



Université de Lille, Faculté des Sciences et Technologies
Ecole Doctorale Biologie Santé de Lille EDBSL (ED n°446)
Unité de Glycobiologie Structurale et Fonctionnelle (UGSF), CNRS UMR 8576

Diplôme National de DOCTORAT

Spécialité : Aspects Moléculaires et Cellulaires de la Biologie

Traçage des protéines O-GlcNAcylées par la stratégie de suppression d'Ambre

Présentée et soutenue publiquement le 27 Juin 2024 à 14h30 à Lille par :

Monsieur Christer ABOU ANNY

Thèse dirigée par Pr. Tony LEFEBVRE et Pr. Christophe BIOT ; Membres du Jury :

Pr. GROUX-DEGROOTE Sophie	Université de Lille	PRÉSIDENTE
Pr. LAUZIER Benjamin	Université de Nantes	RAPPORTEUR
Dr. LOPEZ Marie	Université de Montpellier	RAPPORTEURE
Dr. CAMBERLEIN Emilie	Université de Nantes	EXAMINATRICE
Dr. NOUAILLE Sébastien	Université de Toulouse	EXAMINATEUR
Pr. LEFEBVRE Tony	Université de Lille	DIRECTEUR DE THÈSE
Pr. BIOT Christophe	Université de Lille	CO-DIRECTEUR DE THÈSE
Dr. OLIVIER-VAN STICHELEN Stéphanie	Medical College of Wisconsin	MEMBRE INVITÉE
Dr. FAURE Régis	Université de Toulouse	MEMBRE INVITÉ

« La patience est un élément clé de la réussite »

Bill Gates

« Rien ne s'obtient qu'avec effort, tout a son sacrifice »

Gustave Flaubert

Remerciements :

Je tiens tout d'abord à remercier l'unité de Glycobiologie Structurale et Fonctionnelle (UGSF), dirigée par le Dr. **Yann GUERARDEL**, de m'avoir accueilli durant ces années ainsi que l'ensemble du personnel y travaillent.

Je souhaite également remercier l'ensemble des membres de mon jury de thèse, la présidente la Pr. **Sophie GROUX-DEGROOTE**, les rapporteurs la Dr. **Marie LOPEZ** et le Dr. **Benjamin LAUZIER** et les examinateurs la Dr. **Emilie CAMBERLEIN** et le Dr. **Sébastien NOUAILLE**, ainsi que les membres invités la Dr. **Stéphanie OLIVIER-VAN STICHELEN** et le Dr. **Régis FAURE** pour avoir accepté d'évaluer mon travail de thèse et pour l'intérêt qu'ils y ont porté.

Je tiens à exprimer ma sincère gratitude en premier lieu envers mon directeur de thèse le Pr. **Tony LEFEBVRE**, et mon co-directeur de thèse le Pr. **Christophe BIOT**. Je suis reconnaissant pour cette chance d'avoir eu des encadrants aussi exceptionnels et de faire partie des équipes qui sont devenus pour moi ma deuxième famille. J'aimerais bien vous remercier pour votre encadrement, votre soutien inestimable, votre gentillesse, votre disponibilité constante et votre encouragement qui m'ont donné la force de persévérer durant les moments les plus difficiles du parcours doctoral, et surtout pour le temps que vous m'avez consacré tout au long de ces années et plus encore lors de la rédaction de ce mémoire. Votre expertise académique et professionnelle a été une source d'inspiration et un modèle à suivre. Votre empathie et votre patience qui m'ont permis de surmonter les obstacles avec confiance, votre communication claire et vos conseils éclairés ont éclairé le chemin vers la réussite.

Tous mes sincères remerciements vont également à tous les membres des deux équipes, grâce à qui j'ai passé de très bons moments. Pr. **Ikram**, merci pour vos conseils, vos questions lors des réunions ainsi que pour vos encouragements. Dr. **Anne-Sophie**, merci beaucoup pour votre aide précieuse, votre support et conseils qui m'ont beaucoup aidé à trouver des solutions aux problèmes que j'ai rencontrés lors mes manipulations durant ma thèse. Dr. **Stéphan**, merci beaucoup pour votre aide précieuse et vos conseils en biologie moléculaire qui m'ont permis de trouver des solutions et des

protocoles adaptés surtout dans les PCR. Dr. **Quentin**, merci beaucoup pour votre gentillesse, votre aide précieuse dans le laboratoire, j'apprécie beaucoup votre sens de l'organisation. **Céline**, je vous remercie pour votre aide précieuse et formations ainsi que vos conseils en Biologie Moléculaire. **Dimitri**, merci beaucoup pour votre sympathie, gentillesse, votre aide lorsque j'avais besoin de cellules ainsi que votre discussion enrichissante en sciences ou sur les sujets de la vie; bon courage pour la suite de votre doctorat, future collaborateur. **Marlène**, merci beaucoup pour votre gentillesse. **Jodie**, merci beaucoup pour votre gentillesse et votre aide dans le laboratoire ; bon courage pour la suite de votre doctorat. **Ferdinand**, merci pour votre sympathie, bonne chance pour la suite de votre doctorat, **Awatef**, merci pour votre gentillesse, je vous souhaite une bonne continuation après le doctorat. Dr. **Sadia**, ça m'a fait plaisir de faire votre connaissance et d'avoir eu la chance de discuter avec vous sur pas mal de sujets. Dr. **Vanessa**, merci pour votre gentillesse et votre discussion enrichissante en sciences ou autres sujets de la vie, Dr. **Isabelle**, merci beaucoup pour votre gentillesse, votre support et formation dans la purification des protéines, **Dorothée**, je tiens à vous remercier pour ta gentillesse et pour m'avoir si bien formé en culture cellulaire. **Anastasia, Maxime et Thomas**, bonne chance pour votre master 2. Dr. **Olivier Vidal**, merci pour votre sympathie et vos conseils, Dr. **Christophe Mariller**, merci pour vos conseils et discussion surtout dans le processus de la concentration des protéines après purification, **Peggy**, merci pour votre gentillesse et sympathie. **Vincent, Angelina, Réginald, Yannick, Mathieu, Audrey**, merci pour les discussions partagées et votre sympathie.

Je tiens à remercier Dr. **Corentin SPRIET** et **Thomas BANCE**, pour leur enthousiasme, leur soutien, leurs formations en microscopie confocale; leurs discussions stimulantes ont considérablement enrichi mon expérience dans le domaine de l'imagerie des cellules.

Je tiens à remercier également nos chers collaborateurs du TBI de Toulouse, le Dr. **Régis FAURE** et le Dr. **Sébastien NOUILLE** par ce petit mot à vous remercier de votre temps, patience et dévouement à mon égard durant mes deux séjours, qui pour moi ont été, et grâce, à vous, riches en expériences d'apprentissage, expériences

encouragées par vos précieuses recommandations et instructions qui pour moi ont été très formatrices !

Agnès, Fabrice, Mathieu, Fred Chirat, Laurence, Martine, Audrey, Hammou, Marie-Ange, Dave, merci pour votre sympathie.

Je tiens à exprimer ma gratitude au programme PEARL (*Programme for EARly-stage Researchers in Lille*), dirigée par M. **Alexis Boulet** qui m'a donné l'opportunité à effectuer ma recherche interdisciplinaire, de suivre des formations dans le cadre de mon doctorat. En particulier, je remercie M. Alexis pour son assistance et suivi pendant les quatre années du programme de thèse et sa précieuse aide dans les démarches administratives. Je tiens à remercier également le coordinateur du programme PEARL Pr. **Franck DUMEIGNIL** et directeur général de la fondation I-SITE le Pr. **Régis BORDET**.

Mes parents, mes frères **Paul, Patrick** et **Charbel** et sœurs **Ramona** et **Rita**, mon cousin **Robert**, mon beau-frère **Souheil**, mes neveux et nièces **Carelle, Paul, Charbel et Georges**, malgré les kilomètres de distance vous restez toujours proche de mon cœur et dans ma mémoire. Vous me manquez énormément! Aujourd'hui, alors que je termine cette étape importante de ma vie académique, je tiens à vous remercier du fond du cœur. Je porte avec moi toujours le souvenir précieux de notre amour et de notre unité. Que cette réussite soit la vôtre autant que la mienne, car c'est grâce à vous que j'ai pu atteindre ces sommets. Votre amour et votre soutien ont été les fondations sur lesquelles j'ai bâti mon succès, et je suis profondément reconnaissant de vous avoir à mes côtés.

Je tiens à exprimer mes plus sincères remerciements à ma famille, et plus particulièrement à ma mère et ma sœur Ramona, que j'aime profondément.

Maman, grâce à tes prières, j'ai eu la chance de rencontrer des personnes formidables, que ce soit dans mon travail ou dans ma vie en France. Tu avais toujours ce sentiment que je devais bâtir mon avenir ailleurs, et tu avais raison. En France, j'ai commencé à bâtir la vie magnifique dont j'avais rêvé.

Ramona, ma chère sœur, tu as raison de dire que la vie peut changer pour le mieux, et que même après les moments difficiles, le soleil finit toujours par briller. Tes paroles d'encouragement ont été comme une bouffée d'air frais pendant les tempêtes de ma vie. Souviens-toi de notre enfance, nos espoirs de grandir pour faire une différence dans la vie des moins fortunés; là je suis sur le bon chemin. Tu es ma source d'inspiration, et je te suis infiniment reconnaissant pour tout ce que tu as fait pour moi. Je te souhaite également la réussite dans ton doctorat et dans ta vie au Liban.

A tous ceux qui de près ou de loin ont contribué d'une façon ou d'une autre à la réussite de mon projet. Un grand merci à vous.



Marie Skłodowska-Curie
Actions

Ce projet a reçu un financement du programme de recherche et d'innovation Horizon 2020 de l'Union Européenne dans le cadre de l'accord de subvention Marie Skłodowska-Curie n° 847568.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 847568.

Résumé :

Traçage des protéines *O*-GlcNAcylées par la stratégie de suppression d'Ambre

La *O*-GlcNAcylation est une modification post-traductionnelle (MPT) régulée par deux enzymes antagonistes, l'OGT (*O*-GlcNAc transférase), qui transfère le résidu GlcNAc à partir de l'UDP-GlcNAc sur le groupe hydroxyle d'une sérine ou d'une thréonine des protéines cibles, et l'OGA (*O*-GlcNAcase) qui hydrolyse ce résidu. Cette MPT est étroitement liée à la voie de biosynthèse des hexosamines qui produit l'UDP-GlcNAc et fait intervenir différents nutriments tels que le glucose ou autres monosaccharides, les acides gras, l'UTP et l'ensemble des acides aminés, la glutamine en particulier. L'émergence de certaines pathologies, notamment le cancer, est associée à une *O*-GlcNAcylation aberrante. En ce sens, les travaux antérieurs de l'équipe ont indiqué une augmentation des niveaux d'expression de la β -caténine *O*-GlcNAcylée dans les cancers épithéliaux. Quatre sites de *O*-GlcNAcylation ont été identifiés au niveau de l'extrémité *N*-terminale de la β -caténine (S23/T40/T41/T112). La T41 revêt un intérêt particulier car elle influence la sensibilité au protéasome de la β -caténine selon qu'elle soit *O*-GlcNAcylée ou phosphorylée, les deux MPTs antagonistes jouant des fonctions opposées sur l'expression de la protéine. Cependant, il était difficile voire impossible de suivre spécifiquement cette protéine sous sa forme *O*-GlcNAcylée ou phosphorylée dans la même cellule. Pour pallier cette lacune, nous avons combiné la technologie d'expansion du code génétique avec la chimie *click* pour suivre la *O*-GlcNAcylation de la T41 en utilisant une *S*-GlcNAc, un analogue enzymatiquement stable de la *O*-GlcNAc. Une forme non modifiée en T41 ainsi qu'une forme phosphomimétique (T41E) de la β -caténine ont été générées. Nous avons synthétisé un ensemble de cassettes d'expression β -caténine-XFP (fusionnées ou non à un TAT-HA) et introduit le codon ambre TAG à la position T41 par mutagenèse dirigée du site pour produire la forme *S*-GlcNAcylée. Chaque forme (wt, T41*S*-GlcNAc et T41E) présentant une fluorescence spécifique a été produite chez *E. coli*, purifiée puis introduite en cellule humaine par transduction cellulaire ou par lipofection. Nos résultats indiquent que selon l'isoforme, les protéines n'occupent pas exactement le même espace dans la cellule, aucune superposition totale n'ayant été observée. La forme *S*-GlcNAcylée de la β -caténine en T41 semble plus stable dans le temps que la forme phosphomimétique, renforçant le rôle protecteur de la *O*-GlcNAcylation de la β -caténine vis-à-vis de la dégradation au contraire de sa phosphorylation.

La technologie que nous avons développée doit permettre de fabriquer à façon des protéines modifiées par des groupements chimiques différents sur des sites bien précis, et d'en étudier l'impact sur le comportement de la protéine en réponse à divers stimuli.

Mots clés : *O*-GlcNAcylation, β -caténine, thiosucre, assemblage moléculaire Gibson, mutagenèse dirigée, expansion du code génétique, incorporation d'acides aminés non naturels, Chimie *click*, transduction cellulaire, transfection cellulaire.

Abstract:

Tracing the *O*-GlcNAcylated proteins by the strategy of Amber suppression

O-GlcNAcylation is a post-translational modification (PTM) regulated by two antagonistic enzymes, OGT (*O*-GlcNAc transferase), which transfers the GlcNAc residue from UDP-GlcNAc to the hydroxyl group of a serine or a threonine of target proteins, and OGA (*O*-GlcNAcase) which hydrolyzes this residue. This PTM is closely linked to the hexosamine biosynthesis pathway that produces UDP-GlcNAc by using different nutrients such as glucose or other monosaccharides, fatty acids, UTP and all amino acids, glutamine in particular. The emergence of certain pathologies, cancer in particular, is associated with aberrant *O*-GlcNAcylation. In this sense, the previous work of the team showed an increase in the expression levels of *O*-GlcNAcylated β -catenin in epithelial cancers. Four *O*-GlcNAcylation sites were identified at the N-terminus of β -catenin (S23/T40/T41/T112). T41 is of particular interest because it influences the proteasome sensitivity of β -catenin depending on whether it is *O*-GlcNAcylated or phosphorylated, the two antagonistic PTMs playing opposite functions on the expression of the protein. However, it was difficult or impossible to specifically track this protein in its *O*-GlcNAcylated or phosphorylated form in the same cell. To fill this gap, we combined genetic code expansion technology with click chemistry to monitor *O*-GlcNAcylation of T41 using *S*-GlcNAc, an enzymatically stable analogue of *O*-GlcNAc. An unmodified form at T41 as well as a phosphomimetic form (T41E) of β -catenin were generated. We synthesized a set of β -catenin-XFP expression cassettes (fused or not to a TAT-HA) and introduced the amber TAG codon at position T41 by site-directed mutagenesis to produce the *S*-GlcNAcylated form. Each form (wt, T41S-GlcNAc and T41E) presenting a specific fluorescence was produced in *E. coli*, purified and then introduced into human cells by cell transduction or by lipofection. Our results indicate that depending on the isoform, the proteins do not occupy exactly the same space in the cell, no total overlap having been observed. The *S*-GlcNAcylated form of β -catenin at T41 seems more stable over time than the phosphomimetic form, reinforcing the protective role of *O*-GlcNAcylation of β -catenin against degradation as opposed to its phosphorylation. The technology that we have developed should make it possible to custom manufacture proteins modified by different chemical groups on specific sites, and to study their impact on the behavior of the protein in response to various stimuli.

Keywords: *O*-GlcNAcylation, β -catenin, thiosugar, Gibson molecular assembly, site-directed mutagenesis, genetic code expansion, incorporation of unnatural amino acids, click chemistry, cellular transduction, cellular transfection.

Table des matières

Avant-propos	18
Chapitre I. INTRODUCTION	20
Généralités	20
Partie 1 : la O-GlcNAcylation : une modification post-traductionnelle monosaccharidique unique	22
1. Mécanismes moléculaires sous-jacents de la O-GlcNAcylation	22
2. La S-GlcNAcylation : une forme alternative stable à la O-GlcNAcylation	26
3. O-GlcNAcylation et phosphorylation : un duo en compétition	27
4. O-GlcNAcylation et cancer : un duo malfaisant.	30
Partie 2 : Approches et méthodes de détection et d'étude de la O-GlcNAcylation	33
1. Approches traditionnelles de détection de la O-GlcNAcylation	33
2. Inhibiteurs pharmacologiques des enzymes de la O-GlcNAcylation et de la voie des hexosamines	35
3. Rapporteurs chimiques et métaboliques utilisés dans le marquage de la O-GlcNAcylation.	39
4. Imagerie spécifique des protéines O-GlcNAcylées <i>via</i> l'utilisation de la stratégie FRET	46
Projet de thèse	58
Objectifs des travaux de thèse	58
Chapitre II : Matériels et Méthodes	64
1. Matériels des expériences supplémentaires	64
1.1. Incorporation de la Phényl-L-sélénocystéine à la place du codon <i>ambre</i> TAG et S-GlcNAcylation	64
1.2. Transduction et transfection cellulaire	65

1.3. Immunofluorescence	66
2. Méthodes des manip supplémentaires	67
2.1. S-GlcNAcylation de la β -caténine en T41	67
2.2. Stratégie de double transfection de plasmides et de protéines recombinantes en cellules eucaryotes par lipofection	67
2.3. Immunofluorescence	69
Chapitre III : Résultats	70
1. Clonage et assemblage moléculaire des constructions TAT-HA	72
2. Traçage des isoformes de la β -caténine par transduction cellulaire	85
3. Contrôle de la β -caténine transduite par le peptide TAT.	89
4. Clonage et assemblage moléculaire des constructions sans TAT-HA	92
5. Evaluation des β -caténines transfectées sans TAT-HA	94
6. Impact de la S-GlcNAcylation sur l'expression de la β -caténine.	100
Chapitre IV : Discussion	104
Chapitre V : Conclusion et perspectives	110
Annexes	113
Références bibliographiques	

Table des figures et des tableaux

Figure 1. La *O*-GlcNAcylation, une modification post-traductionnelle (MPT) associée au métabolisme des nutriments

Figure 2. Régulation des mécanismes moléculaires et cellulaires par la *O*-GlcNAcylation.

Figure 3. Principales pathologies liées à des dérégulations de la *O*-GlcNAcylation.

Figure 4. Les sucres *O*-hétérosides vs *S*-hétérosides

Figure 5. La *S*-GlcNAc, une forme alternative stable de la *O*-GlcNAc.

Figure 6. Pathologies associées à la perturbation de l'équilibre homéostasique entre la *O*-GlcNAcylation et la phosphorylation

Figure 7. Implication de la *O*-GlcNAcylation dans l'émergence du cancer.

Figure 8. Méthodes et approches utilisées dans l'analyse de la *O*-GlcNAcylation.

Figure 9. Perspectives de l'ingénierie oligosaccharidique métabolique (MOE) sur la détection des voies de glycosylation cellulaire

Figure 10. Stratégie des rapporteurs chimiques métaboliques utilisés pour le marquage de la *O*-GlcNAcylation.

Figure 11. Imagerie basée sur FRET du double marquage des protéines et des glycanes.

Figure 12. Rôle des ARNt orthogonaux dans l'expansion du code génétique.

Figure 13. Stratégie de l'expansion du code génétique. L'AANN est introduit dans les cellules hôtes contenant l'aminacyl-ARNt synthétase orthogonale (aaRS) et l'ARNt orthogonal.

Figure 14. Application de l'expansion du code génétique à notre projet de thèse.

Figure 15. Formation du *S*-GlcNAc.

Figure 16. Production de la forme phosphomimétique en T41 de la β -caténine (T41E).

Figure 17. Stratégie et méthodologie suivies au cours du projet de thèse.

Figure 18. Construction des vecteurs plasmidiques par assemblage des fragments d'ADN.

Figure 19. Caractérisation des produits de PCR par électrophorèse en gel d'agarose.

Figure 20. Génération du plasmide codant la synthétase orthogonale aaRS spécifique de la PhSeCys.

Figure 21. Résumé graphique des différentes constructions plasmidiques avec TAT-HA utilisées dans ce projet.

Figure 22. Traçage des différentes isoformes de la TAT-HA- β -caténine par transduction des cellules Hep3B et CCD841CoN.

Figure 23. Contrôle de co-marquage de la forme phosphomimétique TAT-HA- β -caténine-mCherry avec les marqueurs d'endosome précoce Rab5 et EEA1 en cellules Hep3B.

Figure 24. Élimination de la partie TAT-HA des différentes constructions par PCR à l'aide du couple d'amorces P25-P26.

Figure 25. Caractérisation des différentes constructions délétées de la séquence TAT-HA par électrophorèse en gel d'agarose.

Figure 26. Résumé graphique des différentes constructions plasmidiques sans TAT-HA utilisées dans ce projet.

Figure 27. Approche combinant la lipofection d'une β -caténine-EGFP (wt) sans TAT-HA et la transduction cellulaire d'une TAT-HA- β -caténine-mCherry (wt) en Hep3B (A) et HeLa.

Figure 28. Expression d'une β -caténine codée par un vecteur plasmidique versus β -caténine produite chez *E. coli* et lipofectée.

Figure 29. Évaluation de l'expression des trois isoformes de la β -caténine avant et après purification.

Figure 30. Introduction des trois isoformes de la β -caténine dans les cellules Hep3B et HeLa par lipofection.

Figure 31. Évaluation de viabilité cellulaire.

Figure 32. Suivi de la S-GlcNAcylation et de la forme phosphomimétique du résidu T41 de la β -caténine-XFP pendant 24 et 48 heures

Tableau 1. Inhibiteurs d'OGT, d'OGA et des enzymes de la voie de biosynthèse des hexosamines.

Tableau 2. Les différentes catégories de métabolites utilisés pour marquer les protéines O-GlcNAcylées.

Tableau 3. Exemples d'acides aminés accompagnés de leur paire orthogonale aaRS/ARNt.

Tableau 4. Souches bactériennes et plasmides.

Tableau 5. Amorces utilisées pour générer l'ensemble des constructions.

Tableau 6. Composition et conditions de PCR réalisée par l'ADN polymérase haute-fidélité Q5.

Tableau 7. Caractéristiques des composants de la PCR.

Tableau 8. Conditions d'assemblage de l'ADN plasmidique.

Tableau 9. Assemblage d'ADN plasmidique par la méthode Gibson.

Liste des abréviations et nomenclatures

A

aaRS: *aminoacyl-tRNA synthetase*

AGM1 ou **PGM:** Phospho-GlcNAc Mutase

ATP: Adénosine Triphosphate

C

CuAAC: *Copper-catalyzed Azide-Alkyne Cycloaddition*

CRISPR/Cas9: *Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein 9*

D

DAPI: 4',6-diamidino-2-phénylindole

Dha: Déshydroalanine

DMEM: *Dulbecco's Modified Eagle Medium*

E

EEA1: *Early Endosome Antigen 1*

EGFP: *Enhanced Green Fluorescent Protein*

eRF1: *eukaryotic Release Factor 1*

ESRRB: *Estrogen-Related Receptor Beta*

F

FRET: *Fluorescence Resonance Energy Transfer*

Fru6P: Fructose-6-phosphate

G

GCE: *Genetic Code Expansion*

GFAT: Glutamine : fructose-6-phosphate amidotransférase

Glc: Glucose

GlcNAc1P: N-acétylglucosamine-1-phosphate

GlcNAc6P: N-acétylglucosamine-6-phosphate

GlcNH₂6P: Glucosamine-6-phosphate

Glc6P: Glucose-6-phosphate

GalNAc: N-acétylgalactosamine

GNA: GlcNH₂-6-phosphate-N-acétyltransférase

GPI: Glucose-6-phosphate isomérase

gT41: Thréonine 41 O-GlcNAcylée

H

HA: *Hemagglutinin tag*

HBP: *Hexosamine biosynthetic pathway*

6xHis tag: *6x Histidine tag*

HK: Hexokinase

I

IMAC: *Immobilized Metal Affinity Chromatography*

IP: Immunoprécipitation

M

MEM: *Minimum Essential Medium*

PTM: *Post-translational modification*

MOE: *Metabolic Oligosaccharide Engineering*

mTagBFP: Monomeric Tag Blue Fluorescent Protein

N

Nup153: Nucleoporin 153

O

OGA: O-GlcNAcase ou O- β -D-N-acétylglucosaminidase

O-GlcNAcylation: O- β -N-acétylglucosaminylation

OGT: O-GlcNAc transférase ou polypeptide O- β -D-N-acétylglucosaminyltransférase

OSMI: *OGT small molecule inhibitor*

P

PAF: Paraformaldéhyde

P53: *Tumor protein 53*

PCR: *Polymerase Chain Reaction*

PFK1: Phosphofructokinase-1

PhSeCys: Phénylsélocystéine

pT41: Thréonine 41 phosphorylée

R

Rab5: *Ras-related protein Rab-5*

RF1: *Release Factor 1*

S

SDS-PAGE: *Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis*

Ser: Sérine

S-GlcNAc: S-linked N-acetylglucosamine

S-GlcNAcylation: S- β -N-acetylglucosaminylation

SPAAC: Strain-Promoted Azide-Alkyne Cycloaddition

T

T41: Thréonine 41

TAT: *Transactivator of transcription*

TCF: *T-cell factor*

T41E: Thréonine 41 phospho-mimétique

T41 S-GlcNAc: Thréonine 41 S-GlcNAcylée

U

UDP: Uridine diphosphate

UDP-GlcNAc PPase: UDP-N-acetylglucosamine pyrophosphorylase

UGPP: UDP-GlcNAc Pyrophosphorylase

UTP: Uridine triphosphate

W

WB: *Western Blot*

Liste des produits chimiques

A

Ac₃6AlkGlcNAc: 6-alkynyl-6-deoxy-N-acetylglucosamine per-*O*-acetylated

Ac₃2AzGlc: 2-azido-2-deoxy-3-*O*-acetylglucose

Ac₃6AzGlcNAc: 6-azido-6-deoxy-N-acetylglucosamine per-*O*-acetylated

Ac₃4dGlcNAz: 4-deoxy-N-azidoacetylglucosamine per-*O*-acetylated

Ac₄GlcNCyoc: N-cyclooctynyl-N-acetylglucosamine per-*O*-acetylated

Ac₄GlcNAIk: N-pentynylglucosamine per-*O*-acetylated

Ac₄GlcNAz: N-azidoacetylglucosamine per-*O*-acetylated

Ac₄-5S-GlcNAc: 2-acetamido-1,3,4,6-tetra-*O*-acetyl-2-deoxy-5-thio- α -D-glucopyranose

B

BZX: benzoxazolinone

P

PhSeCys: Phenylselenocysteine

1,3-Pr₂GalNAz: 1,3-di-*O*-propionyl-N-azidoacetylgalactosamine

1,6-Pr₂GalNAz: 1,6-di-*O*-propionyl-N-azidoacetylgalactosamine

S

5S-GlcNHex: 2-deoxy-2-N-hexanamide-5-thio-D-glucopyranose

SH-GlcNAc: Sulfhydryl-N-Acetylglucosamine

T

TCEP :Tris(2-carboxyethyl)phosphine

Avant-propos

Cette thèse de doctorat a été réalisée au sein de l'Unité de Glycobiologie structurale et fonctionnelle (UGSF-UMR CNRS 8576) à Villeneuve d'Ascq dans l'équipe « Régulation des mécanismes moléculaires et cellulaires par O-GlcNAcylation » sous la direction du Pr. Tony Lefebvre, et dans l'équipe « Chemical GlycoBiology (CheGB) », sous la direction du le Pr. Christophe Biot. Une collaboration dans le cadre d'un projet PEARL (Programme for EARly-stage Researchers in Lille) a été réalisée avec l'Institut Biotechnologique de Toulouse (TBI) au sein des équipes EADs 1 et 4 sous la responsabilité du Dr. Sébastien Nouaille et du Dr. Régis Fauré.

Publication :

Christer Abou Anny, Sébastien Nouaille, Régis Fauré, Céline Schulz, Corentin Spriet, Isabelle Huvent, Christophe Biot, Tony Lefebvre (2024). A Step-by-Step Guide for the Production of Recombinant Fluorescent TAT-HA-tagged proteins and their Transduction into Mammalian Cells. *Curr Protoc.* 2024 Mar;4(3):e1016. doi: 10.1002/cpz1.1016.

Communications orales:

Christer Abou Anny “Monitoring of O-GlcNAcylated Proteins in Living Cells Through the Use of Non-Canonical Amino Acid”. The 10th Annual Conference of the International Chemical Biology Society - ICBS 2021: Looking towards the next decade of chemical biology. (Présentation en anglais), 10-13 Novembre 2021. Atlanta, GA, Etats-Unis.

Christer Abou Anny “Monitoring of O-GlcNAcylated Proteins in Living Cells Through the Use of Non-Canonical Amino Acid”. Journée André Verbert 2022 - Colloque des doctorants. 15 Novembre 2022. Lille, France.

Communication par poster :

Christer Abou Anny “Asset of genetic code expansion for the use of the enzymatically stable analogue S-GlcNAc to track O-GlcNAcylated β -catenin in living cells”. The 47th

FEBS Congress (FEBS 2023)-"Together in bioscience for a better future. (Présentation en anglais), 8-12 Juillet 2023. Tours, France.

Conférences :

Journée André Verbert 2023 - Colloque des doctorants. 15 Novembre 2023. Lille, France.

Journée André Verbert 2022 - Colloque des doctorants. 15 Novembre 2022. Lille, France.

7ème Forum de Recherche en Oncologie Clinique FRON 2020 - 'COVID-19 et Cancer'. 11 Septembre 2020. Université Saint Joseph, Faculté de Médecine, Beyrouth, Liban.

Chapitre I. INTRODUCTION

Généralités

L'espérance de vie a connu une croissance continue au cours des deux derniers siècles. Cependant, un ensemble de troubles compromet à terme l'objectif de "vivre plus longtemps en bonne santé". Ces troubles incluent les neuropathologies, les maladies cardiovasculaires, les cancers ainsi que des problèmes liés aux perturbations du métabolisme et de la nutrition tels que le diabète, l'obésité et l'hypertension. Ces pathologies, quelles qu'elles soient, sont le résultat de dérégulations moléculaires qui affectent le fonctionnement des cellules, des tissus et des organes, mettant ainsi en danger l'organisme dans son ensemble. De nombreux états pathologiques sont en partie causés par la dérégulation des modifications post-traductionnelles (MPT). Cependant, il n'existe actuellement aucune technologie permettant de suivre "une protéine d'intérêt" avec "une MPT spécifique" au sein de cellules, ce qui constitue un obstacle à l'étude de ces modifications dans divers contextes physiopathologiques. Parmi ces MPT, la *O*-GlcNAcylation, dont le niveau est étroitement lié à la disponibilité en nutriments, est largement répandue dans le règne animal (Hanover et al., 2012; Hart, 2014). Ce type de glycosylation consiste en l'ajout réversible *via* une liaison β -*O*-glycosidique d'un unique résidu de *O*-N-acétyl-D-glucosamine (GlcNAc) sur le groupement hydroxyle d'une sérine ou d'une thréonine de protéines confinées dans le cytosol, le noyau et la mitochondrie. Cette modification et les enzymes régulant son dynamisme, à savoir la *O*-GlcNAc transférase (OGT) et la *O*-GlcNAcase (OGA), sont perturbées dans nombre de pathologies, comme les maladies cardiovasculaires et les cancers. A ce jour, plusieurs milliers de protéines ont été identifiées comme étant *O*-GlcNAcylées, parmi celles-ci la β -caténine (Sayat et al., 2008; Olivier-Van Stichelen et al., 2014).

La β -caténine remplit de nombreuses fonctions essentielles chez les mammifères. Par exemple, la signalisation Wnt/ β -caténine, dont la β -caténine est l'acteur central, joue un rôle actif dans le développement embryonnaire. De plus, elle est nécessaire pour maintenir l'intégrité de la muqueuse et favoriser la prolifération des cellules épithéliales.

Aussi, sa dérégulation est responsable de la tumorigenèse de certains tissus épithéliaux (Monga, 2015; Mohammed et al., 2016). De manière intéressante, notre équipe de recherche a démontré que les nutriments impactent l'expression de la β -caténine en lui conférant une résistance au protéasome esquissant un nouveau mode de régulation de cette protéine et proposant un lien entre environnement nutritionnel et tumorigenèse (Olivier-Van Stichelen et al., 2014).

Partie 1 : la O-GlcNAcylation : une modification post-traductionnelle monosaccharidique unique

1. Mécanismes moléculaires sous-jacents de la O-GlcNAcylation

La O-GlcNAcylation est une MPT impliquant l'addition réversible d'un unique monosaccharide, la N-acétylglucosamine (GlcNAc), sur des groupes hydroxyles (OH) des résidus de sérine ou de thréonine de protéines exclusivement nucléaires, cytoplasmiques et mitochondriales (Yang and Qian, 2017). La O-GlcNAcylation est régulée par deux enzymes : la O-GlcNAc transférase (OGT), ajoutant le résidu de GlcNAc sur les protéines cibles, et la O-GlcNAcase (OGA), qui l'hydrolyse (Stephen et al., 2021; Wu et al., 2024). La O-GlcNAcylation est donc une modification dynamique comme la phosphorylation, avec laquelle elle peut entrer en compétition (Hart et al., 2011).

La nature du cycle de la O-GlcNAc repose sur plusieurs composants, comprenant l'uridine-5'-diphosphate N-acétylglucosamine (UDP-GlcNAc), l'OGT, l'OGA, et les protéines substrats. Contrairement à la phosphorylation, impliquant des centaines de kinases et de phosphatases (Ubersax and Ferrell Jr, 2007; Johnson, 2009), le cycle de la O-GlcNAc des protéines est caractérisé par l'activité exclusive d'une paire d'enzymes, l'OGT et l'OGA (Parker et al., 2021; Stephen et al., 2021). Ces deux enzymes, hautement conservées au cours de l'évolution, jouent un rôle essentiel chez les eucaryotes supérieurs. L'UDP-GlcNAc, issu de la voie de biosynthèse des hexosamines (HBP, *hexosamine biosynthetic pathway*) (Paneque et al., 2023), agit en tant que substrat donneur de résidus GlcNAc pour l'OGT. La découverte simultanée au début des années 1980 de ces composants clés, à savoir OGT, OGA, et HBP, a considérablement enrichi notre compréhension de la nature chimique de la réaction de O-GlcNAcylation (Hart, 2014).

Au cours des quelques décennies écoulées depuis ces premières découvertes, la recherche a révélé que des processus de O-GlcNAcylation aussi bien au niveau global que sur des sites bien spécifiques, influent sur tout un ensemble de processus

physiologiques et pathologiques, des altérations pouvant résulter du flux de la voie des hexosamines que de la modulation des enzymes impliquées dans le cycle de la O-GlcNAcylation. Au niveau moléculaire, la O-GlcNAc modifie des milliers de protéines, représentant presque toutes les classes d'un point de vue fonctionnel. À noter particulièrement, la O-GlcNAc est un senseur nutritionnel indirect, car la concentration intracellulaire d'UDP-GlcNAc répond fortement aux fluctuations des nutriments *via* la voie des hexosamines (**Figure 1**).

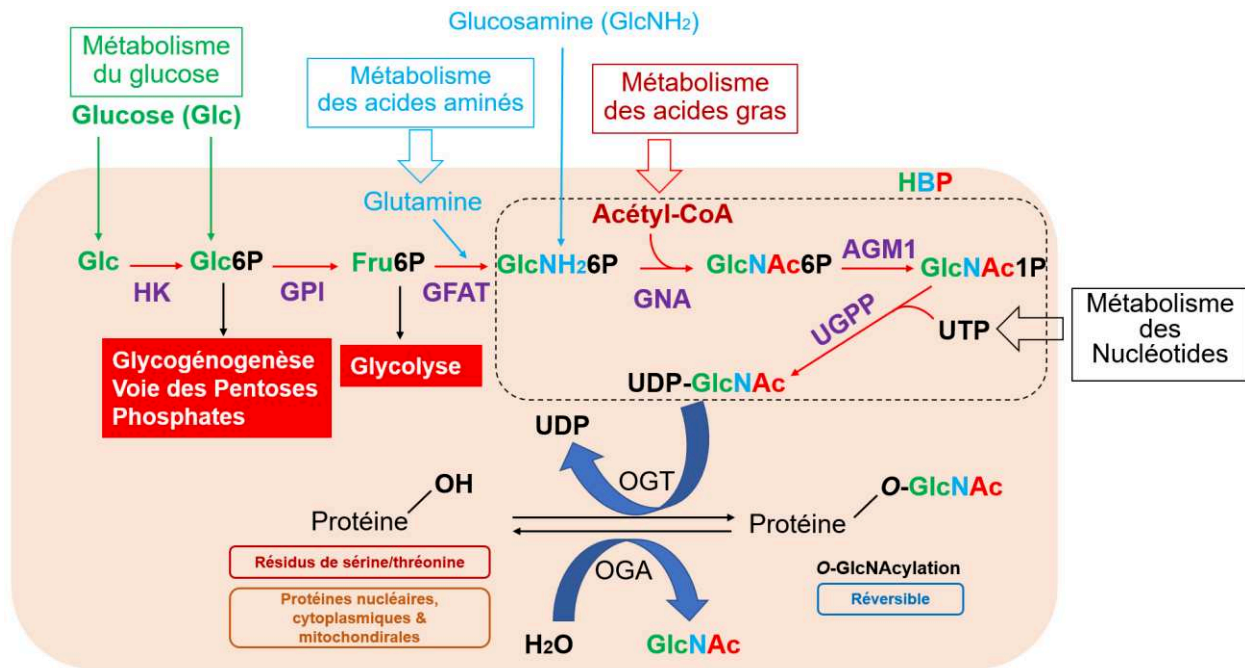


Figure 1. La O-GlcNAcylation, une modification post-traductionnelle (MPT) associée au métabolisme des nutriments. Cette MPT est étroitement liée à la voie de biosynthèse des hexosamines (HBP, *hexosamine biosynthetic pathway*), qui fournit le nucléotide sucre donneur, l'UDP-GlcNAc, nécessaire aux processus de glycosylation, notamment la O-GlcNAcylation. La voie des hexosamines est ainsi un indicateur de l'état nutritionnel en raison de son flux métabolique alimenté par diverses sources de nutriments.

AGM1, phospho-GlcNAc mutase ; GFAT, glutamine : fructose-6-phosphate amido-transférase ; GPI, glucose-6-phosphate isomérase ; GNA, GlcNH₂-6-phosphate-N-acétyltransférase ; HK, hexokinase ; UGPP, UDP-GlcNAc pyrophosphorylase.

Au cours des trois dernières décennies, il a été révélé que les mécanismes moléculaires sous-jacents à cette MPT étaient extrêmement complexes, et que celle-ci intervenait dans une grande diversité de processus cellulaires (Ma et al., 2021): le cycle cellulaire (Tan et al., 2017; Saunders et al., 2023), la différenciation cellulaire (Sheikh et

al., 2021), la prolifération cellulaire (Levine et al., 2021), les voies de signalisation (Zachara and Hart, 2006; Zhang et al., 2021) ou les voies métaboliques (Lefebvre and Issad, 2015). Par ailleurs, cette modification influence de manière significative la transcription génique en modulant l'activité de facteurs de transcription et en affectant la structure de la chromatine, contribuant ainsi à la maintenance du génome et à la régulation épigénétique (Hanover et al., 2012; Hardivillé and Hart, 2016; Wu et al., 2017). En outre, la O-GlcNAcylation s'avère essentielle dans la réponse au stress cellulaire, plus précisément dans des conditions de stress métabolique (Groves et al., 2013) et au cours de l'apoptose (Pellegrini et al., 2023). La O-GlcNAcylation régule plus largement sur le plan physiologique, l'appétit (Dai et al., 2018), les rythmes circadiens (Kim et al., 2012) et la mémoire (Wheatley et al., 2019) (**Figure 2**).

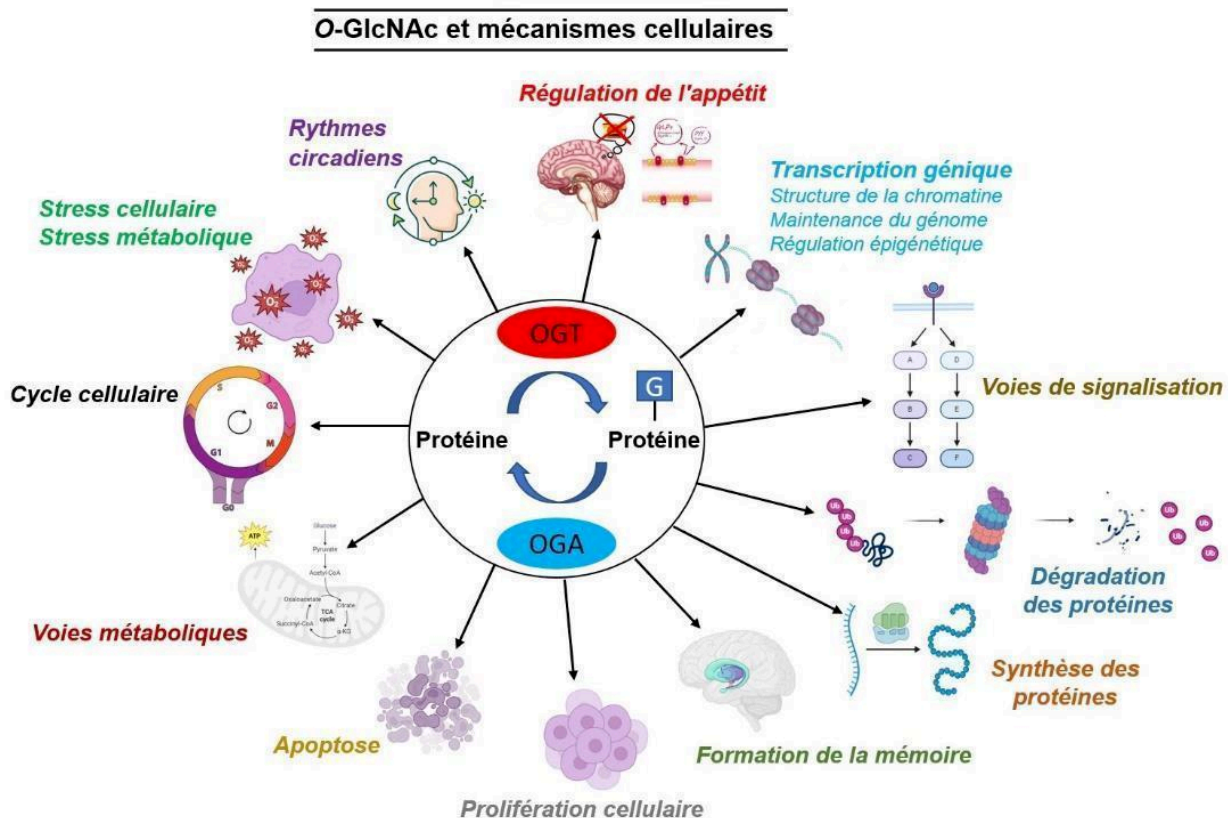


Figure 2. Régulation des mécanismes moléculaires et cellulaires par la O-GlcNAcylation.

Des études ont également établi des liens entre des niveaux altérés de la O-GlcNAcylation et diverses pathologies, notamment le cancer (Fardini et al., 2013; Parker et al., 2021), le diabète (Fricovsky et al., 2012; Peterson and Hart, 2016), les maladies neurodégénératives (Balana and Pratt, 2021; Domenico et al., 2020), les désordres musculaires (Lambert et al., 2018) et cardiovasculaires (Wright et al., 2017; Bolanle et al., 2021), soulignant ainsi son rôle critique dans les contextes physiologiques et pathologiques (**Figure 3**).

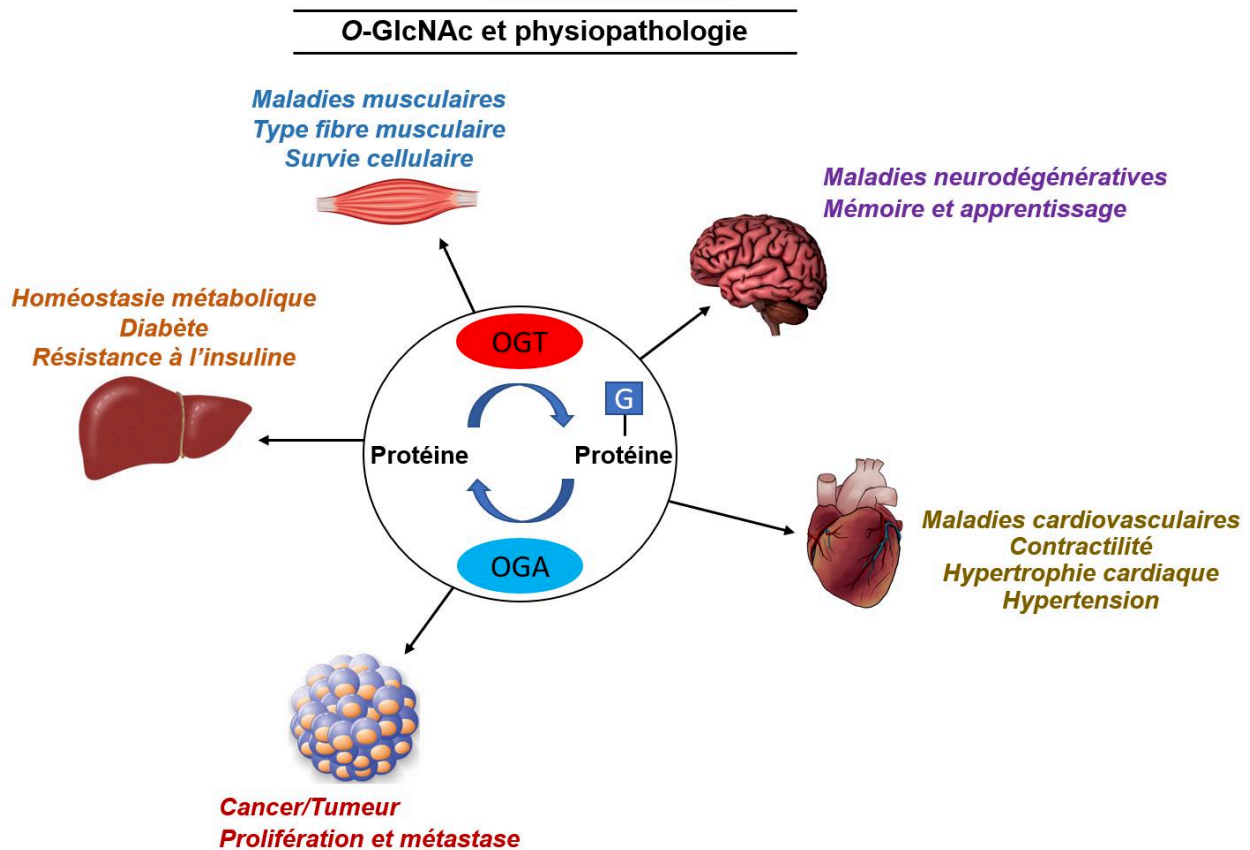


Figure 3. Principales pathologies liées à des dérégulations de la O-GlcNAcylation.

2. La S-GlcNAcylation : une forme alternative stable à la O-GlcNAcylation

Il a été récemment identifié la S-GlcNAcylation, caractérisée par l'addition de GlcNAc sur des résidus de cystéine, comme une forme alternative stable de la O-GlcNAcylation (Buchowiecka, 2023). Le mécanisme moléculaire de la S-GlcNAcylation est similaire à celui de la O-GlcNAcylation. La S-GlcNAcylation intervient dans la modulation de la structure et de la fonction des protéines, régulant ainsi divers processus cellulaires tels que la signalisation cellulaire, la réponse au stress et la prolifération cellulaire. L'addition de GlcNAc aux résidus de cystéine est catalysée par l'OGT (Mitchell et al., 2023).

Les thiosucres représentent des analogues de sucres naturels où l'oxygène anomérique est remplacé par un atome de soufre. Au niveau structural, la liaison C–S est plus longue que la liaison C–O (1,78 Å contre 1,41 Å). De plus, les angles formés par les deux carbones adjacents à l'hétéroatome diffèrent dans les séries O- ou S-glycoside (C–S–C est de 98° contre 117° pour C–O–C) (**Figure 4**).

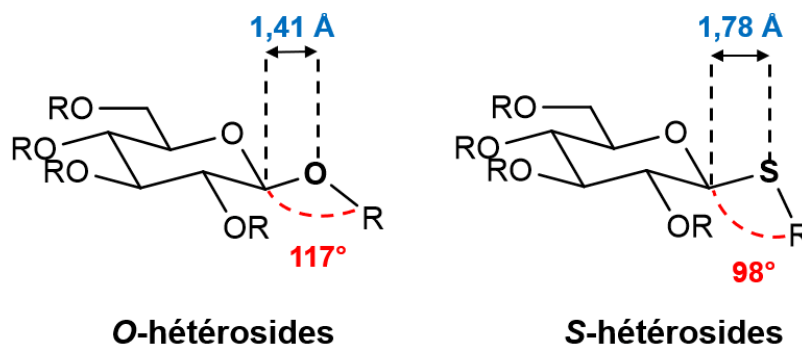


Figure 4. Les sucres O-hétérosides vs S-hétérosides

Des études complémentaires réalisées par De Lion et al., alliant résonance magnétique nucléaire (RMN) et modélisation moléculaire, ont conforté le fait que la S-GlcNAc est une alternative stable de la O-GlcNAc partageant de nombreuses similitudes avec cette dernière (De Leon et al., 2017) (**Figure 5**).

Dans le cadre de mon projet de thèse, nous avons utilisé la S-GlcNAc pour suivre la O-GlcNAcylation sur le site T41 de la β -caténine à l'aide de la bioimagerie.

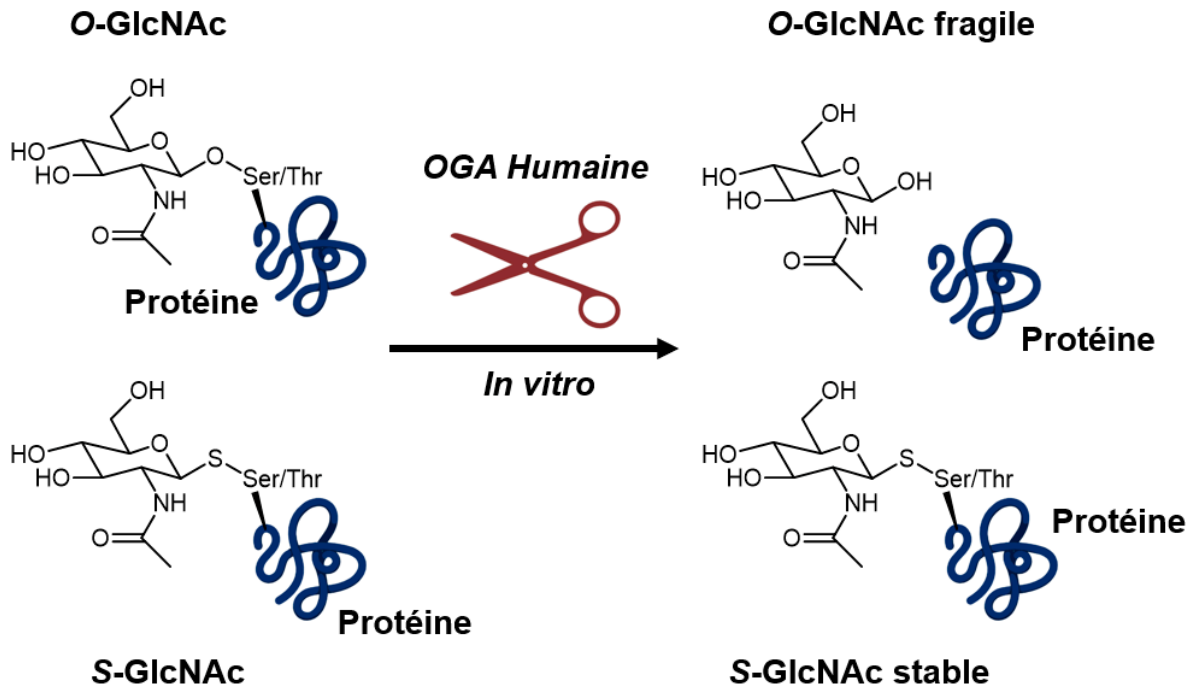


Figure 5. La S-GlcNAc, une forme alternative stable de la O-GlcNAc.

