expression. *Cell* 2004; **117**: 721–733.
- Schubeler D, MacAlpine DM, Scalzo D, Wirbelauer C, Kooperberg C, van Leeuwen F, Gottschling DE, O'Neill LP, Turner BM, Delrow J, Bell SP, Groudine M. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* 2004; **18**: 1263–1271.
- Ura K, Kurumizaka H, Dimitrov S, Almouzni G, Wolffe AP. Histone acetylation: influence on transcription, nucleosome mobility and positioning, and linker histone-dependent transcriptional repression. *EMBO J.* 1997; **16**: 2096–2107.
- Benecke A. Chromatin code, local non-equilibrium dynamics, and the emergence of transcription regulatory programs. *Eur. Phys. J. E Soft matter* 2006; **19**: 353–366.
- Lau OD, Kundu TK, Soccio RE, Ait-Si-Ali S, Khalil EM, Vassilev A, Wolffe AP, Nakatani Y, Roeder RG, Cole PA. HATs off: selective synthetic inhibitors of the histone acetyltransferases p300 and PCAF. *Mol. Cell* 2000; **5**: 589–595.
- Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 1996; **87**: 953–959.
- Schiltz RL, Mizzen CA, Vassilev A, Cook RG, Allis CD, Nakatani Y. Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. *J. Biol. Chem.* 1999; **274**: 1189–1192.
- Tini M, Benecke A, Um SJ, Torchia J, Evans RM, Chambon P. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. *Mol. Cell* 2002; **9**: 265–277.
- Smith CM. Quantification of acetylation at proximal lysine residues using isotopic labeling and tandem mass spectrometry. *Methods* 2005; **36**: 395–403.
- Smith CM, Gafken PR, Zhang Z, Gottschling DE, Smith JB, Smith DL. Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Anal. Biochem.* 2003; **316**: 23–33.
- Brownell JE, Allis CD. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr. Opin. Genet. Dev.* 1996; **6**: 176–184.
- Chen D, Lucey MJ, Phoenix F, Lopez-Garcia J, Hart SM, Losson R, Buluwela L, Coombes RC, Chambon P, Schar P, Ali S. T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor alpha. *J. Biol. Chem.* 2003; **278**: 38586–38592.
- Gallinari P, Jiricny J. A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. *Nature* 1996; **383**: 735–738.
- Hardeland U, Bentele M, Jiricny J, Schar P. Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. *J. Biol. Chem.* 2000; **275**: 33449–33456.
- Hardeland U, Bentele M, Jiricny J, Schar P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, *Drosophila* and fission yeast orthologs. *Nucleic Acids Res.* 2003; **31**: 2261–2271.
- Lucey MJ, Chen D, Lopez-Garcia J, Hart SM, Phoenix F, Al-Jehani R, Alao JP, White R, Kindle KB, Losson R, Chambon P, Parker MG, Schar P, Heery DM, Buluwela L, Ali S. T:G mismatch-specific thymine-DNA glycosylase (TDG) as a coregulator of transcription interacts with SRC1 family members through a novel tyrosine repeat motif. *Nucleic Acids Res.* 2005; **33**: 6393–6404.



- 41 Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, Wiebauer K, Jiricny J. Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J. Biol. Chem.* 1996; **271**: 12767–12774.
- 42 Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* 1992; **2**: 661–665.
- 43 Wishart DS, Sykes BD, Richards FM. The chemical shift index : a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* 1992; **31**: 1647–1651.
- 44 Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD.  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* 1995; **5**: 67–81.
- 45 Marsh JA, Singh VK, Jia Z, Forman-Kay JD. Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein : implications for fibrillation. *Prot. Sci.* 2006; **15**: 2795–2804.
- 46 Smet-Nocca C, Wieruszeski J-M, Chaar V, Leroy A, Benecke A. The Thymine-DNA Glycosylase regulatory domain: residual structure and DNA binding. *Biochemistry* 2008; **47**: 6519–6530.
- 47 Iwahara J, Jung YS, Clore GM. Heteronuclear NMR spectroscopy for lysine NH(3) groups in proteins: unique effect of water exchange on  $(^{15}\text{N})$  transverse relaxation. *J. Am. Chem. Soc.* 2007; **129**: 2971–2980.
- 48 Bussell R, Jr, Eliezer D. Residual structure and dynamics in Parkinson's disease-associated mutants of alpha-synuclein. *J. Biol. Chem.* 2001; **276**: 45996–46003.
- 49 Lippens G, Sillen A, Smet C, Wieruszeski JM, Leroy A, Buee L, Landrieu I. Studying the natively unfolded neuronal Tau protein by solution NMR spectroscopy. *Protein Pept. Lett.* 2006; **13**: 235–246.
- 50 Smet C, Leroy A, Sillen A, Wieruszeski JM, Landrieu I, Lippens G. Accepting its random coil nature allows a partial NMR assignment of the neuronal Tau protein. *ChemBioChem* 2004; **5**: 1639–1646.
- 51 Sung YH, Eliezer D. Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Mol. Biol.* 2007; **372**: 689–707.
- 52 Wang X, Hayes JJ. Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating the stability of higher-order chromatin structure. *Mol. Cell. Biol.* 2008; **28**: 227–236.
- 53 Shogren-Knaak MA, Fry CJ, Peterson CL. A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J. Biol. Chem.* 2003; **278**: 15744–15748.
- 54 Shogren-Knaak MA, Peterson CL. Creating designer histones by native chemical ligation. *Methods Enzymol.* 2004; **375**: 62–76.
- 55 Muir TW. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* 2003; **72**: 249–289.
- 56 Muralidharan V, Muir TW. Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat. Methods* 2006; **3**: 429–438.
- 57 Otomo T, Ito N, Kyogoku Y, Yamazaki T. NMR observation of selected segments in a larger protein: central-segment isotope labeling through intein-mediated ligation. *Biochemistry* 1999; **38**: 16040–16044.
- 58 Landrieu I, Lacosse L, Leroy A, Wieruszeski JM, Trivelli X, Sillen A, Sibille N, Schwalbe H, Saxena K, Langer T, Lippens G. NMR analysis of a Tau phosphorylation pattern. *J. Am. Chem. Soc.* 2006; **128**: 3575–3583.
- 59 Landrieu I, Leroy A, Smet-Nocca C, Huvent I, Amniai L, Hamdane M, Sibille N, Buée L, Wieruszeski J-M, Lippens G. Nuclear Magnetic Resonance Spectroscopy of the neuronal Tau protein: normal function and implication in Alzheimer's disease. *Biochem. Soc. Trans.* 2010; **38** (in press) .

Cite this: DOI: 10.1039/c0mb00337a

www.rsc.org/molecularbiosystems

PAPER

## Identification of O-GlcNAc sites within peptides of the Tau protein and their impact on phosphorylation†

Caroline Smet-Nocca,<sup>\*a</sup> Malgorzata Broncel,<sup>b</sup> Jean-Michel Wieruszeski,<sup>a</sup>  
Caroline Tokarski,<sup>c</sup> Xavier Hanouille,<sup>a</sup> Arnaud Leroy,<sup>a</sup> Isabelle Landrieu,<sup>a</sup>  
Christian Rolando,<sup>c</sup> Guy Lippens<sup>a</sup> and Christian P. R. Hackenberger<sup>\*b</sup>

Received 17th December 2010, Accepted 25th January 2011

DOI: 10.1039/c0mb00337a

Phosphorylation of the microtubule-associated Tau protein plays a major role in the regulation of its activity of tubulin polymerization and/or stabilization of microtubule assembly. A dysregulation of the phosphorylation/dephosphorylation balance leading to the hyperphosphorylation of Tau proteins in neurons is thought to favor their aggregation into insoluble filaments. This in turn might underlie neuronal death as encountered in many neurodegenerative disorders, including Alzheimer's disease. Another post-translational modification, the O-linked  $\beta$ -N-acetylglucosaminylation (O-GlcNAcylation), controls the phosphorylation state of Tau, although the precise mechanism is not known. Moreover, analytical difficulties have hampered the precise localization of the O-GlcNAc sites on Tau, except for the S400 site that was very recently identified on the basis of ETD-FT-MS. Here, we identify three O-GlcNAc sites by screening a library of small peptides sampling the proline-rich, the microtubule-associated repeats and the carboxy-terminal domains of Tau as potential substrates for the O- $\beta$ -N-acetylglucosaminyltransferase (OGT). The *in vitro* activity of the nucleocytoplasmic OGT was assessed by tandem mass spectrometry and NMR spectroscopy. Using phosphorylated peptides, we establish the relationship between phosphate and O-GlcNAc incorporation at these sites. Phosphorylation of neighboring residues S396 and S404 was found to decrease significantly S400 O-GlcNAcylation. Reciprocally, S400 O-GlcNAcylation reduces S404 phosphorylation by the CDK2/cyclinA3 kinase and interrupts the GSK3 $\beta$ -mediated sequential phosphorylation process.

### Introduction

One of major hallmarks in Alzheimer's disease (AD) is the formation of neurofibrillary tangles (NFTs), composed mainly of the Tau protein in an aggregated and hyperphosphorylated state. Although not clear whether the latter hyperphosphorylation results from an imbalance between the kinases/phosphatases involved, recent results suggest that it also might stem from a dysregulation of another post-translational modification (PTM), the O-linked  $\beta$ -N-acetylglucosaminylation (O-GlcNAcylation).

This PTM of serine or threonine is believed to (in)directly compete with phosphorylation,<sup>1</sup> as the phosphorylation levels for many Tau phosphorylation sites were found to be reciprocally regulated by increasing or decreasing O-GlcNAc levels in cultured cells or brain homogenates.<sup>2</sup> The reversibility of O-GlcNAcylation is ensured by two antagonistic enzymes, each encoded by a single gene in mammals. The O- $\beta$ -N-acetylglucosaminyltransferase (OGT) and the O- $\beta$ -N-acetylglucosaminylase (OGlcNAcase) adds and removes the O-GlcNAc moiety, respectively.<sup>3</sup> Three OGT isoforms, ncOGT for nucleocytoplasmic, mOGT for mitochondrial and sOGT for shorter OGT isoforms differ in their number of tetratricopeptide repeats (TPRs)<sup>4,5</sup> and target different cellular compartments.<sup>4</sup> The carboxy-terminal domain ensures the O-GlcNAc transferase activity.<sup>6</sup> At least twelve O-GlcNAc sites have been counted in Tau by [<sup>3</sup>H]-galactose labeling and RP-HPLC tryptic mapping.<sup>7</sup> However, their exact localization remained unclear until very recently, when the S400 and T123 residues were identified as glycosylation sites by ETD MS (Fig. 1).<sup>8</sup>

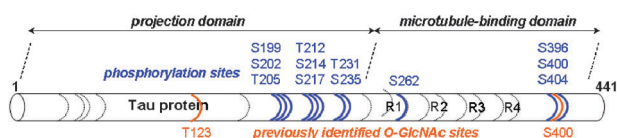
In the current study, we focus on the identification of O-GlcNAcylation sites of Tau by probing *in vitro* activity of OGT isoforms on recombinant Tau as well as synthetic

<sup>a</sup> Unité de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, IFR 147, Université de Lille1, 59655 Villeneuve d'Ascq Cedex, France.  
E-mail: caroline.smet@univ-lille1.fr; Fax: +33 (0)3 2043 6555;  
Tel: +33 (0)3 2043 4997

<sup>b</sup> Institut für Chemie und Biochemie, Freie Universität Berlin, Takustrasse 3, 14195 Berlin, Germany.  
E-mail: hackenbe@chemie.fu-berlin.de; Fax: +49 (0)3 08385 2551;  
Tel: +49 (0)3083852451

<sup>c</sup> Miniaturisation pour l'Analyse, la Synthèse & la Protéomique, USR CNRS 3290, IFR 147, Université de Lille1, 59655 Villeneuve d'Ascq Cedex, France

† Electronic supplementary information (ESI) available: Experimental procedures and analytical data. See DOI: 10.1039/c0mb00337a



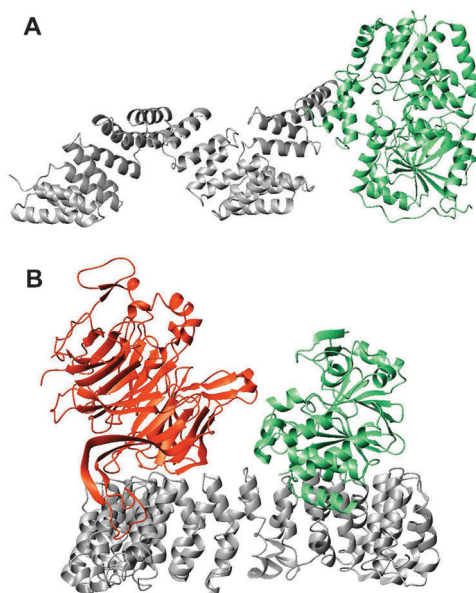
**Fig. 1** Known O-GlcNAc and phosphorylation sites in Tau investigated in this study.

peptides derived from this intrinsically unstructured protein. Glycosylation products were then analyzed by either exploiting the *per residue* resolution of NMR spectroscopy or mass spectrometry to detect O-GlcNAc sites. We previously used NMR to identify phosphorylation sites of Tau after phosphorylation with different purified kinases (PKA, CDK2/cyclinA3, GSK3 $\beta$ ).<sup>9,10</sup> In all cases, we found a pattern of phosphorylation similar to those that has been identified using phospho-Tau specific antibodies, mass spectrometry or peptide sequencing,<sup>11</sup> which confirms both the activity of recombinant kinases and the applicability of a NMR-based strategy at the qualitative level. The additional information provided by NMR spectroscopy came from the ability to directly evaluate the phosphorylation levels and to monitor the kinetics of the phosphorylation reactions at each site.<sup>9,10</sup> Here, the use of peptide fragments carrying phosphorylated residues as OGT substrates further allows us to monitor directly the interference of phosphorylation on additional glycosylation or phosphorylation events in the C-terminal part of Tau.

## Results and discussion

### *In vitro* O-GlcNAcylation of Tau protein with recombinant OGT isoforms

We first incubated <sup>15</sup>N-labeled recombinant full-length Tau with ncOGT or sOGT enzymes. Although we verified the activity of our enzymes towards a peptide substrate from the CKII kinase, with a quantitative O-GlcNAcylation after 30 min at 37 °C (Fig. S1, ESI<sup>†</sup>),<sup>12</sup> the <sup>1</sup>H-<sup>15</sup>N and <sup>1</sup>H-<sup>13</sup>C HSQC spectra of a Tau sample incubated with ncOGT in the presence or absence of UDP-GlcNAc showed no difference whatsoever. Despite the smaller shifts for the glycosylated Ser or Thr residues compared to their phosphorylated counterparts were expected, the resolution of the Tau spectrum at a high field (900 MHz) should allow to detect those shifts if present. The same was true for a shorter Tau fragment encompassing the proline-rich domain (Tau[185-245]), with longer incubation times (up to two days) or with the sOGT isoform. Western blots of NMR samples with either WGA or CTD110.2 antibody further indicated the absence of any O-GlcNAc (Fig. S1, ESI<sup>†</sup>). Consequently, the enzymatic O-GlcNAcylation of the full-length recombinant Tau appears to be highly inefficient with the tested OGT isoforms, which could arise from limited accessibility or incorrect anchoring of the enzyme. A parallel can be drawn here with the PP2A phosphatase – the major dephosphorylating enzyme of Tau, which is down-regulated in AD-brain,<sup>13</sup> and has a similar domain structure as the OGT (Fig. 2).<sup>14</sup> PP2A is a multimeric enzyme composed of a dimeric core (PP2A<sub>D</sub>) that can associate with a B-type regulatory subunit to form the trimeric PP2A<sub>T55</sub> complex (Fig. 2B).<sup>15</sup> This B-type subunit indeed regulates



**Fig. 2** Comparison between the three-dimensional structures of a model of ncOGT<sup>5,6</sup> (A, PDB ID: 1W3B and 2VSN) and the PP2A phosphatase<sup>14</sup> (B, PDB ID: 3DW8). The scaffold subunits are depicted in grey, the catalytic subunit in green and the B55 $\alpha$  regulatory subunit of PP2A in red.

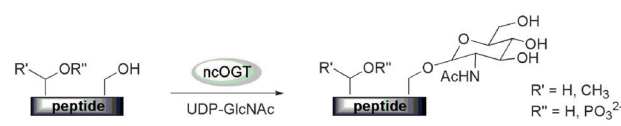
the catalytic activity and specificity of PP2A<sub>D</sub> towards the target phospho-proteins.<sup>16</sup> Different regulatory subunits of OGT have been suggested, but the correct one for Tau as a substrate is not yet identified.<sup>17</sup>

In order to get around this specific recognition issue between the full-length Tau and our OGT enzymes without the corresponding subunit, we decided to further assess putative GlcNAc sites by peptide mapping. The intrinsic unfolded nature of Tau protein argues in favor for the use of peptides as substrates for studying Tau posttranslational modifications and (dys)functions.<sup>18</sup> Furthermore, peptides carrying posttranslational modifications are easily accessible by solid phase peptide synthesis (SPPS), which allows to probe the influence of a specific modification on additional ones.

### *In vitro* O-GlcNAcylation of Tau peptides and identification of O-GlcNAc sites

For screening purposes, several unmodified sequences of overlapping peptides encompassing the proline-rich, the first repeat microtubule-binding (R1) and the carboxy-terminal domains of Tau protein (Table 1 and Table S1, ESI<sup>†</sup>) were obtained by SPPS. The primary sequence of these peptides bears the known phosphorylation sites of Tau including the AT8 (S199/S202/T205), AT100 (T212/S214/T217), AT180 (T231/S235), 12E8 (S262) and PHF-1 (S396/S400/S404) epitopes (Fig. 1).

The synthetic Tau peptides were incubated at 37 °C for 2 days with recombinant ncOGT and a 10-fold excess of UDP-GlcNAc or without UDP-GlcNAc as a control. Screening of O-GlcNAc transferase activity was performed by both MALDI-TOF MS and reverse phase chromatography analyses of the crude reaction mixtures. Five peptides encompassing three distinct regions gave an additional *m/z* signal as

**Table 1** Tau peptides for the screening of O-GlcNAc transferase activity


Peptide name <sup>a</sup>	Peptide sequence	% O-GlcNAc
Tau[194-209]	RSGYSSPGSPGTPGS <sup>208</sup> R	ca. 5
Tau[194-209]pS202	RSGYSSPGpSPGTPGS <sup>208</sup> R	n.d.
Tau[194-209]pT205	RSGYSSPGSPGpTPGS <sup>208</sup> R	n.d.
Tau[198-217]	SSPGSPGTPGS <sup>208</sup> RSRTPSLPT	< 1
Tau[205-220]	TPGSRSRTPSLPTPPT	—
Tau[224-239]	KKVAVVRTPPKSPSS <sup>238</sup> A	ca. 10
Tau[229-244]	VRTPPKSPSS <sup>238</sup> AKSRLQ	< 1
Tau[354-369]	KNVKSKIGSTENLKHQ	—
Tau[392-411]	IVYKSPVVS <sup>400</sup> GDTSPRHLSNV	8.2
Tau[392-411]pS396	IVYKpSPVVS <sup>400</sup> GDTSPRHLSNV	4.7
Tau[392-411]pS400	IVYKSPVpSGDTSPRHLSNV	—
Tau[392-411]pS404	IVYKSPVVS <sup>400</sup> GDTpSPRHLSNV	2.7

<sup>a</sup> Sequences of peptides used in this study and position of the first and the last peptide residues in the full-length Tau protein (Tau441 isoform numbering). Phosphorylation sites probed in this study are indicated (pS/pT). The O-GlcNAc transferase activity is given (% O-GlcNAc) and modified sites are underlined in the peptide sequence. Note that ncOGT activity on the Tau[392-411] series was performed with 0.1 mM of peptide instead of 1 mM for the others. In these latter conditions, Tau[392-411] was found to be 27% glycosylated. N.d., not determined.

compared to control experiments, with a mass increment of +203.08 Da matching the incorporation of an O-GlcNAc group (Fig. S2 and Table S2, ESI<sup>†</sup>). The peptides that revealed activity towards ncOGT included the Tau[194-209]/Tau[198-217] and the Tau[224-239]/Tau[229-244] overlapping regions of the proline-rich domain, as well as the Tau[392-411] region of the carboxy-terminal domain (Table 1).

RP-HPLC of the crude reaction mixtures revealed a separate peak in the chromatogram only when the Tau[392-411] peptide was used as a substrate. Quantification of this additional peak, which was assigned to the O-GlcNAcylated peptide by MALDI-TOF MS (Fig. S3, ESI<sup>†</sup>), indicated an overall O-GlcNAc content of 8.2% (Table 1). ECD-coupled MS/MS analysis of the purified O-GlcNAc fraction (Fig. 3C) and subsequent NMR spectroscopy using <sup>1</sup>H-<sup>1</sup>H TOCSY and <sup>1</sup>H-<sup>13</sup>C HMQC experiments (Fig. 4A and B) unambiguously confirmed the GlcNAcylation of the S400 residue. It is important to note that this residue, recently identified by MS on brain purified Tau,<sup>8</sup> was the unique O-GlcNAc site among a total of five Ser or Thr residues on the C-terminal peptide sequence.

The covalent attachment of the O-GlcNAc moiety to the S400 residue was further supported by significant shifts of the S400 <sup>13</sup>C $\alpha$  and <sup>13</sup>C $\beta$  signals (Fig. 4B). Moreover, chemical shift perturbations of amide signals upon O-GlcNAcylation were detected for the residues V399 to S404 (Fig. 4C and D). No modification was observed for remote residues, indicating that the O-GlcNAc moiety has a limited impact on the overall peptide structure. The observation of the sole sequential H<sub>N</sub>/H $\alpha$  NOEs contacts along the entire peptide sequence indicates an extended structure that remains unchanged upon S400 O-GlcNAcylation.

For the Tau[194-209]/Tau[198-217] and Tau[224-239]/Tau[229-244] overlapping peptides, only poorly resolved additional signals were observed by RP-HPLC (Fig. S3, ESI<sup>†</sup>). O-GlcNAc-enriched fractions were obtained by collecting the

head fraction of the main peak and analyzed by ECD MS/MS (Fig. 3A and B) and NMR. Similarly to S400, S208 and S238 were identified as two additional O-GlcNAc sites in Tau[194-209] and Tau[224-239] peptides, respectively, and were confirmed in the overlapping Tau[198-217] and Tau[229-244] peptides (Fig. S4, ESI<sup>†</sup>).

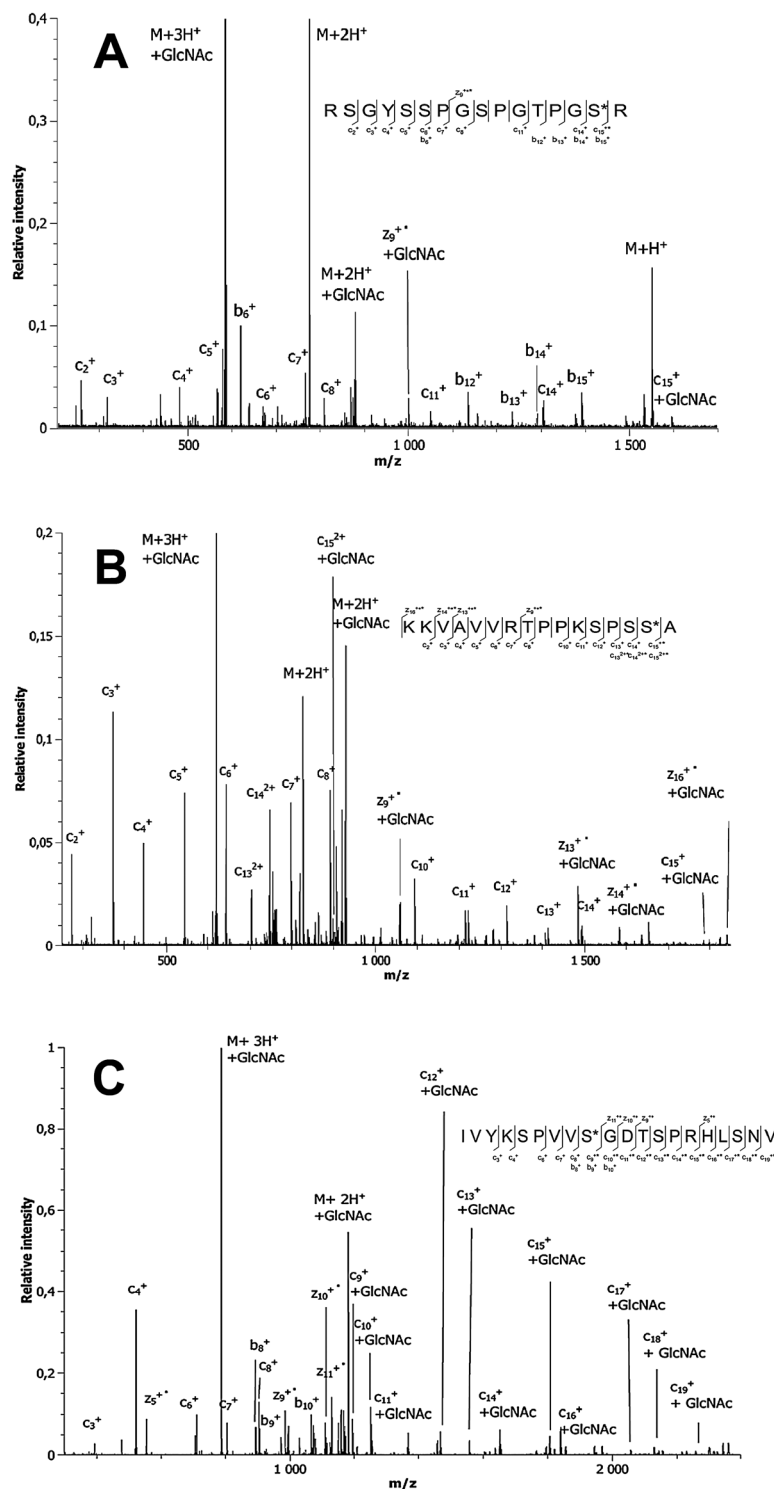
The finding of two novel O-GlcNAc sites in the proline-rich region of Tau which is implicated in the regulation of microtubule binding through the repeat domain and contains a number of phosphorylation sites, in particular AD-specific phospho-epitopes such as the AT100 (T212/S214), AT180 (T231/S235) and AT8 (S202/T205),<sup>19</sup> potentially indicates a regulation of the phosphorylation state by both O-GlcNAcylated S208 and S238 located in their near vicinity.

#### Reciprocal down-regulation of phosphorylation and O-GlcNAcylation in the C-terminal region of Tau

Subsequently we focused on the reciprocal impact of phosphorylation and O-GlcNAc modification in the Tau[392-411] peptide. Here, the use of peptide substrates allows for a direct identification and quantification of posttranslational modification sites and levels. It is important to note that recombinant enzymes used in this study exhibit the same site-specificities in the peptides as in the full-length Tau protein.<sup>9,10</sup>

#### Effect of constitutive S396, S400 or S404 phosphorylation on Tau[392-411] O-GlcNAcylation by ncOGT

O-GlcNAc incorporation in synthetic phospho-peptides containing either pS396, pS400 or pS404 with the recombinant ncOGT was assessed by MALDI-TOF MS and RP-HPLC. As expected, phosphorylation of S400 completely abolished O-GlcNAcylation while phosphorylation of S396 and S404 equally reduced the O-GlcNAcylation level by a factor of 2 and 4, respectively (Table 1 and Fig. S5, ESI<sup>†</sup>). Consequently,

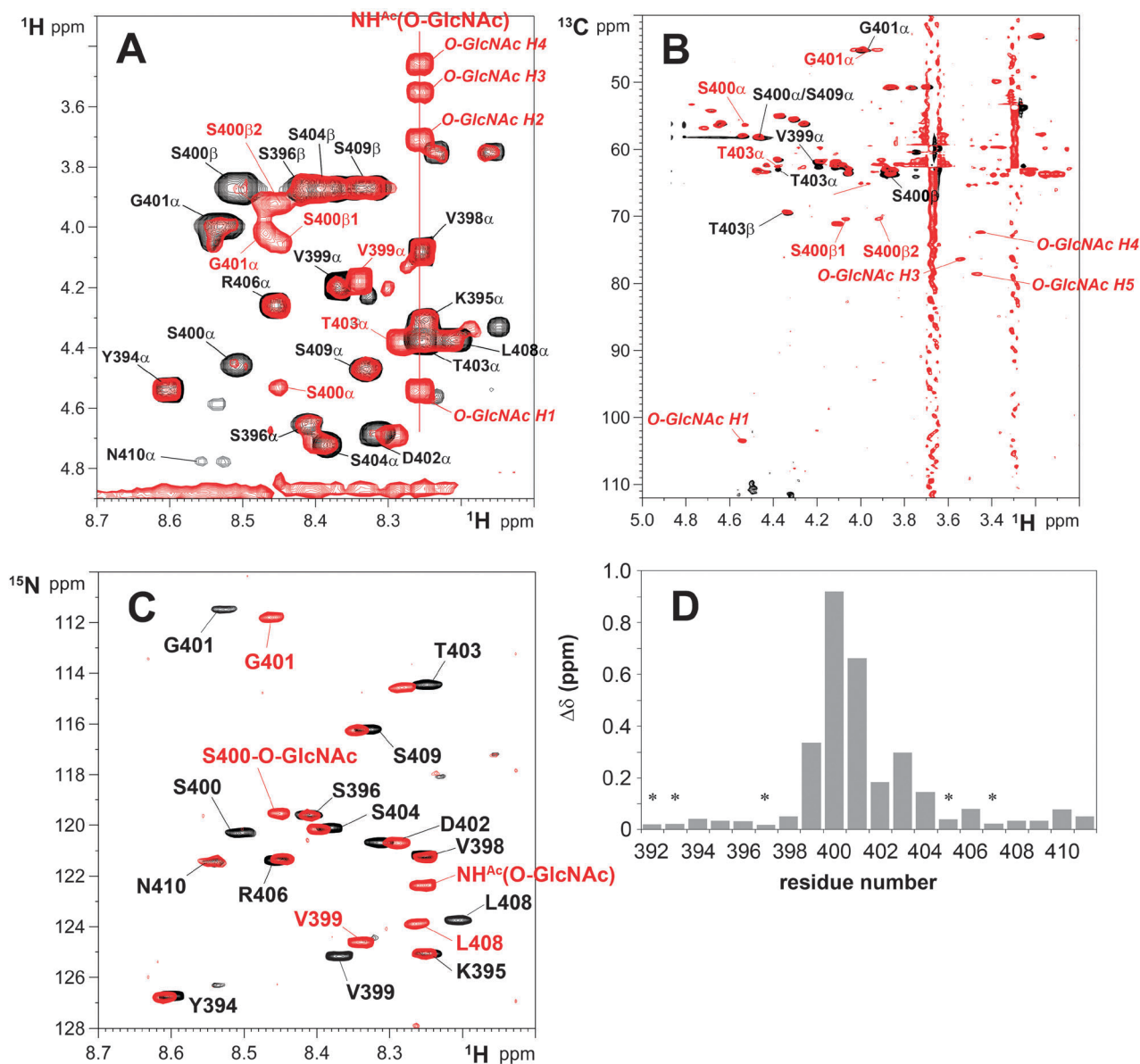


**Fig. 3** Localization of O-GlcNAcylation sites in Tau[194-209] (A), Tau[224-239] (B) and Tau[392-411] (C) using  $c^+$  and  $z^+$  ions produced by ECD-based experiments on nanoESI-Qh-FT-ICR mass spectrometer. Triply charged ions ( $m/z$  584.943 in A,  $m/z$  619 in B, and  $m/z$  786.747 in C with  $\Delta m = 1$  ppm) including GlcNAc modifications were selected for ECD spectra. The fragment ions incremented by the mass of a GlcNAc moiety allow precise localization of the modified amino acids, *i.e.* S208, S238 and S400 in Tau[194-209] (A), Tau[224-239] (B) and Tau[392-411] (C), respectively. \* Annotation indicates the presence of the GlcNAc moieties on peptide sequences and on fragment ions.  $M + 3H^+ + \text{GlcNAc}$  and  $M + 2H^+ + \text{GlcNAc}$  represent the peptide ions including the GlcNAc moieties.  $M + 2H^+$  and  $M + H^+$  represent the peptide ions without the GlcNAc moieties.

we have shown that phosphorylation of adjacent residues reduces the S400 O-GlcNAcylation in a direct manner.

To further evaluate the impact of O-GlcNAcylation on the phosphorylation levels, *in vitro* phosphorylation experiments





**Fig. 4** Identification of S400 as an O-GlcNAc site in Tau[392-411] peptide by NMR spectroscopy. Overlay of  $^1\text{H}$ - $^1\text{H}$  TOCSY (A),  $^1\text{H}$ - $^{13}\text{C}$  HMQC (B) and  $^{15}\text{N}$ - $^1\text{H}$  HSQC (C) spectra of the peptide in its non-glycosylated (black) and O-GlcNAcylated (red) forms. (D) Mapping of S400 O-GlcNAcylation in Tau[392-411] peptide by combination of  $\text{H}_\text{N}$  and  $\text{H}\alpha$  chemical shift variations ( $\Delta\delta$ ). Asterisks indicate values obtained from  $\text{H}\alpha$  resonances only due to a lack of detectable  $\text{H}_\text{N}$  signals.

**Table 2** Reduction of CDK- and GSK3 $\beta$  kinase activity by S400 O-GlcNAcylation

Peptide name	CDK phosphorylation on S404 <sup>a</sup> (%)	GSK3 $\beta$ phosphorylation on S396/S400 (%)
Tau[392-411]	61.3	<1
Tau[392-411]pS396	45.0	—
Tau[392-411]pS400	12.5	—
Tau[392-411]pS404	—	55
Tau[392-411]gS400	41.5	0
Tau[392-411]gS400/pS404	—	0

<sup>a</sup> CDK- and GSK3 $\beta$ -mediated phosphorylation of the peptide from the C-terminus of Tau in either its non-modified, phosphorylated (pS) or glycosylated (gS) forms. O-GlcNAc peptides were obtained by enzymatic treatment with ncOGT of the non-modified peptide and subsequent RP-HPLC purification. Phosphorylation of the gS400 peptide on S404 was performed by incubation with the CDK2/cyclinA3 kinase and RP-HPLC purification.

of Tau[392-411] peptides with recombinant CDK2/cyclinA3 or GSK3 $\beta$  kinases, known to phosphorylate Tau *in vivo*, were performed.

### Effect of S400 O-GlcNAcylation on S404 CDK2/cyclinA3-directed phosphorylation in Tau[392-411] peptide

In analogy to preliminary experiments, the purified CDK2/cyclinA3 complex phosphorylates S404 in the Tau[392-411] peptide to a level of 61.3% after overnight incubation at 22 °C (Fig. S6, ESI $\dagger$ ). This fraction decreased to 45% and 12.5% when we used phospho-peptides modified synthetically at the S396 or S400 position, respectively (Table 2 and Fig. S6, ESI $\dagger$ ). S400 O-GlcNAcylation reduced similarly the S404 phosphorylation level to 41.5% (Table 2 and Fig. S7A, ESI $\dagger$ ). Subsequent kinetic experiments demonstrate that S400 O-GlcNAcylation decreases the S404 phosphorylation rate by 27.5% (Fig. S7B, ESI $\dagger$ ).

### Effect of S400 O-GlcNAcylation on GSK3 $\beta$ -mediated phosphorylation in Tau[392-411] peptide

To probe the activity of an AD-relevant kinase which was found to phosphorylate Tau in cell or mouse models as well as *in vitro*<sup>20</sup> on modified peptide substrates, we examined the GSK3 $\beta$ -mediated phosphorylation of four different Tau[392-411] peptides. These included the non-modified peptide, as well as the peptides containing a phosphorylation at S404, a O-GlcNAcylation at S400 as well as both modifications, which were incubated at 30 °C overnight with recombinant GSK3 $\beta$ . An additional peak with a significantly shorter retention time was detected by RP-HPLC exclusively for the peptide with pS404, which was identified as the triple phosphorylated peptide at S404/S400/S396 by MALDI-TOF MS (Table 2 and Fig. S7C, ESI $\dagger$ ). S404 priming appears to be essential for the GSK3 $\beta$  kinase activity in this Tau region, since the non-modified peptide did not exhibit comparable phosphorylation levels. Additionally, S400 O-GlcNAcylation completely abolishes the GSK3 $\beta$ -mediated phosphorylation by interrupting the sequential phosphorylation process (Fig. S7C, ESI $\dagger$  and Fig. 5).

Here, we have shown that O-GlcNAc linkage on S400 both negatively regulates priming on S404 by CDK2/cyclinA3 and suppresses sequential phosphorylation on S404/S400/S396 by GSK3 $\beta$ . Implications of such a disruption in the physiopathological phosphorylation processes could be of utmost importance in the context of AD, since GSK3 $\beta$  is a major kinase for Tau phosphorylation at AD-relevant Pro-directed sites<sup>21</sup> where CDK5 plays a role as a priming kinase for GSK3 $\beta$ .<sup>22</sup>

### S208 O-GlcNAcylation has no effect on CDK2/cyclinA3-directed phosphorylation in Tau[194-209] peptide

As for the C-terminal peptide, preliminary phosphorylation experiments with the purified CDK2/cyclinA3 kinase were performed with the Tau[194-209] series (unmodified or quantitatively phosphorylated on S202 or T205 by SPPS) using RP-HPLC analysis of crude phosphorylation mixtures and subsequent MALDI-TOF MS analysis of isolated fractions. A more complex phosphorylation pattern was observed in this case due to the presence of multiple Ser/Thr-Pro substrates (S199, S202 and T205) that have previously been identified on full-length Tau using NMR spectroscopy (Fig. S8, ESI $\dagger$ ).<sup>9</sup> Low levels of mono-phosphorylated peptides (6.95% for the non phosphorylated substrate) were detected after overnight incubation at 22°C with 2 mM ATP. In contrast, a mixture of three doubly and one triply phosphorylated peptides were generated in significant amounts (53.7% and 39.4%, respectively, for the non-phosphorylated substrate). Furthermore, S208 O-GlcNAcylation does not induce any significant modification of the CDK-mediated phosphorylation pattern (Fig. S9, ESI $\dagger$ ) or any change in the formation kinetics of the various phosphorylated species (data not shown).

In contrast to the C-terminal region encompassed in the Tau[392-411] peptide where S400 O-GlcNAcylation down-regulates phosphorylation of the downstream S404 site, S208 O-GlcNAcylation has no significant impact on CDK activity towards the upstream S199/S202/T205 epitope. Based on the structure of the CDK2/cyclinA3 in complex with a model peptide substrate,<sup>24</sup> the regulation of CDK activity by upstream glycosylation could be explained by a preferential recognition of the substrate through the N-terminus of its phosphorylation-prone Ser/Thr-Pro motif whereas the C-terminal part points away from the catalytic site except a basic residue in the P + 3 position, which makes contacts with the conserved phospho-T160. Hence, these observations indicate that the impact of an O-GlcNAc moiety on phosphorylation patterns seems to be controlled by the sequence of the substrate, *i.e.* the position of phosphorylation-prone sites and by the kinase specificities.

## Material and methods

### Peptide synthesis, purification and characterization

Peptides were synthesized on a SyroXP-I peptide synthesizer (Multi-SynTech, Witten, Germany) according to standard Fmoc/tBu chemistry using TBTU/HOBt and preloaded resins (Novabiochem). Fmoc-Ser/Thr(PO(OBzl)OH)-OH (Bachem) was activated with HATU/DIEA and coupled manually to the

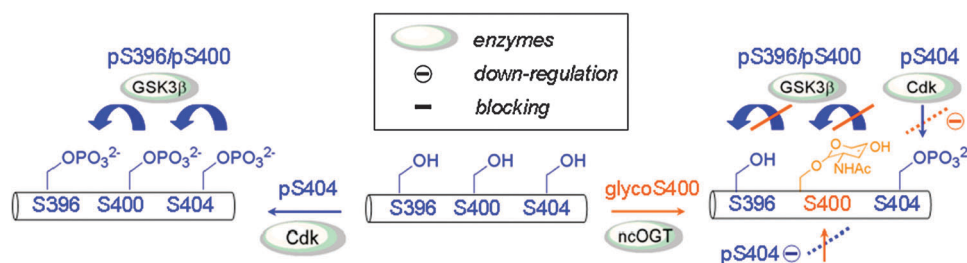


Fig. 5 Reciprocal relationship between phosphorylation and O-GlcNAcylation in Tau[392-411] peptide.



resin. DIEA was added in 3-fold excess with respect to the amino acid and HATU. The reaction time was extended to 6 hours. A mixture of DBU and piperidine (2% each) in DMF was used for Fmoc deprotection. Peptides were cleaved from the resin by treatment with TFA/TIS/H<sub>2</sub>O (95/2.5/2.5) for 3 hours followed by precipitation with cold diethyl ether. Purification was carried out by preparative reversed phase HPLC on a Knauer Smartline system (Knauer, Berlin, Germany) equipped with a Luna C8 column (10  $\mu$ m, 250  $\times$  21.20 mm; Phenomenex, Torrance, CA, USA) running with acetonitrile/0.1% TFA and water/0.1% TFA gradient at 20 ml min<sup>-1</sup>. Purified peptides were characterized by analytical HPLC and HRMS. The analytical HPLC was carried out with a VWR-Hitachi Elite LaChrome system (VWR, Darmstadt, Germany) equipped with a Luna C8 column (5  $\mu$ m, 250  $\times$  4.6 mm; Phenomenex, Torrance, CA, USA). Peptide mass to charge ratios were measured by using an Agilent 6210 ESI-TOF (Agilent Technologies, Santa Clara, CA, USA) (Table S1, ESI<sup>†</sup>).

### Expression and purification of the human O- $\beta$ -N-acetylglucosaminyl-transferase

BL21(DE3) *E. coli* strains were transformed with the pET24d plasmid (Novagen, Merck) carrying the human *ogt*.<sup>23</sup> Plasmid for the recombinant production of human ncOGT in *Escherichia coli* was obtained from the laboratory of S. Walker (Harvard Medical School). Freshly plated bacteria were picked up to inoculate a 20 ml culture which was grown at 37 °C overnight. The 20 ml culture then inoculated 2 l culture in LB rich medium that allowed to grow at 37 °C for 3 hours until OD at 600 nm reached 0.8–0.9, then the culture was cooled down to 16 °C. The protein induction was performed at 16 °C for 24 hours upon addition of 0.2 mM IPTG. Cells were harvested at 6000 $\times$ g for 30 min at 4 °C and the pellet was resuspended in 100 ml extraction buffer (50 mM KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub> pH 7.60, 250 mM NaCl, 40 mM imidazole, 1% Triton X-100 complemented with a protease inhibitor cocktail). Cell lysis was performed by incubation of the cell suspension with 0.3 mg DNase I and 0.1 mg RNase A at 4 °C for 20 min followed by a sonication step. Soluble proteins were isolated from the bacterial extract by centrifugation at 25 000  $\times$  g for 25 min at 4 °C. An affinity purification step was performed on a 5 ml-HiTrap Chelating column (GE Healthcare) charged with Ni<sup>2+</sup> ions. The full-length OGT was eluted with 250 mM imidazole while a truncated form was eluted with 110 mM imidazole. Homogenous fractions containing the full-length recombinant ncOGT as checked on a 10% polyacrylamide SDS gel were pooled (final volume 10 ml) and dialysed at 4 °C overnight in 2 l of sample buffer (50 mM KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub> pH 7.60, 150 mM NaCl, 1 mM EDTA). After dialysis, fractions were supplemented with 0.5 mM Tris(hydroxypropyl)phosphine (THP) and stored at –80 °C. Starting from 2 l cultures and following this purification scheme we obtained 4.5 mg of ncOGT (Fig. S1, ESI<sup>†</sup>).

### In vitro O-GlcNAcylation of Tau peptides and full-length Tau protein by ncOGT

Tau peptides or full-length protein were incubated at 2 mM or 200  $\mu$ M, respectively, with ncOGT at 0.5 mg ml<sup>-1</sup> (*i.e.* 4.7  $\mu$ M) in the reaction buffer (50 mM KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub> pH 7.60,

150 mM NaCl, 1 mM EDTA, 0.5 mM THP, 12.5 mM MgCl<sub>2</sub>, 10 mM UDG-GlcNAc). Reactions were performed at 37 °C for 24 hours then an equal volume of ncOGT solution and 10 mM UDP-GlcNAc were added and the reactions were allowed to continue for 24 hours at 37 °C. Control experiments were performed under the same conditions without UDP-GlcNAc (Fig. S3, ESI<sup>†</sup>). For the determination of O-GlcNAcylation levels of Tau[392–411] peptides in their non- or mono-phosphorylated (at S396 or S400 or S404) forms, three independent experiments were performed with incubation of peptides at 200  $\mu$ M with ncOGT at 0.5 mg ml<sup>-1</sup> for 1 day at 37 °C with 1 mM UDP-GlcNAc, then an equal volume of ncOGT and 1 mM UDP-GlcNAc were added and the reaction allowed to continue at 37 °C for 1 day. Control experiments were performed under the same conditions without UDP-GlcNAc.

### Expression and purification of the CDK2/cyclinA3 complex and GSK3 $\beta$ kinase

CDK2 phosphorylated at T160 was produced in *E. coli* by coexpression of human GST-CDK2 and *S. cerevisiae* GST-Cak1 using a pGEX vector (GE Healthcare), as described.<sup>24,25</sup> Human CycA3 (residues 174–432) lacking the destruction box can activate CDK2.<sup>24,25</sup>

BL21(DE3) *E. coli* strains are transformed with a pGEX vector (GE Healthcare) carrying the human *gsk3 $\beta$* . The culture was performed in 0.5 l LB at 37 °C for 3 hours until O.D. at 600 nm reached 0.8–0.9, then the culture was cooled down to 20 °C and protein induction was performed at 20 °C overnight upon addition of 0.2 mM IPTG. Cells were harvested at 6000 $\times$ g for 30 min at 4 °C and the pellet was resuspended in 40 ml extraction buffer (PBS, 10% glycerol, 1% Triton X-100, 10 mM EDTA, 2 mM DTT complemented with protease inhibitor cocktail). Cell lysis was performed by incubation of the cell suspension with 5 mg lysozyme, 0.15 mg DNase I and 0.05 mg RNase A at 4 °C for 20 min followed by a sonication step. Soluble proteins were isolated by centrifugation at 25 000 $\times$ g for 25 min at 4 °C. An affinity purification step was performed using 0.3 ml of glutathione sepharose resin (GE Healthcare). Resin beads were incubated with the soluble extract at 4 °C for 3 hours and extensively washed with the extraction buffer, then equilibrated in PBS, 1 mM EDTA. The GST-GSK3 $\beta$  fusion protein on resin beads was freshly used for the phosphorylation reactions of Tau[392–411] peptides.

### In vitro phosphorylation of Tau peptides by CDK2/cyclinA3 and GSK3 $\beta$

Tau peptides were incubated at final concentrations of 0.1 mM with the recombinant CDK2/cyclinA3 complex and 1 mM ATP at 22 °C overnight or for 1, 2, 6, 8, 16 and 24 hours for kinetics measurement. Tau[392–411] peptide in either its non-phosphorylated form or mono-phosphorylated on S396, S400 or S404 was used in preliminary experiments to determine the phosphorylation site(s) (Fig. S6, ESI<sup>†</sup>). The detection by reverse phase chromatography at 215 nm of a unique additional peak for all except the phospho-S404 Tau[392–411] peptide indicates the presence of a unique phosphorylation site. The retention time of the corresponding peak matches with that of the constitutively phospho-S404 Tau[392–411]

peptide obtained by solid phase synthesis (Fig. S6, ESI†) indicating that the CDK-mediated phosphorylation site is the S404 residue. Furthermore, the MALDI-TOF MS analysis of the isolated corresponding fraction shows a mass increment of +80 Da as compared to the mass of the non-phosphorylated peptide confirming the presence of a phosphate moiety (Fig. S6, ESI†). CDK-mediated phosphorylation of Tau[392-411] peptides either in its non-glycosylated or S400-O-GlcNAc state was performed in the same conditions (Fig. S7A,B, ESI†) in three independent experiments.

Tau[194-209] peptides were incubated at final concentrations of 0.1 mM with recombinant CDK2/cyclinA3 and 2 mM ATP at 22 °C for 24 hours for kinase activity measurement. Tau[194-209] peptide in either its non-phosphorylated form or mono-phosphorylated on S202 or T205 was used in preliminary experiments to characterize the phosphorylation pattern (Fig. S8, ESI†). Mono-, di- and tri-phosphorylated peptides were found as assessed by MALDI-TOF MS through the detection of mass increments of +80, +160 and +240 Da, respectively, as compared to the mass of the non-phosphorylated peptide (Fig. S8, ESI†). CDK-mediated phosphorylation of Tau[194-209] peptide either in its non-glycosylated or S208-O-GlcNAc state was performed under the same conditions (Fig. S9, ESI†).

Tau peptides were incubated at final concentrations of 0.1 mM with the recombinant GSK3β kinase and 4 mM ATP at 30 °C overnight. GSK3β phosphorylation reactions of Tau[392-411] peptides in its non-modified, phospho-S404, S400-O-GlcNAc or S400-O-GlcNAc/phospho-S404 states were performed in two independent experiments as well as controls without ATP (Fig. S7C, ESI†).

### NMR spectroscopy

For NMR experiments, peptides were dissolved at 2 mM in a buffer containing 25 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 6.8, 25 mM NaCl, 2.5 mM EDTA and 5% D<sub>2</sub>O. All <sup>1</sup>H spectra were calibrated with 1 mM sodium 3-trimethylsilyl-3,3',2,2'-d<sub>4</sub>-propionate as a reference. For peptide assignment, standard water-flipback NOESY and TOCSY experiments were recorded with 400 and 69 ms mixing times, respectively, with 2048 and 512 points in both proton dimensions. <sup>1</sup>H-<sup>15</sup>N HSQC spectra at nitrogen natural abundance were recorded with 64 scans per increment, with 2048 and 256 points in the proton and nitrogen dimensions, respectively. <sup>1</sup>H-<sup>13</sup>C HSQC spectra were recorded with 864 scans per increment, with 1024 and 512 points in the proton and carbon dimensions, respectively.