3. O-GlcNAcylation et phosphorylation : un duo en compétition

Depuis des décennies, la recherche s'est concentrée sur la compréhension d'une interaction dynamique et compétitive : il s'agit du duo O-GlcNAcylation/phosphorylation. Ces deux MPT, réversibles, jouent un rôle crucial dans la régulation coordonnée de nombreuses protéines cellulaires. Elles partagent souvent des sites de résidus similaires, créant ainsi un scénario de compétition.

Plusieurs études ont rapporté la réciprocity O-GlcNAc/phosphorylation sur des sites spécifiques de certaines protéines d'intérêts. Il a été révélé que cette compétition pouvait agir comme interrupteur moléculaire, régulant l'activité, la localisation cellulaire, la stabilité ou les interactions protéine-protéine. Il existe des protéines présentant à la fois des sites de phosphorylation et de O-GlcNAcylation: parmi ces protéines se trouve la β -caténine. La β -caténine est une protéine impliquée dans la régulation de la signalisation cellulaire, notamment la voie de signalisation Wnt/ β -caténine, cruciale dans le développement embryonnaire et la régénération tissulaire chez les adultes. Des

études ont montré que cette protéine présente plusieurs sites de *O*-GlcNAcylation, dont certains chevauchent des sites de phosphorylation, impactant par conséquent sa fonction, sa localisation et son interaction avec différents partenaires.

Dans cette partie, nous allons présenter quelques études réalisées dans le but d'identifier et d'analyser des sites de réciprocity *O*-GlcNAc/phosphorylation de protéines considérées comme importantes sur le niveau physiologique et métabolique, comme la β -caténine, la protéine tau, la protéine CREB et la protéine c-jun.

Une étude réalisée par Olivier-Van Stichelen et al., a permis d'identifier le site thréonine 41 (T41) de la partie N-terminale de la β -caténine comme un site de réciprocity *O*-GlcNAc/phosphorylation. Il a été montré que ce site est également impliqué dans la sensibilité au protéasome, la localisation et l'interaction avec des partenaires protéiques. Cela suppose que toute dérégulation de la réciprocity *O*-GlcNAcylation/phosphorylation du site T41 peut affecter sa conformation, son interaction avec divers partenaires et par conséquent ses fonctions. Des altérations de cet antagonisme *O*-GlcNAcylation/phosphorylation doivent être répercutées par des dérégulations de voies de signalisation impliquées dans la prolifération cellulaire, et au final participer à l'émergence de diverses pathologies et cancers (Olivier-Van Stichelen et al., 2014). Pour cela, il est particulièrement intéressant d'explorer le potentiel de la *O*-GlcNAcylation de la β -caténine en utilisant l'ingénierie et l'imagerie basées sur la fluorescence. Une étude réalisée par Kasproicz et al. a appliqué cette technologie pour suivre cette protéine en cellules vivantes en utilisant une stratégie de rapporteur chimique bioorthogonal. Cette stratégie est basée sur le double marquage de la β -caténine avec une protéine fluorescente verte (GFP) pour la séquence protéique combinée à une sonde d'imagerie chimiquement cliquée pour la MPT, se caractérisant par un FRET qualitatif rapide et facile à quantifier (Kasproicz et al., 2020). Étant donné la réciprocity entre la *O*-GlcNAcylation et la phosphorylation du site T41 de la β -caténine, ce site a constitué l'objet central de mon projet de thèse. Pour réaliser notre étude, nous avons ainsi mis en place une technologie novatrice permettant de suivre à la fois la *O*-GlcNAcylation et la phosphorylation de T41 de la β -caténine par bio-imagerie.

D'autres sites de réciprocité *O*-GlcNAcylation/phosphorylation sont connus. C'est le cas du site sérine 396 (S396) de la protéine tau (Gatta et al., 2016), connue pour son rôle dans la stabilisation des microtubules dans les neurones. Sous une forme hyperphosphorylée, cette protéine peut être liée à la dégénérescence neuronale dans la pathologie Alzheimer. Le site S133 de la protéine CREB (cAMP response element-binding protein), un facteur de transcription impliqué dans la régulation de l'expression génique en réponse à divers signaux cellulaires, a été identifié comme un site de réciprocité *O*-GlcNAc/phosphorylation (Rexach et al., 2012). Il a été suggéré qu'une dérégulation au niveau de ce site serait impliquée en partie dans des maladies métaboliques, neurodégénératives, cardiovasculaires ou cancers. Une étude réalisée par Kim et al., a montré que la sérine 63 (S63) du facteur de transcription c-jun, est un site de réciprocité *O*-GlcNAc/phosphorylation (Kim et al., 2018), la dérégulation au niveau de ce site affectant le comportement et la fonction de c-jun. Il est supposé que ce déséquilibre soit en partie responsable de certains cancers ou autres maladies métaboliques. Autre site de réciprocité *O*-GlcNAc/phosphorylation, la sérine 276 (S276) de la protéine NF- κ B (facteur nucléaire kappa B), un facteur de transcription impliqué dans la régulation de la réponse immunitaire et l'inflammation. Petta et al., ont démontré que la phosphorylation ou la *O*-GlcNAcylation aberrante de ce site se répercute sur des altérations de voies cellulaires et à un changement d'activité transcriptionnelle de NF- κ B impactant divers processus biologiques comme la survie, la réponse immunitaire, la prolifération cellulaire et l'inflammation (Petta et al., 2016).

Au niveau moléculaire, des études ont montré que la présence d'un groupe *O*-GlcNAc peut bloquer la phosphorylation du même résidu ou d'un site adjacent, et *vice versa*. Par conséquent, la dérégulation du duo *O*-GlcNAcylation/phosphorylation affecte directement la protéostase cellulaire, pouvant contribuer à l'émergence de pathologies métaboliques, neurodégénératives, cancéreuses, cardiovasculaires, hépatiques et musculaires (**Figure 6**).

Nous pouvons conclure qu'il existe de nombreux sites de réciprocité *O*-GlcNAc/phosphorylation sur un grand nombre de protéines cellulaires jouant un rôle crucial dans la signalisation cellulaire et les voies métaboliques. Ces sites sont donc

essentiels pour le bon fonctionnement des cellules et des tissus de l'organisme, garantissant ainsi un maintien de l'homéostasie cellulaire optimale. Cependant, toute altération des MPTs au niveau des protéines cellulaires, ou des enzymes impliquées dans la O-GlcNAcylation (OGT/OGA), ou la phosphorylation/déphosphorylation (kinases/phosphatases), peut avoir des conséquences dramatiques sur l'intégrité de l'organisme.

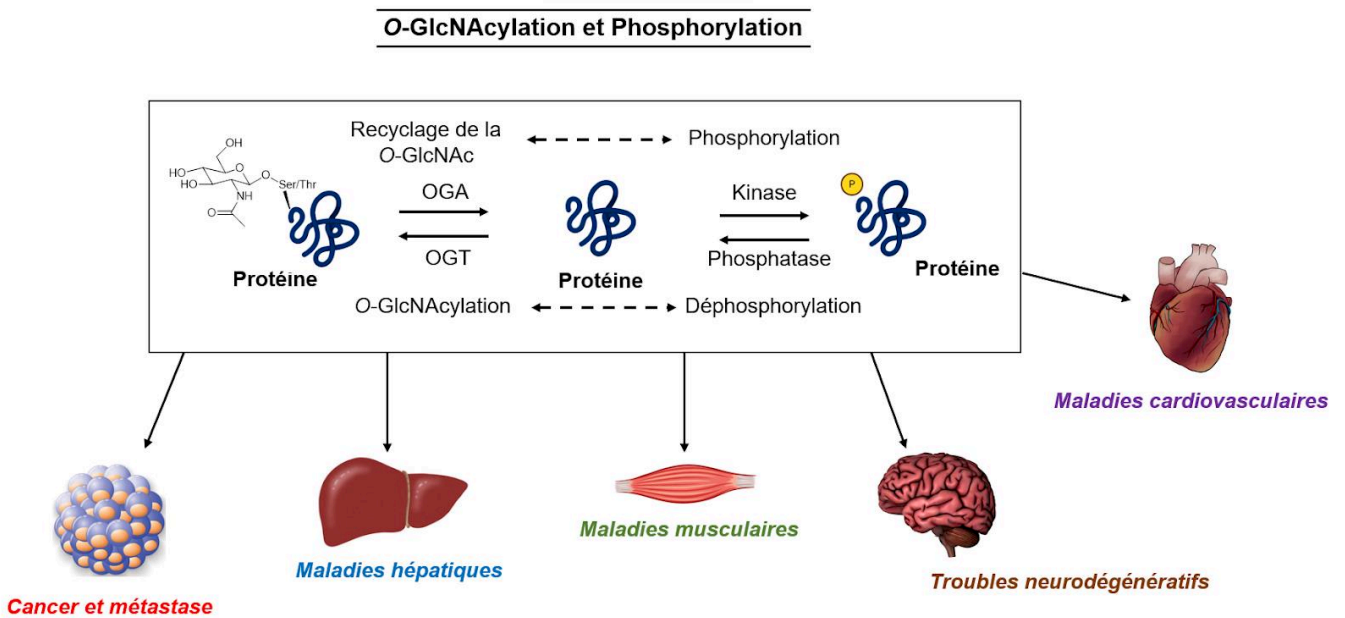


Figure 6. Pathologies associées à la perturbation de l'équilibre homéostasique entre la O-GlcNAcylation et la phosphorylation.

4. O-GlcNAcylation et cancer : un duo malfaisant.

Le cancer est une maladie dont l'impact s'étend au travers de l'histoire de l'humanité. Récemment, un article publié par le Centre International de Recherche sur le Cancer (CIRC), une agence de l'Organisation Mondiale de la Santé (OMS), a confirmé que le cancer figure parmi les principales causes de mortalité à l'échelle mondiale. Selon une étude menée en 2022 dans 185 pays, les trois types de cancer les plus fréquents sont le cancer du poumon, le cancer du sein et le cancer colorectal. Les statistiques issues de l'Observatoire Mondial du Cancer (GCO, *Global Cancer Observatory*) indiquent

qu'en 2022, les nouveaux cas enregistrés comprennent ces mêmes cancers, cancer du poumon (12,4%), cancer du sein (11,6%) et cancer colorectal (9,6%), mais également le cancer de la prostate (7,3%) et le cancer de l'estomac (4,9%) (WHO, 2024, who.net) **(Figure 7)**.

Un lien a été établi entre la *O*-GlcNAcylation, l'émergence et la progression du cancer (Lee et al., 2021). Ces études ont révélé que cette MPT peut altérer l'activité et la stabilité de nombreuses protéines impliquées dans le cancer, telles que les oncogènes incluant la β -caténine (Olivier-Van Stichelen et al., 2014; Gao et al., 2019) et la protéine c-Myc (Sodi et al., 2015; Makwana et al., 2019), et les suppresseurs de tumeurs comme p53 (Li and Yi, 2014; Queiroz et al., 2016).

D'autres études ont montré que la *O*-GlcNAcylation pouvait aussi influencer des processus cellulaires cruciaux pour la croissance tumorale, notamment une augmentation du flux glycolytique et une résistance accrue à l'apoptose, ainsi qu'une régulation plus forte de voies de signalisation telles que la voie de l'insuline et la signalisation associée aux récepteurs de facteurs de croissance (Issad et al., 2010).

De plus, la capacité de la *O*-GlcNAcylation à influencer le microenvironnement tumoral pourrait affecter la réponse immunitaire et favoriser la migration cellulaire, des éléments clés dans l'invasion et la métastase tumorale. Dans le domaine thérapeutique, la plupart des recherches visent à cibler cette *O*-GlcNAcylation anormale en développant et testant des inhibiteurs spécifiques d'OGT et d'OGA (Alteen et al., 2021) **(Figure 7)**.

Nous pouvons conclure que dans des conditions physiologiques, la *O*-GlcNAcylation peut avoir un effet protecteur sur les diverses fonctions de l'organisme humain. Cependant, lorsque cette MPT est altérée en raison de facteurs génétiques, épigénétiques ou environnementaux, des dérégulations dans le comportement des protéines modifiées perturbent les voies de signalisation cellulaires et autres processus biologiques associés. Ces dysfonctionnements peuvent contribuer au développement de maladies, en particulier le cancer.

O-GlcNAc et cancer

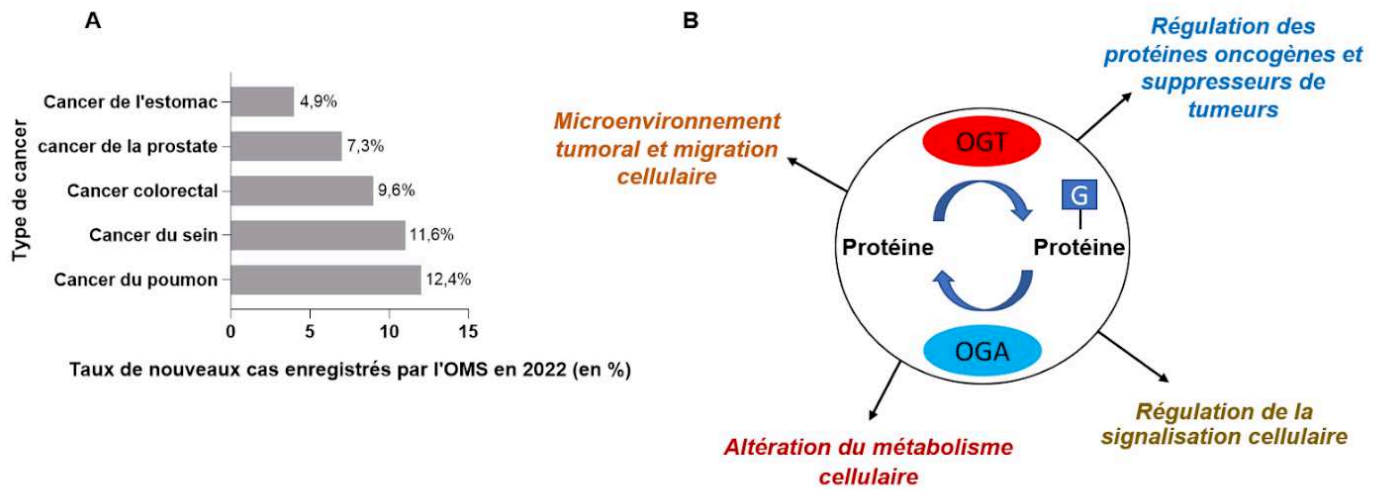


Figure 7. Implication de la O-GlcNAcylation dans l'émergence du cancer.

Partie 2 : Approches et méthodes de détection et d'étude de la O-GlcNAcylation

Dans cette partie, nous allons illustrer et discuter des avancées et des études récentes portant sur les différentes approches, stratégies et outils moléculaires appliqués à la détection de la O-GlcNAcylation (**Figure 8**).

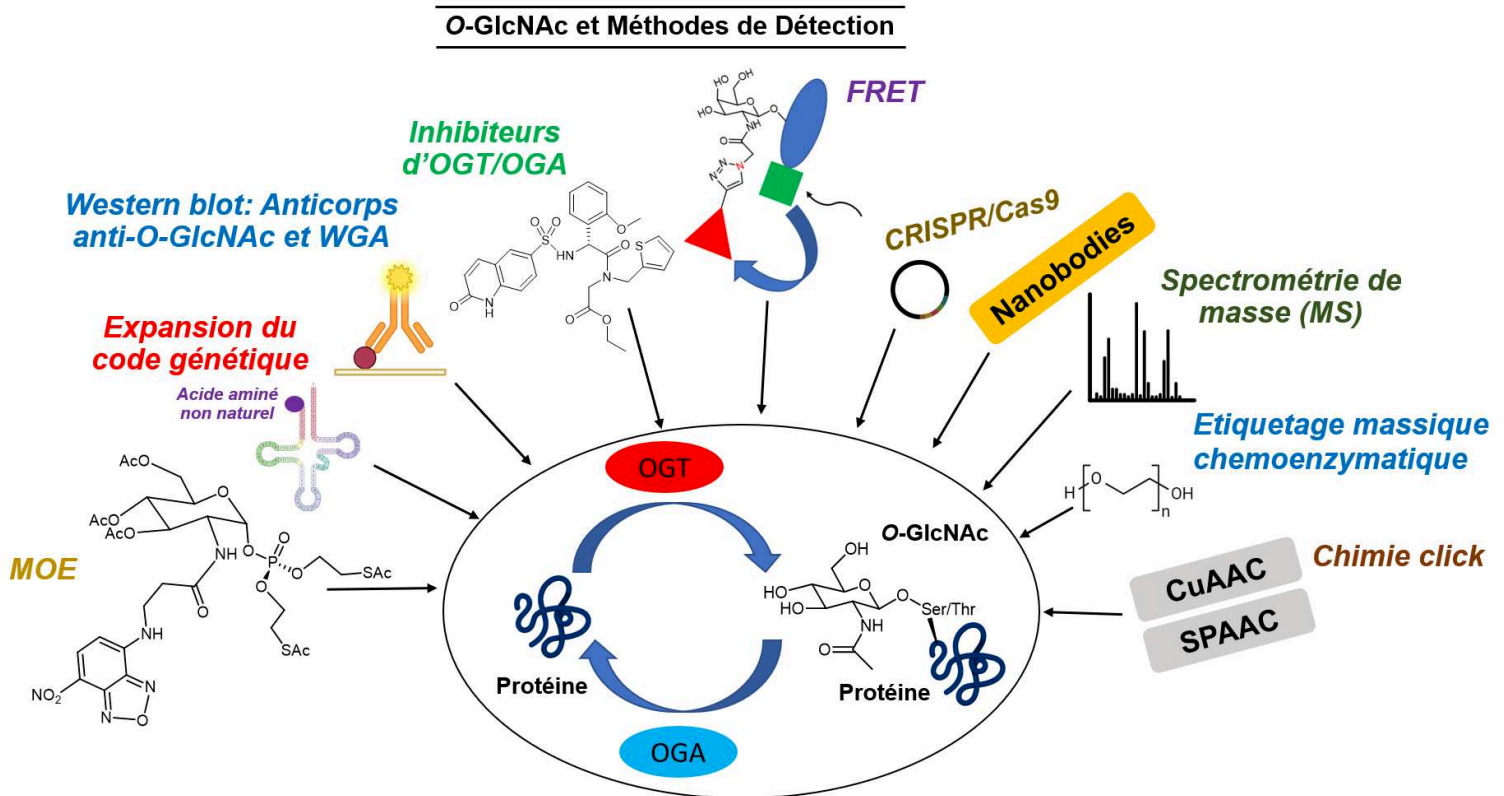


Figure 8. Méthodes et approches utilisées dans l'analyse de la O-GlcNAcylation

1. Approches traditionnelles de détection de la O-GlcNAcylation

La détection et l'analyse de la O-GlcNAcylation des protéines peuvent être réalisées en utilisant diverses approches traditionnelles parmi lesquelles l'utilisation d'anticorps anti-O-GlcNAc et, bien avant que l'emploi de ces anticorps ne devienne routinier, de lectines telles les agglutinines du germe de blé (WGA, *Wheat Germ Agglutinin*), en Western blot (WB) et immunoprécipitation (IP).

Parmi les anticorps spécifiques des résidus *O*-GlcNAc le CTD110.6, un anticorps de type IgM, a été développé contre la répétition d'un peptide *O*-GlcNAcylé dans la partie C-terminale de l'ARN polymérase II (Comer et al., 2001). L'anticorps RL2, quant à lui, est un anticorps de type IgG dirigé un groupe de nucléoporines *O*-GlcNAcylées (Holt and Hart, 1986; Comer et al., 2001). Il a été montré que les anticorps RL2 et CTD110.6 présentaient des réactivités croisées avec d'autres motifs porteurs de résidus GlcNAc terminaux, tels que la *S*-GlcNAc pour le RL2 (Gorelik and Aalten, 2020). Outre le CTD110.6 et le RL2 qui sont les plus connus et les plus utilisés, il existe d'autres anticorps anti-*O*-GlcNAc tels que les anticorps HGAC39 et HGAC85 (Lee et al., 2016) et 9D1.E4 (Saha et al., 2021).

Des anticorps "site-spécifiques" ont été développés: nous pouvons citer par exemple les anticorps ciblant tau-Ser400 (Cameron et al., 2013), l'histone H2A-Ser40/Thr101 (Vaidyanathan et al., 2014), c-Myc-Thr58 (Swamy et al., 2016), TAB1-Ser395 (Rafie et al., 2017), CRMP2-Ser517 (Muha et al., 2019) et TAB3-Ser408 (Kim, 2021). Ces outils ont permis des investigations précises sur la fonction de la *O*-GlcNAc des protéines ainsi ciblées. A ce niveau, il nous paraît intéressant de citer le travail de Lee et al. dans la recherche des mécanismes par lesquels la *O*-GlcNAc favorise la survie cellulaire au cours d'un stress oxydatif aigu induit par H₂O₂ (Lee et al., 2016). Ainsi, en combinant quatre anticorps spécifiques de la *O*-GlcNAc (CTD110.6, RL2, HGAC39 et HGAC85), la lectine WGA et un marquage de type SILAC (*stable isotope labelling by amino acids in cell culture*), les auteurs ont développé une nouvelle stratégie d'immunoprécipitation appelée "G5-lectibody". Grâce à cette démarche, ils ont identifié une vingtaine de protéines *O*-GlcNAcylées et validé l'idée selon laquelle la *O*-GlcNAcylation induite par le stress contrôle des voies cellulaires diverses favorisant la survie cellulaire.

Pour l'identification des sites *O*-GlcNAcylés des protéines, la spectrométrie de masse représente la méthode la plus utilisée et la mieux adaptée. Généralement, cette approche s'accompagne d'un enrichissement des protéines ou de la protéine candidate par les anticorps décrits ci-dessus par IP, la WGA ou même par chimie *click* que nous décrirons plus loin.

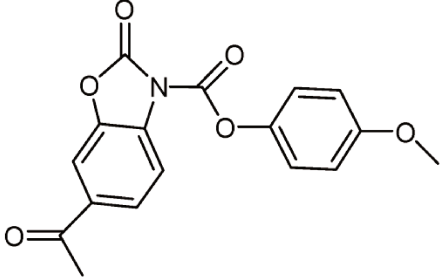
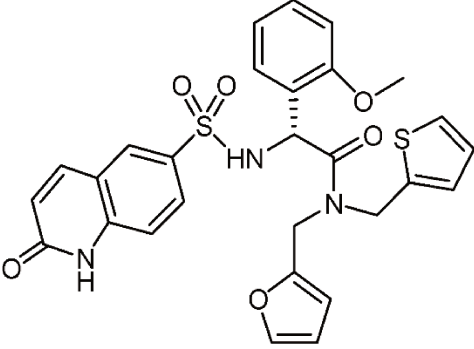
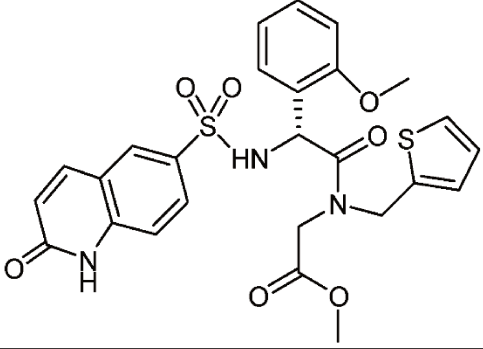
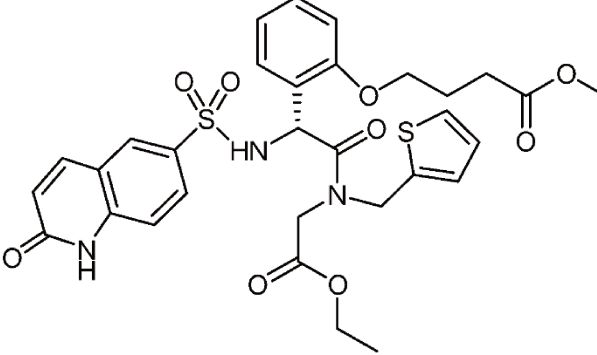
Dans le cadre de mon projet de thèse, nous avons utilisé préférentiellement l'anticorps RL2 par Western blot O-GlcNAc (RL2) pour valider la S-GlcNAcylation de T41 de la β -caténine.

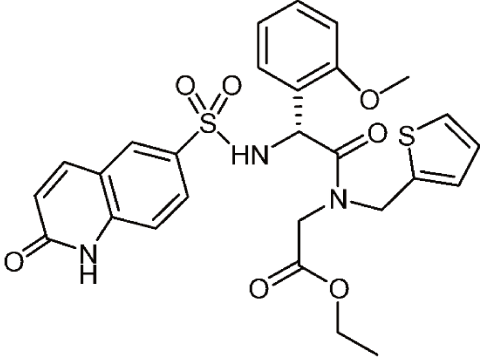
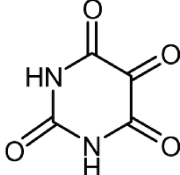
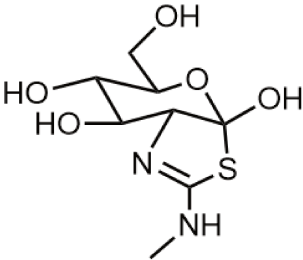
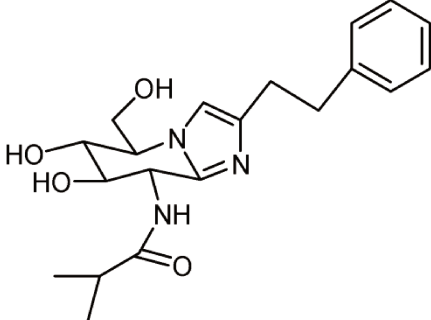
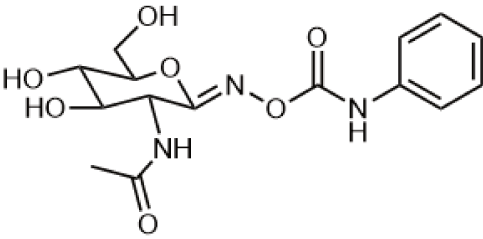
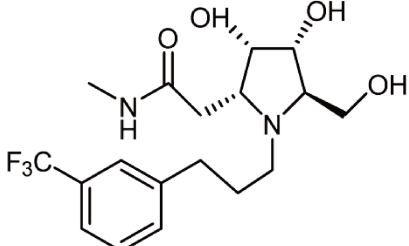
2. Inhibiteurs pharmacologiques des enzymes de la O-GlcNAcylation et de la voie des hexosamines

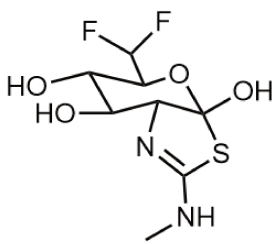
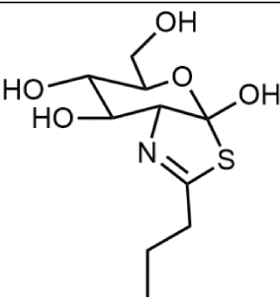
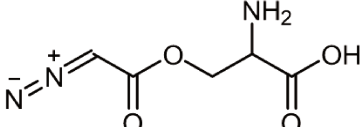
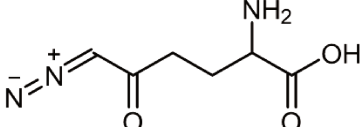
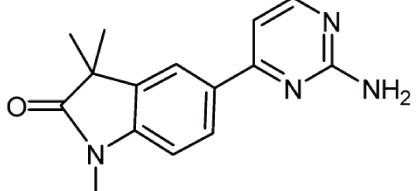
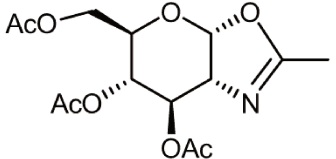
Nombre d'inhibiteurs des enzymes de la O-GlcNAc, OGT et OGA, ainsi que des enzymes de la voie des hexosamines ont été développés au cours des dernières décennies. Par simplicité et commodité, ces molécules sont listées dans le **tableau 1**.

Tableau 1. Inhibiteurs d'OGT, d'OGA et des enzymes de la voie de biosynthèse des hexosamines (adapté de Ju Kim, 2020 et Saha et al., 2021). (Ju Kim, 2020; Saha et al., 2021).

Enzymes	Nom/Caractéristiques	Structure chimique
OGT	<p>Ac₄-5S-GlcNAc Lipophile Faible IC₅₀ = 5 μM EC₅₀ = 5 μM Actif sous la forme 5S-GlcNAc-6-P</p>	
	<p>5S-GlcNHex Inhibiteur amphiphile IC₅₀ = 11 μM EC₅₀ = 5 μM</p>	
	<p>BZX1 Inhibiteur irréversible IC₅₀ = 10 μM</p>	

<p>BZX2 Inhibiteur irréversible $IC_{50} < 10 \mu M$</p>	
<p>OSMI-1 Inhibition partielle Faible solubilité aqueuse $IC_{50} = 2,7 \mu M$</p>	
<p>OSMI-2 $K_d = 140 \text{ nM}$</p>	
<p>OSMI-3 $K_d = 5 \text{ nM}$</p>	

	<p>OSMI-4 Inhibiteur le plus puissant d'OGT IC₅₀ = 60 nM, EC₅₀ ~ 3 μM Kd = 8 nM</p>	
	<p>Alloxane Demi-vie courte Faible spécificité IC₅₀ élevée = 100 μM</p>	
<p>OGA</p>	<p>Thiamet-G Inhibiteur très puissant Assez sélectif Passe la barrière hémato-encéphalique EC₅₀ = 30 nM) IC₅₀ = 1,2 nM)</p>	
	<p>GlcNAcstatine Inhibiteur très puissant Sélectivité modérée EC₅₀ = 20 nM)</p>	
	<p>PUGNAc Manque de spécificité EC₅₀ = 3 μM Ki = 46 nM</p>	
	<p>W347 Inhibiteur très puissant Assez sélectif Passe la barrière hémato-encéphalique Ki = 7.6 nM</p>	

	<p>MK-8719 Inhibiteur très puissant Assez sélectif Passe la barrière hémato-encéphalique $EC_{50} = 52,7 \text{ nM}$</p>	
	<p>NButGT Inhibiteur très puissant Assez sélectif Instable en solution $EC_{50} = 8 \text{ nM}$ $K_i = 230 \text{ nM}$</p>	
GFAT	<p>Azasérine Inhibiteur non spécifique Irréversible</p>	
	<p>DON Inhibiteur non spécifique Irréversible</p>	
UDP-GlcNAc-PPase	<p>GAL-012 Inhibiteur non spécifique $IC_{50} = 30 \mu\text{M}$</p>	
GlcNAcPM	<p>FR054 Inhibiteur spécifique</p>	

3. Rapporteurs chimiques et métaboliques utilisés dans le marquage de la O-GlcNAcylation.

L'ingénierie métabolique des oligosaccharides (*Metabolic Engineering of Oligosaccharides en anglais ou MOE*) ou marquage métabolique des glycanes est une technologie consistant à intégrer des analogues de monosaccharides non naturels dans des systèmes biologiques (Gilormini et al., 2018; Rigolot et al., 2021).

Ces analogues sont introduits dans les cellules, les organismes ou des modèles biologiques où ils interfèrent avec des voies spécifiques de glycosylation biosynthétique, et sont ensuite incorporés dans des oligosaccharides. Ainsi, ils peuvent influencer diverses activités biologiques ou servir "d'étiquettes" pour des réactions de ligation chimiosélectives de type chimie *click* ou chimie bioorthogonale (**Figures 9 et 10**).

La chimie *click* désigne un ensemble de réactions chimiques hautement efficaces et sélectives qui génèrent de nouvelles molécules en joignant de petites unités ensemble en une seule étape. Parmi ces réactions, on peut noter la réaction de cycloaddition de type CuAAC (*copper(I)-catalyzed alkyne-azide cycloaddition*), également connue sous le nom de Huisgen. Il s'agit d'une réaction entre un azide et un alcyne catalysée par le cuivre(I) (Gupta et al., 2005; Hein and Fokin, 2010). La SPAAC (*Strain-promoted alkyne-azide cycloaddition*) quant à elle se caractérise par la réaction entre un azide et un alcyne mais sans l'aide d'un catalyseur métallique (Agard et al., 2004).

Un rapporteur chimique de qualité se distingue par une solubilité élevée dans le milieu de culture cellulaire ainsi que par sa capacité à traverser la membrane cellulaire. Parmi les rapporteurs chimiques on trouve dès lors tout un ensemble d'analogues de monosaccharides tels que l'acide *N*-acétylneuraminique (Neu5Ac), la *N*-acétylmannosamine (ManNAc), la *N*-acétylgalactosamine (GalNAc), le fucose (Fuc), ou, nous concernant plus particulièrement, la *N*-acétylglucosamine (GlcNAc) (**figure 9**). Une fois introduit dans la cellule, le rapporteur subit plusieurs étapes de transformation avant de marquer ces cibles. Tout d'abord, une fois introduit dans la cellule, le rapporteur chimique subit une métabolisation cellulaire en ciblant une voie métabolique

spécifique qui se caractérise par plusieurs réactions enzymatiques successives. Il est donc très important de choisir le bon rapporteur pour cibler une voie métabolique spécifique. Ensuite, le rapporteur s'incorpore au niveau de la molécule cible qui sera exportée à la membrane (protéine membranaire), excrétée dans le milieu extracellulaire, confinée dans les compartiments cytoplasmique et nucléaire, ou exportée dans tout autre compartiment intracellulaire. Une molécule sonde tel qu'un fluorophore portant une étiquette chimique bioorthogonale est ajoutée dans le milieu de culture. Une réaction de ligation spécifique se réalise entre le rapporteur et la sonde. En se basant sur la nature de la sonde, la technique de détection ou de visualisation est alors choisie, un fluorophore pour une détection directe (**Figure 9**) ou une biotine pour une détection indirecte (**Figure 10**) (J. Pedowitz and R. Pratt, 2021).

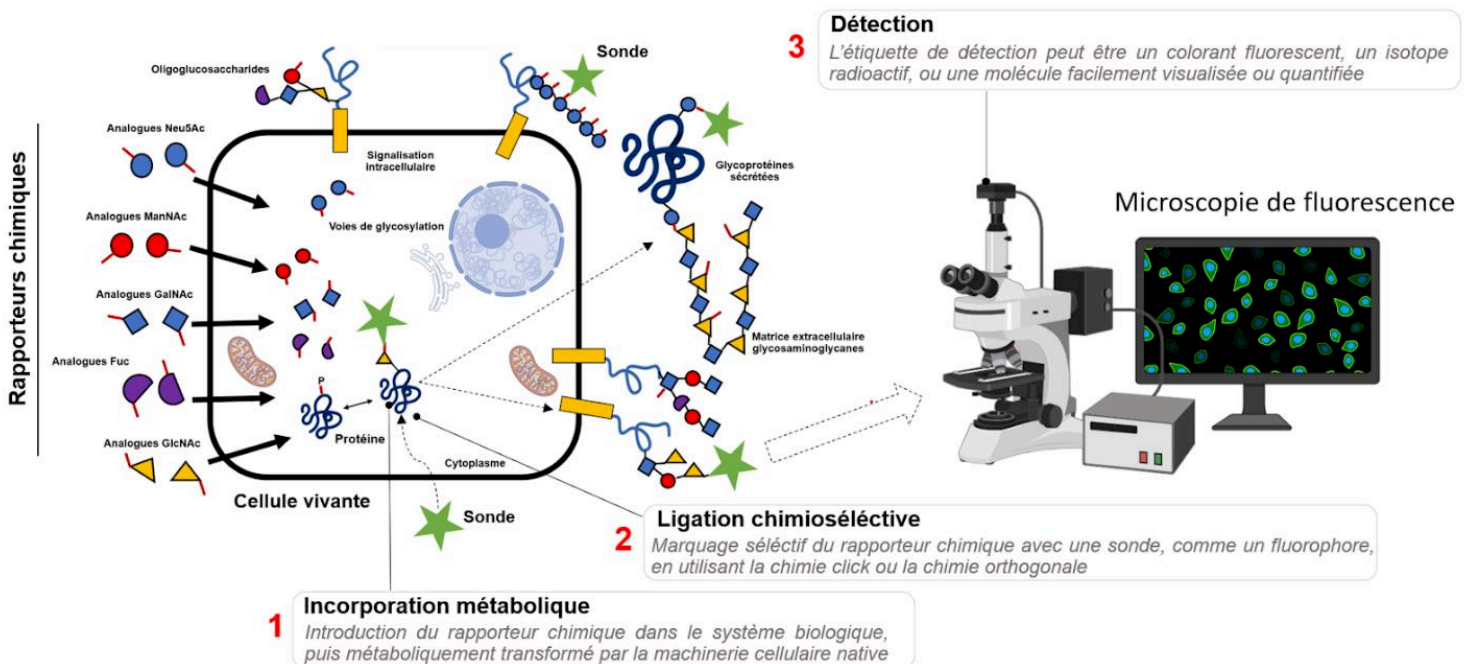
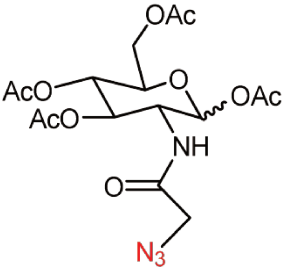
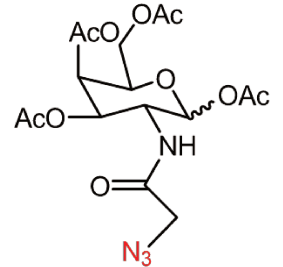
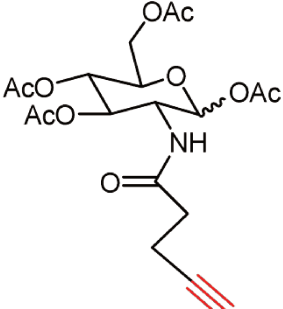


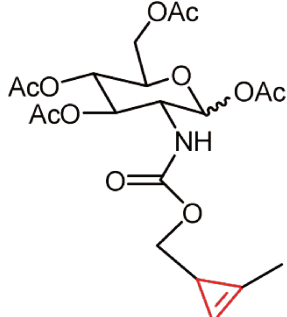
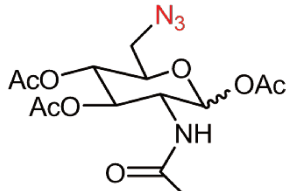
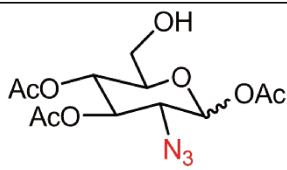
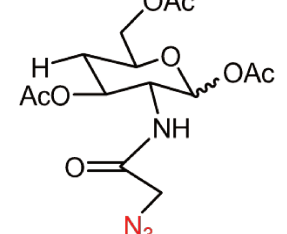
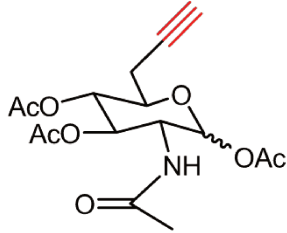
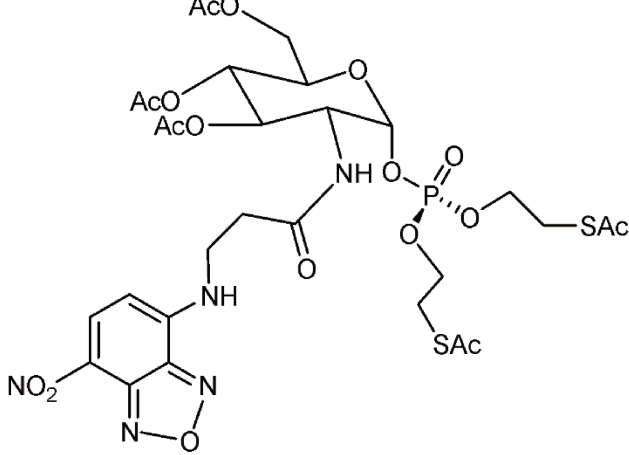
Figure 9. Perspectives de l'ingénierie oligosaccharidique métabolique (MOE) sur la détection des voies de glycosylation cellulaire. Neu5Ac, acide *N*-acétylneuraminique ou NANA ; ManNAc, *N*-acétylmannosamine ; GalNAc, *N*-acétylgalactosamine ; Fuc, fucose ; GlcNAc, *N*-acétylglucosamine.

Cette stratégie est largement utilisée en chémoglycobiologie pour étudier les protéines *O*-GlcNAcylées, la voie des hexosamines et la voie de sauvetage de la GlcNAc ou de la GalNAc (J. Pedowitz and R. Pratt, 2021).

Dans le but de réduire les problèmes rencontrés au cours du traitement avec ces précurseurs métaboliques et d'assurer toutes les exigences d'une bonne perméabilité cellulaire, d'une non-cytotoxicité, d'une spécificité optimale pour la protéine cible ainsi qu'une efficacité du marquage, les chimobiologistes ont synthétisé et développé de nouvelles molécules avec des structures chimiques plus adaptables et plus optimisées. Nous illustrons dans le **tableau 2**, les différentes catégories de ces analogues de GlcNAc en précisant leur objectif d'utilisation et leur structure chimique.

Tableau 2. Les différentes catégories de métabolites utilisés pour marquer les protéines O-GlcNAcylées.

Objectif de l'utilisation	Formule	Structure chimique
Métabolites de marquage de la O-GlcNAcylation	Ac ₄ GlcNAz	
	Ac ₄ GalNAz	
	Ac ₄ GlcNAIk	

	Ac ₄ GlcNCyoc	
	Ac ₃ 6AzGlcNAc	
	Ac ₃ 2AzGlc	
	Ac ₃ 4dGlcNAz	
	Ac ₃ 6AlkGlcNAc	
Métabolites fluorescents	Ac ₃ GlcN-β-Ala-NBD-1-P(Ac-SATE) ₂	

	<p>1,3-Pr₂GalNAz : R₁ = H et R₂ = Pr 1,6-Pr₂GalNAz : R₁ = Pr et R₂ = H</p>	
Métabolites photo-actives	<p>Ac₃GlcNDAz-1- P(Ac-SATE)₂</p>	

Dans la **figure 10**, nous présentons un schéma de réaction bioorthogonale entre la glycoprotéine chimiquement marquée et des sondes fluorescentes et/ou biotinylées. Ces approches permettent de visualiser et/ou d'isoler des protéines *O*-GlcNAcylées pour des analyses ultérieures, par spectrométrie de masse par exemple.

Il existe d'autres dérivés de sucres modifiés - ou rapporteurs chimiques - largement utilisés en MOE. Nous allons détailler les progrès réalisés dans le développement de ces rapporteurs chimiques.

Le groupe de Carolyn Bertozzi a utilisé le 2-azidoacétamido-2-désoxy-glucopyranose (Ac₄GlcNAz) pour réaliser un marquage métabolique de la *O*-GlcNAcylation en cellules vivantes (Boyce et al., 2011). Il a été montré que le groupement azoture biologiquement inerte incorporé chimiquement dans le groupement acétamido GlcNAc était bien toléré par les enzymes biosynthétiques de la voie de récupération de la GlcNAc. Cela conduit à la formation d'UDP-GlcNAz utilisé par l'OGT pour incorporer le *O*-GlcNAz sur des résidus Ser/Thr, générant ainsi des protéines chimiquement marquées portant une fonctionnalité azoture pour une détection ultérieure par des réactions bioorthogonales. Malgré ces avantages, il a été observé que ce rapporteur présentait un marquage faible et un degré relativement élevé de glycosylation non spécifique (Boyce et al., 2011).

En outre, le groupe de David Vocadlo a utilisé le rapporteur acétylgalactosamine *N*-azidoacétylée per-*O*-acétylée (Ac₄GalNAz) à la place du Ac₄GlcNAz. Le fait que le traitement avec l'Ac₄GalNAz présente un marquage plus élevé par rapport à l'Ac₄GlcNAz provient d'une conversion isomérique efficace d'UDP-GalNAz en UDP-GlcNAz sans se soucier d'étapes inefficaces dans la voie de sauvetage de la GalNAc contrairement au traitement avec la Ac₄GlcNAz (Liu et al., 2017).

Pour éviter l'utilisation des rapporteurs Ac₄GlcNAz et Ac₄GalNAz, Cioce et collaborateurs ont utilisé la *N*-pentynylglucosamine per-*O*-acétylée (Ac₄GlcNAIk) (Cioce et al., 2021), Haiber et collaborateurs ont utilisé le 6-azido-6-désoxy-GlcNAc per-*O*-acétylé (Ac₃6AzGlcNAc) (Haiber, 2023), et Doll et collaborateurs ont utilisé le 6-alcynyl-6-désoxy-GlcNAc per-*O*-acétylé (Ac₃6AlkGlcNAc) (Doll, 2018). Les rapporteurs métaboliques utilisés par ces différents auteurs - Ac₄GlcNAIk, Ac₃6AzGlcNAc et Ac₃6AlkGlcNAc - présentent une spécificité améliorée par rapport à l'Ac₄GlcNAz et l'Ac₄GalNAz.

Le groupe de Sheibani a utilisé quant à lui le rapporteur chimique Ac₄GlcNAIk révélant l'effet anti-apoptotique de la *O*-GlcNAcylation associé à la protéine p53 au niveau de péricytes rétiens (Gurel et al., 2014). Pratt et ses collaborateurs ont utilisé ce même rapporteur chimique. Il a été constaté que le GlcNAIk était incorporé et ôté des protéines cibles à des niveaux similaires à ceux du GlcNAz, et était difficilement interconverti en GalNAIk *via* GALE, conduisant à un faible niveau de glycoprotéines de type mucines *O*-liées marquées. Le GlcNAIk est donc un rapporteur plus spécifique des processus de *O*-GlcNAcylation en comparaison aux rapports chimiques cités précédemment (Chuh et al., 2017).

L'Ac₃6AzGlcNAc a été utilisé dans le but de cibler la *O*-GlcNAcylation de la protéine Blimp-1 (Chen et al., 2022). Il a été montré que la *O*-GlcNAcylation de cette protéine bloque son activité transcriptionnelle, et est associée à la migration et à l'invasion des cellules cancéreuses du sein. Chuh et al. ont montré l'effet anti-apoptotique de la *O*-GlcNAcylation sur la caspase-8 en utilisant ce même rapporteur (Chuh et al., 2017). Ces deux études montrent que le rapporteur Ac₃6AlkGlcNAc présente un meilleur rapport signal sur bruit, et une capacité de détection plus élevée suite à des réactions

bioorthogonales par cycloaddition azoture-alcyne catalysée par le cuivre par rapport à l'Ac₃6AzGlcNAc.

Dans d'autres études ayant toujours pour objectif de suivre sélectivement les protéines *O*-GlcNAcylées, Wang et collaborateurs ont utilisé le rapporteur 4-désoxy-N-azidoacétylglucosamine per-*O*-acétylée (Ac₃4dGlcNAz) (Li et al., 2016), alors que le groupe de Pratt a utilisé le glucose 2-azido-2-désoxy tri-*O*-acétylé (Ac₃2AzGlc) (Zaro et al., 2017). Ces deux groupes ont montré que ces deux rapporteurs chimiques possèdent une haute résistance à l'hydrolyse par l'OGA. Par contre, il a été montré que la forme per-*O*-acétylée du 2-Azido-Glc à concentration élevée (~ 200 µM) a un effet toxique sur les cellules, tandis que sa toxicité est très faible lorsque l'acétyle du C6 est enlevé pour libérer le groupe hydroxyle.

En adoptant un rapporteur chimique contenant du cyclopropényle peracétylé (Ac₄GlcNCyoc), il a été révélé une amélioration de l'efficacité du marquage par rapport aux rapporteurs mentionnés dans les études précédentes (Hassenrück and Wittmann, 2019). Malgré cet avantage, il a été reporté que la per-*O*-acétylation de ce rapporteur pouvait induire une *S*-glycomodification non spécifique des protéines cibles. Pour résoudre le problème de *S*-glycomodification secondaire, le groupe de Chen a développé le rapporteur 1,6-di-*O*-propionyl-N-azidoacétylgalactosamine (1,6-Pr2GalNAz) (Hao et al., 2023). L'utilisation d'un dérivé proche, la 1,3-di-*O*-propionyl-N-azidoacétylgalactosamine partiellement protégée par les groupes acétyle (1,3-Pr2GalNAz), a montré que la *O*-GlcNAcylation favorise la stabilité, l'activité transcriptionnelle et l'interaction du facteur ESRRB impliquée dans la pluripotence (Hao et al., 2019).

Pour conclure sur cette partie, la stratégie des rapporteurs chimiques métaboliques représente une approche puissante et présentant des avantages tels que la spécificité et la sensibilité des rapporteurs chimiques, et le suivi en temps réel des protéines *O*-GlcNAcylées. Malgré ces qualités, les chimobiologistes cherchent toujours à produire un rapporteur chimique doté d'une bonne perméabilité cellulaire, d'une non-cytotoxicité, d'une spécificité optimale pour la protéine cible ainsi qu'une efficacité du marquage.

L'idée principale de cette partie est alignée avec l'objectif de ma thèse : le suivi de la O-GlcNAcylation des protéines en temps réel.

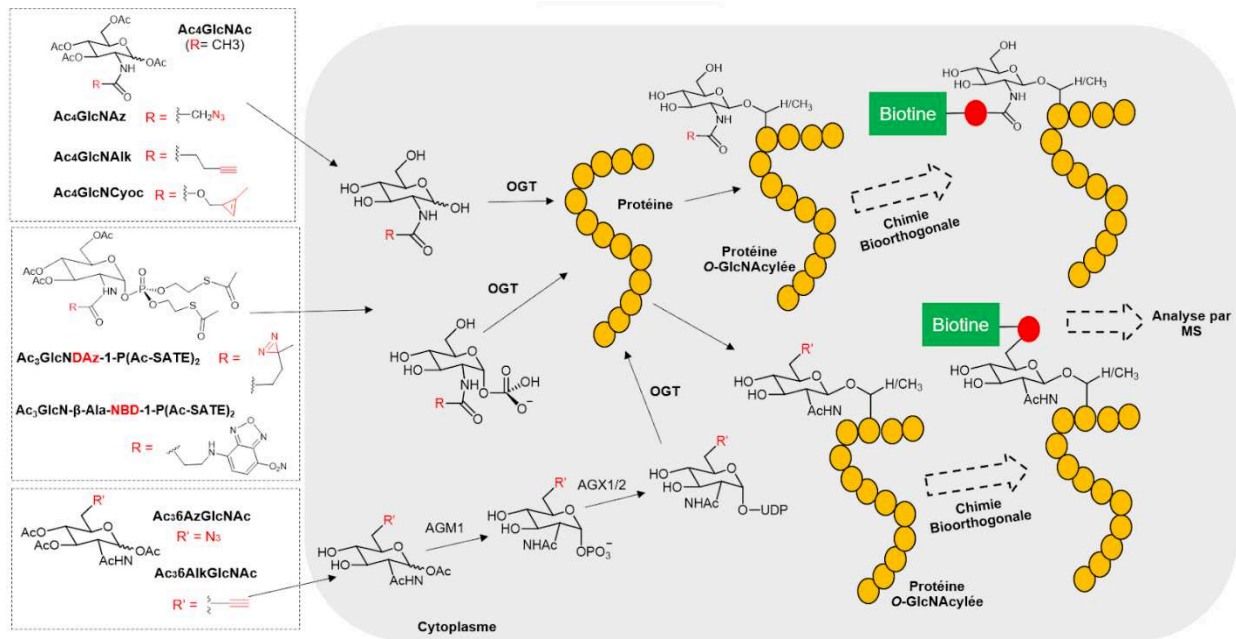


Figure 10. Stratégie des rapporteurs chimiques métaboliques utilisés pour le marquage de la O-GlcNAcylation. Les analogues de GlcNAc modifiés sont introduits dans la cellule et convertis en dérivés de l'UDP-GlcNAc. La réaction catalysée par l'OGT permet le marquage des protéines ainsi modifiées par des mimétiques de la O-GlcNAc qui peuvent être identifiées ou tracées par diverses sondes telle que la biotine.