The chemical shift perturbations of amide individual resonances in peptides were calculated with eqn (1) taking into account the relative dispersion of the proton and nitrogen chemical shifts (1 ppm and 20 ppm, respectively).

$$\Delta\delta \text{ (ppm)} = \sqrt{[\Delta\delta(^1\text{H})]^2 + 0.05[\Delta\delta(^{15}\text{N})]^2} \quad (1)$$

### Reverse phase chromatography and MALDI-TOF mass spectrometry analyses of crude reaction mixtures

Crude reaction mixtures were analyzed (after protein precipitation at 75 °C for 15 min) by reverse phase chromatography on a C18 column (μRPC C2/C18 4.6 × 100 mm, GE Healthcare)

equilibrated in 0.1% TFA or 10 mM ammonium acetate buffer pH 5.5. 10 nmoles of crude peptide mixtures were injected and separated by a linear gradient of 0–60% acetonitrile in 20 ml (*i.e.* 12 column volume) at room temperature. Determination of the O-GlcNAc content was estimated by peak integrations at 215 nm of the O-GlcNAcylated and unmodified peptides. The signal integration of the control experiment was subtracted to the peak integration value corresponding to the O-GlcNAc-modified peptide to take into account the signal corresponding to peptide contaminants that elute at the same retention volume than the O-GlcNAcylated peptide. The crude reaction mixtures were analyzed by MALDI-TOF MS (Voyager DE-STR, Applied Biosystems) on a α-cyano-4-hydroxycinnamic acid matrix after ZipTip<sup>®</sup>-C18 desalting (Millipore). O-GlcNAc-enriched fractions obtained by reverse phase chromatography of the crude reaction mixtures were analyzed by mass spectrometry after lyophilization and dissolution in 20 μl deionized water.

### Identification of O-GlcNAc sites by Electron Capture Dissociation mode on a nanoESI-Qh-FT-ICR mass spectrometer

The samples were analyzed with an Apex Qe 9.4 T Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (Bruker Daltonics, Bremen, Germany). The FT-ICR mass spectrometer is equipped with a nano-electrospray source. Detection was carried out in a positive mode. A potential of 1.1 kV was applied on the needle (PicoTip Emitter, New Objective, Woburn, USA). The detection parameters were broadband detection, 512 K acquisition size. Ions were accumulated in the hexapole during 1 s, and in the collision cell 0.01 s and 2 s, respectively, for MS and MS/MS. Electron Capture Dissociation mode was employed for MS/MS experiments using following parameters: ECD pulse length 0.03 s, ECD bias 2.0 V, ECD heater 20.0 V and ECD heater 1.7 A.

The O-GlcNAc-enriched fraction of each Tau[392-411], Tau[194-209] and Tau[224-239] peptides were analyzed using Electron Capture Dissociation mode on a nanoESI-Qh-FT-ICR mass spectrometer in order to identify the O-GlcNAcylation sites. S400 residue was identified in Tau[392-411] peptide as the O-GlcNAcylation site using fragment ion series observed on the ECD spectrum (Fig. 2C). In details, b<sup>+</sup>, c<sup>+</sup> and z<sup>+•</sup> fragment ion series were detected as shown in Fig. 2C; more interestingly, c<sup>+</sup> ions incremented by GlcNAc modification from c<sub>9</sub><sup>+</sup> to c<sub>19</sub><sup>+</sup> were identified allowing the localization of the modification on the S400 amino acid (9th serine of the studied sequence). It can be noted that despite the difficulty reported in the literature<sup>26</sup> to obtain c<sup>+</sup> and z<sup>+•</sup> fragments near proline residue, in this spectrum the c<sub>6</sub><sup>+</sup> and c<sub>14</sub><sup>+</sup> + GlcNAc fragment ions could be observed. Relating to Tau[194-209] peptide, c<sup>+</sup>, b<sup>+</sup> and z<sup>+•</sup> fragment ions were detected in the ECD spectrum (Fig. 2A); in particular, two ions, z<sub>9</sub><sup>+•</sup> and c<sub>15</sub><sup>+</sup> were identified with mass shift corresponding to GlcNAc moieties. Using the informative c<sub>15</sub><sup>+</sup> + GlcNAc ion (*versus* c<sub>14</sub><sup>+</sup> ion), it can be concluded that the last serine of the sequence, S208, is glycosylated. Finally, c<sup>+</sup>, c<sup>2+</sup> and z<sup>+•</sup> fragment ions were highlighted on the ECD spectrum related to the glycosylated Tau[224-239] peptide (Fig. 2B).

The informative  $c_{15}^+$  and  $c_{15}^{2+}$  ions implemented with a mass shift related to the GlcNAc moiety allowed identifying the serine S238 glycosylated.

## Conclusions

Although O-GlcNAcylation of Tau was described more than ten years ago, the first O-GlcNAc site (S400) was directly identified only very recently.<sup>8</sup> This fact underscores the technical difficulties associated with the labile character and, probably, the low natural occurring O-GlcNAc modification rate for unambiguous identification by MS/MS methods.<sup>27</sup> In our studies, we have found that full-length Tau and a Tau fragment containing the proline-rich domain were not detectably modified by recombinant OGTs using NMR spectroscopy. Using small peptide substrates, we were able to verify the previously identified O-GlcNAc site S400. Additionally, we have located two novel O-GlcNAc sites in Tau peptides from the proline-rich domain at S208 and S238. Finally, we have shown the impact of O-GlcNAcylated S400 on further phosphorylation events in a peptide from the C-terminal part of Tau, which resembles an AD-specific phospho-epitope. Specifically, a reciprocity between phosphorylation and O-GlcNAcylation in the S396/S400/S404 region was determined.

Current efforts in our laboratory aim to access the modified full length Tau protein by semi-synthetic strategies,<sup>28</sup> which allows us to probe the activity of kinases with homogeneously modified Tau protein for detailed biophysical analysis.

## Acknowledgements

We thank Dr S. Walker (Harvard University, MA, USA) for the gift of the OGT plasmids.<sup>18</sup> The NMR (600 and 900 MHz) facilities are funded by the European community, the Centre National de la Recherche Scientifique (CNRS), the Région Nord-Pas de Calais (France), the University of Lille 1 and the Institut Pasteur de Lille. Financial support from the TGE RMN THC Fr3050 for conducting the research is gratefully acknowledged. The Mass Spectrometry facilities used for this study are funded by the European community (FEDER), the Région Nord-Pas de Calais (France), the IBISA (Infrastructures en Biologie Santé et Agronomie) network, the CNRS, and the University of Lille 1. M.B. and C.P.R.H. acknowledge the DFG (Emmy Noether grant HA 4468/2-1), the SFB 765 and the Innovationsfonds of the FU Berlin for financial support.

## Notes and references

- G. W. Hart, M. P. Housley and C. Slawson, *Nature*, 2007, **446**, 1017.
- F. Liu, J. Shi, H. Tanimukai, J. Gu, J. Gu, I. Grundke-Iqbal, K. Iqbal and C. X. Gong, *Brain*, 2009, **132**, 1820; S. A. Yuzwa, M. S. Macauley, J. E. Heinonen, X. Shan, R. J. Dennis, Y. He, G. E. Whitworth, K. A. Stubbs, E. J. McEachern, G. J. Davies and D. J. Vocadlo, *Nat. Chem. Biol.*, 2008, **4**, 483; T. Lefebvre, S. Ferreira, L. Dupont-Wallois, T. Bussiere, M. J. Dupire, A. Delacourte, J. C. Michalski and M. L. Caillet-Boudin, *Biochim. Biophys. Acta*, 2003, **1619**, 167; F. Liu, K. Iqbal, I. Grundke-Iqbal, G. W. Hart and C. X. Gong, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 10804; P. M. Fischer, *Nat. Chem. Biol.*, 2008, **4**, 448; C. X. Gong, F. Liu, I. Grundke-Iqbal and K. Iqbal, *J. Alzheimer's Dis.*, 2006, **9**, 1; Y. Liu, F. Liu, K. Iqbal, I. Grundke-Iqbal and C. X. Gong, *FEBS Lett.*, 2008, **582**, 359; L. A. Robertson, K. L. Moya and K. C. Breen, *J. Alzheimer's Dis.*, 2004, **6**, 489.
- C. Martinez-Fleites, Y. He and G. J. Davies, *Biochim. Biophys. Acta*, 2010, **1800**, 122; Y. Gao, L. Wells, F. I. Comer, G. J. Parker and G. W. Hart, *J. Biol. Chem.*, 2001, **276**, 9838.
- J. A. Hanover, S. Yu, W. B. Lubas, S. H. Shin, M. Ragano-Caracciola, J. Kochran and D. C. Love, *Arch. Biochem. Biophys.*, 2003, **409**, 287.
- M. Jinek, J. Rehwinkel, B. D. Lazarus, E. Izaurralde, J. A. Hanover and E. Conti, *Nat. Struct. Mol. Biol.*, 2004, **11**, 1001.
- A. J. Clarke, R. Hurtado-Guerrero, S. Pathak, A. W. Schuttelkopf, V. Borodkin, S. M. Shepherd, A. F. Ibrahim and D. M. van Aalten, *EMBO J.*, 2008, **27**, 2780; C. Martinez-Fleites, M. S. Macauley, Y. He, D. L. Shen, D. J. Vocadlo and G. J. Davies, *Nat. Struct. Mol. Biol.*, 2008, **15**, 764.
- C. S. Arnold, G. V. Johnson, R. N. Cole, D. L. Dong, M. Lee and G. W. Hart, *J. Biol. Chem.*, 1996, **271**, 28741.
- Z. Wang, N. D. Udeshi, M. O'Malley, J. Shabanowitz, D. F. Hunt and G. W. Hart, *Mol. Cell. Proteomics*, 2010, **9**, 153; S. A. Yuzwa, A. K. Yadav, Y. Skorobogatko, T. Clark, K. Vosseller and D. J. Vocadlo, *Amino Acids*, 2010, DOI: 10.1007/s00726-010-0705-1.
- L. Amniai, P. Barbier, A. Sillen, J. M. Wieruszkeski, V. Peyrot, G. Lippens and I. Landrieu, *FASEB J.*, 2009, **23**, 1146.
- I. Landrieu, L. Lacosse, A. Leroy, J. M. Wieruszkeski, X. Trivelli, A. Sillen, N. Sibille, H. Schwalbe, K. Saxena, T. Langer and G. Lippens, *J. Am. Chem. Soc.*, 2006, **128**, 3575; I. Landrieu, A. Leroy, C. Smet-Nocca, I. Huvent, L. Amniai, M. Hamdane, N. Sibille, L. Buée, J.-M. Wieruszkeski and G. Lippens, *Biochem. Soc. Trans.*, 2010, **38**, 1006; A. Leroy, I. Landrieu, I. Huvent, D. Legrand, B. Codeville, J.-M. Wieruszkeski and G. Lippens, *J. Biol. Chem.*, 2010, **285**, 33435.
- K. S. Kosik and H. Shimura, *Biochim. Biophys. Acta*, 2005, **1739**, 298; M. Morishima-Kawashima, M. Hasegawa, K. Takio, M. Suzuki, H. Yoshida, K. Titani and Y. Ihara, *J. Biol. Chem.*, 1995, **270**, 823; M. Morishima-Kawashima, M. Hasegawa, K. Takio, M. Suzuki, H. Yoshida, A. Watanabe, K. Titani and Y. Ihara, *Neurobiol. Aging*, 1995, **16**, 365; N. Sergeant, A. Delacourte and L. Buee, *Biochim. Biophys. Acta*, 2005, **1739**, 179; W. H. Stoothoff and G. V. Johnson, *Biochim. Biophys. Acta*, 2005, **1739**, 280.
- V. Dehennaut, X. Hanouille, J. F. Bodart, J. P. Vilain, J. C. Michalski, I. Landrieu, G. Lippens and T. Lefebvre, *Biochem. Biophys. Res. Commun.*, 2008, **369**, 539.
- G. Drewes, E. M. Mandelkow, K. Baumann, J. Goris, W. Merlevede and E. Mandelkow, *FEBS Lett.*, 1993, **336**, 425; M. Goedert, R. Jakes, Z. Qi, J. H. Wang and P. Cohen, *J. Neurochem.*, 1995, **65**, 2804; C. X. Gong, I. Grundke-Iqbal and K. Iqbal, *Neuroscience*, 1994, **61**, 765.
- Y. Xu, Y. Chen, P. Zhang, P. D. Jeffrey and Y. Shi, *Mol. Cell*, 2008, **31**, 873.
- V. Janssens and J. Goris, *Biochem. J.*, 2001, **353**, 417; V. Janssens, S. Longin and J. Goris, *Trends Biochem. Sci.*, 2008, **33**, 113.
- D. M. Virshup and S. Shenolikar, *Mol. Cell*, 2009, **33**, 537-545.
- W. D. Cheung, K. Sakabe, M. P. Housley, W. B. Dias and G. W. Hart, *J. Biol. Chem.*, 2008, **283**, 33935.
- N. L. Daly, R. Hoffmann, L. Otvos, Jr. and D. J. Craik, *Biochemistry*, 2000, **39**, 9039; W. J. Goux, L. Kopplin, A. D. Nguyen, K. Leak, M. Rutkofsky, V. D. Shanmuganandam, D. Sharma, H. Inouye and D. A. Kirschner, *J. Biol. Chem.*, 2004, **279**, 26868; F. A. Rojas Quijano, D. Morrow, B. M. Wise, F. L. Brancia and W. J. Goux, *Biochemistry*, 2006, **45**, 4638; C. Smet, J.-M. Wieruszkeski, L. Buée, I. Landrieu and G. Lippens, *FEBS Lett.*, 2005, **579**, 4159; C. Smet, A. Leroy, A. Sillen, J.-M. Wieruszkeski, I. Landrieu and G. Lippens, *ChemBioChem*, 2004, **5**, 1639.
- M. Goedert, R. Jakes, R. A. Crowther, P. Cohen, E. Vanmechelen, M. Vandermeeren and P. Cras, *Biochem. J.*, 1994, **301**(Pt 3), 871; M. Goedert, R. Jakes and E. Vanmechelen, *Neurosci. Lett.*, 1995, **189**, 167; M. G. Spillantini, R. A. Crowther and M. Goedert, *Acta Neuropathol.*, 1996, **92**, 42.
- J. Brownlee, N. G. Irving, J. P. Brion, B. J. Gibb, U. Wagner, J. Woodgett and C. C. Miller, *NeuroReport*, 1997, **8**, 3251; D. P. Hanger, K. Hughes, J. R. Woodgett, J. P. Brion and B. H. Anderton, *Neurosci. Lett.*, 1992, **147**, 58; S. Lovestone,

- C. H. Reynolds, D. Latimer, D. R. Davis, B. H. Anderton, J. M. Gallo, D. Hanger, S. Mulot, B. Marquardt and S. Stabel, *et al.*, *Curr. Biol.*, 1994, **4**, 1077; J. J. Lucas, F. Hernandez, P. Gomez-Ramos, M. A. Moran, R. Hen and J. Avila, *EMBO J.*, 2001, **20**, 27; K. Spittaels, C. Van den Haute, J. Van Dorpe, H. Geerts, M. Mercken, K. Bruynseels, R. Lasrado, K. Vandezande, I. Laenen, T. Boon, J. Van Lint, J. Vandenneede, D. Moechars, R. Loos and F. Van Leuven, *J. Biol. Chem.*, 2000, **275**, 41340–41349.
- 21 T. Li and H. K. Paudel, *Biochemistry*, 2006, **45**, 3125.
- 22 T. Li, C. Hawkes, H. Y. Qureshi, S. Kar and H. K. Paudel, *Biochemistry*, 2006, **45**, 3134.
- 23 B. J. Gross, B. C. Kraybill and S. Walker, *J. Am. Chem. Soc.*, 2005, **127**, 14588.
- 24 N. R. Brown, M. E. Noble, J. A. Endicott and L. N. Johnson, *Nat. Cell Biol.*, 1999, **1**, 438.
- 25 J. Welburn and J. Endicott, *Methods Mol. Biol. (Clifton, N.J.)*, 2005, **296**, 219.
- 26 N. Leymarie, E. A. Berg, M. E. McComb, P. B. O'Connor, J. Grogan, F. G. Oppenheim and C. E. Costello, *Anal. Chem.*, 2002, **74**, 4124.
- 27 N. Khidekel, S. B. Ficarro, P. M. Clark, M. C. Bryan, D. L. Swaney, J. E. Rexach, Y. E. Sun, J. J. Coon, E. C. Peters and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2007, **3**, 339; J. E. Rexach, P. M. Clark and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2008, **4**, 97.
- 28 C. P. R. Hackenberger and D. Schwarzer, *Angew. Chem., Int. Ed.*, 2008, **47**, 10030.



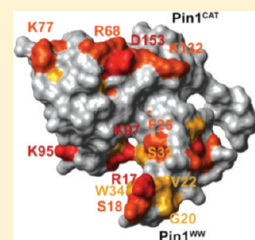
## Molecular Basis for an Ancient Partnership between Prolyl Isomerase Pin1 and Phosphatase Inhibitor-2

Furqan Sami,<sup>†</sup> Caroline Smet-Nocca,<sup>‡</sup> Meera Khan,<sup>‡</sup> Isabelle Landrieu,<sup>‡</sup> Guy Lippens,<sup>‡</sup> and David L. Brautigan<sup>\*,†</sup>

<sup>†</sup>Center for Cell Signaling and Department of Microbiology, University of Virginia School of Medicine, Box 800577-MSB7225, Charlottesville, Virginia 22908, United States

<sup>‡</sup>UMR 8576 CNRS-Unité de Glycobiologie Structurale et fonctionnelle RMN Biologique, Université des Sciences et Technologies de Lille, Bâtiment C9, 59655 Villeneuve d'Ascq Cedex, France

**ABSTRACT:** Pin1 is a prolyl isomerase that recognizes phosphorylated Ser/Thr-Pro sites, and phosphatase inhibitor-2 (I-2) is phosphorylated during mitosis at a PSpTP site that is expected to be a Pin1 substrate. However, we previously discovered I-2, but not phospho-I-2, bound to Pin1 as an allosteric modifier of Pin1 substrate specificity [Li, M., et al. (2008) *Biochemistry* 47, 292]. Here, we use binding assays and NMR spectroscopy to map the interactions on Pin1 and I-2 to elucidate the organization of this complex. Despite having sequences that are ~50% identical, human, *Xenopus*, and *Drosophila* I-2 proteins all exhibited identical, saturable binding to GST-Pin1 with  $K_{0.5}$  values of 0.3  $\mu$ M. The <sup>1</sup>H–<sup>15</sup>N heteronuclear single-quantum coherence spectra for both the WW domain and isomerase domain of Pin1 showed distinctive shifts upon addition of I-2. Conversely, as shown by NMR spectroscopy, specific regions of I-2 were affected by addition of Pin1. A single-residue I68A substitution in I-2 weakened binding to Pin1 by half and essentially eliminated binding to the isolated WW domain. On the other hand, truncation of I-2 to residue 152 had a minimal effect on binding to the WW domain but eliminated binding to the isomerase domain. Size exclusion chromatography revealed that wild-type I-2 and Pin1 formed a large (>300 kDa) complex and I-2(I68A) formed a complex of half the size that we propose are a heterotetramer and a heterodimer, respectively. Pin1 and I-2 are conserved among eukaryotes from yeast to humans, and we propose they make up an ancient partnership that provides a means for regulating Pin1 specificity and function.



Pin1 is the most widely studied member of the parvulin family of peptidyl-prolyl *cis*–*trans* isomerase proteins, with the unique feature of selectively recognizing the phosphorylated Ser/Thr-Pro motif in protein.<sup>1,2</sup> First characterized in 1996 as a regulator of the mitotic kinase NIMA in the fungus *Aspergillus nidulans*,<sup>3</sup> Pin1 is homologous to the previously reported *essential* gene in *Saccharomyces cerevisiae* (Ess1)<sup>4</sup> and to the *Drosophila* gene *dodo*.<sup>5</sup> The Pin1 protein comprises 163 amino acids in two distinct domains, an N-terminal WW domain (Pin1<sub>ww</sub>, residues 1–39) and a C-terminal catalytic domain (Pin1<sub>cat</sub>, residues 45–163), and three-dimensional structures have been determined by crystallography<sup>6</sup> and NMR spectroscopy.<sup>7</sup> Whereas the crystal structure shows the domains compressed together in a single, overall globular shape, the solution structure determined by NMR spectroscopy depicts the domains separated from one another in a more extended configuration. Isomerization of phosphoproteins by Pin1 is considered to modulate many biological processes, especially those involving key proteins such as cyclin D1,<sup>8</sup> c-jun,<sup>9</sup> c-Myc,<sup>10</sup> p53,<sup>11,12</sup> and tau.<sup>13–15</sup> The overexpression of Pin1 in various human tumors<sup>9,16–20</sup> and the recognition that its depletion from cells induces mitotic arrest have made Pin1 an attractive therapeutic target for drug development.<sup>21–23</sup> A first Pin1 inhibitor is juglone, which covalently modifies the Cys in the active site of the isomerase domain<sup>24</sup> but as a reactive compound lacks selectivity. Other Pin1 inhibitors have been described<sup>25,26</sup> but, as far as we know,

have not yet entered clinical trials. Because both WW and isomerase domains can bind phospho sites, it has been proposed that Pin1 uses simultaneous interaction with two different phospho sites in a particular substrate.<sup>27</sup> However, Ser16 in the WW domain is phosphorylated by PKA,<sup>28</sup> which then occupies the site where a sulfate ion binds in the three-dimensional (3D) structure.<sup>6</sup> Moreover, dynamic measurements by NMR spectroscopy indicate that dually phosphorylated peptides tend to interact only with the isomerase domain in Pin1.<sup>29</sup> The presumption has been that Pin1 acts as a monomer to isomerize phosphorylated sites in many different proteins, yet the basis for Pin1 substrate specificity is poorly understood.<sup>29</sup>

Inhibitor-2 (I-2) was discovered in 1976<sup>30</sup> as a thermostable protein that inhibited protein phosphatase activity and later was used to distinguish type 1 (I-2 sensitive) from type 2 (I-2 insensitive) protein Ser/Thr phosphatases.<sup>31</sup> I-2 is the most ancient of the more than 200 PP1 binding proteins, with recognized homologues in yeast (Glc8), *Drosophila*, *Xenopus*, and all mammals.<sup>32</sup> The most conserved feature of eukaryotic I-2 proteins is a Pro-X-Thr-Pro (PXTTP) sequence motif. The heterodimer of I-2 with PP1 was studied as an “MgATP-dependent phosphatase”, wherein the phosphorylation and

Received: April 12, 2011

Revised: June 21, 2011

Published: June 28, 2011

dephosphorylation of Thr73 in the PXTTP motif cause conformational activation of the bound PP1.<sup>33</sup> GSK3, MAPK, and CDK kinases phosphorylate this motif in biochemical assays.<sup>34–37</sup> We discovered a >25-fold increase in the level of PXTTP phosphorylation during mitosis,<sup>38</sup> catalyzed by CDK1:cyclinB1<sup>36</sup> in a reaction enhanced by Suc1,<sup>32</sup> making this site a potential Pin1 substrate. However, T73-phosphorylated I-2 was not a substrate and did not bind to Pin1, but the negative controls in this assay, using unphosphorylated or T73A I-2, showed formation of a Pin1–I-2 complex.<sup>39</sup> In the presence of I-2, compared to serum albumin as a control, GST–Pin1 binding to a panel of known mitotic phosphoprotein substrates was allosterically modified, with evidence of both enhanced and restricted binding.<sup>39</sup> These results showed that association of Pin1 with I-2 does not occlude its phosphopeptide binding sites but does alter substrate binding specificity. Our hypothesis is that functions of Pin1 and I-2 are interdependent.

Recent results have shown that I-2 acts as a critical regulator of cellular events related to mitosis. The protein is localized at centrosomes and serves to activate Nek2 kinase by inhibition of associated PP1.<sup>40</sup> Centrosomes lie at the base of the primary cilium, and I-2 is concentrated in the cilium, as seen by immunofluorescence microscopy.<sup>41</sup> Knockdown of I-2 prevents formation of the cilium and reduces the level of acetylation of tubulin in the cilium.<sup>41</sup> I-2 is a maternal gene in *Drosophila*, and the I-2 protein is concentrated in oocytes and is required for the proper execution of early mitosis in the syncytial embryo.<sup>42</sup> Hypomorphic embryos exhibit abnormal mitotic spindles and lagging chromosomes, with lethal consequences. The defects are rescued by dose-dependent transgenic expression of DI-2.<sup>42</sup> Likewise, in human epithelial cells, knockdown of I-2 caused defects in mitosis, with the appearance of lagging chromosomes and failed cytokinesis, resulting in multinucleated cells, phenotypes that were rescued by coexpression of I-2.<sup>43</sup> Direct binding of I-2 that activates Aurora A kinase<sup>44</sup> reduces the threshold for CDK activation at entry into mitosis.<sup>45</sup> The biological functions of I-2 can in part be attributed to inhibition of PP1, but that alone seems to be inadequate to account for these phenotypes.

In this study, we have explored the binding of I-2 to Pin1 at atomic resolution. We found that the binding of I-2 to Pin1 is conserved among species, despite the sequences of the proteins being only ~50% identical. We used NMR spectroscopy and reciprocal labeling to titrate and map the interactions on the surfaces of these two proteins. When we found that a single residue substitution in I-2 gave 50% binding at saturation compared to that of the wild type, we realized that multimers must be involved and used size exclusion chromatography to demonstrate that Pin1 and I-2 form a tetramer in solution. We used truncated and mutated proteins in binding assays to assign which regions of I-2 interact with the WW and isomerase domains in Pin1 to generate a model for this complex. We propose that Pin1 and I-2 functions are, like the proteins themselves, inextricably intertwined, and this opens new directions for future research.

## MATERIALS AND METHODS

Restriction enzymes and reagents for polymerase chain reaction were purchased from New England BioLabs. Oligonucleotides were synthesized by Integrated DNA Technologies. Affinity-purified sheep anti-HI-2, anti-GST, anti-XI-2, and anti-DI-2 antibodies have been described previously.<sup>42</sup> Buffers and

chemicals were purchased from Thermo Fisher Scientific. Phospho-S16Pin1 antibodies were from Cell Signaling Technologies. Alexa Fluor 680 goat anti-rabbit IgG was purchased from Invitrogen, and donkey anti-rabbit IRDye 800CW was purchased from LI-COR Biosciences (Lincoln, NE).

**Cloning and Bacterial Expression of Proteins.** Cloning of all I-2 proteins used in this study was performed according to the methods described previously.<sup>32</sup> Human Pin1, the WW domain (residues 1–44), and the isomerase domain (residues 46–163) were cloned in pGEX-4T-1 (Amersham Pharmacia Biotech, Freiburg, Germany). In vitro phosphorylation of purified recombinant GST–Pin1 and GST–WW domain was performed using the pure PKA catalytic subunit with a previously described method.<sup>46</sup>

*Escherichia coli* strain BL21-CodonPlus (DE3)-RIPL (Agilent Technologies) was transformed with pET-I-2 vectors or pGEX-4T-1 (Pin1) bacterial expression vectors and grown overnight at 37 °C in 10 mL of TB medium (1.2% tryptone, 2.4% yeast extract, 2% glucose, 0.017 M KH<sub>2</sub>PO<sub>4</sub>, and 0.072 M K<sub>2</sub>HPO<sub>4</sub>), 30 µg/mL kanamycin (for pET vectors), 30 µg/mL ampicillin (for pGEX vectors), and 20 µg/mL chloramphenicol. The culture was inoculated into 1.0 L of TB medium, including 30 µg/mL kanamycin or ampicillin and 20 mL of ethanol. The transformed cells were grown to an A<sub>600</sub> of 0.6 at 37 °C, and the expression of I-2 proteins was induced by addition of isopropyl 1-thio-β-D-galactopyranoside to the culture at a final concentration of 1 mM for 12–16 h at 18 °C. The bacteria were collected by centrifugation at 5000g for 10 min, and the cell pellet was washed by suspension in ice-cold PBS, centrifuged at 5000g for 10 min, and frozen at –80 °C.

**Purification of His6-Tagged I-2 Proteins.** The pellet was suspended by a buffer containing 50 mM MOPS–NaOH (pH 7.4), 300 mM NaCl, 10 mM imidazole, 10% glycerol, 1 mM DDT, one Complete EDTA-free tablet (Roche Applied Science) per 30 mL of buffer, and 1.0 mg/mL lysozyme and frozen at –80 °C. After being completely frozen, the pellet was thawed for 10 min in a 37 °C bath and sonicated for 30 s pulses eight times to lyse the cells. The homogenates were centrifuged at 20000g for 1 h at 4 °C, and the supernatant was heated for 10 min in boiling water and centrifuged for 20 min at 20000g. The supernatant was transferred into a 50 mL tube containing 3 mL of prewashed TALON His-Tag resin (Clontech) to adsorb protein for 1 h at 4 °C. The beads were packed into a column. The column was washed extensively with lysis buffer, and bound proteins eluted with buffer containing 150 mM imidazole hydrochloride. The eluted proteins were dialyzed against 10 mM Tris–HCl (pH 7.4), and a Complete EDTA-free protease inhibitor tablet was added to the dialyzed proteins that were stored at –80 °C.

**Purification of GST and GST-Tagged Proteins.** Pelleted bacteria were suspended in a buffer containing 100 mM Tris–HCl (pH 8.0), 200 mM NaCl, 10% glycerol, one Complete EDTA-free tablet per 30 mL of buffer, and 1.0 mg/mL lysozyme and frozen at –80 °C. The frozen pellet was thawed for 10 min in a 37 °C bath and sonicated for 30 s eight times for lysis of the cells. The homogenate was centrifuged at 20000g for 1 h at 4 °C, and the supernatant was transferred into a 50 mL tube containing 3 mL of prewashed Glutathione Sepharose 4B resin (GE Healthcare Life Sciences and Laboratories). After 1 h at 4 °C, the resin was packed into a column and washed with lysis buffer, and the bound proteins eluted using 10 mM reduced glutathione (GSH, Sigma Aldrich) in the same buffer. Eluted proteins were dialyzed against 10 mM Tris–HCl (pH 7.4), and a Complete

EDTA-free tablet was added to the dialyzed proteins that were stored at  $-80^{\circ}\text{C}$ .

**Site-Directed Mutagenesis.** Site-directed mutagenesis of HI-2 was performed using the Phusion Site Directed Mutagenesis Kit (New England Biolabs) according to the manufacturer's protocol. Mutagenic primers were designed using Primer X software available online. The mutations were all verified by automated DNA sequencing in the core facility of the University of Virginia School of Medicine.

**Protein-Protein Binding Assays.** The binding of purified I-2 to GST-Pin1 was performed using pull-down assays, quantified by fluorescence immunoblotting. A key way to reduce background was saturation of beads, accomplished by incubating  $10\ \mu\text{L}$  of prewashed glutathione Sepharose beads with excess recombinant GST-Pin1 in binding buffer containing 50 mM HEPES (pH 7.4), 100 mM NaCl, 10 mM  $\text{MgCl}_2$ , 0.5 mM  $\text{MnCl}_2$ , 5 mM EGTA, 1 mg/mL BSA, 0.5 mM DDT, and 1 mM Pefabloc (Roche Applied Sciences). The GST-Pin1 beads were pelleted by centrifugation and washed once with the same buffer. Various concentrations of recombinant I-2 proteins were added to the GST-Pin1 beads in a total volume of 1 mL with binding buffer and incubated for 1 h at room temperature. The beads were pelleted and washed three times by centrifugation with 1 mL of the same buffer, without BSA. Bound proteins were eluted with  $2\times$  SDS sample buffer and subjected to 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis SDS-PAGE and Western blotting as described previously.<sup>47</sup>

**Expression and Purification of Pin1 for NMR Spectroscopy Experiments.** The BL21(DE3) *E. coli* strain was transformed with the pET15b plasmid (Merck) carrying *pin1*. The recombinant strain was grown at  $37^{\circ}\text{C}$  in LB medium until  $A_{600}$  reached  $\sim 0.8$ , and then the induction phase was started by addition of 0.5 mM IPTG, for 3 h at  $31^{\circ}\text{C}$ . The cells were harvested by centrifugation, and the pellet was resuspended in lysis buffer [50 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$  (pH 8.0) and 300 mM NaCl (buffer A)], with 10 mM imidazole, 1 mM DTT, 0.1% NP40, and a protease inhibitor cocktail (Roche). Cell lysis was performed by incubation with 10 mg of lysozyme and 0.5 mg of DNase I per liter of culture, followed by a brief sonication, and an extract prepared by centrifugation. The soluble extract was loaded on a nickel affinity column (Chelating Sepharose Fast Flow 5 mL, GE Healthcare). Unbound proteins were washed away with 20 mM imidazole in buffer A, and the protein of interest was eluted with 250 mM imidazole in buffer A. The same protocol was performed for the production of  $^{15}\text{N}$ -labeled Pin1, with the use of M9 minimal medium complemented with 1 g of  $^{15}\text{NH}_4\text{Cl}$  (Cambridge Isotope Laboratories, Cambridge, MA) as a nitrogen source, 4 g/L glucose, 1 mM  $\text{MgSO}_4$ , 100 mg/L ampicillin, and MEM vitamin cocktail (Sigma). Homogeneous fractions were buffer-exchanged on a PD10 G-25 column (GE Healthcare) equilibrated in the NMR spectroscopy sample buffer [50 mM Tris- $d_{11}$  (pH 6.4), 25 mM NaCl, 1 mM DTT, and 1 mM EDTA].

**Expression and Purification of I-2 for NMR Spectroscopy Experiments.** The BL21(DE3) *E. coli* strain was transformed with the pET28a plasmid (Merck) encoding HI-2 and grown at  $37^{\circ}\text{C}$  in LB medium until  $A_{600}$  reached  $\sim 0.8$ , and then 0.5 mM IPTG was added for 16 h at  $20^{\circ}\text{C}$ . The cells were harvested by centrifugation, and the pellet was resuspended in lysis buffer [50 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$  (pH 8.0) and 300 mM NaCl (buffer A)], with added 10 mM imidazole, 1 mM DTT, 0.1% NP40, and a protease inhibitor cocktail (Roche). Cell lysis was

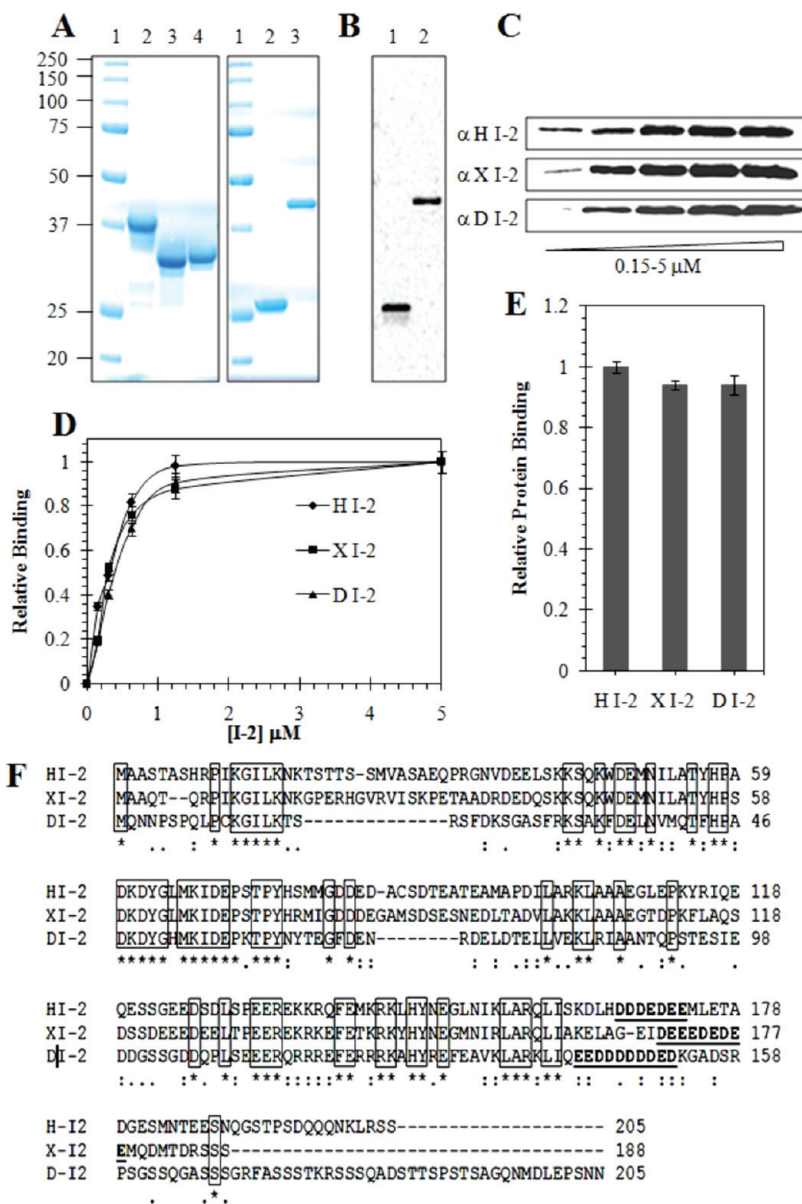
performed by incubation with 5 mg of lysozyme per liter of culture followed by sonication and centrifugation. The soluble extract was incubated at  $75^{\circ}\text{C}$  for 15 min and centrifuged at 15000g. The protein in the supernatant was purified on a 1 mL column of Chelating Sepharose Fast Flow (GE Healthcare) following the same procedure that was used for Pin1. The same protocol was performed for the production of  $^{15}\text{N}$ -labeled or  $^{15}\text{N}$ - and  $^{13}\text{C}$ -labeled I-2 with the use of M9 minimal medium complemented with 1 g of  $^{15}\text{NH}_4\text{Cl}$  (Cambridge Isotope Laboratories, Cambridge, MA) as the nitrogen source, 4 g/L glucose or 2 g of [ $^{13}\text{C}_6$ ]glucose, respectively, 1 mM  $\text{MgSO}_4$ , 20 mg/L kanamycin, and MEM vitamin cocktail (Sigma). The same  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum was obtained for I-2, regardless of whether the heating step of the soluble extract at  $75^{\circ}\text{C}$  was performed during the purification, indicating that heating did not modify the natively unfolded conformation of the I-2 protein. The eluted fractions containing most of the I-2 were pooled and purified by gel filtration (HiLoad 16/60 Superdex75 prep grade, GE Healthcare) in 50 mM ammonium bicarbonate. Fractions were checked on a 12% SDS-polyacrylamide gel, pooled, and lyophilized before being redissolved in NMR spectroscopy sample buffer [50 mM Tris- $d_{11}$  (pH 6.4), 25 mM NaCl, 1 mM DTT, and 1 mM EDTA] or directly in a Pin1 solution for titration.

**NMR Spectroscopy.** Interactions between I-2 and Pin1 were mapped either with  $^{15}\text{N}$ -labeled I-2 at 100  $\mu\text{M}$  and unlabeled Pin1 added to reach 0.5, 1, 2, or 12 molar equiv of Pin1 or  $^{15}\text{N}$ -labeled Pin1 at 62.5  $\mu\text{M}$  and unlabeled I-2 as lyophilized aliquots to reach 0.5 or 1 molar equiv of I-2. Spectra were recorded at 293 or 277 K on a 600 or 900 MHz spectrometer equipped with a cryogenic probehead. All  $^1\text{H}$  spectra were calibrated with 1 mM sodium 3-trimethylsilyl-3,3',2,2'- $d_4$ -propionate as a reference.  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded with at least 2048 points in the proton and 256 points in the nitrogen dimensions. The chemical shift perturbations of individual amide resonances in I-2 protein were calculated with eq 1, taking into account the relative dispersion of the proton and nitrogen chemical shifts (1 and 20 ppm, respectively). Instead, a correction coefficient of 0.2 was applied to the nitrogen dimension for the mapping of Pin1 resonances.

$$\Delta\delta(\text{ppm}) = \sqrt{[\Delta\delta(^1\text{H})]^2 + 0.05[\Delta\delta(^{15}\text{N})]^2} \quad (1)$$

**Superose 12 Size Exclusion Chromatography.** Chromatography on a Superose 12 HR 10/300 column was performed to study formation of the Pin1-I-2 complex. The GST tag was cleaved from Pin1 by on-column thrombin digestion using the previously described method.<sup>5</sup> Pin1 was concentrated on a centrifugal ultrafilter (10 kDa molecular mass cutoff, Millipore), mixed with a final HI-2 concentration of 1.25  $\mu\text{M}$ , and incubated for 1 h at room temperature. An aliquot (200  $\mu\text{L}$ ) was applied to the column and eluted with 50 mM HEPES (pH 7.4), 100 mM NaCl, 10 mM  $\text{MgCl}_2$ , 0.5 mM  $\text{MnCl}_2$ , 5 mM EGTA, 1 mg/mL BSA, 0.5 mM dithiothreitol, and 1 mM Pefabloc at a flow rate of 0.6 mL/min. Fractions of 1 min (0.6 mL) were collected and elution profiles developed by dot-blot immunoblotting. Samples (0.2 mL) of individual fractions were adsorbed on nitrocellulose using a 96-well vacuum manifold and developed as immunoblots for I-2 and Pin1. The staining intensities were quantified by fluorescence scanning and plotted as an elution profile.



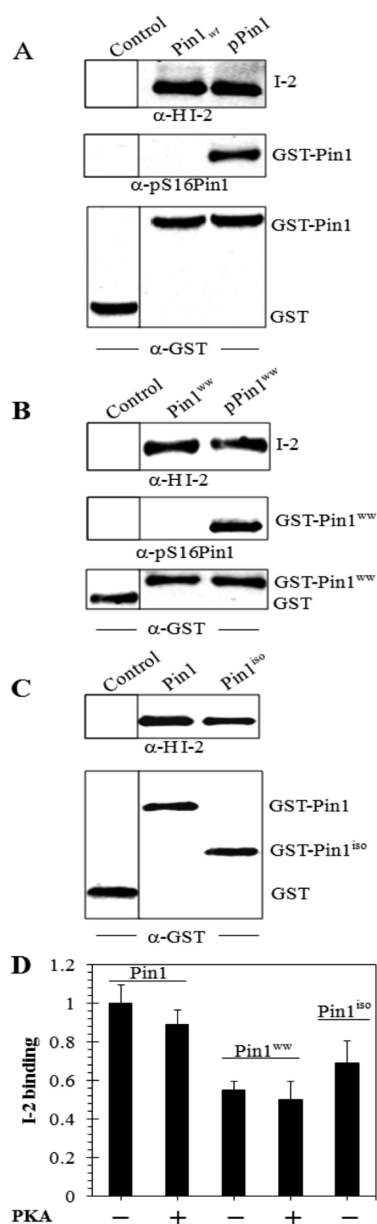


**Figure 1.** Binding of human, *Xenopus*, and *Drosophila* I-2 proteins to GST-Pin1. (A) Coomassie staining of purified proteins used in binding assays after SDS–PAGE. Left panel: lane 1, molecular size standard proteins; lane 2, human I-2 (HI-2); lane 3, *Xenopus* I-2 (XI-2); and lane 4, *Drosophila* I-2 (DI-2). Right panel: lane 1, molecular size standards; lane 2, GST control protein; and lane 3, GST-Pin1. (B) Immunoblotting of GST protein and GST-Pin1 using the anti-GST antibody. (C) Immunoblotting of different I-2 forms using species-specific antibodies for an assay of direct binding of recombinant His<sub>6</sub>-I-2 proteins to GST-Pin1 in a pull-down assay, as described in Materials and Methods. (D) Concentration dependence for I-2 from different species: (◆) HI-2, (■) XI-2, and (▲) DI-2. (E) Quantification of the amounts of different I-2 proteins bound to GST-Pin1, relative to recombinant protein standards. (F) Sequence alignment for human I-2 (HI-2), *Xenopus* I-2 (XI-2), and *Drosophila* I-2 (DI-2) using ClustalW 2.1. Conserved residues are marked with an asterisk and a box drawn around them. The polyacidic region is shown in boldface and underlined. HI-2 is 60% identical to XI-2 and 40% identical to DI-2.

**RESULTS**

**Binding of Inhibitor-2 Proteins with Pin1.** To define the structural requirements for interaction, we compared binding of I-2 proteins from different species [human (HI-2), *Xenopus* (XI-2), and *Drosophila* (DI-2)] to human GST-Pin1 in a pull-down assay. The recombinant proteins were expressed in bacteria and extensively purified by affinity chromatography, and the final products were analyzed by Coomassie staining following SDS–PAGE (Figure 1A). GST and GST-Pin1 were detected

as single bands by anti-GST immunoblotting (Figure 1B). The binding assay used a range of concentrations of added I-2 proteins (0.1–5  $\mu$ M) mixed with GST-Pin1 that was bound to glutathione Sepharose beads. The bound proteins were quantified by fluorescence immunoblotting with species-specific anti-I-2 antibodies. Because the I-2 sequences are so divergent, no single anti-I-2 antibody reacts with these three proteins; therefore, we prepared and affinity purified species-specific anti-HI-2, anti-XI-2, and anti-DI-2 antibodies for immunoblotting detection (Figure 1C). Staining intensities on immunoblots were



**Figure 2.** Comparison of I-2 binding to Pin1, phospho-Pin1, and separate WW and isomerase domains. GST-Pin and GST-Pin<sub>ww</sub> and their PKA-phosphorylated forms, and GST-Pin1<sub>iso</sub>, were prepared as described in Materials and Methods. (A) Glutathione Sepharose beads were saturated with GST (Control), wild-type GST-Pin1 (Pin1<sub>wt</sub>), or phosphorylated GST-Pin1 (pPin1) and subjected to a pull-down assay with human I-2 protein (1.25 μM) as described in Materials and Methods. Proteins were eluted and immunoblotted with anti-human I-2 (top), anti-phospho-S16Pin1 (middle), and anti-GST (bottom) antibodies. (B) GST (Control), GST-Pin1<sub>ww</sub> (Pin1<sup>ww</sup>), and PKA-phosphorylated GST-Pin1<sub>ww</sub> (pPin1<sup>ww</sup>) were incubated for 1 h in with HI-2 protein and processed as described for panel A. (C) Direct binding of HI-2 to GST (Control, left lane), GST-Pin1 (center lane), and the GST-Pin1 isomerase domain (right lane). The top panel shows immunoblots for bound HI-2 and the bottom panel immunoblots for GST as a loading control. (D) Quantification of pull-down assays depicted in panels A–C, showing average values and the standard deviation for three independent experiments.

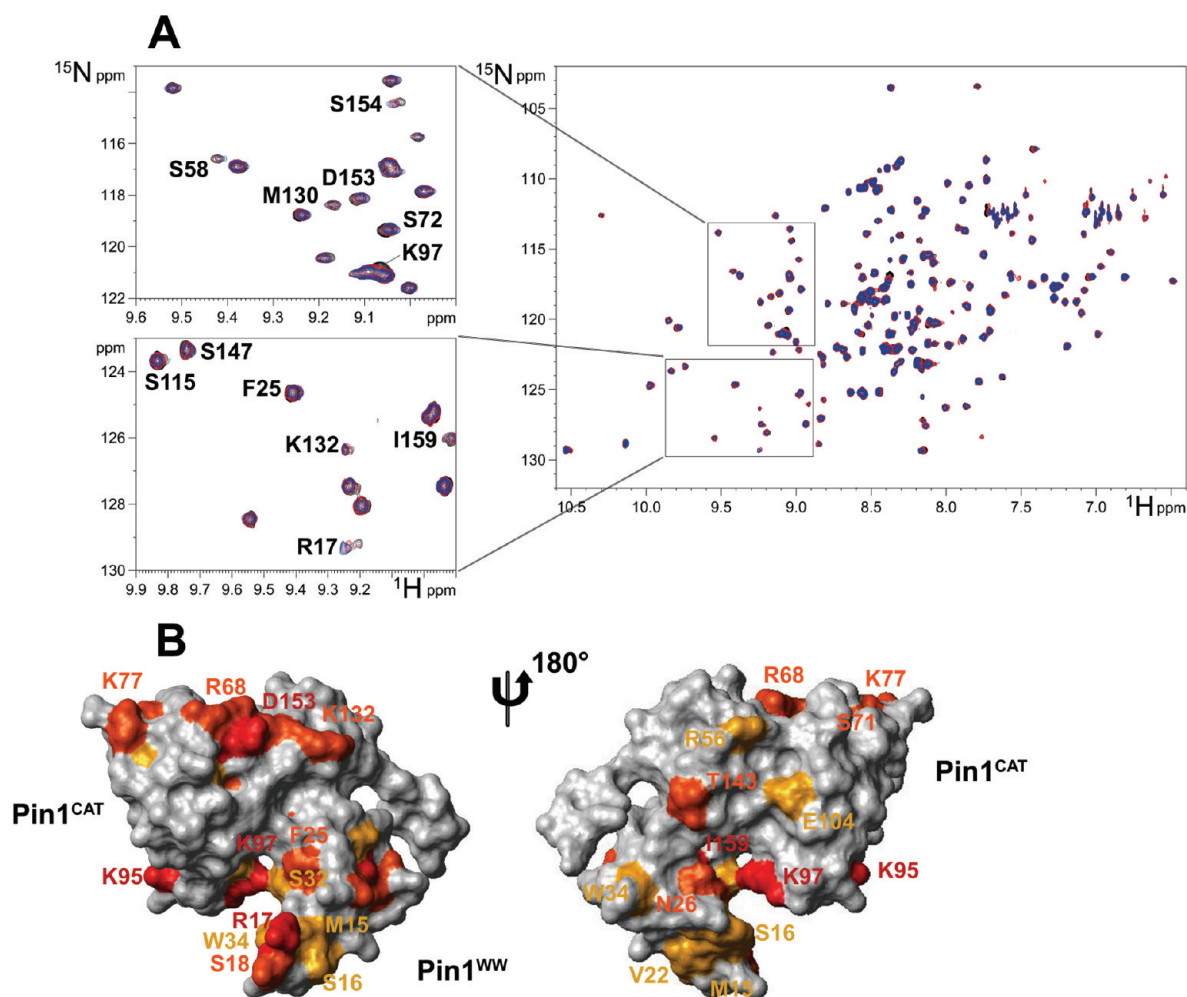
quantified and calibrated over a range of concentrations, using recombinant I-2 standards with the absolute amount of the

proteins determined by A<sub>280</sub>. The binding of HI-2, XI-2, and DI-2 to GST-Pin1 was indistinguishable, with a K<sub>0.5</sub> of ~0.3 μM, and all showed saturation at 1.25–5 μM (Figure 1D), which is well below the estimated cellular concentration of I-2.<sup>43</sup> The total amount of each I-2 protein bound to GST-Pin1 at saturation was identical (Figure 1E). GST was used as the negative control, and the assay was optimized so there was essentially no I-2 binding to GST beads. This showed that the I-2 proteins were binding to the Pin1 portion of the fusion protein, not to GST itself. These results demonstrated that stable, saturable association with Pin1 is a conserved function of the I-2 family of proteins. Multiple-sequence alignment generated for these I-2 proteins revealed the location and clustering of identical residues (boxed in Figure 1F) that are the most probable regions for interaction with human Pin1.

**Association of HI-2 with Phospho-Pin1 and Separate Pin1 Domains.** Binding of HI-2 to Pin1 used proteins expressed in bacteria; therefore, they were both devoid of phosphorylations that occur in live animal cells, such as phosphorylation of Ser16 in the Pin1 WW domain.<sup>28</sup> We reacted full-length GST-Pin1 with purified PKA and showed phosphorylation of Ser16 by immunoblotting with a phospho site-specific antibody (Figure 2A). Binding of HI-2 to the PKA-phosphorylated GST-Pin1 was weakened only slightly compared to the binding of HI-2 with nonphosphorylated GST-Pin1 (Figure 2A,D). These results show that phosphorylation of Pin1 by PKA, involving at least Ser16, did not affect the formation of a protein–protein complex with HI-2.

We expressed the N-terminal WW domain of Pin1 (Pin1<sub>ww</sub>) and C-terminal isomerase catalytic domain (Pin1<sub>iso</sub>) fused to GST and used them in pull-down assays with purified HI-2. The amount of HI-2 bound to Pin1<sub>ww</sub> was approximately half of that bound to Pin1 (Figure 2B,D). Phosphorylation of Pin1<sub>ww</sub> by PKA was demonstrated by immunoblotting with the pSer16 phospho site antibody (Figure 2B). Phosphorylation of the Pin1<sub>ww</sub> domain resulted in no difference in binding of HI-2 in the pull-down assay, reinforcing the results with full-length Pin1. These results showed that the WW domain of Pin1 alone was sufficient for interaction with HI-2, albeit at approximately half the level relative to that of full-length Pin1. Assays with the isomerase domain of Pin1 (Pin1<sub>iso</sub>) revealed that the amount of HI-2 bound was ~30% lower than the amount of full-length Pin1 (Figure 2C,D), indicating that the isomerase domain also was sufficient for association with I-2. The results indicate that I-2 independently associates with both the WW domain and the isomerase domain of Pin1.

**NMR Spectroscopy Analysis of Association of I-2 with Pin1.** Pin1 and I-2 proteins were separately labeled with <sup>15</sup>N by expression in bacteria and purified, and their interaction in solution was investigated by NMR spectroscopy. The two-dimensional <sup>15</sup>N–<sup>1</sup>H heteronuclear single-quantum coherence (HSQC) spectra of 100 μM <sup>15</sup>N-labeled Pin1 alone and mixed with purified unlabeled HI-2 were recorded and superimposed. The spectrum of Pin1 alone (black) and that after the addition of 0.5 (red) and 1.0 (blue) molar equiv of HI-2 showed weak but reproducible shifts of selected resonances in Pin1 (Figure 3A), indicative of physical contacts with HI-2. To delineate the HI-2 binding sites on Pin1, the residues with induced chemical shift perturbations were mapped on the Pin1 3D structure from the Protein Data Bank (PDB) (entry 1PIN), depicted front and back in Figure 3B. Cross-peak movements that correspond to residues with combined chemical shift variations



**Figure 3.** NMR spectroscopy mapping of surface residues in Pin1 that contact I-2. (A)  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -labeled Pin1 alone (black) overlaid with spectra after addition of 0.5 (red) and 1.0 molar equiv (blue) of unlabeled I-2 showing the shifts of resonances, marked with single-letter code and residue number. The magnified areas are enlarged to the left side for closer examination. (B) Mapping of the I-2 interaction sites on the Pin1 structure (PDB entry 1PIN). The isomerase domain (Pin1<sup>CAT</sup>) is shown at the top left and the WW domain (Pin1<sup>WW</sup>) at the bottom right. The structure on the right is rotated 180° to show the reverse side. Residues with combined chemical shift variations (calculated as described in Materials and Methods) greater than 0.04 ppm are colored red, those in the 0.02–0.04 ppm range orange, and those in the 0.015–0.02 ppm range yellow.

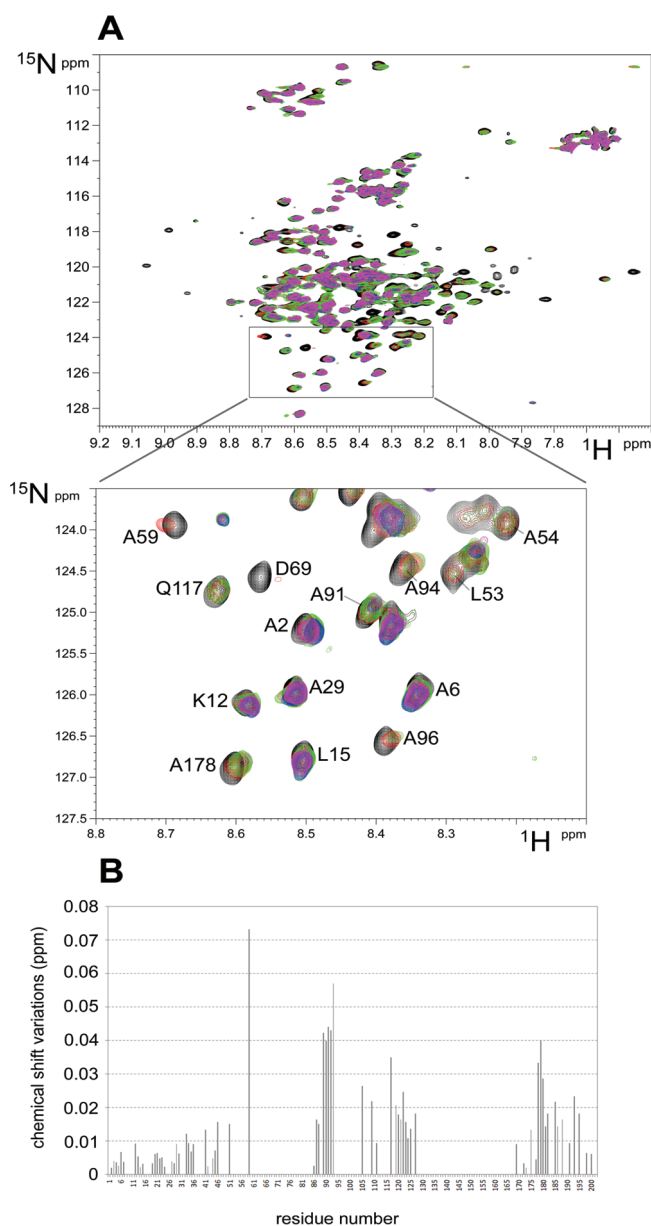
greater than 0.04 ppm (red), in the 0.02–0.04 ppm range (orange), and in the 0.015–0.02 ppm range (yellow) mapped residues in contiguous areas on the surfaces of both the WW and catalytic domains of Pin1. The structure is rotated 180° to view the opposite side, where relatively fewer interactions appeared. These images support our results in binding assays in which I-2 interacts with the separate, isolated Pin1<sub>WW</sub> and Pin1<sub>iso</sub> domains. The contacts of I-2 with the WW domain involved R17, S18, and S32, and several hydrophobic side chains, including V22, F25, and W34. On the other hand, interactions with the isomerase catalytic domain were dominated by charged side chains, K77, R68, D153, and K132. The I-2 contact surface on the Pin1 catalytic domain includes the putative active site residue R68 but does not appear to occlude the active site to interfere with substrate binding. This is consistent with our previous report that I-2 allosterically modifies Pin1 specificity.

Reciprocal labeling experiments involved the titration of purified  $^{15}\text{N}$ -labeled HI-2 with unlabeled Pin1 and were conducted under different conditions (both 4 and 20 °C) and different molar ratios of the proteins (from 0.5 to 12). The

HSQC spectra of HI-2 (Figure 4A) in the presence of increasing amounts of Pin1 were superimposed over the spectrum of HI-2 alone (black) to show significant perturbations that correspond to residues in HI-2 in physical contact with Pin1. HI-2 is known to be an extended polypeptide in solution and other than one short helical region is devoid of secondary or tertiary structure, so the results could not be mapped on a three-dimensional structure as for Pin1 (see above); instead, we plotted shifts as a function of sequence position in HI-2 (Figure 4B). This analysis reveals multiple regions in I-2 appear to be in contact with Pin1. These include the N- and C-termini, residues 85–95, and residues 105–125. The region of residues 60–85 was so broadened by addition of excess Pin1 the delta could not be calculated. These relatively small variations could be reproduced under different conditions, and our interpretation was that the extreme flexibility of I-2 did not allow for larger changes. These NMR spectroscopy results with labeled I-2 show binding with Pin1 involved multiple sites of protein–protein interaction.

**Defining Regions in HI-2 Involved in Binding to Pin1.** To further dissect the structure–function relationship for I-2, we





**Figure 4.** NMR spectroscopy mapping of I-2 residues that contact Pin1. (A) Complete (top) and magnified (bottom)  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -labeled I-2 alone (black) overlaid with spectra recorded after addition of 0.5 (red), 1.0 (green), 2.0 (blue), and 12 (magenta) molar equiv of unlabeled Pin1, recorded at 4 °C showing the shifts of specific residues. (B) Graphical representation of the chemical shift variations in the HSQC spectrum of I-2 resonances ( $\Delta\delta$ ) upon addition of excess Pin1, as done for panel A.

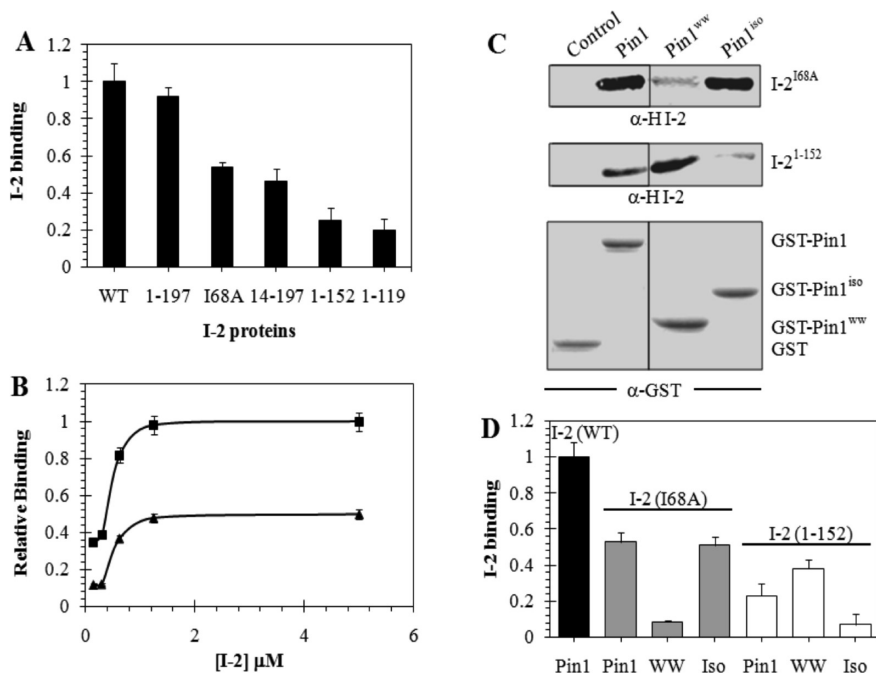
produced truncated and mutated forms of HI-2 and assayed binding to GST-Pin1 (Figure 5). The following HI-2 variants were used: full-length wild type I-2 (WT), C-terminally truncated I-2 (1–197), substitution of a conserved aliphatic side chain, I-2 (I68A), deletion of both N- and C-terminal regions (14–197), and larger deletions of the C-terminus, (1–152 and 1–119). Recombinant I-2 proteins were affinity purified and assayed at a final concentration of 1.25  $\mu\text{M}$ . The results (Figure 5A) showed no significant difference between the binding of I-2(205) and I-2(197), showing that the extreme C-terminal region was not required for association of HI-2 with

GST-Pin1. In contrast, truncation of the 13 N-terminal residues that includes the SILK motif common to many PP1 binding proteins reduced by half the amount of binding of I-2 to Pin1. Further truncation of I-2 from the C-terminus to residue 152, or residue 119, resulted in a significant 75% loss of binding to GST-Pin1. These results reinforced the NMR spectroscopy assignment of a site in I-2 encompassing residues 170–200 as strongly interacting with Pin1. We suspect involvement of an uninterrupted stretch of acidic residues, which is a conserved feature among I-2 proteins that does not align exactly in the same residue positions. It is important to note here that 1–152 was tested because its inhibitory potency for PP1C is identical to that of full-length I-2.<sup>48</sup>

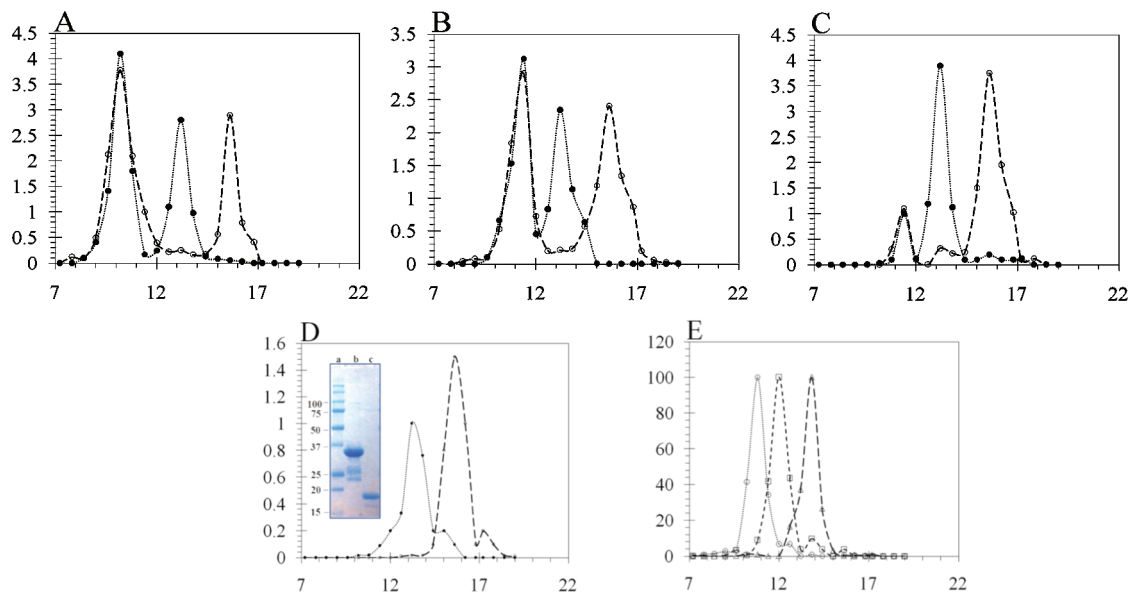
We had reported association of I-2 with Pin1 was exceedingly sensitive to disruption by detergents,<sup>39</sup> presumably because of hydrophobic interaction between the proteins. The highly conserved region of residues 60–85 of I-2 is likely engaged with Pin1 on the basis of results of NMR spectroscopy. Within this region, most of the conserved residues are charged, with the exception of Ile68, which is a strictly conserved, aliphatic residue that we suspected could be involved in a hydrophobic protein–protein interface. Therefore, we produced an I68A substituted form of I-2 and found its level of binding with Pin1 was reduced to half of that of wild-type HI-2 when both proteins were compared at 1.25  $\mu\text{M}$  (Figure 5A). We assayed for binding to Pin1 over a range of concentrations and found a reduced slope in the initial portion (<1  $\mu\text{M}$ ) of the saturation curve (Figure 5B). However, more provocative was the observation that when binding of I-2(I68A) reached saturation at approximately the same concentration as wild-type I-2, there was only 50% of the amount of I-2 protein bound. This suggested to us that there were two binding sites for wild-type I-2 binding to Pin1, and I-2(I68A) bound to only one of them.

We compared binding of I-2, I-2(I68A), and I-2(152) to full-length Pin1, Pin1<sub>ww</sub>, and Pin1<sub>iso</sub> in pull-down assays (Figure 5C, D). Mutation I68A resulted in essentially a complete loss of binding to the WW domain, whereas the binding to the Pin1 isomerase domain was the same as binding to full-length Pin1. Our conclusion was that I68 was critical for binding to the WW domain but was not required for and did not much affect binding of I-2 to the isomerase domain of Pin1. On the other hand, truncation of I-2 to 152 residues essentially eliminated binding to the isomerase domain, while I-2(152) bound to full-length Pin1 and Pin1<sub>ww</sub>. Our hypothesis was that each domain of Pin1 bound separate regions of I-2, and Pin1 bound to two I-2 molecules, one to each domain.

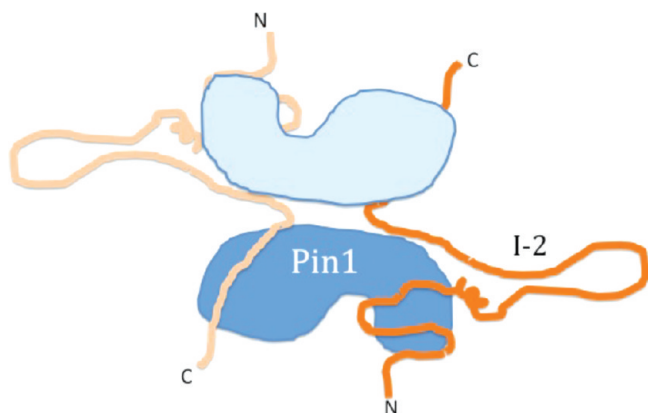
**Analysis of HI-2–Pin1 Complexes by Size Exclusion Chromatography.** We tested our hypothesis using Superose 12 column chromatography to analyze the complexes formed between His<sub>6</sub>-HI-2 and Pin1 (Figure 6). This was a homogeneous solution assay that involved simple mixing of soluble components, compared to the heterogeneous assay that used a GST fusion protein immobilized on beads. Pin1 protein was recovered free of GST by on-column thrombin cleavage of the GST-Pin1 fusion protein purified from bacteria (see the inset of Figure 6D). This preparation of Pin1 eluted as a symmetrical peak centered at 15.6 mL (dashed line in Figure 6D), corresponding to monomeric Pin1 (Figure 6A–D). Purified I-2 eluted from Superose 12 as a peak centered at 13.2 mL (dotted line in Figure 6A–D). Because of its extended and unstructured conformation, I-2 has been known to behave as a larger molecule in solution, and indeed, I-2 eluted before ovalbumin [43 kDa,



**Figure 5.** Binding of I-2 variants to GST-Pin1. (A) Binding of wild-type (WT) and truncated and substituted HI-2 to GST-Pin1 in a pull-down assay as described in Materials and Methods. The recovery of bound I-2 was quantitated using Odyssey version 1.2 by immunoblotting, with a correction for amount of GST-Pin1 by simultaneous dual-wavelength scanning of two different secondary antibodies on the same filter (Odyssey, Li-Cor Industries). Values from three independent experiments were normalized to WT, averaged, and plotted with the standard deviation. (B) Dose-response curve for binding of wild-type HI-2 ( $\blacksquare$ ) and I-2(I68A) ( $\blacktriangle$ ) to GST-Pin1 in a pull-down assay ( $n = 3$ ). Immunoblotting analyzed as described for panel A and data plotted with the standard deviation using Sigma plot 10.0. (C and D) Pull-down assays for binding of 1.25  $\mu$ M HI-2(I68A) (I-2<sup>I68A</sup>) (top gel in panel C, gray bars in panel D) or HI-2(1–152) (I-2<sup>1–152</sup>) (center gel in panels C, white bars in panel D) with GST, GST-Pin1 (Pin1), GST-Pin1<sup>ww</sup> (Pin1<sup>ww</sup> and WW), or GST-Pin1<sup>iso</sup> (Pin1<sup>iso</sup> and Iso). Immunoblotting of GST and fusion proteins is shown in the bottom gel of panel C. For control and normalization of other samples, wild type (WT) HI-2 was pulled down with GST-Pin1. Data ( $n = 3$ ) were analyzed as described for panel A.



**Figure 6.** Size exclusion chromatography analysis of Pin1 and I-2 complexes. Recombinant GST-Pin1 was cleaved by thrombin, and the Pin1 protein was concentrated, incubated with recombinant HI-2, and applied to a Superose 12 HR 10/300 column, as described in Materials and Methods. Aliquots of the eluted fractions were immunoblotted with anti-Pin1 (dashed lines) and anti-HI-2 (dotted lines) antibodies using the dot blot technique with a 96-well vacuum manifold. (A) Elution profile of a mixture of Pin1 and wild-type HI-2. (B) Elution profile of a mixture of Pin1 and the I68A mutant form of HI-2. (C) Elution profile of a mixture of Pin1 and the truncated form of HI-2(1–152). (D) Superimposition of separate elution profiles for Pin1 and wild-type HI-2. The inset is a Coomassie-stained SDS-PAGE gel: lane 1, molecular size standards; lane 2, purified Pin1; and lane 3, purified HI-2 protein used in these experiments. (E) Superimposition of separate elution profiles for purified proteins used as molecular size standards: catalase (260 kDa, 10.8 mL), immunoglobulin (150 kDa, 12.0 mL), and ovalbumin (43 kDa, 13.8 mL).



**Figure 7.** Model of the Pin1–I-2 complex. Cartoon showing the arrangement of two molecules of Pin1 with two molecules of I-2 in a complex formed in solution. Pin1 is colored light and dark blue as two lobes, composed of a smaller WW domain and a larger isomerase catalytic domain. I-2 is colored orange as an unstructured polypeptide with N- and C-termini labeled, using multiple regions for binding to separate domains of Pin1.

13.8 mL (Figure 6E)] in size exclusion chromatography. When recombinant I-2 was mixed with Pin1 and the sample resolved on Superose 12, immunoblotting of the fractions for I-2 revealed two prominent peaks: a major symmetrical peak centered at 10.2 mL that contained ~60% of the total I-2 protein and a second peak at 11.4 mL corresponding to unbound monomeric I-2 (Figure 6A). Immunoblotting of the fractions for Pin1 showed exact co-elution of Pin1 with I-2 in the peak at 10.2 mL and also a second peak at 15.6 mL, corresponding to free monomeric Pin1. These results showed the formation of a stable, large complex between purified Pin1 and I-2 that eluted at a size larger than that of catalase [260 kDa, 10.8 mL (Figure 6E)]. On the other hand, when the I68A mutant of I-2 was mixed with Pin1, the complex that formed eluted at 11.4 mL (Figure 6B). This complex is smaller than catalase but still larger than immunoglobulin [150 kDa, 12.0 mL (Figure 6E)] and corresponds to approximately half the relative size of the wild-type I-2–Pin1 complex. The distribution of I-2 protein between the Pin1 complexes and the free monomer was approximately the same for I-2(I68A) and wild-type I-2 (Figure 6A,B), and our interpretation was that these different I-2 proteins had approximately the same affinity for Pin1. Analysis of I-2(152) binding to Pin1 showed a very low yield of the complex between the proteins in solution, with most of the proteins eluting separately as monomers (Figure 6C). What little complex was formed eluted at 11.4 mL, not 10.2 mL. The excluded volume ( $V_0$ ) of this column is 9 mL; thus, the complexes formed were discrete molecular species and not simply polymers. Because the complexes eluting at 10.2 and 11.4 mL both contained equivalent amounts of I-2 and Pin1, we concluded that these peaks represented tetrameric  $(\text{Pin1})_2(\text{I-2})_2$  and dimeric  $(\text{Pin1})(\text{I-2})$  complexes, respectively.

## DISCUSSION

The results of this study reveal molecular interactions between two ancient and essential proteins, Pin1 and I-2. Purified recombinant I-2 proteins from divergent species exhibited saturable binding to Pin1 with submicromolar affinity. Estimates for the intracellular concentrations of Pin1 and I-2

are micromolar,<sup>43,49</sup> which would predict that the proteins form complexes together in living cells and tissues. Our results argue that binding to Pin1 is an evolutionarily conserved function, and our hypothesis was that interactions involved regions of the I-2 protein identical in sequence among the different species. Indeed, NMR spectroscopy of reciprocally labeled proteins dynamically interacting in solution showed that multiple conserved sequence regions in I-2 were in contact with surface sites on Pin1, in both the WW and isomerase domains. One key conclusion from the results of NMR spectroscopy is that the I-2 interactions mapped on the surfaces of Pin1 do not occlude the active site. This supports the concept that I-2 does not interfere with substrate binding but acts as an allosteric modulator of Pin1 substrate specificity, as we previously demonstrated using a panel of mitotic phosphoproteins.<sup>39</sup>

Mapping the sites of I-2 interaction on Pin1 by NMR spectroscopy revealed surface regions on the WW domain and on the isomerase domain. Residues in the WW domain that contacted I-2 did not include Ser16, consistent with unimpeded binding of I-2 to PKA-phosphorylated Pin1. Phosphorylation by PKA involves at least Ser16, and this phosphoryl group is thought to preclude the binding of phosphorylated substrates to the WW domain.<sup>28</sup> Binding of I-2 was insensitive to PKA phosphorylation of either the full-length Pin1 protein or the WW domain, so the phosphoryl group at Ser16 did not exert steric or electrostatic constraints upon association with I-2. These results are important in devising a model for the complex between Pin1 and I-2 and potentially will influence our understanding of how signaling events regulate Pin1 function. The NMR spectroscopy analysis also identified several hydrophobic residues in the Pin1 WW domain that made contact with I-2. Pin1 was already known to elaborate a hydrophobic surface on the WW domain that faces the isomerase domain. This forms a hydrophobic cleft that binds polyethylene glycol during crystallization, as visualized in the 3D structure determined by X-ray crystallography.<sup>6</sup> There are relatively few conserved hydrophobic residues in I-2 that would be available for generating an aliphatic protein–protein interface with Pin1. Searching among the conserved sequence regions in I-2 that NMR spectroscopy identified as Pin1 contact sites, we selected Ile68 as a likely candidate. The I68A substitution was a relatively minor modification in the entire 205-residue protein but had profound effects on binding to Pin1. The level of binding was reduced ~50%, but more importantly, the amount of I-2(I68A) bound at saturating concentrations was only half as much as the amount of wild type. This was unexpected and forced us to consider the possibility that there were two binding sites for I-2 on Pin1, only one of which bound I-2(I68A).

This idea was validated by an assay of I-2 binding to Pin1 and the separate WW and isomerase domains. The I68A mutation essentially eliminated binding of I-2 to the isolated WW domain but did not equally impair binding to the Pin1 isomerase domain. On this basis, we assign the I68 and surrounding region of I-2 as the part of the protein engaged with the WW domain of Pin1. On the other hand, truncation of I-2 to residue 152 eliminates the Pin1 interacting site at residue 170–190 seen by NMR spectroscopy, which we suspect involves a long uninterrupted stretch of acidic residues that appears as a conserved feature among I-2 forms from different species (boldface and underlined in Figure 1). I-2(1–152) bound more weakly to Pin1; however, the level of binding to Pin1 and its isolated WW domain was approximately the same, whereas there was almost no binding to the isomerase domain. Our interpretation of these



results is that the polyacidic site in I-2 probably interacts with the multiple adjacent basic residues across the surface of the isomerase domain in Pin1. Assays with mutated I-2 and Pin1 separate domains were consistent with the results of mapping with NMR spectroscopy and allowed us to assign interactions of different I-2 regions to individual Pin1 domains.

The results led us to consider the idea that if two molecules of I-2 were bound to a single Pin1, at different sites, in the two domains, then it is possible that a second molecule of Pin1 could contact these same two I-2 proteins to form a sandwich, with the Pin1 proteins positioned opposite one another bridged by two molecules of I-2 that are predominantly unstructured and extended (Figure 7). The  $[\text{Pin1}]_2[\text{I-2}]_2$  heterotetramer is proposed as the native active form of Pin1. Bringing two molecules of Pin1 together positions two phosphoSer/Thr binding sites in the proximity, and the involvement of I-2 in the complex generates a composite protein surface surrounding the Pin1 isomerase domain, which is predicted to dictate the substrate specificity. This tetramer represents a quite different target for drug development compared to monomeric Pin1.

Furthermore, our model accommodates phosphorylation as a means of Pin1 regulation. While PKA phosphorylation of S16 in Pin1 does not affect association with I-2, this phosphorylation of two sites on the surface of the  $[\text{Pin1}]_2[\text{I-2}]_2$  complex could serve to alter binding to different substrates. Phosphorylation of Thr73 in I-2, which is catalyzed by CDK/cyclinB during mitosis,<sup>36</sup> prevents I-2 from binding to Pin1.<sup>39</sup> Therefore, we predict upon entry into mitosis phosphorylation of I-2 in the  $[\text{Pin1}]_2[\text{I-2}]_2$  complex will cause dissociation into separate components, releasing monomeric Pin1 to interact with mitotic phosphoproteins. Mitotic exit will involve dephosphorylation of I-2 and reassembly with Pin1 into the heterotetramer. This model implies dynamic regulation of Pin1 complexes that would affect substrate specificity at different stages of mitosis.

We note that I-2 residues 85–95 and 105–125 are engaged by Pin1. For I-2 bound to PP1C, the segment of residues 95–105 is looped out and does not form contacts with PP1C, and NMR spectroscopy shows some tendency for this region to be in an  $\alpha$  helix.<sup>50</sup> In a recent publication, the authors speculated that this putative helix could possibly contact another I-2 partner, such as Aurora A or Pin1.<sup>51</sup> This would fit with our results. The region of residues 105–125 we see affected by Pin1 does not overlap with the region of residues 130–142 of I-2 that is 70% populated in a helix conformation, as observed by NMR spectroscopy.<sup>50</sup> The segment of residues 130–142 of I-2 was visualized as a helix bound across the PP1C active site in the cocrystal structure, with only 25% of I-2 showing discernible electron density.<sup>52</sup> Even when in complex with its primary partner, PP1C, most of I-2 is thought to remain as flexible as it is when it is alone in solution. Overall, we are impressed that our NMR spectroscopy results and those of Peti and collaborators map different I-2 residues interacting with Pin1 compared to those that contact the PP1C monomer, or PP1C complexed to neurabin.<sup>51</sup> Although different regions of I-2 are used in these various complexes, we do not imagine this implies assembly of supercomplexes with multiple partners, but rather assembly of I-2 into multiple complexes with different partners. Our previous research has revealed that, in addition to PP1C, I-2 forms complexes with neurabin,<sup>53</sup> KPI-2<sup>54</sup> (also known as LMKT2 and BREK), Aurora A,<sup>44</sup> and, of course, Pin1.<sup>39</sup> This emphasizes the possibilities available to an unstructured protein to employ various sequence segments to engage different protein partners. The ancient and conserved function of

I-2 involves a web of protein–protein interactions, and we speculate that binding to Pin1 may be as ancient, and fundamentally important, as binding to PP1, the property for which I-2 was originally named.

## AUTHOR INFORMATION

### Corresponding Author

\*Phone: (434) 924-5892. Fax: (434) 924-1236. E-mail: db8g@virginia.edu.

### Funding Sources

This work was supported by National Institutes of Health Grant GM56362 (to D.L.B.) and by CNRS (to G.L.).

## ACKNOWLEDGMENT

We thank Drs. Jean-Michel Wieruszkeski and Bernd Fritzinger for their support in recording the NMR spectra. We dedicate the article to the memory of Professor Emanuel Margoliash, a pioneer in the use of homologous proteins from various species to provide insights into structure and function. The NMR spectroscopy facility was supported by the TGE RMN THC (FR-3050, France).

## ABBREVIATIONS

GST, glutathione S-transferase; HI-2, human phosphatase inhibitor-2; XI-2, *Xenopus* phosphatase inhibitor-2; DI-2, *Drosophila* phosphatase inhibitor-2; Pin1<sub>ww</sub>, human Pin1 WW domain (residues 1–44); Pin1<sub>iso</sub> or Pin1<sub>cat</sub>, Pin1 isomerase domain (residues 46–163); HSQC, heteronuclear single-quantum coherence.

## REFERENCES

- (1) Shen, M., Stukenberg, P. T., Kirschner, M. W., and Lu, K. P. (1998) The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins. *Genes Dev.* 12, 706–720.
- (2) Schutkowski, M., Bernhardt, A., Zhou, X. Z., Shen, M., Reimer, U., Rahfeld, J. U., Lu, K. P., and Fischer, G. (1998) Role of phosphorylation in determining the backbone dynamics of the serine/threonine-proline motif and Pin1 substrate recognition. *Biochemistry* 37, 5566–5575.
- (3) Lu, K. P., Hanes, S. D., and Hunter, T. (1996) A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* 380, 544–547.
- (4) Hanes, S. D., Shank, P. R., and Bostian, K. A. (1989) Sequence and mutational analysis of ESS1, a gene essential for growth in *Saccharomyces cerevisiae*. *Yeast* 5, 55–72.
- (5) Maleszka, R., Lupas, A., Hanes, S. D., and Miklos, G. L. (1997) The dodo gene family encodes a novel protein involved in signal transduction and protein folding. *Gene* 203, 89–93.
- (6) Ranganathan, R., Lu, K. P., Hunter, T., and Noel, J. P. (1997) Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* 89, 875–886.
- (7) Bayer, E., Goettsch, S., Mueller, J. W., Griewel, B., Guiberman, E., Mayr, L. M., and Bayer, P. (2003) Structural analysis of the mitotic regulator hPin1 in solution: Insights into domain architecture and substrate binding. *J. Biol. Chem.* 278, 26183–26193.
- (8) Li, H., Wang, S., Zhu, T., Zhou, J., Xu, Q., Lu, Y., and Ma, D. (2006) Pin1 contributes to cervical tumorigenesis by regulating cyclin D1 expression. *Oncol. Rep.* 16, 491–496.
- (9) Wulf, G. M., Ryo, A., Wulf, G. G., Lee, S. W., Niu, T., Petkova, V., and Lu, K. P. (2001) Pin1 is overexpressed in breast cancer and cooperates with Ras signaling in increasing the transcriptional activity of c-Jun towards cyclin D1. *EMBO J.* 20, 3459–3472.

- (10) Yeh, E., Cunningham, M., Arnold, H., Chasse, D., Monteith, T., Ivaldi, G., Hahn, W. C., Stukenberg, P. T., Shenolikar, S., Uchida, T., Counter, C. M., Nevins, J. R., Means, A. R., and Sears, R. (2004) A signalling pathway controlling c-Myc degradation that impacts oncogenic transformation of human cells. *Nat. Cell Biol.* 6, 308–318.
- (11) Wulf, G. M., Liou, Y. C., Ryo, A., Lee, S. W., and Lu, K. P. (2002) Role of Pin1 in the regulation of p53 stability and p21 transactivation, and cell cycle checkpoints in response to DNA damage. *J. Biol. Chem.* 277, 47976–47979.
- (12) Zheng, H., You, H., Zhou, X. Z., Murray, S. A., Uchida, T., Wulf, G., Gu, L., Tang, X., Lu, K. P., and Xiao, Z. X. (2002) The prolyl isomerase Pin1 is a regulator of p53 in genotoxic response. *Nature* 419, 849–853.
- (13) Hamdane, M., Dourlen, P., Bretteville, A., Sambo, A. V., Ferreira, S., Ando, K., Kerdraon, O., Begard, S., Geay, L., Lippens, G., Sergeant, N., Delacourte, A., Maurage, C. A., Galas, M. C., and Buee, L. (2006) Pin1 allows for differential Tau dephosphorylation in neuronal cells. *Mol. Cell. Neurosci.* 32, 155–160.
- (14) Lim, J., and Lu, K. P. (2005) Pinning down phosphorylated tau and tauopathies. *Biochim. Biophys. Acta* 1739, 311–322.
- (15) Zhou, X. Z., Kops, O., Werner, A., Lu, P. J., Shen, M., Stoller, G., Kullertz, G., Stark, M., Fischer, G., and Lu, K. P. (2000) Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins. *Mol. Cell* 6, 873–883.
- (16) Ayala, G., Wang, D., Wulf, G., Frolov, A., Li, R., Sowadski, J., Wheeler, T. M., Lu, K. P., and Bao, L. (2003) The prolyl isomerase Pin1 is a novel prognostic marker in human prostate cancer. *Cancer Res.* 63, 6244–6251.
- (17) Wulf, G., Ryo, A., Liou, Y. C., and Lu, K. P. (2003) The prolyl isomerase Pin1 in breast development and cancer. *Breast Cancer Res.* 5, 76–82.
- (18) Bao, L., Kimzey, A., Sauter, G., Sowadski, J. M., Lu, K. P., and Wang, D. G. (2004) Prevalent overexpression of prolyl isomerase Pin1 in human cancers. *Am. J. Pathol.* 164, 1727–1737.
- (19) Kim, C. J., Cho, Y. G., Park, Y. G., Nam, S. W., Kim, S. Y., Lee, S. H., Yoo, N. J., Lee, J. Y., and Park, W. S. (2005) Pin1 overexpression in colorectal cancer and its correlation with aberrant  $\beta$ -catenin expression. *World J. Gastroenterol.* 11, 5006–5009.
- (20) Zhou, C. X., and Gao, Y. (2006) Aberrant expression of  $\beta$ -catenin, Pin1 and cyclin D1 in salivary adenoid cystic carcinoma: Relation to tumor proliferation and metastasis. *Oncol. Rep.* 16, 505–511.
- (21) Ryo, A., Liou, Y. C., Lu, K. P., and Wulf, G. (2003) Prolyl isomerase Pin1: A catalyst for oncogenesis and a potential therapeutic target in cancer. *J. Cell Sci.* 116, 773–783.
- (22) Lu, K. P. (2003) Prolyl isomerase Pin1 as a molecular target for cancer diagnostics and therapeutics. *Cancer Cell* 4, 175–180.
- (23) Wildemann, D., Erdmann, F., Alvarez, B. H., Stoller, G., Zhou, X. Z., Fanghanel, J., Schutkowski, M., Lu, K. P., and Fischer, G. (2006) Nanomolar inhibitors of the peptidyl prolyl cis/trans isomerase Pin1 from combinatorial peptide libraries. *J. Med. Chem.* 49, 2147–2150.
- (24) Hennig, L., Christner, C., Kipping, M., Schelbert, B., Rucknagel, K. P., Grabley, S., Kullertz, G., and Fischer, G. (1998) Selective inactivation of parvulin-like peptidyl-prolyl cis/trans isomerases by juglone. *Biochemistry* 37, 5953–5960.
- (25) Xu, G. G., and Etkorn, F. A. (2009) Pin1 as an anticancer drug target. *Drug News Perspect.* 22, 399–407.
- (26) Zhao, S., and Etkorn, F. A. (2007) A phosphorylated prodrug for the inhibition of Pin1. *Bioorg. Med. Chem. Lett.* 17, 6615–6618.
- (27) Yaffe, M. B., Schutkowski, M., Shen, M., Zhou, X. Z., Stukenberg, P. T., Rahfeld, J. U., Xu, J., Kuang, J., Kirschner, M. W., Fischer, G., Cantley, L. C., and Lu, K. P. (1997) Sequence-specific and phosphorylation-dependent proline isomerization: A potential mitotic regulatory mechanism. *Science* 278, 1957–1960.
- (28) Lu, P. J., Zhou, X. Z., Liou, Y. C., Noel, J. P., and Lu, K. P. (2002) Critical role of WW domain phosphorylation in regulating phosphoserine binding activity and Pin1 function. *J. Biol. Chem.* 277, 2381–2384.
- (29) Lippens, G., Landrieu, I., and Smet, C. (2007) Molecular mechanisms of the phospho-dependent prolyl cis/trans isomerase Pin1. *FEBS J.* 274, 5211–5222.
- (30) Huang, F. L., and Glinsmann, W. H. (1976) Separation and characterization of two phosphorylase phosphatase inhibitors from rabbit skeletal muscle. *Eur. J. Biochem.* 70, 419–426.
- (31) Ingebritsen, T. W., and Cohen, P. (1983) The protein phosphatases involved in cellular regulation. I. Classification and substrate specificities. *Eur. J. Biochem.* 132, 255–261.
- (32) Li, M., Satinover, D. L., and Brautigan, D. L. (2007) Phosphorylation and functions of inhibitor-2 family of proteins. *Biochemistry* 46, 2380–2389.
- (33) Ballou, L. M., and Fischer, E. H. (1986) Phosphoprotein Phosphatases. In *The Enzymes* (Boyer, P. D., and Krebs, E. G., Eds.) pp 311–361, Academic Press, Inc., Orlando, FL.
- (34) Hemmings, B. A., Resink, T. J., and Cohen, P. (1982) Reconstitution of a Mg-ATP-dependent protein phosphatase and its activation through a phosphorylation mechanism. *FEBS Lett.* 150, 319–324.
- (35) Puntoni, F., and Villa-Moruzzi, E. (1995) Phosphorylation of the inhibitor-2 of protein phosphatase-1 by cdc2-cyclin B and GSK3. *Biochem. Biophys. Res. Commun.* 207, 732–739.
- (36) Li, M., Stefansson, B., Wang, W., Schaefer, E. M., and Brautigan, D. L. (2006) Phosphorylation of the Pro-X-Thr-Pro site in phosphatase inhibitor-2 by cyclin-dependent protein kinase during M-phase of the cell cycle. *Cell. Signalling* 18, 1318–1326.
- (37) Wang, Q. M., Guan, K. L., Roach, P. J., and DePaoli-Roach, A. A. (1995) Phosphorylation and activation of the ATP-Mg-dependent protein phosphatase by the mitogen-activated protein kinase. *J. Biol. Chem.* 270, 18352–18358.
- (38) Leach, C., Shenolikar, S., and Brautigan, D. L. (2003) Phosphorylation of phosphatase inhibitor-2 at centrosomes during mitosis. *J. Biol. Chem.* 278, 26015–26020.
- (39) Li, M., Stukenberg, P. T., and Brautigan, D. L. (2008) Binding of phosphatase inhibitor-2 to prolyl isomerase Pin1 modifies specificity for mitotic phosphoproteins. *Biochemistry* 47, 292–300.
- (40) Eto, M., Elliott, E., Prickett, T. D., and Brautigan, D. L. (2002) Inhibitor-2 regulates protein phosphatase-1 complexed with NimA-related kinase to induce centrosome separation. *J. Biol. Chem.* 277, 44013–44020.
- (41) Wang, W., and Brautigan, D. L. (2008) Phosphatase inhibitor 2 promotes acetylation of tubulin in the primary cilium of human retinal epithelial cells. *BMC Cell Biol.* 9, 62.
- (42) Wang, W., Cronmiller, C., and Brautigan, D. L. (2008) Maternal phosphatase inhibitor-2 is required for proper chromosome segregation and mitotic synchrony during *Drosophila* embryogenesis. *Genetics* 179, 1823–1833.
- (43) Wang, W., Stukenberg, P. T., and Brautigan, D. L. (2008) Phosphatase inhibitor-2 balances protein phosphatase 1 and Aurora B kinase for chromosome segregation and cytokinesis in human retinal epithelial cells. *Mol. Biol. Cell* 19, 4852–4862.
- (44) Satinover, D. L., Leach, C. A., Stukenberg, P. T., and Brautigan, D. L. (2004) Activation of Aurora-A kinase by protein phosphatase inhibitor-2, a bifunctional signaling protein. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8625–8630.
- (45) Satinover, D. L., Brautigan, D. L., and Stukenberg, P. T. (2006) Aurora-A Kinase and Inhibitor-2 Regulate the Cyclin Threshold for Mitotic Entry in *Xenopus* Early Embryonic Cell Cycles. *Cell Cycle* 5, 2268–2274.
- (46) Liu, J., Wu, J., Oliver, C., Shenolikar, S., and Brautigan, D. L. (2000) Mutations of the serine phosphorylated in the protein phosphatase-1-binding motif in the skeletal muscle glycogen-targeting subunit. *Biochem. J.* 346, 77–82.
- (47) Hall, E. H., Daugherty, A. E., Choi, C. K., Horwitz, A. F., and Brautigan, D. L. (2009) Tensin1 requires protein phosphatase-1 $\alpha$  in addition to RhoGAP DLC-1 to control cell polarization, migration, and invasion. *J. Biol. Chem.* 284, 34713–34722.
- (48) Huang, H. B., Horiuchi, A., Watanabe, T., Shih, S. R., Tsay, H. J., Li, H. C., Greengard, P., and Nairn, A. C. (1999) Characterization of the inhibition of protein phosphatase-1 by DARPP-32 and inhibitor-2. *J. Biol. Chem.* 274, 7870–7878.

(49) Winkler, K. E., Swenson, K. I., Kornbluth, S., and Means, A. R. (2000) Requirement of the prolyl isomerase Pin1 for the replication checkpoint. *Science* 287, 1644–1647.

(50) Dancheck, B., Nairn, A. C., and Peti, W. (2008) Detailed structural characterization of unbound protein phosphatase 1 inhibitors. *Biochemistry* 47, 12346–12356.

(51) Dancheck, B., Ragusa, M. J., Allaire, M., Nairn, A. C., Page, R., and Peti, W. (2011) Molecular investigations of the structure and function of the protein phosphatase 1-spinophilin-inhibitor 2 heterotrimeric complex. *Biochemistry* 50, 1238–1246.

(52) Hurley, T. D., Yang, J., Zhang, L., Goodwin, K. D., Zou, Q., Cortese, M., Dunker, A. K., and DePaoli-Roach, A. A. (2007) Structural basis for regulation of protein phosphatase 1 by inhibitor-2. *J. Biol. Chem.* 282, 28874–28883.

(53) Terry-Lorenzo, R. T., Elliot, E., Weiser, D. C., Prickett, T. D., Brautigam, D. L., and Shenolikar, S. (2002) Neurabins recruit protein phosphatase-1 and inhibitor-2 to the actin cytoskeleton. *J. Biol. Chem.* 277, 46535–46543.

(54) Wang, H., and Brautigam, D. L. (2002) A novel transmembrane Ser/Thr kinase complexes with protein phosphatase-1 and inhibitor-2. *J. Biol. Chem.* 277, 49605–49612.

# Molecular Implication of PP2A and Pin1 in the Alzheimer's Disease Specific Hyperphosphorylation of Tau

Isabelle Landrieu<sup>1\*</sup>, Caroline Smet-Nocca<sup>1</sup>, Laziza Amniai<sup>1</sup>, Justin Vijay Louis<sup>2</sup>, Jean-Michel Wieruszeski<sup>1</sup>, Jozef Goris<sup>2</sup>, Veerle Janssens<sup>2</sup>, Guy Lippens<sup>1</sup>

**1** CNRS-UMR 8576, Federative Research Institute IFR 147, University of Lille-North of France, Villeneuve d'Ascq, France, **2** Protein Phosphorylation and Proteomics Group, Molecular Cell Biology Department, Faculty of Medicine of Leuven University, Leuven, Belgium

## Abstract

**Background:** Tau phosphorylation and dephosphorylation regulate in a poorly understood manner its physiological role of microtubule stabilization, and equally its integration in Alzheimer disease (AD) related fibrils. A specific phospho-pattern will result from the balance between kinases and phosphatases. The heterotrimeric Protein Phosphatase type 2A encompassing regulatory subunit PR55/B $\alpha$  (PP2A<sub>T55 $\alpha$</sub> ) is a major Tau phosphatase *in vivo*, which contributes to its final phosphorylation state. We use NMR spectroscopy to determine the dephosphorylation rates of phospho-Tau by this major brain phosphatase, and present site-specific and kinetic data for the individual sites including the pS202/pT205 AT8 and pT231 AT180 phospho-epitopes.

**Methodology/Principal Findings:** We demonstrate the importance of the PR55/B $\alpha$  regulatory subunit of PP2A within this enzymatic process, and show that, unexpectedly, phosphorylation at the pT231 AT180 site negatively interferes with the dephosphorylation of the pS202/pT205 AT8 site. This inhibitory effect can be released by the phosphorylation dependent prolyl *cis/trans* isomerase Pin1. Because the stimulatory effect is lost with the dimeric PP2A core enzyme (PP2A<sub>D</sub>) or with a phospho-Tau T231A mutant, we propose that Pin1 regulates the interaction between the PR55/B $\alpha$  subunit and the AT180 phospho-epitope on Tau.

**Conclusions/Significance:** Our results show that phosphorylation of T231 (AT180) can negatively influence the dephosphorylation of the pS202/pT205 AT8 epitope, even without an altered PP2A pool. Thus, a priming dephosphorylation of pT231 AT180 is required for efficient PP2A<sub>T55 $\alpha$</sub> -mediated dephosphorylation of pS202/pT205 AT8. The sophisticated interplay between priming mechanisms reported for certain Tau kinases and the one described here for Tau phosphatase PP2A<sub>T55 $\alpha$</sub>  may contribute to the hyperphosphorylation of Tau observed in AD neurons.

**Citation:** Landrieu I, Smet-Nocca C, Amniai L, Louis JV, Wieruszeski J-M, et al. (2011) Molecular Implication of PP2A and Pin1 in the Alzheimer's Disease Specific Hyperphosphorylation of Tau. PLoS ONE 6(6): e21521. doi:10.1371/journal.pone.0021521

**Editor:** Jurgen Gotz, The University of Sydney, Australia

**Received:** March 23, 2011; **Accepted:** May 30, 2011; **Published:** June 23, 2011

**Copyright:** © 2011 Landrieu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by ANR-05-blanc-6320-01 grant. The NMR facilities were funded by the Région Nord, CNRS, Pasteur Institute of Lille, European Community (FEDER), French Research Ministry and the University of Sciences and Technologies of Lille I. Financial support from the TGE RMN THC Fr3050 for conducting the research is gratefully acknowledged. L.A. was supported by a pre-doctoral fellowship of the CNRS/Région Nord-Pas de Calais (France). J.V.L., J.G. and V.J. are supported by grants of the 'Geconcerteerde OnderzoeksActies' of the Flemish government, the 'Interuniversity Attraction Poles' of the Belgian Science Policy P6/28 and the 'Fonds voor Wetenschappelijk Onderzoek-Vlaanderen'. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: isabelle.landrieu@univ-lille1.fr

## Introduction

Phosphorylation/dephosphorylation of the neuronal microtubule associated protein Tau regulates in a complex manner its capacity to assemble tubulin into microtubules. It is also associated with the presence of pathological fibrils in neurons of AD patients, which are mainly composed of hyperphosphorylated Tau. Monoclonal antibodies such as AT180 and AT8, recognizing respectively the pT231 [1] and pS202/pT205 [2] Tau phospho-motifs, are available for post-mortem diagnostics of the disease progression, and can define the neurofibrillary lesions at different stages of the disease [3]. The spatial hierarchy observed is equally accompanied by a temporal progression of the phosphorylation pattern of Tau. The T231 site, for example, becomes phosphorylated early in the

disease, and precedes phosphorylation at the pS202/pT205 AT8 site [4], [5].