4. Imagerie spécifique des protéines O-GlcNAcylées via l'utilisation de la stratégie FRET

La visualisation de la O-GlcNAc d'une protéine d'intérêt en cellule individuelle peut être réalisée par une technique basée sur le transfert d'énergie par résonance de Förster (FRET, *Förster Resonance Energy Transfer*). Cette méthode représente une alternative aux anticorps à condition que les deux fluorophores soient à proximité, c'est-à-dire à une distance de moins de 10 nm (Filius et al., 2024).

L'idée principale de cette technique réside dans le marquage d'une protéine d'intérêt avec une protéine fluorescente telle que l'EGFP (*enhanced green fluorescent protein*), et la O-GlcNAcylation de cette protéine avec un second fluorophore par marquage métabolique alliant réactivité chimique et bioorthogonalité (par exemple, Ac₄GalNAz et

alcyne-TAMRA). Dans ce cas, les deux fluorophores (par exemple, EGFP et TAMRA) installés sur la même protéine forment une paire donneur-accepteur FRET, et l'occurrence du FRET intramoléculaire conduit au raccourcissement du temps de vie de fluorescence (τ) du premier fluorophore (EGFP). En utilisant cette stratégie, Lin et collaborateurs ont visualisé par imagerie le statut de *O*-GlcNAcylation de la protéine tau et de la β -caténine à l'intérieur même des cellules (Lin et al., 2015).

Dans ce même contexte, notre équipe a validé une technologie innovante pour visualiser le statut de la *O*-GlcNAcylation de la β -caténine en cellules HeLa par imagerie basée sur le FRET. Cette stratégie de rapporteur chimique bioorthogonal basée sur le double marquage de la β -caténine avec une protéine fluorescente verte (GFP) pour la séquence protéique combinée à une sonde d'imagerie chimiquement liée pour la MPT a généré un signal FRET qualitatif, rapide, et facile à suivre (Kasprowicz et al., 2020) **(Figure 11)**.

Cette stratégie présente donc beaucoup d'avantages pour suivre la *O*-GlcNAcylation des protéines en temps réel dans les cellules vivantes. Parmi ces avantages, on peut noter une spécificité de détection de la protéine cible (Kasprowicz et al., 2020), une sensibilité élevée offerte par la réaction FRET (Liu et al., 2024), ainsi qu'une rapidité d'exécution d'expérience (Kasprowicz et al., 2020). L'objectif de ma thèse a été de mettre en place une technologie permettant également de suivre la *O*-GlcNAcylation de la β -caténine mais sur un site précis, la thréonine 41.

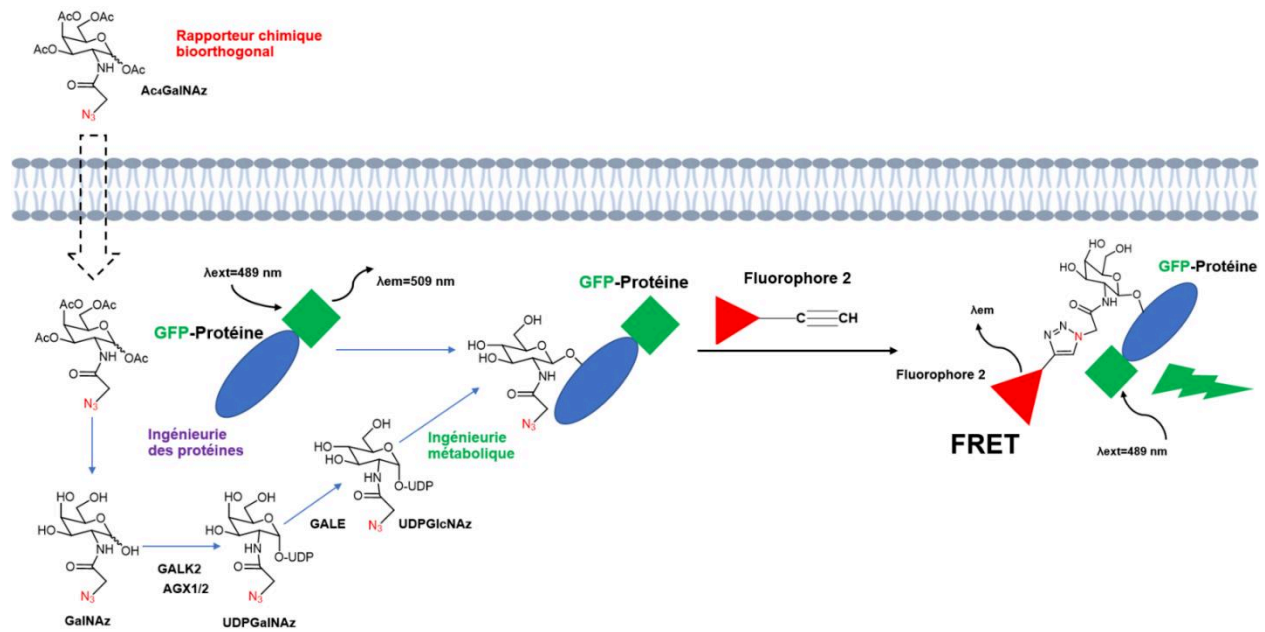


Figure 11. Imagerie basée sur FRET du double marquage des protéines et des glycanes. FRET, transfert d'énergie de résonance de Förster. Adapté de (Kasprowicz et al., 2020). Ac₄GalNAz, 1,3,4,6-Tétra-O-acétyl-N-azidoacétylgalactosamine ; GalNAz, N-azidoacétylgalactosamine ; UDPGalNAz, UDP-N-azidoacétylgalactosamine ; UDPGlcNAz, UDP-N-azidoacétylglucosamine ; GALK2, galactokinase 2 ; AGX1/2, uridylypyrophosphorylase 1/2 ; GALE, UDP-galactose 4'-épipimérase.

5. Expansion du code génétique dans l'optique de décrypter les processus de O-GlcNAcylation

Peter Schultz est un pionnier de la chimobiologie et de la chimie des protéines, reconnu pour ses contributions majeures dans le développement de technologies visant à étendre le code génétique des organismes vivants. Sa recherche a ouvert de nouvelles voies pour l'introduction d'acides aminés artificiels dans les protéines, élargissant ainsi considérablement le potentiel fonctionnel des systèmes biologiques. Cette expansion du code génétique (*Genetic Code Expansion (GCE) en anglais*) a des implications importantes dans de nombreux domaines, y compris la biologie synthétique, la médecine et les biotechnologies, ouvrant la voie à une gamme diversifiée d'applications, telles que la conception de médicaments, la biocatalyse et la création de matériaux fonctionnels.

L'expansion du code génétique est une technique novatrice qui permet l'introduction précise d'acides aminés chimiquement modifiés sur des sites spécifiques de protéines cibles en leur donnant des propriétés chimiques, structurales et fonctionnelles uniques. La technique a été inspirée dans les années 1990 par l'utilisation des ARNt suppresseurs non-sens naturellement présents. Cette stratégie repose sur trois composants clés : l'aminacyl ARNt synthétase (aaRS), l'ARNt et l'acide aminé non canonique. Dans cette stratégie, un aaRS charge exclusivement l'acide aminé désiré sur un ARNt spécifique (Shandell et al., 2021) (**Figure 12**).

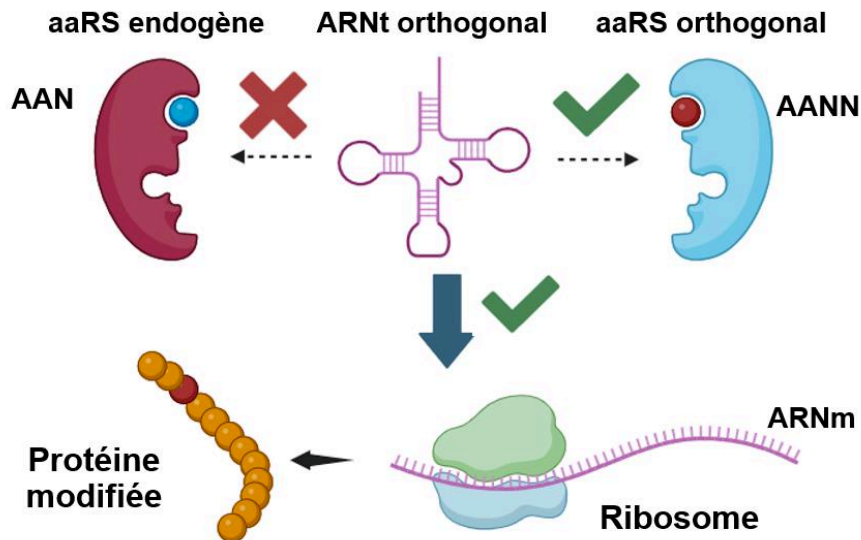


Figure 12. Rôle des ARNt orthogonaux dans l'expansion du code génétique. AAN: Acide aminé naturel ; AANN: Acide aminé non naturel.

Dans cette section, nous détaillerons la méthode d'expansion du code génétique et son importance dans l'étude et l'analyse de la *O*-GlcNAcylation. Cette technologie repose sur l'utilisation d'une paire orthogonale aminoacyl-ARNt synthétase (aaRS)/ARNt pour guider l'incorporation d'un acide aminé non canonique en des sites spécifiques de protéines d'intérêt. Elle est rendue possible par l'utilisation d'un codon d'arrêt *ambre*, UAG, introduit au site désiré dans un gène donné (Hao et al., 2024) (**Figure 13**).

Cette technologie exige que l'orthogonalité d'une paire aaRS/ARNt soit définie par rapport à chaque hôte, que ce soit une bactérie comme *Escherichia coli* ou des cellules eucaryotes (Abdulsalam et al., 2024). De nombreuses études ont réussi à incorporer des centaines d'acides aminés structurellement divers dans des protéines avec une grande efficacité (Francis, 2013; Doura et al., 2024).

Plusieurs facteurs ont été identifiés comme ayant un impact sur l'efficacité (le rapport entre la protéine complète et la protéine tronquée) et le rendement (quantité totale de protéine produite) des systèmes d'expression des acides aminés non canoniques au niveau des cellules hôtes. Parmi ces facteurs, on peut citer la souche d'expression, le contexte de séquence du codon *ambre*, l'expression et l'activité de l'aaRS et de l'ARNt, ainsi que l'expression d'autres composants. Ainsi, le rendement varie généralement de

10 à 100%. Par conséquent, l'efficacité et le rendement diffèrent selon les protéines et les sites au sein d'une même protéine. Diverses stratégies ont été rapportées pour améliorer l'efficacité, le rendement et la spécificité de l'incorporation de l'acide aminé non canonique, notamment l'optimisation des séquences orthogonales d'aaRS et d'ARNt et du niveau d'expression (Kim et al., 2024), ainsi que des modifications du ribosome (Dunkelmann et al., 2024), des facteurs d'élongation et des facteurs de libération (RF1 chez *E. coli* et eRF1 chez les eucaryotes) mettant fin à la synthèse polypeptidique due aux codons *ambre* (Perdiguero and Liang, 2024).

Cette stratégie a été appliquée non seulement sur des bactéries, mais également sur des cellules eucaryotes animales et végétales. Pour améliorer l'incorporation d'un acide aminé non canonique sur plusieurs sites de codons *ambre* dans un seul polypeptide, ainsi que d'améliorer et empêcher la diminution des rendements en protéines, Y. Kim et al. et Perdiguero et Liang ont développé un ribosome orthogonal 10 à 100 fois plus efficace que le ribosome naturel (Kim et al., 2024; Perdiguero and Liang, 2024). Ainsi, le problème de compétition entre le décodage des codons *ambre*, pour l'incorporation d'un acide aminé non canonique, et les facteurs de libération, a-t-il été en partie résolu par la création de plusieurs souches d'*E. coli* sans RF1.

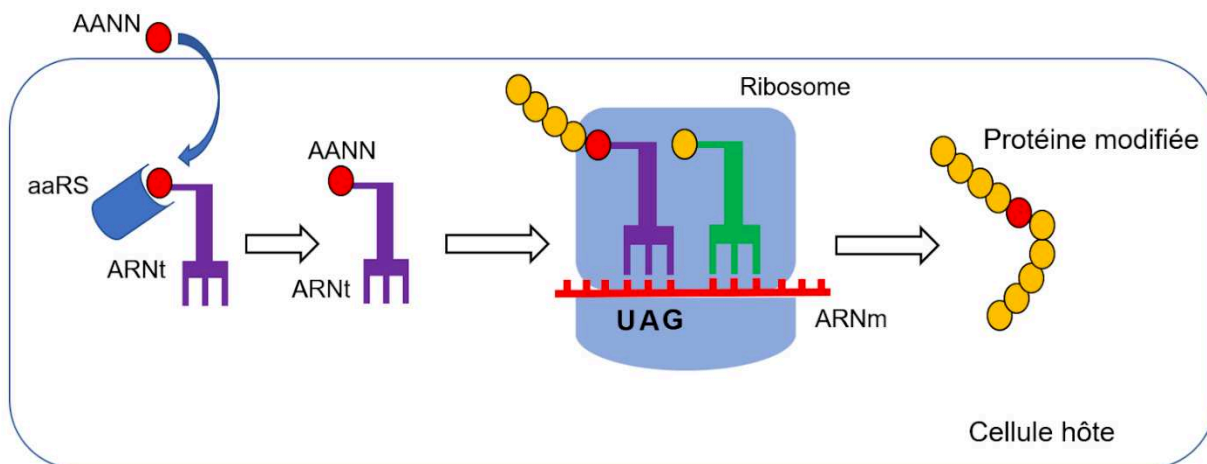


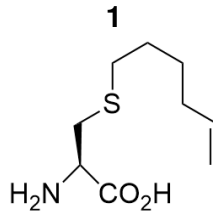
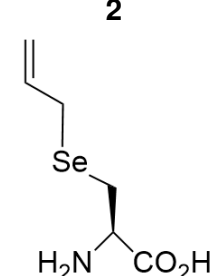
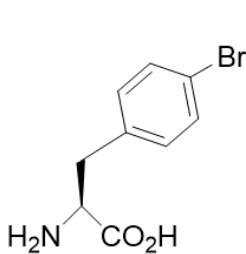
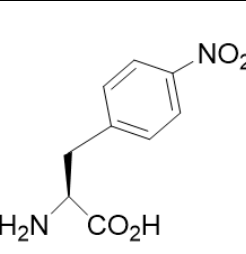
Figure 13. Stratégie de l'expansion du code génétique. L'AANN est introduit dans les cellules hôtes contenant l'aminoacyl-ARNt synthétase orthogonale (aaRS) et l'ARNt orthogonal. L'aaRS orthogonal transfère l'acide aminé sur l'ARNt, et l'ARNt amino-acylé est utilisé pour décoder les codons UAG dans l'ARNm, produisant une protéine modifiée contenant un acide aminé non canonique.

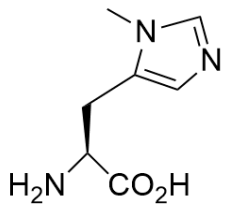
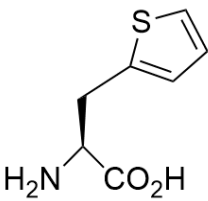
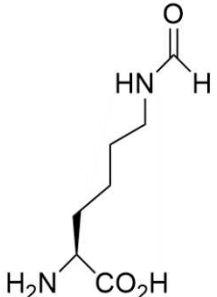
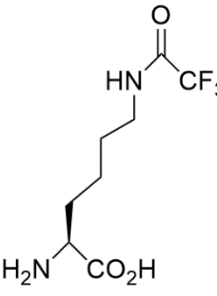
L'évolution de cette stratégie a connu beaucoup de progrès au travers du maintien de l'efficacité des paires aaRS/ARNt de différents hôtes autres qu'*Escherichia coli*. Cependant, les ARNt doivent répondre à deux critères essentiels : l'orthogonalité aux aaRS de la cellule hôte et la coopérativité avec le reste de la machinerie cellulaire.

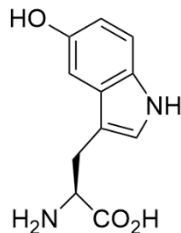
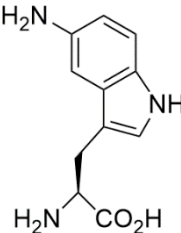
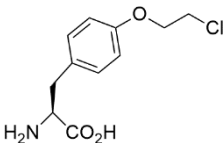
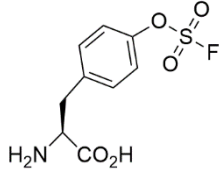
Par définition, un acide aminé non naturel ou non canonique est un composé chimique de synthèse arborant une structure similaire aux acides aminés protéinogènes mais ne pouvant pas être introduit dans les protéines au cours de la traduction. De nombreux acides aminés non canoniques ont été introduits avec succès dans tout un ensemble de protéines produites en cellules procaryotes, mais également en cellules eucaryotes telles que les levures (Nödling et al., 2019), quelques-uns d'entre eux sont présentés dans le **tableau 3**.

Pour mon projet de thèse, nous avons choisi l'acide aminé non canonique phénylsélénocystéine (PhSeCys) pour l'introduire en position T41 de la β -caténine en cellules procaryotes *E. coli*. BL21 (DE3) grâce à une paire aaRS (PhSeCys)/ARNt.

Tableau 3. Exemples d'acides aminés accompagnés de leur paire orthogonale aaRS/ARNt. Adapté de (Nödling et al., 2019).

Acide aminé non canonique	aaRS	Mutations	ARNt
Dérivés - Cystéine et sélénocystéine			
<p>1</p> 	<i>EcLeuRS</i>	M40I, T252A Y499I, Y527A, H529G	<i>EctRNA</i> ^{Leu} _{CUA}
<p>2</p> 	<i>MbPylRS</i>	C313W W382T	<i>MbPylRS</i> tRNA
Dérivés Phénylalanine			
<p>3</p> 	<i>EcTyrRS</i>	Y37V, D182S, F183M et D265R	<i>EctRNA</i> ^{Tyr} _{CUA} <i>BstRNA</i> ^{Tyr} _{CUA}
<p>4</p> 	<i>EcTyrRS</i>	Y37V, D182S F183M et D265R	<i>EctRNA</i> ^{Tyr} _{CUA} <i>BstRNA</i> ^{Tyr} _{CUA}
Dérivés Histidine			
<p>5</p>	<i>MaPylRS</i>	L121M L125I Y126F M129A V168F	<i>MatRNA</i> ^{Pyl} _{CUA}
	<i>MbPylRS</i>	L270I Y271F L274G C313F Y349F	<i>MbtRNA</i> ^{Pyl} _{CUA}

			
6 	MbPylRS	L270I, Y271F, L274G C313F, Y349F	<i>MbtRNA</i> ^{Pyl} _{CUA}
Dérivés Lysine			
7 	MbPylRS	L266M L270I Y271F L274A C313F	<i>MbtRNA</i> ^{Pyl} _{CUA}
8 	MmPylRS	L301M, Y306L L309A, C348F	<i>MmtRNA</i> ^{Pyl} _{CUA}
Dérivés Tryptophane			
9	<i>EcTrpRS</i>	S8A V144S V146A S8A V144G V146C	<i>EctRNA</i> ^{Trp} _{CUA}

			
10 	<i>Ec</i> TrpRS	S8A V144S V146A	<i>Ect</i> RNA ^{Trp} _{CUA}
Dérivés Tyrosine			
11 	<i>Ec</i> TyrRS	Y37V D182S F183M D265R, Y37V, D165G D182S, F183M, L186A, D265R	<i>Ect</i> RNA ^{Tyr} _{CUA}
12 	<i>Mm</i> PylRS	N346T C348I Y384L W417K	<i>Mmt</i> RNA ^{Pyl} _{CUA}

Un acide aminé non canonique peut être reconnu par une aaRS orthogonale (acides aminés #1-4 et #6-12, **tableau 3**) ou plusieurs aaRS orthogonales (acide aminé #5) pour être fixé sur un ARNt orthogonal (acides aminés #1, 2, 6-12) ou plusieurs ARNt orthogonaux (acides aminés #3, 4 et 5), pour être incorporé dans la protéine d'intérêt en cours de synthèse dans la cellule hôte. Une aaRS peut même reconnaître plusieurs acides aminés différents, voire des acides aminés n'appartenant pas à la même famille: c'est le cas par exemple de la synthétase *Mb*PylRS qui reconnaît les acides aminés #5-6, dérivés de l'histidine, et l'acide aminé #7, dérivé de la lysine, ou la synthétase

EcTyrRS qui reconnaît les acides aminés #3-4, dérivés de la phénylalanine, et les acides aminés #11-12, dérivés de la tyrosine. De plus, différents acides aminés non canoniques peuvent se fixer sur un même ARNt, c'est le cas de l'acide aminé #5, dérivé de l'histidine, et des acides aminés #6 et 7, dérivés de la lysine qui se fixent sur le même ARNt *MbtRNA*(Plyl/CUA).

Nous allons aborder maintenant les différentes technologies permettant d'étendre le code génétique d'organismes vivants appelée stratégie d'expansion du code génétique. Parmi ces technologies on retrouve la mutagenèse dirigée, la suppression du codon stop *ambre* et l'utilisation de ribosomes modifiés.

Liu et Naismith ont développé un protocole en une seule étape appelé "mutagenèse dirigée sur site simple ou multiple", en introduisant des oligonucléotides par PCR dirigée sur un ou plusieurs sites spécifiques de l'ADN d'intérêt de sorte à introduire un acide aminé non canonique *via* les paires aaRS/ARNt orthogonales (Liu and Naismith, 2008).

Brown et collaborateurs ont récemment adapté le concept de suppression d'*ambre* (UAG) pour introduire des acides aminés non canoniques chez les embryons de poissons-zèbres (Brown et al., 2024). Pour ce faire, ils ont injecté l'ARNm codant la synthétase orthogonale PyIRS, l'ADN codant la protéine d'intérêt comportant un codon *ambre*, l'ARNt orthogonal PyIT, et l'acide aminé non canonique directement dans l'ovocyte fécondé. Ils ont obtenu par cette technique une expression importante en protéines modifiées dans les embryons de grenouille. Par conséquent, la technique de suppression d'*ambre* permet de produire des protéines avec des séquences et des fonctions personnalisées *in vivo*.

Il est également possible de produire des ribosomes modifiés pour introduire des acides aminés non canoniques par expansion du code génétique. Sidney Hecht a développé cette technique en introduisant des plasmides codant des ARNr 23S bactériens mutés dans la région 2057-2063, centrale pour l'activité peptidyl-transférase en présence d'érythromycine (Hecht, 2022). Ainsi, en présence de concentrations modérées en érythromycine, l'activité des ribosomes sauvages sensibles à l'antibiotique est bloquée, et celle des ribosomes modifiés maximisée.

La stratégie d'expansion du code génétique pourrait être étendue au domaine d'étude des MPT, la O-GlcNAcylation nous concernant. En déterminant la structure d'un peptide synthétique dérivé de la hOGA S-GlcNAcylée en complexe avec une protéine modèle, Gorelik et al. ont démontré que le S-GlcNAc était un mimétique intéressant de la O-GlcNAc (Gorelik et al., 2019). Ces mêmes auteurs ont réussi à introduire génétiquement un résidu S-GlcNAc non hydrolysable par l'OGA par mutagenèse dirigée de résidus hydroxylés en cystéine, et en exploitant le fait que l'OGT possède une activité S-GlcNAc transférase (Gorelik and Aalten, 2020).

Cette stratégie a été utilisée pour évaluer l'impact d'autres MPT, en l'occurrence celui de la phosphorylation et de l'acétylation sur l'interaction de la protéine tau à la tubuline soluble, et de la capacité de ces MPT à polymériser les tubulines en microtubules (Ramirez et al., 2024). Chen et al., ont appliqué cette stratégie à la thiorédoxine réductase 1 et au proto-oncogène Akt pour en étudier la phosphorylation et la méthylation (Chen et al., 2018). Enfin, Schmidt et Summerer, quant à eux, ont tiré avantage de cette approche pour travailler sur la méthylation, l'acétylation et la crotonylation des lysines des protéines histones, ainsi que sur l'acétylation de la lysine des facteurs de transcription comme le facteur p53 (Tumour protein 53) et RcsB (Regulator of capsule synthesis) (Schmidt and Summerer, 2014).

Cette stratégie représente donc un outil prometteur pour l'étude des MPTs en cellules vivantes. C'est pourquoi nous avons opté pour cette approche expérimentale dans mon projet de thèse (**Figure 14**).

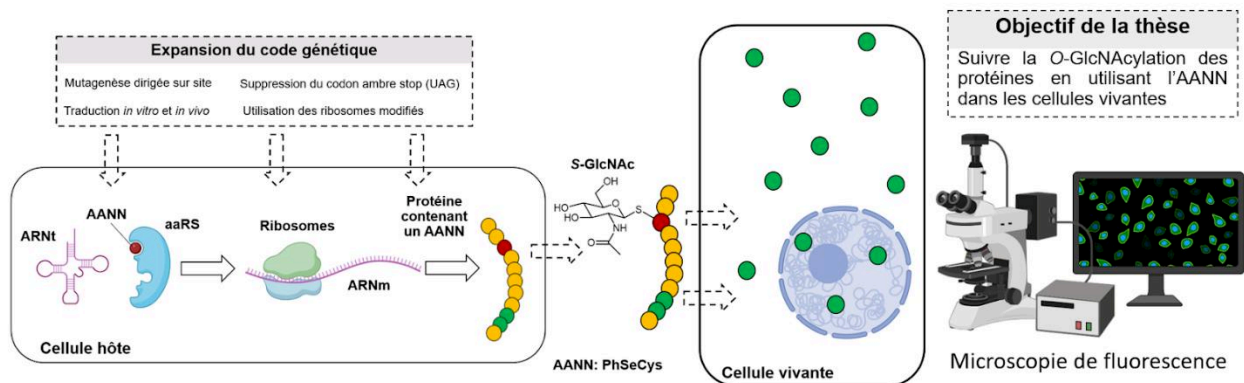


Figure 14. Application de l'expansion du code génétique à notre projet de thèse.

Projet de thèse

Objectifs des travaux de thèse

L'objectif principal de ma thèse a été de développer une méthode pour produire des protéines traçables en fonction de leurs MPTs. La β -caténine a été choisie comme modèle protéique. Le second objectif était d'étudier le comportement de la β -caténine dans divers contextes physio-pathologiques. Nous nous sommes concentrés sur la β -caténine en raison de ses quatre sites de *O*-GlcNAcylation dans la région *N*-terminale (S23, T40, T41 et T112), jouant un rôle dans sa stabilité et son activité biologique (Olivier-Van Stichelen et al., 2014). Dans le contexte de la biologie synthétique, l'incorporation sélective d'un acide aminé non canonique synthétisé chimiquement représente une technologie prometteuse pour la construction de protéines sur mesure. Nous avons spécifiquement ciblé la *O*-GlcNAcylation et la phosphorylation de la thréonine T41 (gT41 et pT41 respectivement) de la β -caténine en raison de leur rôle dans la sensibilité de la protéine au protéasome (Olivier-Van Stichelen et al., 2014; Fu et al., 2024). Pour ce faire, nous avons choisi des formes stables de ces MPT, à savoir *S*-GlcNAc (De Leon et al., 2017; Wang et al., 2018) et phospho-mimétique (T41E) (Rogerson et al., 2015).

Concernant la *S*-GlcNAc, il s'agit de la version la plus proche de la *O*-GlcNAc actuellement décrite. L'utilisation de formes stables non hydrolysables est un préalable essentiel à la mise en œuvre de notre technologie : en effet nous aurions pu envisager de concevoir des protéines modifiées par les formes classiques *O*-GlcNAc et *O*-phosphate. Mais, sitôt internalisées dans les cellules, ces versions naturelles seraient déglycosylées ou déphosphorylées par les enzymes cellulaires, respectivement par l'OGA et les protéines phosphatases, et probablement reglycosylées par l'OGT et/ou rephosphorylées par des kinases sur divers sites, rendant nos protéines fluorescentes indescriptibles en termes de modifications. Les protéines générées seront ainsi glycosylées et phosphorylées de manière stable pour la visualisation dans les cellules grâce à l'incorporation de *S*-GlcNAc et la mutation T41E en lieu et place des résidus *O*-GlcNAc et *O*-phosphate hautement labiles.

Dans ce contexte, nous avons développé cette technologie en utilisant la lignée cellulaire HeLa du fait de sa grande souplesse d'utilisation et de son caractère bien établi.

Ensuite, nous avons étendu cette approche à d'autres modèles cellulaires provenant du côlon (CCD841CoN) et du foie (Hep3B). Ces modèles cellulaires complémentaires permettront d'explorer plus en profondeur les mécanismes moléculaires et les réponses cellulaires associés aux modifications spécifiques des protéines *O*-GlcNAcylées dans ces tissus.

Ainsi, notre stratégie mettant en œuvre une méthodologie innovante et originale pour détecter les *O*-GlcNAc/phospho-isoformes de la β -caténine est basée sur l'expansion du code génétique par incorporation d'un acide aminé non canonique. Cette approche nous a semblé adaptée pour apporter des solutions aux problèmes rencontrés dans le suivi des MPT de protéines modèles : i) cibler précisément un acide aminé à modifier sur une position souhaitée, ii) distinguer plusieurs isoformes de MPT par l'utilisation de différentes protéines de fusion fluorescentes, iii) marquer les protéines avec des MPT stables et non labiles, et iv) travailler en cellule vivante et pas uniquement en cellule fixée augmentant la souplesse de manipulation et la pertinence des observations, notamment pour l'étude des réponses à différents types de stimuli.

La force majeure de cette technologie appelée suppression de codon *ambre* réside dans sa capacité à incorporer en un site spécifique l'acide aminé non canonique dans une protéine à l'aide d'une paire orthogonale aaRS/ARNt exogène réaffectant au codon non-sens TAG (Perdiguero and Liang, 2024), une capacité de codage spécifique à cet acide aminé non canonique.

Le terme « acide aminé non canonique » fait référence à tout acide aminé, modifié ou analogue, qui n'est ni l'un des vingt acides aminés naturels protéinogènes, ni la sélénocystéine ni la pyrrolysine, et qui n'est pas pris en charge par la machinerie de traduction naturelle d'*E. coli*. Un système triptyque est ainsi impliqué : i) l'ARNt suppresseur orthogonal qui ne reconnaît que le codon STOP *ambre* (UAG) alors qu'il n'interfère pas avec le répertoire aaRS endogène ; ii) l'aaRS évoluée orthogonale, qui reconnaît à la fois l'ARNt suppresseur et l'acide aminé non canonique sans interférer

avec le répertoire d'ARNt natifs d'*E. coli* et d'autres acides aminés ; iii) la protéine ciblée ayant un codon *ambre* à la position d'intérêt.

La production *in vivo* (par exemple chez *E. coli*) de la protéine nécessite un apport de l'acide aminé non canonique souhaité dans le milieu de culture. L'ARNt suppresseur d'*ambre* est spécifiquement chargé de l'acide aminé non canonique par son aaRS évoluée orthogonale correspondante. Lorsque le ribosome atteint le codon *ambre*, au lieu de l'arrêt de la traduction, l'ARNt chargé de l'acide aminé non canonique permet son incorporation dans la protéine et la traduction se poursuit jusqu'à la synthèse complète de la protéine. Par conséquent, la protéine recombinante de pleine longueur produite ne contient qu'un seul acide aminé à la position souhaitée. La phénylsélénocystéine (PhSeCys) a été sélectionnée comme l'acide aminé non canonique le plus pratique et approprié en raison de sa capacité à générer de manière stable la S-GlcNAc. La PhSeCys peut être convertie en déshydro-alanine (Dha) par oxydation douce et chimio-sélective avec du peroxyde d'hydrogène (Wang et al., 2007) **(Figure 15)**. Par réaction de chimie click de type addition de Michael de thiol (Nair et al., 2014), la Dha sera par la suite modifiée avec une molécule de (SH)-GlcNAc conduisant à la formation de N-acétyl-D-glucosaminyl-cystéine (De Leon et al., 2017) **(Figure 15)**.

Le couple aaRS/ARNt pour l'incorporation de PhSeCys était déjà disponible à mon arrivée en thèse (collaboration avec les Drs. Sébastien Nouaille et Régis Fauré, Toulouse Biotechnology Institute de Toulouse). Sur la base du niveau d'incorporation de PhSeCys dans la β -caténine-fusionnée-XFP (avec EGFP, mCherry ou BFP pour la partie fluorescente) **(Figure 17)**, la construction plasmidique a été conçue afin d'améliorer davantage l'efficacité d'incorporation. L'utilisation combinée de PhSeCys et la réaction de Michael conduira à des isoformes de β -caténine S-GlcNAcylée avec un minimum de modification structurelle. La question du contrôle de la chiralité reste cependant discutable à ce niveau.

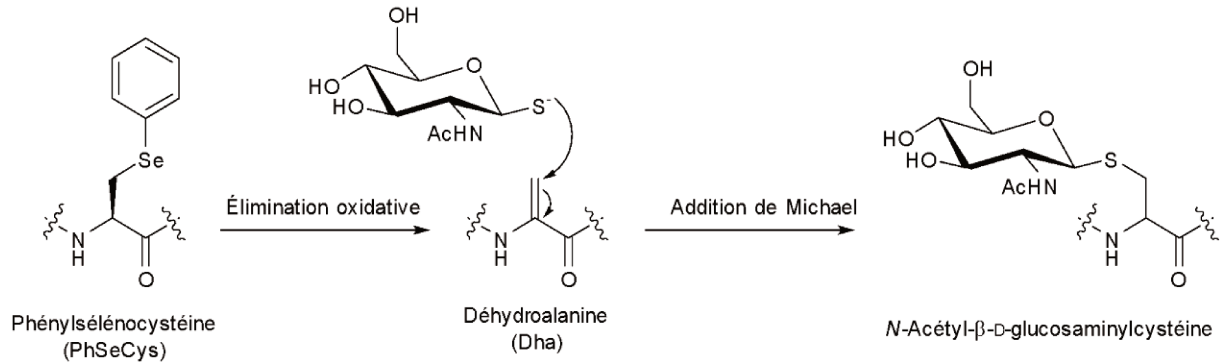


Figure 15. Formation du S-GlcNAc. L'incorporation de la phénylsélocystéine (PhSeCys) est suivie par une réaction d'oxydation et par une addition de Michael du groupement thiol de l'analogue correspondant entraînant la formation de la N-Acetyl-β-D-glucosaminylcystéine.

Par conséquent, nous visons à induire une S-GlcNAcylation constitutive et stable de la β-caténine au niveau du site T41, assurant ainsi la stabilité de la modification GlcNAc sur les formes fluorescentes de la protéine. De plus, pour étudier l'impact de la phosphorylation au niveau de T41, nous introduirons un analogue phosphomimétique non hydrolysable dans notre système expérimental (Rogerson et al., 2015) (**Figures 16 et 17**). Cela nous permettra d'analyser de manière spécifique et précise les effets de ces MPT sur la fonction et le comportement de la β-caténine (Rogerson et al., 2015). Chaque isoforme de β-caténine sera produite sous la forme d'une protéine de fusion avec ou sans TAT (Transactivateur de transcription) et XFP dans des cellules *E. coli* BL21 (**Figure 17**) et purifiée sur colonne IMAC (Chromatographie d'Affinité sur Métal Immobilisé) (Ni^{2+}) grâce à une étiquette 6xHis-Tag. Les protéines ainsi purifiées seront ajoutées au milieu de culture cellulaire et une fois internalisées, visualisées en cellules vivantes grâce à leur fluorophore respectif. En raison de la position stratégique de la β-caténine sur la face interne de la membrane plasmique lui permettant de participer au contrôle des interactions homotypiques E-cadhérine/E-cadhérine responsables de l'intégrité des muqueuses, et dans le noyau où, en partenariat avec le TCF (*T-cell factor*), elle régule l'expression de nombreux gènes cibles impliqués dans le cycle cellulaire, nous pourrions modifier l'environnement cellulaire en inhibant ou en stimulant les cellules avec divers composés chimiques. Aussi, le MG132, inhibiteur du protéasome, devrait nous permettre d'observer une accumulation progressive

d'analogues de la β -caténine dans le noyau et le cytoplasme respectivement. L'utilisation de tels composés permettra de détecter les isoformes gT41 et pT41 de la β -caténine en des emplacements subcellulaires spécifiques et de suivre leur déplacement et leur accumulation cellulaire en fonction de leur statut de modification. En plus de ces composés chimiques, des nutriments tels que les acides gras et les acides aminés pourraient être testés dans notre processus technologique (**Figure 17**).

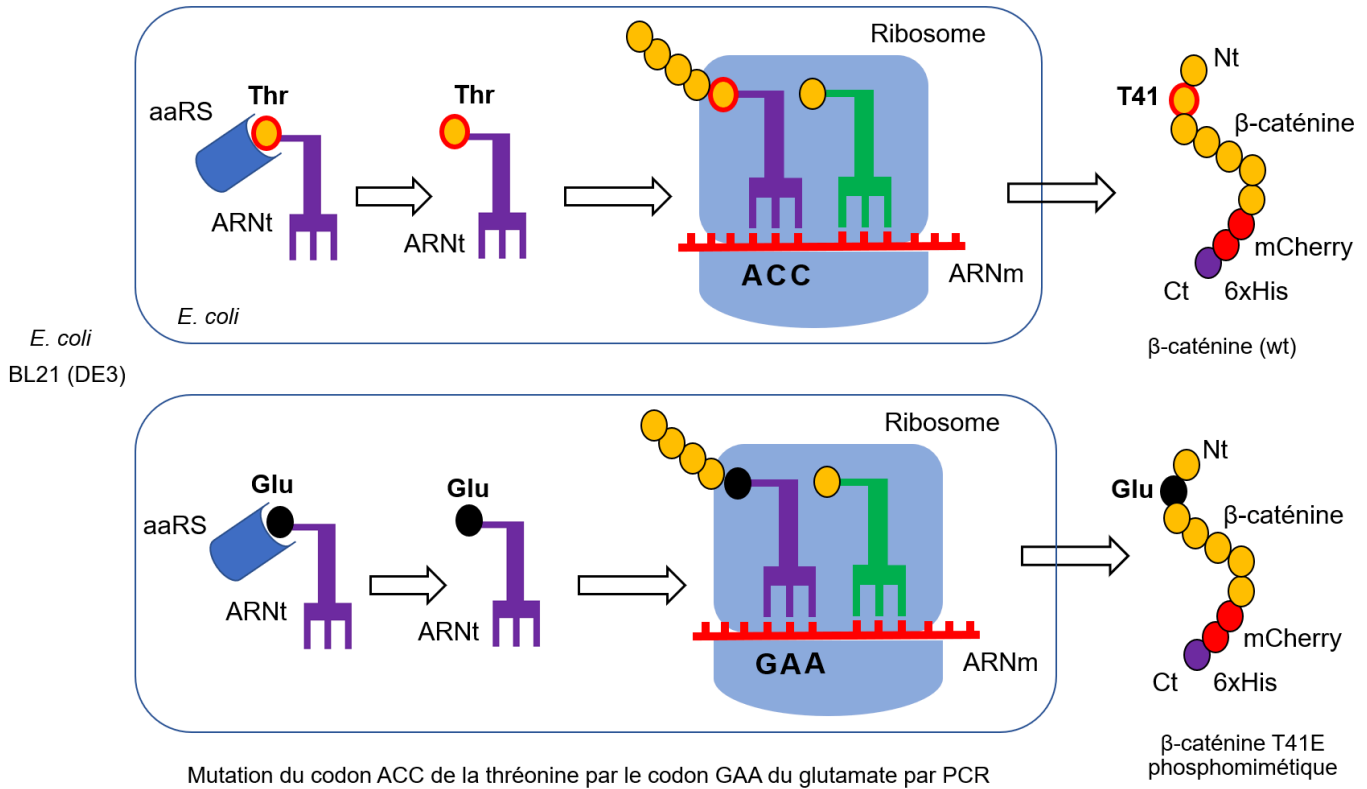


Figure 16. Production de la forme phosphomimétique en T41 de la β -caténine (T41E).

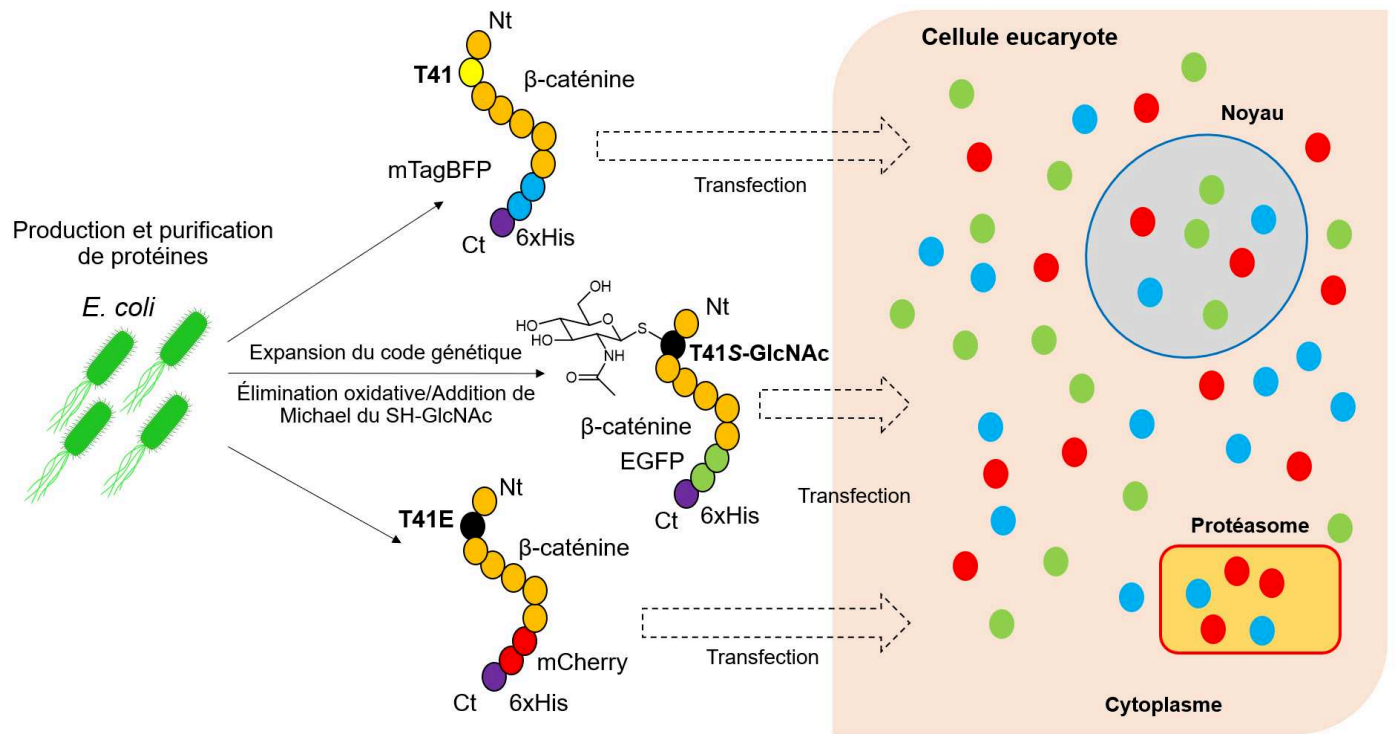


Figure 17. Stratégie et méthodologie suivies au cours du projet de thèse. Les différentes constructions codant la β -caténine-GFP/mCherry/mTagBFP en fusion avec une étiquette 6xHis en C-terminal seront produites chez *E. coli* puis transfectés dans des cellules HeLa, Hep3B et CCD841CoN. Chaque forme de protéine présentera une fluorescence spécifique et sera suivie dans les cellules en réponse à divers produits chimiques. La β -caténine sera S-GlcNAcylée en T41 de manière constitutive et stable. Concernant la phosphorylation en T41, nous muterons la thréonine en acide glutamique, phosphomimétique (Wang et al., 2018). Les protéines β -caténine-GFP/mCherry/mtagBFP seront internalisées grâce à leur fusion avec le tag d'internalisation TAT-HA en N-terminal par transduction cellulaire.

Chapitre II : Matériels et Méthodes

Dans cette partie, nous allons détailler les matériels et méthodes utilisés dans le contexte du projet de ma thèse. Notre article intitulé : **A Step-by-Step Guide for the Production of Recombinant Fluorescent TAT-HA-tagged proteins and their Transduction into Mammalian Cells**. <https://doi.org/10.1002/cpz1.1016> (Anny et al., 2024) résume une partie de ses matériels et méthodes. (Voir la partie “annexes”). Des matériels et méthodes supplémentaires sont également présentés dans cette partie.

1. Matériels des expériences supplémentaires

1.1. Incorporation de la Phényl-L-sélénocystéine à la place du codon *ambre* TAG et S-GlcNAcylation

1.1.1. Incorporation de l'acide aminé : Production de protéines sélénées

Dans le cadre de la production des protéines sélénées, nous avons utilisé le milieu LB (10 g/L de tryptone, 10 g/L de NaCl et 5 g/L d'extrait de levure) supplémentés avec les antibiotiques appropriés, ampicilline à 100 mg/mL (solution mère) diluée à 1/1000 et chloramphénicol à 10 mg/mL (solution mère) diluée à 1/1000. Nous avons utilisé les bactéries *E. coli* BL21-DE3 co-transformées avec le plasmide d'intérêt et le plasmide de l'aaRS de la Se-Phényl-L-sélénocystéine (PhSeCys-aaRS). Nous avons également utilisé l'acide aminé non canonique Se-Phényl-L-sélénocystéine, CAS : 71128-82-0, Cat. No. BD629448 (BLD Pharmatech GmbH, Kaiserslautern, Allemagne). Pour ajuster la neutralité du pH du milieu de culture LB pour la réalisation de l'incorporation de la phénylsélénocystéine, nous avons utilisé : l'acide chlorhydrique 37% (m/v) (HCl 37%) – CAS no 7647-01-0 (Carlo Erba, France) et l'hydroxyde de sodium (NaOH), CAS no 1310-58-3 (Acros organics, Belgique).

1.1.2. Élimination oxydative

Pour réaliser l'élimination oxydative des protéines, nous avons utilisé le surnageant protéique issu de la lyse totale des bactéries, une solution de peroxyde d'hydrogène H₂O₂ 30% (m/v) dans l'eau, référence H1009, CAS-No : 7722-84-1 (Sigmaaldrich, Espagne) et une trizma base, CAS-NO. 77-86-1 (Sigma Life science-France)

1.1.3. Addition de Michael

Pour assurer l'ajout du thiosucre *via* l'addition de Michael, nous avons utilisé les protéines d'intérêt, le 2-acétamido-2-déoxy-b-1-thio- β -D-glucopyranose (HS-GlcNAc) – CAS 781581-10-0, Référence MA44578 (Biosynth Carbosynth, Slovaquie), le Tris(2-carboxyéthyl) phosphine hydrochloride (TCEP) Powder C4706-10G CAS : 51805-45-9 (Aldrich chemistry, France) et Trizma base, CAS-NO. 77-86-1 (Sigma Life Science, France).

1.1.4. Purification par dialyse

Pour assurer la purification des protéines par dialyse, nous avons utilisé le DPBS sans magnésium et sans calcium (Dutsher, cat. no. L0615-500), une membrane de dialyse, le dispositif de dialyse slide-A-Lyzer CAS ; 10 000 MWCO, 0,5-3 mL (Life technologies, cat. no. 66380), un bécher de 1 L, un barreau aimanté, et un agitateur magnétique. Pour assurer le maintien de la stabilité des protéines lors de la purification par dialyse, nous avons placé le bécher à 4°C.

1.2. Transduction et transfection cellulaire

Dans le cadre de la culture cellulaire, nous avons utilisé les lignées suivantes : Hep3B (ATCC, cat. no. HB-8064), HeLa et CCD841CoN. Nous avons également utilisé des flasques T25 et T75 (Corning, cat. no. 430720U), des plaques de 24 puits (Corning, cat. no. 3526) et des lamelles de verre (coverslips)

Dans la salle de culture cellulaire (niveau L2), nous avons utilisé des milieux de cultures adaptés à chaque lignée cellulaire, tels que le Milieu Essentiel Minimum (MEM) de Gibco (Dutsher, cat. no. L0430-50) pour les cellules Hep3B et CCD841CoN, et le Milieu d'Eagle Modifié de Dulbecco (DMEM) à haute teneur en glucose (4.5 g/L), supplémenté par la L-Glutamine et le pyruvate de sodium (Dutsher, cat. no. Réf. L0104-500) pour les cellules HeLa. De plus, pour assurer une survie optimale des cellules, nous avons ajouté du sérum de veau fœtal (Mediatech, Inc. filiale de Corning, cat. no. 35-079-CV), 2 mM de L-glutamine (Gibco par Life Technologies, cat. no. 25030-024), 1 mM de pyruvate de sodium (Gibco par Life Technologies, cat. no. 11360070), 50 mg/L d'antibiotiques (streptomycine, pénicilline) (Gibco par Life Technologies, cat. no. 15070-063). Pour faire le passage des cellules ainsi que leur rinçage, nous avons utilisé le

DPBS sans magnésium et sans calcium (Dutsher, cat. no. L0615-500) et la Trypsine-EDTA 1x dans du PBS sans calcium, sans magnésium, sans rouge de phénol (Dutsher, cat. No. L0940-100). Pour réaliser la transfection ou la transduction des plasmides ou des protéines, nous avons utilisé la Lipofectamine 3000 (Invitrogen, par Thermo Fisher Scientific, cat. no. L3000-015) avec un milieu réduit Opti-MEM (1X) (Gibco par Life Technologies, cat. no. 11058-021) pour les protéines. Nous avons également utilisé le JetOptimus pour les plasmides. Le comptage cellulaire a été réalisé en utilisant une solution de bleu Trypan (Sigma-Aldrich, cat. no. T8154) et une lame de Malassez. La fixation des cellules a été assurée en utilisant le paraformaldéhyde (PAF) à 4% (m/v) (Sigma-Aldrich, cat. no. P6148). Nous avons coloré les noyaux des cellules en bleu en utilisant le DAPI (Dihydrochlorure de 4',6-Diamidino-2-phénylindole) (Sigma-Aldrich, cat. no. D9542). Le montage des lamelles sur les lames a été réalisé en utilisant le milieu de montage Mowiol. Nous avons aussi utilisé du papier Whatman et du parafilm pour confectionner la chambre humide. L'observation des cellules a été assurée par les microscopes optique Leica DM IRB et confocal Zeiss LSM 780. Le traitement des images collectées a été assuré par le logiciel Fiji ImageJ .

1.3. Immunofluorescence

Nous avons réalisé plusieurs expériences en utilisant la technique d'immunofluorescence indirecte. Dans ce contexte, nous avons utilisé plusieurs matériels tels que le PAF à 4% (m/v), de la glycine à 100 mM dans de l'eau ultrapure, du DPBS1x sans magnésium ni calcium un tampon de perméabilisation (0,5% (v/v) de Triton X-100 dans le DPBS), un tampon de blocage, les anticorps primaires anti-flag M2 (souris), anti-Rab5 (souris) et anti-EEA1 (lapin) les anticorps secondaires AlexaFluo488 et AlexaFluo647, le DAPI, et le Mowiol pour le montage des lames.