The phosphorylation of Tau is a reversible process, which implies that the pathological hyperphosphorylation can result from a deregulation of kinase and/or phosphatase activity. Intense research has increased our understanding of the kinases that generate certain phosphorylation events, as well as their complex regulation. Calpain cleavage of the regulatory p35 subunit of CDK5 results in a membrane detachment of the resulting CDK5/p25 complex [6]. Although the intrinsic catalytic activity of this complex towards Tau seems not different from that of CDK5/p35 [7], the cytosolic presence of the CDK5/p25 complex might play a role in Tau's phosphorylation. As a downstream event, T231 phosphorylation by GSK3 $\beta$  requires priming of the S235 site by



CDK5 [8], [9], but CDK5 can equally inhibit directly the GSK3 $\beta$  enzyme through phosphorylation of the latter at its S9 [10].

One major brain phosphatase that dephosphorylates phospho-Tau is the Protein Phosphatase 2A (PP2A) [11], [12], [13], a multimeric enzyme consisting at least of a dimeric core enzyme (hereafter abbreviated PP2A<sub>D</sub>) constituted of a catalytic subunit (C subunit) and a scaffolding subunit (A subunit). This heterodimer (PP2A<sub>D</sub>) can further integrate a regulatory B-type subunit to form a heterotrimeric PP2A complex (PP2A<sub>T</sub>) [14], [15]. The particular regulatory B-type subunit, that exhibits itself specific spatial and temporal expression patterns [16], [17], regulates the catalytic activity and specificity of the heterotrimeric PP2A towards the target phospho-protein(s) [18]. One of the most prominent neuronal forms, highly expressed together with the PP2A catalytic C subunit in neurons, is the B-type subunit PR55/B $\alpha$ , whose crystal structure in the heterotrimeric enzyme (hereafter abbreviated PP2A<sub>T55 $\alpha$</sub> ) has recently been determined [19]. A reduced PP2A activity was shown to induce hyperphosphorylation of Tau, including at the pS202/pT205 AT8 site in transgenic mice [20], and has equally been shown in AD at various levels. Firstly, in the hippocampus, a decrease in mRNA expression of the PP2A catalytic C subunit and various B-type subunits has been observed [21]. Secondly, at the protein level, a decreased level of the PR55/B $\alpha$  subunit due to an increased turnover is associated with the AD pathology [22].

The regulation of PP2A is complex, and involves more than the subunits of the holoenzyme. Polycations can stimulate the activity of PP2A through a direct interaction with the enzyme. As the stimulation is dependent on the type of holoenzyme, the interaction involves the regulatory B-type subunit [23], although an interaction with the substrate has also been reported [24]. An inhibitory interaction was described for  $\alpha$ -endosulfine and Arpp-19, two small proteins that after phosphorylation by the Greatwall kinase interact with the PR55/B $\delta$  regulatory subunit, and thereby shut down the activity of PP2A [25], [26]. Another regulatory factor for PP2A activity is Pin1, a phosphorylation dependent prolyl *cis/trans* isomerase that targets specifically the pS/pT-P motifs. This isomerase was shown to stimulate the dephosphorylation of Tau by PP2A [27]. Pin1<sup>-/-</sup> mice show indeed a decreased phosphatase activity towards pS/pT-P motifs and an accumulation of both the pS202/pT205 AT8 and pT231/pS235 AT180 epitopes [28].

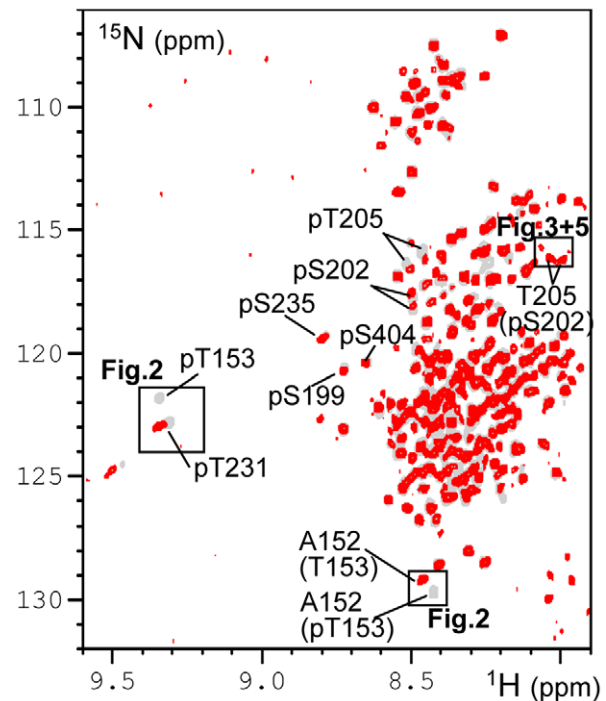
Kinases and phosphatases can hence regulate the phosphorylation/dephosphorylation of Tau in a complex and concerted manner, whereby this intricate feedback can *in vivo* control the final phosphorylation level of Tau [29]. Our present work aims to better define the molecular role of PP2A and Pin1 in this process. We first address the question of whether the different phosphorylation sites of Tau are independent with respect to PP2A catalyzed dephosphorylation, or whether the equivalent of a priming mechanism described for certain kinases exists also for some dephosphorylation reactions. Secondly, we investigate the role of Pin1 in stimulating this PP2A-catalyzed dephosphorylation activity towards specific sites. We use Tau *in vitro* phosphorylated by the activated CDK2/CycA3 complex to study the effect of PP2A on its dephosphorylation. We showed previously that CDK2/CycA3 action at the AT180 epitope is equivalent to that of the combined CDK5/p25 and GSK3 $\beta$  kinases and can generate in a robust manner the pS202/pT205 AT8 and pT231/pS235 AT180 epitopes on Tau [30], [31]. We use NMR spectroscopy as an analytical technique that provides a direct and quantitative view on all phosphorylation events in the full length protein [30], [32]. The dephosphorylation reaction directly performed in the NMR tube allows kinetic monitoring of the individual phosphorylation sites in

one single experiment. We reveal a subtle regulation of PP2A<sub>T55 $\alpha$</sub>  activity towards the Tau pS202/pT205 AT8 site by the phosphorylation status of T231, whereby phosphorylation of the latter T231 site negatively interferes with the PP2A catalyzed dephosphorylation of pS202. In addition, we find that Pin1 releases this negative feedback between the pT231 and pS202/pT205 AT8 epitopes. Our findings define an additional level of molecular regulation of the PP2A phosphatase towards the phosphorylated Tau protein, whereby the interplay between kinase and phosphatase activity can potentially lead to a stable hyperphosphorylated state of Tau that characterizes AD affected neurons.

## Results

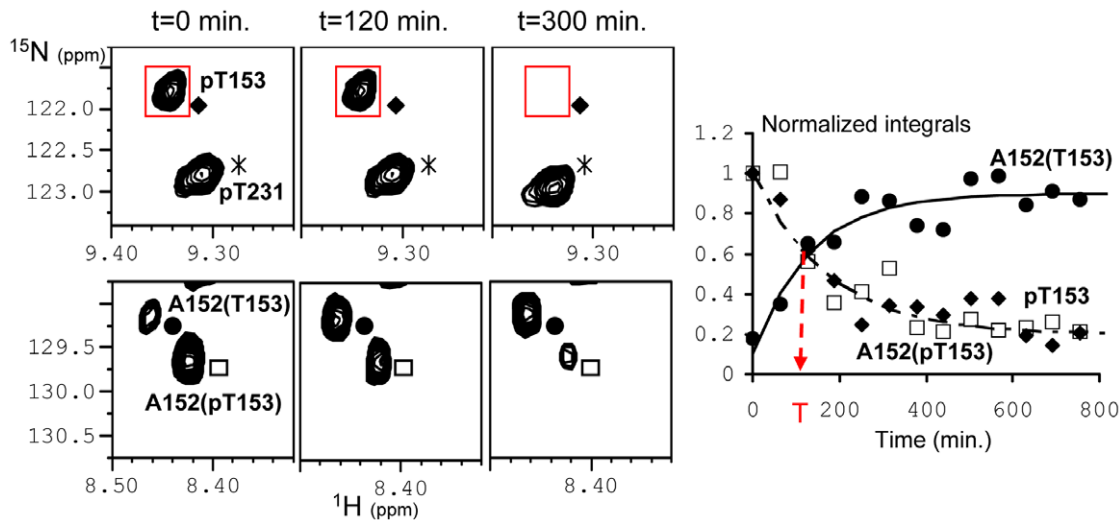
### Specificity of the heterotrimeric PP2A<sub>T55 $\alpha$</sub> for Pro-directed CDK2/CycA3 phospho-sites of Tau

We previously showed that the recombinant CDK2/CycA3 complex generates robust and reproducible phosphorylation of Tau at AD-specific epitopes, with high levels of phosphate incorporation at pS202-pT205 and pT231-pS235 ([30]; **Fig. S1**), recognized respectively by the diagnostic AT8 [1] and AT180 [2] AD-specific monoclonal antibodies. Additional phosphorylation events were observed at position T153 (70–80%), S199 (70–80%) and S404 (50–60%). The heterotrimeric PP2A<sub>T55 $\alpha$</sub>  complex composed of the scaffolding A subunit, the catalytic C subunit and the regulatory PR55/B $\alpha$  subunit, was added to this CDK2/CycA3-phosphorylated Tau441 sample (phospho-Tau) directly in the NMR tube (**Fig. 1**). Consecutive [<sup>1</sup>H,<sup>15</sup>N] Heteronuclear Single Quantum Correlation spectra (HSQC; hereafter named 2D



**Figure 1. Dephosphorylation of phospho-Tau by the heterotrimeric PP2A<sub>T55 $\alpha$</sub> .** Comparison of the [<sup>1</sup>H,<sup>15</sup>N] 2D spectrum of the CDK2/CycA3 phospho-Tau (20  $\mu$ M, in gray) and same sample incubated 16 hours at 25°C (293K) with 1U of PP2A<sub>T55 $\alpha$</sub>  enzyme (superimposed in red). Cross peaks of the amide function of the corresponding phosphorylated residues are labelled as well as A152(T153 or pT153) and T205(pS202) resonances. Regions of the spectrum enlarged in Fig. 2, 3 and 5 are boxed.

doi:10.1371/journal.pone.0021521.g001



**Figure 2. Kinetic of dephosphorylation of phospho-Tau by the heterotrimeric PP2A<sub>T55α</sub>.** Panels from left to right Control 2D spectrum of phospho-Tau before addition of PP2A<sub>T55α</sub> (t = 0 min.), second (t = 120 min.) and sixth 2D spectrum (t = 300 min.) in the serie after addition of PP2A<sub>T55α</sub> (1U). A152(pT153) and A152(T153) stand for the amide resonance of A152 next to a pT153 or a T153 residue in the Tau molecule, respectively. A box is drawn around a peak (see red box around pT153) and is replicated in each spectrum that corresponds to a time-point of the reaction. The surface of the peak inside this box corresponds to an integral value or data-point in the graph on the right. This graph shows normalized integral values plotted as a function of time (min.) for pT153 (black diamonds), A152(pT153) (open squares) and A152(T153) (black dots). Data are fitted with mono-exponential curves characterized by the constant T (min.), indicated on the x-axis by a red arrow and reported in **Table 1**. doi:10.1371/journal.pone.0021521.g002

spectra) yield snapshots of the dephosphorylation reaction with a one hour time resolution, corresponding to the acquisition time of a single 2D spectrum (**Fig. 2, Data S1**). Integration of a peak surface for a given resonance in each 2D spectrum during this *in-spectrometer* dephosphorylation reaction gives information on the kinetics of the modification of the corresponding amino acid (**Fig. 2**). Because phosphorylation does not only affect the amide resonance of the modified amino acid but also that of its neighbours, the same reaction can in favourable cases be followed on different cross peaks. The A152 amide peak resonates for example at slightly different frequencies dependent on the phosphorylation status of T153, and the same enzymatic reaction could be followed as the decrease of the pT153 or the A152(pT153) amide

cross peaks, but equally as the increase of the A152(T153) amide correlation (**Fig. 2**). Dephosphorylation of pT205 also followed an exponential trend with a similar time constant (**Table 1**) that was best monitored by the unphosphorylated T205 resonance(s) (**Fig. 3**), located in a relatively sparse region of the spectrum [33]. Initially, no peak appears at its exact resonance position in the unphosphorylated Tau protein, but we rather observe two intermediate peaks corresponding to T205(pS202) with a further splitting due to the phosphorylation state of S199 (**Fig. 3A**). This assignment was validated by a dephosphorylation experiment on the phospho-TauS199A/S202A double mutant, where the intermediate forms do not appear, but where we recover immediately the T205 resonance at its position in the unmodified Tau spectrum (**Fig. 3B**). When increasing the phosphatase concentration by a factor of 3.5, a partial dephosphorylation of the complete AT8 epitope on wild type phospho-Tau could be obtained, with 30% of the intensity at the position of T205(S202) after 16 hours of incubation at 25°C (**Fig. 3C**). This latter observation also verifies that we do monitor the enzymatic reaction catalysed by PP2A<sub>T55α</sub> and not a spontaneous dephosphorylation under our NMR conditions. In addition to the fast dephosphorylation observed for pT153 and pT205 with the phosphate almost completely removed by 3.5 units of PP2A<sub>T55α</sub> after the first two spectra, two additional classes of residues could be distinguished. The first one contains pT231 and pS404, and is characterized by very slow kinetics with only a marginal reduction of the phospho-resonance after one night. The second group shows intermediate dephosphorylation kinetics, and concerns pS199, pS202 and pS235 (**Fig. 1–4** and **Table 1**).

#### Dephosphorylation of phospho-Tau by the PP2A<sub>D</sub> core enzyme

Several reports have shown the implication of the PR55/B $\alpha$  regulatory subunit in the substrate anchoring and enzyme efficiency of PP2A<sub>T55α</sub> [19], [34]. In order to evaluate whether the dimeric enzyme PP2A<sub>D</sub>, composed of the A and C subunits without any B-type subunit, would maintain a phosphatase

**Table 1. Specificity of dephosphorylation of the CDK2/CycA3 phospho-sites of Tau by the PP2A<sub>T55α</sub> and PP2A<sub>D</sub> phosphatases.**

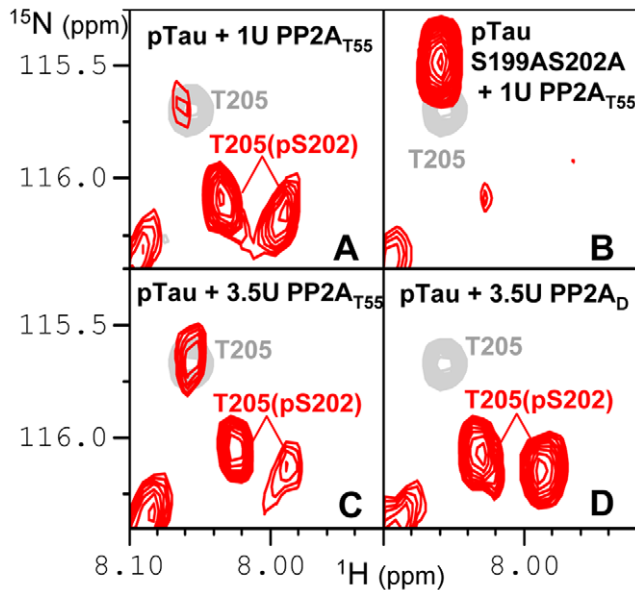
	Tau w.t.			TauT231A
	Trimer (1 U)	Trimer (3.5 U)	Dimer (3.5 U)	Trimer (1 U)
pT153	170	60	200	75
pS199	240 (30%)	220 (60%)	260 (45%)	200 (40%)
pS202	n.d.	300 (25%)	n.d.	200
pT205	160	55	120	65
pT231	n.d.	n.d.	n.d.	n.d.
pS235	n.d.	500 (40%)	n.d.	175 (50%)
pS404	n.d.	n.d.	n.d.	n.d.

Integral values of the resonances during a representative dephosphorylation experiment in the spectrometer are fitted with mono-exponential functions of type  $\text{Integral} = e^{-(t/T)}$  or  $1 - (e^{-(t/T)})$  with time t and constant T in min, reported in the Table.

Percentage in brackets ( ): extent of dephosphorylation if not complete. n.d.: not dephosphorylated.

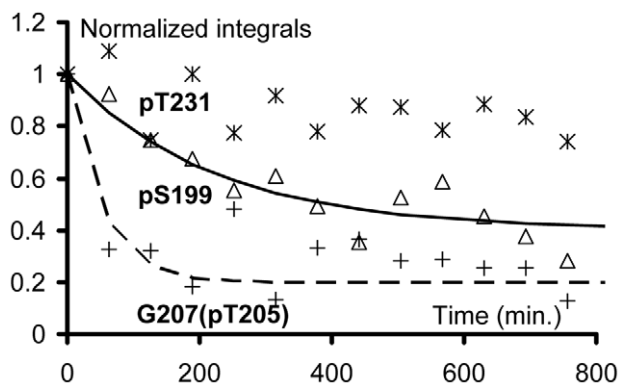
doi:10.1371/journal.pone.0021521.t001





**Figure 3. Regulation of the dephosphorylation of the AT8 epitope.** Details of the 2D spectra (in red) of phospho-Tau (pTau, A, C, D) and phospho-Tau S202A5199A (pTauS202A5199A, B) after 16 hours of incubation at 25°C (293K) with 1U (A, B) or 3.5U PP2A<sub>T55α</sub> (C) and 3.5U of PP2A<sub>D</sub> (D). The spectra are superimposed on the spectrum of the unmodified Tau protein (in gray). T205(pS202) stands for the amide resonance of T205 with a pS202 neighbour in the Tau molecule. doi:10.1371/journal.pone.0021521.g003

activity, we repeated the same experiments as above with PP2A<sub>D</sub>. For an identical enzyme activity (3.5 units) of PP2A<sub>D</sub> heterodimer and PP2A<sub>T55α</sub> heterotrimer as measured with phosphorylase *a* as substrate, the dephosphorylation rate at all phospho-sites of Tau was slower for the PP2A<sub>D</sub> dimer (**Table 1**). For pT153, we found a three-fold slower rate, and a two-fold slow-down for pT205. Significantly, dephosphorylation of pS202 at the AT8 epitope, such as partially (25%) obtained with 3.5 units of PP2A<sub>T55α</sub> after



**Figure 4. Kinetic of dephosphorylation of phospho-Tau by the heterotrimeric PP2A<sub>T55α</sub>.** The dephosphorylation kinetics are obtained by integration of the cross peak of the amide function of the corresponding phospho-residue in the consecutive [<sup>1</sup>H,<sup>15</sup>N] 2D spectra encompassing 60 min. of the *in-spectrometer* reaction at 25°C (293K). The graph shows the differential dephosphorylation by PP2A<sub>T55α</sub> (3.5 units) of phospho-Tau. Fitting of the data by a mono-exponential curve yielded the time constants T (min.) reported in **Table 1**. pT231 is represented by stars, pS199 by triangles and G207(pT205) by crosses. G207 stands for G207 next to a pT205 in the Tau molecule. doi:10.1371/journal.pone.0021521.g004

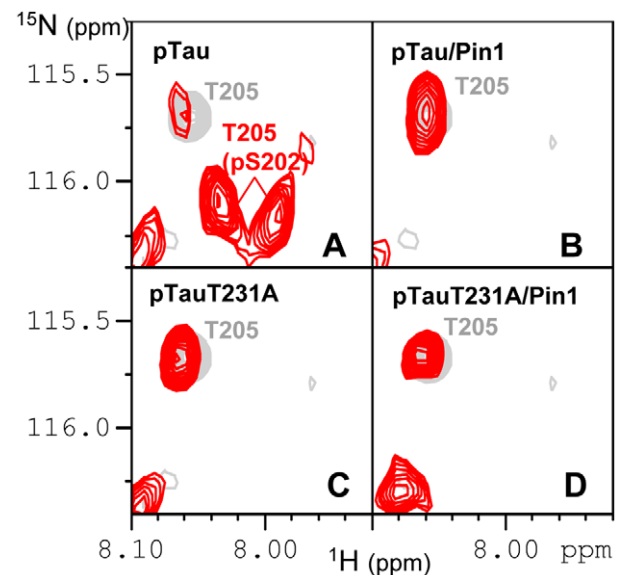
16 hours of incubation at 25°C, did not occur at all with the same amount of PP2A<sub>D</sub> (**Fig. 3D**).

#### Dephosphorylation of the mutant phospho-Tau T231A by PP2A<sub>T55α</sub>

The <sup>224</sup>KKVAVVRTPPKSP<sup>236</sup> peptide of Tau interacts directly with the PR55/Bα subunit, and is able to compete with native Tau for binding to PP2A<sub>T55α</sub> [34]. Concordantly, the substrate binding groove of PR55/Bα shows a negative character, suggesting a charge-charge interaction between both proteins. Charge-inverting mutations on PR55/Bα indeed led to a decreased level of catalytic efficiency of the holoenzyme [19]. Phosphorylation of the T231 residue would be expected to diminish the positive character of the Tau peptide, and might hence also modulate this interaction and possibly the enzymatic efficiency of the holoenzyme. Measuring the dephosphorylation of a phospho-Tau T231A sample with 1 unit of PP2A<sub>T55α</sub> phosphatase, we indeed found an increased dephosphorylation rate compared to the wild type phospho-Tau, with a gain of more than two-fold for pT153 and pT205 (**Table 1**). With this low amount of phosphatase, the pS202 position that was not dephosphorylated in the wild type phospho-Tau sample after 16 hours, becomes almost completely dephosphorylated in the phospho-Tau T231A (**Fig. 5**, compare panels A and C). The phosphorylation status of the AT180 epitope hence influences directly the PP2A<sub>T55α</sub> catalyzed dephosphorylation rate of phospho-Tau and most strikingly of the AT8 epitope.

#### Regulation of the PP2A<sub>T55α</sub> activity by Pin1

Using a *cdc2* phosphorylated Tau as substrate, Zhou et al. described a stimulatory effect of Pin1 on the dephosphorylation activity of PP2A<sub>T55α</sub> [27]. Dephosphorylation of phospho-Tau T231A not being affected by Pin1 led to the hypothesis that the phosphate release assay used in that study monitored the



**Figure 5. Regulation of the dephosphorylation of the AT8 epitope.** Details of the 2D spectrum (in red) of CDK2/CycA3 phospho-Tau (pTau, A and B) and phospho-Tau T231A (pTauT231A, C and D) after 16 hours of incubation at 25°C (293K) with PP2A<sub>T55α</sub> (1U) superimposed on the spectrum of the unmodified Tau protein (in gray). pTau and pTauT231A correspond to dephosphorylation without Pin1 (A and C) and pTau/Pin1 and pTauT231A/Pin1 with an excess of Pin1 (B and D). doi:10.1371/journal.pone.0021521.g005

dephosphorylation at this pT231 site, and suggested a regulatory role of Pin1 through its *cis/trans* isomerase activity at this precise pT231-Pro motif [27]. Our result that pT231 is at best marginally dephosphorylated by PP2A<sub>T55α</sub> (**Fig. 2, 4**) prompted us to re-examine the precise role of Pin1 in the framework of Tau dephosphorylation by PP2A<sub>T55α</sub>. Because the interaction with Pin1 (and most probably its WW domain, [35]) broadens the NMR signals of the phospho-resonances beyond detection, we monitored the PP2A kinetics through the appearance of the non-phosphorylated forms. The rate of dephosphorylation of pT205 (T = 160 min.) was indeed increased 3 fold by the presence of a threefold excess of Pin1 (T = 55 min.), which is the same ratio of Pin1:phospho-Tau as used in the initial report [27]. Remarkably, pS202, although not dephosphorylated at all by 1 unit of PP2A<sub>T55α</sub>, is completely dephosphorylated in the presence of Pin1 (**Fig. 5, compare panels A and B**). This stimulatory effect of Pin1 is however selective, as the dephosphorylation of pT231, monitored by the appearance of the R230(T231) peak, is still negligible.

We next monitored the dephosphorylation of the same phospho-Tau by PP2A<sub>D</sub> in the absence or presence of Pin1. Although already of reduced activity when compared to PP2A<sub>T55α</sub> (**Fig. 3, compare panels C and D**), the efficiency of the core enzyme slowed down even further upon the addition of Pin1 (**Fig. S2**). We finally tested the effect of Pin1 on the enzymatic efficiency of the PP2A<sub>T55α</sub> phosphatase complex towards the phospho-Tau T231A mutant. The dephosphorylation of the complete AT8 epitope was faster than in the wild type phospho-Tau (**Fig. 5, compare panels A and D**), but was not further stimulated by Pin1 (**Fig. 5, compare panels C and D**).

## Discussion

Tau phosphorylation is a reversible process that regulates in a complex manner its physiological role of microtubule stabilization. It is also linked to its pathological role, as the neurofibrillary tangles found inside the neurons of AD patients invariably contain a hyperphosphorylated form of Tau. Analysis with specific antibodies have shown that the latter form is characterized by the simultaneous presence of multiple phosphorylated residues, including the S202/T205 and T231 positions that constitute respectively the AT8 and AT180 epitopes [2]. These phosphorylation events seemingly follow a hierarchical appearance, T231 being one of the earliest sites to be phosphorylated in AD brain before the epitope recognized by the AT8 antibody [4], [36]. In the axon, PP2A<sub>T55α</sub> is associated with the microtubules [34], and its phosphatase activity counteracts the appearance of this multiply phosphorylated Tau [29]. In addition, it has recently been reported that other neuronal PP2A heterotrimers might indirectly contribute to Tau phosphorylation/dephosphorylation by regulating the activities of GSK3β and CDK5 kinases [37]. However, at a certain stage of the disease, this kinase-phosphatase balance breaks down, and one does observe the accumulation of hyperphosphorylated Tau in the somato-dendritic compartment [29], [38], [39].

In this report, we investigate at the molecular level the enzymatic dephosphorylation of Tau by PP2A<sub>T55α</sub>, the major brain isoform of PP2A directly associated to microtubules and Tau [16], [34], and ask whether certain phosphorylation events on Tau might exert an effect towards its activity at other sites. Our *in vitro* set-up with CDK2/CycA3 kinase generating the AT8 and AT180 epitopes on recombinant <sup>15</sup>N-labelled Tau [30] and concomitant NMR analysis gives a global view of the Tau phosphorylation pattern as a function of time. We thereby find that PP2A<sub>T55α</sub> selectively targets certain Tau phosphorylation sites, with fast

dephosphorylation of the pT205 and pT153 sites but hardly any for pT231 (**Table 1**). Whereas inefficient dephosphorylation of pT231 was previously described for PHF-Tau [40], our finding of efficient dephosphorylation for the pT153 site has not yet been described in the literature because of the absence of specific antibody. This underscores the advantage of the global view conferred by our NMR approach.

The dimeric PP2A<sub>D</sub> (PP2A AC) isoform has a lower catalytic efficiency on all Tau phospho-sites compared to the heterotrimeric PP2A<sub>T55α</sub>, emphasizing the crucial role of the B subunit in the function of the PP2A holoenzyme. A correlation between the interaction of Tau with the PR55/Bα subunit and the catalytic efficiency of PP2A<sub>T55α</sub> was previously described [34], [41], and has recently obtained a molecular basis with the crystal structure of the PP2A<sub>T55α</sub> holoenzyme [19]. The β-propeller constituting the PR55/Bα subunit indeed contains an acidic groove constituting a potential binding site for the basic <sub>224</sub>KKVAVVVRTPPKSP<sub>236</sub> Tau peptide that can compete with full-length Tau for binding to PP2A [34]. In this mode of interaction, the PR55/Bα regulatory subunit would thus play a role in anchoring the enzyme to its substrate, or alternatively, would exert an allosteric stimulatory effect as observed for protamine or poly-lysine [23]. We did observe that all sites were more rapidly dephosphorylated in the phosphorylated Tau T231A, implying that phosphorylation at the T231 position reduces the PP2A activity. Phosphorylation at this position could modulate the binding mode of the PR55/Bα subunit leading to the inhibition of the PP2A activity. The recent discovery of α-endosulfine or Arpp-19, that both require phosphorylation of S67 by the Greatwall kinase within the already acidic peptide D<sub>66</sub>SGDD<sub>69</sub> in order to efficiently interact with the PR55/Bδ subunit of the Xenopus mitotic PP2A heterotrimer [26], [25], suggests that different binding modes to the PR55/B regulatory subunits may exist. Phosphorylation of T231, be it by GSK3β after CDK5 priming at the S235 position [8], [9] or by another kinase combination, thus stimulates the phosphorylation at the AT8 epitope via negative feedback on PP2A<sub>T55α</sub> phosphatase activity. This potentially can lead to a stable state with both epitopes phosphorylated [42], as is found in AD neurons [2].

An additional level of regulation has been described for the phospho-dependent prolyl *cis/trans* isomerase Pin1 [43]. Here, we confirm the stimulatory effect of Pin1 on the PP2A<sub>T55α</sub> catalyzed dephosphorylation of wild type phospho-Tau and its absence in the Tau T231A mutant. Stimulation was most striking for the pS202 site, switching from no dephosphorylation with one unit of PP2A<sub>T55α</sub> in the absence of Pin1 to complete dephosphorylation in its presence. The effect is however not homogeneous, with at best a weak activity of PP2A<sub>T55α</sub> towards pT231 despite the presence of Pin1. Moreover, the stimulatory effect directly depends on the PR55/Bα subunit, as Pin1 rather hinders than stimulates the activity of the dimeric core enzyme. Pin1 hence counterbalances the negative regulation of PP2A<sub>T55α</sub> activity towards the AT8 site following phosphorylation of the AT80 pT231 site, in agreement with the inverse correlation between Pin1 expression and actual neurofibrillary degeneration in AD [28]. Because PP2A<sub>T55α</sub> has been shown to regulate the phosphorylation status of Tau *in vivo*, a potential trigger for PHF formation might be a shift of the balance between neuronal kinases and phosphatases. Our results show that phosphorylation of T231 directly influences the dephosphorylation rate of pS202, without necessitating an alteration in the PP2A<sub>T55α</sub> pool. We have reconstituted here the complex system of phosphorylated Tau and the trimeric PP2A<sub>T55α</sub> phosphatase in the NMR tube with several recombinant proteins, and thereby unravelled a subtle regulatory mechanism. An additional level of

regulation has been shown by the prolyl *cis/trans* isomerase Pin1. The phosphorylation/dephosphorylation reactions *in vivo* might even be more complex, because of the presence of tubulin and/or other regulatory factors, but even these can be added in the NMR tube [44]. Whereas we previously showed that NMR observation of Tau in live *Xenopus* oocytes is feasible [45,46], we believe that both approaches will bring novel insights in the complex regulation of multi-site phosphorylation.

## Materials and Methods

### Expression and purification of Recombinant proteins

Preparation of the recombinant Tau proteins, isotopic enrichment and phosphorylation are described in [30], [32]. The Pin1 protein was expressed as a His Tag fusion from a pET15b plasmid in BL21(DE3) *E. coli* strain [47]. The His Tag was not removed before the assays.

### Preparation of CDK2-CycA3 phospho-Tau

CDK2/CycA3 were prepared as described in [48] except that the bacterial extracts containing recombinant GST-CDK2pT160 (Glutathion S transferase fusion protein) and CycA3 were mixed before being applied on a Glutathion sepharose FF (5 ml, GE Healthcare). The CDK2/CycA3 complex was recovered from the resin by proteolysis, performed at 4°C overnight with the Pre-Scission protease (GE Healthcare) directly on the Glutathion Sepharose beads in buffer 50 mM Tris pH 8.0, 20 mM NaCl, 1 mM EDTA, 1 mM DTT. The concentration of the complex was evaluated at 280 nm with an extinction coefficient  $\epsilon$  of  $67420 \text{ M}^{-1}\text{cm}^{-1}$ . The complex was kept frozen at  $-80^\circ\text{C}$  in this buffer until used.

### Purification of dimeric and trimeric PP2A

Purification of PP2A<sub>D</sub> and PP2A<sub>T55 $\alpha$</sub>  from rabbit skeletal muscle is described in [23]. Calibration of the activity was done using the standard PP2A substrate phosphorylase *a* [23]. One unit is defined as the activity corresponding to removal of 1 nmole/min of <sup>32</sup>P-phosphate at 30°C.

### NMR sample preparation

NMR buffer was 25 mM Tris-d11, 25 mM NaCl, 2.5 mM EDTA, 1 mM DTT, pH 6.8. Dephosphorylation of 20  $\mu\text{M}$  of CDK2/CycA3 phospho-Tau (or Tau mutants), as defined by absorption at 280 nm, was performed in a volume of 160  $\mu\text{l}$  in 3 mm NMR tubes. Samples with Pin1 are 20  $\mu\text{M}$  phospho-Tau/60  $\mu\text{M}$  Pin1. The PP2A enzyme diluted in 20  $\mu\text{l}$  buffer was directly added in the NMR tube to 140  $\mu\text{l}$  of the phospho-Tau, immediately before starting acquisition of the 2D spectra series.

### NMR spectroscopy

Spectra were acquired at 293K on a 800 MHz equipped with a 3 mm Probe. Parameters of the [<sup>1</sup>H,<sup>15</sup>N]-HSQC (HeteroNuclear Single Quantum Spectroscopy) were 32 scans, 2048 points in the direct dimension and 128 points in the indirect dimension. An optimized d1 of 0.8 s was used. The [<sup>1</sup>H,<sup>15</sup>N]-HSQC are named 2D spectra in the text.

### Data analysis

The peak integrals (**Data S1**) were normalized against an invariant and intense peak in the spectrum corresponding to A85, set to 1 in the integration. Integration was performed with the Topspin 2.1 software (Bruker, Karlsruhe, Germany). Integral data were fitted with a mono-exponential function.

## Supporting Information

**Figure S1 Identification of the phospho-sites of CDK2/CycA3 phospho-Tau.** Superimposition of the [<sup>1</sup>H,<sup>15</sup>N] 2D spectra of the Tau protein (black) and of the CDK2/CycA3 phospho-Tau protein (red). Inset: the enlarged region of the spectrum illustrates phosphorylation of residues S202 and T205. (PDF)

**Figure S2 Regulation of the dephosphorylation of the pS202/pT205 AT8 epitope.** Details of the 2D spectrum (in red) of CDK2/CycA3 phospho-Tau after 16 hours of incubation at 25°C (293K) with PP2A<sub>D</sub> (3.5 U) superimposed on the spectrum of the unmodified Tau protein (in gray). **A** dephosphorylation without Pin1 and **B** in presence of an excess of Pin1. (PDF)

**Data S1 Step by step procedure of NMR data treatment. I.** A serie of 2D spectra is acquired, in this example we limited the serie to the first five 2D spectra. The first one in the serie is the control experiment without enzyme and is time zero or starting point in the *in spectrometer* kinetics of dephosphorylation. The following 2D spectra, after the addition of the enzyme, are acquired successively while the NMR tube remains in the spectrometer. This experimental dataset corresponds to the one presented in Fig. 2 and Fig. 4. **II.** Each 2D spectrum required 63 minutes for acquisition. Each data-point is thus obtained every 63 minutes: that is the time resolution in this experiment. **III.** The succession of the 2D spectra will allow to follow the kinetics of the reaction. Duration of each spectrum is cumulated. In this example, we cover with the five 2D spectra 252 minutes of the enzymatic reaction. **IV.** The surface of a peak in a boxed area in the 2D spectra (see point I) or Integrals is calculated by TOPSPIN2.1 Bruker software. This value is directly related to the amount of the corresponding amino acid residue in solution. We follow the evolution of the peak surface or Integral for each amino acid residue of interest (I1 to I4) during the reaction, as detected in each successive spectrum. In our example, I1 corresponds to pT153, I2 to pT231 and I3 to A152. I4 is linked to the resonance of amino acid residue A77. **V.** Data of integration is normalized against integral value of A77 (I4), an invariant and intense peak in the spectra, in order to compensate for any variation of intensity due to potential spectrum to spectrum modification of conditions (for example, addition of the enzyme). **VI.** Data are normalized against the highest integral value of the kinetics for each amino acid residue. In that way, the starting point or ending point of the kinetics is set to one, allowing easier comparison of the evolution of the amount of various amino acid residues during the course of the enzymatic reaction. **VII.** A final table is obtained containing the normalized integral values for each time-point in the kinetics. **VIII.** For each amino acid residue, an exponential curve is fitted to the experimental data by adjusting the value T in the exponential formula  $e^{-(t/T)}$ . The value T is found by minimizing the sum of the square differences between the calculated and observed data (compare table values in VII and VIII). This constant characterizes the rate of dephosphorylation and is reported in Table 1. A scaling factor (0.8) as well as a residual value depending of the extent of dephosphorylation (0.2 or 0.1) is also used to adjust the formula to the data. Values in the table IX are calculated based on the formula in the *left column* for pT153 (decreasing exponential) and A152 (increasing exponential). **IX.** Data-points from table in VII are graphically represented with the axis being the time-points (each 2D spectrum) and the ordinate the normalized integral value for each time-point (spectrum) reported for several amino acid residue of interest. In this example, pT153 or

11 (open squares) shows a decrease of its initial integral value or peak surface that can be interpreted as a decrease of the corresponding chemical species in solution, in this case due to the dephosphorylation by PP2A. The resonance of A152 (black dots) shows accordingly an increase of its integral value. pT231 integral value remains unaffected throughout the *in spectrometer* enzymatic reaction (stars). The fitted exponential curves are represented by lines.  
(PDF)

## References

- Goedert M, Jakes R, Vanmechelen E (1995) Monoclonal antibody AT8 recognises tau protein phosphorylated at both serine 202 and threonine 205. *Neurosci Lett* 189: 167–169.
- Goedert M, Jakes R, Crowther RA, Cohen P, Vanmechelen E, et al. (1994) Epitope mapping of monoclonal antibodies to the paired helical filaments of Alzheimer's disease: identification of phosphorylation sites in tau protein. *Biochem J* 301(Pt 3): 871–877.
- Braak H, Alafuzoff I, Arzberger T, Kretschmar H, Del Tredici K (2006) Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol* 112: 389–404.
- Augustinack JC, Schneider A, Mandelkow EM, Hyman BT (2002) Specific tau phosphorylation sites correlate with severity of neuronal cytopathology in Alzheimer's disease. *Acta Neuropathol* 103: 26–35.
- Luna-Munoz J, Garcia-Sierra F, Falcon V, Menendez I, Chavez-Macias L, et al. (2005) Regional conformational change involving phosphorylation of tau protein at the Thr231, precedes the structural change detected by Alz-50 antibody in Alzheimer's disease. *J Alzheimers Dis* 8: 29–41.
- Patrick GN, Zukerberg L, Nikolic M, de la Monte S, Dikkes P, et al. (1999) Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration. *Nature* 402: 615–622.
- Peterson DW, Ando DM, Taketa DA, Zhou H, Dahlquist FW, et al. (2010) No difference in kinetics of tau or histone phosphorylation by CDK5/p25 versus CDK5/p35 *in vitro*. *Proc Natl Acad Sci U S A* 107: 2884–2889.
- Cho JH, Johnson GV (2004) Primed phosphorylation of tau at Thr231 by glycogen synthase kinase 3beta (GSK3beta) plays a critical role in regulating tau's ability to bind and stabilize microtubules. *J Neurochem* 88: 349–358.
- Lin YT, Cheng JT, Liang LC, Ko CY, Lo YK, et al. (2007) The binding and phosphorylation of Thr231 is critical for Tau's hyperphosphorylation and functional regulation by glycogen synthase kinase 3beta. *J Neurochem* 103: 802–813.
- Plattner F, Angelo M, Giese KP (2006) The roles of cyclin-dependent kinase 5 and glycogen synthase kinase 3 in tau hyperphosphorylation. *J Biol Chem* 281: 25457–25465.
- Drewes G, Mandelkow EM, Baumann K, Goris J, Merlevede W, et al. (1993) Dephosphorylation of tau protein and Alzheimer paired helical filaments by calcineurin and phosphatase-2A. *FEBS Lett* 336: 425–432.
- Goedert M, Jakes R, Qi Z, Wang JH, Cohen P (1995) Protein phosphatase 2A is the major enzyme in brain that dephosphorylates tau protein phosphorylated by proline-directed protein kinases or cyclic AMP-dependent protein kinase. *J Neurochem* 65: 2804–2807.
- Gong CX, Grundke-Iqbal I, Iqbal K (1994) Dephosphorylation of Alzheimer's disease abnormally phosphorylated tau by protein phosphatase-2A. *Neuroscience* 61: 765–772.
- Janssens V, Goris J (2001) Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem J* 353: 417–439.
- Janssens V, Longin S, Goris J (2008) PP2A holoenzyme assembly: in cauda venenum (the sting is in the tail). *Trends Biochem Sci* 33: 113–121.
- Schmidt K, Kins S, Schild A, Nitsch RM, Hemmings BA, et al. (2002) Diversity, developmental regulation and distribution of murine PR55/B subunits of protein phosphatase 2A. *Eur J Neurosci* 16: 2039–2048.
- Martens E, Stevens I, Janssens V, Vermeesch J, Gotz J, et al. (2004) Genomic organisation, chromosomal localisation tissue distribution and developmental regulation of the PR61/B' regulatory subunits of protein phosphatase 2A in mice. *J Mol Biol* 336: 971–986.
- Virshup DM, Shenolikar S (2009) From promiscuity to precision: protein phosphatases get a makeover. *Mol Cell* 33: 537–545.
- Xu Y, Chen Y, Zhang P, Jeffrey PD, Shi Y (2008) Structure of a protein phosphatase 2A holoenzyme: insights into B55-mediated Tau dephosphorylation. *Mol Cell* 31: 873–885.
- Kins S, Cramer A, Evans DR, Hemmings BA, Nitsch RM, et al. (2001) Reduced protein phosphatase 2A activity induces hyperphosphorylation and altered compartmentalization of tau in transgenic mice. *J Biol Chem* 276: 38193–38200.
- Vogelsberg-Ragaglia V, Schuck T, Trojanowski JQ, Lee VM (2001) PP2A mRNA expression is quantitatively decreased in Alzheimer's disease hippocampus. *Exp Neurol* 168: 402–412.
- Sontag E, Luangpirom A, Hladik C, Mudrak I, Ogris E, et al. (2004) Altered expression levels of the protein phosphatase 2A A $\beta$ Alphac enzyme are associated with Alzheimer disease pathology. *J Neuropathol Exp Neurol* 63: 287–301.
- Waekens E, Goris J, Merlevede W (1987) Purification and properties of polyclonal-stimulated phosphorylase phosphatases from rabbit skeletal muscle. *J Biol Chem* 262: 1049–1059.
- Cheng Q, Erickson AK, Wang ZX, Killilea SD (1996) Stimulation of phosphorylase phosphatase activity of protein phosphatase 2A1 by protamine is ionic strength dependent and involves interaction of protamine with both substrate and enzyme. *Biochemistry* 35: 15593–15600.
- Mochida S, Maslen SL, Skehel M, Hunt T (2010) Greatwall phosphorylates an inhibitor of protein phosphatase 2A that is essential for mitosis. *Science* 330: 1670–1673.
- Gharbi-Ayachi A, Labbe JC, Burgess A, Vigneron S, Strub JM, et al. (2010) The substrate of Greatwall kinase, Arpp19, controls mitosis by inhibiting protein phosphatase 2A. *Science* 330: 1673–1677.
- Zhou XZ, Kops O, Werner A, Lu PJ, Shen M, et al. (2000) Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins. *Mol Cell* 6: 873–883.
- Liou YC, Sun A, Ryo A, Zhou XZ, Yu ZX, et al. (2003) Role of the prolyl isomerase Pin1 in protecting against age-dependent neurodegeneration. *Nature* 424: 556–561.
- Bertrand J, Plouffe V, Senechal P, Leclerc N (2010) The pattern of human tau phosphorylation is the result of priming and feedback events in primary hippocampal neurons. *Neuroscience* 168: 323–334.
- Amniai L, Barbier P, Sillen A, Wieruszki JM, Peyrot V, et al. (2009) Alzheimer disease specific phosphoepitopes of Tau interfere with assembly of tubulin but not binding to microtubules. *FASEB J* 23: 1146–1152.
- Landrieu I, Leroy A, Smet-Nocca C, Huvent I, Amniai L, et al. (2010) NMR spectroscopy of the neuronal tau protein: normal function and implication in Alzheimer's disease. *Biochem Soc Trans* 38: 1006–1011.
- Landrieu I, Lacosse L, Leroy A, Wieruszki JM, Trivelli X, et al. (2006) NMR analysis of a Tau phosphorylation pattern. *J Am Chem Soc* 128: 3575–3583.
- Smet C, Leroy A, Sillen A, Wieruszki JM, Landrieu I, et al. (2004) Accepting its random coil nature allows a partial NMR assignment of the neuronal Tau protein. *Chembiochem* 5: 1639–1646.
- Sontag E, Numbhakdi-Craig V, Lee G, Brandt R, Kamibayashi C, et al. (1999) Molecular interactions among protein phosphatase 2A, tau, and microtubules. Implications for the regulation of tau phosphorylation and the development of tauopathies. *J Biol Chem* 274: 25490–25498.
- Lu PJ, Zhou XZ, Shen M, Lu KP (1999) Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science* 283: 1325–1328.
- Luna-Munoz J, Chavez-Macias L, Garcia-Sierra F, Mena R (2007) Earliest stages of tau conformational changes are related to the appearance of a sequence of specific phospho-dependent tau epitopes in Alzheimer's disease. *J Alzheimers Dis* 12: 365–375.
- Louis JV, Martens E, Borghgraef P, Lambrecht C, Sents W, et al. (2011) Mice lacking phosphatase PP2A subunit PR61/B' ( $\Delta$  Ppp2r5d) develop spatially restricted tauopathy by deregulation of CDK5 and GSK3 $\beta$ . *Proc Natl Acad Sci U S A* 108: 6957–6962.
- Grundke-Iqbal I, Iqbal K, Tung YC, Quinlan M, Wisniewski HM, et al. (1986) Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc Natl Acad Sci U S A* 83: 4913–4917.
- Gauthier A, Brandt R (2010) Live cell imaging of cytoskeletal dynamics in neurons using fluorescence photoactivation. *Biol Chem* 391: 639–643.
- Yamamoto H, Hasegawa M, Ono T, Tashima K, Ihara Y, et al. (1995) Dephosphorylation of fetal-tau and paired helical filaments-tau by protein phosphatases 1 and 2A and calcineurin. *J Biochem* 118: 1224–1231.
- Sontag E, Numbhakdi-Craig V, Lee G, Bloom GS, Mumby MC (1996) Regulation of the phosphorylation state and microtubule-binding activity of Tau by protein phosphatase 2A. *Neuron* 17: 1201–1207.
- Thomson M, Gunawardena J (2009) Unlimited multistability in multisite phosphorylation systems. *Nature* 460: 274–277.
- Lu PJ, Wulf G, Zhou XZ, Davies P, Lu KP (1999) The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* 399: 784–788.

## Acknowledgments

We thank Dr. Endicott and Dr. Morgan (Cambridge, UK) for the gift of the CDK2 and CycA3 expression clones.

## Author Contributions

Conceived and designed the experiments: IL GL VJ JG. Performed the experiments: IL GL CSN LA JMW. Analyzed the data: IL GL VJ. Contributed reagents/materials/analysis tools: JVL. Wrote the paper: IL GL VJ.

44. Sillen A, Barbier P, Landrieu I, Lefebvre S, Wieruszeski JM, et al. (2007) NMR investigation of the interaction between the neuronal protein tau and the microtubules. *Biochemistry* 46: 3055–3064.
45. Bodart JF, Wieruszeski JM, Ammiat L, Leroy A, Landrieu I, et al. (2008) NMR observation of Tau in *Xenopus* oocytes. *J Magn Reson* 192: 252–257.
46. Lippens G, Landrieu I, Hanouille X (2008) Studying posttranslational modifications by in-cell NMR. *Chem Biol* 15: 311–312.
47. Smet C, Sambo AV, Wieruszeski JM, Leroy A, Landrieu I, et al. (2004) The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau. *Biochemistry* 43: 2032–2040.
48. Welburn J, Endicott J (2005) Methods for preparation of proteins and protein complexes that regulate the eukaryotic cell cycle for structural studies. *Methods Mol Biol* 296: 219–235.

# Cell signaling, post-translational protein modifications and NMR spectroscopy

Francois-Xavier Theillet · Caroline Smet-Nocca · Stamatios Liokatis ·  
Rossukon Thongwichian · Jonas Kosten · Mi-Kyung Yoon ·  
Richard W. Kriwacki · Isabelle Landrieu · Guy Lippens · Philipp Selenko

Received: 28 June 2012 / Accepted: 7 September 2012 / Published online: 26 September 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** Post-translationally modified proteins make up the majority of the proteome and establish, to a large part, the impressive level of functional diversity in higher, multi-cellular organisms. Most eukaryotic post-translational protein modifications (PTMs) denote reversible, covalent additions of small chemical entities such as phosphate-, acyl-, alkyl- and glycosyl-groups onto selected subsets of modifiable amino acids. In turn, these modifications induce highly specific changes in the chemical environments of individual protein residues, which are readily detected by high-resolution NMR spectroscopy. In the following, we provide a concise compendium of NMR characteristics of the main types of eukaryotic PTMs: serine, threonine, tyrosine and histidine phosphorylation, lysine acetylation, lysine and arginine methylation, and serine, threonine O-glycosylation. We further delineate the previously uncharacterized NMR properties of lysine propionylation, butyrylation, succinylation, malonylation and crotonylation, which, altogether, define an initial

reference frame for comprehensive PTM studies by high-resolution NMR spectroscopy.