2. Méthodes des manips supplémentaires

2.1. S-GlcNAcylation de la β -caténine en T41

La conversion de la chaîne latérale de PhSeCys en déhydroalanine (Dha) a été réalisée par un processus d'oxydation chimiosélective et doux utilisant du peroxyde d'hydrogène. Ensuite, pour convertir la Dha en *N*-acétyl- β -D-glucosaminylcystéine, les protéines β -caténine-fusionnées-EGFP contenant de la Dha ont été soumises à une réaction avec du 2-acétamido-2-désoxy-1-thio- β -D-glucopyranose (HS-GlcNAc). Pour la préparation de l'isoforme S-GlcNAcylée de la β -caténine-EGFP à la position T41. La protéine purifiée a été traitée avec 100 mM de peroxyde d'hydrogène dans un tampon Tris 50 mM (pH 8,5) pendant 30 minutes, tout en étant maintenue dans un bain de glace/eau. Ensuite, la protéine a été incubée avec 100 mM de 2-acétamido-2-désoxy-1-thio- β -D-glucopyranose (HS-GlcNAc) dans un tampon Tris 50 mM (pH 8,5), avec l'ajout de 50 mM de TCEP, tout en restant à 4°C et sous agitation pendant une durée de 2 heures. Par la suite, l'échantillon de protéine a été purifié par dialyse toute la nuit à 4°C contre 1 L de DPBS 1X sans magnésium et calcium pour éliminer le peroxyde d'hydrogène résiduel, l'acide phénylsélénique, le HS-GlcNAc non réagi et le TCEP. La confirmation de la S-GlcNAcylation a été réalisée par analyse Western blot en utilisant des anticorps dirigés contre le tag HA et la O-GlcNAc.

2.2. Stratégie de double transfection de plasmides et de protéines recombinantes en cellules eucaryotes par lipofection

Tout d'abord, nous avons cultivé sur lamelles de verre les lignées cellulaires (HeLa, Hep3B et CCD481coN) avec le milieu de croissance approprié (DMEM ou MEM) jusqu'à ce qu'elles atteignent la confluence souhaitée (typiquement 70-80 % pour les cellules adhérentes). Ensuite, nous avons effectué la transfection cellulaire pendant 72 heures du plasmide 2XFlag- β -caténine (wt) en utilisant le JetOptimus. La lipofection de 2 μ M d'isoforme β -caténine-EGFPa été réalisée en utilisant la lipofectamine 3000 directement sur les cellules.

Pour plus de détails, nous avons préparé le complexe JetOptimus-DNA en mélangeant délicatement 1 μ g d'ADN"2xFlag- β -caténine (wt)" avec 1 μ l de réactif JetOptimus dans 100 μ l d'Opti-MEM selon les recommandations du fabricant. Puis, nous avons incubé le

mélange pendant 20 minutes à température ambiante pour former les complexes de transfection. Ensuite, le milieu de culture a été retiré et le mélange de transfection a été ajouté goutte à goutte sur le tapis de cellules adhérentes. Les cellules ont été incubées avec le mélange de transfection pendant la durée recommandée de 6 heures à 37°C dans un incubateur à CO₂. Après cette période, nous avons remplacé le milieu de transfection par un milieu de culture approprié pour permettre aux cellules de récupérer pendant 72h.

Une fois la transfection du plasmide achevée, les protéines ont été introduites et laissées en contact des cellules pendant une durée de 2 heures.

Concernant la lipofection des protéines, nous avons tout d'abord incubé 2 µM de chaque isoforme de β-caténine-XFP (wt, T41E et T41S-GlcNAcylée) avec dans 100 µL d'OptiMEM (Gibco) avec ou sans 2,5 µL de réactif de transfection lipofectamine3000 et 4 µL de P3000 pendant 30 minutes à température ambiante. Puis les cellules ont été incubées à 37 °C et sous atmosphère de 5% (v/v) de CO₂ pendant 2 heures. Après cette période, le milieu a été aspiré, et les cellules rincées trois fois pour éliminer les traces de milieu. Elles ont été ensuite fixées avec le PAF à 4% (m/v) pendant 30 minutes à température ambiante. Puis, trois lavages de 7 minutes chacun avec du DBPS 1x ont été réalisés. Les lamelles sur lesquelles sont fixées les cellules ont alors été transférées dans la chambre humide pour réaliser la coloration des noyaux avec une solution de DAPI pendant 10 minutes, suivie de trois lavages avec du DPBS 1x. Enfin, les lamelles ont été montées sur lame en utilisant une solution de Mowiol (5 µL par lamelle): les lames ont été laissées à l'obscurité à température ambiante toute la nuit. La visualisation des lames a été réalisée le lendemain par microscopie confocale.

2.3. Immunofluorescence

Les cellules ont été cultivées sur des lamelles de verre dans des plaques de culture de 24 puits. Pour démarrer le processus, les cellules ont été rincées délicatement trois fois avec du DPBS1X froid. Ensuite, les cellules ont été fixées en utilisant le PAF à 4% (m/v) pendant 30 minutes à température ambiante avant un lavage au DPBS 1x. L'excès de PAF a été éliminé par une incubation pendant 20 minutes à température ambiante avec une solution de neutralisation (100 mM de glycine), suivie d'un rinçage au PBS trois fois pendant 5 minutes chacun. Les cellules ont été perméabilisées avec dans le tampon de perméabilisation (0,5% (v/v) de Triton X-100 dans du DPBS) pendant 20 minutes. Les sites non spécifiques ont été bloqués avec une solution de blocage IF pendant une heure à température ambiante et à l'obscurité. Les cellules ont ensuite été incubées avec le tampon de blocage IF contenant l'anticorps primaire (dilué au 1:100) pendant la nuit. Après incubation, trois lavages avec du DPBS 1x ont été effectués. Enfin, les lamelles de verre ont été incubées avec des anticorps secondaires conjugués au Texas Red, FITC ou Alexa, dilués au 1:100 dans la solution de blocage IF, pendant une heure à température ambiante à l'abri de la lumière. Pour terminer, les noyaux ont été colorés avec une solution de DAPI pendant 10 minutes suivi de trois lavages avec du DPBS 1x. Les lamelles ont été finalement montées sur lame à l'aide d'une solution de Mowiol.

Chapitre III : Résultats

Une étude publiée en 2014 par l'équipe (Olivier-Van Stichelen et al., 2014) visait à identifier les sites de *O*-GlcNAcylation de la β -caténine. Ainsi quatre sites ont été cartographiés : S23, T40, T41 et T112. Parmi ces sites, T41, au niveau de la boîte de destruction, a plus particulièrement attiré l'attention puisqu'il avait été décrit comme phosphorylé dans la littérature et comme un "hotspot" de mutation dans certains types de cancers épithéliaux (Fukuchi et al., 1998; Takahashi et al., 2000; Liu et al., 2002). Ainsi, Olivier-Van Stichelen et collaborateurs ont montré que la *O*-GlcNAcylation du résidu T41 réduisait sa phosphorylation et était par conséquent impliquée dans la stabilité de la β -caténine *via* une diminution de sa sensibilité au protéasome. Une réduction de la dégradation de la β -caténine par la *O*-GlcNAcylation dépendante des nutriments a alors été proposée. Les résultats de cette même étude ont montré que la *O*-GlcNAcylation régulait non seulement la stabilité de la β -caténine, mais suggérait également qu'elle puisse affecter sa localisation au niveau des jonctions d'adhérence. Le résidu T41 est donc crucial pour l'expression de la β -caténine (Olivier-Van Stichelen et al., 2014).

Il n'existait pas à proprement parler de technologie permettant de suivre une même protéine sous diverses formes de MPT dans une même cellule. De fait, nous avons mis en place une série d'expériences à partir de l'année 2020 pour remédier à ce manque. Au vu des éléments apportés ci-dessus, notre choix s'est tout naturellement porté sur la réciprocity *O*-GlcNAcylation/phosphorylation du résidu T41 de la β -caténine. Le but a été de produire en bactéries *E. coli* BL21 (DE3), trois isoformes de la β -caténine : une forme sauvage (wt) fusionnée à un fluorophore m-TagBFP bleu (λ_{ex} : 405 nm, λ_{em} : 410-470 nm), une forme phosphomimétique fusionnée à un fluorophore mCherry rouge (λ_{ex} : 561 nm, λ_{em} : 570-615 nm) pour laquelle la thréonine 41 est remplacée par un acide glutamique (T41E), et une forme *S*-GlcNAcyliée fusionnée à un fluorophore EGFP en vert (λ_{ex} : 488nm, λ_{em} : 500-543nm) pour laquelle nous désirions introduire un résidu de *S*-GlcNAc en position 41.

Dans un premier temps, nous avons choisi de générer des protéines recombinantes possédant une séquence TAT-HA au niveau de l'extrémité *N*-terminale et une séquence 6xHis tag au niveau de l'extrémité *C*-terminale. En effet, en se basant sur des études antérieures, l'utilisation du peptide TAT, une séquence peptidique courte provenant du virus de l'immunodéficience humaine (VIH), semblait s'avérer utile pour faciliter la transduction cellulaire de notre protéine d'intérêt, permettant ainsi de la faire pénétrer efficacement au travers des membranes cellulaires (Becker-Hapak et al., 2001; Fittipaldi and Giacca, 2005). Aussi, le tag HA est pratique pour marquer et suivre ces protéines d'intérêt (Zhao et al., 2019). Ainsi, l'expression de ces protéines peut être évaluée en ciblant le tag HA par Western blot. Par ailleurs, le choix du tag 6x-His a été fait pour purifier ces protéines par chromatographie d'affinité sur métal immobilisé (IMAC) (Magnusdottir et al., 2009).

1. Clonage et assemblage moléculaire des constructions TAT-HA

Produire des molécules d'ADN recombinantes à façon par clonage moléculaire et les utiliser pour transformer des bactéries *E. coli* DH5 α , représente la première étape du projet. Les fragments d'ADN ont été amplifiés par PCR (*Polymerase Chain Reaction* ou Réaction de Polymérisation en Chaîne), une technique d'amplification enzymatique permettant d'obtenir un grand nombre de copies identiques d'un fragment d'ADN. L'obtention des différents produits de PCR n'a été réalisée qu'après avoir optimisé ses conditions par l'utilisation de l'ADN polymérase Q5, et en modifiant la température d'hybridation et les composants de la PCR.

Daniel G. Gibson et ses collègues ont développé une méthode d'assemblage moléculaire qui s'appelle également assemblage Gibson (*Gibson assembly* en anglais). L'assemblage moléculaire des différents fragments d'ADN a été réalisé en suivant cette stratégie permettant d'assembler plusieurs fragments présentant une homologie de 20 paires de bases en une seule molécule de manière précise et efficace. Elle ne nécessite pas de sites de restriction spécifiques et peut être utilisée pour des fragments d'ADN aux extrémités non complémentaires (Gibson, 2009; Gibson et al., 2009; Bordat et al., 2015)

L'objectif de cette première partie a donc été de construire des molécules d'ADN qui nous serviront pour toute la suite du projet. Les constructions souhaitées, comme préalablement décrit dans la partie "Méthodologie de travail" sont présentées dans la **figure 18** ci-dessous :

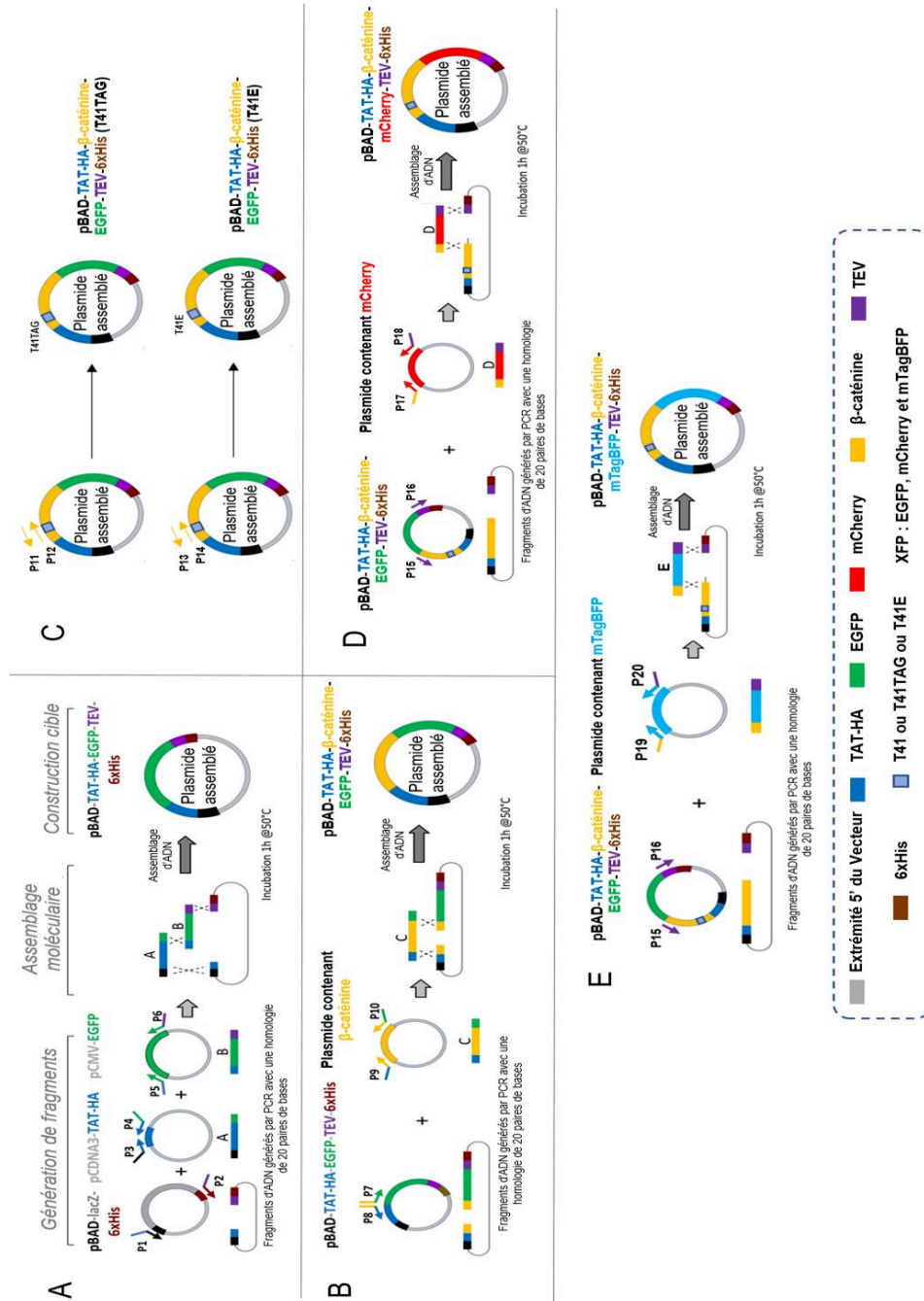


Figure 18. Construction des vecteurs plasmidiques par assemblage des fragments d'ADN. (A) : Utilisation des amorces P1-6 pour générer la construction finale pBAD-TAT-HA-EGFP-TEV-6xHis. (B) : Utilisation des amorces P8-10 pour générer la construction finale pBAD-TAT-HA-β-caténine-EGFP-TEV-6xHis. (C) : Utilisation des amorces P11-P12 et P13-P14 pour remplacer par mutation dirigée la T41 par un codon *ambre stop* TAG (T41TAG) ou un codon GAA codant un acide glutamique (T41E). (D) : Utilisation des amorces P15-P18 pour générer la construction finale pBAD-TAT-HA-β-caténine-mCherry-TEV-6xHis wt, T41TAG et T41E. (E) : Utilisation des amorces P15-P16 et P19-P20 pour générer la construction finale pBAD-TAT-HA-β-caténine-mTagBFP-TEV-6xHis wt, T41TAG et T41E. Les amorces P1-20 représentent les différentes amorces répertoriées dans le Tableau 6 et nécessaires à la génération des différents produits de PCR. Pour plus de détails, voir les Tableaux 4-9. Le code couleur utilisé est le même que celui du Tableau 6.

L'ensemble des composants nécessaires à la PCR, des souches bactériennes, des plasmides, des amorces, et la procédure de PCR sont présentés dans les **tableaux 4-7**. Les résultats de clonage sont présentés dans les **figures 19-21**. La transformation bactérienne a été réalisée en bactérie DH5 α .

Tableau 4. Souches bactériennes et plasmides.

Souche ou plasmide	Caractéristiques	Source
Souche		
<i>E. coli</i> compétentes (haute efficacité)	DH5-alpha	New England Biolabs
<i>E. coli</i> compétentes (haute efficacité)	BL21 (DE3)	New England Biolabs
Plasmide		
pBAD-LacZ-6xHis	Vecteur d'expression pour <i>E. coli</i> , amp ^r	Life science
pcDNA3-Y-TAT-HA	Vecteur d'expression pour <i>E. coli</i> , amp ^r	UGSF Lille (France)
vCMVp-EGFP- β -caténine	Vecteur d'expression pour <i>E. coli</i> , Kan ^r	e-Zyvec, Loos (France)
pBAD-TAT-HA-EGFP-TEV-6xHis	Vecteur d'expression pour <i>E. coli</i> , amp ^r	Cette étude
pBAD-TAT-HA- β -caténine-EGFP-TEV-6xHis	Vecteur d'expression pour <i>E. coli</i> , amp ^r	Cette étude
pCMV-mCherry	Vecteur d'expression pour <i>E. coli</i> , Kan ^r	UGSF, Lille (France)
pCMV-mTagBFP	Vecteur d'expression pour <i>E. coli</i> , Kan ^r	TBI, Toulouse (France)
pUC57-PheSelCys aaRS	Vecteur d'expression pour <i>E. coli</i> , Kan ^r	Life science
pINS-AzF	Vecteur d'expression pour <i>E. coli</i> , amp ^r	TBI, Toulouse (France)

Tableau 5. Amorces utilisées pour générer l'ensemble des constructions (par souci de clarté, le code couleur utilisé pour les séquences est le même que celui utilisé dans la **figure 18**).

Amorces	Séquence (5' → 3')	Construction
P1-TEV-6xHis	GAAAACCTGTA ^{CTTCCAGGGT} AATAGC GCCGTCGACCATCAT	pBAD-TAT-HA-EGFP- TEV-6xHis
P2-TAT-HA-pBAD	AGCCAAGCTTTGGATCCATGGTTAATT CCTCCTGT	
P3-pBAD-TAT-HA	CATGGATCCA ^{AAGCTTGGCTACGGCCG} CAAGAAA	
P4-EGFP-TAT-HA	CCTTGCTCACGGCCATGGAGCCAGCAT AGTCTG	
P5-EGFP-TAT-HA	CTCCATGGCCGTGAGCAAGGGCGAGG AGCTGTTC	
P6-TEV-EGFP	ACCCTGGAAGTACAGGTTTT ^{CTTGTA} CAGCTCGTCCATGCCGAG	
P7-β-caténine-EGFP	TACTGACCTGGT ^{GAGCAAGGGCGAGG} AGCTGTTC	pBAD-TAT-HA-β- caténine-EGFP-TEV- 6xHis
P8-β-caténine-TAT-HA	CTTGAGTAGCGGCCATGGAGCCAGCA TAGTCTG	
P9-TAT-HA-β-caténine	CTCCATGGCCGCTACTCAAGCTGATTT GATGGAGTTG	
P10-EGFP-β-caténine	CCTTGCTCACCAGGTCAGTATCAAACC AGGCCAGC	
P11-β-caténine-T41-TAG Fw	CATTCTGGTGCCACTTAGACAGCTCCT TCT	pBAD-TAT-HA-β- caténine-EGFP-TEV- 6xHis (T41TAG)
P12-β-caténine-T41-TAG Rev	AGAAGGAGCTGTCTAAGTGGCACCAGA ATG	
P-13-β-caténine-T41Glu Fw	CATTCTGGTGCCACTGAAACAGCTCCT TCT	pBAD-TAT-HA-β- caténine-EGFP-TEV- 6xHis (T41E)
P-14-β-caténine-T41Glu Rev	AGAAGGAGCTGTTTCAGTGGCACCAGA ATG	

P15- β -caténine Rev	CAGGTCAGTATCAAACCAGGCCAGC	pBAD-TAT-HA- β -caténine-mcherry-TEV-6xHis (wt)
P16-5'TEV	GAAAACCTGTA CTTCAGGGTAATAGC G	
P17-mcherry	CCTGGTTTGATACTGACCTGGTGAGCA AGGGCGAGGAGGATAACAT	
P18-mcherry	CCCTGGAAGTACAGGTTTTCTTGATC AGCTCGTCCATGCCGCC	
P15- β -caténine Rev	CAGGTCAGTATCAAACCAGGCCAGC	pBAD-TAT-HA- β -caténine-mTagBFP-TEV-6xHis
P16-5'TEV	GAAAACCTGTA CTTCAGGGTAATAGC G	
P19-mTagBFP	CCTGGTTTGATACTGACCTGAGCGAAC TGATCAAAGAGAACATGCACATG	
P20-mTagBFP	CCCTGGAAGTACAGGTTTTCA TTCAGT TTATGACCCAGCTTGCTAGGCA	
P21-pINS-AzF Rev	CGTTTAATCATTTCGAACTCGTCCAT	Plasmide de la synthétase orthogonale aaRS (PhSeCys)
P22-pINS-AzF Fw	GGAACCAATTCGCAAACGTCTGTAA	
P23-pEvolAzfbkb Fw	ATGGACGAGTTCGAAATGATTAACG	
P24-3'aaRS-AzF	TTACAGACGTTTGCGAATTGGTTCC	
P25-3'ATG PBAD	CATGGTTAATTCCTCCTGTTAGC	Toutes les constructions plasmidiques sans TAT-HA
P26-5'- β -caténine Fw	GCTACTCAAGCTGATTTGATGGA	

Tableau 6. Composition et conditions de PCR réalisée par l'ADN polymérase haute-fidélité Q5.

Composition du mélange PCR			
Composant	Solution stock	Volume (μL)	Concentration finale
Tampon de réaction Q5	5x	5 μL	1x
Mélange dNTP	10 mM chacun	0,5 μL	0,2 mM chacun
Amorces (10 μM)	10 μM	1,25 μL pour chaque amorce	0,5 μM chacun
Matrice plasmidique	Variable	Variable	$\sim 4 \text{ pg}/\mu\text{L}$
Q5 polymérase	2 U/ μL	0,3 μL	0,024 U/ μL
Eau ultrapure		25 μL	
Conditions de thermocyclage PCR			
Etape	Température ($^{\circ}\text{C}$)	Temps	Cycle
Dénaturation	98	30 s	1
Dénaturation	98	10 s	x30
Hybridation	60	20 s	
Elongation	72	30 s/kb	
Elongation	72	3 min	1
Refroidissement	14	∞	-

Tableau 7. Caractéristiques des composants de la PCR.

Source d'ADN plasmidique	Paire d'amorces	Taille des amplicons (pb)
pBAD-LacZ-X-6xHis	P1-P2	4047
pCDNA3-Y-TAT-HA	P3-P4	122
vCMV-EGFP- β -caténine	P5-P6	745
pBAD-TAT-HA-EGFP-TEV-6xHis	P7-P8	4873
vCMV-EGFP- β -caténine	P9-P10	2360
Mutation T41 à T41TAG	P11-P12	7193
Mutation T41 à T41E	P13-P14	7193
pBAD-TAT-HA- β -caténine-EGFP-TEV-6xHis	P15-P16	6479
pCMV-mCherry	P17-P18	745
pCMV-mTagBFP	P19-P20	736
pINS-AzF	P21-P22	3066
pUC57-PheSelCys (aaRS)	P23-P24	921
Toutes les constructions plasmidiques avec TAT-HA	P25-P26	7071

Tableau 8. Conditions d'assemblage de l'ADN plasmidique.

Composants	Solution stock	Volume	Concentration finale
PCR purifié	Variable	Variable	5-15 ng/μL
T4 polynucléotide kinase PNK	10 U/μL	1 μL	0.5 U/μL
T4 DNA ligase Buffer with ATP	10x avec 10 mM	1 μL	0.5x avec 0.5 mM
T4 DNA ligase	400 U/μL	1 μL	20 U/μL
Ultrapure water		20 μL	

Tableau 9. Assemblage d'ADN plasmidique par la méthode Gibson.

Composant	Volume (μL)
Fragments PCR purifié	x μL
Enzymes Master Mix 2x HIFI DNA Assemblage MM	5 μL
Eau	10 μL
Conditions de thermocyclage pour l'assemblage de l'ADN	
50°C pendant 1h	
14°C/∞	

En appliquant la procédure de PCR indiquée dans les **tableaux 6-7**, nous avons produit tous les fragments d'ADN plasmidiques présentés dans les **figures 19-21**. Tout d'abord, nous avons produit les fragments pBAD-TEV-6xHis de 4047 pb en utilisant le couple d'oligos P1-P2, β-caténine de 2360 pb en utilisant le couple d'oligos P9-P10, EGFP de 745 pb en utilisant le couple d'oligos P5-P6, et TAT-HA de 122 pb en utilisant le couple d'oligos P3-P4. Puis, en appliquant la procédure selon l'assemblage Gibson

indiquée dans le **tableau 9**, nous avons assemblé les trois fragments pBAD-TEV-6xHis, EGFP et TAT-HA pour générer la construction pBAD-TAT-HA-EGFP-TEV-6xHis de 4873 pb. Nous avons alors transformé des bactéries DH5 α avec ce plasmide par la méthode de choc thermique. Afin de vérifier la séquence du plasmide, nous l'avons extrait et purifié par Mini-prep. Le séquençage moléculaire a confirmé l'obtention d'une construction P_{BAD}-TAT-HA-EGFP-TEV-6xHis (**Figures 19A et 19B**).

La construction pBAD-TAT-HA-EGFP-TEV-6xHis a servi de matrice à une PCR supplémentaire en utilisant le couple d'oligos P7-P8, afin d'ajouter les séquences nucléotidiques codant la β -caténine entre TAT-HA et EGFP pour produire le plasmide pBAD-TAT-HA- β -caténine-EGFP-TEV-6xHis de 7193 pb. Nous avons transformé des bactéries DH5 α avec ce plasmide, puis extrait et purifié ce plasmide par Mini-prep. Le séquençage moléculaire a confirmé l'obtention de la construction pBAD-TAT-HA- β -caténine-EGFP-TEV-6xHis (**Figure 19C**).

En outre, le remplacement du codon T41 de la β -caténine au niveau de la construction pBAD-TAT-HA- β -caténine-EGFP-TEV-His₆ par des codons *ambre* TAG et Glutamate GAA a été réalisé par PCR respectivement en utilisant les couples d'oligos P11-P12 et P13-P14. Afin de circulariser le plasmide, nous avons appliqué les conditions d'assemblage de l'ADN plasmidique mentionnées dans le **tableau 8** en utilisant les enzymes T4 PNK et T4 ADN ligase. Après transformation dans les bactéries DH5 α , le séquençage moléculaire a confirmé cette mutation. Les constructions finales ont été résolues et purifiées par électrophorèse en gel d'agarose: pBAD-TAT-HA- β -caténine-EGFP-TEV-6xHis (T41TAG) et pBAD-TAT-HA- β -caténine-EGFP-TEV-6xHis (T41E) de 7193 pb chacun (**Figure 19D**).

Par la suite, la partie EGFP des constructions a été remplacée par le mCherry ou le mTagBFP de sorte à générer des protéines de fluorescence différente. Pour cela, nous avons produit les fragments pBAD-TAT-HA- β -caténine-TEV-6xHis de 6479 pb en utilisant le couple d'oligos P15-P16, mCherry de 745 pb avec le couple d'oligos P17-P18, et mTagBFP de 736 pb avec le couple d'oligos P19-P20. Les produits de PCR ont été purifiés sur gel avant assemblage Gibson (**Figure 19E**).

La synthèse des plasmides pBAD-TAT-HA- β -caténine-mTagBFP-TEV-6xHis et pBAD-TAT-HA- β -caténine-mCherry-TEV-6xHis pour les trois catégories de plasmides sauvage (wt), muté en T41 (T41TAG) et muté en glutamate GAA (T41E), a été réalisé par PCR puis assemblage Gibson selon les conditions mentionnées dans le **tableau 9**.

Enfin dans le but d'incorporer l'acide aminé non naturel, la phénylsélénocystéine, au niveau du codon *ambre* TAG en T41 de la β -caténine, nous avons produit les fragments pINS de 3066 pb et aaRS (*aminoacyl tRNA synthetase*) de 921 pb par PCR, respectivement en utilisant les couples d'oligos P21-P22 et P23-P24. Après transformation dans des bactéries DH5 α , le séquençage moléculaire a confirmé cet assemblage pour obtenir la construction pINS-aaRS (PhSeCys) de 3987 pb après purification sur gel (**Figure 20**).

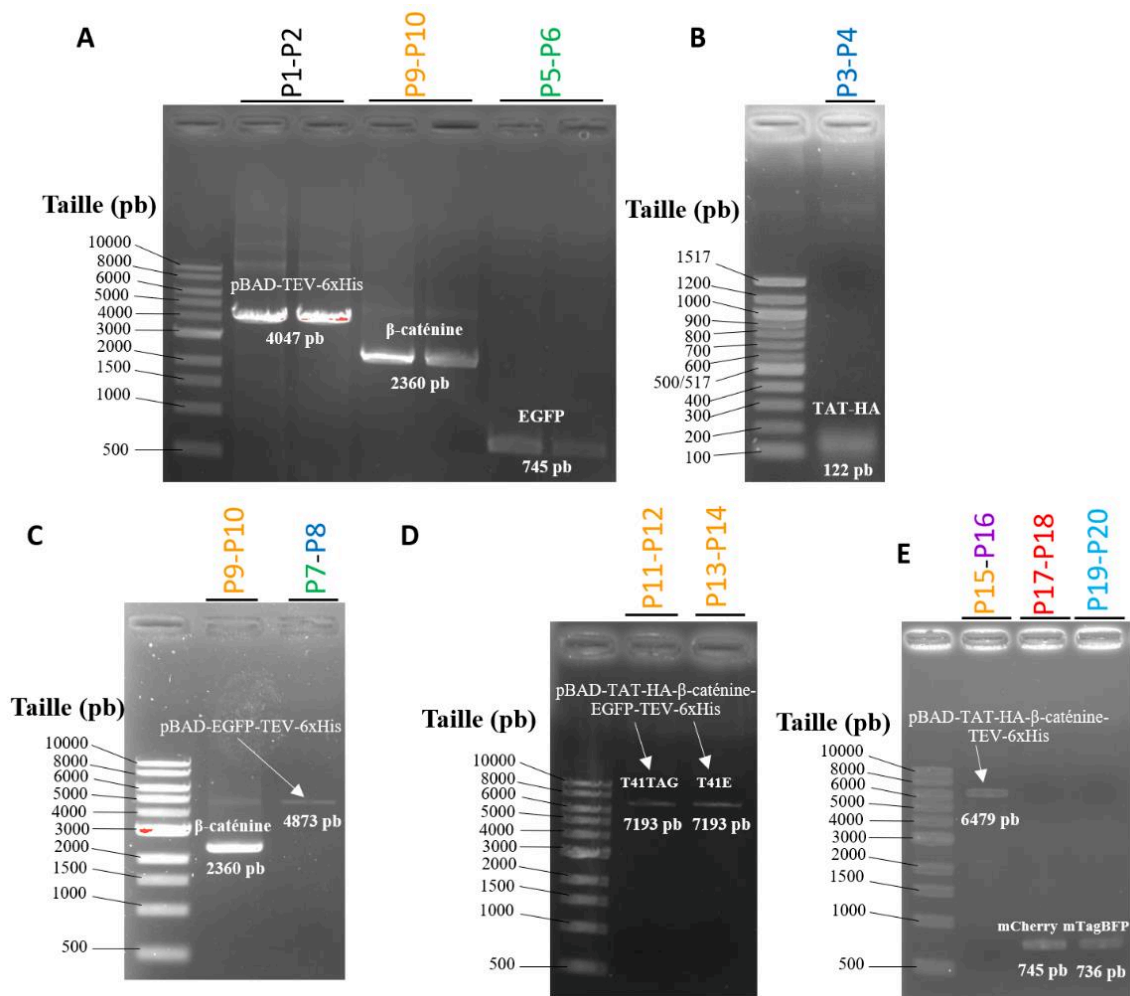


Figure 19. Caractérisation des produits de PCR par électrophorèse en gel d'agarose. Les produits de PCR des différents fragments présentés schématiquement dans la **Figure 18** ont été analysés par électrophorèse en gel d'agarose. En haut de chaque gel sont indiquées les couples d'amorces suivant le même code couleur utilisé dans la **Figure 18**. (A) : Fragments pBAD-TEV-6xHis, β -caténine et EGFP ; (B) : Fragments TAT-HA ; (C) : Fragments β -caténine et pBAD-EGFP-TEV-6xHis ; (D) : Fragments pBAD-TAT-HA- β -caténine-EGFP-TEV-6xHis (T41TAG) et pBAD-TAT-HA- β -caténine-EGFP-TEV-6xHis (T41E) ; (E) : Fragments pBAD-TAT-HA- β -caténine-TEV-6xHis, mCherry et mTagBFP. Les marqueurs de taille sont indiqués à gauche de chaque gel : ladder d'ADN 1 kb Plus pour les panneaux A, C, D et E ; Ladder de 100 pb pour le panneau B.

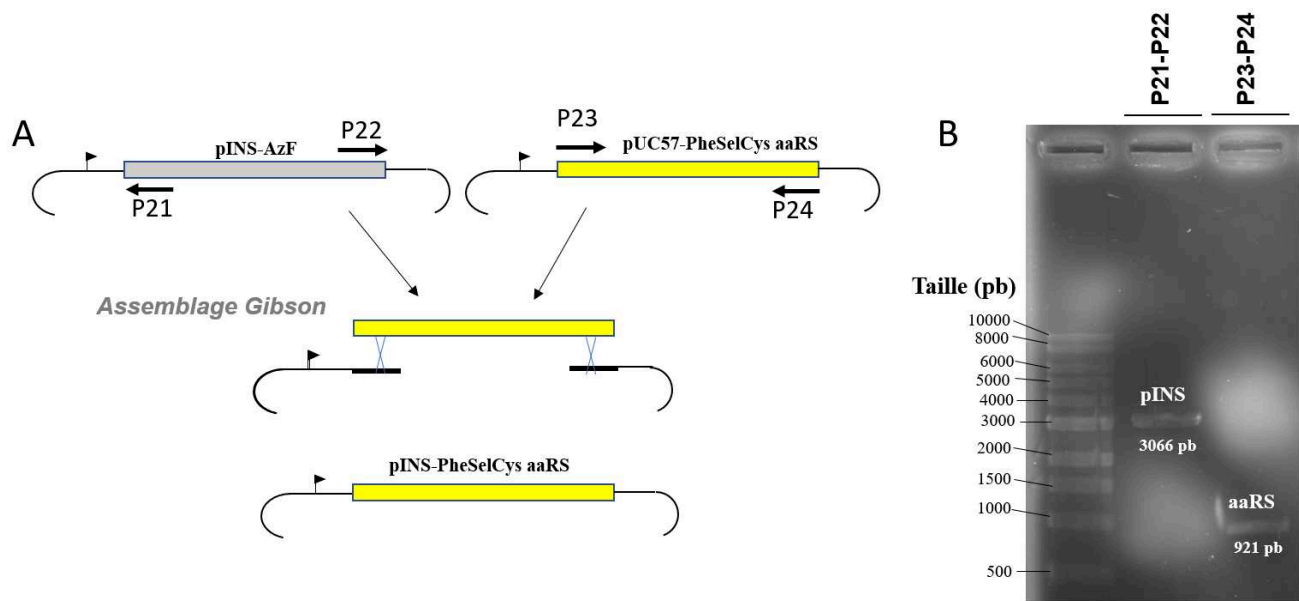


Figure 20. Génération du plasmide codant la synthétase orthogonale aaRS spécifique de la PhSeCys. (A) Stratégie d'assemblage Gibson suivie pour générer le plasmide pINS-PheSelCys aaRS ; (B) : Analyse des fragments de PCR pINS et aaRS par électrophorèse sur gel d'agarose. Les marqueurs de taille sont indiqués à gauche du gel : ladder d'ADN 1 kb Plus.

Après avoir vérifié les séquences correspondant aux deux séries de plasmides sans et avec TAT-HA, nous avons transformé chaque plasmide en bactéries *E. coli* de type BL21-DE3 pour produire les différentes isoformes de la β -caténine-XFP-TEV-6xHis. Nous avons co-transformé les bactéries *E. coli* BL21-DE3 en introduisant le plasmide pBAD-TAT-HA- β -caténine-XFP-TEV-6xHis (T41TAG) avec le plasmide de la synthétase orthogonale T7-pINS-aaRS (PhSeCys), pour incorporer la phénylsélénocystéine au niveau du codon *ambre* TAG pour remplacer la T41 de la β -caténine et produire les isoformes de β -caténine-XFP-TEV-6xHis sélénée en T41.

En résumé, cette partie biologie moléculaire a représenté la base de mon projet de thèse. Nous avons assemblé différents fragments de PCR purifiés pour produire les plasmides pBAD-TAT-HA- β -caténine-XFP-TEV-6xHis de trois catégories : sauvages (wt), mutés par un codon *ambre* T41TAG ou par un codon GAA (T41E). Chaque catégorie de plasmides est représentée en trois exemplaires, chaque plasmide étant fusionné à un fluorophore (XFP) : EGFP, mCherry, ou mTagBFP (**Figure 18A-E**). Pour ce faire, nous avons optimisé le choix des couples d'oligonucléotides utilisés dans chaque PCR, ainsi que le choix des souches bactériennes pour la production des

plasmides et des protéines. Ceci nous a permis de produire la série de plasmides pBAD-TAT-HA- β -caténine-XFP-TEV-6xHis pour les trois formes : wt, T41TAG et T41E (Figure 21).

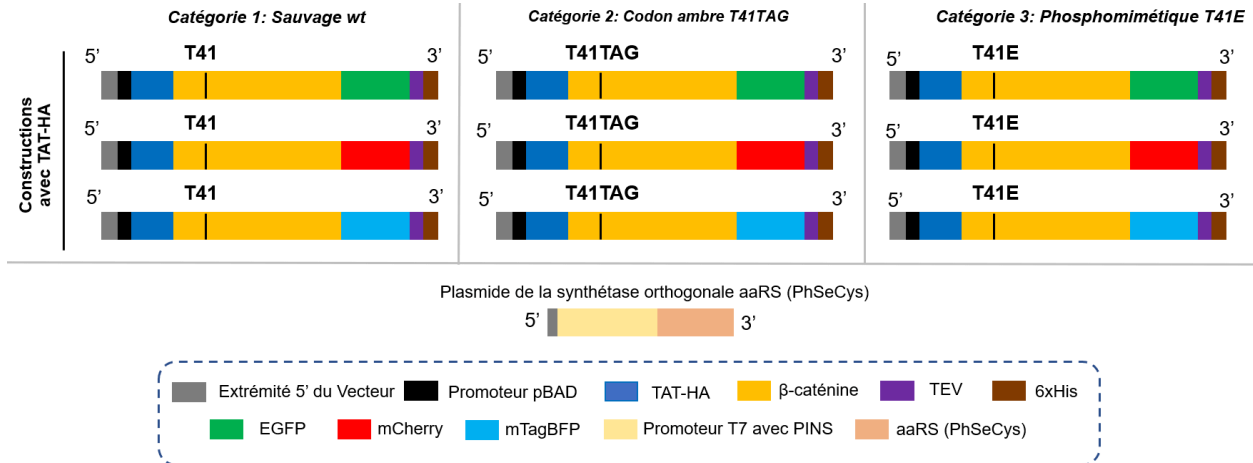


Figure 21. Résumé graphique des différentes constructions plasmidiques avec TAT-HA utilisées dans ce projet. La taille des constructions avec TAT-HA et du plasmide de la synthétase orthogonale aaRS sont respectivement de 7193 pb et 3987 pb.

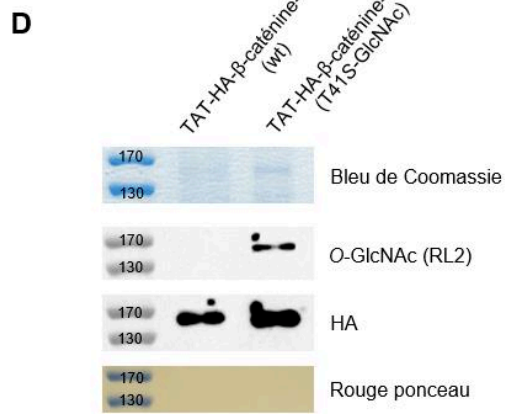
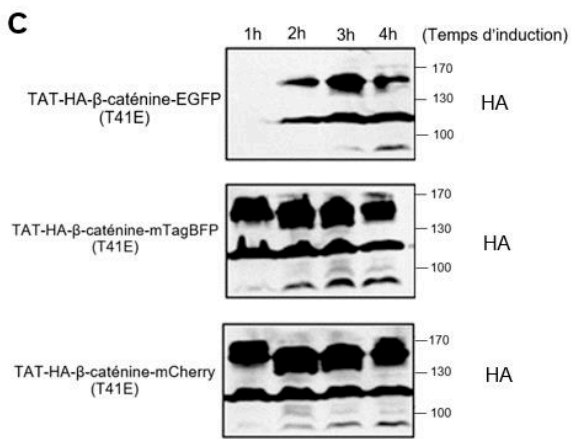
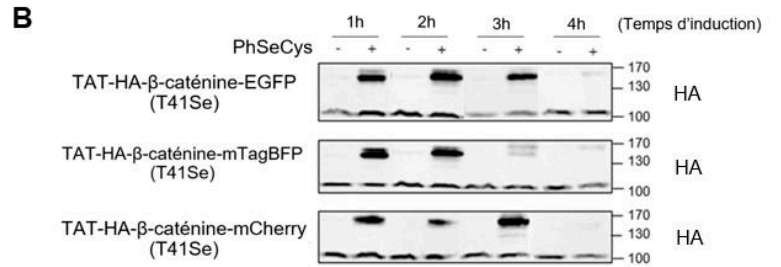
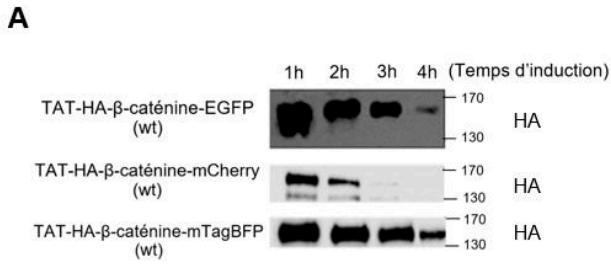
2. Traçage des isoformes de la β -caténine par transduction cellulaire

La production des différentes isoformes TAT-HA- β -caténine-XFP a été évaluée par WB anti-HA. Les **figures 22A, 22B et 22C** montrent respectivement l'expression des formes TAT-HA- β -caténine-XFP sauvage, sélénée et phosphomimétique après quatre heures d'induction avec 0,1% (m/v) de L-arabinose. Nous constatons que la production optimale de ces isoformes correspond au temps d'induction d'une heure par le L-arabinose (à 37°C sous agitation à 150 rpm).

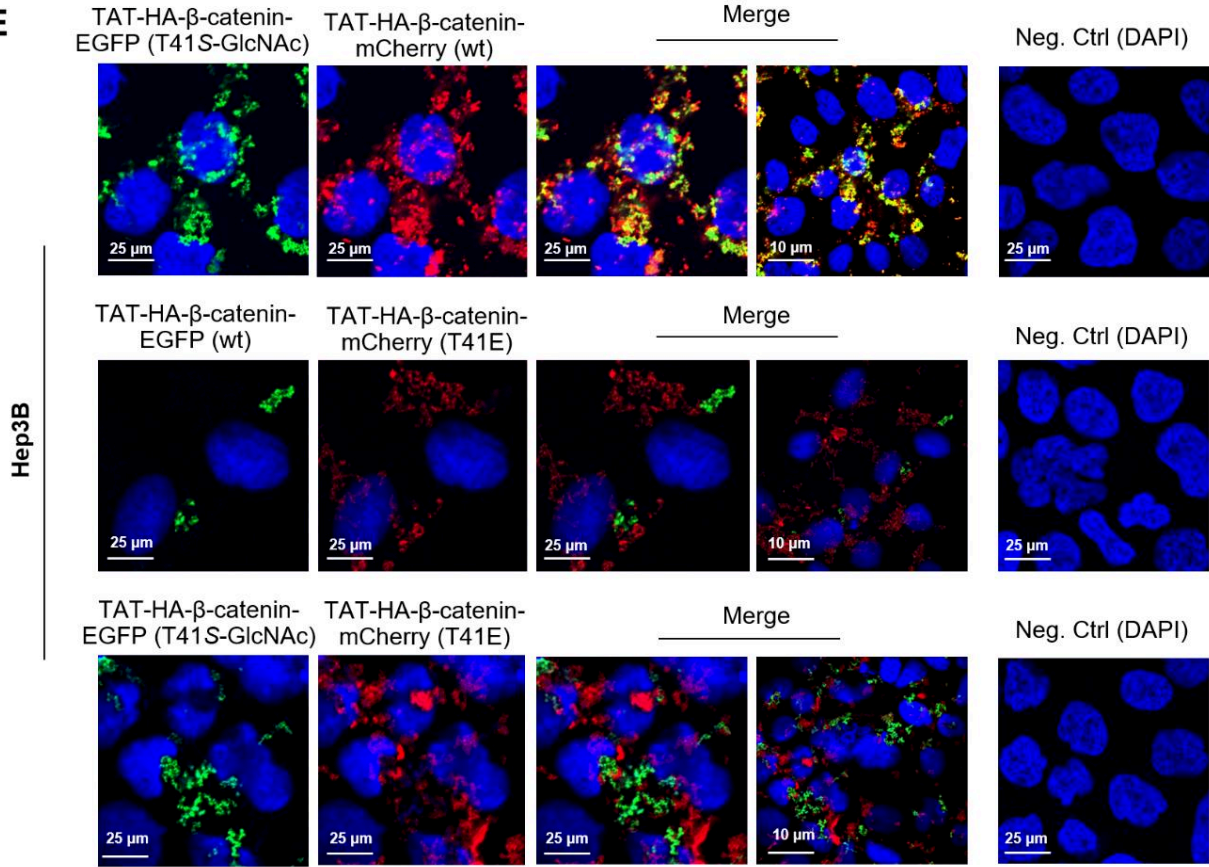
Dans le but de vérifier la *S*-GlcNAcylation de la forme TAT-HA- β -caténine-EGFP, suite aux réactions d'oxydation et d'addition de Michael, une purification des formes sauvage et *S*-GlcNAcylée de cette protéine a été réalisée. Le résultat de cette expérience est présenté en **figure 22D**: il s'agit des WB anti-*O*-GlcNAc et anti-HA, et d'une coloration au bleu de Coomassie des protéines purifiées et résolues par SDS-PAGE. Nous constatons que le WB anti-HA et la coloration au bleu de Coomassie attestent de la production des formes de β -caténine bien que la révélation au bleu de Coomassie soit un peu faible. Le WB anti-*O*-GlcNAc montre quant à lui la *S*-GlcNAcylation de la protéine : absence de bande correspondant à la forme sauvage, et présence d'une bande correspondant à la forme *S*-GlcNAcylée.

Une série d'expériences de transduction cellulaire des trois isoformes de la β -caténine a alors été réalisée dans la lignée cancéreuse de foie Hep3B (**Figure 22E**), et la lignée fœtale de côlon CCD841CoN (**Figure 22F**). Les différentes isoformes ont été transduites deux à deux pendant deux heures à une concentration de 2 μ M chacune: TAT-HA- β -caténine-EGFP (T41 *S*-GlcNAc) vs TAT-HA- β -caténine-mCherry (wt), TAT-HA- β -caténine-EGFP (wt) vs TAT-HA- β -caténine-mCherry (T41E), et TAT-HA- β -caténine-EGFP (T41 *S*-GlcNAc) vs TAT-HA- β -caténine-mCherry (T41E).

Nous pouvons constater que chaque isoforme occupe un espace spécifique et quelque peu différent de l'isoforme co-transduite. Ceci indiquerait qu'en fonction du statut de modification de T41, la β -caténine adopte un comportement différent probablement dû à une modification locale de sa conformation impactant sa localisation subcellulaire.



E



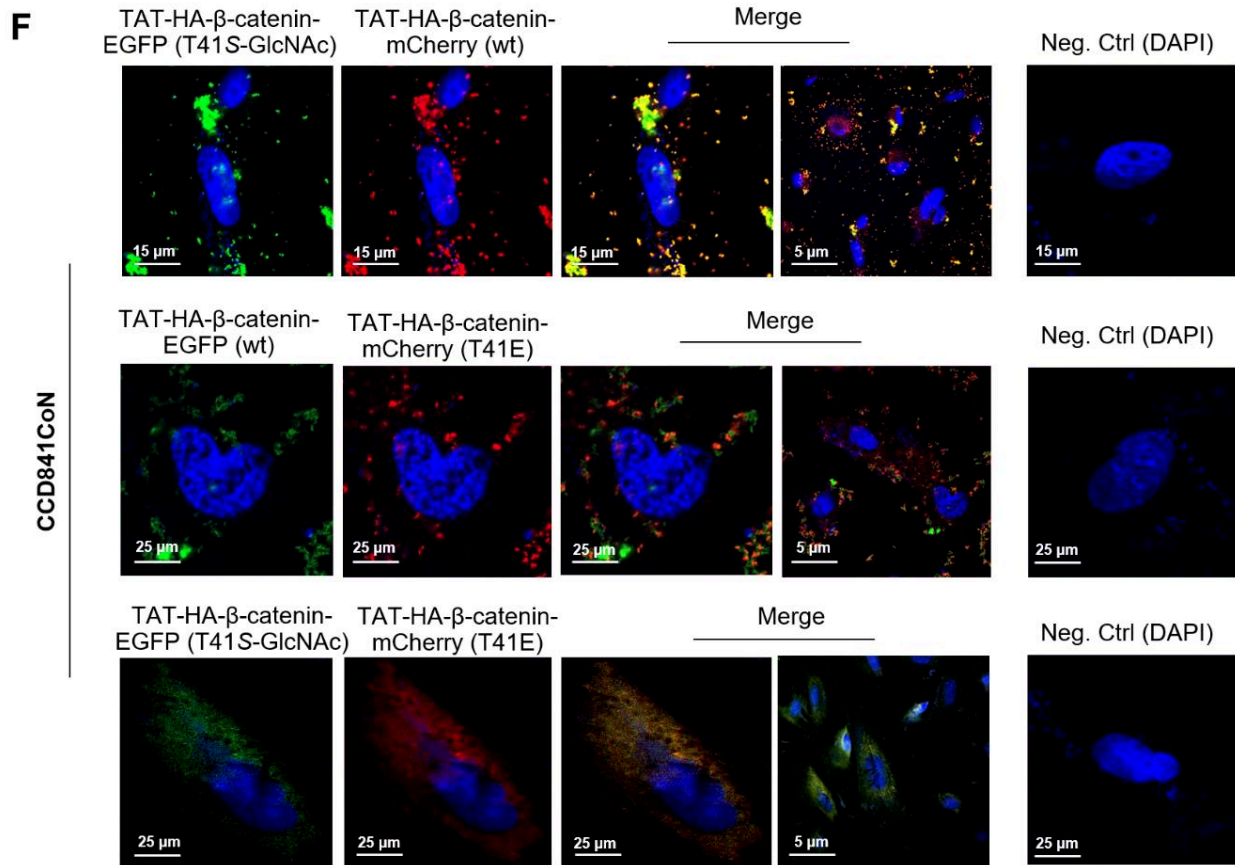
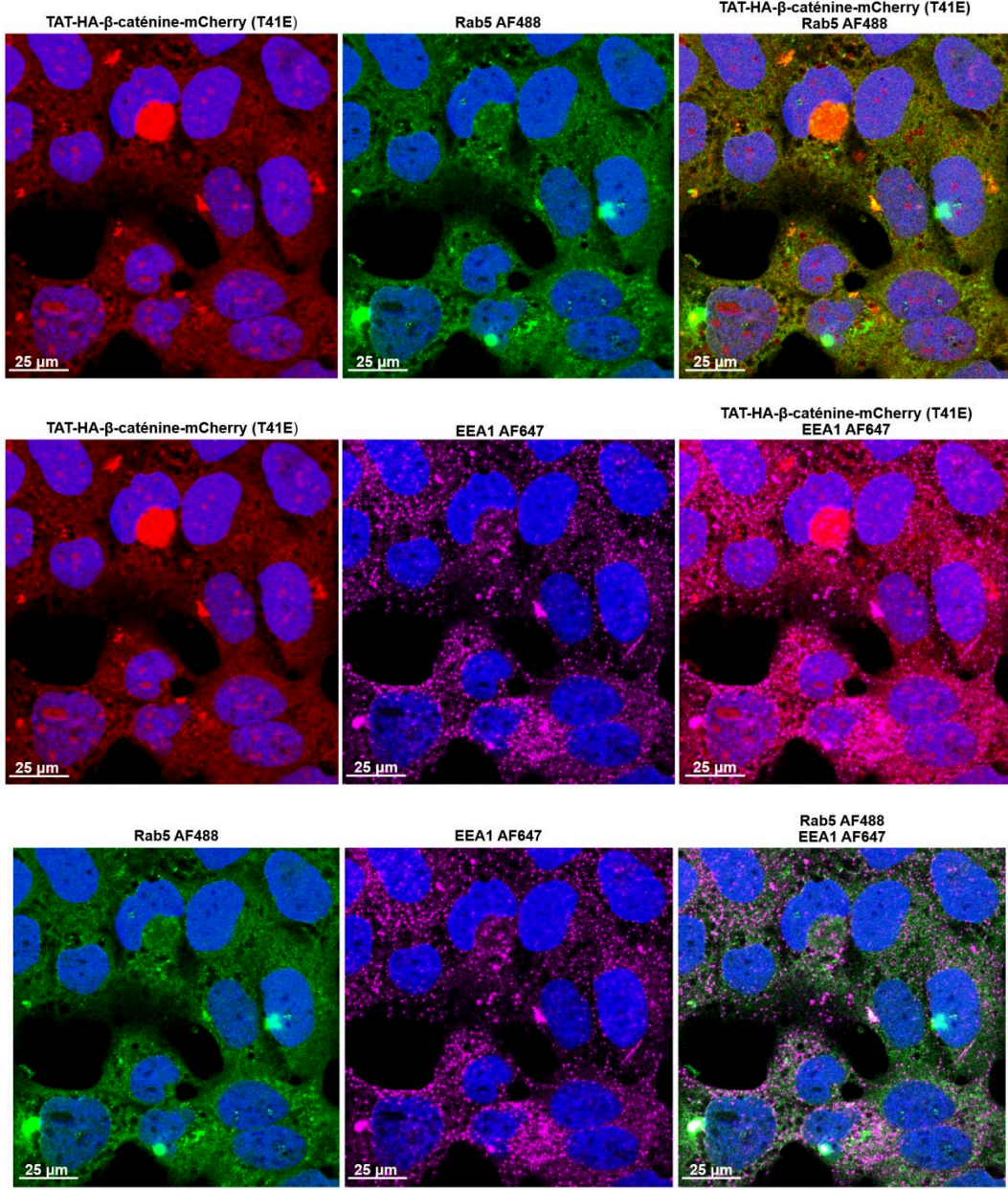


Figure 22. Traçage des différentes isoformes de la TAT-HA- β -caténine par transduction des cellules Hep3B et CCD841CoN. **A-C**, évaluation par Western blot anti-HA de l'expression des différentes isoformes de la TAT-HA- β -caténine dans les bactéries *E. coli* à plusieurs temps d'induction avec 0,1% (m/v) de L-arabinose à 37°C sous agitation à 150 rpm. **D**, contrôle de la S-GlcNAcylation de la TAT-HA- β -caténine après purification, oxydation et addition de Michael. Les échantillons ont été également contrôlés par coloration au bleu de Coomassie. **E-F**, images obtenues par microscopie confocale de cellules Hep3B et CCD841CoN transduites pendant deux heures avec les différentes isoformes de la TAT-HA- β -caténine-XFP (2 μ M chacune). Les noyaux ont été marqués au DAPI. La condition contrôle représente les cellules non transduites. L'expérience a été réalisée en triplicat. Nombre de cellules étudiées : 20 cellules Hep3B et 12 cellules CCD841CoN/ par image. Le nombre de cellules traitées : 20 par image.