**Keywords** Histones · p53 · CBP/p300 · Sic1 · Tau · Integrin b3 · p21 KID · Pyk2 · HPr

## Introduction

Cellular signaling processes heavily rely on reversible post-translational protein modifications (PTMs) in their capacity to rapidly reprogram individual protein functions. PTMs are established and removed in a highly dynamic manner and exist in many different forms and flavors (Walsh et al. 2005). Along with alternative splicing, they provide the proteome with an enormous capacity for biological diversity and regulate virtually every aspect of cellular life, including cell–cell communication, cell growth and differentiation, sensing of metabolic states, mediating intracellular transport and initiating programmed cell death. Errors in PTM establishments and readouts, whether due to hereditary changes or environmental cues, constitute causal agents of many human diseases that include a long list of cancers, heart and brain diseases, diabetes and several metabolic disorders. Thus, the study of PTMs and how they regulate different cellular signaling processes has profound medical implications, both in the preventive and curative sense. PTM detection by high-resolution NMR spectroscopy represents a biophysical extension to studying these signaling marks from an analytical perspective, but also from a mechanistic, functional and structural point of view. The majority of PTMs is brought about by reversible, covalent additions of small, chemical entities, such as phosphate groups, acyl chains, alkyl chains, or various sugars, to the side-chains of

**Electronic supplementary material** The online version of this article (doi:10.1007/s10858-012-9674-x) contains supplementary material, which is available to authorized users.

F.-X. Theillet · S. Liokatis · R. Thongwichian · J. Kosten ·  
P. Selenko (✉)

Department of NMR-Supported Structural Biology, Leibniz  
Institute of Molecular Pharmacology (FMP Berlin), In-cell NMR  
Group, Robert-Roessle Strasse 10, 13125 Berlin, Germany  
e-mail: selenko@fmp-berlin.de

C. Smet-Nocca · I. Landrieu · G. Lippens  
CNRS UMR 8576, Universite Lille Nord de France,  
59655 Villeneuve d'Ascq, France

M.-K. Yoon · R. W. Kriwacki  
Department of Structural Biology, St. Jude Children's Research  
Hospital, Memphis, TN, USA



individual protein residues (Khoury et al. 2011). Others involve the addition of protein modules such as ubiquitin, SUMO, or NEDD to selected target sites. In this article, we describe the NMR characteristics of common types of eukaryotic PTMs that belong to the first class of protein modifications, namely phosphorylation, acylation, alkylation and glycosylation. These PTMs typically occur in ‘regulatory’ protein regions that are intrinsically disordered, including also protein loop regions (Iakoucheva et al. 2004; Xie et al. 2007; Radivojac et al. 2007; Gao and Xu 2012). This, because fast, cellular signaling responses usually require modifying enzymes to rapidly access individual protein PTM sites, which is easier achieved when modifiable amino acids are solvent exposed (i.e. not part of the hydrophobic protein core) and located in parts of the protein that are devoid of regular secondary, or tertiary structure. A substrate’s primary amino acid sequence encodes the specificity determinants for the modifying enzymes and for the protein modules that eventually recognize the different PTMs (Seet et al. 2006). Besides PTM-induced, *functional* modulations in protein–protein interactions (i.e. establishment of new interactions, breaking of existing interactions), PTMs can also mediate a range of *structural* responses that, in turn, differentially regulate functional, biological outcomes (Dyson and Wright 2005) (see below for selected examples).

Some protein residues lend themselves to different forms of modifications at single atom positions, such as lysines for example, which may undergo mono-, di- or trimethylation at the N $\zeta$  position (also referred to as the lysine  $\epsilon$ -amino site), or reversible acetylation of the same site. Similarly, the hydroxyl groups of serines and threonines can be phosphorylated or glycosylated. Other amino acids undergo multiple modifications at different side chain positions. Arginines display ‘regio-specific’ modification patterns, which may be symmetric or asymmetric, as observed in dimethylation reactions for example. This variety of modification states within single amino acid side-chains further increases the scope for diversity and plasticity of protein functions. Combinations of different types of PTMs on the same protein also provide the basis for complex signaling mechanisms via ‘reversible combinatorial codes’ (Jenuwein and Allis 2001) and coupled PTM marks are often established in hierarchical fashions, whereby upstream ‘master switches’ lead to the activation of different downstream signaling cascades. Conversely, co-operative sets of PTMs are frequently laid down in close proximity and allow direct synergistic or antagonistic cross talk between adjacent modification marks (Latham and Dent 2007; Kruse and Gu 2009; Martin et al. 2011).

Protein modifications such as N-terminal acetylation, proline hydroxylation, proline *cis–trans* isomerization, cysteine disulfide bond formation, protein oxidation or

nitrosylation, as well as proteolytic processing will not be discussed at this point, although these PTMs constitute equally abundant and biologically important signaling marks that are well amenable to investigations by NMR spectroscopy.

## PTMs by NMR

Before the advent of recombinant protein expression technologies, selective isotope labeling and multidimensional, hetero-nuclear NMR methods, NMR studies of covalent protein modifications such as phosphorylation or acetylation were restricted to direct, natural abundance readouts of phosphorus, or carbon NMR signals. Protein phosphorylation for example, was assayed by monitoring discrete changes in ATP/ADP  $^{31}\text{P}$  resonances in enzymatic kinase reactions, with respect to increasing phospho-protein signals (Mak et al. 1978; James 1985; Matheis and Whitaker 1984). Similarly, lysine acetylation was observed by directly reacting proteins with ( $1'-^{13}\text{C}$ )-acetylsalicylic acid (Macdonald et al. 1999; Xu et al. 1999), while lysine methylation was chemically established via reactions with  $^{13}\text{C}$  formaldehyde (Ashfield et al. 2000; Macnaughtan et al. 2005; Abraham et al. 2009).

In this article, we restrict ourselves to PTM detection approaches by 2D hetero-nuclear correlation methods i.e.  $^1\text{H}-^{15}\text{N}$  and  $^1\text{H}-^{13}\text{C}$  NMR experiments and isotope-labeled protein samples. Experiments of that sort afford higher resolution insights into PTM reactions and provide residue-resolved, positional information about PTM target sites and about structural PTM consequences (see below). Because PTMs frequently occur in intrinsically disordered protein regions (IDRs), many of the NMR characteristics of protein PTMs described here are deduced from IDR examples. We have nevertheless included examples of PTMs in folded and partially folded protein substrates, whenever possible. We additionally discuss deviations in PTM NMR behaviors of folded proteins in the *Conclusions Section* of the manuscript. In addition, we would like to stress that PTM detection by NMR spectroscopy is subject to the same inherent limitations as all other high-resolution NMR applications. Increasing protein/PTM-substrate sizes inevitably lead to greater spectral complexities and unfavorable NMR relaxation behaviors. Residue-resolved PTM site mapping requires dual isotope labeling ( $^{13}\text{C}/^{15}\text{N}$ ), triple-resonance NMR experiments (3D/4D) and dedicated NMR backbone assignment routines. Nevertheless, NMR detection of PTMs offers several advantages over ‘classical’ analytical methods, which are outlined in the following paragraphs. In addition, qualitative assessments of whether a protein of interest contains PTMs, and what types of PTMs, can be obtained without residue-specific resonance assignments provided that NMR spectra of unmodified

reference states exist (discussed in the concluding remarks of the manuscript).

Covalent PTMs introduce local alterations in the chemical environments of individual protein residues that are readily detected as characteristic chemical shift changes of NMR-observable spin systems in 2D NMR correlation experiments. Because most of the abundant eukaryotic PTMs involve additions of small chemical entities that do not significantly alter the molecular weights of the respectively modified proteins, and are not subject to chemical exchange behavior, they do not compromise size-dependent NMR detection parameters. Knowledge about PTM NMR characteristics enables the correct identification of PTM type(s), as well as to map the corresponding PTM site(s), provided that resonance-specific assignments are available. Protein phosphorylation for example, typically leads to large downfield chemical shift changes of serine/threonine backbone amide resonances ( $^1\text{H}$ - $^{15}\text{N}$ ), while protein acetylation results in smaller upfield chemical shift displacements of lysine backbone amides (see below).

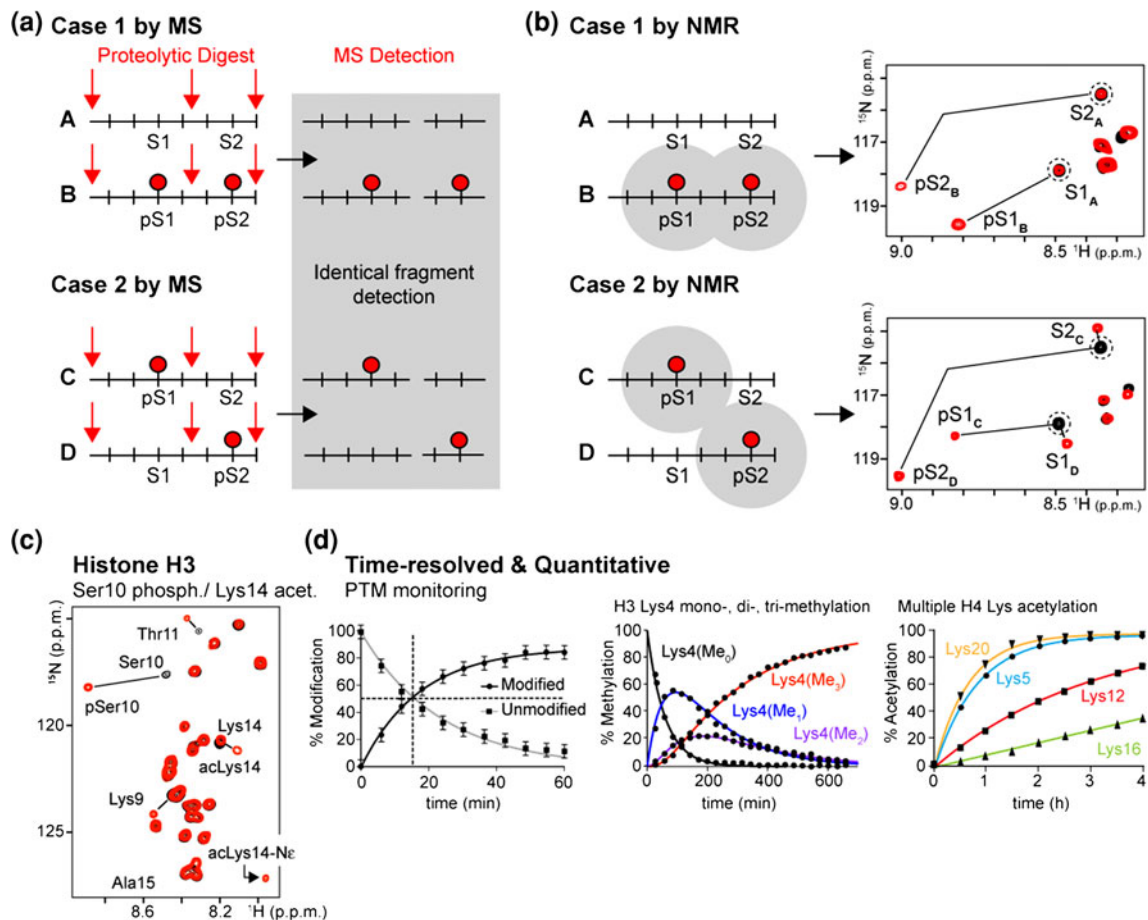
One important feature of PTM detection by NMR spectroscopy is the ability to delineate PTM distributions in proteins modified at multiple sites, provided that the different PTM marks are in close proximity. Site-specific mapping of adjacent protein PTMs is particularly challenging for most analytical methods, especially mass spectrometry (MS), which largely relies on proteolytic processing routines and peptide fragment-based PTM detection. As schematically illustrated in Fig. 1a, identical pairs of PTM/peptide fragments are generated from two different PTM distributions, which, conversely, cannot be distinguished by MS without elaborate identification processes of MS/MS fragmented peptides. In contrast, the corresponding NMR peak patterns unambiguously identify whether both PTMs are present on the same, or on different substrate molecules (Fig. 1b), provided that the presence of one PTM influences the chemical environment, and hence resonance frequency, of the respective other site. Although partial modifications of multiple PTM sites in the course of enzymatic modification reactions can complicate the resulting NMR spectra, their characteristics nevertheless resolve individual PTM distributions (Liokatis et al. 2012). As demonstrated for the doubly phosphorylated TEY fragment of the folded Erk kinase domain activation loop, PTM distributions originally derived from 2D NMR measurements were later confirmed by MS, however only after top-down and fragment-based MS approaches were combined (Prabakaran et al. 2011). Whenever multiple PTMs do not cluster in close proximity, PTM detection by NMR suffers from the same limitations in providing quantitative descriptions of PTM distributions as peptide-based MS approaches.

The non-disruptive nature of high-resolution NMR spectroscopy additionally offers convenient means for time-

resolved NMR measurements of reconstituted PTM reactions *in vitro*, but also of cellular modification events in complex environments such as cell extracts and whole live cells (Sakai et al. 2006; Lippens et al. 2008; Selenko et al. 2008). One such example is provided by the N-terminal ‘tail’ region of histone H3 that is post-translationally modified by endogenous enzymes in extracts of cultured human HeLa cells (Liokatis et al. 2010). 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation experiments revealed phosphorylation of Ser10 and acetylation of Lys14 (Fig. 1c). This example illustrates another advantage of PTM detection by NMR spectroscopy: the unique ability to monitor chemically distinct modification events in parallel (i.e. phosphorylation and acetylation) and without further requirements for selective enrichment or purification procedures, as would be required for MS analyses. The quantitative nature of NMR spectroscopy is another feature that makes it particularly appealing for PTM studies. Because changes in NMR signal intensities of modified and unmodified substrate residues are detected side-by-side, substrate/product concentrations, and time-dependent changes thereof, are readily deduced from simple NMR signal integration routines (Fig. 1d). From these, kinetic reaction parameters can directly be extracted (Dose et al. 2011; Landrieu et al. 2011). Such measurements are particularly useful in providing additional mechanistic insights into stepwise PTM reactions that require completions of certain PTM events, before others can ensue (Selenko et al. 2008; Theillet et al. 2012) (Fig. 1d).

While time-resolved NMR recordings can monitor the incorporation of various PTMs in a quantitative and residue-resolved fashion, PTM removal reactions can be studied equally well (Dose et al. 2011; Landrieu et al. 2011). In contrast to other methods, no changes in experimental setups, assay conditions, or readout parameters are required. Direct observations of reversible PTMs are thus fully compatible with the non-invasive and non-destructive nature of NMR spectroscopy.

A final benefit of multi-dimensional NMR methods for PTM detection is the ability to delineate newly established structural features that result as direct consequences of the respective modification events. Although increases in spectral complexity often result from such structural rearrangements, as the observed chemical shift changes no longer report the modified protein residues alone, PTM-triggered conformational alterations are readily detected by additional resonance peak displacements of ‘PTM-remote’ protein sites. On their own, such ‘long-range’ chemical shift changes do not reveal the details of newly established structural features. They nevertheless enable immediate qualitative assessments of conformational alterations that result as direct consequences of the individual PTM reactions. Protein phosphorylation in particular has long been known to provide the physicochemical basis for well-



**Fig. 1** Post-translational protein modifications by NMR. **a** Schematic outline of exemplary PTM distributions and mass spectrometry (MS) analysis via proteolytic peptide fragmentation and identification. *Note* that the different PTM distributions yield identical sets of peptide fragments. **b** Schematic outline of NMR analyses of the same set of PTMs and correct identification of the respective PTM distributions. NMR spectra reproduced in reference to Liokatis et al. (2012).

**c** Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of unmodified  $^{15}\text{N}$ -labeled histone H3 in vitro (aa1-33, *black*) and in HeLa cell extracts (*red*). Simultaneous NMR detection of H3 Ser10 phosphorylation and Lys14 acetylation by endogenous cellular enzymes. Reproduced from Liokatis et al. (2010). **d** Quantitative PTM monitoring by time-resolved NMR spectroscopy. Reproduced from Theillet et al. (2012) and Dose et al. (2011)

defined structural features (Johnson and Lewis 2001) that include modulations in  $\alpha$ -helix stability via N-cap formation, or C-terminal destabilization (Andrew et al. 2002). Indeed, several NMR studies of phosphorylation-induced conformational changes have been reported (Antz et al. 1999; Patchell et al. 2002; Kar et al. 2002; Bielska and Zondlo 2006; Perez et al. 2009; Tait et al. 2010; Nielsen and Schwalbe 2011; Sibille et al. 2011). A compelling, recent example is provided by phosphorylation of two tyrosine residues within the folded cytoplasmic integrin  $\beta 3$  domain, which result in pronounced structural rearrangements via phospho-tyrosine mediated hydrogen bonds and newly established electrostatic interactions (Deshmukh et al. 2011). Semi-synthetic approaches to introduce site-specific, homogeneous PTM states additionally offer new possibilities for studying long-range conformational response behaviors of modified proteins (Hejjaoui et al. 2012; Fauvet et al. 2012).

## Phosphorylation

### Serine, threonine phosphorylation

Modification of serine and threonine protein residues by reversible phosphorylation constitutes the most abundant PTM in eukaryotes (Cohen 2002b) (Fig. 2a). Phosphorylation is mediated by enzymes collectively referred to as protein kinases, whose own activities are often regulated via reversible phosphorylation (Cohen 2002a). All kinases exploit ATP as the universal phosphate donor. Removal of phosphates from modified protein residues is accomplished by sets of enzymes called phosphatases (Wurzenberger and Gerlich 2011). A number of protein domains specifically interact with phosphorylated serine and threonine residues and thereby enable the switch-like properties that these PTMs bring about (Seet et al. 2006). 14-3-3 domain containing proteins bind phosphorylated serines and threonines

(Gardino and Yaffe 2011). Some members of the WW domain protein family interact with modified serines and threonines followed by a proline (Wintjens et al. 2001; Salah et al. 2012). They are thereby in competition with other domains for the same motifs, such as CKS modules (Landrieu et al. 2001). Fork-head associated (FHA) protein domains selectively recognize phospho-threonines (Mahajan et al. 2008).

As previously mentioned, regulatory protein regions that harbor post-translational modification sites, and phosphorylatable serine/threonine residues in particular, are mostly solvent exposed and intrinsically disordered (Iakoucheva et al. 2004). As a consequence, their NMR characteristics are more readily affected by generic solution conditions such as pH and temperature. 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation experiments are particularly well suited to identify phosphorylated serines and threonines, as these residues experience prominent, modification-induced downfield backbone amide chemical shift changes ( $\Delta\delta \sim 0.5/1.5$  p.p.m.) (Figure 2b). These chemical shift changes are primarily caused by intra-residue hydrogen bonds between amide protons and the phosphate moieties, whenever PTM residues are in extended conformations (Du et al. 2005; Ramelot and Nicholson 2001). Phosphorylation of serines and threonines involved in pre-existing hydrogen bond networks, as it is often encountered in ‘structured’ protein loop regions, results in modulations of these characteristics (see later). Furthermore, the strong pH dependency of phosphorylated serine/threonine backbone amide resonances can directly be exploited to confirm phosphorylation (Bienkiewicz and Lumb 1999; Ramelot and Nicholson 2001; Prabakaran et al. 2011). At pH  $\sim 5$ , below the pKa of phospho-serines and phospho-threonines, phosphorylation-induced  $^1\text{H}$ - $^{15}\text{N}$  chemical shift changes are less pronounced due to the different protonation states of the phosphate group. In addition, high salt concentrations also shield the negative charge of the phosphate group even at pH values well above the respective pKa’s and similarly reduce phospho-serine/threonine chemical shift changes.

Serine/threonine phosphorylation at multiple protein sites can lead to considerable increases in spectral complexity, especially when the individual modification sites are closely spaced and incomplete substrate turnover is encountered. In fact, the large average number of protein serine/threonine residues that are phosphorylated by endogenous kinases in physiological environments such as cell extracts, often result in NMR spectra of that sort. One such example is provided by direct NMR detection of multiple phosphorylation events within the N-terminal, disordered transactivation domain (N-TAD) of human p53, executed by cellular enzymes in nuclear extracts from cultured HeLa cells (Fig. 2b). Other examples include NMR spectra of the multi-site phosphorylated, disordered C-terminus of PTEN, the human nucleolar protein hNIFK

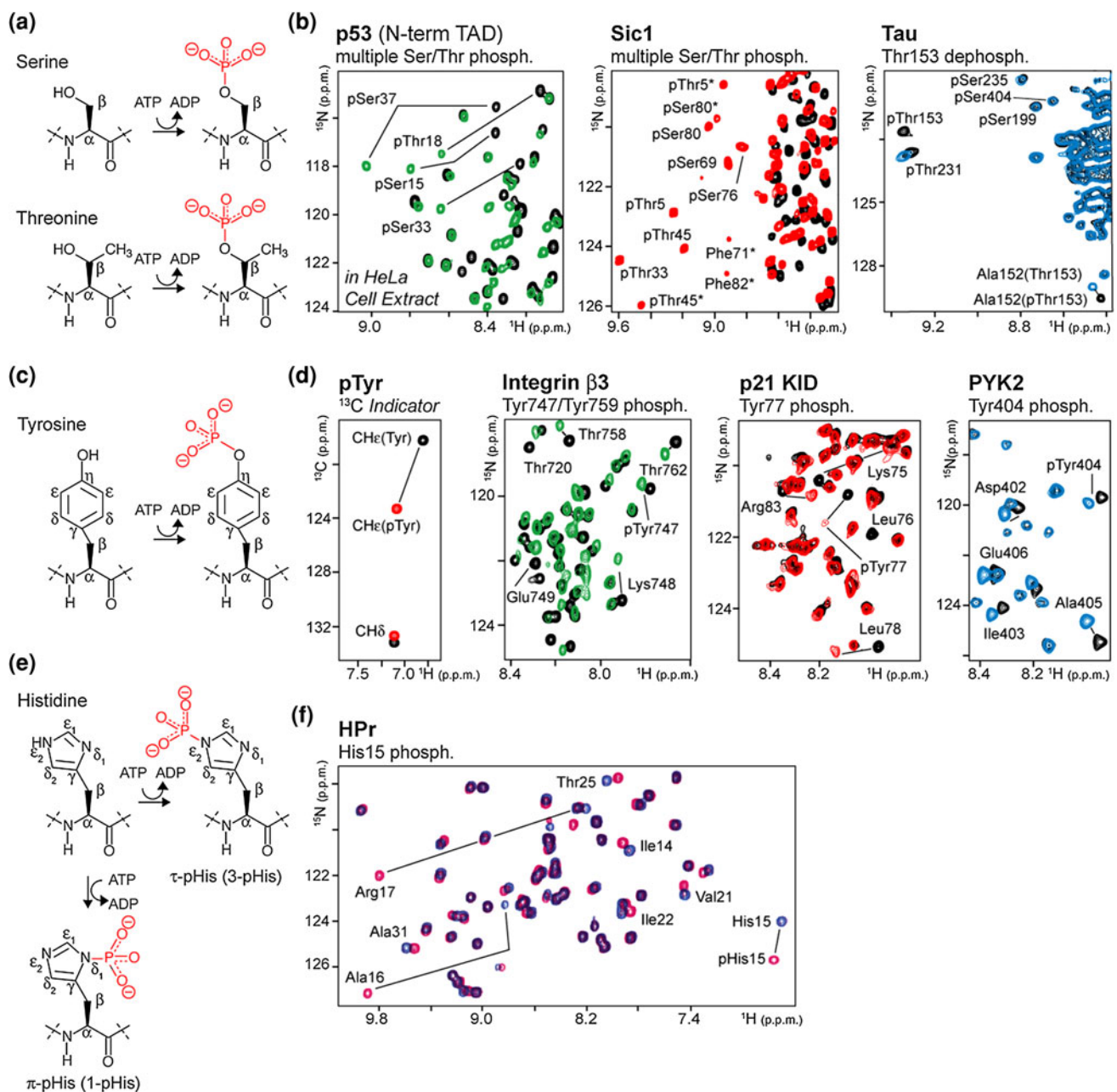
(Byeon et al. 2005), the disordered Tau protein (Landrieu et al. 2006; Leroy et al. 2010; Sibille et al. 2011), the histone H3 tail peptide (Liokatis et al. 2012), or the folded Erk protein kinase domain (Prabakaran et al. 2011). It should be stressed however, that such increases in spectral complexities often provide additional information with regard to different PTM distributions (Amniai et al. 2011; Prabakaran et al. 2011; Liokatis et al. 2012).

Many serine/threonine kinases are proline directed, which means that individual substrate sites are flanked by C-terminal proline residues (Songyang et al. 1996; Lu et al. 2002). The prolyl peptide-bond between serine/threonine and proline residues can exhibit *cis/trans* isomerization (Brown et al. 1999; Weiwad et al. 2000; Zhou et al. 2000; Werner-Allen et al. 2011) and phosphorylation often affects the thermodynamic properties of these isomers (Schutkowski et al. 1998). Moreover, sets of peptidyl-prolyl isomerases (PPIases) that accelerate *cis/trans* inter-conversion have been identified and changes in *cis/trans* equilibria provide additional levels of PTM regulation (Liou et al. 2011). With respect to the NMR chemical shift time scale, prolyl *cis/trans* isomerization is slow and therefore two NMR resonance signals are observed for nuclear spins that are in the proximities of the isomerizing peptidyl-prolyl bonds. This often leads to additional increases in spectral complexity (Andreotti 2003). An example is provided by the NMR study of multi-site, phosphorylation-specific interactions of the disordered cyclin dependent kinase (CDK) inhibitor Sic1 with its receptor Cdc4 (Mittag et al. 2008) (Fig. 2b). In this case, NMR signals of *cis*-Phe71, -Phe82 and of *cis* phospho-Thr5, -Thr45 and -Ser80 further complicate the spectral appearance of phosphorylated Sic1 (marked with asterisks in Fig. 2b). NMR observation of protein dephosphorylation i.e. detection of the ‘reverse’ PTM reaction, is equally well accomplished as illustrated by NMR mapping of site-selective phosphate removal from Thr153 of human Tau by the PP2A phosphatase (Landrieu et al. 2011) (Fig. 2b).

### Tyrosine phosphorylation

Tyrosine phosphorylation has emerged as a fundamentally important mechanism of signal transduction in eukaryotic cells that governs processes such as cell proliferation, cell cycle progression, metabolic homeostasis, transcriptional activation, neuronal transmission, differentiation, development and aging (Hunter 2009). Perturbations in tyrosine phosphorylation underlie many human diseases, in particular cancer, which has prompted the development of tyrosine kinase (TK) inhibitors as prominent drug targets (Kolch and Pitt 2010). Auto-phosphorylation of membrane-bound receptor tyrosine kinases (RTKs) (Lemmon





**Fig. 2** Serine, threonine, tyrosine and histidine phosphorylation. **a** Serine, threonine phosphorylation. **b** *Left panel*: Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of the unmodified (*black*) and HeLa cell extract-phosphorylated (*green*) N-terminal transactivation domain (TAD) of p53 (aa1-63). *Middle panel*: Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of unmodified (*black*) and phosphorylated Sic1 (*red*). Reproduced from Mittag et al. (2008). *Right panel*: Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of phosphorylated Tau (*black*) and upon dephosphorylation by PP2A (*blue*). Reproduced from Landrieu et al. (2011). **c** Tyrosine phosphorylation. **d** *Far left panel*: Superposition of 2D natural abundance  $^1\text{H}$ - $^{13}\text{C}$  NMR spectra (aromatic region) of a

peptide fragment of unmodified (*black*) and phosphorylated (*red*) adenine-specific DNA methyltransferase (aa134-144). *Left panel*: Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of the unmodified (*black*) and phosphorylated (*green*) Integrin  $\beta$ 3 domain. Reproduced from Deshmukh et al. (2011). *Middle panel*: Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of unmodified (*black*) and phosphorylated (*red*) p21 KID (Kriwacki laboratory, unpublished). *Right panel*: Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of unmodified (*black*) and phosphorylated (*blue*) PYK2. **e** Histidine phosphorylation. **d** Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of unmodified (*blue*) and phosphorylated (*pink*) HPr. Reproduced from Rajagopal et al. (1994)

and Schlessinger 2010) upon growth factor stimulation for example, triggers many ‘downstream’, intracellular signaling events that involve serine/threonine- and soluble,

nonreceptor tyrosine-kinases (NRTKs). Phospho-tyrosines are specifically recognized by members of the SH2- and PTB-domain protein families (Yaffe 2002), as well as a

range of protein tyrosine phosphatases (Julien et al. 2011). Importantly, evolution of the SH2 domain family in different organisms correlates with the divergence of downstream signaling networks and appears to recapitulate the complexities of the respective organisms themselves (Liu et al. 2011).

In contrast to serine/threonine phosphorylation, tyrosine phosphorylation does not induce similarly large, downfield backbone-amide chemical shift changes of the modified protein residues (Bienkiewicz and Lumb 1999), which is likely due to the more distal position of the phosphorylatable tyrosine hydroxyl group (Fig. 2c). Tyrosine phosphorylation does, however, lead to large chemical shift changes of aromatic CH $\epsilon$  resonances ( $\Delta\delta \sim 0.3/3$  p.p.m.,  $^1\text{H}$ - $^{13}\text{C}$ ) (Fig. 2d), which function as unambiguous *indicators* for the presence of phospho-tyrosines. Due to the limited chemical shift dispersion of solvent exposed protein tyrosine residues, phospho-site mapping via  $^1\text{H}$ - $^{13}\text{C}$  side-chain resonances, is not easily accomplished. Instead, once the presence of phospho-tyrosines has been confirmed by 2D  $^1\text{H}$ - $^{13}\text{C}$  experiments, their exact positions are mapped via ‘continuous’ backbone amide chemical shift changes of amino acids that surround the respectively modified tyrosine residues and that often display larger chemical shift displacements than the phosphorylated tyrosines themselves. Examples for phospho-tyrosine NMR studies are provided by the mono- and di-modified, folded Integrin  $\beta 3$  domain (Deshmukh et al. 2011) (Fig. 2d), phosphorylation of the folded cell cycle inhibitors p27 (Grimmler et al. 2007) and p21 (Fig. 2d, Kriwacki laboratory, unpublished results) and the disordered activation loop of PYK2 (Fig. 2d, Selenko laboratory, unpublished results). Thus, NMR detection of tyrosine phosphorylation and mapping of tyrosine phosphorylation sites by combinations of 2D  $^1\text{H}$ - $^{13}\text{C}$  and  $^1\text{H}$ - $^{15}\text{N}$  correlation experiments is rather straightforward.

### Histidine phosphorylation

Although histidine phosphorylation was thought to primarily occur in prokaryotic organisms and in plants, it is likely to play an equally important role in mammalian cells (Besant and Attwood 2005). A number of histidine-specific mammalian protein kinases and phosphatases have recently been identified (Attwood et al. 2010) and their particular roles in tissue homeostasis, regeneration and cellular proliferation are currently investigated. Histidines are phosphorylated at the  $\delta 1$ - (1-phospho-histidine) or  $\epsilon 2$ - (3-phospho-histidine) positions (Fig. 2e). Spontaneous histidine dephosphorylation occurs at low pH, and slow inter-conversion of  $\delta 1$ - into  $\epsilon 2$ -phospho-histidines takes place under mild basic conditions. This renders phospho-detection of modified histidine

residue a particularly challenging task for any method (Besant and Attwood 2010; Kee and Muir 2012).

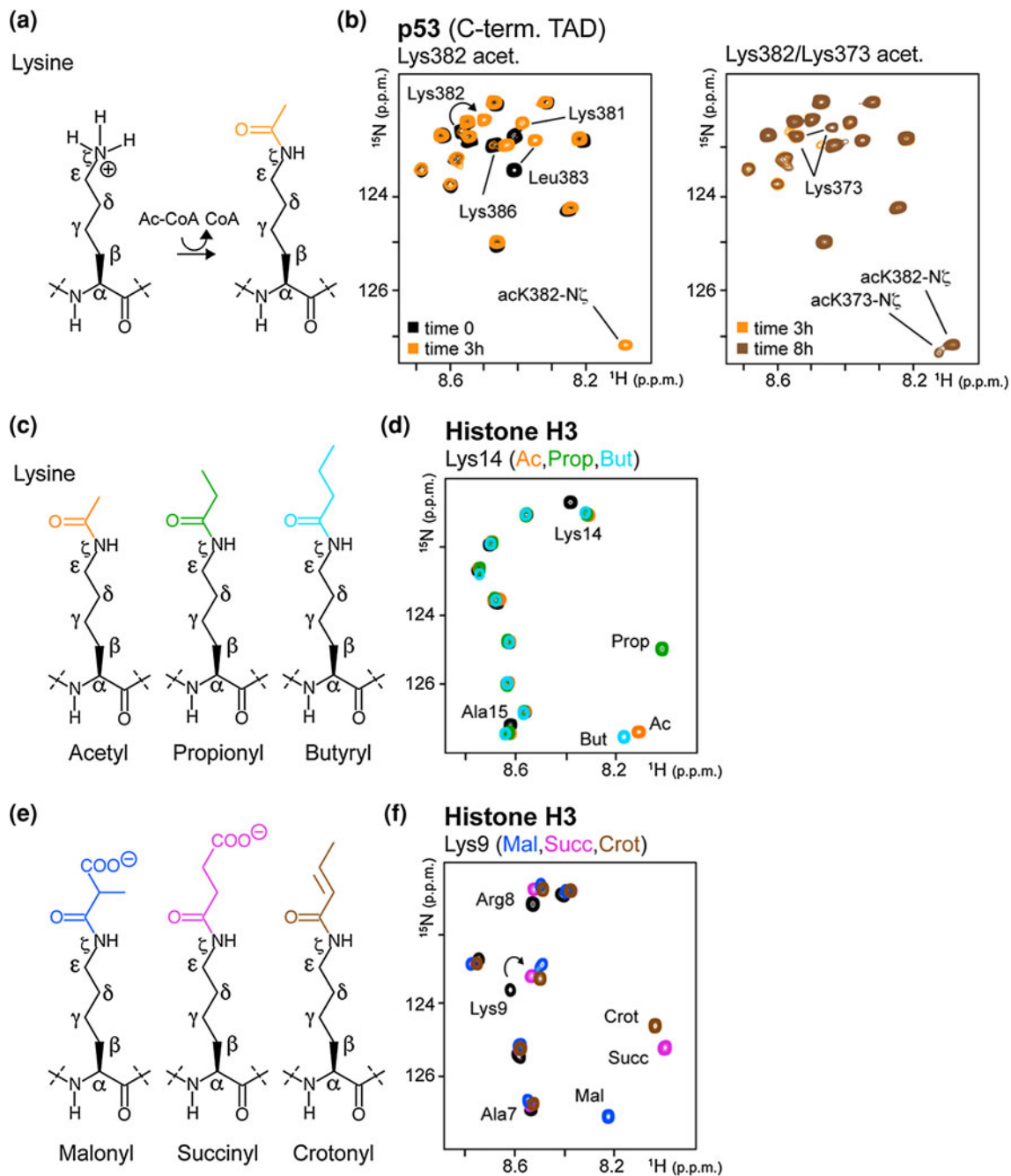
Interestingly, 2D phospho-histidine investigations by hetero-nuclear NMR methods have been reported as early as 1994 (Rajagopal et al. 1994). In their study, the Klevit laboratory generated stably His15-phosphorylated, folded HPr by means of continuous enzymatic regeneration, which counteracted hydrolysis of the modified histidine residue. Phospho-His15 induced minor backbone amide chemical shift changes of the majority of HPr resonances, while the phosphorylated amino acid and neighboring Ala16 and Arg17 experienced large downfield chemical shift changes in the proton and nitrogen dimensions (Fig. 2f). A localized structural rearrangement that was governed by the dianionic phosphoryl-group of His15, which acted as a novel hydrogen bond acceptor for the backbone amide protons of Ala16 and Arg17 and stabilized an  $\alpha$ -helical N-cap position was later shown to be the cause for this behavior (Jones et al. 1997). Similarly, pronounced backbone-amide chemical shift changes were also observed in other phospho-histidine NMR studies despite the absence of phosphorylation-induced structural changes (van Nuland et al. 1995; Garrett et al. 1998; Suh et al. 2008). More recently, HNP-type NMR experiments, based on phospho-histidine  $^1J(^{15}\text{N}/^{31}\text{P})$  coupling constants have been reported for the ‘stereo-specific’ NMR assignment of phosphorylated histidine residues (Himmel et al. 2010).

### Acylation

#### Lysine acetylation

Acetylation of lysine residues constitutes another abundant post-translational protein modification in the eukaryotic proteome (Norris et al. 2009). Differential acetylation of lysine residues in histone proteins establishes, in part, the epigenetic ‘histone code’, which ultimately determines the transcriptional states of entire genomes (Kouzarides 2007). Comprehensive annotation studies have additionally identified over 1,700 acetylated proteins in the human proteome, with functions in a great variety of cellular processes (Choudhary et al. 2009). Acetylation denotes the chemical conversion of the primary  $\text{N}\zeta\text{H}_3^+$ -amino group of lysine side-chains into NH-amide/acetyl moieties (Fig. 3a). Cellular enzymes that catalyze such reactions are collectively referred to as histone acetyltransferases (HATs), all of which employ acetyl-CoA as the ubiquitous acetyl-group donor (Berndsen and Denu 2008). Deacetylation is accomplished by histone deacetylases (HDACs) (Haberland et al. 2009) and both types of enzymes constitute prominent drug targets (Yang and Seto 2007). Acetylated lysine residues are





**Fig. 3** Lysine acylation. **a** Lysine acetylation. **b** Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of unmodified (black) and stepwise acetylated (orange and brown) C-terminal TAD of p53 (aa360-393). **c** Lysine acetylation, propionylation and butyrylation. **d** Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of synthetic unmodified (black) and Lys14 acetylated (orange), Lys14 propionylated (green) and Lys14

butyrylated (blue) histone H3 peptides (aa3-19). **e** Lysine malonylation, succinylation and crotonylation. **f** Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of synthetic unmodified (black) and Lys9 malonylated (blue), Lys9 succinylated (pink), Lys9 crotonylated (brown) histone H3 peptides (aa3-16)

specifically recognized by single-, or multi-copy bromodomain (BRD) containing proteins (Sanchez and Zhou 2009).

In 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra, lysine acetylation in intrinsically disordered protein regions typically results in small backbone amide chemical shift changes ( $\sim \Delta\delta$  0.06/

0.4 p.p.m.) of the respectively modified residues (Fig. 3b). In addition, every acetylation event produces a novel amide resonance signal, which corresponds to the newly established side-chain amide NH $\zeta$  group. For most acetylated lysines this side-chain NH $\zeta$  signal resonates at  $\sim 8.1/127.5$

p.p.m. ( $^1\text{H}$ – $^{15}\text{N}$ ) and therefore constitutes the generic acetylation *indicator*, whereas NMR mapping of acetylation sites relies on chemical shift difference readouts of backbone amide resonances, which serve as specific acetylation site *identifiers* (Liokatis et al. 2010; Smet-Nocca et al. 2010). One example of a dual acetylation reaction in which the two *indicator* signals do not superimpose is the stepwise acetylation of Lys382 and Lys373 of the disordered C-terminal transactivation domain (C-TAD) of human p53 by CBP/p300 shown in Fig. 3b (Selenko laboratory, unpublished results).

#### Lysine propionylation and butyrylation

Propionylated and butyrylated lysine residues (Fig. 3c) were first identified in histone proteins, the transcription factors p53 and CBP/p300 and in the propionyl-CoA synthetase (Chen et al. 2007; Garrity et al. 2007; Zhang et al. 2009; Cheng et al. 2009; Liu et al. 2009). Lysine acetyltransferases such as CBP/p300 and P/CAF were shown to also function as propionyl- and butyryl-transferases (Chen et al. 2007; Cheng et al. 2009; Liu et al. 2009; Leemhuis et al. 2008), while deacetylases SIRT1, SIRT2 and HDAC8 perform the respective de-propionylation and -butyrylation reactions (Riester et al. 2004; Smith and Denu 2007a, b; Cheng et al. 2009; Liu et al. 2009; Bheda et al. 2011). In vitro-, and probably also in vivo-, lysine propionylation and butyrylation are thought to occur via propionyl- and butyryl-CoA metabolites, which are naturally present at high abundance. Although propionylated and butyrylated lysine residues are recognized by acetyl-lysine binding bromodomains (Vollmuth and Geyer 2011), it is not known whether they signal particular biological activities, or whether they merely represent side-products of spontaneous modification reactions by acetyltransferases and propionyl- or butyryl-CoA (Lin et al. 2012).

The NMR characteristics of propionylated and butyrylated lysine residues in intrinsically disordered protein regions are similar to those of acetylated lysines (Fig. 3d). While their  $\text{CH}_2\epsilon$  resonances are indistinguishable from acetylated lysines ( $\sim 3.2/41.5$  p.p.m.,  $^1\text{H}$ – $^{13}\text{C}$ ), their side-chain  $\text{NH}\zeta$  *indicator* signals are well dispersed and unambiguously identify the respective modification types. As shown for butyrylated Lys14 of histone H3, the  $\text{NH}\zeta$  signal ( $^1\text{H}$ – $^{15}\text{N}$ ) is detected at  $\sim 8.15/127.5$  p.p.m., while the propionylated form of this lysine residue resonates at  $\sim 8.0/125.0$  p.p.m.. Backbone amide NMR signals of the respectively modified lysine residue experience similar chemical shift changes.

#### Lysine malonylation, succinylation and crotonylation

In prokaryotic and eukaryotic organisms, lysine residues are also subject to succinylation, malonylation and

crotonylation (Zhang et al. 2011; Peng et al. 2011; Du et al. 2011). These PTMs involve significant chemical and physical changes in the nature of lysine side-chains (Fig. 3e). Malonylation and succinylation are likely to be established via the transfer of a malonyl-, or a succinyl-group from malonyl- or succinyl-CoA, respectively, which are important metabolic intermediates. Cellular enzyme(s) that mediate lysine malonylation and succinylation have not yet been identified, although succinylation is abundant in mammalian proteins, especially in metabolic enzymes (Lin et al. 2012). Moreover, the Sirt5 protein, a *bona fide* member of the HDAC protein family with no known activity as a lysine deacetylase, has been shown to function as a nicotinamide-adenosine dinucleotide (NAD)-dependent lysine de-malonylase and de-succinylase (Peng et al. 2011; Du et al. 2011). Lysine crotonylation has recently been identified as an important histone modification that decorates transcription start sites in active chromatin (Tan et al. 2011). Crotonylating and decrotonylating enzymes remain unknown, while crotonyl-CoA is speculated to constitute the source for the transferred crotonyl group.

$\text{CH}_2\epsilon$  signals of succinylated, malonylated, crotonylated, or acetylated lysine residues in intrinsically disordered protein regions display similar resonances ( $\sim 3.2/41.5$  p.p.m.,  $^1\text{H}$ – $^{13}\text{C}$ ). In contrast, lysine succinylation, malonylation and crotonylation  $\text{NH}\zeta$  *indicator* signals are clearly different from acetylated lysines (Fig. 3f). Exemplified by differentially modified histone H3 Lys9, they resonate at  $\sim 8.0/125.0$  p.p.m.,  $\sim 8.2/127.0$  p.p.m. and  $\sim 8.0/124.5$  p.p.m. ( $^1\text{H}$ – $^{15}\text{N}$ ) respectively. Prominent backbone amide chemical shift changes of the modified residues and adjacent amino acids are additionally detected (Fig. 3f). This last characteristic offers means to easily identify the respective modification site(s), in a manner similar to acetylated lysine residues. NMR detection of acetylation, propionylation, butyrylation, succinylation, malonylation and crotonylation *indicator* cross-peaks in larger proteins may be hampered by signal overlap. However, upon inspection of NMR chemical shift entries of a randomly chosen set of 20 proteins from the Biological Magnetic Resonance Bank (BMRB), we did not detect substantial degrees of signal overlap in this region of the corresponding 2D  $^1\text{H}$ – $^{15}\text{N}$  NMR spectra (Suppl. Figure 1). In our hands, *indicator* cross-peaks of the aforementioned PTMs usually display larger NMR signal intensities than the corresponding lysine amide backbone resonances. This is probably due to comparable water/amide-proton exchange properties of the different types of amide groups, paired with favorable dynamic behaviors of side-chain amides.

## Alkylation

### Lysine methylation

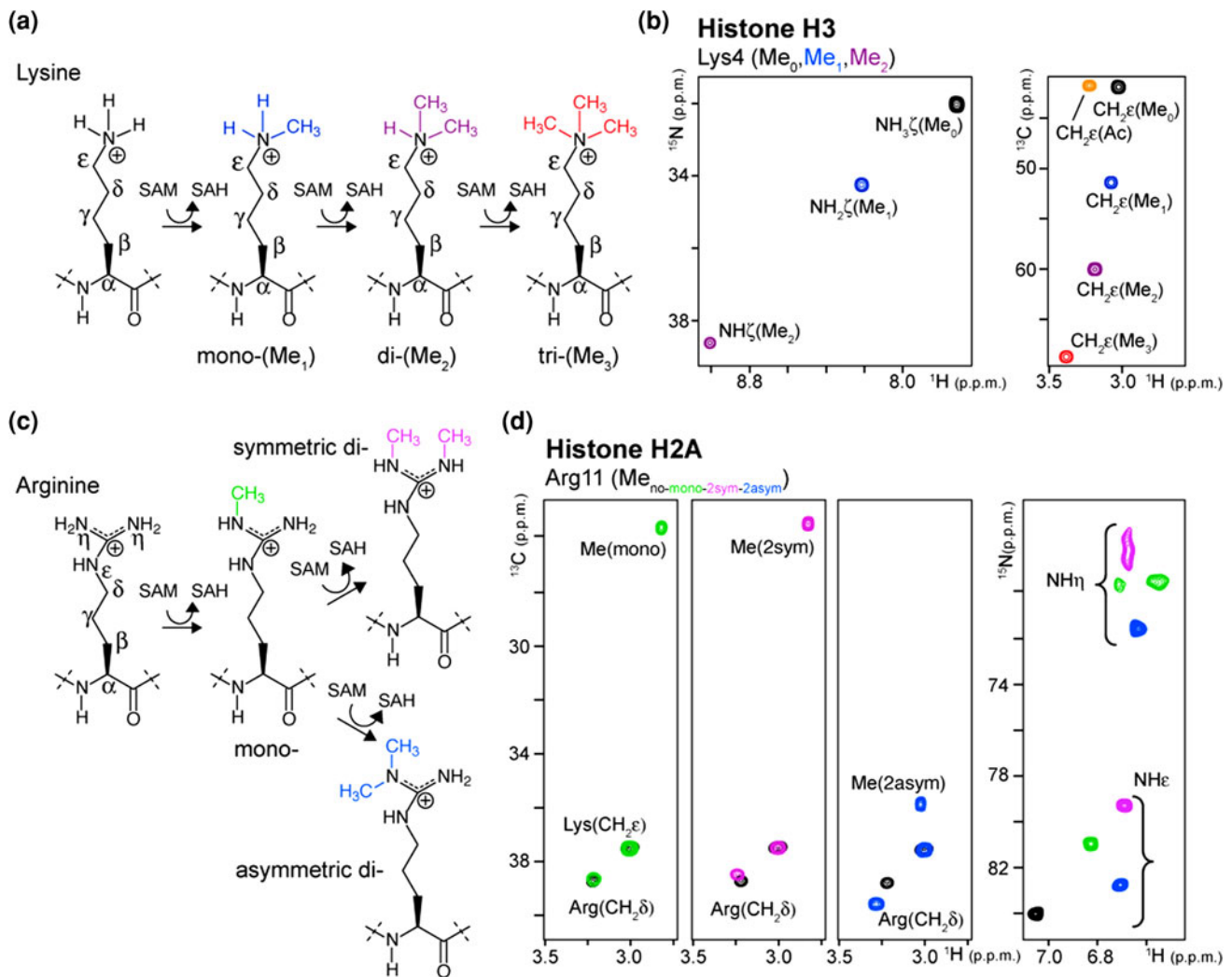
Protein lysine residues are subject to two of the most abundant but chemically distinct PTMs: lysine acetylation (as described above) and lysine methylation. Moreover, one and the same lysine residue may be acetylated, mono-, di-, or trimethylated (Fig. 4a), which produces an impressive range of complexity for switch-like, signaling functions. Indeed, lysine methylation plays pertinent roles in biological processes that include chromatin-mediated signaling (Latham and Dent 2007; Barth and Imhof 2010; Bannister and Kouzarides 2011; Suganuma and Workman 2011) and transcriptional regulation (Egorova et al. 2010; Stark et al. 2011; Lehnertz et al. 2011; Campaner et al. 2011). Conversely, lysine methylation has been linked to carcinogenesis and tumor malignancy (Stark et al. 2011; Fullgrabe et al. 2011; Varier and Timmers 2011), brain function and disorder (Gupta et al. 2010; Peter and Akbarian 2011; Graff et al. 2011), various metabolic pathways (Teperino et al. 2010) and cellular life span (Han and Brunet 2012). In addition, methylated lysines have been identified as major virulence factors and strong immunogens in the *Mycobacterium tuberculosis* heparin binding protein hemagglutinin (HBHA) (Pethe et al. 2002). Methylation is generally accomplished via methyl-transfer of the methylsulfonium moiety of *S*-adenosyl-methionine (SAM) onto the distal N $\zeta$  moieties of lysine side-chains, catalyzed by SET-domain containing enzymes referred to as lysine methyltransferases, or KMTs (Del Rizzo and Trievel 2011). Demethylation is mediated by lysine demethylases, or KDMs. Two classes of KDMs are known: those that contain LSD domains and employ FAD as a cofactor and those that bear Jumonji domains and rely on  $\alpha$ -ketoglutarate for demethylation (Heightman 2011). Because of their prevalent roles in disease relevant biological processes, both KMTs and KDMs constitute prominent drug targets (Kelly et al. 2010; Spannhoff et al. 2009; Copeland et al. 2009). Differentially methylated lysines are specifically recognized by plant homeo- (PHD) (Sanchez and Zhou 2011), chromo- (Yap and Zhou 2011) and MBT-domain containing proteins (Bonasio et al. 2010), which partially also define the ‘Royal Family’ of Tudor-like proteins (see also below) (Maurer-Stroh et al. 2003).