3. Contrôle de la β -caténine transduite par le peptide TAT.

L'aspect des marquages présentés en **figure 22**, en "paquets", nous a quelque peu surpris et laissés perplexes quant à leur apparence, une internalisation par endocytose ayant été envisagée. Pour répondre à cette interrogation, une expérience combinant la transduction cellulaire des protéines à une immunofluorescence indirecte a été réalisée. L'import, le recyclage, le tri et le transport des matériaux cellulaires sont assurés par des protéines impliquées dans la formation et la maturation d'endosomes précoces dans la voie endocytique cellulaire (López-Hernández et al., 2020). Ainsi, nous avons ciblé deux acteurs de l'endocytose précoce de sorte à vérifier que les isoformes de la β -caténine associées au peptide TAT sont transduites dans les cellules sous forme soluble plutôt que vésiculaire : Rab5, une GTPase régulant les premières étapes d'endocytose et contrôlant la fusion des endosomes précoces entre eux et les vésicules endocytiques entrantes (Nagano et al., 2019), et EEA1 (*Early Endosome Antigen 1*), une protéine membranaire périphérique se liant spécifiquement à Rab5-GTP et associée aux endosomes précoces (Ohya et al., 2024). En premier lieu, une transduction de la forme phosphomimétique TAT-HA- β -caténine-mCherry (T41E) a été réalisée dans les cellules Hep3B pendant deux heures, suivie d'une immunofluorescence indirecte ciblant les deux marqueurs d'endosomes précoces en utilisant les anticorps anti-Rab5 et anti-EEA1 puis les anticorps secondaires marqués respectivement avec l'Alexa Fluor 488 vert et l'Alexa Fluor 647 rouge lointain. Les résultats de microscopie (**Figure 23**) montrent l'aspect et la localisation des protéines transduites par rapport aux marquages des endosomes précoces Rab5 et EEA1. Bien que le marquage ne soit pas parfait, il semblerait que la protéine TAT-HA- β -caténine-mCherry phosphomimétique co-localise en partie avec les protéines Rab5 et EEA1, comme indiqué par la superposition partielle de la couleur rouge (TAT-HA- β -caténine-mCherry) avec les couleurs verte et rouge lointain spécifiques des protéines Rab5 et EEA1. Ceci suppose que la forme transduite soit en partie empaquetée dans des vésicules d'endocytose, posant problème pour une protéine cytosoluble. Néanmoins la colocalisation n'étant pas transcendante, il se pourrait que les formes transduites de β -caténine forment des agrégats.



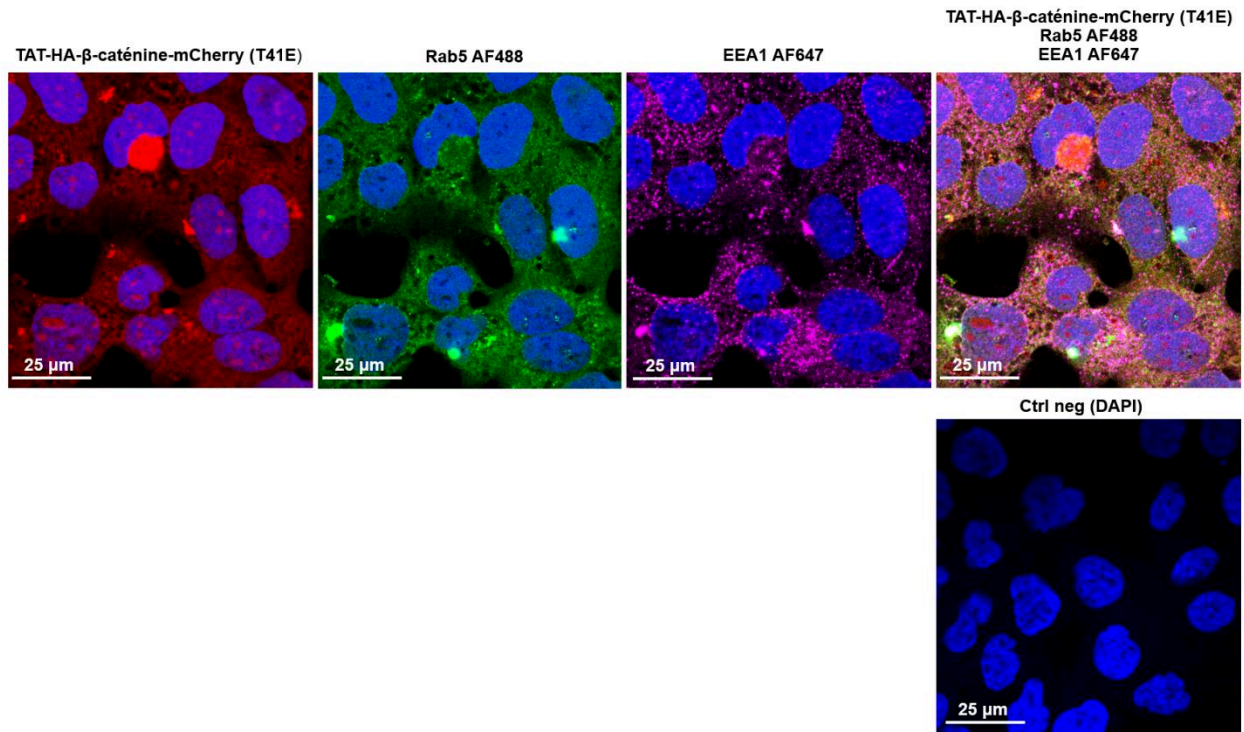


Figure 23. Contrôle de co-marquage de la forme phosphomimétique TAT-HA-β-caténine-mCherry avec les marqueurs d'endosome précoce Rab5 et EEA1 en cellules Hep3B. Des concentrations de 2 μM de TAT-HA-β-caténine-mCherry (T41E) ont été transduites pendant 2 heures. Puis, une immunofluorescence indirecte a été réalisée par marquage de Rab5 en vert (Alexa Fluor 488), et EEA1 en rouge lointain avec (Alexa Fluor 647). Le noyau a été marqué avec le DAPI. Les conditions contrôles comprennent les cellules Hep3B non traitées, ni avec la protéine transduite, ni marquées par immunofluorescence indirecte. L'expérience a été réalisée en triplicat. Les résultats sont représentés par des images issues de la microscopie confocale Zeiss LSM 780. Nombre de cellules traitées : 20 par image.

4. Clonage et assemblage moléculaire des constructions sans TAT-HA

Puisque le peptide TAT-HA semble être à l'origine du problème de l'aspect vésiculaire et/ou agrégats des protéines dans les cellules déjà décrit par le groupe de Patel et ses collègues (Patel et al., 2019), nous avons décidé de générer des constructions sans TAT-HA (**Figure 24**).

Pour déléter le TAT-HA des différentes constructions plasmidiques nous avons utilisé tout simplement la PCR. L'ensemble des composants nécessaires à cette PCR, ainsi que les souches bactériennes, plasmides, amorces, et procédure sont présentés dans les **tableaux 4-7**.

Élimination du TAT-HA de toutes les constructions plasmidiques

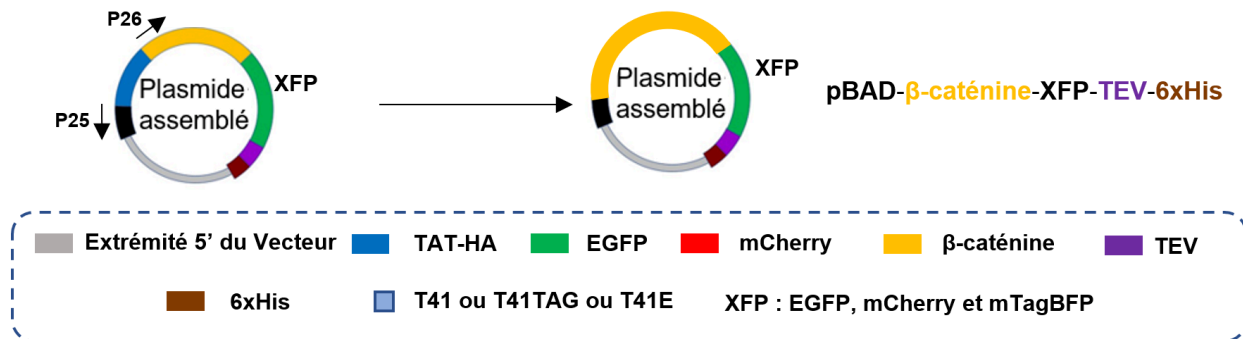


Figure 24. Élimination de la partie TAT-HA des différentes constructions par PCR à l'aide du couple d'amorces P25-P26.

Les conditions d'assemblage de l'ADN plasmidique, ainsi que l'assemblage d'ADN plasmidique par méthode Gibson pour produire les neuf constructions pBAD-TAT-HA-β-caténine-XFP-TEV-6xHis (wt, T41TAG et T41E) de 7071 pb, sans TAT-HA sont présentées dans les **tableaux 6-9**.

Après transformation dans des bactéries DH5α, le séquençage moléculaire a confirmé l'élimination du TAT-HA pour obtenir les constructions finales après purification sur gel (**Figures 25 et 26**).

P25-P26

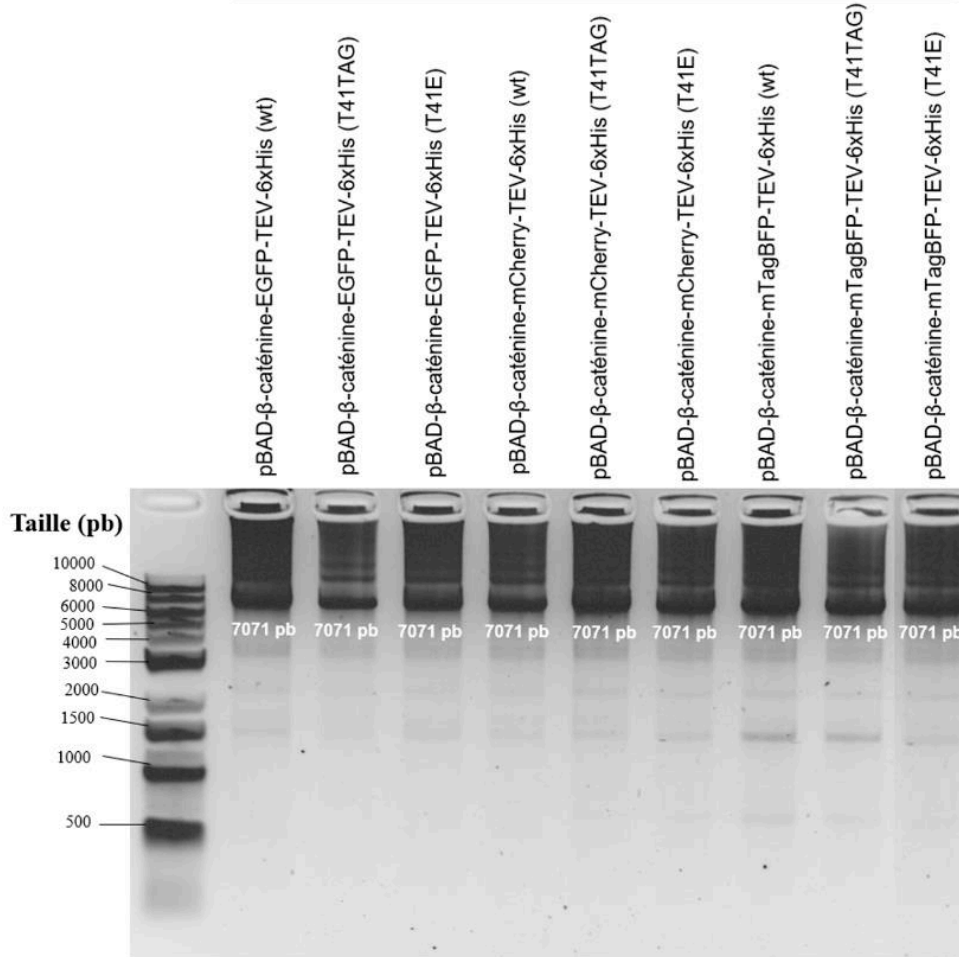


Figure 25. Caractérisation des différentes constructions délétées de la séquence TAT-HA par électrophorèse en gel d'agarose. Les marqueurs de taille sont indiqués à gauche du gel : ladder d'ADN 1 kb Plus.

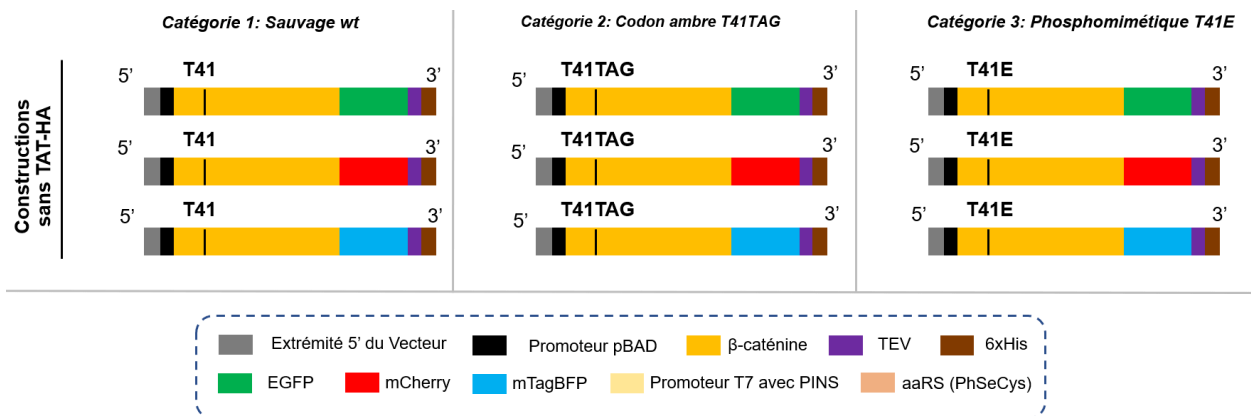


Figure 26. Résumé graphique des différentes constructions plasmidiques sans TAT-HA utilisées dans ce projet. La taille des plasmides est de 7071 pb.

5. Evaluation des β -caténines transfectées sans TAT-HA

Pour évaluer l'impact du peptide de pénétration TAT sur l'aspect de la β -caténine transduite, comme observé dans les **figures 22 et 23**, nous avons réalisé une expérience dans laquelle nous avons comparé la transfection de l'isoforme β -caténine-EGFP (wt) à la transduction de l'isoforme TAT-HA- β -caténine-mCherry (wt) en cellules HeLa et Hep3B (**figure 27**). Pour la lipofection, nous avons choisi la lipofectamine 3000 (Shao et al., 2017). Nous constatons que l'introduction de la β -caténine-EGFP (wt) sans TAT-HA par lipofection règle en grande partie le problème d'agrégation observé avec la protéine TAT-HA- β -caténine-mCherry (wt) introduite par transduction cellulaire.

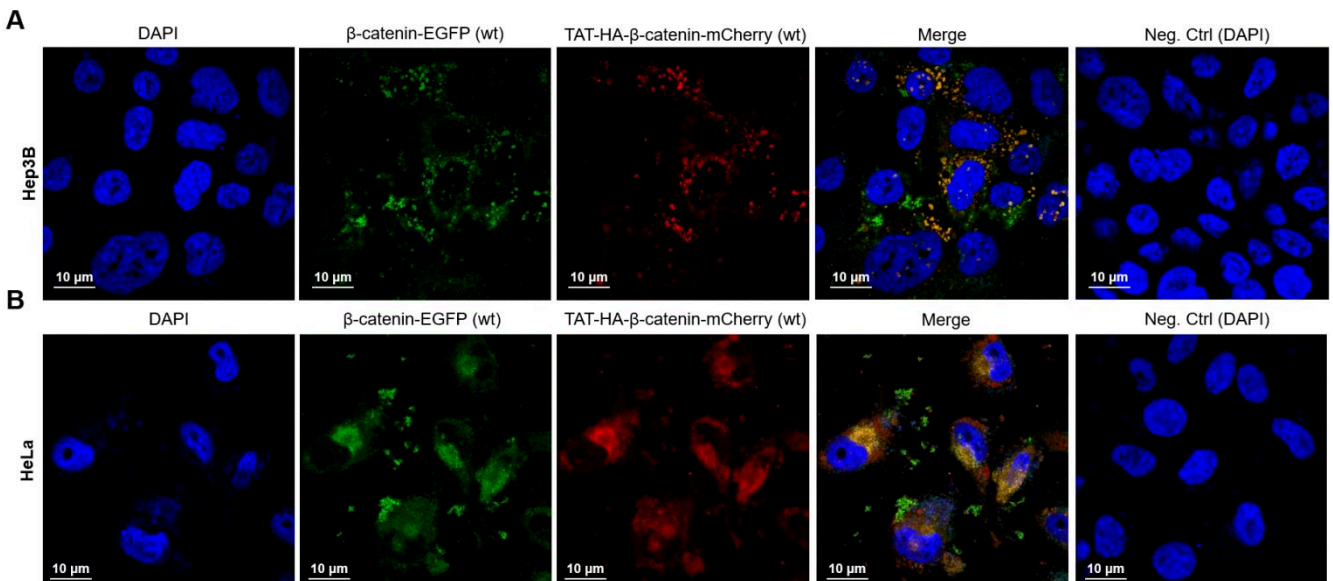


Figure 27. Approche combinant la lipofection d'une β -caténine-EGFP (wt) sans TAT-HA et la transduction cellulaire d'une TAT-HA- β -caténine-mCherry (wt) en Hep3B (A) et HeLa (B). Les noyaux ont été marqués au DAPI. Les images ont été obtenues par microscopie confocale Zeiss LSM 780. L'expérience a été réalisée en triplicat. Nombre de cellules traitées : 20 par image.

Nous avons réalisé deux expériences complémentaires à la lipofection. Pour la première, un plasmide codant la protéine 2xFlag- β -caténine sauvage (wt) non fluorescente a été transfecté dans les cellules Hep3B et HeLa pendant 72 heures, suivie d'une lipofection de la protéine β -caténine-mCherry sauvage (wt) sans TAT-HA

pendant 2 heures. La détection de la protéine 2xFlag- β -caténine (wt) a été réalisée par immunofluorescence indirecte en utilisant des anticorps anti-flag M2 (souris), et des anticorps secondaires Alexa-fluo 488 (souris). Dans la seconde expérience, un plasmide fluorescent codant la protéine EGFP- β -caténine sauvage (wt) a été transfecté dans les cellules Hep3B et HeLa pendant une période de 72 heures. Ensuite, la protéine β -caténine-mCherry sauvage (wt) sans TAT-HA a été introduite dans les mêmes cellules par lipofection. Les résultats des deux expériences sont présentés en **figure 28**.

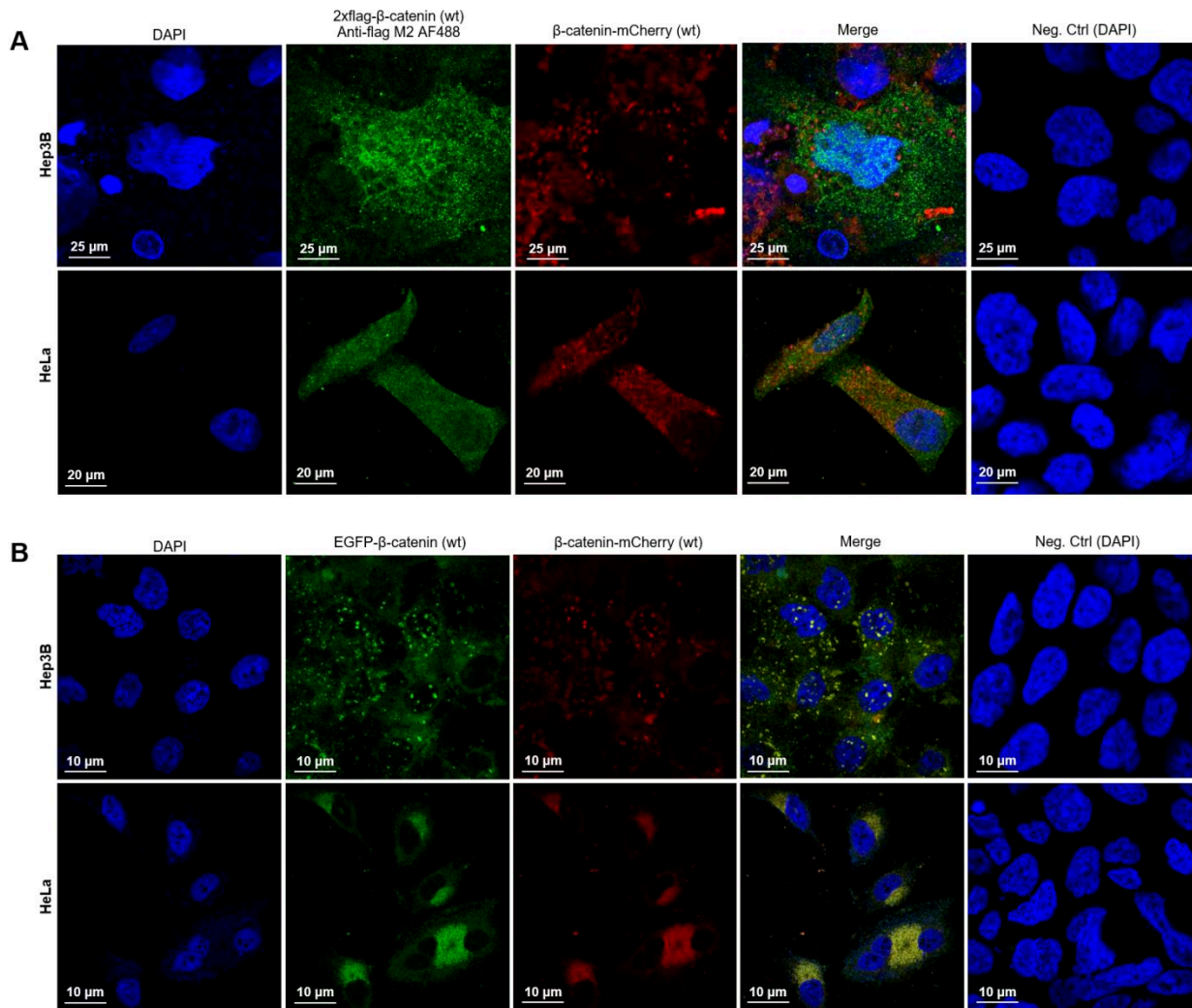


Figure 28. Expression d'une β -caténine codée par un vecteur plasmidique versus β -caténine produite chez *E. coli* et lipofectée. A, transfection pendant 72 heures d'un plasmide codant une 2x-flag- β -caténine (wt), suivie de deux heures de transfection de 2 μ M de β -caténine-mCherry (wt), respectivement dans les cellules Hep3B et les cellules HeLa, accompagnées d'une immunofluorescence indirecte d'un tag 2xflag en vert (anticorps anti-flag

M2 suivi d'un anticorps secondaire Alexa Fluor 488). **B**, transfection pendant 72 heures d'un plasmide codant une protéine EGFP- β -caténine (wt), suivie de deux heures de transfection de 2 μ M de β -caténine-mCherry (wt), respectivement dans les cellules Hep3B et les cellules HeLa. Les noyaux ont été marqués au DAPI. La condition contrôle négatif représente les cellules non transfectées ni avec les plasmides et ni avec les protéines. Les résultats d'imagerie sont issus de la microscopie confocale Zeiss LSM 780. L'expérience a été réalisée en triplicat. Nombre de cellules traitées : 20 par image.

Les résultats de la **figure 28** indiquent que les protéines β -caténine-mCherry (wt) introduites dans les cellules Hep3B et HeLa présentent la même distribution que celles couplées EGFP et codées par un vecteur plasmidique. Les résultats sont moins probants concernant l'expérience mettant en jeu la β -caténine non fluorescente et la β -caténine détectée par immunofluorescence indirecte.

Nous pouvons conclure que l'introduction des protéines par transfection résout en partie le problème lié à l'agrégation observée lors de la transduction cellulaire des protéines *via* le peptide TAT-HA en cellules Hep3B et HeLa.

Nous avons produit les trois isoformes de la β -caténine sans TAT-HA, à savoir sauvage non modifiée (wt), *S*-GlcNAcylée en T41 (T41*S*-GlcNAc), et phosphomimétique en T41 (T41E) respectivement fusionnées avec les fluorophores mTagBFP, EGFP, et mCherry. L'obtention des protéines purifiées a été confirmée par SDS-PAGE suivi d'une coloration au bleu de Coomassie ou d'un Western Blot avec les anticorps anti-6xHis, anti- β -caténine E5, et anti-*O*-GlcNAc (RL2) (**Figures 29A et 29B**).

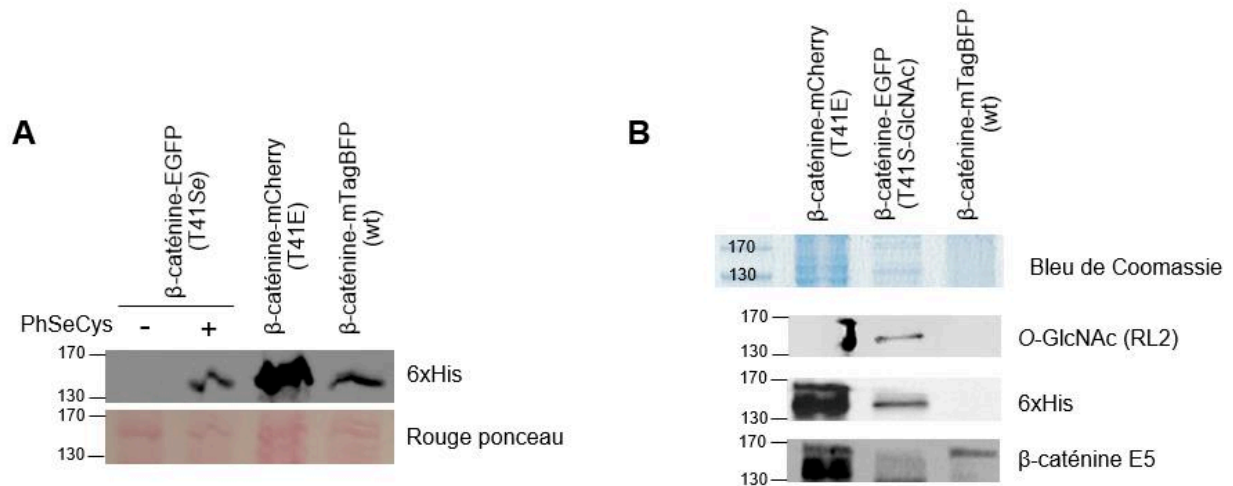


Figure 29. Évaluation de l'expression des trois isoformes de la β -caténine avant et après purification. **A**, Western blot 6xHis évaluant l'expression des trois isoformes de la β -caténine-XFP dans les bactéries *E. coli* après une heure d'induction avec 0,1% (m/v) de L-arabinose à 37°C sous agitation à 150 rpm. **B**, Western blot anti-O-GlcNAc, anti-6x-His et anti- β -caténine-E5 et coloration au bleu de Coomassie d'un gel SDS-PAGE dans lequel ont été résolues les trois isoformes de β -caténine après purification.

La **figure 29A** atteste de l'expression des trois isoformes de la β -caténine avant toute réaction chimique sur la forme sélénée. La **figure 29B** atteste quant à elle de l'obtention des trois isoformes de la β -caténine après purification et réaction chimique sur la forme sélénée. En effet, la S-GlcNAcylation de la β -caténine-EGFP en T41 est avérée par le WB anti-O-GlcNAc (RL2): présence d'une bande correspondant à la protéine β -caténine-EGFP (T41S-GlcNAc), et absence de bande pour les protéines β -caténine-mTagBFP (wt) et β -caténine-mCherry (T41E).

La lipofection des protéines a ensuite été réalisée avec 2 μ M de chaque protéine pendant 2 heures (**Figure 30**). Chaque expérience a été réalisée en triplicat en utilisant la lipofectamine 3000 dans le but de suivre le comportement et la localisation de chaque isoforme de la β -caténine dans les cellules cancéreuses de foie Hep3B et de col de l'utérus HeLa (**Figure 30**).

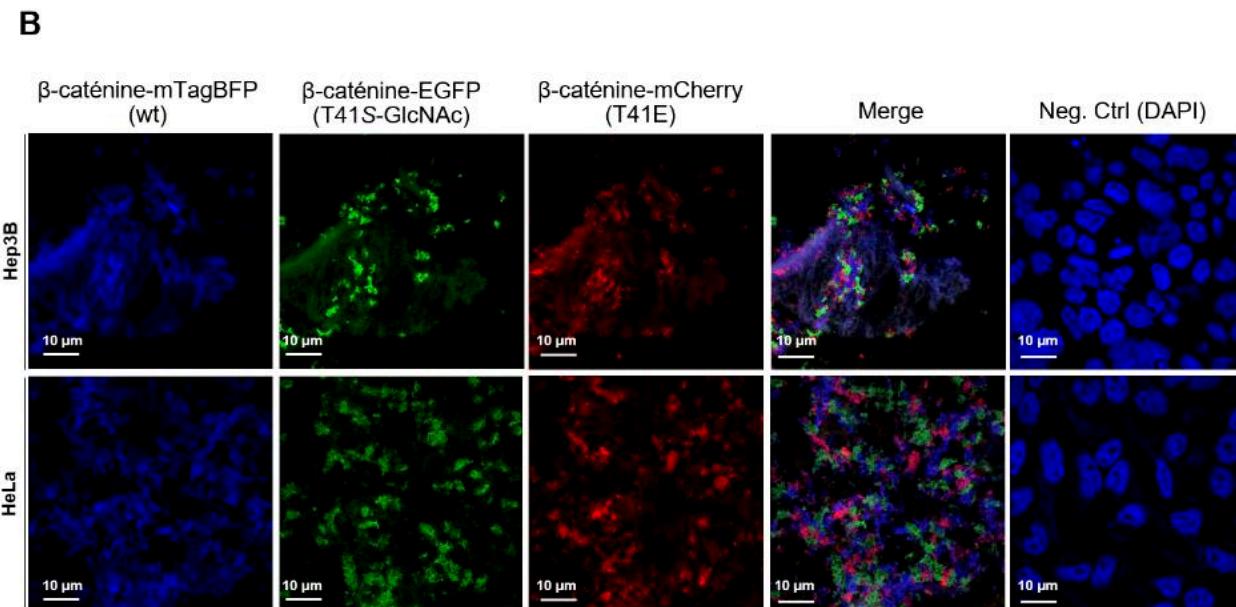
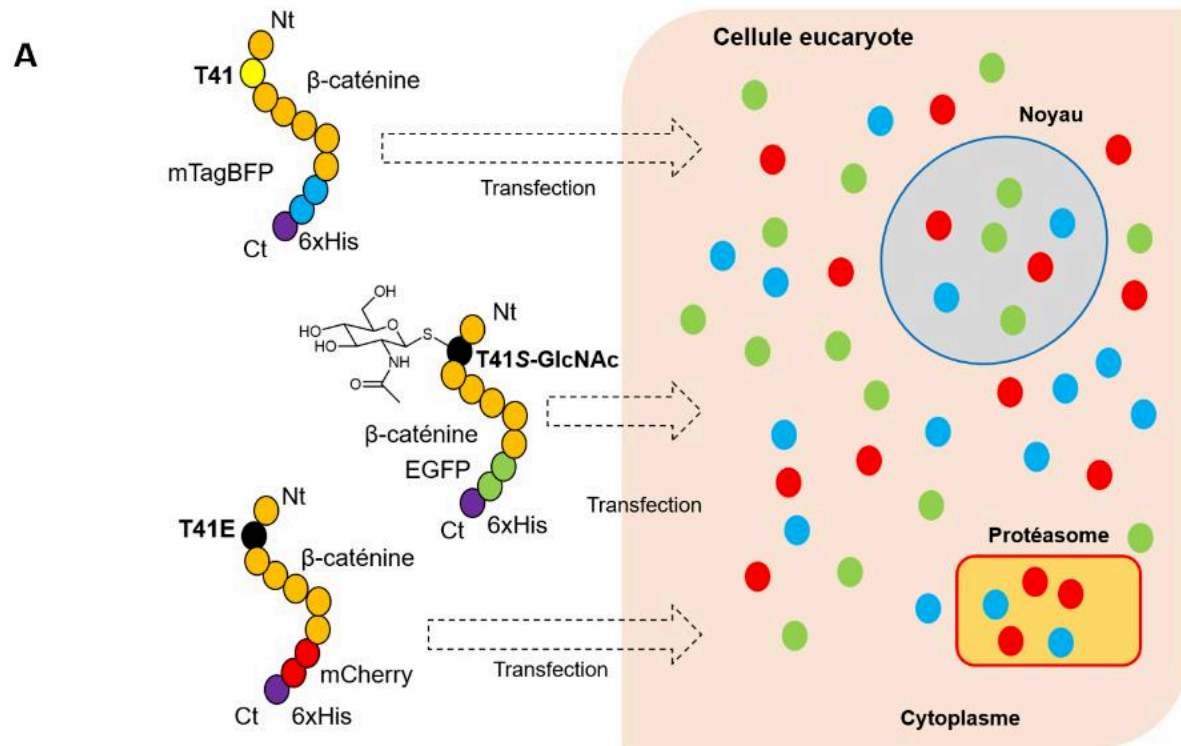


Figure 30. Introduction des trois isoformes de la β -caténine dans les cellules Hep3B et HeLa par lipofection. **A**, procédure de transfection des protéines par lipofection. **B**, résultats de microscopie confocale (Zeiss LSM 780) obtenus après lipofection des différentes isoformes de la β -caténine à 2 μ M chacune dans les cellules Hep3B et les cellules HeLa pendant une durée de 2 heures. Les noyaux ont été marqués au DAPI. La condition contrôle représente les cellules non transduites. Le test a été réalisé en triplicat avec un nombre de cellules de 20 par image.

Les résultats d'imagerie issus de la microscopie confocale exposés en **figure 30** montrent une répartition et une localisation différentes des isoformes de la β -caténine (wt) en bleu, (T41S-GlcNAc) en vert et (T41E) en rouge. Ceci semblerait indiquer que la modification en T41 de la β -caténine, O-GlcNAcylation ou phosphorylation respectivement mimés par le résidu S-GlcNAc et la mutation T41E, affecte leur comportement au sein même de la cellule, que ce soit la lignée Hep3B ou la lignée HeLa. Ces résultats indiquent également que la lipofection est plus adaptée à ce type d'approche expérimentale en comparaison à la transduction cellulaire.

Un test de viabilité cellulaire MTS a été réalisé en triplicat pour s'assurer de la non toxicité des protéines recombinantes après lipofection en cellules Hep3B et HeLa (**Figure 31**).

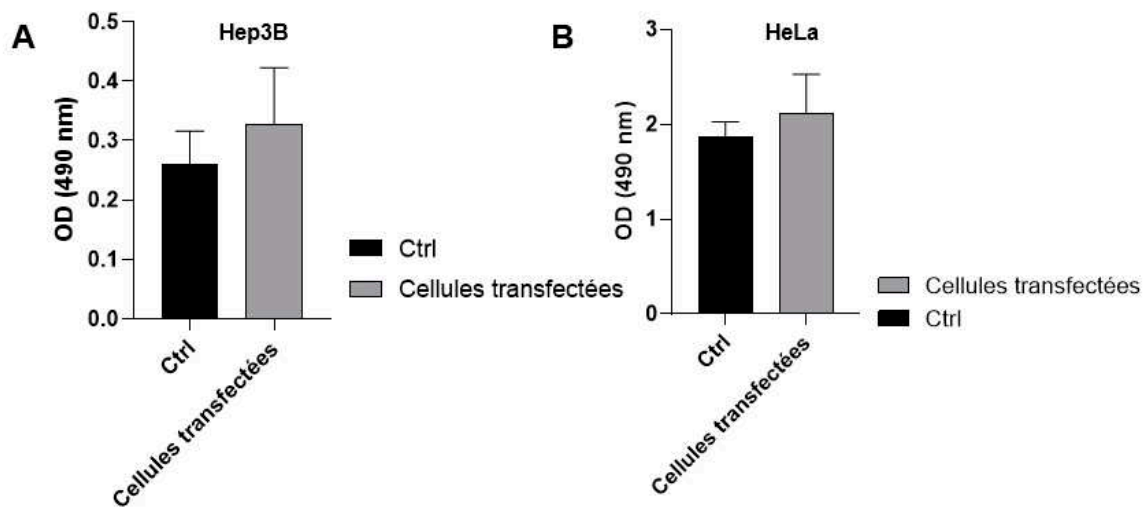


Figure 31. Évaluation de viabilité cellulaire. Des tests de viabilité cellulaire par la méthode au MTS (DO à 490 nm) ont été réalisés respectivement sur les cellules Hep3B (**A**) et HeLa (**B**) traitées avec les trois isoformes de la β -caténine à une concentration de 2 μ M chacune pendant 2 heures. Le test a été réalisé en triplicat.

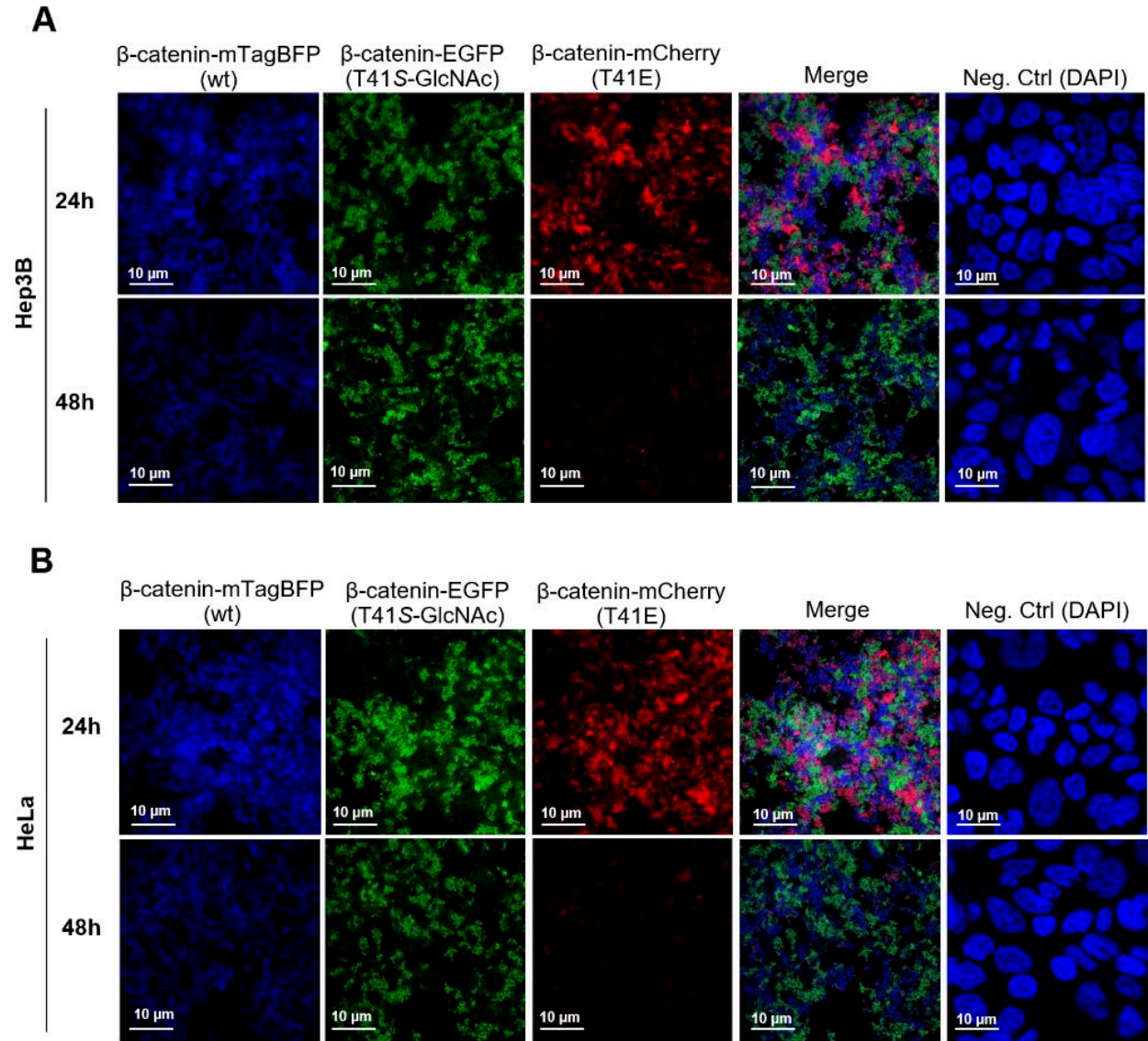
Au vu des résultats présentés en **figure 31**, nous pouvons conclure que la lipofection des différentes isoformes de β -caténine n'affecte pas significativement la survie cellulaire des lignées Hep3B et HeLa. Ces protéines ne présentent donc pas d'effet toxique sur les cellules, tout au moins sur des temps courts d'exposition.

En résumé, dans cette partie, nous avons produit, purifié et introduit les trois isoformes de la β -caténine dans des cellules en culture après 2 h de lipofection, tout en montrant que ces protéines n'avaient pas d'effet toxique significatif sur les cellules.

6. Impact de la S-GlcNAcylation sur l'expression de la β -caténine.

Dans le but de mieux comprendre l'effet de la O-GlcNAcylation en T41 de la β -caténine, nous avons prolongé le temps d'incubation des trois isoformes de la β -caténine.

La lipofection des protéines a été effectuée en triplicat en utilisant la lipofectamine 3000 à une concentration de 2 μ M pour chaque protéine, pendant 24 et 48 heures, dans le but de suivre le niveau moyen de fluorescence, et la localisation de chaque isoforme dans les cellules Hep3B et HeLa (**Figure 32**).



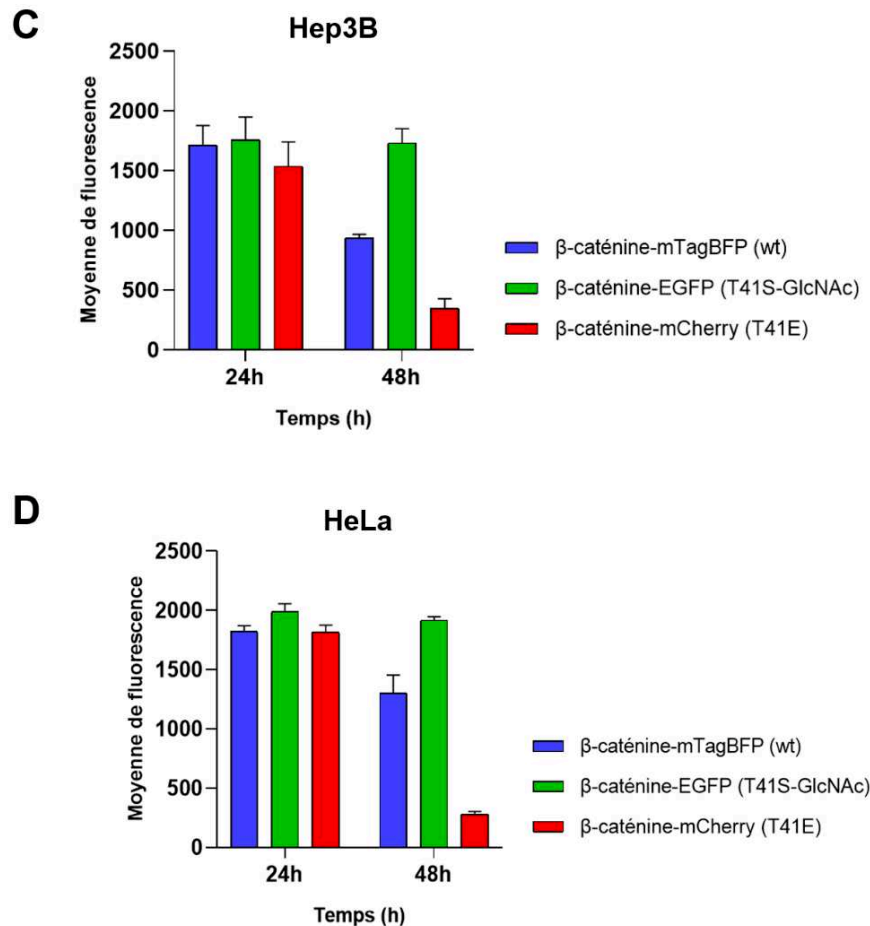


Figure 32. Suivi de la S-GlcNAcylation et de la forme phosphomimétique du résidu T41 de la β -caténine-XFP pendant 24 et 48 heures. **A** et **B**, résultats de microscopie confocale Zeiss LSM 780 obtenus après lipofection des différentes isoformes de la β -caténine (2 μ M chacune) respectivement dans les cellules Hep3B et les cellules HeLa pendant une durée de 24 et 48 heures. Les noyaux ont été marqués au DAPI. La condition contrôle représente les cellules non transfectées. **C** et **D** histogrammes représentatifs de la fluorescence des différentes isoformes de β -caténine en cellules Hep3B et HeLa. L'expérience a été réalisée en triplicat. Nombre de cellules traitées : 20 par image.

Les résultats d'imagerie par microscopie confocale de la **figure 32** montrent ici encore une répartition et localisation différentes des isoformes de la β -caténine (wt) en bleu, (T41S-GlcNAc) en vert et (T41E) en rouge. D'autre part, nous observons une diminution très significative de l'intensité du fluorophore mCherry (en rouge) fusionnée à la forme phosphomimétique (T41E) de la β -caténine par rapport à l'intensité des fluorophores mTagBFP et EGFP fusionnés respectivement aux formes sauvage (wt) et T41S-GlcNAcyliée de β -caténine entre des temps de lipofection de 24 h et de 48 h. Il semblerait que l'intensité du fluorophore EGFP reste quasi constante pendant toute la

durée de l'expérience, suggérant une résistance à la dégradation protéasomale de la forme *S*-GlcNAcylée en comparaison à la forme phosphomimétique (T41E). Une diminution significative, mais moindre, de l'intensité du fluorophore mTagBFP fusionnée à la forme sauvage (wt) a été également observée par rapport à la forme *S*-GlcNAcylée en T41. Ceci indique que la β -caténine non modifiée en T41 est dégradée par le protéasome après 48 heures de traitement, possiblement *via* la phosphorylation de T41 par la machinerie cellulaire et polyubiquitination subséquente des résidus K19 et K49. Ceci renforce l'idée selon laquelle la *S*-GlcNAcylation en T41 de la β -caténine stabilise et protège cette dernière de la dégradation protéasomale observée pour les formes phosphomimétique (T41E) et sauvage (wt). De plus, la forme phosphomimétique T41E semble plus susceptible à la dégradation protéasomale que la forme sauvage (wt) (non modifiée en T41).

Chapitre IV : Discussion

A l'heure actuelle, la santé publique mondiale est confrontée à des défis importants liés à l'augmentation de l'incidence du cancer et des pathologies métaboliques. Ces problèmes peuvent être associés à des perturbations de l'homéostasie protéique cellulaire et des voies de signalisation découlant de dysfonctionnements des MPT. Mon projet de thèse s'est inscrit dans ce contexte, visant à étudier l'impact et le suivi de la *O*-GlcNAcylation et de la phosphorylation du résidu thréonine 41 de la β -caténine, une protéine dont la dysrégulation est impliquée dans différents types de cancers.

Bien qu'il existe de nombreuses approches et stratégies pour détecter la *O*-GlcNAcylation des protéines cellulaires, l'expansion du code génétique, décrit notamment par Peter Schultz (Wang and Schultz, 2005), n'a pour le moment jamais été utilisée dans cette optique: c'est pourquoi, combiner cette approche au traçage des protéines *O*-GlcNAcylées a constitué l'objectif principal de mon projet de thèse. Nous avons adapté cette méthode en ciblant spécifiquement la *O*-GlcNAcylation de la T41 de la β -caténine. Pour cela, nous avons dû surmonter de nombreux problèmes qui à l'issue de cette thèse permettent de proposer cette approche expérimentale comme une alternative directe à d'autres approches comme la stratégie des rapporteurs métaboliques basée sur l'ingénierie métabolique des oligosaccharides, MOE, décrit par les groupes de Carolyn R. Bertozzi (Goon and Bertozzi, 2002), Benjamin Schumann (Cioce et al., 2021) et Christophe Biot (Gilormini et al., 2018).