While lysine acetylation is manifested by the aforementioned chemical shift changes in 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra, lysine mono-, di-, and trimethylation in intrinsically disordered protein regions does not yield observable perturbations of backbone amide resonance signals (Theillet et al. 2012). Lysine mono-, and dimethylation does however produce characteristic side-chain  $^1\text{H}$ - $^{15}\text{N}$   $\zeta$  indicator resonances (Fig. 4b), but their fast

chemical exchange properties at physiological temperatures and pH make NMR detection impracticable. Instead, lysine methylation is best observed via 2D  $^1\text{H}$ - $^{13}\text{C}$  correlations, which display unique chemical shift changes of  $\text{CH}_2\epsilon$  side-chain resonances for the different methylation states (Fig. 4b). Specifically,  $\text{CH}_2\epsilon$  signals of unmodified lysines resonate at  $\sim 3.0/42.0$  p.p.m. ( $^1\text{H}$ - $^{13}\text{C}$ ), while mono-, di- and trimethylated lysines experience large downfield chemical shift changes ( $\sim \Delta\delta$  0.1/9.0 p.p.m.,  $^1\text{H}$ - $^{13}\text{C}$ ) and thus display characteristic resonance frequencies at  $\sim 3.1/51.0$  p.p.m.,  $\sim 3.2/60.0$  p.p.m. and  $\sim 3.4/68.5$  p.p.m. ( $^1\text{H}$ - $^{13}\text{C}$ ) respectively. The added methyl groups of mono-, di-, or trimethylated lysines are detected at  $\sim 2.7/35.5$  p.p.m.,  $\sim 2.9/45.5$  p.p.m. and  $\sim 3.1/55.5$  p.p.m. ( $^1\text{H}$ - $^{13}\text{C}$ ) (Theillet et al. 2012). Because most modifiable lysine residues in folded and intrinsically disordered proteins are solvent exposed, they sample similar chemical environments and the above NMR characteristics are generally preserved. Hence, lysine  $\text{CH}_2\epsilon$  resonances unambiguously determine whether a respective residue is methylated and, if so, in what form. Because NMR detection of lysine methylation via proton-carbon correlations does not involve exchangeable protons that could be subject to differential chemical exchange behaviors, time-resolved NMR measurements of methylation reactions can be performed in a broad range of in vitro conditions, or directly in complex environments such as cellular extracts (Theillet et al. 2012). The advantage of observing non-exchangeable  $^1\text{H}$ - $^{13}\text{C}$  correlations is offset by the chemical shift degeneracy of lysine side-chain resonances, which makes NMR mapping of lysine methylation sites difficult. Residue-selective isotope labeling and dedicated 2D ( $\text{HC}\epsilon(\text{Me}_x)$ -TOCSY- $\text{C}\alpha$ )NH pulse schemes that exploit selective methyl-lysine  $\text{C}\epsilon$  excitations and correlations to well-resolved lysine backbone amide ( $^1\text{H}$ - $^{15}\text{N}$ ) resonances, obliterate these problems (Theillet et al. 2012). Different to methylated lysine  $\text{CH}_2\epsilon$  signals,  $\text{CH}_2\epsilon$  resonances of acetylated lysines are detected at  $\sim 3.15/42.0$  p.p.m. (Fig. 4b) and can therefore be monitored simultaneously with methylated lysines (Theillet et al. 2012).

### Arginine methylation

Methylation of arginine residues is yet another abundant and biologically important PTM (Bedford and Richard 2005). Initially described in histone- and splicing-proteins, arginine methylation occurs in numerous other polypeptides that exert vital functions in signal transduction, transcription and translation (Bedford and Clarke 2009; Teyssier et al. 2010; Parry and Ward 2010; Erce et al. 2012). Arginine methylation involves the covalent addition of one, or two methyl groups to either one, or both distal guanidino N $\eta$  nitrogens of arginine side-chains (Fig. 4c). In contrast to



**Fig. 4** Lysine and arginine alkylation. **a** Lysine mono-, di-, and trimethylation. **b** *Left panel*: Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra of unmodified (black), Lys4 monomethylated (blue) and Lys4 dimethylated (purple) histone H3 peptides (aa1-15) at pH 4.5. *Right panel*: Superposition of 2D  $^1\text{H}$ - $^{13}\text{C}$  NMR spectra ( $\text{CH}_2\epsilon$  region, selective  $^{13}\text{C}$  lysine labeling) of unmodified (black) and Lys4 monomethylated (blue), Lys4 dimethylated (purple), Lys4 trimethylated (red) and Lys4 acetylated (orange) histone H3 peptides. Reproduced from Theillet et al. (2012). **c** Arginine mono- and

dimethylation (symmetric/asymmetric). **d** *Left panels*: Superpositions of 2D  $^1\text{H}$ - $^{13}\text{C}$  NMR spectra ( $\text{CH}_2\epsilon$  region) of synthetic unmodified (black) and Arg11 monomethylated (green), Arg11 symmetrically dimethylated (pink) and Arg11 asymmetrically dimethylated (blue) histone H2A peptides (aa4-11). *Right panel*: Superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  NMR spectra ( $\text{NH}\epsilon$  region) of synthetic unmodified (black) and Arg11 monomethylated (green), Arg11 symmetrically dimethylated (pink), Arg11 asymmetrically dimethylated (blue) histone H2A peptides (aa4-11) at pH 6.4 and 275 K

lysine acetylation, but similar to lysine methylation, arginine methylation preserves the overall positive charge of the residue. Arginine dimethylation exhibits stereo-specific chemical properties and occurs in either a symmetric (SDMA), or asymmetric (ADMA) form. Methylation of arginine residues is mediated by sets of enzymes called peptidylarginine methyltransferases, or PRMTs (Wolf 2009). Symmetrically dimethylating PRMTs are referred to as Class I enzymes. Asymmetric dimethylation is established by Class II enzymes. For both classes of PRMTs, monomethylation typically occurs as an intermediate step

*en route* to dimethylation. All PRMTs employ SAM as a cofactor and methyl-group donor. Demethylation is accomplished by peptidylarginine demethylases, or PRDMs (Smith and Denu 2009; Di Lorenzo and Bedford 2011). Dedicated methyl-arginine binding is primarily mediated by proteins of the Tudor domain family (Chen et al. 2011). Arginine methylation and aberrant PRMT and PRDM functions are implicated in a number of human diseases, including several forms of cancer, which has spurred interest in PRMTs and PRDMs as novel drug targets (Spannhoff et al. 2009; Lakowski et al. 2010; Luo 2012).



2D  $^1\text{H}$ - $^{13}\text{C}$  correlation spectra of non-methylated, mono- and symmetric di-methylated arginines in intrinsically disordered protein regions display minute differences in their  $\text{CH}_2\delta$  resonance signals, which superimpose at  $\sim 3.25/39.0$  p.p.m. ( $^1\text{H}$ - $^{13}\text{C}$ ), clearly offset from lysine  $\text{CH}_2\epsilon$  resonances (Fig. 4d). As shown for histone H2A Arg11 in comparison, asymmetric dimethylated arginines exhibit pronounced downfield chemical shift changes in the carbon dimension and resonate at  $\sim 3.25/40.0$  p.p.m. ( $^1\text{H}$ - $^{13}\text{C}$ ) (Fig. 4d). Methyl-group correlation signals of mono- and symmetric di-methylated arginines superimpose at  $\sim 2.75/26.0$  p.p.m. ( $^1\text{H}$ - $^{13}\text{C}$ ), well set apart from arginine  $\text{CH}_2\delta$  resonances. NMR signals of asymmetric dimethyl-groups are detected at a uniquely different resonance frequency at  $\sim 3.0/36.0$  p.p.m. ( $^1\text{H}$ - $^{13}\text{C}$ ). The poor chemical shift dispersion of non-, mono- and symmetric dimethylated arginine  $\text{CH}_2\delta$  signals limits the usefulness of  $^1\text{H}$ - $^{13}\text{C}$  correlation experiments in identifying these particular PTM states. Instead, 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra of methylated arginine residues reveal a great spread of their  $\text{NH}\epsilon$  resonance signals, depending on their individual modification states (Fig. 4d):  $\text{NH}\epsilon$  cross-peaks of non-methylated arginines are detected at  $\sim 7.05/84.0$  p.p.m. ( $^1\text{H}$ - $^{15}\text{N}$ ), while mono- ( $\sim 6.85/81.0$  p.p.m.), symmetric di- ( $\sim 6.65/79.0$  p.p.m.) and asymmetric dimethylated arginines ( $\sim 6.7/83.0$  p.p.m.) display uniquely different resonance frequencies. In addition, arginine  $^1\text{H}$ - $^{15}\text{N}$   $\text{NH}\eta$  signals exhibit characteristic chemical shift values in their differentially methylated forms and because most modifiable arginines in PRMT substrates are solvent exposed and sample similar chemical environments, these NMR characteristics are generally preserved. However, NMR detection of solvent accessible protein arginine  $\text{NH}\epsilon$  and  $\text{NH}\eta$  resonances is only feasible at a pH lower than 6.5, because of fast water/guanidinium proton chemical exchange (Liepinsh and Otting 1996). This precludes NMR measurements of enzymatic arginine methylation reactions under truly physiological conditions (i.e. at pH 7.0–7.5). Nevertheless, qualitative information about the presence of methylated arginine residues can be obtained at a pH below 6.5 and by low temperature NMR measurements as shown in Fig. 4, while residue-resolved NMR mapping of arginine methylation sites requires additional side-chain/backbone amide correlation experiments. As stated before, NMR detection of arginine methylation may become increasingly more difficult in proteins of larger sizes. However, methylated arginines are usually located in disordered protein regions (Gao and Xu 2012), which-, paired with enhanced side-chain dynamics-, offers additional advantages for low temperature detection routines that usually suffer from unfavorable increases in NMR correlation times in folded proteins.

## Glycosylation

Protein glycosylation refers to a large number of chemically distinct modifications that are overall classified based on the chemical nature of their protein–sugar linkages: N-glycosylation and O-glycosylation (Spiro 2002; Cummings 2009; Larkin and Imperiali 2011). Glycosyltransferase enzymes employ UDP-, GDP- or CMP-‘activated’ sugars as cofactors, from which they transfer the respective carbohydrate entities onto substrate proteins (Ohtsubo and Marth 2006). Glycosylation is abundant in viruses, prokaryotes, archaea and eukaryotes (Vigerust and Shepherd 2007; Eichler and Adams 2005; Bhat et al. 2011; Cummings 2009; Khoury et al. 2011; Hart and Copeland 2010), where it is involved in nutrient sensing, transcription, translation, signal transduction, organelle transport and cell–cell communication (Roth 2002; Hart et al. 2011; Marth and Grewal 2008). Conversely, these functions are often hijacked by pathogens for invasive mechanisms of cell entry (Varki 2008; Vigerust and Shepherd 2007). Glyco-mediated self, non-self recognition and pathological aberrations thereof are implicated in a number of human auto-immune diseases (Alavi and Axford 2008; Arnold et al. 2007) and speculated to be involved in diabetes and cancer (Slawson et al. 2010; Slawson and Hart 2011). O-GlcNAc glycans have additionally been shown to prevent aggregation-prone proteins from oligomerization and fibrillization (Yuzwa et al. 2012; Liang et al. 2006; Yu et al. 2008). Structural studies of glycosylated proteins generally require homogeneous glycans, which are difficult to produce especially when extended structures are desired. Recent advances in genetic engineering of bacterial and yeast glycosylation pathways for *in vivo* glycoprotein production have greatly improved this task (Rich and Withers 2009). Genetically engineered bacteria can also be employed to produce homogeneous glycans, which can then be linked via chemical, or enzymatic reactions to proteins of interest (Skrisovska et al. 2010). These strategies allow alternative isotope labeling schemes for protein and glycan moieties that permit isotope-filtered/edited NMR experiments (Slynko et al. 2009). In addition, dedicated protocols for the production of specifically isotope-labeled and glycosylated antibodies using hybridoma cell lines have been reported (Yamaguchi and Kato 2010).

Glycans typically display high internal mobility (DeMarco and Woods 2008), which hampers their characterization by X-ray crystallography (Wormald et al. 2002; Meyer and Moller 2007). In many instances, glycan sugar moieties retain their high degree of internal mobility when they are covalently attached to the respective protein targets. This, in turn, renders them amenable to high-resolution NMR studies, as exemplified by recent work on the glycosylated, 55 kDa Fc fragment of immunoglobulin

G (Barb et al. 2011). Thus, NMR constitutes the preferred tool for characterizing the structural and dynamic properties of sugar moieties in glyco-proteins (Fletcher et al. 1994; Wyss et al. 1995; Metzler et al. 1997; Erbel et al. 2000; Slynko et al. 2009; Barb and Prestegard 2011). On the protein side, NMR has also been used to investigate the conformational properties of glycosylated-, and neighboring protein residues. Specifically, preferred protein backbone conformations have been correlated with glycosidic peptide-sugar linkages of modified serine and threonine residues (Corzana et al. 2006b, 2007). However, due to the great variety in glycan residues and in glycosylation-induced changes of protein backbone conformations, it is not possible to pin down common glycan features that define the NMR characteristics of individual glycosylation events. Despite that, residue-resolved NMR measurements of glycan modification kinetics are easily accessible, because of the large chemical shift differences between free and polymerized carbohydrate entities (Barb et al. 2011).

#### N-glycosylation

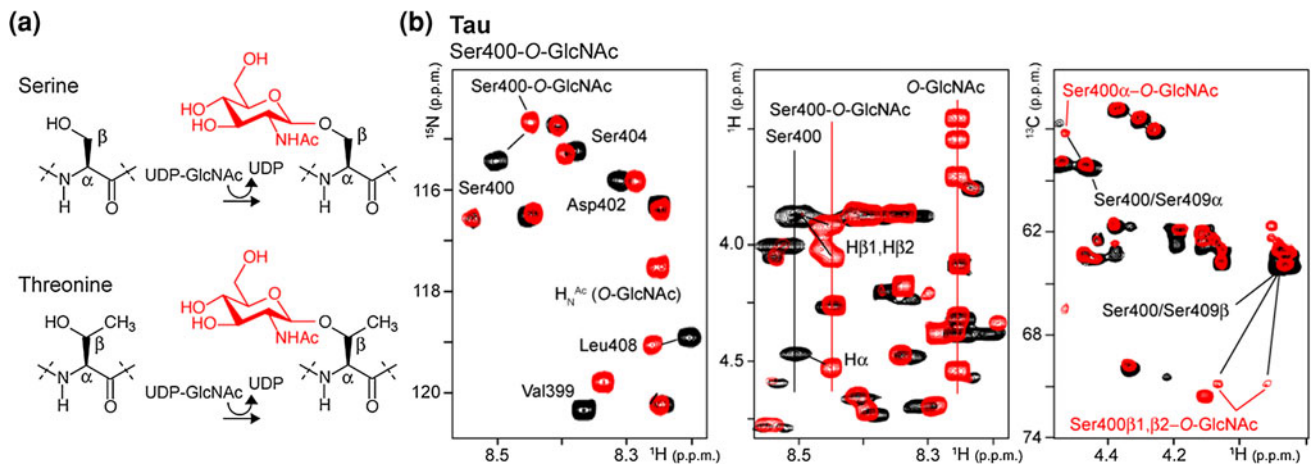
N-linked glycans contain  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$  as the basic building block, which is covalently added onto the N $\delta$  position of asparagine side-chains within the Asn-X-Ser/Thr consensus sequence (X must not be a proline) (Stanley et al. 2009). Starting from this primary structure, additional carbohydrate moieties (fucose, GalNAc, sialic acid, galactose) are progressively added, or, in turn, removed to yield the final N-glycan products. Cellular N-glycan maturation occurs in multiple, spatially separated reaction steps in the endoplasmic reticulum (ER) and Golgi apparatus. Once established, N-glycosylation itself is rather long lived, whereas individual glycan structures may experience dynamic compositional changes during a protein's lifetime (Stanley et al. 2009; Schwarz and Aebi 2011). With regard to glycosylation-induced changes in local protein backbone conformations, N-glycosylation has been reported to promote  $\beta$ -turn conformations (Meyer and Moller 2007), stabilize folded protein domains (Wyss et al. 1995) and increase the degree of order in human chorionic gonadotropin (Erbel et al. 2000). Immunoglobulin G N-glycans exhibit conformational exchange between two states, one giving rise to contacts between the glycan chain and the immunoglobulin, the other one preventing such contacts (Barb and Prestegard 2011; Barb et al. 2012). In contrast, a single, well-defined structure of the N-glycan chain was delineated for the adhesion domain of human CD2 and for a model glycoprotein from *Campylobacter jejuni* (Wyss et al. 1995; Slynko et al. 2009).

#### O-glycosylation

O-glycans are established via direct glycosylation of serine, threonine, or tyrosine side-chain hydroxyl groups and can be found in prokaryotes and eukaryotes (Brockhausen et al. 2009; Freeze and Haltiwanger 2009; Hart and Akimoto 2009). The structural and chemical diversity of O-glycans is much higher than for N-glycans, and many different sugars, such as *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), galactose, glucose, sialic acid, fucose or xylose are commonly incorporated. O-glycosylation has been shown to induce various protein conformational changes. Serine, threonine O-glycosylation with  $\alpha$ -D-GalNAc for example, was reported to decrease the  $\alpha$ -helical content of the modified peptide hormone calcitonin (Tagashira et al. 2002), or even elicit  $\beta$ -like, extended structures in glycopeptides, which, in turn, were detected by large, residue-specific chemical shift changes (Coltart et al. 2002; Hashimoto et al. 2011). In contrast, it has been observed that O-glycosylation with  $\beta$ -D-glucose increases  $\alpha$ -helicity (Corzana et al. 2006a), while O-glycosylation with  $\beta$ -D-GlcNAc was reported to induce turn-like structures in several glycopeptides (Simanek et al. 1998; Wu et al. 1999). These structures were further disrupted by alternative phosphorylation, as has been shown for the N-terminus of the murine estrogen receptor beta, for example (Chen et al. 2006). NMR was also used to decipher a glycosylation-dependent decrease in proline *cis* isomer content at positions C-terminal to modified serine, or threonine residues (Narimatsu et al. 2010). Finally, O-glycosylation by  $\alpha$ -D-Gal or  $\beta$ -D-Gal impacts *cis/trans* isomerization of (2S,4S)-4-hydroxyproline (Owens et al. 2009), but not of (2S,4R)-4-hydroxyproline (Owens et al. 2007), whereas glycosylation of poly-hydroxyprolines induces stable poly-proline type II helices (Owens et al. 2010). O-linked  $\beta$ -*N*-acetylglucosamination is highly dynamic and as abundant as protein phosphorylation, or acetylation (Hart and Akimoto 2009; Khoury et al. 2011) (Fig. 5a). In many instances, individual serine and threonine residues in eukaryotic proteins compete in phosphorylation/glycosylation reactions (Hart et al. 2007; Hart et al. 2011). Because no consensus sequences have yet been identified for O-linked *N*-acetylglucosamine transferase (OGT) enzymes, even MS detection of protein O-GlcNAc sites has proven difficult. Sophisticated enrichment routines for O-GlcNAc peptides via combined enzymatic and chemical reactions, in combination with soft ionization modes that preserve the labile O-GlcNAc groups on serines and threonines have to be employed (Wang et al. 2010).

$^{13}\text{C}\beta$  chemical shift values of O-GlcNAcylated or O-GalNAcylated serines and threonines ( $\sim 71.0$  and  $\sim 78.0$  p.p.m. respectively) give rise to well-separated resonance signals in 2D  $^1\text{H}$ - $^{13}\text{C}$  correlation experiments, which can





**Fig. 5** Serine, threonine glycosylation. **a** Serine, threonine O-GlcNAc modification. **b** *Left panel*: Superposition of 2D <sup>1</sup>H-<sup>15</sup>N NMR spectra of unmodified (*black*) and Ser400 O-GlcNAcylated (*red*) Tau (aa392-411). *Middle panel*: Superposition of 2D <sup>1</sup>H-<sup>15</sup>N

NMR spectra of unmodified (*black*) and Ser400 O-GlcNAcylated (*red*) Tau. *Right panel*: Superposition of 2D <sup>1</sup>H-<sup>13</sup>C NMR spectra of unmodified (*black*) and Ser400 O-GlcNAcylated (*red*) Tau. Reproduced from Dehennaut et al. (2008), Smet-Nocca et al. (2011)

serve as unique O-glycan *indicators* (Corzana et al. 2007; Smet-Nocca et al. 2011). Characteristic, anomeric O-Glycan <sup>1</sup>H<sub>1</sub>-<sup>13</sup>C<sub>1</sub> correlation signals at ~4.3-5.0/99.0–105.0 p.p.m. could further function as O-glycosylation *indicators*. Direct NMR identification of protein O-glycosylation sites by homonuclear <sup>1</sup>H-<sup>1</sup>H correlation experiments is not straightforward, because no cross peaks between GlcNAc protons and the modified serine or threonine residues are detected (Smet-Nocca et al. 2011; Dehennaut et al. 2008). In the case of O-GlcNAc modified Tau for example, 2D <sup>1</sup>H-<sup>15</sup>N correlation experiments enabled chemical shift difference mapping of the glycosylated protein region, but failed to identify the respective modification site(s), because of large chemical shift changes of two serine and one threonine residues, Ser400, Thr403 and Ser404, and of neighboring Val399, Gly401, Asp402 (Fig. 5b). NMR assignment of the modification site was achieved via a combination of 2D <sup>1</sup>H-<sup>15</sup>N HSQC, <sup>1</sup>HN-<sup>1</sup>H TOCSY and <sup>1</sup>H-<sup>13</sup>C HSQC experiments. Based on the large C $\alpha$  and C $\beta$  chemical shift changes of O-glycosylated protein residues ( $\Delta\delta \sim 2.0$  p.p.m. and  $\sim 6.0$  p.p.m., respectively), corresponding H $\alpha$  and H $\beta$  chemical shifts were extracted from 2D <sup>1</sup>H-<sup>13</sup>C spectra and correlated to HN resonances via HN-H $\alpha/\beta$  signals from 2D TOCSY NMR spectra (Fig. 5b). Thereby, O-GlcNAc modification of Tau Ser400 was confirmed (Dehennaut et al. 2008; Smet-Nocca et al. 2011).

## Conclusions

The growing demand for quantitative methods to annotate cellular signaling states on systems levels has been met by the development of analytical tools that enable direct

observations of cellular PTMs in unperturbed environments. Mapping of protein PTM sites, as well as in situ deductions of mechanistic properties of cellular PTM reactions -directly obtainable from such analyses-, are critically required to further our understanding about how these processes are modulated under different health and disease conditions. With this article, we hope to have conveyed strong arguments in favor of high-resolution NMR spectroscopy as highly useful in providing such information.

What would be the requirements for an *ideal* analytical tool in eukaryotic PTM research? Above all, it ought to be able to generically and qualitatively report whether a protein of interest is post-translationally modified and if so, to identify what kind of PTMs are present and at which residue positions. In most instances, it will be important to address such questions in cellular contexts and without defined information about the nature of the modifying enzymes and their respective activities. Here, direct NMR measurements of isotope-labeled proteins in different cell extracts can provide valuable first insights. While *in extract* NMR approaches may only be feasible for reasonably-sized (<20 kDa) proteins, and for one isotope-labeled protein at a time, in most instances simple 2D NMR correlation (<sup>1</sup>H-<sup>15</sup>N and <sup>1</sup>H-<sup>13</sup>C) experiments may prove sufficient to qualitatively identify which types of PTMs are present, even without the necessity for NMR resonance assignments. As we have outlined throughout the text, most of the predominant eukaryotic PTMs display characteristic NMR *indicator* properties that make their identification straightforward. Serine/threonine and histidine phosphorylation in intrinsically disordered protein regions results in large downfield chemical shift changes of backbone amide

resonances for example, which are easily discernable in 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra. Tyrosine phosphorylation is less pronounced by 2D  $^1\text{H}$ - $^{15}\text{N}$  measures, but phosphotyrosines display unique *indicator* properties in the aromatic region of 2D  $^1\text{H}$ - $^{13}\text{C}$  NMR correlations. Similarly, different acylation events (i.e. acetylation, malonylation, succinylation, crotonylation, propionylation and butyrylation) produce unique  $\text{HN}\zeta$  *indicator* signals in 2D  $^1\text{H}$ - $^{15}\text{N}$  experiments, which can simultaneously be detected with most phosphorylation modifications. Characteristic lysine  $\text{CH}_2\epsilon$  resonances in 2D  $^1\text{H}$ - $^{13}\text{C}$  NMR correlations unambiguously function as *indicators* for mono-, di- and trimethylation and transferring the protein mixture to a low pH environment enables NMR recordings of unique *indicator* resonances of different arginine methylation states by 2D  $^1\text{H}$ - $^{15}\text{N}$  experiments. At the same time, sets of 2D  $^1\text{H}$ - $^{15}\text{N}$  and  $^1\text{H}$ - $^{13}\text{C}$  correlation experiments provide qualitative *indications* for possible glycosylation events at serine/threonine positions. Thereby, combinations of ‘simple’ 2D correlation experiments ( $^1\text{H}$ - $^{15}\text{N}$  and  $^1\text{H}$ - $^{13}\text{C}$ ) can be used to identify the most common types of eukaryotic PTMs.

While we have focused our article on NMR characteristics of eukaryotic PTMs in disordered, regulatory protein regions, we wish to explicitly stress that PTMs within folded proteins, or protein domains, may exhibit deviations from the canonical NMR properties described above (Lippens and Selenko laboratories, unpublished observations). Especially for cases in which post-translationally modified amino acids are involved in hydrogen bond networks, PTM-induced NMR behaviors can differ substantially from disordered, solvent exposed PTM sites. In addition, whenever more global backbone amide chemical shifts changes are observed, NMR mapping of individual PTM sites may require more dedicated in vitro experimental setups and additional 3D NMR experiments. These potential drawbacks are contrasted by the unique ability of NMR spectroscopy to provide time-resolved, quantitative information about individual modification levels (i.e. ratios of modified versus unmodified substrate sites and molecules), as well as about individual PTM distributions in the case of closely spaced modification sites. Time-resolved NMR spectroscopy does provide high-resolution insights into hierarchical properties of processive PTM events at multiple protein sites, which ideally complements PTM data from proteome-wide MS studies. Combined with direct NMR readouts in complex environments such as cell extracts and intact cells, it offers the advantage to quickly and comparatively analyze PTMs under different in vitro and in vivo conditions. The scope of PTM induced structural rearrangements, which is not easily accessible with classical in vitro methods in structural biology and particularly important for intrinsically disordered, regulatory

protein regions, provides yet another area for unique NMR input. For these reasons, we believe that NMR spectroscopy will become an increasingly important tool for deciphering the full biological range of signaling-mediated, cellular processes. Here, we have provided an initial NMR reference frame for the most abundant eukaryotic post-translation protein modifications. Future studies will likely reveal an even greater chemical repertoire of cellular PTMs, but given the inherent physical nature of high-resolution NMR spectroscopy and its unique ability to report changes in the chemical environments of individual atomic nuclei, it is well poised to face these challenges with ease.

**Acknowledgments** We would like to thank Rachel Klevit, Olga Vinogradova, Tanja Mittag and Julie Forman-Kay for providing original NMR spectra for reproduction in this manuscript. F.X.T. acknowledges support from the Association pour la Recherche contre le Cancer (ARC). P.S. acknowledges funding by an Emmy Noether research grant (SE1794/1-1) from the Deutsche Forschungsgemeinschaft (DFG). R. W. K. acknowledges support from NIH core grant P30CA21765 (to St. Jude Children’s Research Hospital) and 5R01CA082491 (to R. W. K.), and the American Lebanese Syrian Associated Charities (ALSAC) of St. Jude Children’s Research Hospital. We further express our gratitude to Angela Gronenborn and Georges Mer for expert advice and stimulating discussions in the course of writing the paper.

## References

- Abraham SJ, Kobayashi T, Solaro RJ, Gaponenko V (2009) Differences in lysine pKa values may be used to improve NMR signal dispersion in reductively methylated proteins. *J Biomol NMR* 43(4):239–246
- Alavi A, Axford JS (2008) Sweet and sour: the impact of sugars on disease. *Rheumatology (Oxf)* 47(6):760–770
- Amniai L, Lippens G, Landrieu I (2011) Characterization of the AT180 epitope of phosphorylated Tau protein by a combined nuclear magnetic resonance and fluorescence spectroscopy approach. *Biochem Biophys Res Commun* 412(4):743–746
- Andreotti AH (2003) Native state proline isomerization: an intrinsic molecular switch. *Biochemistry* 42(32):9515–9524
- Andrew CD, Warwicker J, Jones GR, Doig AJ (2002) Effect of phosphorylation on alpha-helix stability as a function of position. *Biochemistry* 41(6):1897–1905
- Antz C, Bauer T, Kalbacher H, Frank R, Covarrubias M, Kalbitzer HR, Ruppertsberg JP, Baukrowitz T, Fakler B (1999) Control of K<sup>+</sup> channel gating by protein phosphorylation: structural switches of the inactivation gate. *Nat Struct Biol* 6(2):146–150
- Arnold JN, Wormald MR, Sim RB, Rudd PM, Dwek RA (2007) The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* 25:21–50
- Ashfield JT, Meyers T, Lowne D, Varley PG, Arnold JR, Tan P, Yang JC, Czaplowski LG, Dudgeon T, Fisher J (2000) Chemical modification of a variant of human MIP-1alpha; implications for dimer structure. *Protein Sci* 9(10):2047–2053
- Attwood PV, Ludwig K, Bergander K, Besant PG, Adina-Zada A, Kriegelstein J, Klumpp S (2010) Chemical phosphorylation of histidine-containing peptides based on the sequence of histone H4 and their dephosphorylation by protein histidine phosphatase. *Biochim Biophys Acta* 1804(1):199–205

- Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. *Cell Res* 21(3):381–395
- Barb AW, Prestegard JH (2011) NMR analysis demonstrates immunoglobulin G *N*-glycans are accessible and dynamic. *Nat Chem Biol* 7(3):147–153
- Barb AW, Freedberg DI, Battistel MD, Prestegard JH (2011) NMR detection and characterization of sialylated glycoproteins and cell surface polysaccharides. *J Biomol NMR* 51(1–2):163–171
- Barb AW, Meng L, Gao Z, Johnson RW, Moremen KW, Prestegard JH (2012) NMR Characterization of immunoglobulin G Fc glycan motion on enzymatic sialylation. *Biochemistry* 46:18–4626
- Barth TK, Imhof A (2010) Fast signals and slow marks: the dynamics of histone modifications. *Trends Biochem Sci* 35(11):618–626
- Bedford MT, Clarke SG (2009) Protein arginine methylation in mammals: who, what, and why. *Mol Cell* 33(1):1–13
- Bedford MT, Richard S (2005) Arginine methylation an emerging regulator of protein function. *Mol Cell* 18(3):263–272
- Berndsen CE, Denu JM (2008) Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr Opin Struct Biol* 18(6):682–689
- Besant PG, Attwood PV (2005) Mammalian histidine kinases. *Biochim Biophys Acta* 1754(1–2):281–290
- Besant PG, Attwood PV (2010) Histidine phosphorylation in histones and in other mammalian proteins. *Methods Enzymol* 471:403–426
- Bhat AH, Mondal H, Chauhan JS, Raghava GP, Methi A, Rao A (2011) ProGlycProt: a repository of experimentally characterized prokaryotic glycoproteins. *Nucleic Acids Res* 40(Database issue) D388–D393
- Bheda P, Wang JT, Escalante-Semerena JC, Wolberger C (2011) Structure of Sir2Tm bound to a propionylated peptide. *Protein Sci* 20(1):131–139
- Bielska AA, Zondlo NJ (2006) Hyperphosphorylation of Tau induces local polyproline II helix. *Biochemistry* 45(17):5527–5537
- Bienkiewicz EA, Lumb KJ (1999) Random-coil chemical shifts of phosphorylated amino acids. *J Biomol NMR* 15(3):203–206
- Bonasio R, Lecona E, Reinberg D (2010) MBT domain proteins in development and disease. *Semin Cell Dev Biol* 21(2):221–230
- Brockhausen I, Schachter H, Stanley P (2009) *O*-GalNAc glycans. In: Varki A, Cummings RD, Esko JD et al (eds) *Essentials of glycobiology*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Brown NR, Noble ME, Endicott JA, Johnson LN (1999) The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat Cell Biol* 1(7):438–443
- Byeon IJ, Li H, Song H, Gronenborn AM, Tsai MD (2005) Sequential phosphorylation and multisite interactions characterize specific target recognition by the FHA domain of Ki67. *Nat Struct Mol Biol* 12(11):987–993
- Campaner S, Spreafico F, Burgold T, Doni M, Rosato U, Amati B, Testa G (2011) The methyltransferase Set7/9 (Setd7) is dispensable for the p53-mediated DNA damage response in vivo. *Mol Cell* 43(4):681–688
- Chen YX, Du JT, Zhou LX, Liu XH, Zhao YF, Nakanishi H, Li YM (2006) Alternative *O*-GlcNAcylation/*O*-phosphorylation of Ser(16) induce different conformational disturbances to the N terminus of murine estrogen receptor beta. *Chem Biol* 13(9):937–944
- Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, Falck JR, Peng J, Gu W, Zhao Y (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* 6(5):812–819
- Chen C, Nott TJ, Jin J, Pawson T (2011) Deciphering arginine methylation: Tudor tells the tale. *Nat Rev Mol Cell Biol* 12(10):629–642
- Cheng Z, Tang Y, Chen Y, Kim S, Liu H, Li SS, Gu W, Zhao Y (2009) Molecular characterization of propionyllysines in non-histone proteins. *Mol Cell Proteomics* 8(1):45–52
- Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen JV, Mann M (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325(5942):834–840
- Cohen P (2002a) Protein kinases—the major drug targets of the twenty-first century? *Nat Rev Drug Discov* 1(4):309–315
- Cohen P (2002b) The origins of protein phosphorylation. *Nat Cell Biol* 4(5):E127–130
- Coltart DM, Royyuru AK, Williams LJ, Glunz PW, Sames D, Kuduk SD, Schwarz JB, Chen XT, Danishefsky SJ, Live DH (2002) Principles of mucin architecture: structural studies on synthetic glycopeptides bearing clustered mono-, di-, tri-, and hexasaccharide glycodomains. *J Am Chem Soc* 124(33):9833–9844
- Copeland RA, Solomon ME, Richon VM (2009) Protein methyltransferases as a target class for drug discovery. *Nat Rev Drug Discov* 8(9):724–732
- Corzana F, Busto JH, Engelsen SB, Jimenez-Barbero J, Asensio JL, Peregrina JM, Avenoza A (2006a) Effect of beta-O-glucosylation on L-Ser and L-Thr diamides: a bias toward alpha-helical conformations. *Chem-Eur J* 12(30):7864–7871
- Corzana F, Busto JH, Jimenez-Oses G, Asensio JL, Jimenez-Barbero J, Peregrina JM, Avenoza A (2006b) New insights into alpha-GalNAc-Ser motif: influence of hydrogen bonding versus solvent interactions on the preferred conformation. *J Am Chem Soc* 128(45):14640–14648
- Corzana F, Busto JH, Jimenez-Oses G, de Luis MG, Asensio JL, Jimenez-Barbero J, Peregrina JM, Avenoza A (2007) Serine versus threonine glycosylation: the methyl group causes a drastic alteration on the carbohydrate orientation and on the surrounding water shell. *J Am Chem Soc* 129(30):9458–9467
- Cummings RD (2009) The repertoire of glycan determinants in the human glycome. *Mol Biosyst* 5(10):1087–1104
- Dehennaut V, Hanouille X, Bodart JF, Vilain JP, Michalski JC, Landrieu I, Lippens G, Lefebvre T (2008) Microinjection of recombinant *O*-GlcNAc transferase potentiates *Xenopus* oocytes M-phase entry. *Biochem Biophys Res Commun* 369(2):539–546
- Del Rizzo PA, Trievel RC (2011) Substrate and product specificities of SET domain methyltransferases. *Epigenetics* 6(9):1059–1067
- DeMarco ML, Woods RJ (2008) Structural glycobiology: a game of snakes and ladders. *Glycobiology* 18(6):426–440
- Deshmukh L, Meller N, Alder N, Byzova T, Vinogradova O (2011) Tyrosine phosphorylation as a conformational switch: a case study of integrin beta3 cytoplasmic tail. *J Biol Chem* 286(47):40943–40953
- Di Lorenzo A, Bedford MT (2011) Histone arginine methylation. *FEBS Lett* 585(13):2024–2031
- Dose A, Liokatis S, Theillet FX, Selenko P, Schwarzer D (2011) NMR profiling of histone deacetylase and acetyl-transferase activities in real time. *ACS Chem Biol* 6(5):419–424
- Du JT, Li YM, Wei W, Wu GS, Zhao YF, Kanazawa K, Nemoto T, Nakanishi H (2005) Low-barrier hydrogen bond between phosphate and the amide group in phosphopeptide. *J Am Chem Soc* 127(47):16350–16351
- Du J, Zhou Y, Su X, Yu JJ, Khan S, Jiang H, Kim J, Woo J, Kim JH, Choi BH, He B, Chen W, Zhang S, Cerione RA, Auwerx J, Hao Q, Lin H (2011) Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science* 334(6057):806–809
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6(3):197–208
- Egorova KS, Olenkina OM, Olenina LV (2010) Lysine methylation of nonhistone proteins is a way to regulate their stability and function. *Biochemistry (Mosc)* 75(5):535–548
- Eichler J, Adams MW (2005) Posttranslational protein modification in Archaea. *Microbiol Mol Biol Rev* 69(3):393–425
- Erbel PJA, Karimi-Nejad Y, van Kuik JA, Boelens R, Kamerling JP, Vliegthart JFG (2000) Effects of the *N*-linked glycans on the 3D structure of the free alpha-subunit of human chorionic gonadotropin. *Biochemistry* 39(20):6012–6021

- Erce MA, Pang CN, Hart-Smith G, Wilkins MR (2012) The methylproteome and the intracellular methylation network. *Proteomics* 12:1–23
- Fauvet B, Fares MB, Samuel F, Dikiy I, Tandon A, Eliezer D, Lashuel HA (2012) Characterization of semisynthetic and naturally N alpha-acetylated alpha-synuclein in vitro and in intact cells: implications for aggregation and cellular properties of alpha-synuclein. *J Biol Chem* 287(34):28243–28262
- Fletcher CM, Harrison RA, Lachmann PJ, Neuhaus D (1994) Structure of a soluble, glycosylated form of the human-complement regulatory protein Cd59. *Structure* 2(3):185–199
- Freeze HH, Haltiwanger RS (2009) Other classes of ER/golgi-derived glycans. In: Varki A, Cummings RD, Esko JD et al (eds) *Essentials of glycobiology*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Fullgrabe J, Kavanagh E, Joseph B (2011) Histone onco-modifications. *Oncogene* 30(31):3391–3403
- Gao J, Xu D (2012) Correlation between posttranslational modification and intrinsic disorder in protein. *Pac Symp Biocomput* 94–103
- Gardino AK, Yaffe MB (2011) 14-3-3 proteins as signaling integration points for cell cycle control and apoptosis. *Semin Cell Dev Biol* 22(7):688–695
- Garrett DS, Seok YJ, Peterkofsky A, Clore GM, Gronenborn AM (1998) Tautomeric state and pKa of the phosphorylated active site histidine in the N-terminal domain of enzyme I of the *Escherichia coli* phosphoenolpyruvate: sugar phosphotransferase system. *Protein Sci* 7(3):789–793
- Garrity J, Gardner JG, Hawse W, Wolberger C, Escalante-Semerena JC (2007) *N*-lysine propionylation controls the activity of propionyl-CoA synthetase. *J Biol Chem* 282(41):30239–30245
- Graff J, Kim D, Dobbin MM, Tsai LH (2011) Epigenetic regulation of gene expression in physiological and pathological brain processes. *Physiol Rev* 91(2):603–649
- Grimmler M, Wang Y, Mund T, Cilensek Z, Keidel EM, Waddell MB, Jakel H, Kullmann M, Kriwacki RW, Hengst L (2007) Cdk-inhibitory activity and stability of p27Kip1 are directly regulated by oncogenic tyrosine kinases. *Cell* 128(2):269–280
- Gupta S, Kim SY, Artis S, Molfese DL, Schumacher A, Sweatt JD, Paylor RE, Lubin FD (2010) Histone methylation regulates memory formation. *J Neurosci* 30(10):3589–3599
- Haberland M, Montgomery RL, Olson EN (2009) The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat Rev Genet* 10(1):32–42
- Han S, Brunet A (2012) Histone methylation makes its mark on longevity. *Trends Cell Biol* 22(1):42–49
- Hart GW, Akimoto Y (2009) The *O*-GlcNAc modification. In: Varki A, Cummings RD, Esko JD et al (eds) *Essentials of glycobiology*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Hart GW, Copeland RJ (2010) Glycomics hits the big time. *Cell* 143(5):672–676
- Hart GW, Housley MP, Slawson C (2007) Cycling of *O*-linked beta-*N*-acetylglucosamine on nucleocytoplasmic proteins. *Nature* 446(7139):1017–1022
- Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O (2011) Cross talk between *O*-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu Rev Biochem* 80:825–858
- Hashimoto R, Fujitani N, Takegawa Y, Kuroguchi M, Matsushita T, Naruchi K, Ohyabu N, Hinou H, Gao XD, Manri N, Satake H, Kaneko A, Sakamoto T, Nishimura SI (2011) An efficient approach for the characterization of mucin-type glycopeptides: the effect of *O*-glycosylation on the conformation of synthetic mucin peptides. *Chem Eur J* 17(8):2393–2404
- Heightman TD (2011) Chemical biology of lysine demethylases. *Curr Chem Genomics* 5(Suppl 1):62–71
- Hejjaoui M, Butterfield S, Fauvet B, Vercruyse F, Cui J, Dikiy I, Prudent M, Olschewski D, Zhang Y, Eliezer D, Lashuel HA (2012) Elucidating the role of C-terminal post-translational modifications using protein semisynthesis strategies: alpha-synuclein phosphorylation at tyrosine 125. *J Am Chem Soc* 134(11):5196–5210
- Himmel S, Wolff S, Becker S, Lee D, Griesinger C (2010) Detection and identification of protein-phosphorylation sites in histidines through HNP correlation patterns. *Angew Chem Int Ed Engl* 49(47):8971–8974
- Hunter T (2009) Tyrosine phosphorylation: thirty years and counting. *Curr Opin Cell Biol* 21(2):140–146
- Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* 32(3):1037–1049
- James TL (1985) Phosphorus-31 NMR as a probe for phosphoproteins. *CRC Crit Rev Biochem* 18(1):1–30
- Jenuwein T, Allis CD (2001) Translating the histone code. *Science* 293(5532):1074–1080
- Johnson LN, Lewis RJ (2001) Structural basis for control by phosphorylation. *Chem Rev* 101(8):2209–2242
- Jones BE, Rajagopal P, Klevit RE (1997) Phosphorylation on histidine is accompanied by localized structural changes in the phosphocarrier protein. HPr from *Bacillus subtilis*. *Protein Sci* 6(10):2107–2119
- Julien SG, Dube N, Hardy S, Tremblay ML (2011) Inside the human cancer tyrosine phosphatome. *Nat Rev Cancer* 11(1):35–49
- Kar S, Sakaguchi K, Shimohigashi Y, Samaddar S, Banerjee R, Basu G, Swaminathan V, Kundu TK, Roy S (2002) Effect of phosphorylation on the structure and fold of transactivation domain of p53. *J Biol Chem* 277(18):15579–15585
- Kee JM, Muir TW (2012) Chasing phosphohistidine, an elusive sibling in the phosphoamino acid family. *ACS Chem Biol* 7(1):44–51
- Kelly TK, De Carvalho DD, Jones PA (2010) Epigenetic modifications as therapeutic targets. *Nat Biotechnol* 28(10):1069–1078
- Khoury GA, Baliban RC, Floudas CA (2011) Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep* 1:90
- Kolch W, Pitt A (2010) Functional proteomics to dissect tyrosine kinase signalling pathways in cancer. *Nat Rev Cancer* 10(9):618–629
- Kouzarides T (2007) Chromatin modifications and their function. *Cell* 128(4):693–705
- Kruse JP, Gu W (2009) Modes of p53 regulation. *Cell* 137(4):609–622
- Lakowski TM, Hart P, Ahern CA, Martin NI, Frankel A (2010) Neta-substituted arginyl peptide inhibitors of protein arginine *N*-methyltransferases. *ACS Chem Biol* 5(11):1053–1063
- Landrieu I, Odaert B, Wieruszkeski JM, Drobecq H, Rousselot-Pailley P, Inze D, Lippens G (2001) p13(SUC1) and the WW domain of PIN1 bind to the same phosphothreonine-proline epitope. *J Biol Chem* 276(2):1434–1438
- Landrieu I, Lacosse L, Leroy A, Wieruszkeski JM, Trivelli X, Sillen A, Sibille N, Schwalbe H, Saxena K, Langer T, Lippens G (2006) NMR analysis of a Tau phosphorylation pattern. *J Am Chem Soc* 128(11):3575–3583
- Landrieu I, Smet-Nocca C, Amniai L, Louis JV, Wieruszkeski JM, Goris J, Janssens V, Lippens G (2011) Molecular implication of PP2A and Pin1 in the Alzheimer's disease specific hyperphosphorylation of Tau. *PLoS ONE* 6(6):e21521
- Larkin A, Imperiali B (2011) The expanding horizons of asparagine-linked glycosylation. *Biochemistry* 50(21):4411–4426
- Latham JA, Dent SY (2007) Cross-regulation of histone modifications. *Nat Struct Mol Biol* 14(11):1017–1024
- Leemhuis H, Packman LC, Nightingale KP, Hollfelder F (2008) The human histone acetyltransferase P/CAF is a promiscuous histone propionyltransferase. *Chem Bio Chem* 9(4):499–503

- Lehnertz B, Rogalski JC, Schulze FM, Yi L, Lin S, Kast J, Rossi FM (2011) p53-dependent transcription and tumor suppression are not affected in Set7/9-deficient mice. *Mol Cell* 43(4):673–680
- Lemmon MA, Schlessinger J (2010) Cell signaling by receptor tyrosine kinases. *Cell* 141(7):1117–1134
- Leroy A, Landrieu I, Huvent I, Legrand D, Codeville B, Wieruszkeski JM, Lippens G (2010) Spectroscopic studies of GSK3{beta} phosphorylation of the neuronal Tau protein and its interaction with the N-terminal domain of apolipoprotein E. *J Biol Chem* 285(43):33435–33444
- Liang FC, Chen RP, Lin CC, Huang KT, Chan SI (2006) Tuning the conformation properties of a peptide by glycosylation and phosphorylation. *Biochem Biophys Res Commun* 342(2):482–488
- Liepinsh E, Otting G (1996) Proton exchange rates from amino acid side chains—implications for image contrast. *Magn Reson Med* 35(1):30–42
- Lin H, Su X, He B (2012) Protein lysine acylation and cysteine succination by intermediates of energy metabolism. *ACS Chem Biol* 7(6):947–960
- Liokatis S, Dose A, Schwarzer D, Selenko P (2010) Simultaneous detection of protein phosphorylation and acetylation by high-resolution NMR spectroscopy. *J Am Chem Soc* 132(42):14704–14705
- Liokatis S, Stuetzter A, Elsaesser S, Theillet FX, Klingberg R, van Rossum B, Schwarzer D, Allis CD, Fischle W, Selenko P (2012) Phosphorylation of histone H3 Serine 10 establishes a hierarchy for subsequent intramolecular modification events. *Nat Struct Mol Biol* 19(8):819–823
- Liou YC, Zhou XZ, Lu KP (2011) Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins. *Trends Biochem Sci* 36(10):501–514
- Lippens G, Landrieu I, Hanouille X (2008) Studying posttranslational modifications by in-cell NMR. *Chem Biol* 15(4):311–312
- Liu B, Lin Y, Darwanto A, Song X, Xu G, Zhang K (2009) Identification and characterization of propionylation at histone H3 lysine 23 in mammalian cells. *J Biol Chem* 284(47):32288–32295
- Liu BA, Shah E, Jablonowski K, Stergachis A, Engelmann B, Nash PD (2011) The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. *Sci Signal* 4(202):83
- Lu KP, Liou YC, Zhou XZ (2002) Pinning down proline-directed phosphorylation signaling. *Trends Cell Biol* 12(4):164–172
- Luo M (2012) Current chemical biology approaches to interrogate protein methyltransferases. *ACS Chem Biol* 443–463
- Macdonald JM, LeBlanc DA, Haas AL, London RE (1999) An NMR analysis of the reaction of ubiquitin with [acetyl-1-13C] aspirin. *Biochem Pharmacol* 57(11):1233–1244
- Macnaughtan MA, Kane AM, Prestegard JH (2005) Mass spectrometry assisted assignment of NMR resonances in reductively 13C-methylated proteins. *J Am Chem Soc* 127(50):17626–17627
- Mahajan A, Yuan C, Lee H, Chen ES, Wu PY, Tsai MD (2008) Structure and function of the phosphothreonine-specific FHA domain. *Sci Signal* 1(51):re12
- Mak A, Smillie LB, Barany M (1978) Specific phosphorylation at serine-283 of alpha tropomyosin from frog skeletal and rabbit skeletal and cardiac muscle. *Proc Natl Acad Sci USA* 75(8):3588–3592
- Marth JD, Grewal PK (2008) Mammalian glycosylation in immunity. *Nat Rev Immunol* 8(11):874–887
- Martin L, Latypova X, Terro F (2011) Post-translational modifications of Tau protein: implications for Alzheimer's disease. *Neurochem Int* 58(4):458–471
- Matheis G, Whitaker JR (1984) 31P NMR chemical shifts of phosphate covalently bound to proteins. *Int J Biochem* 16(8):867–873
- Maurer-Stroh S, Dickens NJ, Hughes-Davies L, Kouzarides T, Eisenhaber F, Ponting CP (2003) The Tudor domain 'royal family': Tudor, plant agenet, chromo, PWWP and MBT domains. *Trends Biochem Sci* 28(2):69–74
- Metzler WJ, Bajorath J, Fenderson W, Shaw SY, Constantine KL, Naemura J, Leytze G, Peach RJ, Lavoie TB, Mueller L, Linsley PS (1997) Solution structure of human CTLA-4 and delineation of a CD80/CD86 binding site conserved in CD28. *Nat Struct Biol* 4(7):527–531
- Meyer B, Moller H (2007) Conformation of glycopeptides and glycoproteins. *Top Curr Chem* 267:187–251
- Mittag T, Orlicky S, Choy WY, Tang X, Lin H, Sicheri F, Kay LE, Tyers M, Forman-Kay JD (2008) Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc Natl Acad Sci USA* 105(46):17772–17777
- Narimatsu Y, Kubota T, Furukawa S, Morii H, Narimatsu H, Yamasaki K (2010) Effect of glycosylation on Cis/trans isomerization of prolines in IgA1-hinge peptide. *J Am Chem Soc* 132(16):5548–5549
- Nielsen G, Schwalbe H (2011) NMR spectroscopic investigations of the activated p38alpha mitogen-activated protein kinase. *Chem Bio Chem* 12(17):2599–2607
- Norris KL, Lee J-Y, Yao T-P (2009) Acetylation goes global: the emergence of acetylation biology. *Sci Signal* 2(97):pe76
- Ohtsubo K, Marth JD (2006) Glycosylation in cellular mechanisms of health and disease. *Cell* 126(5):855–867
- Owens NW, Braun C, O'Neil JD, Marat K, Schweizer F (2007) Effects of glycosylation of (2S,4R)-4-hydroxyproline on the conformation, kinetics, and thermodynamics of prolyl amide isomerization. *J Am Chem Soc* 129(38):11670–11671
- Owens NW, Lee A, Marat K, Schweizer F (2009) The implications of (2S,4S)-hydroxyproline 4-O-glycosylation for prolyl amide isomerization. *Chem Eur J* 15(40):10649–10657
- Owens NW, Stetefeld J, Lattova E, Schweizer F (2010) Contiguous O-galactosylation of 4(R)-hydroxy-L-proline residues forms very stable polyproline II helices. *J Am Chem Soc* 132(14):5036–5042
- Parry RV, Ward SG (2010) Protein arginine methylation: a new handle on T lymphocytes? *Trends Immunol* 31(4):164–169
- Patchell VB, Vorotnikov AV, Gao Y, Low DG, Evans JS, Fattoum A, El-Mezgueldi M, Marston SB, Levine BA (2002) Phosphorylation of the minimal inhibitory region at the C-terminus of caldesmon alters its structural and actin binding properties. *Biochim Biophys Acta* 1596(1):121–130
- Peng C, Lu Z, Xie Z, Cheng Z, Chen Y, Tan M, Luo H, Zhang Y, He W, Yang K, Zwaans BM, Tishkoff D, Ho L, Lombard D, He TC, Dai J, Verdin E, Ye Y, Zhao Y (2011) The first identification of lysine malonylation substrates and its regulatory enzyme. *Mol Cell Proteomics* 10(12):M111.012658
- Perez Y, Gairi M, Pons M, Bernado P (2009) Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. *J Mol Biol* 391(1):136–148
- Peter CJ, Akbarian S (2011) Balancing histone methylation activities in psychiatric disorders. *Trends Mol Med* 17(7):372–379
- Pethe K, Bifani P, Drobecq H, Sergheraert C, Debrie AS, Loch C, Menozzi FD (2002) Mycobacterial heparin-binding hemagglutinin and laminin-binding protein share antigenic methyllysines that confer resistance to proteolysis. *Proc Natl Acad Sci USA* 99(16):10759–10764
- Prabakaran S, Everley RA, Landrieu I, Wieruszkeski JM, Lippens G, Steen H, Gunawardena J (2011) Comparative analysis of Erk phosphorylation suggests a mixed strategy for measuring phospho-form distributions. *Mol Syst Biol* 7:482
- Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK (2007) Intrinsic disorder and functional proteomics. *Biophys J* 92(5):1439–1456
- Rajagopal P, Waygood EB, Klevit RE (1994) Structural consequences of histidine phosphorylation: NMR characterization of the

- phosphohistidine form of histidine-containing protein from *Bacillus subtilis* and *Escherichia coli*. *Biochemistry* 33(51):15271–15282
- Ramelot TA, Nicholson LK (2001) Phosphorylation-induced structural changes in the amyloid precursor protein cytoplasmic tail detected by NMR. *J Mol Biol* 307(3):871–884
- Rich JR, Withers SG (2009) Emerging methods for the production of homogeneous human glycoproteins. *Nat Chem Biol* 5(4):206–215
- Riester D, Wegener D, Hildmann C, Schwienhorst A (2004) Members of the histone deacetylase superfamily differ in substrate specificity towards small synthetic substrates. *Biochem Biophys Res Commun* 324(3):1116–1123
- Roth J (2002) Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chem Rev* 102(2):285–303
- Sakai T, Tochio H, Tenno T, Ito Y, Kokubo T, Hiroaki H, Shirakawa M (2006) In-cell NMR spectroscopy of proteins inside *Xenopus laevis* oocytes. *J Biomol NMR* 36(3):179–188
- Salah Z, Alian A, Aqeilan RI (2012) WW domain-containing proteins: retrospectives and the future. *Front Biosci* 17:331–348
- Sanchez R, Zhou MM (2009) The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Devel* 12(5):659–665
- Sanchez R, Zhou MM (2011) The PHD finger: a versatile epigenome reader. *Trends Biochem Sci* 36(7):364–372
- Schutkowski M, Bernhardt A, Zhou XZ, Shen M, Reimer U, Rahfeld JU, Lu KP, Fischer G (1998) Role of phosphorylation in determining the backbone dynamics of the serine/threonine-proline motif and Pin1 substrate recognition. *Biochemistry* 37(16):5566–5575
- Schwarz F, Aebi M (2011) Mechanisms and principles of N-linked protein glycosylation. *Curr Opin Struct Biol* 21(5):576–582
- Seet BT, Dikic I, Zhou MM, Pawson T (2006) Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol* 7(7):473–483
- Selenko P, Frueh DP, Elsaesser SJ, Haas W, Gygi SP, Wagner G (2008) In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat Struct Mol Biol* 15(3):321–329
- Sibille N, Huvent I, Fauquant C, Verdegem D, Amniai L, Leroy A, Wieruszkeski JM, Lippens G, Landrieu I (2011) Structural characterization by nuclear magnetic resonance of the impact of phosphorylation in the proline-rich region of the disordered Tau protein. *Proteins* 80:454–462
- Simanek EE, Huang DH, Pasternack L, Machajewski TD, Seitz O, Millar DS, Dyson HJ, Wong CH (1998) Glycosylation of threonine of the repeating unit of RNA polymerase II with beta-linked N-acetylglucosamine leads to a turnlike structure. *J Am Chem Soc* 120(45):11567–11575
- Skrisovska L, Schubert M, Allain FHT (2010) Recent advances in segmental isotope labeling of proteins: NMR applications to large proteins and glycoproteins. *J Biomol NMR* 46(1):51–65
- Slawson C, Hart GW (2011) O-GlcNAc signalling: implications for cancer cell biology. *Nat Rev Cancer* 11(9):678–684
- Slawson C, Copeland RJ, Hart GW (2010) O-GlcNAc signaling: a metabolic link between diabetes and cancer? *Trends Biochem Sci* 35(10):547–555
- Slyno V, Schubert M, Numao S, Kowarik M, Aebi M, Allain FHT (2009) NMR structure determination of a segmentally labeled glycoprotein using in vitro glycosylation. *J Am Chem Soc* 131(3):1274–1281
- Smet-Nocca C, Wieruszkeski JM, Melnyk O, Benecke A (2010) NMR-based detection of acetylation sites in peptides. *J Pept Sci* 16(8):414–423
- Smet-Nocca C, Broncel M, Wieruszkeski JM, Tokarski C, Hanouille X, Leroy A, Landrieu I, Rolando C, Lippens G, Hackenberger CP (2011) Identification of O-GlcNAc sites within peptides of the Tau protein and their impact on phosphorylation. *Mol BioSyst* 7(5):1420–1429
- Smith BC, Denu JM (2007a) Acetyl-lysine analog peptides as mechanistic probes of protein deacetylases. *J Biol Chem* 282(51):37256–37265
- Smith BC, Denu JM (2007b) Sir2 deacetylases exhibit nucleophilic participation of acetyl-lysine in NAD<sup>+</sup> cleavage. *J Am Chem Soc* 129(18):5802–5803
- Smith BC, Denu JM (2009) Chemical mechanisms of histone lysine and arginine modifications. *Biochim Biophys Acta* 1789(1):45–57
- Songyang Z, Lu KP, Kwon YT, Tsai LH, Filhol O, Cochet C, Brickey DA, Soderling TR, Bartleson C, Graves DJ, DeMaggio AJ, Hoekstra MF, Blenis J, Hunter T, Cantley LC (1996) A structural basis for substrate specificities of protein Ser/Thr kinases: primary sequence preference of casein kinases I and II, NIMA, phosphorylase kinase, calmodulin-dependent kinase II, CDK5, and Erk1. *Mol Cell Biol* 16(11):6486–6493
- Spannhoff A, Hauser AT, Heinke R, Sippl W, Jung M (2009) The emerging therapeutic potential of histone methyltransferase and demethylase inhibitors. *Chem Med Chem* 4(10):1568–1582
- Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 12(4):43R–56R
- Stanley P, Schachter H, Taniguchi N (2009) N-Glycans. In: Varki A, Cummings RD, Esko JD et al (eds) *Essentials of glycobiology*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Stark GR, Wang Y, Lu T (2011) Lysine methylation of promoter-bound transcription factors and relevance to cancer. *Cell Res* 21(3):375–380
- Suganuma T, Workman JL (2011) Signals and combinatorial functions of histone modifications. *Annu Rev Biochem* 80:473–499
- Suh JY, Cai M, Clore GM (2008) Impact of phosphorylation on structure and thermodynamics of the interaction between the N-terminal domain of enzyme I and the histidine phosphocarrier protein of the bacterial phosphotransferase system. *J Biol Chem* 283(27):18980–18989
- Tagashira M, Iijima H, Toma K (2002) An NMR study of O-glycosylation induced structural changes in the alpha-helix of calcitonin. *Glycoconj J* 19(1):43–52
- Tait S, Dutta K, Cowburn D, Warwicker J, Doig AJ, McCarthy JE (2010) Local control of a disorder-order transition in 4E-BP1 underpins regulation of translation via eIF4E. *Proc Natl Acad Sci USA* 107(41):17627–17632
- Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B, Zhao Y (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146(6):1016–1028
- Teperino R, Schoonjans K, Auwerx J (2010) Histone methyltransferases and demethylases; can they link metabolism and transcription? *Cell Metab* 12(4):321–327
- Teyssier C, Le Romancer M, Sentis S, Jalaguier S, Corbo L, Cavailles V (2010) Protein arginine methylation in estrogen signaling and estrogen-related cancers. *Trends Endocrinol Metab* 21(3):181–189
- Theillet FX, Liokatis S, Jost JO, Bekei B, Rose HM, Binolfi A, Schwarzer D, Selenko P (2012) Site-specific mapping and time-resolved monitoring of lysine methylation by high-resolution NMR spectroscopy. *J Am Chem Soc* 134(18):7616–7619
- van Nuland NA, Boelens R, Scheek RM, Robillard GT (1995) High-resolution structure of the phosphorylated form of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from NMR-NOE data. *J Mol Biol* 246(1):180–193
- Varier RA, Timmers HT (2011) Histone lysine methylation and demethylation pathways in cancer. *BBA-Rev Cancer* 1815(1):75–89



- Varki A (2008) Sialic acids in human health and disease. *Trends Mol Med* 14(8):351–360
- Vigerust DJ, Shepherd VL (2007) Virus glycosylation: role in virulence and immune interactions. *Trends Microbiol* 15(5): 211–218
- Vollmuth F, Geyer M (2011) Interaction of propionylated and butyrylated histone H3 lysine marks with Brd4 bromodomains. *Angew Chem Int Ed Engl* 49(38):6768–6772
- Walsh CT, Garneau-Tsodikova S, Gatto GJ Jr (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl* 44(45):7342–7372
- Wang Z, Udeshi ND, O'Malley M, Shabanowitz J, Hunt DF, Hart GW (2010) Enrichment and site mapping of *O*-linked *N*-acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation mass spectrometry. *Mol Cell Proteomics* 9(1):153–160
- Weiwad M, Kullertz G, Schutkowski M, Fischer G (2000) Evidence that the substrate backbone conformation is critical to phosphorylation by p42 MAP kinase. *FEBS Lett* 478(1–2):39–42
- Werner-Allen JW, Lee CJ, Liu P, Nicely NI, Wang S, Greenleaf AL, Zhou P (2011) *cis*-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. *J Biol Chem* 286(7):5717–5726
- Wintjens R, Wieruszkeski JM, Drobecq H, Rousselot-Pailley P, Buee L, Lippens G, Landrieu I (2001) 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. *J Biol Chem* 276(27):25150–25156
- Wolf SS (2009) The protein arginine methyltransferase family: an update about function, new perspectives and the physiological role in humans. *Cell Mol Life Sci* 66(13):2109–2121
- Wormald MR, Petrescu AJ, Pao YL, Glithero A, Elliott T, Dwek RA (2002) Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. *Chem Rev* 102(2):371–386
- Wu WG, Pasternack L, Huang DH, Koeller KM, Lin CC, Seitz O, Wong CH (1999) Structural study on *O*-glycopeptides: glycosylation-induced conformational changes of *O*-GlcNAc, *O*-LacNAc, *O*-sialyl-LacNAc, and *O*-sialyl-lewis-X peptides of the mucin domain of MAdCAM-1. *J Am Chem Soc* 121(11):2409–2417
- Wurzenberger C, Gerlich DW (2011) Phosphatases: providing safe passage through mitotic exit. *Nat Rev Mol Cell Biol* 12(8): 469–482
- Wyss DF, Choi JS, Li J, Knoppers MH, Willis KJ, Arulanandam AR, Smolyar A, Reinherz EL, Wagner G (1995) Conformation and function of the *N*-linked glycan in the adhesion domain of human CD2. *Science* 269(5228):1273–1278
- Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN (2007) Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J Proteome Res* 6(5):1917–1932
- Xu AS, Macdonald JM, Labotka RJ, London RE (1999) NMR study of the sites of human hemoglobin acetylated by aspirin. *Biochim Biophys Acta* 1432(2):333–349
- Yaffe MB (2002) Phosphotyrosine-binding domains in signal transduction. *Nat Rev Mol Cell Biol* 3(3):177–186
- Yamaguchi Y, Kato K (2010) Dynamics and interactions of glycoconjugates probed by stable-isotope-assisted NMR spectroscopy. *Method Enzymol* 478:305–322
- Yang XJ, Seto E (2007) HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* 26(37):5310–5318
- Yap KL, Zhou MM (2011) Structure and mechanisms of lysine methylation recognition by the chromodomain in gene transcription. *Biochemistry* 50(12):1966–1980
- Yu CH, Si T, Wu WH, Hu J, Du JT, Zhao YF, Li YM (2008) *O*-GlcNAcylation modulates the self-aggregation ability of the fourth microtubule-binding repeat of Tau. *Biochem Biophys Res Commun* 375(1):59–62
- Yuzwa SA, Shan X, Macauley MS, Clark T, Skorobogatko Y, Vosseller K, Vocadlo DJ (2012) Increasing *O*-GlcNAc slows neurodegeneration and stabilizes Tau against aggregation. *Nat Chem Biol* 8(4):393–399
- Zhang K, Chen Y, Zhang Z, Zhao Y (2009) Identification and verification of lysine propionylation and butyrylation in yeast core histones using PTMap software. *J Proteome Res* 8(2): 900–906
- Zhang Z, Tan M, Xie Z, Dai L, Chen Y, Zhao Y (2011) Identification of lysine succinylation as a new post-translational modification. *Nat Chem Biol* 7(1):58–63
- Zhou XZ, Kops O, Werner A, Lu PJ, Shen M, Stoller G, Kullertz G, Stark M, Fischer G, Lu KP (2000) Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and Tau proteins. *Mol Cell* 6(4):873–883

# Unraveling a phosphorylation event in a folded protein by NMR spectroscopy: phosphorylation of the Pin1 WW domain by PKA

Caroline Smet-Nocca · Hélène Launay ·  
Jean-Michel Wieruszeski · Guy Lippens ·  
Isabelle Landrieu

Received: 9 October 2012 / Accepted: 15 February 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** The Pin1 protein plays a critical role in the functional regulation of the hyperphosphorylated neuronal Tau protein in Alzheimer's disease and is by itself regulated by phosphorylation. We have used Nuclear Magnetic Resonance (NMR) spectroscopy to both identify the PKA phosphorylation site in the Pin1 WW domain and investigate the functional consequences of this phosphorylation. Detection and identification of phosphorylation on serine/threonine residues in a globular protein, while mostly occurring in solvent-exposed flexible loops, does not lead to chemical shift changes as obvious as in disordered proteins and hence does not necessarily shift the resonances outside the spectrum of the folded protein. Other complications were encountered to characterize the extent of the phosphorylation, as part of the  $^1\text{H}$ ,  $^{15}\text{N}$  amide resonances around the phosphorylation site are specifically

broadened in the unphosphorylated state. Despite these obstacles, NMR spectroscopy was an efficient tool to confirm phosphorylation on S16 of the WW domain and to quantify the level of phosphorylation. Based on this analytical characterization, we show that WW phosphorylation on S16 abolishes its binding capacity to a phosphorylated Tau peptide. A reduced conformational heterogeneity and flexibility of the phospho-binding loop upon S16 phosphorylation could account for part of the decreased affinity for its phosphorylated partner. Additionally, a structural model of the phospho-WW obtained by molecular dynamics simulation and energy minimization suggests that the phosphate moiety of phospho-S16 could compete with the phospho-substrate.

**Keywords** NMR spectroscopy · Post-translational modifications · Phosphorylation · Protein · Pin1 WW binding module

**Electronic supplementary material** The online version of this article (doi:10.1007/s10858-013-9716-z) contains supplementary material, which is available to authorized users.

C. Smet-Nocca · H. Launay · J.-M. Wieruszeski · G. Lippens · I. Landrieu  
Institut Fédératif de Recherches 147, CNRS UMR 8576,  
Université de Lille-Nord de France, Villeneuve d'Ascq, France

C. Smet-Nocca (✉)  
Unité de Glycobiologie Structurale et Fonctionnelle (CNRS UMR 8576), Université de Lille-Nord de France, Cité Scientifique Bâtiment C9, 59655 Villeneuve d'Ascq, France  
e-mail: caroline.smet@univ-lille1.fr

I. Landrieu (✉)  
Unité de Glycobiologie Structurale et Fonctionnelle (CNRS UMR 8576), Interdisciplinary Research Institute, Université de Lille-Nord de France, 50 Avenue de Halley, 59658 Villeneuve d'Ascq, France  
e-mail: isabelle.landrieu@univ-lille1.fr

## Abbreviations

AD	Alzheimer's disease
IDPs	Intrinsically disordered proteins
Pin1	Protein interacting with NIMA-1
NIMA	Never in mitosis gene A
PKA	Protein kinase A
MALDI-TOF MS	Matrix-assisted laser desorption ionization- time of flight mass spectrometry
PTMs	Post translational modifications
RP-HPLC	Reverse phase high pressure liquid chromatography
WW	Trp-Trp binding module
WW <sup>PKA</sup>	PKA-phosphorylated WW domain
Phospho-WW	Phosphorylated form of the WW domain

## Introduction

Protein phosphorylation is a major regulatory process within eukaryotic cells. Characterization of phosphorylation sites is an important issue to understand the impact of phosphorylation on protein activity and function (Johnson and Barford 1993). To allow phosphorylation-dependent signalling pathways, proteins incorporate modular domains enabling specific recognition of other proteins in their phosphorylated states. The human peptidyl-prolyl *cis/trans* isomerase Pin1 was first implicated in the regulation of cell cycle through interactions with mitotic phospho-proteins by one of these modules, the WW domain (Lu et al. 1996; Yaffe et al. 1997). More recently, functional interactions with phosphorylated neuronal Tau and APP proteins have linked Pin1 to Alzheimer's disease (AD)-neurodegenerative processes (Arosio et al. 2012; Balastik et al. 2007; Bulbarelli et al. 2009; Lonati et al. 2011; Lu et al. 1999; Ma et al. 2012; Pastorino et al. 2006; Sultana et al. 2006). In particular, Tau phosphorylation level regulates its physiological function of microtubule binding and tubulin polymerization as well as the formation of the pathological Tau aggregates, one of the hallmarks of AD. Pin1 was shown to play a critical role in the functional regulation of the hyperphosphorylated neuronal Tau protein in AD neurons as Pin1<sup>-/-</sup> mice develop an age-related tauopathy (Lu et al. 1999; Liou et al. 2003). For these reasons, molecular characterization of the Pin1 and phospho-Tau interactions has gained attention (Hamdane et al. 2002; Smet et al. 2004, 2005a; b). Interestingly, Pin1 is regulated by phosphorylation itself (Lu et al. 2002; Rangasamy et al. 2012). High levels of phosphorylation of the Pin1 protein are detected in AD affected brain tissues compared to control tissues (Ando et al. 2012). Additionally, phosphorylation of Pin1 was previously shown to affect its cellular function (Lu et al. 2002).

Here, we investigate the phosphorylation by the PKA kinase of the WW domain of Pin1, a small module folded into a triple-stranded  $\beta$ -sheet, localized in the amino-terminal part of the Pin1 protein (Ranganathan et al. 1997; Verdecia et al. 2000; Wintjens et al. 2001). The WW domain of Pin1 exhibits a strong specificity for phosphorylated substrates and is responsible for the binding of phospho-proteins through phospho-S/T-P motifs generated by proline-directed kinases (Yaffe et al. 1997). Phosphorylation of the Pin1 WW domain by PKA was shown to modify the interaction with its partners and leads to a change in sub-cellular localization (Lu et al. 2002). Firstly indirectly identified through experiments involving alanine mutations, the S16 residue located in the phospho-binding loop was proposed to be the phosphorylation site (Lu et al. 2002). Additional evidence comes from the detection of Pin1 hyperphosphorylation at the S16 residue in AD brain

tissues when compared with controls (Ando et al. 2012). The investigation of the functional consequences of the Pin1 WW domain phosphorylation required first the analytical characterization of the WW PKA-phosphorylation site(s) in the Nuclear Magnetic Resonance (NMR) sample.

Whereas immunochemistry and mass spectrometry (MS) have been traditionally the methods of choice to detect PTMs (Post-Translational Modifications), both at the level of individual proteins or more largely on a proteome-wide scale, NMR as an analytical technique has recently entered this field (Theillet et al. 2012). Although limited by the requirement of a stable isotope-labeled protein and low sensitivity, its capacity to detect even in a complex mixture a variety of PTMs, its quantitative nature and non-destructive character have all led to propose this biophysical technique as a useful alternative to the above mentioned techniques. In order to increase its general use, we have previously proposed a reference library of the NMR spectral signatures that the various PTMs might cause (Theillet et al. 2012). These signatures can be divided in “*indicators*”, spectral characteristics that show that a certain PTM is present in the sample, and “*identifying signals*” that allow the site-specific assignment of the PTM to a given residue. As stated, these spectral signatures are quite generally valid for the PTM of amino acids embedded in intrinsically disordered proteins (IDPs). The “*indicators*” for phosphorylation sites are based on <sup>1</sup>H-<sup>15</sup>N chemical shift perturbations, while the “*identifying signals*” correspond to the C $\alpha$  and C $\beta$  chemical shifts of the phospho-S/T residues (Bienkiewicz and Lumb 1999; Theillet et al. 2012). Whereas disordered regions often tend to harbor regulatory PTMs (Iakoucheva et al. 2004), many examples equally exist whereby PTMs are found in the context of globular proteins. Although often located on residues in accessible and flexible loops of proteins, PTMs are embedded in these structured proteins and therefore, their NMR-based detection and characterization do not necessarily obey the same rules as for IDPs. One major issue for the identification of a modification site in a globular protein is to differentiate the chemical shift changes that are related to the secondary or tertiary structural changes from those related to the modification itself. A second issue concerns the determination of the extent of modification as phosphorylation can have an effect on both water exchange and relaxation rates complicating the quantification due to specific line broadening. This effect probably is quite general, as phosphorylation has been shown to be responsible for stabilization/destabilization of local secondary structure (Andrew et al. 2002; Baker et al. 2007; Pufall et al. 2005) and introduction of negative charges through the phosphate moiety will result in water exchange rate perturbations (Bai et al. 1993).

Despite numerous NMR studies that have carefully evaluated the role of phosphorylation on protein structures

and on the modulation of protein–protein interfaces, the modification site is generally not identified by NMR, or at best only assessed by  $^1\text{H}$ - $^{15}\text{N}$  chemical shift mapping (Eto et al. 2007; Ohki et al. 2001; Pullen et al. 1995; Teriete et al. 2009; Wittekind et al. 1989). The analytical characterization of the PKA phosphorylation of the WW domain is thus here addressed in a systematic manner. Based on these results, we show that the PKA-phosphorylated WW is unable to bind a phosphorylated Tau peptide substrate. We next investigate the structural and dynamical perturbations that explain the loss of function due to the phosphorylation. Whereas the triple-stranded conformation of the WW is not changed upon phosphorylation of S16, it does however induce a loss of flexibility of the phospho-binding region as evidenced by changes in NMR relaxation parameters. A model of the phospho-WW domain shows a novel hydrogen bond network as a consequence of the phosphate moiety on S16 that stabilizes the binding-loop.

## Materials and methods

### Production of $^{15}\text{N}/^{13}\text{C}$ - or $^{15}\text{N}$ -labeled WW domain of Pin1 in *Escherichia coli*

The WW domain (residue 5–41 of human Pin1) was produced fused to a N-terminus GST tag in *E. coli* BL21(DE3) strain carrying the WW-pGEX-4T-3 recombinant plasmid. Cells were grown at 37 °C in a M9 minimal medium containing 4 g/L glucose or 2 g/L [ $^{13}\text{C}_6$ ]-glucose for expression of  $^{13}\text{C}$ -labeled proteins, 1 g/L  $^{15}\text{N}$ -ammonium chloride, 1 mM  $\text{MgSO}_4$ , MEM vitamin cocktail (Sigma) and ampicilline (100 mg/L). The induction phase was initiated by addition of 0.5 mM Isopropyl-1-thio- $\beta$ -D-galactopyranoside and continued at 31 °C for 3 h. Cells were harvested by centrifugation and the pellet was suspended in Phosphate Buffer Saline, 5 % glycerol, 2 mM dithiothreitol, 10 mM EDTA, 1 % Triton X-100 pH 7.2 complemented with a protease inhibitor cocktail (Complete<sup>TM</sup>, Roche). The cell lysate was obtained by incubation of 0.25 mg/ml lysozyme with the cell suspension in extraction buffer complemented with RNase and DNase and followed by a brief sonication step. The soluble extract was isolated by centrifugation. The GST-fusion protein was purified on Glutathione Sepharose (GE Healthcare). Soluble extracts were incubated for 3 h at 4 °C with 25–100  $\mu\text{l}$  resin per milliliter of soluble extracts. Unbound proteins were extensively washed away with a wash buffer (PBS, 5 % glycerol, 1 % Triton X-100, 10 mM EDTA) and proteins were eluted by digestion with Thrombin protease at 0.1 unit/mg fusion protein in one bead-volume of elution buffer (50 mM Tris-Cl pH 8.0, 150 mM NaCl, 2 % glycerol, 0.1 % NP-40, 5 mM EDTA, 5 mM DTT) at 20 °C for

20 h. Elution from the beads was then done twice with one bead-volume of elution buffer. The pooled fractions were concentrated and purified by reverse phase chromatography on a Source 5RPC column (GE Healthcare) equilibrated in 0.05 % TFA aqueous solution (solution A). Separation of proteins was carried out at room temperature at a flow rate of 2 mL/min using a linear acetonitrile gradient over 40 min from 5 to 40 % solution B (80 % acetonitrile in solution A). Homogeneous fractions as checked by 15 % denaturing polyacrylamide gel electrophoresis and MALDI-TOF MS were pooled and lyophilized.

### Phosphorylation of the WW domain of Pin1 by PKA

The WW domain was dissolved at 10  $\mu\text{M}$  in 10 mL of phosphorylation buffer (50 mM Hepes.KOH pH 8.0, 50 mM NaCl, 12.5 mM  $\text{MgCl}_2$ , 1 mM EDTA) with 0.1  $\mu\text{M}$  PKA enzyme. The phosphorylation reaction was performed at 30 °C for 8 h. The mixture was next purified by reverse phase chromatography under the same conditions as for the WW domain. Homogeneous fractions were lyophilized. The protein sample was estimated to be 99 % WW proteins.

### NMR spectroscopy

For NMR experiments, the WW domains were dissolved at 0.3–0.4 mM in a buffer containing 50 mM  $\text{d}_{11}$ -Tris pH 6.8, 100 mM NaCl, 1 mM EDTA, 1 mM DTT and 5 %  $\text{D}_2\text{O}$ . As a dissociation constant of 2 mM was measured between Pin1 and phosphate ions (Bayer et al. 2003), we use in this study the deuterated Tris buffer instead of the usual phosphate buffer for this pH range.  $^1\text{H}$  spectra were calibrated with 1 mM sodium 3-trimethylsilyl-3,3',2,2'-d4-propionate (TSP) as a reference. It should be noticed that 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) reference is the more widely used. Since the chemical shift difference between TSP and DSS is a constant under given conditions (TSP resonances are pH-dependent), corrections of chemical shifts given in this study using TSP referencing are of  $-0.015$  ppm for protons as compared to DSS (0 ppm) and  $-0.12$  ppm for  $^{13}\text{C}$  nuclei (Wishart et al. 1995).  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded at 600 MHz on  $^{15}\text{N}$ - or  $^{15}\text{N}/^{13}\text{C}$ -labeled proteins with 64 scans per increment, with 2048 and 256 points in the proton and nitrogen dimension, respectively. Three-dimensional HNCACB and HNCO experiments were recorded for both the WW<sup>PKA</sup> and the unmodified WW domain as a control, with 8 and 4 scans per increment, respectively, and with 2048 and 94 points in the proton and nitrogen dimensions. Proton and nitrogen spectral widths were 13.9 and 30 ppm centered on 4.7 and 120.0 ppm, respectively. In the carbon dimension, spectral widths of 65.4 and 12.0 ppm centered on 39.7 and

172.5 ppm, sampled with 196 and 112 points, were used for the HNCACB and the HNCOC experiments, respectively. HNHA experiments were recorded with 8 scans, and 2048, 208, 114 points in the  $^1\text{H}$ ,  $^{15}\text{N}$  and  $^1\text{H}$  dimensions, respectively. Spectral widths were 14.0 and 11.9 ppm centered on 4.7 ppm for the proton, and 24.9 ppm centered on 118.5 ppm for the nitrogen dimension.

The chemical shift perturbations ( $\Delta\delta$ ) of individual amide resonances were calculated with the Eq. (1) taking into account the relative dispersion of the proton and nitrogen chemical shifts.

$$\Delta\delta(\text{ppm}) = \sqrt{\Delta\delta(^1\text{H})^2 + 0.2\Delta\delta(^{15}\text{N})^2} \quad (1)$$

A titration with a phospho-peptide was carried out at 293 K with the WW<sup>PKA</sup> domain at 0.2 mM and increasing concentrations of substrate at 0, 0.1, 0.2, 0.4, 0.6, 1, 2, and 4 mM on a 300 MHz-spectrometer equipped with a BBI probe.

A pH titration was performed at 293 K on a 600 MHz-spectrometer on a mixture of  $^{15}\text{N}$ -WW<sup>PKA</sup> and  $^{15}\text{N}/^{13}\text{C}$ -WW domains at a concentration of 0.1 and 0.12 mM, respectively. The acquisition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra was performed at three different pH values (6.8, 5.86 and 4.86) in a 50 mM d<sub>11</sub>-Tris buffer where pH were adjusted with concentrated HCl. The subspectra corresponding to each differentially labeled species were obtained from a sequence which uses a double INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) transfer with isotopical discrimination between  $^{15}\text{N}\{^{12}\text{CO}\}$  and  $^{15}\text{N}\{^{13}\text{CO}\}$  (Golovanov et al. 2007) and water suppression using a watergate (WATER suppression by Gradient Tailored Excitation) sequence (Piotto et al. 1992).

Quantification of phosphorylation level was calculated from peak integration of resonances in two different WW<sup>PKA</sup> samples at pH 6.8. To monitor the effect of line broadening on the determination of the percentage of phospho-WW, resonances of WW and phospho-WW were integrated in  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of a 1:0.8 mixture of WW:WW<sup>PKA</sup> at different pH. Given that the WW<sup>PKA</sup> sample contains both the phospho-WW and WW forms, the determination of the phosphorylation level from  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments of the WW:WW<sup>PKA</sup> mixture follows the Eq. (2), where P is the overall phosphorylation level in the WW: WW<sup>PKA</sup> mixture, p is the phosphorylation level of the WW<sup>PKA</sup> sample, [WW<sup>PKA</sup>] and [WW] are the concentrations of WW<sup>PKA</sup> and WW proteins in the mixture, respectively.

$$p = P \times \frac{[\text{WW}^{\text{PKA}}] + [\text{WW}]}{[\text{WW}^{\text{PKA}}]} \quad (2)$$

P is calculated with the integration of the peaks corresponding to the phosphorylated and the unphosphorylated forms of the WW as in the Eq. (3).

$$P = \frac{A^{\text{phospho-WW}}}{A^{\text{phospho-WW}} + A^{\text{WW}}} \quad (3)$$

$A^{\text{phospho-WW}}$  and  $A^{\text{WW}}$  correspond to the integration of the resonance of a given residue in the phosphorylated and the unphosphorylated form, respectively.

For the 2D NOESY, relaxation and water exchange experiments, WW and WW<sup>PKA</sup> samples were prepared at 660 and 940  $\mu\text{M}$ , respectively, in 50 mM d<sub>11</sub>-Tris.Cl pH 6.0, 25 mM NaCl, 2.5 mM EDTA, 10 % D<sub>2</sub>O. For  $^{15}\text{N}$  relaxation and water exchange experiments, data were acquired at 293 K with spectral widths of 14.0 and 52.0 ppm centered on 4.7 and 106.0 ppm in  $^1\text{H}$  and  $^{15}\text{N}$  dimensions, respectively, with 2048 points in the detection dimension with 8 scans (for T1 measurements), 16 scans (for T2 and  $^{15}\text{N}$ -NOEs measurements) or 32 scans (for water exchange measurements). Two-dimensional  $^1\text{H}$ - $^1\text{H}$  NOESY experiments with 300 ms mixing time, and spectral width of 14.0 and 12.6 ppm in F2 and F1 dimensions, respectively, both centered on 4.7 ppm, were acquired with 2048 points in both dimensions at 293 K with 16 scans and 16 dummy scans. For T1 measurements, mixing times of 10, 100, 200, 400, 600, 800, 1,000, 1,200, 1,500 and 2,000 ms were used. For T2 measurements, mixing times of 16, 32, 63, 95, 126, 158, 189 and 252 ms were used. A proton presaturation delay of 4 s was used for  $^{15}\text{N}$ -heteronuclear NOE experiments. The water exchange rate of backbone amide protons and arginine H $\epsilon$  side chain protons have been measured using the phase-modulated CLEAN chemical Exchange (CLEANEX-PM) experiment (Hwang et al. 1998) with mixing times of 1, 3, 8, 16, 32, 48, 64, 81, 97, 114, 130, 146, 162, 178, 194 ms. The exchange rate constants k were derived from the equation as described (Hwang et al. 1998). Data were fitted to the Eq. (4) where  $R_{1\text{water}}$  is the longitudinal relaxation rate of water,  $R_{1\text{app}}$  the apparent longitudinal relaxation rate and k the exchange rate constant. Here, the water R1 was set to 9 s<sup>-1</sup> ( $T_{1\text{water}} = 111$  ms) allowing the calculation of the k values (Figure S1) for each exchangeable amide proton or arginine side chain.

$$\text{intensity}(t) = \text{Cte} \times \left[ \exp^{-R_{1\text{water}} \cdot t} - \exp^{-(R_{1\text{app}} + k) \cdot t} \right] \quad (4)$$

Trypsin digests of WW and PKA-phosphorylated WW domains, and MALDI-TOF MS analyses

Samples of WW domain in its un- and PKA-phosphorylated forms were dissolved at 1 mg/ml in 50 mM ammonium bicarbonate and incubated for 1, 2, 6 or 18 h at 37 °C with 1:100<sup>o</sup> trypsin (w/w). Samples were then lyophilized, dissolved in water and analyzed with the  $\alpha$ -cyano-4-hydroxycinnamic acid matrix. MALDI-TOF MS analyses were performed on a Voyager DE-STR spectrometer



(Applied Biosystems). Detection was carried out in a positive ion mode for the unmodified WW domain and both positive and negative modes for the phosphorylated WW samples.

#### Molecular dynamics simulation of the phospho-binding loop of the phospho-WW domain

Simulations were performed using the CAChe software (Fujitsu Limited, Tokyo) with the MM3 force field. Starting from the NMR structure of the apo-WW domain of Pin1 (PDB ID: 1I6C) (Wintjens et al. 2001) and incorporation of the di-anionic phosphoryl group on S16, preliminary energy minimization was accomplished with 1,000 steps of conjugate-gradient algorithm on the protein loop from residues M15–Y23, the rest of the protein being locked. A first molecular dynamics (MD) simulation was then performed *in vacuo* for 50 ps, whereas the final simulated annealing steps (20 ps at high temperature and 50 ps at lower temperature) were carried out using an explicit water model.

## Results

### Identification of the PKA phosphorylation site

The phosphorylated form of the Pin1 WW domain, hereby called the phospho-WW domain, was obtained *in vitro* by incubation of the WW domain with purified PKA enzyme for 8 h at 30 °C followed by a purification of the crude phosphorylation mixture by reverse phase chromatography. Analysis of the purified WW<sup>PKA</sup> domain by MALDI-TOF MS indicates the presence of a single phosphorylation site (Fig. 1a), but does not identify the modified residue. MALDI-TOF MS analysis of trypsin digests performed on both the unmodified and PKA-phosphorylated WW domains (Fig. 1b) on the other hand confirm the phosphorylation of S16 (Lu et al. 2002). Signals at *m/z* values of 393.14 and 549.27 Da corresponding to MS<sup>16</sup>R and RMS<sup>16</sup>R peptides around the S16 residue, respectively, are detected in the unmodified WW domain, the former being also detected in the PKA-phosphorylated WW sample. Signals of peptides with a mass increment of +80 Da were detected only in the PKA-phosphorylated WW sample both in positive and negative ion modes and match with the incorporation of a phosphate group. These results unambiguously confirm the presence of a phosphate group on the S16 residue (Fig. 1; Table 1). It should be noted that in case of phosphorylation on S18 or S19 we would not have been able to distinguish one or the other by this method, as trypsin digest could not generate individual peptides in the absence of an arginine residue separating both serine.

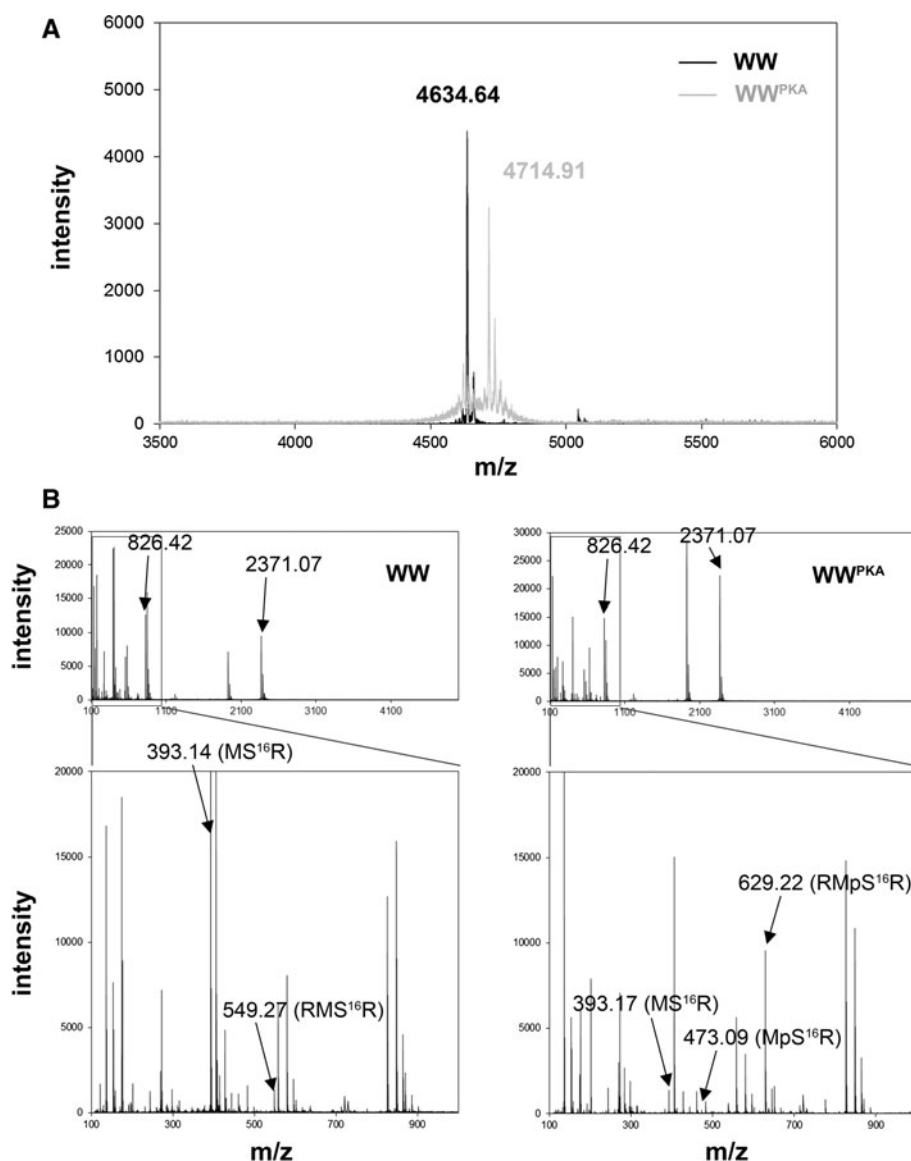
The <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of the WW<sup>PKA</sup> domain shows two sets of resonances corresponding to the phosphorylated and the unphosphorylated forms of the WW domain. The subspectrum of the unmodified form is significantly weaker confirming that phosphorylation by PKA is almost complete. A mapping of the chemical shift perturbations upon PKA phosphorylation was performed at a pH value of 5.86 (Fig. 2) since severe broadening of the WW resonances at pH values higher than 6.4 leads to a loss of R17 to S19 resonances. Line broadening of these loop resonances and changes related to the phosphorylation state of the WW domain have been further analyzed into details (see below). The chemical shift mapping shows that several residues, grouped in two main regions are affected (Fig. 2).

A first region is the whole phospho-binding loop between β1 and β2 strands, encompassing residues M15 to R21. A second group contains resonances corresponding to residues Y23 and E35, localized in the β2 strand and in the C-terminal loop. In peptides or disordered proteins, phosphorylation induces a large downfield shift of the amide protons of the phosphorylated residues (Theillet et al. 2012), allowing for their easy detection in an otherwise empty region of the <sup>1</sup>H-<sup>15</sup>N HSQC spectrum. However, in the case of a folded protein such as the WW domain, the corresponding resonances still remain inside the spectral window of amide backbone resonances, and a full assignment of the spectrum is therefore necessary. Based on this assignment, we unexpectedly find that the <sup>1</sup>H-<sup>15</sup>N largest chemical shift perturbations are detected for S18 and S19 residues and not for S16 (Fig. 2). Shifts of the <sup>1</sup>H-<sup>15</sup>N cross-peaks compared to the unmodified domain hence cannot unambiguously identify the phosphorylation site of the WW domain.

Comparison with the random coil carbon chemical shift values of either the un- or phosphorylated serine/threonine equally does not lead in a straightforward manner to the identification of the phosphorylation site. As all three serine residues of the loop (S16, S18 and S19) exhibit deviations in <sup>13</sup>Cα values from their random coil values (Fig. 3), it precludes the use of the <sup>13</sup>Cα value as a reliable sensor to probe the phosphorylation site. However, a clear difference in the <sup>13</sup>Cβ chemical shift of the sole S16 residue between both states points to the latter as the phosphorylation site (Fig. 3). Therefore, in contrast to the situation of IDPs, where the amide proton chemical shift and/or the Cα/Cβ values are unambiguous parameters for identifying the phosphorylated residue, in the case of the folded WW domain, the Cβ value of the phospho-S16 acts as the unique identifying parameter. NMR and MS data thereby confirm the previously identified S16 residue as the unique PKA phosphorylation site within the Pin1 WW domain (Lu et al. 2002).



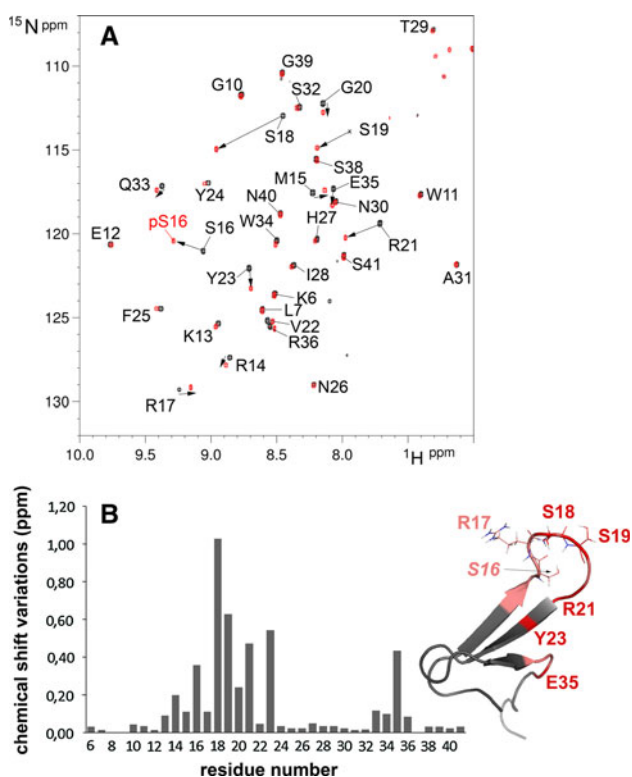
**Fig. 1** Detection and identification of the PKA phosphorylation site on the WW domain by mass spectrometry and trypsin digest. **a** MALDI-TOF mass spectra of the non- (black) and PKA-phosphorylated (grey) forms of the  $^{15}\text{N}/^{13}\text{C}$ -labeled WW domain. **b** MALDI-TOF MS analyses of trypsin digests of the unphosphorylated (WW) and PKA-phosphorylated ( $\text{WW}^{\text{PKA}}$ ) WW domains in positive ion mode showing the  $m/z$  region from 100 to 5000 Da (upper panels) and zooms from 100 to 1000 Da (lower panels). The molecular ions annotated are  $[\text{M} + \text{H}]^+$  species. Peaks are labeled with their monoisotopic masses



**Table 1** Summary of trypsin fragments taking into account no missed cleavage or one missed cleavage in the region of interest

Non phosphorylated		Phosphorylated		Primary sequence	Number of missed cleavage
Average	Monoisotopic	Average	Monoisotopic		
290.319	290.159	–	–	GSK	0
392.479	392.184	472.459	472.150	$\text{MS}^{16}\text{R}$	0
405.411	405.197	485.391	485.163	$\text{S}^{18}\text{S}^{19}\text{GR}$	0
548.667	548.285	628.647	628.251	$\text{RMS}^{16}\text{R}$	1
779.875	779.370	859.855	859.337	$\text{MS}^{16}\text{RS}^{18}\text{S}^{19}\text{GR}$	1
825.963	825.438	–	–	LPPGWEK	0
2370.523	2369.082	–	–	VYYFNHITNASQWERPSGNS	0

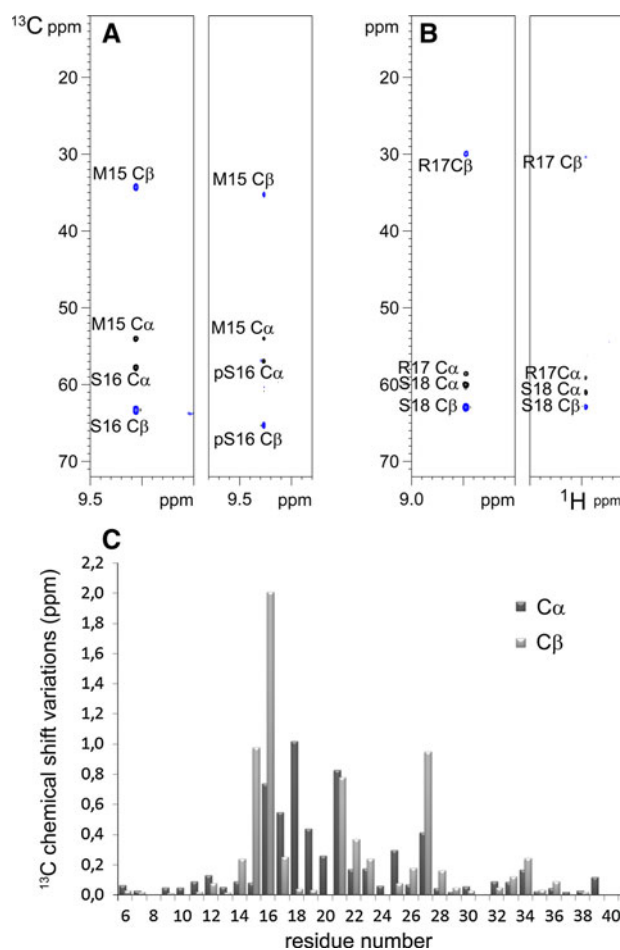
Average and monoisotopic masses are given. Masses of the corresponding fragments in their phosphorylated form were calculated only for those from the loop containing the three potential phosphorylated serine residues at positions 16, 18 or 19



**Fig. 2** Mapping of the PKA phosphorylation site of Pin1 WW domain. **a**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the WW (black) and the PKA-phosphorylated WW (red) domain acquired at pH 5.86 allowing the detection of the loop resonances otherwise unobservable at higher pH in the unphosphorylated form. **b**  $^1\text{H}$ - $^{15}\text{N}$  combined chemical shift variations upon PKA phosphorylation of WW domain along the primary sequence, calculated using Eq. (1) and mapped on the WW 3D structure (PDB ID 116C). Residues with variations greater than 0.4 ppm are colored in red and those comprised between 0.1 and 0.4 ppm in pink

#### Alternative detection of phosphorylation in WW domain and identification of phosphorylation-related residues by pH titration

The phosphate group carried by serine or threonine residues can experience different ionization states. The transition from the mono- to the dibasic form is pH-dependent and characterized by a pKa value that falls around 6.03 for phospho-serine and 6.34 for phospho-threonine (Bienkiewicz and Lumb 1999). Thus, NMR spectra of phosphorylated peptides undergo pH-dependent shifts attributed to protonation changes of the phospho-S/T side chain. Such pH dependent variations of the amide proton chemical shift have proven valuable indicators of phosphorylation in peptides derived from the Erk activation loop (Prabakaran et al. 2011). We therefore tried the same strategy to characterize the phospho-WW domain. We prepared a mixture of unmodified  $^{15}\text{N}/^{13}\text{C}$ -labeled and PKA-phosphorylated  $^{15}\text{N}$ -labeled WW domains to ensure identical pH conditions during the pH titration. For each pH



**Fig. 3** Effect of phosphorylation on  $^{13}\text{C}$  chemical shifts. **a** and **b**  $^{13}\text{C}$  strips from HNCACB experiments acquired on  $^{15}\text{N}/^{13}\text{C}$ -WW (**a** and **b**, left panels) and on  $^{15}\text{N}/^{13}\text{C}$ -WW<sup>PKA</sup> (**a** and **b**, right panels) for Ser16 (**a**) and Ser18 (**b**) residues. **c** Graphical representation of differences of  $^{13}\text{C}\alpha$  (dark grey bars) and  $^{13}\text{C}\beta$  (light grey bars) chemical shifts between the non- and the PKA-phosphorylated forms of the WW domain along the primary sequence

value (i.e. pH 6.8, 5.86 and 4.86), a  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum with isotopic discrimination between  $^{15}\text{N}\{^{12}\text{CO}\}$  and  $^{15}\text{N}\{^{13}\text{CO}\}$  was recorded enabling the extraction of sub-spectra corresponding to each differentially labeled species. Effect of pH changes on  $^1\text{H}$ - $^{15}\text{N}$  cross-correlation peaks concerns two regions of the WW domain. First, residues around H27 exhibit pronounced chemical shift modifications with identical behavior in both proteins, in agreement with histidine ionization being pH-sensitive in this range of pH values (Fig. 4, left panels). These His-linked pH perturbations are also visible in both proteins in the region center around residue 13 that is in spatial proximity to H27. When we compared the  $^1\text{H}$ - $^{15}\text{N}$  chemical shifts in the HSQC spectrum of the phospho-WW domain, we observed that resonances in the phospho-binding loop are specifically affected by pH changes between 4.86 and 5.86. We were not able to extend the comparison for pH changes towards pH

values of 6.8 for all residues in this loop due to their broadening in the WW domain (Fig. 4). Together with residues S16–R21 of the loop, both Y23 and E35 were also specifically perturbed upon pH titration (Fig. 4, right panels). The reduced changes between pH 6.8 and 5.86 and the marked shift when going from pH 5.86 to 4.86 indicate that the pK<sub>a</sub> value of phospho-S16 in the WW domain is comprised in this latter range. However, phospho-S16 itself does not exhibit the largest shift when compared to amide resonances of S18, S19, G20 and R21. These data show that a pH titration experiment leaves a potential ambiguity to detect and identify phosphorylation events in folded proteins, as resonances of other residues can be even more affected by changes of the phosphate protonation state.