Le choix de cette stratégie s'explique par une très simple comparaison aux stratégies existantes. Contrairement à l'utilisation des rapporteurs métaboliques, notre approche permet une modification cible spécifique et contrôlée de n'importe quelle protéine telle que nous l'avons développée autour de la T41 de la β -caténine. Ceci évite tout problème de spécificité et manque de précision des rapporteurs chimiques pouvant réagir avec divers groupes fonctionnels présents dans la cellule, entraînant la détection de faux positifs. En outre, l'incorporation de rapporteurs chimiques dans des protéines peut potentiellement perturber leur conformation, leur fonction ainsi que leur interaction

avec d'autres molécules de l'environnement. Ces problèmes liés à l'utilisation de la stratégie de rapporteurs chimiques ont été décrits par Nichole J. Pedowitz et Matthew R. Pratt (J. Pedowitz and R. Pratt, 2021). Notre stratégie permet une incorporation d'un acide aminé non naturel, la phénylsélocystéine dans notre cas, sur un site cible de la protéine d'intérêt tel que la T41, minimisant toute modification possible affectant la structure et la fonction de la protéine d'intérêt comme décrit par Peter Schultz (Wang and Schultz, 2005).

De même, cette méthode peut être combinée avec d'autres approches expérimentales tels que l'utilisation d'anticorps anti-*O*-GlcNAc (Holt and Hart, 1986; Comer et al., 2001) et de lectines telle la WGA (Kubota et al., 2017) en WB ou en enrichissement, la spectrométrie de masse (Ma and Hart, 2017) ou le double marquage basé sur le FRET (Kasprowicz et al., 2020), pour accéder à des informations supplémentaires sur la *O*-GlcNAcylation de sites spécifiques de protéines d'intérêts. L'idée principale de ce projet de thèse a été de produire des molécules d'ADN recombinantes à façon par clonage moléculaire et de les utiliser pour transformer des souches bactériennes *E. coli* DH5 α (**Figure 18**). Les fragments d'ADN ont été amplifiés par PCR, une technique d'amplification enzymatique permettant d'obtenir un grand nombre de copies identiques d'un fragment d'ADN (Erlich, 1989). L'obtention des différents produits de PCR n'a été réalisée qu'après avoir optimisé les conditions tels que le choix de l'ADN polymérase Q5, la concentration du plasmide matrice, la séquence des oligonucléotides, ou encore la température d'hybridation (**Tableaux 4-7**). Nous avons également optimisé les conditions d'assemblage moléculaire décrit par Gibson et ses collègues (Gibson, 2009; Gibson et al., 2009) pour obtenir la construction finale (**Tableau 9**) et les conditions de recircularisation des plasmides linéaires (**Tableau 8**).

En résumé de cette première partie de thèse (**Figures 18-21**), nous avons réussi à produire une série de plasmides pBAD-TAT-HA- β -caténine-XFP-TEV-6xHis, de 7193 pb chacun, contenant un TAT-HA. Cet ensemble de plasmides comporte une version sauvage (wt) de la β -caténine, une version mutée en position T41 avec un codon *ambre*, et une version mutée sur la même position mais avec un codon GAA codant un acide glutamique. Chaque version arbore une couleur spécifique grâce à un fluorophore

de type EGFP, mCherry ou mTagBFP (**Figure 18, 19 et 21**). Nous avons également construit le plasmide codant la synthétase orthogonale aaRS (PhSeCys). (**Figures 20 et 21**)

Pour tenter de répondre à des questions relatives à la stabilité des protéines S-GlcNAcylées et de la version phosphomimétique au niveau du résidu T41 de la β -caténine, nous avons opté pour une méthode d'introduction de ces protéines dans les cellules cancéreuses HeLa et Hep3B, et non pathologiques CCD841CoN : la transduction cellulaire (**Figures 22E, 22F et 23**).

La technologie de suppression de codon stop ambre TAG (expansion du code génétique) est largement décrite dans la littérature, notamment par le groupe de Peter G. Schultz et ses collègues (Wang et al., 2001; Wang and Schultz, 2005). En appliquant cette technologie, nous avons incorporé la phénylsélocystéine au codon T41, dans une β -caténine-EGFP exprimée dans une souche de bactéries *E. coli* de type BL21 (DE3) (**Figure 22B**). Nous avons également produit les isoformes sauvage (wt) et phosphomimétique en T41 de la β -caténine (**Figures 22A et 22C**). La production de ces trois isoformes a été optimisée pour une induction d'une heure avec 0,1% (m/v) de L-arabinose à 37°C, sous agitation à 150 rpm. Nous avons également dû optimiser d'autres paramètres tels que la composition des tampons de lyse et d'élution, la dialyse, la concentration des protéines pour pouvoir faire les tests biologiques *in vitro*.

La S-GlcNAcylation de la β -caténine-EGFP a été confirmée par WB anti-O-GlcNAc (RL2) (**Figure 22D**). Ces différentes isoformes de type TAT-HA- β -caténine-XFP ont alors été transduites sur une durée de deux heures à une concentration de 2 μ M chacune dans les cellules Hep3B et les cellules "normales" de côlon CCD481CoN (**Figures 22E et 22F**) comme décrit par les groupes de Steven F. Dowdy, Tarun B. Patel et Yifu Guan (Vocero-Akbani et al., 2000; Yigzaw et al., 2001; Guo et al., 2012).

Pour tenter d'identifier le problème lié à l'agrégation probable des différentes β -caténines suite à leur transduction cellulaire et possiblement provoqué par le peptide de pénétration TAT, nous avons couplé la transduction cellulaire à une immunofluorescence indirecte pour cibler les protéines impliquées dans la formation

d'endosomes précoces, Rab5 et EEA1. Les résultats de ces analyses montrent que la protéine TAT-HA- β -caténine-mCherry phosphomimétique co-localise en partie avec les protéines Rab5 et EEA1 dans les cellules Hep3B, comme l'indique la superposition partielle de la couleur rouge liée à la construction TAT-HA- β -caténine-mCherry aux couleurs verte et rouge lointain spécifiques des protéines Rab5 et EEA1. Ceci suppose que la forme transduite est en partie empaquetée dans des vésicules d'endocytose posant problème pour une protéine cytosoluble (**Figure 23**). Cette colocalisation très partielle présuppose que les formes de β -caténine fluorescentes couplées au peptide TAT forment des agrégats potentiellement insolubles comme décrits dans l'étude de Patel et collaborateurs (Patel et al., 2019). Ces observations nous ont incités à produire des constructions dépourvues de séquence TAT-HA, évitant leur encapsulation dans des vésicules une fois introduites dans les cellules, et leur agrégation malgré une bonne efficacité de transduction cellulaire (**Figures 24-26**). Nous avons en conséquence produit et purifié les différentes nouvelles isoformes, et à *S*-GlcNAcyler la β -caténine-EGFP (**Figure 29B**).

Du fait d'une apparente agrégation des protéines ainsi produites, purifiées et transduites en cellules eucaryotes, nous avons construit une nouvelle série de plasmides de 7071 pb chacun, dépourvus de séquence TAT-HA (**Figures 24-26**). Toutes les constructions plasmidiques ont été confirmées par séquençage.

L'introduction de ces différentes isoformes de β -caténine en cellules eucaryotes a été réalisée par lipofection en utilisant la lipofectamine 3000 (Rust et al., 2015). Grâce aux expériences de contrôle présentées dans les **figures 27 et 28**, nous avons mieux assuré l'aspect cytosoluble des β -caténines sans TAT-HA dans les cellules HeLa et Hep3B contrairement à l'aspect des protéines avec TAT-HA observée au niveau cellulaire dans les **figures 22 et 23**. Ceci indique que l'introduction des protéines sans TAT-HA par transfection résout en partie le problème lié à l'agrégation observée lors de la transduction cellulaire des protéines *via* le peptide TAT-HA en cellules Hep3B et HeLa comme décrit par Patel et collaborateurs (Patel et al., 2019). L'efficacité de transfection semble avoir été optimale et de manière très intéressante nous avons pu observer que les différentes isoformes de β -caténine occupaient des zones différentes

dans les cellules Hep3B et HeLa, après 2 heures de traitement, ce qui à notre connaissance est totalement nouveau voire sans précédent dans le domaine des isoformes de MPT (**Figure 30**). Ceci indique que la MPT de T41 affecte le comportement géographique de la β -caténine, quelque-soit le type cellulaire, Hep3B ou HeLa, et renforce l'importance de T41 dans le comportement de la β -caténine, tel que préalablement décrit par Stéphanie Olivier-Van Stichelen dans notre équipe de recherche (Olivier-Van Stichelen et al., 2014). Ces résultats confirment également que la lipofection décrite par Bazbek Davletov et ses collègues (Rust et al., 2015), est une excellente technique d'internalisation de protéines en cellules, alternative à la transduction cellulaire.

Mosmann et ses collègues ont été les premiers à décrire l'évaluation de la viabilité cellulaire et de la toxicité *via* le test MTT/MTS à la suite d'un traitement cellulaire avec une protéine exogène ou bien une molécule bioactive. Depuis leur publication initiale en 1983, l'utilisation des tests de viabilité cellulaire est devenue un pilier essentiel en biologie cellulaire. La potentielle toxicité des différentes constructions a été testée *via* le test MTS : aucune différence significative n'a été relevée entre les cellules traitées et les cellules non traitées, quelle que soit l'isoforme considérée (**Figure 31**).

Des temps d'incubation des trois isoformes de la β -caténine plus longs, 24 et 48 heures, a permis de renforcer les observations à 2 heures de traitement quant à la répartition des différentes isoformes de la β -caténine, et de conforter l'effet stabilisant de la *O*-GlcNAcylation *via* la *S*-GlcNAcylation de la β -caténine-EGFP en comparaison aux isoformes sauvage et phosphomimétique, comme préalablement rapporté et décrit par le groupe de Matthew R. Pratt (De Leon et al., 2017). Nous avons en effet montré que le niveau protéique des formes sauvage et phosphomimétique est plus bas à 48 h en comparaison au temps de traitement 48 h, alors que la forme *S*-GlcNAcylée reste constante (**Figure 32**). Cette expérience est fondamentale dans notre projet puisqu'elle valide tout le potentiel de la stratégie mise en place : nous avons en effet re-démontré *via* un système très artificialisé, la dualité du site T41 sur la destinée de la β -caténine. La *S*-GlcNAcylation en T41 de la β -caténine semble avoir stabilisé et protégé cette dernière de la dégradation protéasomale, alors que la forme phosphomimétique (T41E)

et, dans une moindre mesure, la forme sauvage sont plus susceptibles à la dégradation très probablement protéasomale, point qu'il nous faudra éclaircir très prochainement.

L'aspect de distribution cellulaire des trois isoformes de la β -caténine dépourvues de TAT-HA (**Figures 30 et 32**) est particulièrement intrigant et ne trouve pas, pour le moment, d'explication. Des études complémentaires sont nécessaires pour tenter d'élucider ce phénomène. Il est tout à fait possible que selon l'isoforme de MPT, la β -caténine interagisse avec des partenaires différents la confinant dans des endroits différents de la cellule. Une des expériences qu'il sera intéressant de faire consiste en l'identification d'un interactome différentiel par protéomique comme décrit par le groupe de Anne-Claude Gingras (Lambert et al., 2013). La comparaison des partenaires d'interaction pourra orienter notre réflexion sur une partition des différentes isoformes en fonction des besoins de la cellule à un instant donné. Des expériences de PLA (*Proximity Ligation Assay*) (Söderberg et al., 2008; Alam, 2018) entre chacune des isoformes et le partenaire considéré devrait permettre de valider ou d'invalider l'hypothèse selon laquelle une protéine s'oriente vis-à-vis d'un gradient comme le fait toute molécule soluble ou selon un partenaire d'arrimage à un compartiment subcellulaire (lysosome, peroxyosome, vésicules de sécrétion...) ou à un organelle spécifique (ribosome, protéasome, granules de stress...).

Enfin, une des approches que nous n'avons malheureusement pas eu le temps de tester avec notre système expérimental est la vidéomicroscopie. L'idée est de travailler en cellules vivantes transfectées avec les trois isoformes de β -caténine et d'observer sur plusieurs heures leur déplacement dans la cellule jusqu'à leur emplacement de destination. L'application de tout un ensemble de drogues visant à interférer avec divers processus cellulaires tel que les mouvements entre le cytosol et le noyau, la dissociation des jonctions d'adhérence ou la désorganisation du cytosquelette, sont autant d'exemples d'application simple que notre stratégie expérimentale pourrait permettre de réaliser.

Chapitre V : Conclusion et perspectives

Au cours de ma thèse, je me suis attelé à la mise au point d'une stratégie expérimentale permettant de suivre une protéine sous diverses formes de MPT dans une même cellule. Pour ce faire, la thréonine 41 de β -caténine a été choisie comme modèle, ce résidu pouvant subir une réaction de *O*-GlcNAcylation par l'OGT ou de phosphorylation par GSK3 β . J'ai ainsi fabriqué tout un ensemble de vecteurs plasmidiques codant différentes isoformes de β -caténine: une forme *O*-GlcNAcylée "stable" en T41 grâce à une réaction de *S*-GlcNAcylation, une forme phosphomimétique (T41E), et une forme dite sauvage, non modifiée. Une bonne partie de mes travaux ont donc porté sur la construction de ces vecteurs plasmidiques, la production de protéines bactériennes (*E. coli* de type BL21 - DE3), l'incorporation d'acides aminés non canoniques (la phénylsélocystéine), la purification de protéines recombinantes et l'ajout du résidu *S*- β -GlcNAc à la position T41 par une réaction d'addition de thiol (addition de Michael). J'ai en seconde partie de thèse adapté des méthodes de transduction et de lipofection de ces protéines recombinantes dans les cellules cancéreuses HeLa et Hep3B, ainsi que dans les cellules normales du côlon CCD841CoN. Par la suite, j'ai utilisé la bioimagerie pour analyser la distribution et l'expression des différentes formes de β -caténine sur des cellules fixées. Ainsi, nous avons pu observer que la *S*-GlcNAcylation stabilise les protéines en les protégeant très probablement de la dégradation protéasomale par rapport aux formes sauvage (wt) et phosphomimétique. Par ailleurs, toute modification du site T41 de la β -caténine a impacté sur sa distribution subcellulaire pour des raisons que nous n'expliquons pas pour le moment.

En raison de la position stratégique de la β -caténine sur la face interne de la membrane plasmique où elle participe activement au contrôle des interactions homotypiques E-cadhérine/E-cadhérine responsables de l'intégrité des muqueuses, et dans le noyau où, en partenariat avec le TCF (T-cell factor), elle régule l'expression de nombreux gènes cibles impliqués entre autres dans le cycle cellulaire, nous pourrions modifier l'environnement cellulaire en inhibant ou en stimulant les cellules avec divers composés

chimiques. Ainsi, l'utilisation du LPA (acide lysophosphatidique) qui par perturbation des jonctions d'adhérence libère la β -caténine de la queue cytoplasmique de la E-cadhérine, permettrait de relarguer une large partie de la protéine dans le cytoplasme: la β -caténine serait alors libre de se déplacer vers le noyau. L'application de notre stratégie expérimentale permettrait d'octroyer à chacune des isoformes de la β -caténine (wt, S-GlcNAcylée ou phosphomimétique) une fonction bien spécifique dans ce processus fondamental à la transition épithélio-mésenchymateuse (TEM): doit-on s'attendre à ce que la forme S-GlcNAcylée et donc, par extension la forme naturelle O-GlcNAcylée, soit la forme actrice majoritaire de ce processus ? Par ailleurs, La leptomycine, et la lactacystine (ou le MG132), respectivement inhibiteurs de l'export nucléaire et du protéasome, devraient nous permettre d'observer une accumulation progressive d'analogues de la β -caténine dans le noyau et le cytoplasme respectivement. L'utilisation de tels composés permettrait de détecter les isoformes gT41 et pT41 de la β -caténine en des emplacements subcellulaires spécifiques et de suivre leur déplacement et leur accumulation cellulaire en fonction de leur statut de modification. En plus de ces composés chimiques, des nutriments tels que les acides gras et les acides aminés pourraient être testés dans le cadre de travaux ultérieurs.


La validation de cette nouvelle technologie basée sur l'expansion du code génétique permettrait d'ouvrir tout un champ d'investigation en chémobiologie et en bioimagerie. Cela pourrait inclure l'utilisation de techniques de microscopie avancées, telle que la microscopie à super-résolution, pour visualiser et quantifier les MPT de diverses protéines avec une résolution spatiale accrue.

Notre stratégie expérimentale est une nouvelle opportunité pour explorer les implications physiopathologiques et les conséquences fonctionnelles des MPT, telles que celles dévoilées sur la β -caténine dans divers modèles cellulaires et animaux pertinents dans le contexte de pathologies comme le cancer et les maladies métaboliques. Cette approche pourrait permettre de mieux comprendre le rôle de la β -caténine dans la pathogenèse de ces affections et ouvrir la voie au développement de nouvelles stratégies thérapeutiques.

D'autre part, cela offrirait des avancées notables et approfondies dans tout un ensemble de mécanismes moléculaires sous-jacents aux innombrables activités protéiques telles que celles faisant intervenir des facteurs de transcription, des composants de voies de signalisation et de voies métaboliques, ou même de suivre spécifiquement des variants pathologiques *versus* physiologiques d'acteurs majeurs de l'homéostasie cellulaire.

Annexes

A Step-by-Step Guide for the Production of Recombinant Fluorescent TAT-HA-Tagged Proteins and their Transduction into Mammalian Cells

Christer Abou Anny,¹ Sébastien Nouaille,² Régis Fauré,² Céline Schulz,¹ Corentin Spriet,^{1,3} Isabelle Huvent,¹ Christophe Biot,¹ and Tony Lefebvre^{1,4} 

¹Université de Lille, CNRS, UMR 8576 - UGSF, Lille, France

²TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

³Université de Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000, Lille, France

⁴Corresponding author: tony.lefebvre@univ-lille.fr

Published in the Molecular Biology section

Investigating the function of target proteins for functional prospection or therapeutic applications typically requires the production and purification of recombinant proteins. The fusion of these proteins with tag peptides and fluorescently derived proteins allows the monitoring of candidate proteins using SDS-PAGE coupled with western blotting and fluorescent microscopy, respectively. However, protein engineering poses a significant challenge for many researchers. In this protocol, we describe step-by-step the engineering of a recombinant protein with various tags: TAT-HA (trans-activator of transduction-hemagglutinin), 6×His and EGFP (enhanced green fluorescent protein) or mCherry. Fusion proteins are produced in *E. coli* BL21(DE3) cells and purified by immobilized metal affinity chromatography (IMAC) using a Ni-nitrilotriacetic acid (NTA) column. Then, tagged recombinant proteins are introduced into cultured animal cells by using the penetrating peptide TAT-HA. Here, we present a thorough protocol providing a detailed guide encompassing every critical step from plasmid DNA molecular assembly to protein expression and subsequent purification and outlines the conditions necessary for protein transduction technology into animal cells in a comprehensive manner. We believe that this protocol will be a valuable resource for researchers seeking an exhaustive, step-by-step guide for the successful production and purification of recombinant proteins and their entry by transduction within living cells. © 2024 The Authors. Current Protocols published by Wiley Periodicals LLC.

Basic Protocol 1: DNA cloning, molecular assembly strategies, and protein production

Basic Protocol 2: Protein purification

Basic Protocol 3: Protein transduction in mammalian cells

Keywords: DNA molecular assembly • fluorescent protein • fusion protein • Ni-NTA-based protein purification • protein transduction • recombinant protein production • tagged protein

How to cite this article:

Anny, C. A., Nouaille, S., Fauré, R., Schulz, C., Spriet, C., Huvent, I., Biot, C., & Lefebvre, T. (2024). A step-by-step guide for the production of recombinant fluorescent TAT-HA-Tagged proteins and their transduction into mammalian cells. *Current Protocols*, 4, e1016. doi: 10.1002/cpz1.1016

INTRODUCTION

Over the years, the interest in innovative approaches for the expression, purification, and characterization of recombinant proteins has increased. It empowers scientists to create, engineer, and produce proteins with diverse functions, characteristics, and applications (Assenberg et al., 2013). Thus, production of recombinant proteins represents a cornerstone for various applications in drug discovery (Andersen & Krummen, 2002; Overton, 2014), therapeutic proteins (Burnett & Burnett, 2020; Wurm, 2004) or vaccine production (Farzaneh et al., 2017; Joung et al., 2016). Recombinant DNA encoding proteins are typically generated through the use of recombinant DNA technology from historical methods (e.g., amplifications and use of restriction enzymes) to recently developed sophisticated technologies (e.g., Gibson assembly method for DNA cloning) (Bordat et al., 2015; Casini et al., 2015; Cohen, 2013; Ferro et al., 2019; Gibson et al., 2009; Jang & Magnuson, 2013; Kostylev et al., 2015; Li et al., 2019; Li et al., 2020; Li et al., 2018; Norris et al., 2015; Roberts, 2019; Rudenko & Barnes, 2018; Sands & Brent, 2016; Thomas et al., 2015; Wang et al., 2015). One of the most significant breakthroughs in recombinant protein technology has been the development of expression systems for the efficient production of complex proteins by selecting an appropriate method in microbial hosts (Brondyk, 2009), such as *Escherichia coli* (Baneyx, 1999; Baneyx & Mujacic, 2004; Chen, 2012; Gopal & Kumar, 2013; Rosano & Ceccarelli, 2014; Sørensen & Mortensen, 2005), yeast (Cregg et al., 2000; Mattanovich et al., 2012; Porro et al., 2005), *Streptomyces spp.* (Anné et al., 2012), mammalian cells (Andersen & Krummen, 2002; Barnes et al., 2003; Hopkins et al., 2012; Jäger et al., 2015; Sunley & Butler, 2010) or plants (Anné et al., 2012; Newman et al., 2011; Nienhaus & Nienhaus, 2014; Noguchi & Matsumoto, 2006). These systems provide researchers with the ability to produce proteins that may be challenging to obtain by fastidious protein purification from various biological sources (Brondyk, 2009). Among these recombinant proteins, fluorescent proteins have proven to be invaluable tools for monitoring and visualizing cellular processes in real-time (Ai et al., 2014; Nienhaus & Nienhaus, 2014; Wang et al., 2008; Wiedenmann et al., 2009). Fluorescent proteins are derived from naturally occurring proteins with intrinsic fluorescence, such as the Green Fluorescent Protein (GFP) originally isolated from *Aequorea victoria* (Jakobs et al., 2000; McRae et al., 2005; Nolte et al., 2001) or the red fluorescent protein mCherry, derived from DsRed isolated from *Discosoma* sea anemones (Shen et al., 2017). These color-differentiated fluorescent proteins can be used together to perform multicolor imaging experiments in molecular and cellular biology research in order to label specific cellular structures, organelles, or proteins for visualization and tracking within living cells (Helmuth et al., 2009; Kirchhausen, 2009; McDonald et al., 2002; Newman et al., 2011; Rizzuto et al., 1995; Salomonsson et al., 2012; Subach et al., 2009; Yin et al., 2013).

This protocol is designed to provide a step-by-step guide to successfully express recombinant fluorescent proteins in bacteria and subsequently visualize them in animal cells (Fig. 1). The successful execution of this protocol involves several steps, starting with the design and construction of recombinant DNA vectors using the Gibson assembly method to express the target protein fused with the HIV-1 trans-activator peptide called TAT-tag

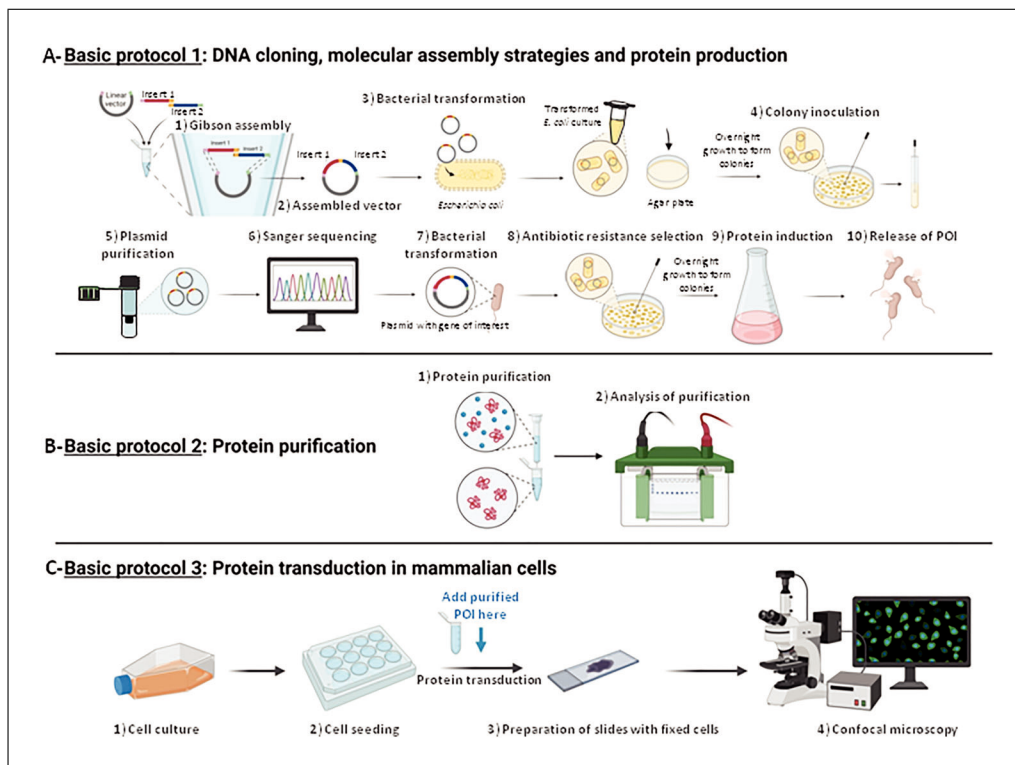


Figure 1 Schematic overview of protocols workflow. This Current Protocol article is divided into three distinct Basic Protocols aimed at cloning by molecular assembly a POI with various tags and producing it in *E. coli* cells (Basic Protocol 1, panel A), to purify this protein (Basic Protocol 2, panel B), and transducing it in eukaryotic animal cells for further investigations (Basic Protocol 3, panel C). POI, Protein of Interest.

(Trans-Activator of Transduction), a fluorescent protein, and a TEV protease recognition sequence to eliminate the downstream 6×His-tag. This method is used for the purpose of gene cloning and expression, through constructing plasmids for protein expression. These assembled vectors are then transformed into a suitable host organism, typically *E. coli* BL21(DE3), to produce recombinant fluorescent proteins. Recombinant protein production is carried out through transcription induction by L-arabinose thanks to the P_{BAD} bacterial promoter (Greenblatt & Schleif, 1971; Schleif, 2000; Schleif et al., 1973). The protein purification is performed via immobilized metal affinity chromatography (IMAC) thanks to the 6×His-tag (Block et al., 2009; Gaberc-Porekar & Menart, 2001; Porath, 1992; Sun et al., 2005). The TAT-tag, due to its transduction property, enables the fused protein to enter animal cells (Becker-Hapak et al., 2001; Eguchi et al., 2001; Fittipaldi & Giacca, 2005; Ford et al., 2001; Gump & Dowdy, 2007; Gump et al., 2010; Guo et al., 2012; Kabouridis, 2003; Kaplan et al., 2005; Leifert et al., 2002; MacKay & Szoka, 2003; Noguchi & Matsumoto, 2006; Pawson & Nash, 2000; Schwarze et al., 2000; Shokolenko et al., 2005; Wadia & Dowdy, 2002; Zhang et al., 2012). Once inside the living cells, the recombinant protein can be monitored and tracked according to its localization and behavior by its fluorescent signal.

The production of proteins in bacteria and their delivery to animal cells by a penetration peptide has the advantage of working with well-defined quantities of proteins, obtaining a homogeneous distribution of the latter throughout the cell population and to be a vector-free method. The transient transfection of vectors, e.g., plasmids, makes it impossible to control the quantity of protein expressed within the cell and very often leads to heterogeneous expression within the cell population, as well as being cell type dependent.

We offer the reader, with a range of readily accessible techniques and devices found in research laboratories or with a simple Ni-NTA column, the opportunity to produce

fluorescent proteins using three distinct but complementary protocols. Basic Protocol 1 details, thanks to Gibson-type molecular assembly, the construction of custom plasmid vectors making it possible to express a specific protein of interest (POI) in phase with any fluorescent protein and various tags, and to produce it in *E. coli* bacteria. Basic Protocol 2 presents the purification of the POI by taking advantage of a poly-His-tag by using a preparative chromatography system with a fast protein liquid chromatography (FPLC) Ni-nitrilotriacetic acid (NTA) agarose column, and its quality control by SDS-PAGE coupled to western blotting. Lastly, Basic Protocol 3 exposes a very simple experimental strategy for protein transduction into animal cells using a cell penetration peptide. We use direct fluorescence microscopy to avoid long, complex, and often costly indirect immunofluorescence steps.

BASIC PROTOCOL 1

DNA CLONING, MOLECULAR ASSEMBLY STRATEGIES, AND PROTEIN PRODUCTION

In this protocol, we describe a procedure to construct a specific DNA sequence encompassing the different blocks needed for expression of the final protein containing the protein of interest (POI) moiety (Fig. 1A and Fig. 2). For that purpose, the choice of the gene

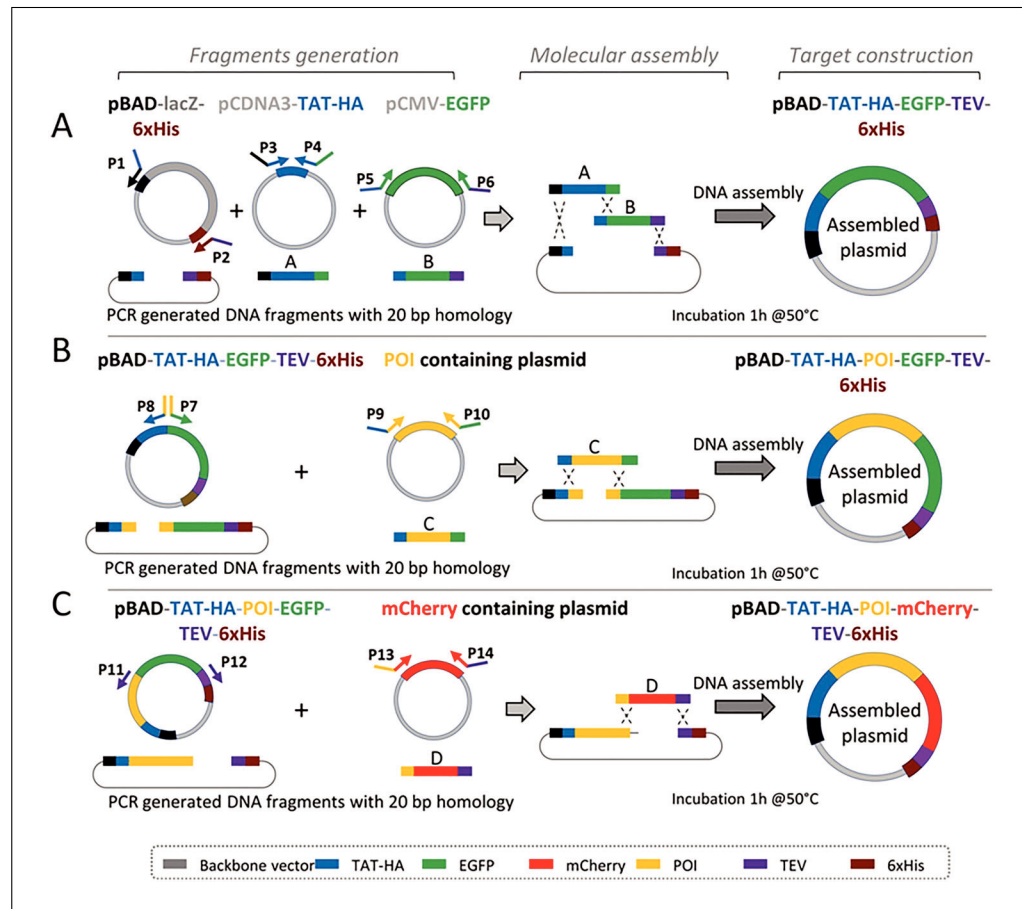


Figure 2 DNA assembly towards the overall plasmid constructions. **(A)** Use of P1-P6 primers to generate the pBAD-TAT-HA-EGFP-TEV-6×His final construction. **(B)** Use of P8-P10 primers to generate the pBAD-TAT-HA-POI-EGFP-TEV-6×His final construction. **(C)** Use of P11-P14 primers to generate the pBAD-TAT-HA-POI-mCherry-TEV-6×His final construction. P1-P14 represent the different primers listed in Table 3 needed to generate the different PCR products that are then PCR-amplified to step-by-step reach the final constructions. For more details see Tables 1–7. The color code used is the same as in Table 3. EGFP, Enhanced Green Fluorescent Protein, LacZ, β-galactosidase encoding gene; Ori, Origin of replication; pBAD, promoter of *araB*, *araA* and *araD*; POI, Protein Of Interest; TAT-HA, Trans-Activator of Transduction-HemAgglutinin TEV, recognition sequence by the Tobacco Etch Virus protease; vCMV, Cytomegalovirus vector.

Table 1 Characteristics of the PCR Components

Plasmid DNA source	Primer pair	Amplicon size (bp)
pBAD-LacZ-X-6×His	P1-P2	4047
pCDNA3-Y-TAT-HA	P3-P4	122
vCMV-EGFP-POI	P5-P6	745
pBAD-TAT-HA-EGFP-TEV-6×His	P7-P8	4873
vCMV-EGFP-POI	P9-P10	2360
pBAD-TAT-HA-POI-EGFP-TEV-6×His	P11-P12	6479
pCMV-mCherry	P13-P14	745

Table 2 Bacterial Strains and Plasmids Used in this Study

Strain or Plasmid	Characteristics	Source
Strain		
Competent <i>E. coli</i> (High Efficiency)	DH5-alpha	New England BioLabs
Competent <i>E. coli</i> (High Efficiency)	BL21(DE3)	New England BioLabs
Plasmid		
pBAD-LacZ-6×His	<i>E. coli</i> expression vector, amp ^r	Life Science
pCDNA3-Y-TAT-HA	<i>E. coli</i> expression vector, amp ^r	UGSF, Lille (France)
vCMV _p -EGFP-POI	<i>E. coli</i> expression vector, kan ^r	e-Zyvec, Loos (France)
pBAD-TAT-HA-EGFP-TEV-6×His	<i>E. coli</i> expression vector, amp ^r	This study
pBAD-TAT-HA-POI-EGFP-TEV-6×His	<i>E. coli</i> expression vector, amp ^r	This study
pCMV-mCherry	<i>E. coli</i> expression vector, kan ^r	UGSF, Lille (France)

encoding the POI, the plasmid vector and its features highly depend on specific needs of the experiment, including the type of host organism (e.g., *E. coli*) and the size of the DNA to be inserted. For the expression in *E. coli*, the desired level of gene expression is dependent on the choice of the appropriate promoter (e.g., P_{BAD} promoter, used here). These choices greatly influence the success of gene expression studies (for detailed information see Tables 1–3). The plasmids undergo construction through Gibson molecular assembly of purified PCR fragments. The advantage with this approach is that it is possible, with a whole set of constructions either available in the laboratory or obtained elsewhere, to recover a fragment of interest (comparable to a block or a brick from a game assembly) and nesting it with other fragments to custom manufactured plasmids expressing the POI in phase with any desired tag or fluorescent protein.

Materials

- Plasmids, or other source of DNA insert (see Table 2 for detailed information)
- Custom primers (see Table 3 for primers used in this study)
- Q5 high-fidelity DNA polymerase, 2000 U/ml (New England BioLabs, cat. no. M0491S)
- Q5 reaction buffer, 5× (New England BioLabs, cat. no. B9027S)
- Deoxynucleotide (dNTP) solution mix, 10 mM (New England BioLabs, cat. no. N0447S)
- Ultrapure DNase-free H₂O
- Agarose powder
- Tris-acetate-EDTA (TAE) buffer, 0.5× (Euromedex)
- Gel loading buffer (New England Biolabs, cat. no. B7025)

Table 3 Primers Used in this Study^a

Primer	Sequence (5' → 3')	Construct
P1-TEV-6×His	GAAAACCTGTACTTCCAGGGTAATAGCGCCGTCGACCATCAT	pBAD-TAT-
P2-TAT-HA-pBAD	AGCCAAGCTTTGGATCCATGGTTAATTCCTCCTGT	HA-EGFP-
P3-pBAD-TAT-HA	CATGGATCCA AAGCTTGGCTACGGCCGCAAGAAA	TEV-6×His
P4-EGFP-TAT-HA	CCTTGCTCACGGCCATGGAGCCAGCATAGTCTG	
P5-EGFP-TAT-HA	CTCCATGGCCGTGAGCAAGGGCGAGGAGCTGTTC	
P6-TEV-EGFP	ACCCTGGAAGTACAGGTTTTCCTTGACAGCTCGTCCATGCCGAG	
P7-POI-EGFP	NNNNNNNNNNGTGAGCAAGGGCGAGGAGCTGTTC	pBAD-TAT-
P8-POI-TAT-HA	NNNNNNNNNNGGCCATGGAGCCAGCATAGTCTG	HA-POI-
P9-TAT-HA-POI	CTCCATGGCCNNNNNNNNNNNNNNNNNNNNNNNN	EGFP-TEV-
P10-EGFP-POI	CCTTGCTCACNNNNNNNNNNNNNNNNNNNNNNNN	6×His
hP11-POI Rv	CAGGTCAGTATCAAACCAGGCCAGC	pBAD-TAT-
P12-5'TEV	GAAAACCTGTACTTCCAGGGTAATAGCG	HA-POI-
P13-mcherry	CCTGGTTTGATACTGACCTGGTGAGCAAGGGCGAGGAGGATAACAT	mcherry-TEV-
P14-mcherry	CCCTGGAAGTACAGGTTTTCCTTGACAGCTCGTCCATGCCGCC	6×His

^aFor clarity the color code used for sequences is the same as that used in Figure 2 for fragments.

1 kb Plus DNA ladder, 1000 µg/ml (New England BioLabs, cat. no. N3200S)
 100 bp DNA ladder, 500 µg/ml (New England BioLabs, cat. no. N3231S)
 Ethidium bromide (Thermo Fisher Scientific, cat. no. 15585011)
 QIAquick gel extraction kit (Qiagen, cat. no. 28706)
 NEBuilder HiFi DNA assembly master mix (New England BioLabs, cat. no. E2621S)
 Competent 5-alpha *E. coli* (New England BioLabs, cat. no. C2987H)
 SOC medium (New England BioLabs, cat. no. B9020S)
 LB agar plates: LB medium plus 15 g/L agar
 LB medium: 10 g/L tryptone, 10 g/L NaCl, and 5 g/L yeast extract supplemented with the appropriate antibiotics, e.g., 100 mg/ml ampicillin (stock solution) diluted at 1/1000
 QIAprep spin miniprep kit (Qiagen, cat. no. 27104)
 Competent BL21(DE3) *E. coli* (New England BioLabs, cat. no. C2527H)
 L-arabinose (Sigma-Aldrich, cat. no. A3256)
 Complete EDTA-free protease inhibitor cocktail tablets (Sigma-Aldrich, cat. no. 11873580001)
 Buffer A (see recipe)
 DNase I (3000 U/mg) 20,000 U (Euromedex, cat. no. 1307)

0.2-ml PCR tubes (Dutsher, cat. no. 010207)
 Benchtop centrifuge
 Thermal cycler (Eppendorf Mastercycler)
 Microwave oven or hotplate
 Agarose gel electrophoresis system
 E-Box gel documentation imaging system (Vilber Lourmat, France)
 Benchtop UV transilluminator machine
 Scalpel
 NanoDrop ND-1000 spectrophotometer
 Petri dishes (TRP, cat. no. 067002)
 1.5-ml microcentrifuge tubes (Thermo Fisher Scientific, cat. no. 69715)
 Ice
 Heat block (Eppendorf ThermoMixer) or water bath
 Shaking incubator
 Sterile toothpicks

Table 4 PCR Components and Procedure for Q5 High-Fidelity DNA Polymerase

Composition of PCR mixture			
Component	Stock solution	Volume (μ l)	Final concentration
Q5 reaction buffer	5 \times	5 μ l	1 \times
dNTP mix	10 mM each	0.5 μ l	0.2 mM each
Primers (10 μ M)	10 μ M	1.25 μ l for each primer	0.5 μ M each
Plasmid matrix	Variable	Variable	\sim 4 pg/ μ l
Q5 polymerase	2 U/ μ l	0.3 μ l	0.024 U/ μ l
Ultrapure water		To 25 μ l	
PCR Thermocycling conditions			
Step	Temperature ($^{\circ}$ C)	Time	Cycle
Denaturation	98	30 s	1
Denaturation	98	10 s	
Annealing	60	20 s	\times 30
Elongation	72	30 s/kb	
Elongation	72	3 min	1
Cooling	14	∞	-

1-L glass Erlenmeyer flasks

Centrifuge Beckman-Coulter Avanti J-26S XP (with a rotor adapted for 1-L volumes)

50-ml sterile conical tubes (Falcon, cat. no. 352070)

Sonicator

Centrifuge, rotor 19776H, angle rotor, for six 50-ml culture tubes

0.45- μ m filter (Dutscher, cat. no. 033942)

CAUTION: Ethidium bromide is a putative carcinogen compound (the dye intercalates between the stacked bases of DNA fragments) and must be handled with gloves. Do not use ethidium bromide powder to avoid irritation by inhalation and never use a microwave oven to heat solutions containing ethidium bromide. It is preferable to use the less toxic SYBR safe stain gel to visualize DNA fragments in agarose gels. When using the transilluminator to visualize DNA bands, it is essential to protect yourself from aggressive UV rays with plastic glasses, gloves and sleeves that must cover the integrity of the arms.

Amplification of the various PCR products

1. In PCR tubes, mix the Q5 reaction buffer (adjusted with DNase-free ultrapure water) with the DNA templates, primers (designed as P1 to P14 in Table 1 and Fig. 2), dNTP and the Q5 high-fidelity DNA polymerase at desired concentrations (Table 4). Mix and pulse-centrifuge to ensure the total liquid is located at the bottom of the tube.

Primers P1 and P2 serve to recover the P_{BAD} -LacZ-6 \times His part; primers P3 and P4, the TAT-HA penetrating peptide encoding part; primers P5 and P6, the EGFP encoding part; primers P7 and P8 serve to recover the p_{BAD} -TAT-HA-EGFP-TEV-6 \times His by adding 15 to 20 bp of the POI between the TAT-HA and EGFP sequences; primers P9 and P10, the POI encoding part; primers P11 and P12 serve to recover the p_{BAD} -TAT-HA-POI-TEV-6 \times His by excluding the EGFP part; and primers P13 and P14 serve to recover the mCherry encoding part (Fig. 2).

2. Place the PCR tubes in a thermal cycler and apply the corresponding PCR amplification program (Table 4).

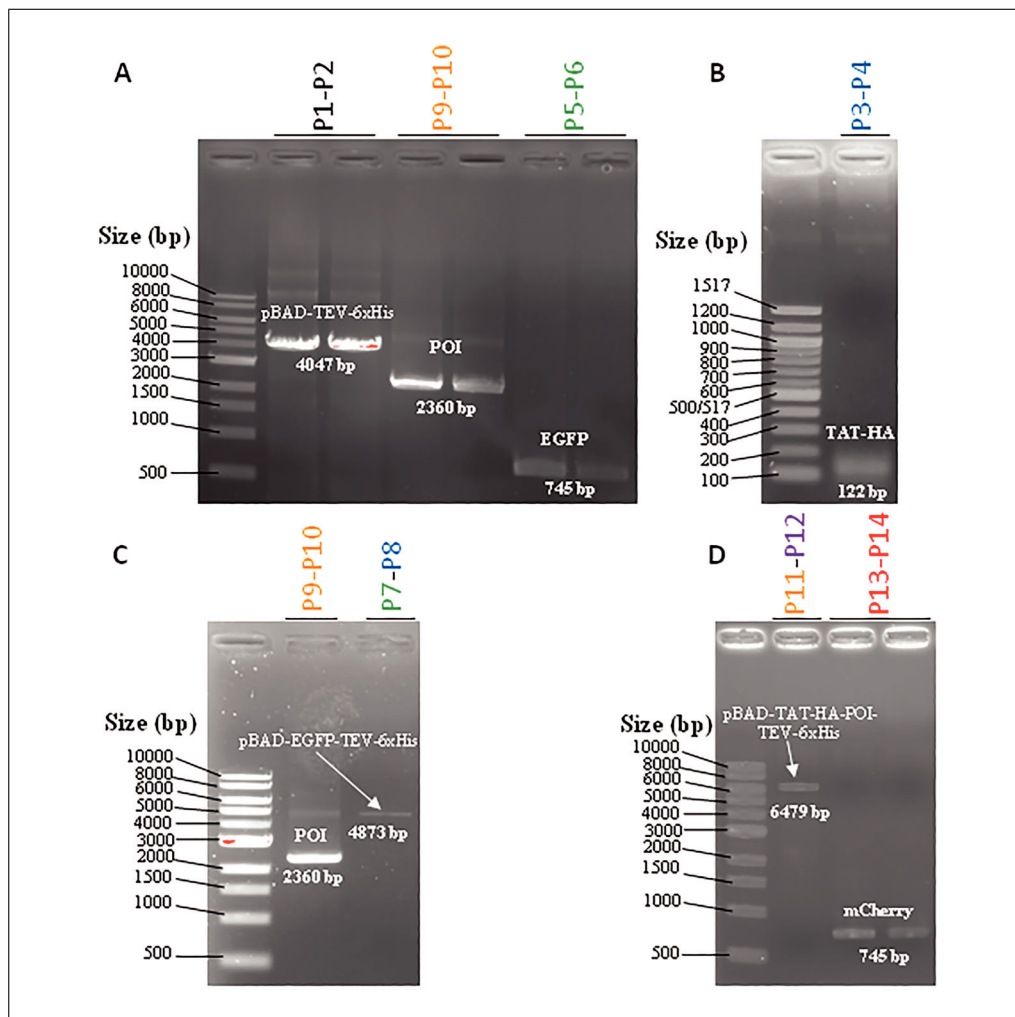


Figure 3 PCR electrophoresis gels. The PCR products of the different fragments shown schematically in Figure 2 were analyzed by electrophoresis on agarose gel. At the top of each gel are indicated the primer pairs following the same color code that used in Figure 2. **(A)** pBAD-TEV-6×His, POI and EGFP fragments. **(B)** TAT-HA fragment. **(C)** POI and pBAD-EGFP-TEV-6×His fragments. **(D)** pBAD-TAT-HA-POI-TEV-6×His and mCherry fragments. The size markers are indicated at the left of each gel: 1 kb Plus DNA ladder for panels **(A)**, **(C)** and **(D)**; 100 bp ladder for panel **(B)**.

3. Prepare the gel by mixing agarose at a concentration of 1% (w/v) in 0.5× TAE. Heat the solution in a microwave oven or on a hotplate until the agarose has dissolved.
4. Pour the gel in the electrophoresis casting tray and let it solidify. Remove the comb and put the gel in the electrophoresis chamber containing 0.5× TAE as migration buffer.
5. Mix the 5 µl PCR reaction to 2 µl gel loading buffer, load on the gel along with 5 µl DNA size marker to monitor DNA size after migration. Migrate at 100 V for 30 min.
6. After migration, incubate the gel for 10 min in water containing 0.5 µg/ml ethidium bromide. Wash the excess of ethidium bromide by incubation in water for 5 min.
7. Visualize the PCR products (Fig. 3) by exposition of the gel to a UV transilluminator, such as the E-box gel documentation imaging system.
8. To purify the amplicon of interest, cut out the UV-detected DNA bands using a scalpel blade under a UV transilluminator instrument.

Table 5 Plasmid DNA Assembly Conditions

Components	Stock solution	Volume	Final concentration
Purified PCR	Variable	Variable	5-15 ng/ μ l
T4 polynucleotide kinase PNK	10 U/ μ l	1 μ l	0.5 U/ μ l
T4 DNA ligase Buffer with ATP	10 \times with 10 mM	1 μ l	0.5 \times with 0.5 mM
T4 DNA ligase	400 U/ μ l	1 μ l	20 U/ μ l
Ultrapure water		To 20 μ l	

Table 6 Gibson Method Plasmid DNA Assembly

Component	Volume (μ l)
Purified PCR fragments	x μ l
Enzymes, 2 \times Hi-Fi DNA assembly master mix	5 μ l
Water	Up to 10 μ l
DNA assembly thermocycling conditions	
50°C for 1 hr	
14°C/ ∞	

9. Weigh the DNA containing the gel slice and add 3 volumes (v/w) of QG buffer. Purify by using QIAquick gel extraction kit following the manufacturer's instructions. Elute with 50 μ l elution buffer. Quantify the DNA using a NanoDrop ND-1000 spectrophotometer.

DNA molecular assembly

10. Mix each purified PCR fragment, i.e, TAT-HA, POI, pBAD-TEV-6 \times His, and EGFP or mCherry for the construction of pBAD-TAT-HA-POI-EGFP-TEV-6 \times His and pBAD-TAT-HA-POI-mCherry-TEV-6 \times His, respectively (Fig. 2), with NEBuilder HiFi DNA assembly 2 \times master mix at a ratio (1:1) (Tables 5 and 6). Complement with water to a final volume of 10 μ l. Incubate the mixture at 50°C for 1 hr in a thermocycler.
11. Use 5 μ l assembly reaction for bacterial transformation in *E. coli* 5-alpha chemically competent cells (Froger & Hall, 2007). Mix DNA to 50 μ l cells in 1.5-ml microcentrifuge tubes and mix gently by pipetting up and down or flicking the tube 4 to 5 times (do not vortex).
12. After 20 min on ice, perform heat shock using a heat block or water bath set to 42°C for 30 s and immediately transfer cells on ice for 3 min.
13. Add 450 μ l SOC medium and incubate the mixture under shaking (150 rpm) for 1 hr at 37°C.
14. Spread 10 and 100 μ l on LB agar plates containing the appropriate antibiotic and incubate at 37°C overnight.
15. Pick single colonies using toothpicks and cultivate to saturation in 5 ml LB medium containing antibiotic (37°C, 150 rpm) overnight.
16. Extract plasmid DNA using a QIAprep spin miniprep kit following the manufacturer's protocols. Elute the plasmid DNA in 50 μ l elution buffer and quantify with a NanoDrop ND-1000 spectrophotometer. Check and confirm the fidelity of the construction by DNA sequencing (Eurofins).

Table 7 Amino Acid Composition of the Recombinant Protein Isoforms^a

Recombinant protein	Profiling amino acid sequence
TAT-HA- POI-EGFP-TEV-6×His	MDPKLGYGRKKRRQRRRGGSTMSGYPYDVPDYAGSMA[POI]VSKGEELFTGVVP ILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKICTTGKLPVWPVTLVTTLTLYGVQ CFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIE LKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGVSVQLA DHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITLGMD ELYKENLYFQGN SAVD HHHHHH
TAT-HA- POI-mCherry-TEV-6×His	MDPKLGYGRKKRRQRRRGGSTMSGYPYDVPDYAGSMA[POI]VSKGEEDNMAIHK EFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQ FMYGSKAYVKHPADIPDYLLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFI YKVKLRGTNFPDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLDGGH YDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMD ELYKENLYFQGN SAVD HHHHHH

^aFor clarity the color code used for sequences is the same as that used in Figure 2 for fragments.

Transformation and selection procedure

17. Perform the transformation with the plasmid expressing the POI by using the heat shock method described in step 11. Thaw *E. coli* BL21(DE3) competent cells on ice.
18. Add 0.1 µg purified DNA plasmid to 50 µl competent cells and mix gently by pipetting up and down or flicking the tube 4 to 5 times (do not vortex).
19. After 20 min on ice, perform heat shock at 42°C for 45 s and immediately transfer cells on ice for 3 min.
20. Add 450 µl SOC medium and incubate the mixture under shaking (150 rpm) for 1 hr at 37°C.
21. Spread 10 and 100 µl of the transformation mixture on separate LB agar plates (Petri dishes) supplemented with the antibiotic.
22. Incubate overnight at 37°C.