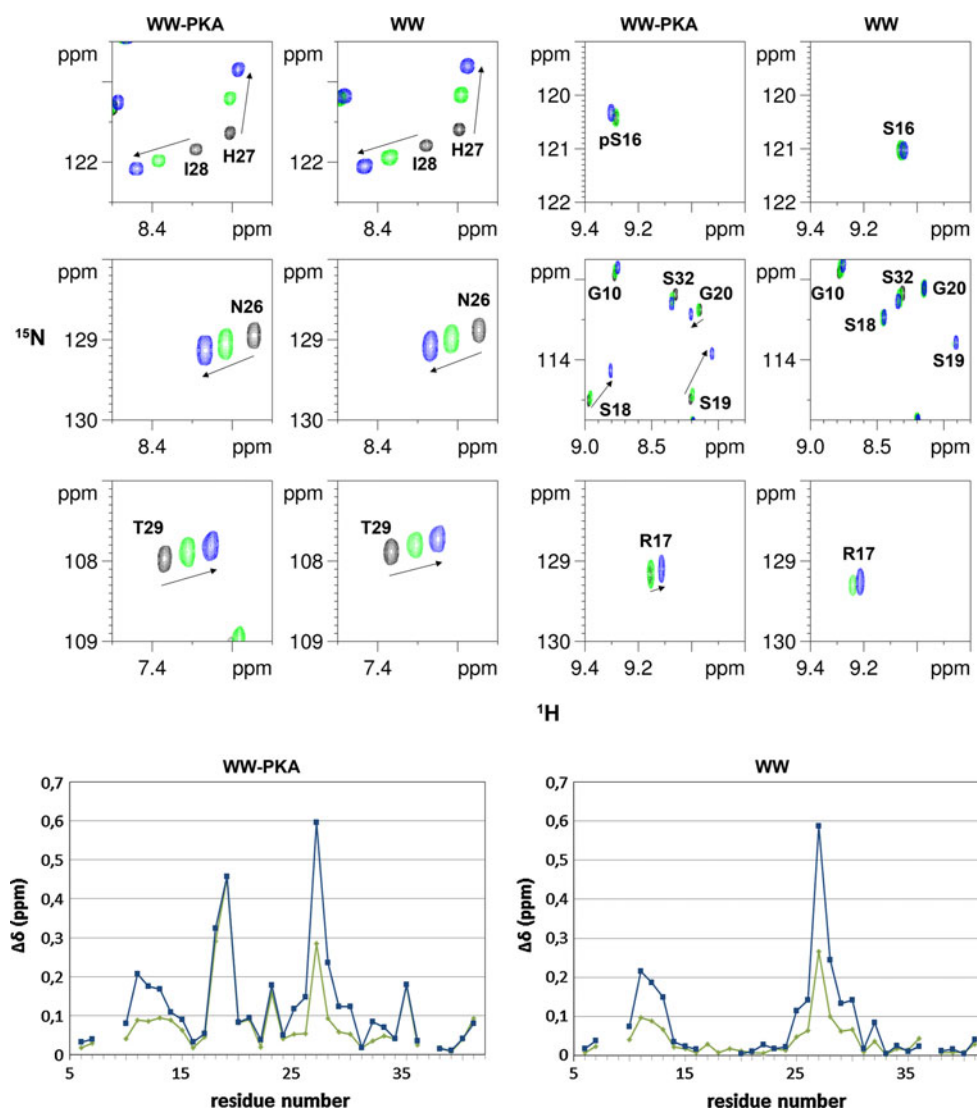
#### Quantification of the phosphorylation level

In the WW<sup>PKA</sup> sample, integration of the MALDI-TOF MS signals corresponding to WW and phospho-WW at m/z of

4,634.64 and 4,714.91, respectively, gives 85.3 % of phosphorylated form (see Fig. 1a). The MALDI-TOF MS analysis of trypsin digests that identified the phosphorylation site on the phospho-WW domain did not give a similar level of phosphorylation (Fig. 1b). Phosphorylation potentially has an influence on both the trypsin digestion profile and the ionization properties of peptides, and these factors preclude comparison of the MS signals of modified and unmodified peptides as a basis for quantification of the phosphorylation level of the folded protein.

One of the advantages of NMR spectroscopy is its ability to extract peak intensities or integrals for individual resonances in the 1D or 2D HSQC spectra. This can in principle be used to estimate the stoichiometry of the PTM. Here, we investigate this approach on a folded protein domain by measuring the relative integrals for peaks corresponding to the WW and phospho-WW domain in the HSQC spectra of the PKA-phosphorylated WW domain. Broadening of resonances of the loop region in the WW

**Fig. 4** pH dependency of the phospho-residue and its neighbours in the phospho-WW domain. (*Upper panels*) Overlay of <sup>1</sup>H-<sup>15</sup>N HSQC spectra of either the WW<sup>PKA</sup> or the WW domain at three different pH: 6.8 (*black*), 5.86 (*green*) and 4.86 (*blue*). *Left panels* residues surrounding H27 that have a similar behavior upon pH variations for both proteins. *Right panels* residues surrounding S16 that exhibit a distinct behavior upon pH variations in phospho-WW and WW proteins. (*Lower panels*) Graphical representation of chemical shift displacements for phospho-WW (*left*) and WW (*right*) domains upon pH variation from 4.86 to 5.68 (*green curves*) and from 4.86 to 6.8 (*blue curves*). Residues without data points correspond either to proline residues (position 8, 9 and 37) or residues of the loop (from R17 to S19) that are too broad to be detected at pH 6.8



domain posed an obvious challenge for the quantification, as one could indeed conclude on the basis of these peaks that the sample only contains phospho-WW (Fig. 5). The line broadening is moreover pH-dependent, because lowering the pH restores some intensity for peaks corresponding to the loop residues in the unmodified WW domain. When we take peaks that are distinguishable between both forms but that are further away from the loop, we arrive at an overall phosphorylation level of  $90 \pm 3$  % phosphorylation in the WW<sup>PKA</sup> domain, in agreement with the MS data on the intact domain. In the <sup>1</sup>H-<sup>15</sup>N HSQC of binary mixture (1:0.8 WW:WW<sup>PKA</sup>) used to monitor the pH titration, and using the spectrum at the lowest pH, we estimated in a similar manner a phosphorylation level of  $35 \pm 5$  % (Fig. 5b) that indicates a phosphorylation level of  $80 \pm 10$  % for the initial WW<sup>PKA</sup> sample.

#### Effect of WW phosphorylation on its binding activity

The WW domain of Pin1 binds phosphorylated peptide substrates, with a marked specificity for phospho-S/T-P motifs generated by proline-directed kinases. In a previous work, we have evaluated the dissociation constants of several peptides from the neuronal Tau protein for either the full-length Pin1 or the isolated WW domain, and found no selectivity of Pin1 binding for a particular Tau phospho-epitope (Smet et al. 2004). Here, we measure by NMR spectroscopy the binding constant of phospho-WW for a phospho-peptide of the human Tau protein (residues 208–221) phosphorylated on residue T212. We performed a titration of <sup>15</sup>N-labeled WW<sup>PKA</sup> domain at a constant concentration of 200  $\mu$ M with increasing amounts of the phospho-peptide. The dissociation constant of the complex was beyond the values accessible by NMR measurements which would need concentrations largely above the limit of solubility of the peptide to reach saturation. In contrast, a binding constant of 100  $\mu$ M had been measured for the WW domain under the same conditions (Smet et al. 2004) (Fig. 6). Phosphorylation of the WW hence leads to a loss of function on phospho-substrate binding, thereby rationalizing the loss of Pin1 function upon PKA phosphorylation (Lu et al. 2002).

#### Dynamical and structural changes of the phospho-binding loop upon phosphorylation

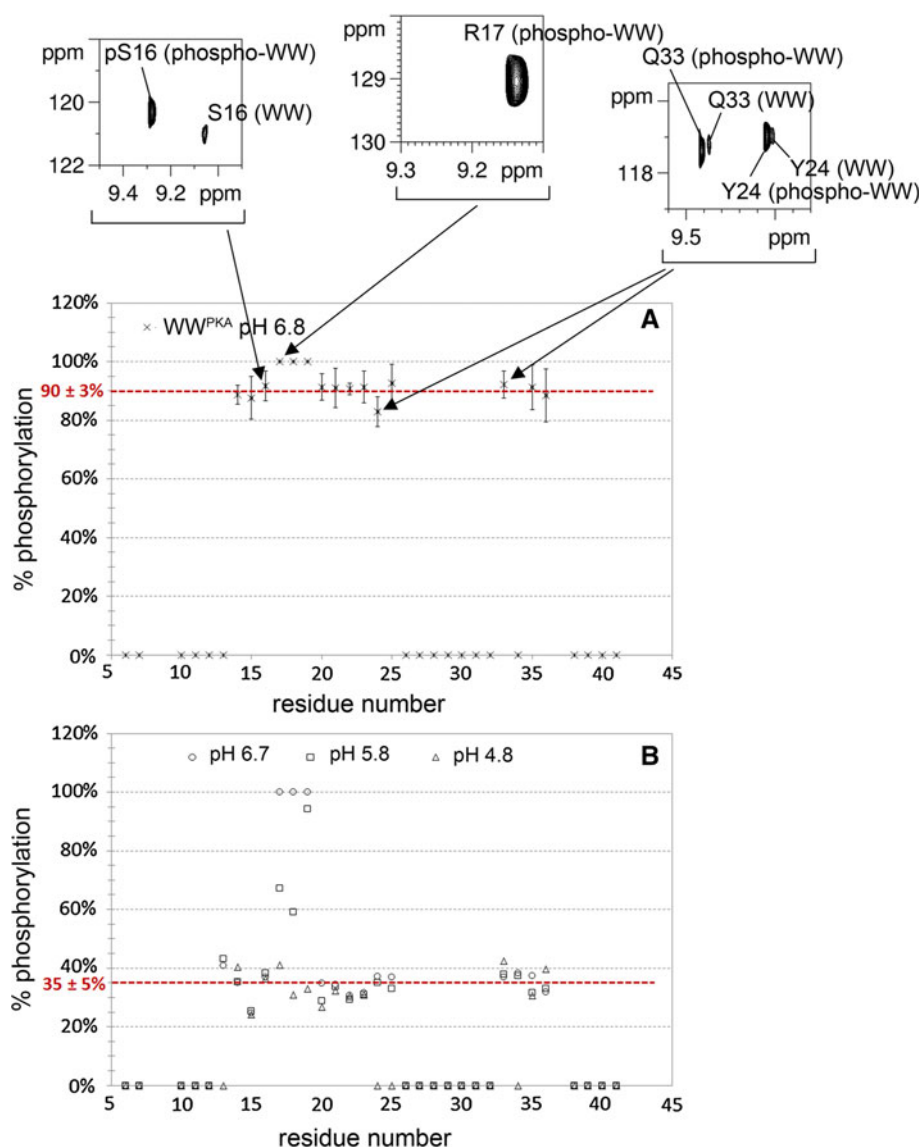
To explore the relationship between WW phosphorylation and its loss of binding capacity, we first considered the structural consequences of phosphate incorporation at position S16. Based on the comparison of C $\alpha$  and C $\beta$  secondary chemical shifts and <sup>3</sup>J H<sub>N</sub>-H $\alpha$  couplings between the unphosphorylated and the PKA-phosphorylated forms of the WW domain (Fig. 6), we conclude that the triple  $\beta$ -sheet

structure is maintained upon PKA phosphorylation. Comparison of the NOE patterns of phospho-WW and WW additionally indicates that no major structural change is induced by phosphorylation (Fig. 7a). Without obvious structural modifications of the WW backbone upon phosphorylation, we compared the relaxation parameters at pH 6.0 (Fig. 8) where all resonances in both WW and phospho-WW domains <sup>1</sup>H-<sup>15</sup>N HSQC can be detected. As we estimated the pK<sub>a</sub> of phospho-S16 between pH 4.86 and 5.86, the phospho-S16 is present in its physiological di-anionic state at pH 6.0.

<sup>15</sup>N-R1 and <sup>1</sup>H-<sup>15</sup>N heteronuclear NOE values of the backbone amide resonances allow distinguishing the strands from the loop regions, but are not significantly different upon phosphorylation. Whereas this is in agreement with the conservation of the overall fold upon phosphorylation, the slight decrease of the overall <sup>15</sup>N-R1 values points towards a higher compactness of the phospho-WW (Fig. 8). However, the selective line broadening of the loop resonances in the unmodified domain is reflected in the selective decrease in <sup>15</sup>N-R<sub>2</sub> values of those same loop residues in the phosphorylated WW domain at pH 6.0 (Fig. 8). Because the overall tumbling is not expected to change in view of the overall structure conservation, phosphorylation of S16 hence seems to result in a lessening of exchange broadening on the microsecond-millisecond timescale. This effect is not limited to next neighbours in the WW binding loop but equally expands throughout the  $\beta$ 1 strand (Fig. 8).

Comparative measurements of water exchange rates of backbone amide and side chain protons using the phase-modulated CLEAN chemical EXchange experiment show a significant decrease of water exchange upon phosphorylation for most backbone amides from the phospho-binding loop (R17 to S19) but also for the side chain protons of the arginine residues (Fig. 8 and see Supplementary Materials, Figure S1). These reduced water exchange rates could be solely due to the introduction of negative charges in the binding loop of the phosphorylated domain. However, part of the observed changes in water exchange rates could also reflect reduced solvent accessibility accompanying the conformational stabilization. Rearrangement of the hydrogen network previously described for the WW domain when bound to its substrate (Peng et al. 2007) is a third factor in agreement with the reduced water exchange we observe. Both the decrease of microsecond-millisecond movements and the slower water exchange for backbone amides of residues in the binding-loop, and more specifically for residues S16, R17, S18 and S19, account for the decreased line-width of the corresponding resonances in the <sup>1</sup>H-<sup>15</sup>N HSQC of the phosphorylated domain.

A model of the phospho-WW domain (Fig. 7) shows compaction of the binding-loop region, as a consequence of



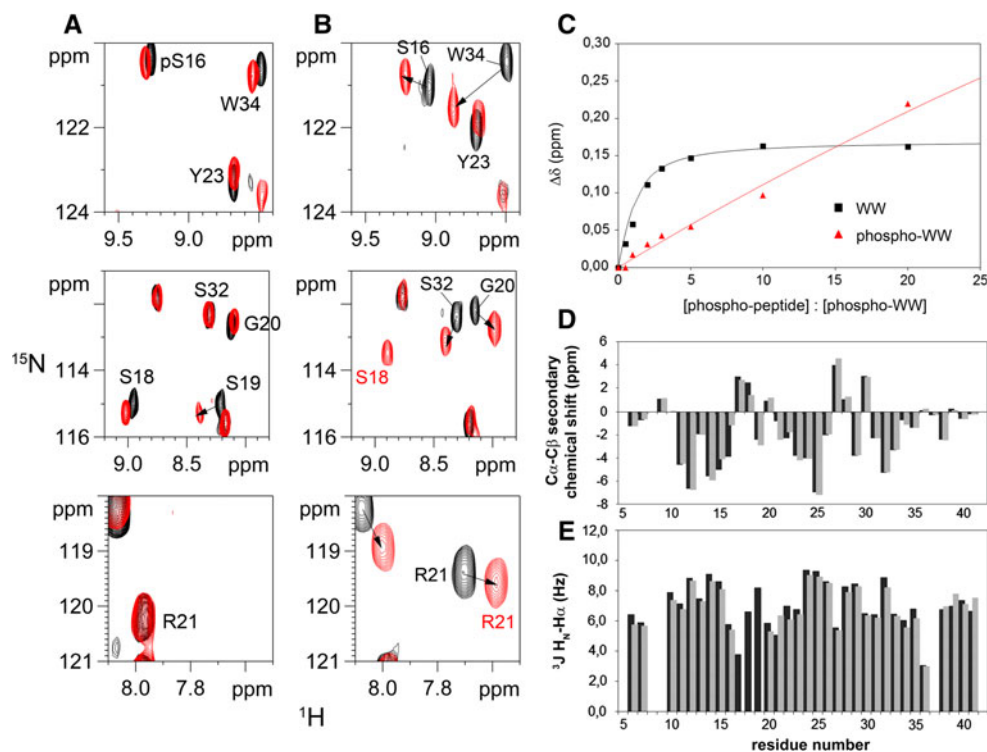
**Fig. 5** Quantification of PKA phosphorylation level in WW domain by NMR. Determination of the phosphorylation level based on integrations of signals corresponding to WW and phospho-WW in the  $\text{WW}^{\text{PKA}}$  sample at pH 6.8 in  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments (a) or in  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments of a WW:  $\text{WW}^{\text{PKA}}$  mixture (ratio 1: 0.8) acquired at different pH (b). a Resonances of S16, R17, Y24 and Q33 residues from the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of the  $\text{WW}^{\text{PKA}}$  sample corresponding either to the phosphorylated (phospho-WW) or unphosphorylated (WW) forms are shown. Calculation of phosphorylation level was done for two different phosphorylation reactions with PKA and mean values are indicated for each residue. Zero values are arbitrarily assigned to residues that have the same chemical shift in both the unphosphorylated and PKA-phosphorylated forms. For

residues of the phospho-binding loop (R17–S19), despite different chemical shifts in both forms, the absence of detectable signals at pH 6.8 in the unphosphorylated form lead to an apparent phosphorylation level of 100% based on these residues. b Data acquired at pH 6.7, 5.86 and 4.86 are depicted by open circles, squares and triangles, respectively. The mean phosphorylation level and root mean square deviation were calculated with data excluding values corresponding either to residues that give the same resonances in both WW and  $\text{WW}^{\text{PKA}}$  or residues of the  $\beta$ -hairpin loop that experience severe line broadening (R17–S19). The mean phosphorylation level of 35 ± 5% was calculated at pH 6.7 and similar values were found at pH 5.86 or 4.86

novel H-bonds between the phosphate moiety and the side chains of R17, R21 and Y23. These could account for the conformational stabilization as witnessed by reduced  $^{15}\text{N}$ -R2 rates and for the reduction of water exchange rates for residues in the binding loop. Moreover, conformational rearrangements of the phospho-binding loop and R21 side

chain in this model are in good agreement with the significant chemical shift perturbations observed for the S18, S19, R21 and Y23  $^1\text{H}$ ,  $^{15}\text{N}$  resonances upon PKA phosphorylation (Fig. 2). The novel H-bond of the side chain of R21 to the phospho-S16 in our model will turn it away from the side chain of E35, and hence could additionally





**Fig. 6** Binding activity of the phospho-WW domain of Pin1 to phospho-substrates. **(a, b)** Titration of the  $^{15}\text{N}$ -labeled WW<sup>PKA</sup> **(a)** and WW **(b)** domains with a peptide from human Tau protein (residues 208–221) containing the phospho-Thr212-Pro motif (Smet et al. 2004).  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra depicted in *black* represent the apo proteins, WW<sup>PKA</sup> **(a)** or WW **(b)**, spectra in *red* represent the proteins in the presence of a 20-fold excess of phospho-peptide. **c** Titration curves for the non- (*black*) and PKA-phosphorylated (*red*) WW

explain the chemical shift perturbation observed for the E35  $^1\text{H}$ ,  $^{15}\text{N}$  upon phosphorylation of S16. As the pH value will determine the value of the net charge of the phospho-S16 residue, we furthermore observe for all these residues networked to the phospho-serine that the amide  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts titrate with pH (Fig. 4). The conformation of the R17, S18, S19 and Y23 residues in the phospho-WW is similar to the one they adopt in the loop of the WW domain bound to a phospho-peptide substrate (Fig. 7d and see Supplementary Materials, Figure S2). In support of this, chemical shift variations for residues R17 to S19 are similar in direction when we compare the phospho-WW domain and the WW domain complexed to a phospho-peptide (see Figs. 2a and 6b for comparison). They are however distinct for Y23 and R21, which can be rationalized by the role that Y23 plays in accommodating the proline moiety of the phospho-S-P ligand (Verdecia et al. 2000; Wintjens et al. 2001) while its side chain directly points to the S16 phosphate moiety in our model structure. Equally, the R21 side chain conformation in the structure of the complex points away from the phosphorylated serine side chain of the peptide ligand (Verdecia et al. 2000;

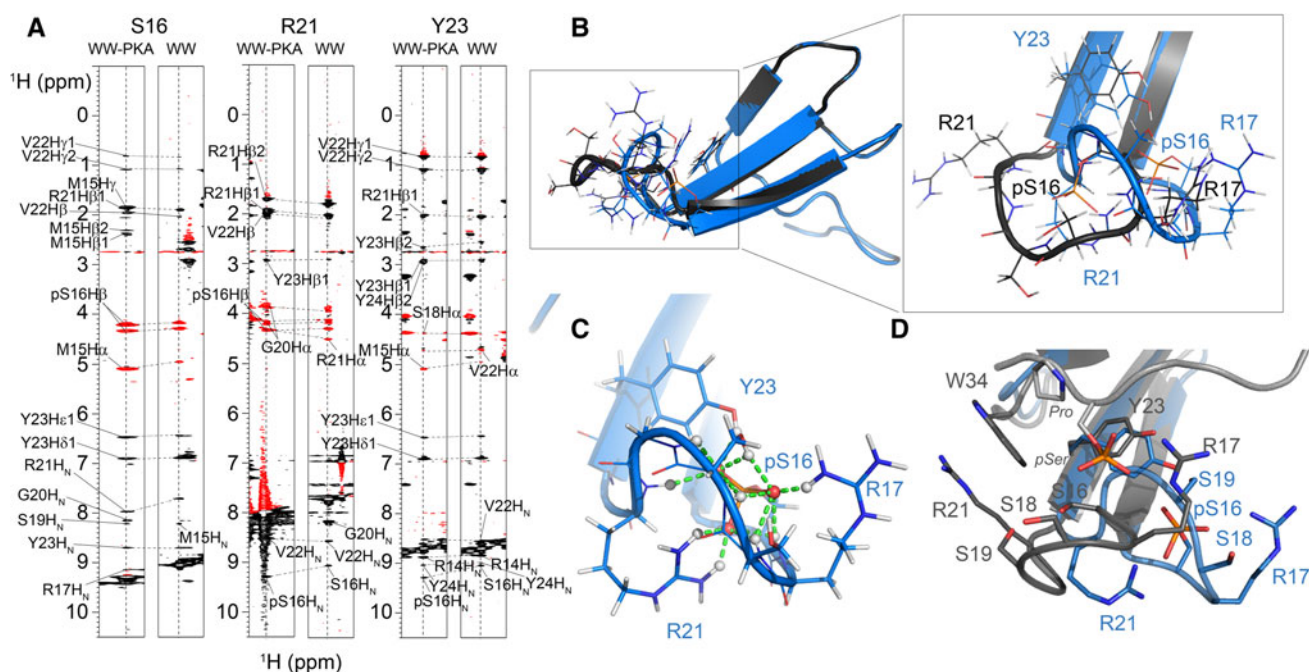
domain allowing for the calculation of dissociation constants. **d**  $\text{C}\alpha - \text{C}\beta$  secondary chemical shifts of the phospho-WW and unmodified WW domains.  $^{13}\text{C}$  chemical shifts values used as references are described in (Wishart and Sykes 1994). **e**  $^3\text{J}$  scalar couplings between backbone amide and alpha protons calculated from HNHA experiments. Data of phospho-WW domain are represented in *black bars* and those of unmodified WW domain in *grey bars*

Wintjens et al. 2001) whereas in our model, it is involved in the positioning of the phospho-S16 side chain.

## Discussion

To better understand the mechanism of regulation of Pin1 function through phosphorylation of its WW domain, and increase our general understanding of functional regulation by phosphorylation, we present here a detailed structural and functional study of its phosphorylated WW domain. At first, we evaluated the use of NMR to detect, identify and quantify the phosphorylation event after *in vitro* phosphorylation by the PKA kinase. Although phosphorylation of globular proteins is mostly directed to disordered regions or flexible loops, the fact that those are embedded in the three-dimensional structure renders the characterization of phosphorylation less straightforward than in case of IDPs (Landrieu et al. 2006; 2010; Lippens et al. 2008; Theillet et al. 2012). Indeed, we found that no significant shift of  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  goes significantly beyond the usual range observed in folded proteins, and hence can





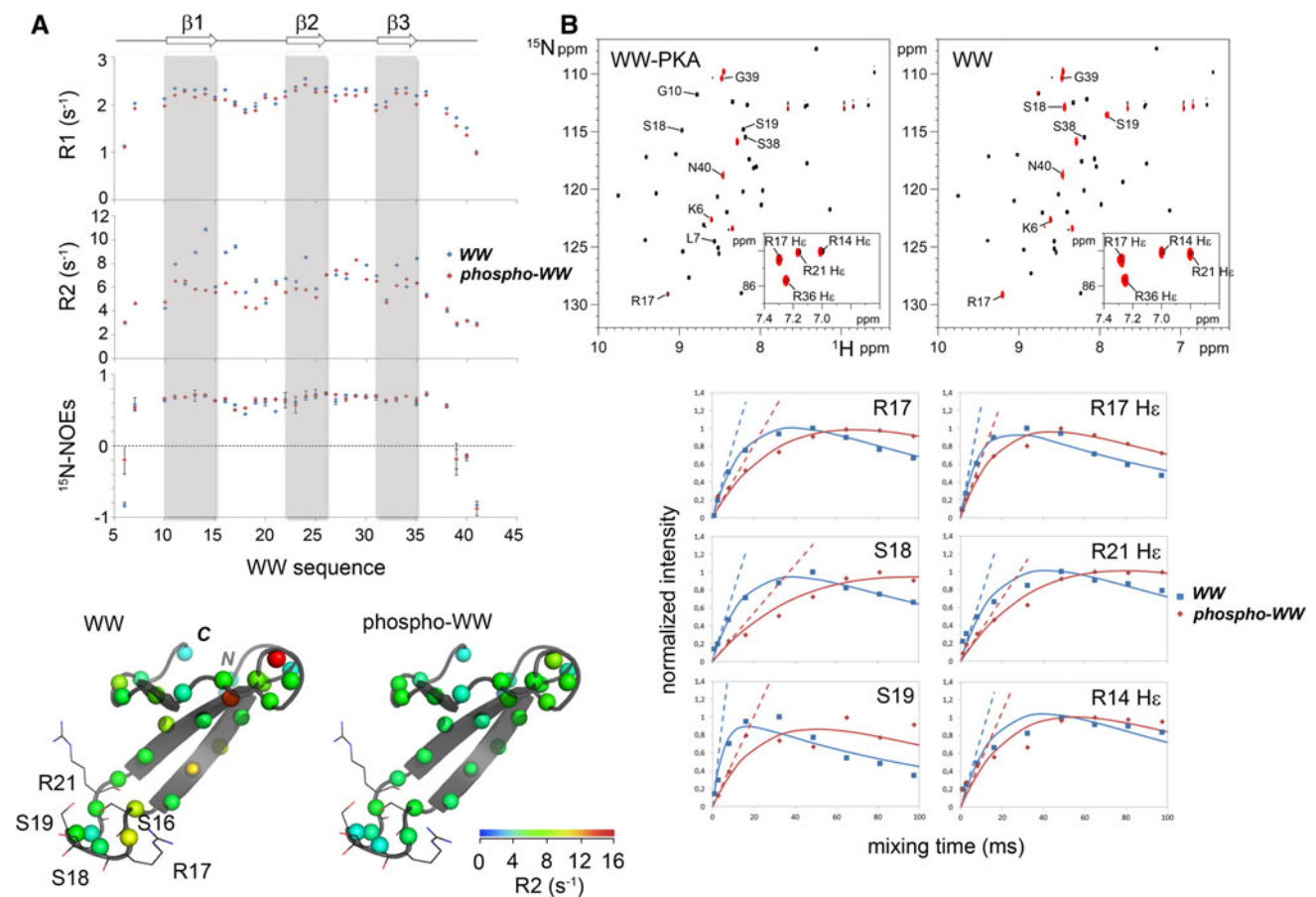
**Fig. 7** Structural model of the phospho-WW domain. **a** Comparison of  $^1\text{H}$  strips from the  $^1\text{H}$ - $^1\text{H}$   $^{15}\text{N}$ -filtered NOESY experiments for the residues S16, R21 and Y23 of the WW (*right panels*) and the WW-PKA domains (*left panels*). **b** Comparison of the structure of the WW (*grey*, PDB ID : 1I6C) (Wintjens et al. 2001) and the model structure of the phospho-WW based on the structure of the unphosphorylated form and molecular dynamics simulations restrained to M15 to Y23 residues (*blue*) highlighting a compaction of the phospho-binding loop upon PKA phosphorylation. **c** Close-up of the phospho-binding loop of the phospho-WW domain showing the intramolecular contacts with the phosphate moiety. Both R17 and R21 side chains guanidium moieties establish intramolecular contacts with

the phospho-S16 together with the phenol group of the Y23 that probably participate to the compactness of the phosphorylated WW domain. This conformation explains the decrease of water exchange rates for the residues of the loop in the phospho-WW as compared to the WW domain as well as the overall R1 decrease and the reduced R2 values of the residues of the loop. **d** Comparison of the modeled structure of phospho-WW (*blue*) with the structure of the WW domain bound to a phospho-peptide (*grey*, PDB ID: 1F8A) (Verdecia et al. 2000). In this latter structure, residues of the WW domain are indicated in regular *grey* characters and those of the phospho-S-P motif of the peptide substrate in *italic*

unambiguously identify a specific serine residue as the phosphorylation site. The  $^1\text{H}$ - $^{15}\text{N}$  chemical shift mapping when comparing both the unmodified and PKA-phosphorylated WW resonances only highlighted the phospho-binding loop as the region containing the phosphorylation site. However, all the three serine residues present in this loop show pronounced backbone amide chemical shift changes upon PKA phosphorylation, with S18 residue exhibiting the largest chemical shift variations. This clearly argues against the identification of the phospho-site on the basis of the sole  $^1\text{H}$ - $^{15}\text{N}$  chemical shift mapping (Fig. 2). The site identification came from the precise comparison of  $^{13}\text{C}$  values between the unmodified and phosphorylated WW domains, whereby the  $\text{C}\beta$  of S16 appears to be solely affected by phosphorylation (Fig. 3). Whereas this result is in agreement with MS, we should notice that this latter approach would not necessarily be able to discriminate between S18 and S19 residues as putative phosphorylation sites, given their adjacent positions. While peptide-based MS is conditioned by the presence of a protease cleavage site between the putative phosphorylation positions, NMR

identification based on the characteristic shift of the phospho-serine  $^{13}\text{C}\beta$  carbon should function independently of the primary sequence, in a non-destructive manner.

Since  $^1\text{H}_\text{N}$  and  $^{15}\text{N}$  chemical shifts are not valuable *indicators* of phosphorylation events in folded proteins, we investigated whether the pH dependency of protein resonances in the range of pH around the pKa value of phospho-S/T residues could play this role. In the absence of stable tertiary structure, the intramolecular hydrogen bond between the phospho-serine phosphate moiety and its own amide hydrogen is the underlying factor that determines the pH dependence of the amide proton chemical shift (Du et al. 2005). In the case of the WW loop, we observe the most pronounced pH dependence for the amide proton of S18, suggesting that the tertiary structure imposes an inter- rather than intra-residue hydrogen bond (see the modeled structure, Fig. 7c). As a result, the pH titration behaviour of the backbone amide resonance is a second invalid parameter to identify the site of phosphate incorporation. Finally, NMR as a tool to measure the stoichiometry of the PTM equally requires some caution in the case of a folded



**Fig. 8** Dynamics of the phospho-WW domain. **a** Comparison of dynamics parameters of the WW and the phospho-WW domains. *Upper panel*, representation of R1, R2 and  $^{15}\text{N}$ -NOEs values along the Pin1 WW primary sequence for WW (blue) and WW-PKA (red) domains. Dynamics parameters were measured at pH 6.0 allowing the detection of the resonances of the phospho-binding loop (residue 16–21) in both samples. *Arrows* above the graphics indicate the position of each  $\beta$ -strand. Note that the R2 values for N26 and T29 residues in the WW domain are not shown and are of 14.5 and 16.5  $\text{s}^{-1}$ , respectively. *Lower panel*, the R2 values color coded on the WW structure (PDB ID 1F8A) where each sphere is a backbone amide nitrogen. **b** Water exchange measurements using the CLEAN

chemical EXchange (CLEANEX) experiments. *Upper panel*, comparison of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC of the WW and WW-PKA domains (black) with the CLEANEX experiments with a mixing time of 50 ms (red). *Lower panel*, normalized intensity variations of the signals of WW (blue) versus phospho-WW (red) for different mixing times for the loop NH signals (R17, S18, S19) and the H $\epsilon$  arginine side chains (R17, R21 and R14). Fitted CLEANEX exchange curves (plain lines) were calculated as described (Hwang et al. 1998). Initial velocities (dotted lines) were calculated from fitted curves allowing the determination of the water exchange rate parameter  $k$  (see Supplementary Materials, Figure S1)

domain, especially when it induces modifications such as the loop stabilization and decrease in water exchange rate of backbone amides that we observe for the phospho-WW domain. However, when we compare peak integrals in  $^1\text{H}$ - $^{15}\text{N}$  HSQC for the residues outside of this loop region in the WW<sup>PKA</sup> sample, we found good agreement with the MALDI-TOF MS on the intact proteins, and deduced a phosphorylation yield of  $90 \pm 3\%$ .

In addition to its analytical capacities, NMR spectroscopy is powerful to characterize the structural, dynamical and functional consequences of phosphorylation. PKA phosphorylation of the WW leads to a loss of its binding capacity to a phosphorylated Tau peptide, despite the overall triple  $\beta$ -sheet structure of the WW module being

conserved (Fig. 6). This suggests modulation of the WW binding activity at another level. Protein motions such as loop flexibility are known to be involved in the control of enzymatic functions, ligand binding or protein-protein interactions (Baldwin and Kay 2009; Eisenmesser et al. 2005; Kay 2005; Mittermaier and Kay 2006). Conformational exchange and flexibility of the binding-loop region of the WW domain were indeed shown to be an important parameter for the binding affinity and specificity of the WW domain for phosphorylated substrates (Peng et al. 2007; Sharpe et al. 2007). The observed restriction in the backbone dynamics of this region once the WW domain is phosphorylated on S16, most pronounced on the  $\mu\text{s}$ -ms timescale, is thus a first factor that might participate in the

observed decrease of affinity for binding. Introduction of a phosphate moiety on S16 further limits the availability of R17, the essential residue for the phosphate binding of the protein partner. Our data hence suggest that the conformational and dynamical changes of the phospho-binding loop upon S16 phosphorylation resemble the structural adaptation occurring when a phosphorylated peptide substrate binds to the WW domain (see Fig. 7d and Supplementary Materials, Figure S2). Such conformational mimicking could explain the loss of function of the Pin1 WW domain in its PKA-phosphorylated state. Both mechanisms of gain of conformational rigidity and compactness together with the phosphate moiety itself occupying the position of the incoming phospho-S/T will lead to a phospho-WW domain unable to bind other phospho-proteins.

**Acknowledgments** We thank Professor H. Schwalbe (Frankfurt, Germany) for a kind gift of the purified PKA enzyme. The NMR facilities are funded by the the European community, the Centre National de la Recherche Scientifique (CNRS), the Région Nord-Pas de Calais (France), the University of Lille 1 and the Institut Pasteur de Lille. The Mass Spectrometry facilities are funded by the European community (FEDER), the Région Nord-Pas de Calais (France), the IBISA network, the CNRS, and the University of Lille 1.

## References

- Ando K, Dourlen P, Sambo AV, Bretteville A, Belarbi K, Vingtdoux V, Eddarkaoui S, Drobecq H, Ghestem A, Begard S, Demey-Thomas E, Melnyk P, Smet C, Lippens G, Mauraage CA, Caillet-Boudin ML, Verdier Y, Vinh J, Landrieu I, Galas MC, Blum D, Hamdane M, Sergeant N, Buee L (2012) Tau pathology modulates Pin1 post-translational modifications and may be relevant as biomarker. *Neurobiol Aging* 34(3):757–769
- Andrew CD, Warwicker J, Jones GR, Doig AJ (2002) Effect of phosphorylation on alpha-helix stability as a function of position. *Biochemistry* 41(6):1897–1905
- Arosio B, Bulbarelli A, Bastias Candia S, Lonati E, Mastronardi L, Romualdi P, Candeletti S, Gussago C, Galimberti D, Scarpini E, Dell'Osso B, Altamura C, MacCarrone M, Bergamaschini L, D'Addario C, Mari D (2012) Pin1 contribution to Alzheimer's disease: transcriptional and epigenetic mechanisms in patients with late-onset Alzheimer's disease. *Neuro Degener Dis* 10(1–4): 207–211
- Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17(1): 75–86
- Baker JM, Hudson RP, Kanelis V, Choy WY, Thibodeau PH, Thomas PJ, Forman-Kay JD (2007) CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat Struct Mol Biol* 14(8):738–745
- Balastik M, Lim J, Pastorino L, Lu KP (2007) Pin1 in Alzheimer's disease: multiple substrates, one regulatory mechanism? *Biochim Biophys Acta* 1772(4):422–429
- Baldwin AJ, Kay LE (2009) NMR spectroscopy brings invisible protein states into focus. *Nat Chem Biol* 5(11):808–814
- Bayer E, Goettsch S, Mueller JW, Griewel B, Guiberman E, Mayr LM, Bayer P (2003) Structural analysis of the mitotic regulator hPin1 in solution: insights into domain architecture and substrate binding. *J Biol Chem* 278(28):26183–26193
- Bienkiewicz EA, Lumb KJ (1999) Random-coil chemical shifts of phosphorylated amino acids. *J Biomol NMR* 15(3):203–206
- Bulbarelli A, Lonati E, Cazzaniga E, Gregori M, Masserini M (2009) Pin1 affects Tau phosphorylation in response to Abeta oligomers. *Mol Cell Neurosci* 42(1):75–80
- Du JT, Li YM, Wei W, Wu GS, Zhao YF, Kanazawa K, Nemoto T, Nakanishi H (2005) Low-barrier hydrogen bond between phosphate and the amide group in phosphopeptide. *J Am Chem Soc* 127(47):16350–16351
- Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature* 438(7064): 117–121
- Eto M, Kitazawa T, Matsuzawa F, Aikawa S, Kirkbride JA, Iozumi N, Nishimura Y, Brautigam DL, Ohki SY (2007) Phosphorylation-induced conformational switching of CPI-17 produces a potent myosin phosphatase inhibitor. *Structure* 15(12): 1591–1602
- Golovanov AP, Blankley RT, Avis JM, Bermel W (2007) Isotopically discriminated NMR spectroscopy: a tool for investigating complex protein interactions in vitro. *J Am Chem Soc* 129(20): 6528–6535
- Hamdane M, Smet C, Sambo AV, Leroy A, Wieruszeski JM, Delobel P, Mauraage CA, Ghestem A, Wintjens R, Begard S, Sergeant N, Delacourte A, Horvath D, Landrieu I, Lippens G, Buee L (2002) Pin1: a therapeutic target in Alzheimer neurodegeneration. *J Mol Neurosci* 19(3):275–287
- Hwang TL, van Zijl PC, Mori S (1998) Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J Biomol NMR* 11(2): 221–226
- Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* 32(3): 1037–1049
- Johnson LN, Barford D (1993) The effects of phosphorylation on the structure and function of proteins. *Annu Rev Biophys Biomol Struct* 22:199–232
- Kay LE (2005) NMR studies of protein structure and dynamics. *J Magn Reson* 173(2):193–207
- Landrieu I, Lacosse L, Leroy A, Wieruszeski JM, Trivelli X, Sillen A, Sibille N, Schwalbe H, Saxena K, Langer T, Lippens G (2006) NMR analysis of a Tau phosphorylation pattern. *J Am Chem Soc* 128(11):3575–3583
- Landrieu I, Leroy A, Smet-Nocca C, Huvent I, Amniai L, Hamdane M, Sibille N, Buee L, Wieruszeski JM, Lippens G (2010) NMR spectroscopy of the neuronal tau protein: normal function and implication in Alzheimer's disease. *Biochem Soc Trans* 38(4): 1006–1011
- Liou YC, Sun A, Ryo A, Zhou XZ, Yu ZX, Huang HK, Uchida T, Bronson R, Bing G, Li X, Hunter T, Lu KP (2003) Role of the prolyl isomerase Pin1 in protecting against age-dependent neurodegeneration. *Nature* 424(6948):556–561
- Lippens G, Landrieu I, Hanouille X (2008) Studying posttranslational modifications by in-cell NMR. *Chem Biol* 15(4):311–312
- Lonati E, Masserini M, Bulbarelli A (2011) Pin1: a new outlook in Alzheimer's disease. *Curr Alzheimer Res* 8(6):615–622
- Lu KP, Hanes SD, Hunter T (1996) A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* 380 (6574): 544–547
- Lu PJ, Wulf G, Zhou XZ, Davies P, Lu KP (1999) The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* 399(6738):784–788

- Lu PJ, Zhou XZ, Liou YC, Noel JP, Lu KP (2002) Critical role of WW domain phosphorylation in regulating phosphoserine binding activity and Pin1 function. *J Biol Chem* 277(4):2381–2384
- Ma SL, Pastorino L, Zhou XZ, Lu KP (2012) Prolyl isomerase Pin1 promotes amyloid precursor protein (APP) turnover by inhibiting glycogen synthase kinase-3beta (GSK3beta) activity: novel mechanism for Pin1 to protect against Alzheimer disease. *J Biol Chem* 287(10):6969–6973
- Mittermaier A, Kay LE (2006) New tools provide new insights in NMR studies of protein dynamics. *Science* 312(5771):224–228
- Ohki S, Eto M, Kariya E, Hayano T, Hayashi Y, Yazawa M, Brautigam D, Kainosho M (2001) Solution NMR structure of the myosin phosphatase inhibitor protein CPI-17 shows phosphorylation-induced conformational changes responsible for activation. *J Mol Biol* 314(4):839–849
- Pastorino L, Sun A, Lu PJ, Zhou XZ, Balastik M, Finn G, Wulf G, Lim J, Li SH, Li X, Xia W, Nicholson LK, Lu KP (2006) The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-beta production. *Nature* 440(7083):528–534
- Peng T, Zintsmaster JS, Namanja AT, Peng JW (2007) Sequence-specific dynamics modulate recognition specificity in WW domains. *Nat Struct Mol Biol* 14(4):325–331
- Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2(6):661–665
- Prabakaran S, Everley RA, Landrieu I, Wieruszeski JM, Lippens G, Steen H, Gunawardena J (2011) Comparative analysis of Erk phosphorylation suggests a mixed strategy for measuring phospho-form distributions. *Mol Syst Biol* 7:482
- Pufall MA, Lee GM, Nelson ML, Kang HS, Velyvis A, Kay LE, McIntosh LP, Graves BJ (2005) Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science* 309(5731):142–145
- Pullen K, Rajagopal P, Branchini BR, Huffine ME, Reizer J, Saier MH Jr, Scholtz JM, Klevit RE (1995) Phosphorylation of serine-46 in HPr, a key regulatory protein in bacteria, results in stabilization of its solution structure. *Protein Sci* 4(12):2478–2486
- Ranganathan R, Lu KP, Hunter T, Noel JP (1997) Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* 89(6):875–886
- Rangasamy V, Mishra R, Sondarva G, Das S, Lee TH, Bakowska JC, Tzivion G, Malter JS, Rana B, Lu KP, Kanthasamy A, Rana A (2012) Mixed-lineage kinase 3 phosphorylates prolyl-isomerase Pin1 to regulate its nuclear translocation and cellular function. *Proc Natl Acad Sci USA* 109(21):8149–8154
- Sharpe T, Jonsson AL, Rutherford TJ, Daggett V, Fersht AR (2007) The role of the turn in beta-hairpin formation during WW domain folding. *Protein Sci* 16(10):2233–2239
- Smet C, Duckert JF, Wieruszeski JM, Landrieu I, Buee L, Lippens G, Deprez B (2005a) Control of protein–protein interactions: structure-based discovery of low molecular weight inhibitors of the interactions between Pin1 WW domain and phosphopeptides. *J Med Chem* 48(15):4815–4823
- Smet C, Sambo AV, Wieruszeski JM, Leroy A, Landrieu I, Buee L, Lippens G (2004) The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau. *Biochemistry* 43(7):2032–2040
- Smet C, Wieruszeski JM, Buee L, Landrieu I, Lippens G (2005b) Regulation of Pin1 peptidyl-prolyl cis/trans isomerase activity by its WW binding module on a multi-phosphorylated peptide of Tau protein. *FEBS Lett* 579(19):4159–4164
- Sultana R, Boyd-Kimball D, Poon HF, Cai J, Pierce WM, Klein JB, Markesbery WR, Zhou XZ, Lu KP, Butterfield DA (2006) Oxidative modification and down-regulation of Pin1 in Alzheimer's disease hippocampus: a redox proteomics analysis. *Neurobiol Aging* 27(7):918–925
- Teriete P, Thai K, Choi J, Marassi FM (2009) Effects of PKA phosphorylation on the conformation of the Na, K-ATPase regulatory protein FXYD1. *Biochim Biophys Acta* 1788(11):2462–2470
- Theillet FX, Smet-Nocca C, Liokatis S, Thongwichian R, Kosten J, Yoon MK, Kriwacki RW, Landrieu I, Lippens G, Selenko P (2012) Cell signaling, post-translational protein modifications and NMR spectroscopy. *J Biomol NMR* 54(3):217–236
- Verdecia MA, Bowman ME, Lu KP, Hunter T, Noel JP (2000) Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol* 7(8):639–643
- Wintjens R, Wieruszeski JM, Drobecq H, Rousselot-Pailley P, Buee L, Lippens G, Landrieu I (2001) 1H NMR study on the binding of Pin1 Trp–Trp domain with phosphothreonine peptides. *J Biol Chem* 276(27):25150–25156
- Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD (1995) 1H, 13C and 15 N chemical shift referencing in biomolecular NMR. *J Biomol NMR* 6(2):135–140
- Wishart DS, Sykes BD (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J Biomol NMR* 4(2):171–180
- Wittekind M, Reizer J, Deutscher J, Saier MH, Klevit RE (1989) Common structural changes accompany the functional inactivation of HPr by seryl phosphorylation or by serine to aspartate substitution. *Biochemistry* 28(26):9908–9912
- Yaffe MB, Schutkowski M, Shen M, Zhou XZ, Stukenberg PT, Rahfeld JU, Xu J, Kuang J, Kirschner MW, Fischer G, Cantley LC, Lu KP (1997) Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. *Science* 278(5345):1957–1960