Recombinant protein expression

23. Pick single colonies and cultivate to saturation in 5 ml LB containing antibiotic (37°C, 150 rpm) overnight.
24. Add pre-culture of transformed *E. coli* BL21(DE3) to 1 L of LB medium in 1-L glass Erlenmeyer flasks containing the antibiotic to reach a starting optical density at 600 nm (OD₆₀₀) of 0.1. Monitor the OD₆₀₀ until 0.6; then induce the recombinant POI expression (Table 7) for 4 h by adding 0.1% (w/v) L-arabinose (final concentration).
25. To ensure that the conditions for bacterial growth are optimal for the expression of the recombinant protein, monitor the bacterial growth by measuring the OD₆₀₀ during the culture and at the end of the expression phase (4 hr post-induction) (Fig. 4).
26. For cell harvesting, centrifuge the 1 L of LB culture medium 30 min at 6000 × g, 4°C, using the Avanti J-26S XP centrifuge (Fig. 5A).
27. Discard the supernatant LB and add an anti-protease complete EDTA-free protease inhibitor cocktail tablet to the bacterial pellet (one tablet per L culture medium).
28. Resuspend the bacterial pellet in 20 ml buffer A and store the 50-ml tube at −20°C overnight.

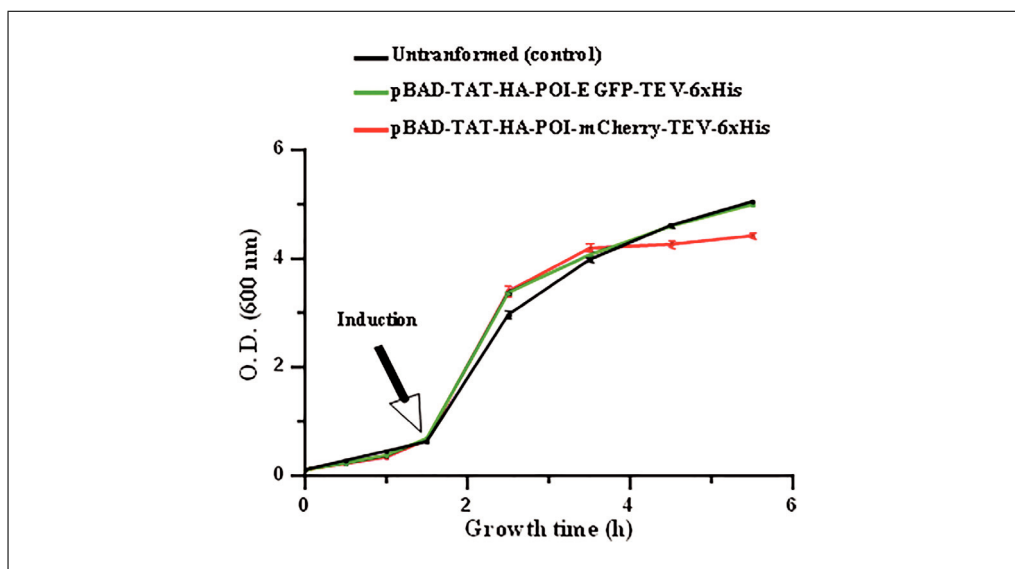


Figure 4 Growth profiles of empty *E. coli* BL21(DE3) and *E. coli* transformed with the pBAD-TAT-HA-POI-EGFP-TEV-6×His or pBAD-TAT-HA-POI-mCherry-TEV-6×His plasmids measured as function of time at OD₆₀₀ (induction of POI expression by L-arabinose addition is indicated by the arrow) (n = 3). POI, Protein of Interest.

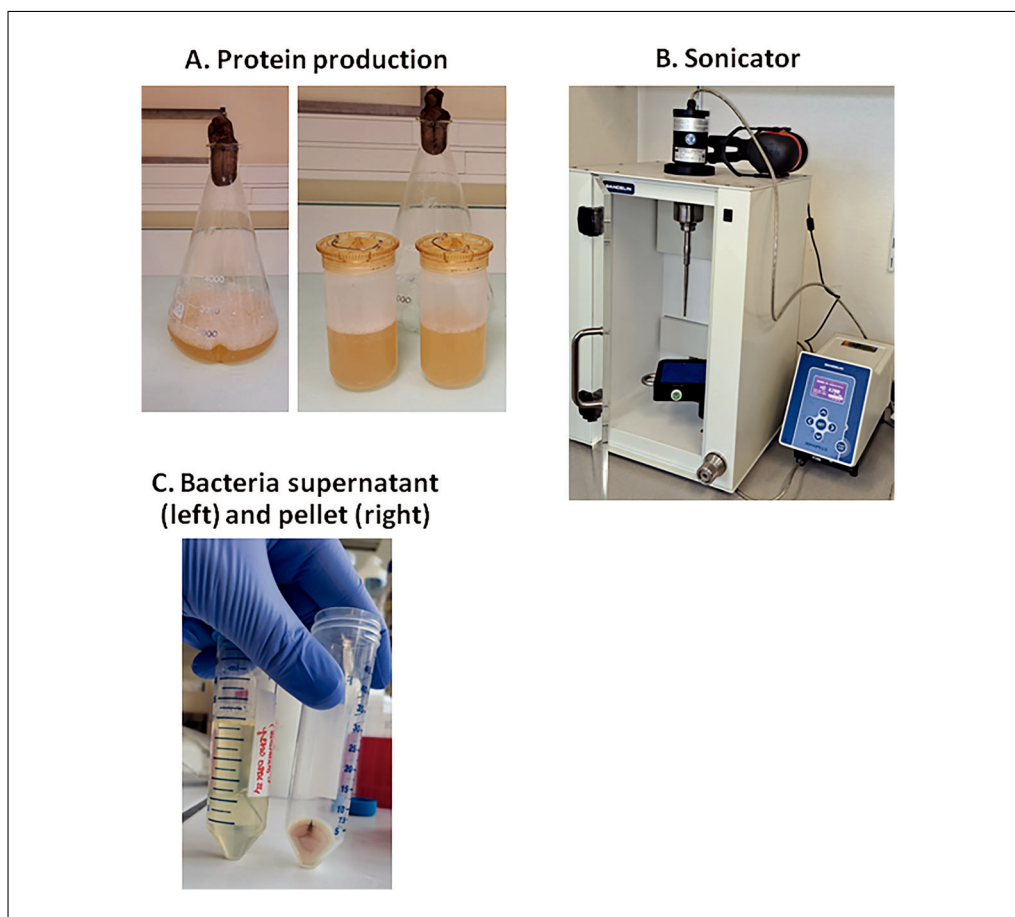


Figure 5 Efficient cell lysis for protein extraction by sonication. **(A)** LB culture medium containing the *E. coli* BL21 cells after induction with L-arabinose (left) to promote protein production. Distribution of the culture medium in two centrifuge bottles before centrifugation (right) (Basic Protocol 1, step 26). **(B)** Optimizing protein extraction with the Bandelin Electronic Sonicator HD 4200 in cell lysis protocols (Basic Protocol 1, step 31). **(C)** Recovery of the bacterial pellet (right) and supernatant containing the POI (left) after centrifugation of bacteria following sonication (Basic Protocol 1, step 32). POI, Protein of Interest.

29. Thaw the frozen bacterial pellet by placing the 50-ml tube containing the bacterial pellet in a water bath at room temperature.
30. Add 100 μ l DNase I for every 1 L bacterial culture to degrade DNA that will be released by the bacterial lysis.
31. Perform bacterial lysis by sonication on ice, employing cycles of 30 s ON and 15 s OFF during 5 min, with the probe operating at 70% of maximal power (Fig. 5B).
32. Using a centrifuge and rotor 19776H, centrifuge the lysate 15 min at 12,000 \times g, 4°C (Fig. 5C).
33. Filter the supernatant through a 0.45- μ m filter. The supernatant containing the POI is ready for protein purification (Basic Protocol 2).

BASIC PROTOCOL 2

PROTEIN PURIFICATION

In this protocol, we describe the procedure of production and purification of recombinant protein isoforms (Fig. 1B). The remaining supernatant (from Basic Protocol 1, step 33) is then subjected to IMAC thanks to the 6 \times His-tag. Subsequent wash steps remove impurities and then the POI is eluted using a buffer that disrupts the binding interaction. Further purification steps may include size-exclusion chromatography and/or ion-exchange chromatography to achieve higher purity. Throughout the protocol protein concentration is monitored and quality is assessed using SDS-PAGE and a brilliant Coomassie blue staining and by western blot. If conducted properly, the final purified POI should exhibit the expected molecular weight, high purity, and biological activity.

Materials

- Buffer A (see recipe)
- Buffer B (see recipe)
- Ultrapure H₂O
- Desalting buffer: Dulbecco's phosphate-buffered saline (DPBS), w/o Mg, w/o Ca (Dutscher, cat. no. L0615-500)
- Laemmli buffer, 5 \times (see recipe)
- Polyacrylamide gels, 8% separating and 4% stacking (see recipes)
- SDS-PAGE buffer: 10% (v/v) Tris-glycine-SDS 10 \times (Euromedex, cat. no. EU0510) in distilled H₂O
- PageRuler prestained protein ladder (Thermo Fisher Scientific, cat. no. SM0671)
- Quick Coomassie stain (NeoBiotech, cat. no. NB-45-00078-1litre)
- Transfer buffer: 10% (v/v) Tris-glycine 10 \times (Euromedex, cat. no. EU0550) in distilled water, 20% (v/v) methanol (Carlo Erba, cat. no. 524102)
- Ponceau S red solution (Sigma-Aldrich, cat. no. P7170)
- TBS-Tween buffer (see recipe)
- Blocking buffer (see recipe)
- Mouse monoclonal antibodies:
 - Anti-HA-tag (F-7) (Santa Cruz Biotechnology, cat. no. sc-7392)
 - His-Tag (HIS. H8 / EH158) (Covalab, cat. no. mab90001-P)
 - Anti-POI
 - Anti-EGFP (Roche, cat. no. 45-11814460001)
- ECL anti-mouse IgG (from sheep), horseradish peroxidase-linked whole secondary antibodies (Fisher Scientific, cat. no. NA931V)
- SuperSignal West Pico PLUS (Thermo Scientific, cat. no.34577) or SuperSignal West Femto, maximum sensitivity substrate (Thermo Scientific, cat. no.34095).
- Preparative chromatography system (e.g., ÄKTA start protein purification system associated with UNICORN start control software)



Figure 6 Overview of the protein purification and desalting process. The protein sample containing the POI is injected using an automated injection system (indicated by the red arrowhead) and shown in the inset above-right (sample loading position). The bottles for buffer A (lysis buffer) and buffer B (elution buffer) are in the central polystyrene box. Fractions are collected using an automatic collector at the right of the picture. The process is monitored by measuring absorbance at a specific wavelength of 280 nm thanks to a computer at the left. The device is shown in its desalting configuration: the HiTrap Desalting is indicated by the blue arrowhead.

HisTrap fast flow, 1- or 5-ml, Cytiva (Dutscher, cat. no. 17-5255-01)
 HiTrap desalting, 5-ml, Cytiva (Dutscher, cat. no. 17-1408-01)
 Vivaspin 6, centrifugal concentrator, 50,000 MWCO PES (Grosseron, cat. no. VS0632)
 Benchtop centrifuge
 Heat block (Eppendorf ThermoMixer)
 Buffer tank lid
 Electrode assembly
 Combs electrophoresis Cell, 10-well, 1.5-mm
 Hamilton syringe, 50- μ l with needles
 Power supply generator
 Rotating platform
 Gel cassette
 Gel holder cassette and foam pads
 0.45- μ m pore nitrocellulose membrane (Amersham, cat. no. 10600016)
 Whatman paper sheets, 6 \times 9.5 cm, and sponges
 Electroblotting system
 Fusion Solo machine (Vilber Lourmat) for western blot detection

NOTE: It is possible that the POI aggregates and forms inclusion bodies that are difficult to solubilize and extract from the bacteria. In that case we recommend to process via mechanical rupture of the bacteria by sonication or by using a French press.

Protein fractionation

To purify your recombinant TAT-HA-POI-EGFP/mCherry-TEV-6 \times His protein by IMAC (exploiting the presence of a C-terminal 6 \times His-tag), use a simple benchtop nickel affinity column or combine with a preparative chromatography system, such as the ÄKTA Start system associated with a ÄKTA Start Fraction Collector “Frac30” and a UNICORN Start Software (Fig. 6). This can be carried out by employing an automated protein purification workflow with features like automated sample injection, fractions collection, and real-time monitoring caution.

1. Prepare bottles for buffer A, buffer B, ultrapure water, and the column HisTrap fast flow.
2. Set the IMAC purification program to use both buffer A and buffer B.
3. Equilibrate the column with 20 to 35 ml buffer A.
This step involves filling the column and allowing the buffer to flow through until it is stable and free of air bubbles.
4. Inject your protein sample containing the POI into the equilibrated column (usually done using an automated injection system).
5. Wash the column with 25 ml buffer A to remove unbound proteins and contaminants.
6. Start the imidazole concentration gradient between buffer A and buffer B, up to 100% buffer B, to ensure the controlled release of the POI from the column.
7. Collect 1 ml fractions from the column. Monitor the elution process in real-time by measuring absorbance at a specific wavelength of 280 nm (detection of proteins).

Protein desalting and concentration

8. Use a bottle of 1 × DPBS and a HiTrap desalting column.
9. Set the protein desalting program to only use pump A.
10. In the injection loop, re-inject the fraction containing your purified POI -1 ml per 1 ml and re-collect 5 to 10 desalted fractions of 1 ml each.
11. Concentrate the proteins (sample of 4 ml POI) using a Vivaspin 50 kDa cut-off through successive centrifugations of 5 min at $3500 \times g$, 4°C (Fig. 7A). Prior to use, rinse the concentrator column with water and subsequently with DPBS. Gently mix with a pipette before and after each centrifugation.

Take care not to exceed the specified speed indicated on the centrifugal filters.

12. Quantify protein concentration by measuring absorbance at 280 nm (extinction coefficients of 93,170 and 105,660 $M^{-1}cm^{-1}$ for the EGFP and the mCherry forms of the POI, respectively).

Analysis of recombinant protein expression by SDS-PAGE (Fig. 7B)

13. Load 10 to 20 μg of purified proteins (mixed and boiled 1/5 volume of 5 × Laemmli buffer) by using a Hamilton syringe on an 8% SDS-PAGE gel. Also load 5 μl molecular weight markers (typically ranging from 250 to 10 kDa).

The reticulation of the separating gel must be adjusted according to the molecular weight of the POI. Use a high reticulation for a low molecular weight protein (e.g., 12.5% reticulation for a 40 to 50 kDa protein) and a lower reticulation for a higher molecular weight protein (e.g., 8% reticulation for a 100 to 120 kDa protein).

14. Using a power supply generator, run at 30 mA in 1 × Tris-glycine-SDS buffer until the bromophenol blue line reaches the bottom of the glass plates (~2 hr).
15. For Coomassie blue staining (Figs. 7C and 7D), after migration carefully transfer the gel to a plastic or glass tray.
16. Cover the gel with Quick Coomassie stain with enough volume to allow it to move freely in the tray.
17. Cover with aluminum wrap and place on a platform rotator overnight at room temperature.

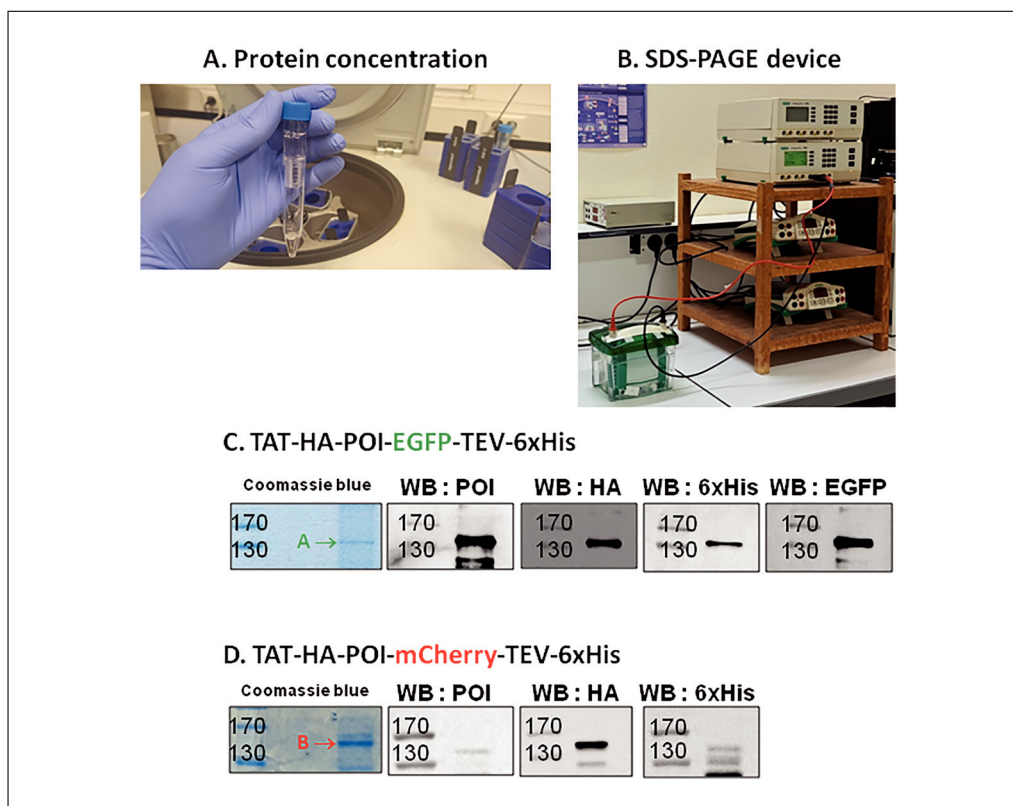


Figure 7 Protein concentration and characterization by SDS-PAGE. **(A)** Concentration of the POI by using a Vivaspin 6 type centrifugal concentrator. **(B)** Characterization of the recombinant POI expression by using an SDS-PAGE device. **(C and D)** Assessed of POI production and purity by Coomassie blue staining and western blotting. Sample A corresponds to the TAT-HA-POI-EGFP-TEV-6×His protein and sample B to the TAT-HA-POI-mCherry-TEV-6×His protein. Molecular weight markers are indicated on the left part of each panel (in kDa).

18. Proceed to the gel destaining step by multiple washes with ultrapure water until the band corresponding to the POI is easily visible.
19. For western blotting (Figs. 7C and 7D), transfer the gel onto nitrocellulose and proceed to electro-blotting in a device dedicated for that purpose (e.g., BioRad blotting system).
20. After transfer, check the transfer efficiency (and if necessary, the equal loading of the samples) by incubating the nitrocellulose membrane in the Ponceau S red solution for 1 min.
21. Discard the excess of solution and wash the membrane with ultrapure water until the band(s) is(are) easily visible.
22. Take a picture of the stained membrane for archiving.
23. Destain the membrane by incubation in the TBS-Tween buffer for a few seconds until gentle shaking.
24. Saturate the membrane by incubation in the blocking buffer for 45 min at room temperature (use a rotating platform for gentle shaking).
25. Discard the blocking buffer and incubate the membrane with the appropriate antibodies (mouse polyclonal anti-HA, mouse monoclonal anti-6×His, mouse monoclonal anti-POI at the appropriate dilution according to the manufacturer's instructions in blocking buffer).
26. Place the membrane on a rotating platform (gentle shaking) overnight at 4°C.

27. Wash the membranes with TBS-Tween 3 times for 7 min each and then incubate with an anti-mouse horseradish peroxidase-labeled secondary antibody (usually at a dilution of 1/10,000) for 1 hr.
28. Wash the membranes with TBS-Tween 3 times for 7 min each.
29. Add a mixed volume of buffers A and B (1:1) of SuperSignal West Pico Plus onto a parafilm sheet placed on the bench.
30. Position the nitrocellulose membrane so that the proteins directly face the revelation liquid. Wait for 5 min in the dark.
31. Transfer the membrane onto a tray for protein detection and place the tray into the Fusion Solo machine.
32. For capturing images using enhanced chemiluminescence, use the Fusion software (chemiluminescence and fluorescence imaging).

In case your POI expression is low, or the signal is weak with the SuperSignal West Pico plus, we recommend using the SuperSignal West Femto, Maximum sensitivity substrate, and repeat the same procedure.

Protein storage

33. According to protein stability, store the POI at -80°C in 10% (v/v) sterile glycerol until use.

PROTEIN TRANSDUCTION IN MAMMALIAN CELLS

In this protocol, we describe the procedure of protein transduction within Hep3B mammalian cells. The investigation of protein entry into mammalian cells is a pivotal aspect of cellular and molecular biology, shedding light on essential physiological processes and enabling the development of targeted therapeutic strategies. This protocol delineates a fundamental approach for the transduction of proteins into mammalian cells, a process to efficiently deliver exogenous recombinant proteins. Furthermore, this protocol provides a detailed guide for monitoring the intracellular fate of transduced proteins through immunofluorescence microscopy, offering a powerful tool for visualizing and analyzing their subcellular localization and dynamics through fluorescent tags, such as EGFP and mCherry.

Materials

- Hep3B cells (ATCC, cat. no. HB-8064)
- Gibco minimum essential medium (MEM) (Dutscher, cat. no. L0430-50), or other cell culture medium adapted to the cell line used
- Fetal calf serum (FCS) (Corning, cat. no. 35-079-CV)
- 2 mM L-glutamine (Gibco by Life Technologies, cat. no. 25030-024)
- 1 mM sodium pyruvate (Gibco by Life Technologies, cat. no. 11360070)
- 50 mg/L antibiotics (streptomycin, penicillin) (Gibco by Life Technologies, cat. no. 15070-063)
- DPBS w/o Mg, w/o Ca (Dutsher, cat. no. L0615-500)
- Trypsin-EDTA, 1× in PBS w/o Ca, w/o Mg w/o phenol red (Dutsher, cat. no. L0940-100)
- Trypan blue solution (Sigma-Aldrich, cat. no. T8154)
- Purified POI in DPBS, 1×
- DAPI (4',6-diamidino-2-phenylindole dihydrochloride) (Sigma-Aldrich, cat. no. D9542)
- 70% ethanol
- Moviol mounting medium (see recipe)
- Water bath
- T25 flasks (Corning, cat. no. 430720U)

Humidified cell culture incubator enriched with 5% (v/v) CO₂
Leica DM IRB optical microscope
T75 flasks (Corning, cat. no. 430641U)
Suction pump
15-ml sterile conical tubes (Falcon, cat. no. 352097)
Malassez slide
Cell counter
24-well plates (Corning, cat. no. 3526)
120-mm coverslips
4% (w/v) paraformaldehyde (PAF) (Sigma-Aldrich, cat. no. P6148)
Humid chamber
Whatman paper
Parafilm
Zeiss LSM 780 confocal Microscope
Fiji ImageJ

CAUTION: Work consistently under the hood (with ventilation) and always wear gloves when handling PAF due to its toxic properties as a CMR substance.

NOTE: In cell culture, it is essential to maintain a sterile working environment. Clean the biological safety cabinet (BSC) thoroughly with 70% ethanol and disinfectant solution before each manipulation or experiment. Thoroughly clean all bottles used in cell culture with 70% ethanol before handling, and always manipulate under the BSC to prevent potential contamination.

Cell culture

1. Prepare the cell culture medium by mixing MEM to 10% (v/v) FCS, 2 mM L-glutamine, 1 mM sodium pyruvate, and 50 mg/L antibiotics. Place the cell culture medium, 1 × DPBS, and trypsin in a water bath at 37°C for ~10 to 20 min to achieve the optimal temperature for cell culture maintenance.
2. Initiate the culture of Hep3B cells in a T25 flask within an incubator set at 37°C with 5% (v/v) CO₂.
3. Monitor cell proliferation using an optical (visible light) microscope to ensure viability; prevent contamination before the experiment and maintain confluency while adjusting the medium if necessary.
4. Once Hep3B cells reach 70% confluence in a T75 flask, discard the cell medium.
5. Wash the cells with DPBS and discard DPBS using a suction pump.
6. Add 2 ml trypsin and place the flask in the cell incubator at 37°C until the cells detach from the flask (~5 to 10 min). Confirm cell detachment under the optical microscope.
7. Add 4 ml cell culture medium, twice the volume of trypsin on the cell monolayer. Discard the total 6 ml and transfer it to a sterile 15-ml sterile conical tube.
8. Take a 10 µl cell sample and mix with 10 µl trypan blue (1:1 ratio).
9. Use a Malassez slide for cell counting, placing 10 µl in each chamber and scan with a cell counting machine. Calculate the volume of cells to add to the culture medium for a final volume of 1 ml per well.

Protein transduction within living cells

10. In a 24-well plate, place a 120-mm coverslip in each well. Seed the cells for 24 hr by depositing ~37,500 cells with cell culture medium to achieve a final volume of 1 ml per well.

11. After 24 hr, the cells should be well adherent and have reached ~ 70% confluence.
12. Discard 1 ml medium, then add 1 to 5 μM of purified TAT-HA-POI-XFP (or POI-XFP if the TEV sequence has been cleaved) onto cells with the required culture medium volume to reach a final volume of 1 ml per well. Incubate the cells at 37°C for 2 hr.
13. Discard the medium, perform three quick washes with cold 1 \times DPBS, and fix the cells on the coverslips by adding 500 μl of 4% (w/v) PAF per well for 30 min at room temperature.

We can consider that cell transduction has occurred after 2 hr, otherwise adjust the time of contact between the POI, which the concentration can also be modulated, and the cells.

Do not exceed a 30 min exposure to 4% (w/v) PAF, as it is toxic to cells.

14. Wash the cells three times with DBPS for 5 min each.

Conduct the initial cell wash with 1 \times DPBS under the hood and in ventilation mode to eliminate any residual 4% (w/v) PAF.
15. Meanwhile, to avoid a drying out of the coverslips, prepare a humid chamber using a support, place a Whatman paper on it, wet it with ultrapure water, and cover it with parafilm.
16. Place the coverslip in the humid chamber and add 100 μl of 1 \times DPBS to cover the cells and prevent drying.
17. Discard DBPS and add DAPI to stain nuclei directly to each coverslip for 10 min at room temperature in three washes of 5 min each. Clean the coverslip with 70% ethanol and let it dry.
18. Add 5 μl Moviol (mounting liquid) to the coverslip and place it directly on the Moviol so that the side containing the fixed cells faces the liquid.
19. Let the coverslip dry at room temperature overnight in the dark. The next day, the coverslip is ready for observation under a confocal microscope.
20. Capture the images by using a confocal microscope:
 - a. Channel 1 (DAPI), λ_{ex} 405 nm, λ_{em} 410 to 470 nm.
 - b. Channel 2 (EGFP), λ_{ex} 488 nm, λ_{em} 500 to 543 nm.
 - c. Channel 3 (mCherry), λ_{ex} 561 nm, λ_{em} 570-615 nm.
21. Images can be analyzed by using Fiji ImageJ software (Fig. 8).

REAGENTS AND SOLUTIONS

Blocking buffer

5% (w/v) non-fatty-acid milk (for food use) or BSA (bovine serum albumin)
(Sigma-Aldrich, cat. no. A9647) in TBS-Tween buffer (see recipe)

Mix well

Prepare fresh

Buffer A

3.15 g Tris (20 mM final; Euromedex, cat. no. EU001B)

17.53 g NaCl (300 mM final; Carlo Erba, cat. no. 479685)

1.36 g imidazole (20 mM final; Sigma-Aldrich, cat. no. I202)

Adjust to 1 L with ultrapure H₂O

Adjust to pH 8.0 with HCl (37% w/v) (Carlo Erba, cat. no. 403871)

Mix well and filter using a 0.22- μm filter

Store up to 1 month at 4°C

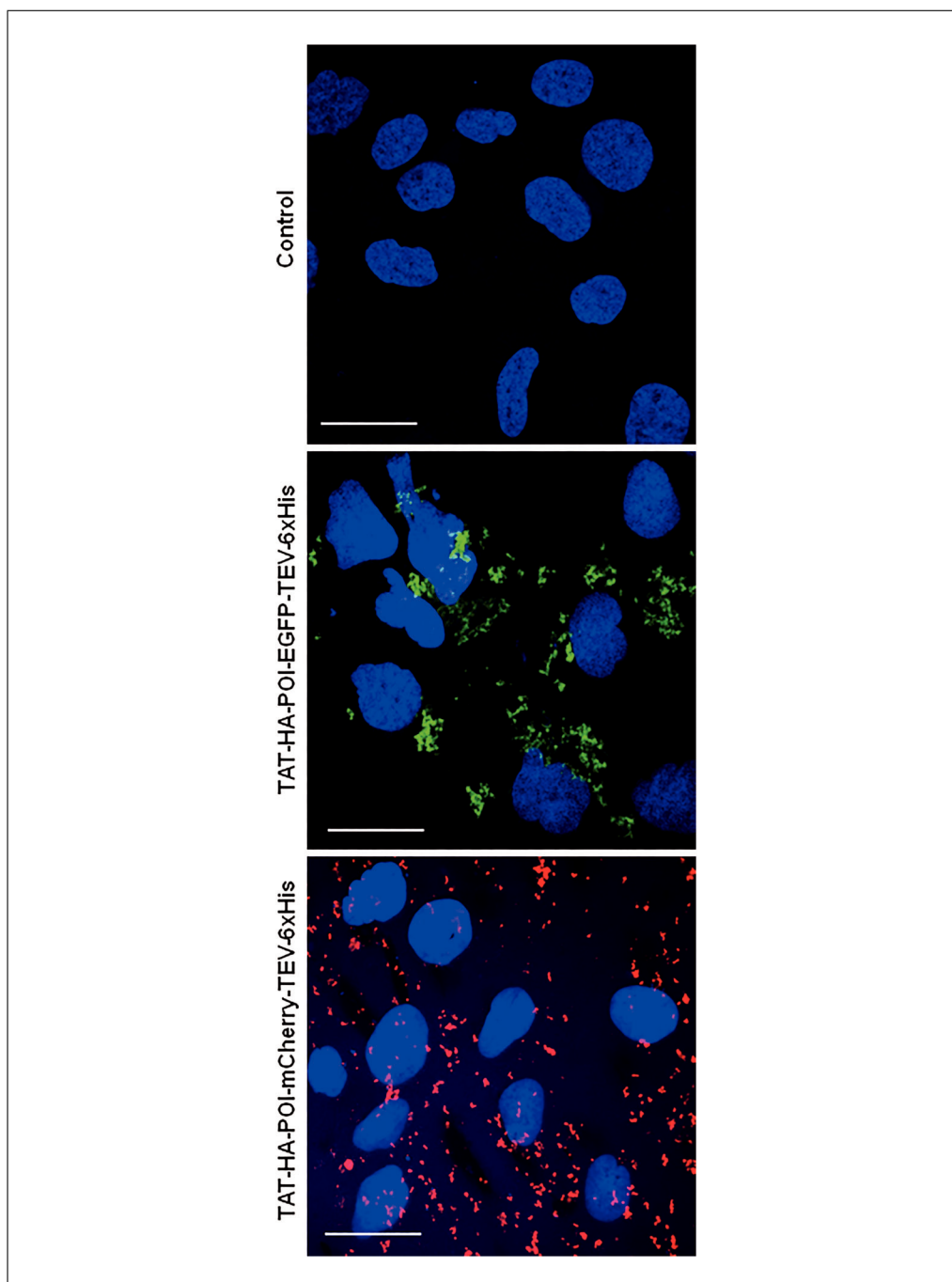


Figure 8 Fluorescence microscopy after treatment for 2 hr of Hep3B cells with 2 μ M of recombinant POI-EGFP and POI-mCherry fusion isoforms. Cells were fixed on coverslips with 4% (w/v) paraformaldehyde. Cells were counterstained for DNA with 1 μ g/ml DAPI (blue fluorescence), washed with PBS and then mounted in Mowiol for microscopy analysis. Scale bar: 20 μ m.

Buffer B

3.15 g Tris (20 mM final; Euromedex, cat. no. EU001B)
 17.53 g NaCl (300 mM final; Carlo Erba, cat. no. 479685)
 34 g imidazole (500 mM final; Sigma-Aldrich, cat. no. I202)
 Adjust to 1 L with ultrapure H₂O
 Adjust to pH 8.0 with HCl (37% w/v) (Carlo Erba, cat. no. 403871)
 Mix well and filter using a 0.22- μ m filter
 Store up to 1 month at 4°C

Laemmli buffer, 5×

12.11 g Tris (250 mM final; Euromedex, cat. no. EU001B)
5% (w/v) SDS (Sigma-Aldrich, cat. no. L3771)
5% (v/v) 2-mercaptoethanol (Sigma-Aldrich, cat. no. M3148)
40% (v/v) glycerol (Sigma-Aldrich, cat. no. G9012)
0.02% (w/v) bromophenol blue (Sigma-Aldrich, cat. no. B0126)
Bring up to 1 L with ultrapure H₂O
Adjust to pH 8.0 with HCl (37% w/v) (Carlo Erba, cat. no. 403871)
Mix well
Store at up to 1 month at 4°C

Polyacrylamide separating gel, 8%

3.75 ml of 1.5 M Tris·HCl pH 8.8 (Current Protocols, 2006)
50 µl of 20% (w/v) SDS (Sigma-Aldrich, cat. no. L3771)
2 ml of 40% (w/v) acrylamide/bis-acrylamide (ratio 29:1) (Sigma-Aldrich, cat. no. A2792)
30 µl Temed (*N,N,N',N'*-tetramethylethylenediamine) (Sigma-Aldrich, cat. no. T9281)
30 µl of 10% (w/v) ammonium persulfate (Sigma-Aldrich, cat. no. A7460)
Bring up to 10 ml with ultrapure H₂O
Mix well
Prepare fresh

Polyacrylamide stacking gel, 4%

1.5 ml of 0.5 M Tris·HCl pH 6.8 (Current Protocols, 2006)
50 µl of 20% (w/v) SDS (Sigma-Aldrich, cat. no. L3771)
1 ml of 40% (w/v) acrylamide/bis-acrylamide (ratio 29:1) (Sigma-Aldrich, cat. no. A2792)
60 µl Temed (*N,N,N',N'*-tetramethylethylenediamine) (Sigma-Aldrich, cat. no. T9281)
50 µl of 10% (w/v) ammonium persulfate (Sigma-Aldrich, cat. no. A7460)
Bring up to 10 ml with ultrapure H₂O
Mix well
Prepare fresh

Mowiol mounting medium

6 g glycerol
2.4 g mowiol 4-88 (biovalley, cat. no. 17951-500)
6 ml distilled H₂O
12 ml 0.2 M Tris·HCl pH 8.5 (Current Protocols, 2006)
Mix well and centrifuge 10 min at 1500 × g, room temperature, to clarify the solution
Store up to 1 month at 4°C or 1 year at −20°C

TBS-Tween buffer

4.2 g Tris (15 mM final; Euromedex, cat. no. EU001B)
8.1 g NaCl (300 mM final; Carlo Erba, cat. no. 479685)
0.05% (v/v) Tween 20 (Fisher Scientific, cat. no. M12247)
Adjust to 1 L with distilled H₂O
Adjust to pH 8.0 with HCl (37% w/v) (Carlo Erba, cat. no. 403871)
Mix well
Store up to 1 week at room temperature

COMMENTARY

Critical Parameters

Several steps in this experimental strategy are critical and therefore can greatly influence its success. It is important to correctly design the primers so that the different PCRs that precede the DNA molecular assembly are as specific as possible to generate a single band visualized on the agarose gel (Basic Protocol 1). This step is fundamental to getting off to a good start. Another crucial point is the bacteria subculturing stage. It is important to pick well-isolated bacterial colonies back into liquid culture so as to not mix different clones (Basic Protocol 1). The last step concerning the transduction of cells is relatively simple because the use of penetration peptides makes it possible to avoid many of the steps inherent in the transfection of plasmid vectors (Basic Protocol 3). Nevertheless, it is important here to test variable quantities of proteins to be transduced to optimize their subsequent study. If the POI is not used quickly after its purification, it is extremely important to store it in suitable conditions. Nevertheless, we encourage users to extemporaneously prepare the protein to transduce it immediately after purification (see Table 8). Other potential critical parameters are listed in Table 8.

Troubleshooting Table

Table 8 lists the most common problems that may be encountered during protein production detailed in this article (Table 8).

Statistical Analysis

GraphPad Prism 9 or any other statistics tool can be used for calculating the standard deviations of the growth curves of bacteria, as those shown in Figure 4.

Understanding Results

Basic Protocols 1 and 2

The data obtained by Basic Protocols 1 and 2 generate agarose gel electrophoretic profiles to assess the quantity and quality of the PCR products, and polyacrylamide gel profiles that can be visualized either by treatment of the gels with a solution of Coomassie blue or by electro-blotting of the resolved proteins by SDS-PAGE onto nitrocellulose. These experiments require carefully supporting the samples to be analyzed with size markers for agarose gels and molecular mass markers for polyacrylamide gels; be careful not to confuse the types of markers, the first being expressed in base pairs (bp), the second in kilodaltons (kDa).

Table 8 Troubleshooting Table for Protein Production

Problem	Possible Cause	Solution
Poor protein stability	Inappropriate storage conditions or buffer	Store protein at recommended temperature; optimize buffer composition; add proteases inhibitors to buffers
Low protein yield	Inefficient cell lysis or purification method	Optimize lysis conditions; consider alternative purification methods
Contaminants in elution	Carryover from previous steps or column issues	Ensure proper column equilibration; optimize washing steps
Low/no signal in western blot results	Poor antibody binding	Ensure primary and secondary antibodies are compatible and correctly diluted
High background or low ratio expected signal/background in western blot	Antibodies too much interact with the membrane or with the blocking reagent	Increase the time and the number of washings; change the blocking reagent; increase the dilution of the primary or/and the secondary antibodies
Weak or no fluorescence signal	Too low protein concentration in contact with the cell	Increase the concentration of the protein
Cell autofluorescence	Presence of residual medium or contaminants	Ensure thorough washing after treatment
Protein aggregation	Improper storage conditions or handling	Store protein appropriately; handle with care

The presentation of the data requires labeling the figures by clearly identifying at the top of the gels the samples loaded on the left or the right lateral side, and the values corresponding to each mark of size (bp) or molecular mass (kDa). By comparing the migration distance of the band corresponding to the PCR product or to the protein produced, it is possible to prove not only that the product is indeed present, but also to estimate its abundance and purity. The western blot provides an advantage since the use of antibodies guarantees an additional argument simply because the apparent molecular mass is indeed what is expected. Make sure to check the supplier's information sheet whether these antibodies can be used for western blotting. Some antibodies are more suitable for immunoprecipitation, ELISA, or immunofluorescence since they recognize proteins in their native form while others are better for western blot because they were generated against a short peptide sequence. This is the case for the latter in this series of protocols, we used antibodies directed against tag "labels" (HA and 6×His), which generally guarantees their use in western blotting, while anti-HA antibodies are better for western blotting than anti-6×His antibodies preferentially used for protein purification. The problem of non-recognition can, however, be encountered with anti-POI antibodies. In fact, in Figure 3 and Figure 7C-D, we can consider on the one hand the good purity of the different PCR products (this must be ensured before moving on to the molecular cloning stage, which will be more perilous if the PCRs are not sufficiently pure) and the protein thus produced. It is common to observe, as we do in Figure 7C-D, bands of molecular mass lower than the expected band; these are generally proteolytic degradation bands generated during production and purification. In our case it turns out that the mCherry form of our POI is less stable than the GFP form (it may be useful in this case to use larger quantities of protease inhibitors during the process).

Basic Protocol 3

Regarding images obtained by microscopy in confocal mode (Basic Protocol 3), they must be clearly labeled by indicating the corresponding condition. It is very useful to merge with the DAPI image, which allows to locate the POI at a glance. Our microscopy images show a cytoplasmic distribution of our POI in both the EGFP and mCherry versions. Finally, a scale bar (20 μm in our experience) makes it possible to judge the distances on

the microscopy images. It is common and even advisable to show separate images and merge for immunofluorescence microscopy experiments.

Time Considerations

Total time required to complete the full experimental strategy: 3 to 4 weeks.

Basic Protocol 1 (7 days)

Amplification of the various PCR products

Samples preparation: 10 min

PCR: 3 hr

Agarose gel preparation and polymerization: 30 min

Agarose gel electrophoresis: 30 min

Ethidium bromide staining: 15 min

Bands visualization by using UV transilluminator: 10 min

Bands excision by using UV transilluminator: 15 min

PCR gel purification: 1 hr

Total: 1 day

DNA molecular assembly

Plasmid assembly procedure: 1 hr

Transformation and selection procedure (*E. coli* 5-alpha): 1 days

DNA extraction and purification: 1 hr

DNA sequencing: 2 days

Total: 4 days

*Transformation and selection procedure in *E. coli* BL21(DE3)*

Total: 1 day

Recombinant protein expression

Total: 1 day

Basic Protocol 2 (5 days)

Protein fractionation

Total: 1 day

Protein desalting, concentration, and quantification

Total: 3 days

Analysis of recombinant protein expression by SDS-PAGE

Coomassie blue staining

Western blot

Total: 1 day

Basic Protocol 3 (3 days)

Cell seeding

Total: 1 day

Protein transduction
Protein transduction: 2 hr
Slide and coverslip preparation
Total: 1 day

Data collection and analysis
Immunofluorescence microscopy: data
collecting
Data analysis: 1 to 2 hr
Total: 1 day

Acknowledgments

The authors are grateful to the European Union's Horizon 2020 research and innovation program that funded this work (Marie Skłodowska-Curie grant agreement No 847568), the Centre National de la Recherche Scientifique (CNRS) and the University of Lille for their recurrent fundings, and the Plateformes Lilloises en Biologie et Santé (PLBS) for help in using fluorescent microscopy.

Author Contributions

Christer Abou Anny: Conceptualization; data curation; formal analysis; methodology; writing original draft. **Sébastien Nouaille:** Conceptualization; investigation; methodology; supervision; validation; writing review and editing. **Régis Fauré:** Conceptualization; investigation; methodology; supervision; validation; writing review and editing. **Céline Schulz:** Conceptualization; formal analysis; methodology; writing review and editing. **Corentin Spriet:** Conceptualization; methodology; writing review and editing. **Isabelle Huvent:** Conceptualization; investigation; methodology; supervision; writing review and editing. **Christophe Biot:** Conceptualization; formal analysis; funding acquisition; investigation; resources; supervision; validation; writing review and editing. **Tony Lefebvre:** Conceptualization; formal analysis; funding acquisition; investigation; project administration; resources; supervision; writing original draft.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created in this study.

Literature Cited

Ai, H., Baird, M. A., Shen, Y., Davidson, M. W., & Campbell, R. E. (2014). Engineering and characterizing monomeric fluorescent proteins for live-cell imaging applications. *Nature Pro-*

ocols, 9(4), 910–928. <https://doi.org/10.1038/nprot.2014.054>

- Andersen, D. C., & Krummen, L. (2002). Recombinant protein expression for therapeutic applications. *Current Opinion in Biotechnology*, 13(2), 117–123. [https://doi.org/10.1016/S0958-1669\(02\)00300-2](https://doi.org/10.1016/S0958-1669(02)00300-2)
- MacKay, A. J., & Szoka, F. C. Jr. (2003). HIV TAT protein transduction domain mediated cell binding and intracellular delivery of nanoparticles. *Journal of Dispersion Science and Technology*, 24(3–4), 465–473. <https://doi.org/10.1081/DIS-120021802>
- Anné, J., Maldonado, B., Van Impe, J., Van Mellaert, L., & Bernaerts, K. (2012). Recombinant protein production and streptomycetes. *Journal of Biotechnology*, 158(4), 159–167. <https://doi.org/10.1016/j.jbiotec.2011.06.028>
- Assenberg, R., Wan, P. T., Geisse, S., & Mayr, L. M. (2013). Advances in recombinant protein expression for use in pharmaceutical research. *Current Opinion in Structural Biology*, 23(3), 393–402. <https://doi.org/10.1016/j.sbi.2013.03.008>
- Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology*, 10(5), 411–421. [https://doi.org/10.1016/S0958-1669\(99\)00003-8](https://doi.org/10.1016/S0958-1669(99)00003-8)
- Baneyx, F., & Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology*, 22(11), 1399–1408. <https://doi.org/10.1038/nbt1029>
- Barnes, L. M., Bentley, C. M., & Dickson, A. J. (2003). Stability of protein production from recombinant mammalian cells. *Biotechnology and Bioengineering*, 81(6), 631–639. <https://doi.org/10.1002/bit.10517>
- Becker-Hapak, M., McAllister, S. S., & Dowdy, S. F. (2001). TAT-mediated protein transduction into mammalian cells. *Methods*, 24(3), 247–256. <https://doi.org/10.1006/meth.2001.1186>
- Block, H., Maertens, B., Priestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J., & Schäfer, F. (2009). Immobilized-metal affinity chromatography (IMAC): A review. *Methods in Enzymology*, 463, 439–473. [https://doi.org/10.1016/S0076-6879\(09\)63027-5](https://doi.org/10.1016/S0076-6879(09)63027-5)
- Bordat, A., Houvenaghel, M.-C., & German-Retana, S. (2015). Gibson assembly: An easy way to clone potyviral full-length infectious cDNA clones expressing an ectopic VPg. *Virology Journal*, 12(1), 89. <https://doi.org/10.1186/s12985-015-0315-3>
- Brondyk, W. H. (2009). Selecting an appropriate method for expressing a recombinant protein. *Methods in Enzymology*, 463, 131–147. [https://doi.org/10.1016/S0076-6879\(09\)63011-1](https://doi.org/10.1016/S0076-6879(09)63011-1)
- Burnett, M. J. B., & Burnett, A. C. (2020). Therapeutic recombinant protein production in plants: Challenges and opportunities. *Plants, People, Planet*, 2(2), 121–132. <https://doi.org/10.1002/ppp3.10073>
- Casini, A., Storch, M., Baldwin, G. S., & Ellis, T. (2015). Bricks and blueprints: Methods and

- standards for DNA assembly. *Nature Reviews Molecular Cell Biology*, 16(9), 568–576. <https://doi.org/10.1038/nrm4014>
- Chen, R. (2012). Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnology Advances*, 30(5), 1102–1107. <https://doi.org/10.1016/j.biotechadv.2011.09.013>
- Cohen, S. N. (2013). DNA cloning: A personal view after 40 years. *Proceedings of the National Academy of Sciences*, 110(39), 15521–15529. <https://doi.org/10.1073/pnas.1313397110>
- Cregg, J. M., Cereghino, J. L., Shi, J., & Higgins, D. R. (2000). Recombinant protein expression in *Pichia pastoris*. *Molecular Biotechnology*, 16(1), 23–52. <https://doi.org/10.1385/MB:16:1:23>
- Current Protocols. (2006). Commonly Used Reagents. *Current Protocols in Microbiology*, 00, A.2A.1–A.2A.15. <https://doi.org/10.1002/9780471729259.mca02as00>
- Eguchi, A., Akuta, T., Okuyama, H., Senda, T., Yokoi, H., Inokuchi, H., Fujita, S., Hayakawa, T., Takeda, K., Hasegawa, M., & Nakanishi, M. (2001). Protein transduction domain of HIV-1 Tat protein promotes efficient delivery of DNA into mammalian cells. *Journal of Biological Chemistry*, 276(28), 26204–26210. <https://doi.org/10.1074/jbc.M010625200>
- Farzaneh, M., Hassani, S.-N., Mozdziak, P., & Baharvand, H. (2017). Avian embryos and related cell lines: A convenient platform for recombinant proteins and vaccine production. *Biotechnology Journal*, 12(5), 1600598. <https://doi.org/10.1002/biot.201600598>
- Ferro, M. M. M., Ramos-Sobrinho, R., Xavier, C. A. D., Zerbini, F. M., Lima, G. S. A., Nagata, T., & Assunção, I. P. (2019). New approach for the construction of infectious clones of a circular DNA plant virus using Gibson assembly. *Journal of Virological Methods*, 263, 20–23. <https://doi.org/10.1016/j.jviromet.2018.10.017>
- Fittipaldi, A., & Giacca, M. (2005). Transcellular protein transduction using the Tat protein of HIV-1. *Advanced Drug Delivery Reviews*, 57(4), 597–608. <https://doi.org/10.1016/j.addr.2004.10.011>
- Ford, K. G., Souberbielle, B. E., Darling, D., & Farzaneh, F. (2001). Protein transduction: An alternative to genetic intervention? *Gene Therapy*, 8(1), 1–4. <https://doi.org/10.1038/sj.gt.3301383>
- Froger, A., & Hall, J. E. (2007). Transformation of plasmid DNA into *E. coli* using the heat shock method. *Journal of Visualized Experiments*, 6, e253. <https://doi.org/10.3791/253>
- Gaberc-Porekar, V., & Menart, V. (2001). Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods*, 49(1), 335–360. [https://doi.org/10.1016/S0165-022X\(01\)00207-X](https://doi.org/10.1016/S0165-022X(01)00207-X)
- Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5), 343–345. <https://doi.org/10.1038/nmeth.1318>
- Gopal, G. J., & Kumar, A. (2013). Strategies for the production of recombinant protein in *Escherichia coli*. *The Protein Journal*, 32(6), 419–425. <https://doi.org/10.1007/s10930-013-9502-5>
- Greenblatt, J., & Schleif, R. (1971). Arabinose C protein: Regulation of the arabinose operon in vitro. *Nature New Biology*, 233(40), 166–170. <https://doi.org/10.1038/newbio233166a0>
- Gump, J. M., & Dowdy, S. F. (2007). TAT transduction: The molecular mechanism and therapeutic prospects. *Trends in Molecular Medicine*, 13(10), 443–448. <https://doi.org/10.1016/j.molmed.2007.08.002>
- Gump, J. M., June, R. K., & Dowdy, S. F. (2010). Revised role of glycosaminoglycans in TAT protein transduction domain-mediated cellular transduction. *Journal of Biological Chemistry*, 285(2), 1500–1507. <https://doi.org/10.1074/jbc.M109.021964>
- Guo, Q., Zhao, G., Hao, F., & Guan, Y. (2012). Effects of the TAT peptide orientation and relative location on the protein transduction efficiency. *Chemical Biology & Drug Design*, 79(5), 683–690. <https://doi.org/10.1111/j.1747-0285.2011.01315.x>
- Helmuth, J. A., Burckhardt, C. J., Greber, U. F., & Sbalzarini, I. F. (2009). Shape reconstruction of subcellular structures from live cell fluorescence microscopy images. *Journal of Structural Biology*, 167(1), 1–10. <https://doi.org/10.1016/j.jsb.2009.03.017>
- Hopkins, R. F., Wall, V. E., & Esposito, D. (2012). Optimizing transient recombinant protein expression in mammalian cells. *Methods in Molecular Biology*, 801, 251–268. https://doi.org/10.1007/978-1-61779-352-3_16
- Jäger, V., Büssov, K., & Schirrmann, T. (2015). Transient recombinant protein expression in mammalian cells. In M. Al-Rubeai (Ed.), *Animal Cell Culture. Cell Engineering* (Vol. 9, pp. 27–64). Springer. https://doi.org/10.1007/978-3-319-10320-4_2
- Jakobs, S., Subramaniam, V., Schönle, A., Jovin, T. M., & Hell, S. W. (2000). EGFP and DsRed expressing cultures of *Escherichia coli* imaged by confocal, two-photon and fluorescence lifetime microscopy. *FEBS Letters*, 479(3), 131–135. [https://doi.org/10.1016/S0014-5793\(00\)01896-2](https://doi.org/10.1016/S0014-5793(00)01896-2)
- Jang, C.-W., & Magnuson, T. (2013). A novel selection marker for efficient DNA cloning and recombineering in *E. coli*. *PLoS ONE*, 8(2), e57075. <https://doi.org/10.1371/journal.pone.0057075>
- Joung, Y. H., Park, S. H., Moon, K.-B., Jeon, J.-H., Cho, H.-S., & Kim, H.-S. (2016). The last ten years of advancements in plant-derived recombinant vaccines against Hepatitis B. *International Journal of Molecular Sciences*, 17(10), 1715. <https://doi.org/10.3390/ijms17101715>

- Kabouridis, P. S. (2003). Biological applications of protein transduction technology. *Trends in Biotechnology*, *21*(11), 498–503. <https://doi.org/10.1016/j.tibtech.2003.09.008>
- Kaplan, I. M., Wadia, J. S., & Dowdy, S. F. (2005). Cationic TAT peptide transduction domain enters cells by macropinocytosis. *Journal of Controlled Release*, *102*(1), 247–253. <https://doi.org/10.1016/j.jconrel.2004.10.018>
- Kirchhausen, T. (2009). Imaging endocytic clathrin structures in living cells. *Trends in Cell Biology*, *19*(11), 596–605. <https://doi.org/10.1016/j.tcb.2009.09.002>
- Kostylev, M., Otwell, A. E., Richardson, R. E., & Suzuki, Y. (2015). Cloning should be simple: *Escherichia coli* DH5 α -mediated assembly of multiple DNA fragments with short end homologies. *PLoS ONE*, *10*(9), e0137466. <https://doi.org/10.1371/journal.pone.0137466>
- Leifert, J. A., Harkins, S., & Whitton, J. L. (2002). Full-length proteins attached to the HIV tat protein transduction domain are neither transduced between cells, nor exhibit enhanced immunogenicity. *Gene Therapy*, *9*(21), 1422–1428. <https://doi.org/10.1038/sj.gt.3301819>
- Li, C., Ji, C., Huguet-Tapia, J. C., White, F. F., Dong, H., & Yang, B. (2019). An efficient method to clone TAL effector genes from *Xanthomonas oryzae* using Gibson assembly. *Molecular Plant Pathology*, *20*(10), 1453–1462. <https://doi.org/10.1111/mpp.12820>
- Li, L., Jiang, W., & Lu, Y. (2018). A modified Gibson assembly method for cloning large DNA fragments with high GC contents. *Methods in Molecular Biology*, *1671*, 203–209. https://doi.org/10.1007/978-1-4939-7295-1_13
- Li, Z.-N., Jelkmann, W., Sun, P., & Zhang, L. (2020). Construction of full-length infectious cDNA clones of apple stem grooving virus using Gibson assembly method. *Virus Research*, *276*, 197790. <https://doi.org/10.1016/j.virusres.2019.197790>
- Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M., & Porro, D. (2012). Recombinant protein production in yeasts. *Methods in Molecular Biology*, *824*, 329–358. https://doi.org/10.1007/978-1-61779-433-9_17
- McDonald, D., Vodicka, M. A., Lucero, G., Svitkina, T. M., Borisy, G. G., Emerman, M., & Hope, T. J. (2002). Visualization of the intracellular behavior of HIV in living cells. *Journal of Cell Biology*, *159*(3), 441–452. <https://doi.org/10.1083/jcb.200203150>
- McRae, S. R., Brown, C. L., & Bushell, G. R. (2005). Rapid purification of EGFP, EYFP, and ECFP with high yield and purity. *Protein Expression and Purification*, *41*(1), 121–127. <https://doi.org/10.1016/j.pep.2004.12.030>
- Newman, R. H., Fosbrink, M. D., & Zhang, J. (2011). Genetically encodable fluorescent biosensors for tracking signaling dynamics in living cells. *Chemical Reviews*, *111*(5), 3614–3666. <https://doi.org/10.1021/cr100002u>
- Nienhaus, K., & Nienhaus, G. U. (2014). Fluorescent proteins for live-cell imaging with super-resolution. *Chemical Society Reviews*, *43*(4), 1088–1106. <https://doi.org/10.1039/C3CS60171D>
- Noguchi, H., & Matsumoto, S. (2006). Protein transduction technology: A novel therapeutic perspective. *Acta Medica Okayama*, *60*(1), 1–11. <http://doi.org/10.18926/AMO/30757>
- Nolte, C., Matyash, M., Pivneva, T., Schipke, C. G., Ohlemeyer, C., Hanisich, U.-K., Kirchhoff, F., & Kettenmann, H. (2001). GFAP promoter-controlled EGFP-expressing transgenic mice: A tool to visualize astrocytes and astrogliosis in living brain tissue. *Glia*, *33*(1), 72–86.
- Norris, A. D., Kim, H.-M., Colaiácovo, M. P., & Calarco, J. A. (2015). Efficient genome editing in *Caenorhabditis elegans* with a toolkit of dual-marker selection cassettes. *Genetics*, *201*(2), 449–458. <https://doi.org/10.1534/genetics.115.180679>
- Overton, T. W. (2014). Recombinant protein production in bacterial hosts. *Drug Discovery Today*, *19*(5), 590–601. <https://doi.org/10.1016/j.drudis.2013.11.008>
- Pawson, T., & Nash, P. (2000). Protein-protein interactions define specificity in signal transduction. *Genes & Development*, *14*(9), 1027–1047. <https://doi.org/10.1101/gad.14.9.1027>
- Porath, J. (1992). Immobilized metal ion affinity chromatography. *Protein Expression and Purification*, *3*(4), 263–281. [https://doi.org/10.1016/1046-5928\(92\)90001-D](https://doi.org/10.1016/1046-5928(92)90001-D)
- Porro, D., Sauer, M., Branduardi, P., & Mattanovich, D. (2005). Recombinant protein production in yeasts. *Molecular Biotechnology*, *31*(3), 245–259. <https://doi.org/10.1385/MB:31:3:245>
- Rizzuto, R., Brini, M., Pizzo, P., Murgia, M., & Pozzan, T. (1995). Chimeric green fluorescent protein as a tool for visualizing subcellular organelles in living cells. *Current Biology*, *5*(6), 635–642. [https://doi.org/10.1016/S0960-9822\(95\)00128-X](https://doi.org/10.1016/S0960-9822(95)00128-X)
- Roberts, M. A. J. (2019). Recombinant DNA technology and DNA sequencing. *Essays in Biochemistry*, *63*(4), 457–468. <https://doi.org/10.1042/EBC20180039>
- Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology*, *5*, 172. <https://www.frontiersin.org/articles/10.3389/fmicb.2014.00172>
- Rudenko, O., & Barnes, A. C. (2018). Gibson assembly facilitates bacterial allelic exchange mutagenesis. *Journal of Microbiological Methods*, *144*, 157–163. <https://doi.org/10.1016/j.mimet.2017.11.023>
- Salomonsson, E., Mihalko, L. A., Luker, K. E., Verkhusha, V. V., & Luker, G. D. (2012). Cell-based and *in vivo* spectral analysis of fluorescent proteins for multiphoton microscopy. *Journal of Biomedical Optics*, *17*(9),

096001. <https://doi.org/10.1117/1.JBO.17.9.096001>
- Sands, B., & Brent, R. (2016). Overview of post Cohen-Boyer methods for single segment cloning and for multisegment DNA assembly. *Current Protocols in Molecular Biology*, *113*(1), 3.26.1–3.26.20. <https://doi.org/10.1002/0471142727.mb0326s113>
- Schleif, R. (2000). Regulation of the L-arabinose operon of *Escherichia coli*. *Trends in Genetics*, *16*(12), 559–565. [https://doi.org/10.1016/S0168-9525\(00\)02153-3](https://doi.org/10.1016/S0168-9525(00)02153-3)
- Schleif, R., Hess, W., Finkelstein, S., & Ellis, D. (1973). Induction kinetics of the L-arabinose operon of *Escherichia coli*. *Journal of Bacteriology*, *115*(1), 9–14. <https://doi.org/10.1128/jb.115.1.9-14.1973>
- Schwarze, S. R., Hruska, K. A., Dowdy, S. F., Schwarze, S. R., Hruska, K. A., & Dowdy, S. F. (2000). Protein transduction: Unrestricted delivery into all cells? *Trends in Cell Biology*, *10*(7), 290–295. [https://doi.org/10.1016/S0962-8924\(00\)01771-2](https://doi.org/10.1016/S0962-8924(00)01771-2)
- Shen, Y., Chen, Y., Wu, J., Shaner, N. C., & Campbell, R. E. (2017). Engineering of mCherry variants with long Stokes shift, red-shifted fluorescence, and low cytotoxicity. *PLoS ONE*, *12*(2), e0171257. <https://doi.org/10.1371/journal.pone.0171257>
- Shokolenko, I. N., Alexeyev, M. F., LeDoux, S. P., & Wilson, G. L. (2005). TAT-mediated protein transduction and targeted delivery of fusion proteins into mitochondria of breast cancer cells. *DNA Repair*, *4*(4), 511–518. <https://doi.org/10.1016/j.dnarep.2004.11.009>
- Sørensen, H. P., & Mortensen, K. K. (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology*, *115*(2), 113–128. <https://doi.org/10.1016/j.jbiotec.2004.08.004>
- Subach, F. V., Patterson, G. H., Manley, S., Gillette, J. M., Lippincott-Schwartz, J., & Verkhusha, V. V. (2009). Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nature Methods*, *6*(2), 153–159. <https://doi.org/10.1038/nmeth.1298>
- Sun, X., Chiu, J.-F., & He, Q.-Y. (2005). Application of immobilized metal affinity chromatography in proteomics. *Expert Review of Proteomics*, *2*(5), 649–657. <https://doi.org/10.1586/14789450.2.5.649>
- Sunley, K., & Butler, M. (2010). Strategies for the enhancement of recombinant protein production from mammalian cells by growth arrest. *Biotechnology Advances*, *28*(3), 385–394. <https://doi.org/10.1016/j.biotechadv.2010.02.003>
- Thomas, S., Maynard, N. D., & Gill, J. (2015). DNA library construction using Gibson assembly®. *Nature Methods*, *12*(11), i–ii. <https://doi.org/10.1038/nmeth.f.384>
- Wadia, J. S., & Dowdy, S. F. (2002). Protein transduction technology. *Current Opinion in Biotechnology*, *13*(1), 52–56. [https://doi.org/10.1016/S0958-1669\(02\)00284-7](https://doi.org/10.1016/S0958-1669(02)00284-7)
- Wang, J. W., Wang, A., Li, K., Wang, B., Jin, S., Reiser, M., & Lockey, R. F. (2015). CRISPR/Cas9 nuclease cleavage combined with Gibson assembly for seamless cloning. *Biotechniques*, *58*(4), 161–170. <https://doi.org/10.2144/000114261>
- Wang, Y., Shyy, J. Y. J., & Chien, S. (2008). Fluorescence proteins, live-cell imaging, and mechanobiology: Seeing is believing. *Annual Review of Biomedical Engineering*, *10*(1), 1–38. <https://doi.org/10.1146/annurev.bioeng.010308.161731>
- Wiedenmann, J., Oswald, F., & Nienhaus, G. U. (2009). Fluorescent proteins for live cell imaging: Opportunities, limitations, and challenges. *IUBMB Life*, *61*(11), 1029–1042. <https://doi.org/10.1002/iub.256>
- Wurm, F. M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology*, *22*(11), 1393–1398. <https://doi.org/10.1038/nbt1026>
- Yin, J., Zhu, D., Zhang, Z., Wang, W., Fan, J., Men, D., Deng, J., Wei, H., Zhang, X.-E., & Cui, Z. (2013). Imaging of mRNA–protein interactions in live cells using novel mCherry trimolecular fluorescence complementation systems. *PLoS ONE*, *8*(11), e80851. <https://doi.org/10.1371/journal.pone.0080851>
- Zhang, H., Ma, Y., Gu, J., Liao, B., Li, J., Wong, J., & Jin, Y. (2012). Reprogramming of somatic cells via TAT-mediated protein transduction of recombinant factors. *Biomaterials*, *33*(20), 5047–5055. <https://doi.org/10.1016/j.biomaterials.2012.03.061>

Références bibliographiques

- Abdulsalam, M., Ubah, F.Z.Y., Ahmed, H.U., Tafida, U.A., Nasir, A.W., Abdulsalam, M., Ubah, F.Z.Y., Ahmed, H.U., Tafida, U.A., Nasir, A.W., **2024**. Deciphering the Genetic Code: Mechanisms, Evolution, and Implications for Biotechnology. *World J. Adv. Res. Rev.* 21, 858–868. <https://doi.org/10.30574/wjarr.2024.21.1.2195>
- Agard, N.J., Prescher, J.A., Bertozzi, C.R., **2004**. A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* 126, 15046–15047. <https://doi.org/10.1021/ja044996f>
- Alam, M.S., **2018**. Proximity Ligation Assay (PLA). *Curr. Protoc. Immunol.* 123, e58. <https://doi.org/10.1002/cpim.58>
- Alteen, M.G., Tan, H.Y., Vocadlo, D.J., **2021**. Monitoring and modulating O-GlcNAcylation: assays and inhibitors of O-GlcNAc processing enzymes. *Curr. Opin. Struct. Biol., Protein-Carbohydrate Complexes and Glycosylation Sequences and Topology* 68, 157–165. <https://doi.org/10.1016/j.sbi.2020.12.008>
- Anny, C.A., Nouaille, S., Fauré, R., Schulz, C., Spriet, C., Huvent, I., Biot, C., Lefebvre, T., **2024**. A Step-by-Step Guide for the Production of Recombinant Fluorescent TAT-HA-Tagged Proteins and their Transduction into Mammalian Cells. *Curr. Protoc.* 4, e1016. <https://doi.org/10.1002/cpz1.1016>
- Balana, A.T., Pratt, M.R., **2021**. Mechanistic roles for altered O-GlcNAcylation in neurodegenerative disorders. *Biochem. J.* 478, 2733–2758. <https://doi.org/10.1042/BCJ20200609>
- Becker-Hapak, M., McAllister, S.S., Dowdy, S.F., **2001**. TAT-Mediated Protein Transduction into Mammalian Cells. *Methods* 24, 247–256. <https://doi.org/10.1006/meth.2001.1186>
- Bolanle, I.O., Riches-Suman, K., Williamson, R., Palmer, T.M., **2021**. Emerging roles of protein O-GlcNAcylation in cardiovascular diseases: Insights and novel

therapeutic targets. *Pharmacol. Res.* 165, 105467.
<https://doi.org/10.1016/j.phrs.2021.105467>

Bordat, A., Houvenaghel, M.-C., German-Retana, S., **2015**. Gibson assembly: an easy way to clone potyviral full-length infectious cDNA clones expressing an ectopic VPg. *Viol. J.* 12, 89. <https://doi.org/10.1186/s12985-015-0315-3>

Boyce, M., Carrico, I.S., Ganguli, A.S., Yu, S.-H., Hangauer, M.J., Hubbard, S.C., Kohler, J.J., Bertozzi, C.R., **2011**. Metabolic cross-talk allows labeling of O-linked β -N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. *Proc. Natl. Acad. Sci.* 108, 3141–3146.
<https://doi.org/10.1073/pnas.1010045108>

Brown, W., Davidson, L.A., Deiters, A., **2024**. Expanding the Genetic Code of *Xenopus laevis* Embryos. *ACS Chem. Biol.* 19, 516–525.
<https://doi.org/10.1021/acscchembio.3c00686>

Buchowiecka, A.K., **2023**. Protein cysteine S-glycosylation: oxidative hydrolysis of protein S-glycosidic bonds in aqueous alkaline environments. *Amino Acids* 55, 61–74. <https://doi.org/10.1007/s00726-022-03208-7>

Cameron, A., Giacomozzi, B., Joyce, J., Gray, A., Graham, D., Ousson, S., Neny, M., Beher, D., Carlson, G., O'Moore, J., Shearman, M., Hering, H., **2013**. Generation and characterization of a rabbit monoclonal antibody site-specific for tau O-GlcNAcylated at serine 400. *FEBS Lett.* 587, 3722–3728.
<https://doi.org/10.1016/j.febslet.2013.09.042>

Chen, H., Venkat, S., McGuire, P., Gan, Q., Fan, C., **2018**. Recent Development of Genetic Code Expansion for Posttranslational Modification Studies. *Molecules* 23, 1662. <https://doi.org/10.3390/molecules23071662>

Chen, Y.-F., Shao, G.-C., Li, Jie, Yang, A.-Q., Li, Jing, Ye, X.-S., **2022**. O-GlcNAcylation of Blimp-1 in Lymphocytes Inhibits Its Transcriptional Function and Is Associated with Migration and Invasion of Breast Cancer Cells. *Mol. Cancer Res.* 20, 650–660. <https://doi.org/10.1158/1541-7786.MCR-21-0405>

Chuh, K.N., Batt, A.R., Zaro, B.W., Darabedian, N., Marotta, N.P., Brennan, C.K., Amirhekmat, A., Pratt, M.R., **2017**. The New Chemical Reporter 6-Alkynyl-6-deoxy-GlcNAc Reveals O-GlcNAc Modification of the Apoptotic Caspases That

Can Block the Cleavage/Activation of Caspase-8. *J. Am. Chem. Soc.* 139, 7872–7885. <https://doi.org/10.1021/jacs.7b02213>

Cioce, A., Bineva-Todd, G., Agbay, A.J., Choi, J., Wood, T.M., Debets, M.F., Browne, W.M., Douglas, H.L., Roustan, C., Tastan, O.Y., Kjaer, S., Bush, J.T., Bertozzi, C.R., Schumann, B., **2021**. Optimization of Metabolic Oligosaccharide Engineering with Ac4GalNAIk and Ac4GlcNAIk by an Engineered Pyrophosphorylase. *ACS Chem. Biol.* 16, 1961–1967. <https://doi.org/10.1021/acscchembio.1c00034>

Comer, F.I., Vosseller, K., Wells, L., Accavitti, M.A., Hart, G.W., **2001**. Characterization of a Mouse Monoclonal Antibody Specific for O-Linked *N*-Acetylglucosamine. *Anal. Biochem.* 293, 169–177. <https://doi.org/10.1006/abio.2001.5132>

Dai, C., Gu, J., Liu, F., Iqbal, K., Gong, C.-X., **2018**. Neuronal *O*-GlcNAc transferase regulates appetite, body weight, and peripheral insulin resistance. *Neurobiol. Aging* 70, 40–50. <https://doi.org/10.1016/j.neurobiolaging.2018.05.036>

De Leon, C.A., Levine, P.M., Craven, T.W., Pratt, M.R., **2017**. The Sulfur-Linked Analogue of *O*-GlcNAc (*S*-GlcNAc) Is an Enzymatically Stable and Reasonable Structural Surrogate for *O*-GlcNAc at the Peptide and Protein Levels. *Biochemistry* 56, 3507–3517. <https://doi.org/10.1021/acs.biochem.7b00268>

Doll, F., **2018**. Visualizing Protein-Specific Post-Translational Modifications with FLIM-FRET Microscopy. *Doctoral Thesis*.

Domenico, F.D., Zuliani, I., Perluigi, M., Lanzillotta, C., Barone, E., Tramutola, A., **2020**. Aberrant protein *O*-GlcNAcylation drives AD-like neurodegeneration in DS mice. *Alzheimers Dement.* 16, e039361. <https://doi.org/10.1002/alz.039361>

Doura, T., Matsuoka, Y., Kiyonaka, S., **2024**. Hijacking endogenous mRNA for genetic code expansion. *Nat. Chem. Biol.* 1–2. <https://doi.org/10.1038/s41589-024-01574-9>

Dunkelmann, D.L., Piedrafita, C., Dickson, A., Liu, K.C., Elliott, T.S., Fiedler, M., Bellini, D., Zhou, A., Cervettini, D., Chin, J.W., **2024**. Adding α,α -disubstituted and β -linked monomers to the genetic code of an organism. *Nature* 625, 603–610. <https://doi.org/10.1038/s41586-023-06897-6>

- Erlich, H.A. (Ed.), **1989**. PCR Technology: Principles and Applications for DNA Amplification. *Palgrave Macmillan* UK, London. <https://doi.org/10.1007/978-1-349-20235-5>
- Fardini, Y., Dehennaut, V., Lefebvre, T., Issad, T., **2013**. O-GlcNAcylation: A New Cancer Hallmark? *Front. Endocrinol.* 4. <https://doi.org/10.3389/fendo.2013.00099>
- Filius, M., van Wee, R., de Lannoy, C., Westerlaken, I., Li, Z., Kim, S.H., de Agrela Pinto, C., Wu, Y., Boons, G.-J., Pabst, M., de Ridder, D., Joo, C., **2024**. Full-length single-molecule protein fingerprinting. *Nat. Nanotechnol.* 1–8. <https://doi.org/10.1038/s41565-023-01598-7>
- Fittipaldi, A., Giacca, M., **2005**. Transcellular protein transduction using the Tat protein of HIV-1. *Adv. Drug Deliv. Rev.*, Protein- and Peptide-Mediated Transduction: Mechanisms and Implications for Drug Delivery 57, 597–608. <https://doi.org/10.1016/j.addr.2004.10.011>
- Francis, B.R., **2013**. Evolution of the Genetic Code by Incorporation of Amino Acids that Improved or Changed Protein Function. *J. Mol. Evol.* 77, 134–158. <https://doi.org/10.1007/s00239-013-9567-y>
- Fricovsky, E.S., Suarez, J., Ihm, S.-H., Scott, B.T., Suarez-Ramirez, J.A., Banerjee, I., Torres-Gonzalez, M., Wang, H., Ellrott, I., Maya-Ramos, L., Villarreal, F., Dillmann, W.H., **2012**. Excess protein O-GlcNAcylation and the progression of diabetic cardiomyopathy. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* 303, R689–R699. <https://doi.org/10.1152/ajpregu.00548.2011>
- Fu, B., Fang, L., Wang, R., Zhang, X., **2024**. Inhibition of Wnt/ β -catenin signaling by monensin in cervical cancer. *Korean J. Physiol. Pharmacol. Off. J. Korean Physiol. Soc. Korean Soc. Pharmacol.* 28, 21–30. <https://doi.org/10.4196/kjpp.2024.28.1.21>
- Fukuchi, T., Sakamoto, M., Tsuda, H., Maruyama, K., Nozawa, S., Hirohashi, S., **1998**. β -Catenin Mutation in Carcinoma of the Uterine Endometrium1. *Cancer Res.* 58, 3526–3528.
- Gao, S., Miao, Y., Liu, Y., Liu, X., Fan, X., Lin, Y., Qian, P., Zhou, J., Dai, Y., Xia, L., Zhu, P., Zhu, J., **2019**. Reciprocal Regulation Between O-GlcNAcylation and β -

Catenin Facilitates Cell Viability and Inhibits Apoptosis in Liver Cancer. *DNA Cell Biol.* 38, 286–296. <https://doi.org/10.1089/dna.2018.4447>

Gatta, E., Lefebvre, T., Gaetani, S., Dos Santos, M., Marrocco, J., Mir, A.-M., Cassano, T., Maccari, S., Nicoletti, F., Mairesse, J., **2016**. Evidence for an imbalance between tau O-GlcNAcylation and phosphorylation in the hippocampus of a mouse model of Alzheimer's disease. *Pharmacol. Res.* 105, 186–197. <https://doi.org/10.1016/j.phrs.2016.01.006>

Gibson, D.G., **2009**. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res.* 37, 6984–6990. <https://doi.org/10.1093/nar/gkp687>

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., **2009**. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. <https://doi.org/10.1038/nmeth.1318>

Gilormini, P.-A., R. Batt, A., R. Pratt, M., Biot, C., **2018**. Asking more from metabolic oligosaccharide engineering. *Chem. Sci.* 9, 7585–7595. <https://doi.org/10.1039/C8SC02241K>

Goon, S., Bertozzi, C.R., **2002**. Metabolic Substrate Engineering as a Tool for Glycobiology. *J. Carbohydr. Chem.* 21, 943–977. <https://doi.org/10.1081/CAR-120016493>

Gorelik, A., Aalten, D.M.F. van, **2020**. Tools for functional dissection of site-specific O-GlcNAcylation. *RSC Chem. Biol.* 1, 98–109. <https://doi.org/10.1039/D0CB00052C>

Gorelik, A., Bartual, S.G., Borodkin, V.S., Varghese, J., Ferenbach, A.T., van Aalten, D.M.F., **2019**. Genetic recoding to dissect the roles of site-specific protein O-GlcNAcylation. *Nat. Struct. Mol. Biol.* 26, 1071–1077. <https://doi.org/10.1038/s41594-019-0325-8>

Groves, J.A., Lee, A., Yildirim, G., Zachara, N.E., **2013**. Dynamic O-GlcNAcylation and its roles in the cellular stress response and homeostasis. *Cell Stress Chaperones* 18, 535–558. <https://doi.org/10.1007/s12192-013-0426-y>

- Guo, Q., Zhao, G., Hao, F., Guan, Y., **2012**. Effects of the TAT Peptide Orientation and Relative Location on the Protein Transduction Efficiency. *Chem. Biol. Drug Des.* 79, 683–690. <https://doi.org/10.1111/j.1747-0285.2011.01315.x>
- Gupta, S.S., Kuzelka, J., Singh, P., Lewis, W.G., Manchester, M., Finn, M.G., **2005**. Accelerated Bioorthogonal Conjugation: A Practical Method for the Ligation of Diverse Functional Molecules to a Polyvalent Virus Scaffold. *Bioconjug. Chem.* 16, 1572–1579. <https://doi.org/10.1021/bc050147l>
- Gurel, Z., Zaro, B.W., Pratt, M.R., Sheibani, N., **2014**. Identification of O-GlcNAc Modification Targets in Mouse Retinal Pericytes: Implication of p53 in Pathogenesis of Diabetic Retinopathy. *PLOS ONE* 9, e95561. <https://doi.org/10.1371/journal.pone.0095561>
- Haiber, L.M., **2023**. Metabolic Glycoengineering: A Tool to Study Protein O-GlcNAcylation. *Doctoral Thesis*.
- Hanover, J.A., Krause, M.W., Love, D.C., **2012**. linking metabolism to epigenetics through O-GlcNAcylation. *Nat. Rev. Mol. Cell Biol.* 13, 312–321. <https://doi.org/10.1038/nrm3334>
- Hao, M., Ling, X., Sun, Y., Wang, X., Li, W., Chang, L., Zeng, Z., Shi, X., Niu, M., Chen, L., Liu, T., **2024**. Tracking endogenous proteins based on RNA editing-mediated genetic code expansion. *Nat. Chem. Biol.* 1–11. <https://doi.org/10.1038/s41589-023-01533-w>
- Hao, Y., Fan, X., Shi, Y., Zhang, C., Sun, D., Qin, K., Qin, W., Zhou, W., Chen, X., **2019**. Next-generation unnatural monosaccharides reveal that ESRRB O-GlcNAcylation regulates pluripotency of mouse embryonic stem cells. *Nat. Commun.* 10, 4065. <https://doi.org/10.1038/s41467-019-11942-y>
- Hao, Y., Li, X., Qin, K., Shi, Y., He, Y., Zhang, C., Cheng, B., Zhang, X., Hu, G., Liang, S., Tang, Q., Chen, X., **2023**. Chemoproteomic and Transcriptomic Analysis Reveals that O-GlcNAc Regulates Mouse Embryonic Stem Cell Fate through the Pluripotency Network. *Angew. Chem.* 135, e202300500. <https://doi.org/10.1002/ange.202300500>
- Hardivillé, S., Hart, G.W., **2016**. Nutrient regulation of gene expression by O-GlcNAcylation of chromatin. *Curr. Opin. Chem. Biol.*, Chemical genetics and

Hart, G.W., **2014**. Three Decades of Research on O-GlcNAcylation – A Major Nutrient Sensor That Regulates Signaling, Transcription and Cellular Metabolism. *Front. Endocrinol.* 5. <https://doi.org/10.3389/fendo.2014.00183>

Hart, G.W., 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. <https://doi.org/10.1146/annurev-biochem-060608-102511>

Hassenrück, J., Wittmann, V., **2019**. Cyclopropene derivatives of aminosugars for metabolic glycoengineering. *Beilstein J. Org. Chem.* 15, 584–601. <https://doi.org/10.3762/bjoc.15.54>

Hecht, S.M., **2022**. Expansion of the Genetic Code Through the Use of Modified Bacterial Ribosomes. *J. Mol. Biol.*, Recent Advances in Genetic Code Expansion: from Cell Engineering to Protein Design 434, 167211. <https://doi.org/10.1016/j.jmb.2021.167211>

Hein, J.E., Fokin, V.V., **2010**. Copper-catalyzed azide–alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides. *Chem. Soc. Rev.* 39, 1302–1315. <https://doi.org/10.1039/B904091A>

Holt, G.D., Hart, G.W., **1986**. The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharide linkage, O-linked GlcNAc. *J. Biol. Chem.* 261, 8049–8057. [https://doi.org/10.1016/S0021-9258\(19\)57510-X](https://doi.org/10.1016/S0021-9258(19)57510-X)

Issad, T., Masson, E., Pagesy, P., **2010**. O-GlcNAc modification, insulin signaling and diabetic complications. *Diabetes Metab.* 36, 423–435. <https://doi.org/10.1016/j.diabet.2010.09.001>

Johnson, L.N., **2009**. The regulation of protein phosphorylation. *Biochem. Soc. Trans.* 37, 627–641. <https://doi.org/10.1042/BST0370627>

- J. Pedowitz, N., R. Pratt, M., **2021**. Design and synthesis of metabolic chemical reporters for the visualization and identification of glycoproteins. *RSC Chem. Biol.* 2, 306–321. <https://doi.org/10.1039/D1CB00010A>
- Ju Kim, E., **2020**. O-GlcNAc Transferase: Structural Characteristics, Catalytic Mechanism and Small-Molecule Inhibitors. *ChemBioChem* 21, 3026–3035. <https://doi.org/10.1002/cbic.202000194>
- Kasproicz, A., Spriet, C., Terryn, C., Rigolot, V., Hardville, S., Alteen, M.G., Lefebvre, T., Biot, C., **2020**. Exploring the Potential of β -Catenin O-GlcNAcylation by Using Fluorescence-Based Engineering and Imaging. *Molecules* 25, 4501. <https://doi.org/10.3390/molecules25194501>
- Kim, E.J., **2021**. Advances in Strategies and Tools Available for Interrogation of Protein O-GlcNAcylation. *ChemBioChem* 22, 3010–3026. <https://doi.org/10.1002/cbic.202100219>
- Kim, E.Y., Jeong, E.H., Park, S., Jeong, H.-J., Edery, I., Cho, J.W., **2012**. A role for O-GlcNAcylation in setting circadian clock speed. *Genes Dev.* 26, 490–502. <https://doi.org/10.1101/gad.182378.111>
- Kim, S., Maynard, J.C., Strickland, A., Burlingame, A.L., Milbrandt, J., **2018**. Schwann cell O-GlcNAcylation promotes peripheral nerve remyelination via attenuation of the AP-1 transcription factor JUN. *Proc. Natl. Acad. Sci.* 115, 8019–8024. <https://doi.org/10.1073/pnas.1805538115>
- Kim, Y., Cho, S., Kim, J.-C., Park, H.-S., **2024**. tRNA engineering strategies for genetic code expansion. *Front. Genet.* 15. <https://doi.org/10.3389/fgene.2024.1373250>
- Kubota, Y., Fujioka, K., Takekawa, M., **2017**. WGA-based lectin affinity gel electrophoresis: A novel method for the detection of O-GlcNAc-modified proteins. *PLOS ONE* 12, e0180714. <https://doi.org/10.1371/journal.pone.0180714>
- Lambert, J.-P., Ivosev, G., Couzens, A.L., Larsen, B., Taipale, M., Lin, Z.-Y., Zhong, Q., Lindquist, S., Vidal, M., Aebersold, R., Pawson, T., Bonner, R., Tate, S., Gingras, A.-C., **2013**. Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nat. Methods* 10, 1239–1245. <https://doi.org/10.1038/nmeth.2702>

- Lambert, M., Bastide, B., Cieniewski-Bernard, C., **2018**. Involvement of O-GlcNAcylation in the Skeletal Muscle Physiology and Physiopathology: Focus on Muscle Metabolism. *Front. Endocrinol.* 9. <https://doi.org/10.3389/fendo.2018.00578>
- Lee, A., Miller, D., Henry, R., Paruchuri, V.D.P., O'Meally, R.N., Boronina, T., Cole, R.N., Zachara, N.E., **2016**. Combined Antibody/Lectin Enrichment Identifies Extensive Changes in the O-GlcNAc Sub-proteome upon Oxidative Stress. *J. Proteome Res.* 15, 4318–4336. <https://doi.org/10.1021/acs.jproteome.6b00369>
- Lee, J.B., Pyo, K.-H., Kim, H.R., **2021**. Role and Function of O-GlcNAcylation in Cancer. *Cancers* 13, 5365. <https://doi.org/10.3390/cancers13215365>
- Lefebvre, T., Issad, T., **2015**. 30 Years Old: O-GlcNAc Reaches the Age of Reason – Regulation of Cell Signaling and Metabolism by O-GlcNAcylation. *Front. Endocrinol.* 6. <https://doi.org/10.3389/fendo.2015.00017>
- Levine, Z.G., Potter, S.C., Joiner, C.M., Fei, G.Q., Nabet, B., Sonnett, M., Zachara, N.E., Gray, N.S., Paulo, J.A., Walker, S., **2021**. Mammalian cell proliferation requires noncatalytic functions of O-GlcNAc transferase. *Proc. Natl. Acad. Sci.* 118, e2016778118. <https://doi.org/10.1073/pnas.2016778118>
- Li, J., Wang, J., Wen, L., Zhu, H., Li, S., Huang, K., Jiang, K., Li, X., Ma, C., Qu, J., Parameswaran, A., Song, J., Zhao, W., Wang, P.G., **2016**. An OGA-Resistant Probe Allows Specific Visualization and Accurate Identification of O-GlcNAc-Modified Proteins in Cells. *ACS Chem. Biol.* 11, 3002–3006. <https://doi.org/10.1021/acscchembio.6b00678>
- Li, Z., Yi, W., **2014**. Regulation of cancer metabolism by O-GlcNAcylation. *Glycoconj. J.* 31, 185–191. <https://doi.org/10.1007/s10719-013-9515-5>
- Lin, W., Gao, L., Chen, X., **2015**. Protein-Specific Imaging of O-GlcNAcylation in Single Cells. *ChemBioChem* 16, 2571–2575. <https://doi.org/10.1002/cbic.201500544>
- Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G.-H., Tan, Y., Zhang, Z., Lin, X., He, X., **2002**. Control of β -Catenin Phosphorylation/Degradation by a Dual-Kinase Mechanism. *Cell* 108, 837–847. [https://doi.org/10.1016/S0092-8674\(02\)00685-2](https://doi.org/10.1016/S0092-8674(02)00685-2)

- Liu, F.-T., Jiang, P.-F., Wang, Y.-P., Zhao, B.-X., Lin, Z.-M., **2024**. A ratiometric fluorescent probe based on the FRET platform for the detection of sulfur dioxide derivatives and viscosity. *Anal. Chim. Acta* 1288, 342184. <https://doi.org/10.1016/j.aca.2023.342184>
- Liu, H., Naismith, J.H., **2008**. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* 8, 91. <https://doi.org/10.1186/1472-6750-8-91>
- Liu, T.-W., Myschyshyn, M., Sinclair, D.A., Cecioni, S., Beja, K., Honda, B.M., Morin, R.D., Vocadlo, D.J., **2017**. Genome-wide chemical mapping of O-GlcNAcylated proteins in *Drosophila melanogaster*. *Nat. Chem. Biol.* 13, 161–167. <https://doi.org/10.1038/nchembio.2247>
- Ma, J., Hart, G.W., **2017**. Analysis of Protein O-GlcNAcylation by Mass Spectrometry. *Curr. Protoc. Protein Sci.* 87, 24.10.1-24.10.16. <https://doi.org/10.1002/cpps.24>
- Ma, J., Wu, C., Hart, G.W., **2021**. Analytical and Biochemical Perspectives of Protein O-GlcNAcylation. *Chem. Rev.* 121, 1513–1581. <https://doi.org/10.1021/acs.chemrev.0c00884>
- Magnusdottir, A., Johansson, I., Dahlgren, L.-G., Nordlund, P., Berglund, H., **2009**. Enabling IMAC purification of low abundance recombinant proteins from *E. coli* lysates. *Nat. Methods* 6, 477–478. <https://doi.org/10.1038/nmeth0709-477>
- Makwana, V., Ryan, P., Patel, B., Dukie, S.-A., Rudrawar, S., **2019**. Essential role of O-GlcNAcylation in stabilization of oncogenic factors. *Biochim. Biophys. Acta BBA - Gen. Subj.* 1863, 1302–1317. <https://doi.org/10.1016/j.bbagen.2019.04.002>
- Mitchell, C.W., Galan Bartual, S., Ferenbach, A.T., Scavenius, C., van Aalten, D.M.F., **2023**. Exploiting O-GlcNAc transferase promiscuity to dissect site-specific O-GlcNAcylation. *Glycobiology* 33, 1172–1181. <https://doi.org/10.1093/glycob/cwad086>
- Mohammed, M.K., Shao, C., Wang, J., Wei, Q., Wang, X., Collier, Z., Tang, S., Liu, H., Zhang, F., Huang, J., Guo, D., Lu, M., Liu, F., Liu, J., Ma, C., Shi, L.L., Athiviraham, A., He, T.-C., Lee, M.J., **2016**. Wnt/ β -catenin signaling plays an ever-expanding role in stem cell self-renewal, tumorigenesis and cancer

chemoresistance. *Genes Dis.* 3, 11–40.
<https://doi.org/10.1016/j.gendis.2015.12.004>

Monga, S.P., **2015**. β -Catenin Signaling and Roles in Liver Homeostasis, Injury, and Tumorigenesis. *Gastroenterology* 148, 1294–1310.
<https://doi.org/10.1053/j.gastro.2015.02.056>

Muha, V., Williamson, R., Hills, R., McNeilly, A.D., McWilliams, T.G., Alonso, J., Schimpl, M., Leney, A.C., Heck, A.J.R., Sutherland, C., Read, K.D., McCrimmon, R.J., Brooks, S.P., van Aalten, D.M.F., **2019**. Loss of CRMP2 O-GlcNAcylation leads to reduced novel object recognition performance in mice. *Open Biol.* 9, 190192. <https://doi.org/10.1098/rsob.190192>

Nagano, M., Toshima, J.Y., Siekhaus, D.E., Toshima, J., **2019**. Rab5-mediated endosome formation is regulated at the trans-Golgi network. *Commun. Biol.* 2, 1–12. <https://doi.org/10.1038/s42003-019-0670-5>

Nair, D.P., Podgórski, M., Chatani, S., Gong, T., Xi, W., Fenoli, C.R., Bowman, C.N., **2014**. The Thiol-Michael Addition Click Reaction: A Powerful and Widely Used Tool in Materials Chemistry. *Chem. Mater.* 26, 724–744.
<https://doi.org/10.1021/cm402180t>

Nödling, A.R., Spear, L.A., Williams, T.L., Luk, L.Y.P., Tsai, Y.-H., **2019**. Using genetically incorporated unnatural amino acids to control protein functions in mammalian cells. *Essays Biochem.* 63, 237–266.
<https://doi.org/10.1042/EBC20180042>

Ohya, T., Miaczynska, M., Coskun, Ü., Lommer, B., Runge, A., Drechsel, D., Kalaidzidis, Y., Zerial, M., **2024**. Author Correction: Reconstitution of Rab- and SNARE-dependent membrane fusion by synthetic endosomes. *Nature* 626, E8–E9. <https://doi.org/10.1038/s41586-024-07063-2>

Olivier-Van Stichelen, S., Dehennaut, V., Buzy, A., Zachayus, J.-L., Guinez, C., Mir, A.-M., El Yazidi-Belkoura, I., Copin, M.-C., Boureme, D., Loyaux, D., Ferrara, P., Lefebvre, T., **2014**. O-GlcNAcylation stabilizes β -catenin through direct competition with phosphorylation at threonine 41. *FASEB J.* 28, 3325–3338.
<https://doi.org/10.1096/fj.13-243535>

- Paneque, A., Fortus, H., Zheng, J., Werlen, G., Jacinto, E., **2023**. The Hexosamine Biosynthesis Pathway: Regulation and Function. *Genes* 14, 933. <https://doi.org/10.3390/genes14040933>
- Parker, M.P., Peterson, K.R., Slawson, C., **2021**. O-GlcNAcylation and O-GlcNAc Cycling Regulate Gene Transcription: Emerging Roles in Cancer. *Cancers* 13, 1666. <https://doi.org/10.3390/cancers13071666>
- Patel, S.G., Sayers, E.J., He, L., Narayan, R., Williams, T.L., Mills, E.M., Allemann, R.K., Luk, L.Y.P., Jones, A.T., Tsai, Y.-H., **2019**. Cell-penetrating peptide sequence and modification dependent uptake and subcellular distribution of green fluorescent protein in different cell lines. *Sci. Rep.* 9, 6298. <https://doi.org/10.1038/s41598-019-42456-8>
- Pellegrini, F.R., De Martino, S., Fianco, G., Ventura, I., Valente, D., Fiore, M., Trisciuglio, D., Degrassi, F., **2023**. Blockage of autophagosome-lysosome fusion through SNAP29 O-GlcNAcylation promotes apoptosis via ROS production. *Autophagy* 19, 2078–2093. <https://doi.org/10.1080/15548627.2023.2170962>
- Perdiguero, A.N., Liang, A.D., **2024**. Practical Approaches to Genetic Code Expansion with Aminoacyl-tRNA Synthetase/tRNA Pairs. *CHIMIA* 78, 22–31. <https://doi.org/10.2533/chimia.2024.22>
- Peterson, S.B., Hart, G.W., **2016**. New insights: A role for O-GlcNAcylation in diabetic complications. *Crit. Rev. Biochem. Mol. Biol.* 51, 150–161. <https://doi.org/10.3109/10409238.2015.1135102>
- Petta, I., Dejager, L., Ballegeer, M., Lievens, S., Tavernier, J., De Bosscher, K., Libert, C., **2016**. The Interactome of the Glucocorticoid Receptor and Its Influence on the Actions of Glucocorticoids in Combatting Inflammatory and Infectious Diseases. *Microbiol. Mol. Biol. Rev.* 80, 495–522. <https://doi.org/10.1128/mnbr.00064-15>
- Queiroz, R.M. de, Madan, R., Chien, J., Dias, W.B., Slawson, C., **2016**. Changes in O-Linked N-Acetylglucosamine (O-GlcNAc) Homeostasis Activate the p53 Pathway in Ovarian Cancer Cells. *J. Biol. Chem.* 291, 18897–18914. <https://doi.org/10.1074/jbc.M116.734533>

- Rafie, K., Raimi, O., Ferenbach, A.T., Borodkin, V.S., Kapuria, V., van Aalten, D.M.F., **2017**. Recognition of a glycosylation substrate by the O-GlcNAc transferase TPR repeats. *Open Biol.* 7, 170078. <https://doi.org/10.1098/rsob.170078>
- Ramirez, J., Petersson, E.J., Rhoades, E., **2024**. Investigating the differential effects of authentic site-specific post-translational modifications on tau. *Biophys. J.* 123, 70a. <https://doi.org/10.1016/j.bpj.2023.11.501>
- Rexach, J.E., Clark, P.M., Mason, D.E., Neve, R.L., Peters, E.C., Hsieh-Wilson, L.C., **2012**. Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nat. Chem. Biol.* 8, 253–261. <https://doi.org/10.1038/nchembio.770>
- Rigolot, V., Biot, C., Lion, C., **2021**. To View Your Biomolecule, Click inside the Cell. *Angew. Chem. Int. Ed.* 60, 23084–23105. <https://doi.org/10.1002/anie.202101502>
- Rogerson, D.T., Sachdeva, A., Wang, K., Haq, T., Kazlauskaitė, A., Hancock, S.M., Huguenin-Dezot, N., Muqit, M.M.K., Fry, A.M., Bayliss, R., Chin, J.W., **2015**. Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.* 11, 496–503. <https://doi.org/10.1038/nchembio.1823>
- Rust, A., Hassan, H.H.A., Sedelnikova, S., Niranjana, D., Hautbergue, G., Abbas, S.A., Partridge, L., Rice, D., Binz, T., Davletov, B., **2015**. Two complementary approaches for intracellular delivery of exogenous enzymes. *Sci. Rep.* 5, 12444. <https://doi.org/10.1038/srep12444>
- Saha, A., Bello, D., Fernández-Tejada, A., **2021**. Advances in chemical probing of protein O-GlcNAc glycosylation: structural role and molecular mechanisms. *Chem. Soc. Rev.* 50, 10451–10485. <https://doi.org/10.1039/D0CS01275K>
- Saunders, H., Dias, W.B., Slawson, C., **2023**. Growing and dividing: how O-GlcNAcylation leads the way. *J. Biol. Chem.* 299, 105330. <https://doi.org/10.1016/j.jbc.2023.105330>
- Sayat, R., Leber, B., Grubac, V., Wiltshire, L., Persad, S., **2008**. O-GlcNAc-glycosylation of β -catenin regulates its nuclear localization and transcriptional activity. *Exp. Cell Res.* 314, 2774–2787. <https://doi.org/10.1016/j.yexcr.2008.05.017>

- Schmidt, M.J., Summerer, D., **2014**. Genetic code expansion as a tool to study regulatory processes of transcription. *Front. Chem.* 2. <https://doi.org/10.3389/fchem.2014.00007>
- Shandell, M.A., Tan, Z., Cornish, V.W., **2021**. Genetic Code Expansion: A Brief History and Perspective. *Biochemistry* 60, 3455–3469. <https://doi.org/10.1021/acs.biochem.1c00286>
- Shao, S., Chen, B., Cheng, J., Wang, C., Zhang, Y., Shao, L., Hu, Y., Han, Y., Han, F., Li, X., **2017**. A fluorogenic probe for imaging protein S-nitrosylation in live cells. *Biosens. Bioelectron.* 94, 162–168. <https://doi.org/10.1016/j.bios.2017.02.050>
- Sheikh, M.A., Emerald, B.S., Ansari, S.A., **2021**. Stem cell fate determination through protein O-GlcNAcylation. *J. Biol. Chem.* 296. <https://doi.org/10.1074/jbc.REV120.014915>
- Söderberg, O., Leuchowius, K.-J., Gullberg, M., Jarvius, M., Weibrecht, I., Larsson, L.-G., Landegren, U., **2008**. Characterizing proteins and their interactions in cells and tissues using the *in situ* proximity ligation assay. *Methods, Visualization of molecular processes in cells using fluorescence imaging* 45, 227–232. <https://doi.org/10.1016/j.ymeth.2008.06.014>
- Sodi, V.L., Khaku, S., Krutilina, R., Schwab, L.P., Vocadlo, D.J., Seagroves, T.N., Reginato, M.J., **2015**. mTOR/MYC Axis Regulates O-GlcNAc Transferase Expression and O-GlcNAcylation in Breast Cancer. *Mol. Cancer Res.* 13, 923–933. <https://doi.org/10.1158/1541-7786.MCR-14-0536>
- Stephen, H.M., Adams, T.M., Wells, L., **2021**. Regulating the Regulators: Mechanisms of Substrate Selection of the O-GlcNAc Cycling Enzymes OGT and OGA. *Glycobiology* 31, 724–733. <https://doi.org/10.1093/glycob/cwab005>
- Swamy, M., Pathak, S., Grzes, K.M., Damerow, S., Sinclair, L.V., van Aalten, D.M.F., Cantrell, D.A., **2016**. Glucose and glutamine fuel protein O-GlcNAcylation to control T cell self-renewal and malignancy. *Nat. Immunol.* 17, 712–720. <https://doi.org/10.1038/ni.3439>
- Takahashi, M., Nakatsugi, S., Sugimura, T., Wakabayashi, K., **2000**. Frequent mutations of the β -catenin gene in mouse colon tumors induced by

azoxymethane. *Carcinogenesis* 21, 1117–1120.
<https://doi.org/10.1093/carcin/21.6.1117>

Tan, E.P., Duncan, F.E., Slawson, C., **2017**. The sweet side of the cell cycle. *Biochem. Soc. Trans.* 45, 313–322. <https://doi.org/10.1042/BST20160145>

Ubersax, J.A., Ferrell Jr, J.E., **2007**. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* 8, 530–541.
<https://doi.org/10.1038/nrm2203>

Vaidyanathan, K., Durning, S., Wells, L., **2014**. Functional O-GlcNAc modifications: Implications in molecular regulation and pathophysiology. *Crit. Rev. Biochem. Mol. Biol.* 49, 140–163. <https://doi.org/10.3109/10409238.2014.884535>

Vocero-Akbani, A., Lissy, N.A., Dowdy, S.F., **2000**. [44] - Transduction of Full-Length Tat Fusion Proteins Directly into Mammalian Cells: Analysis of T Cell Receptor Activation-Induced Cell Death, in: Reed, J.C. (Ed.), *Methods in Enzymology, Apoptosis*. Academic Press, pp. 508–521. [https://doi.org/10.1016/S0076-6879\(00\)22046-6](https://doi.org/10.1016/S0076-6879(00)22046-6)

Wang, J., Schiller, S.M., Schultz, P.G., **2007**. A Biosynthetic Route to Dehydroalanine-Containing Proteins. *Angew. Chem.* 119, 6973–6975.
<https://doi.org/10.1002/ange.200702305>

Wang, L., Brock, A., Herberich, B., Schultz, P.G., **2001**. Expanding the Genetic Code of *Escherichia coli*. *Science* 292, 498–500.
<https://doi.org/10.1126/science.1060077>

Wang, L., Schultz, P.G., **2005**. Expanding the Genetic Code. *Angew. Chem. Int. Ed.* 44, 34–66. <https://doi.org/10.1002/anie.200460627>

Wang, S.-S., Gao, X., Solar, V.D., Yu, X., Antonopoulos, A., Friedman, A.E., Matich, E.K., Atilla-Gokcumen, G.E., Nasirikenari, M., Lau, J.T., Dell, A., Haslam, S.M., Laine, R.A., Matta, K.L., Neelamegham, S., **2018**. Thioglycosides Are Efficient Metabolic Decoys of Glycosylation that Reduce Selectin Dependent Leukocyte Adhesion. *Cell Chem. Biol.* 25, 1519-1532.e5.
<https://doi.org/10.1016/j.chembiol.2018.09.012>

- Wheatley, E.G., Albarran, E., White, C.W., Bieri, G., Sanchez-Diaz, C., Pratt, K., Snethlage, C.E., Ding, J.B., Villeda, S.A., **2019**. Neuronal O-GlcNAcylation Improves Cognitive Function in the Aged Mouse Brain. *Curr. Biol.* 29, 3359-3369.e4. <https://doi.org/10.1016/j.cub.2019.08.003>
- Wright, J.N., Collins, H.E., Wende, A.R., Chatham, J.C., **2017**. O-GlcNAcylation and cardiovascular disease. *Biochem. Soc. Trans.* 45, 545–553. <https://doi.org/10.1042/BST20160164>
- Wu, C., Li, Jiaheng, Lu, L., Li, M., Yuan, Y., Li, Jing, **2024**. OGT and OGA: Sweet Guardians of the Genome. *J. Biol. Chem.* 107141. <https://doi.org/10.1016/j.jbc.2024.107141>
- Wu, D., Cai, Y., Jin, J., **2017**. Potential coordination role between O-GlcNAcylation and epigenetics. *Protein Cell* 8, 713–723. <https://doi.org/10.1007/s13238-017-0416-4>
- Yang, X., Qian, K., **2017**. Protein O-GlcNAcylation: emerging mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* 18, 452–465. <https://doi.org/10.1038/nrm.2017.22>
- Yigzaw, Y., Cartin, L., Pierre, S., Scholich, K., Patel, T.B., **2001**. The C Terminus of Sprouty Is Important for Modulation of Cellular Migration and Proliferation. *J. Biol. Chem.* 276, 22742–22747. <https://doi.org/10.1074/jbc.M100123200>
- Zachara, N.E., Hart, G.W., **2006**. Cell signaling, the essential role of O-GlcNAc! *Biochim. Biophys. Acta BBA - Mol. Cell Biol.* Lipids, Lipid Signalling in the Nucleus 1761, 599–617. <https://doi.org/10.1016/j.bbaliip.2006.04.007>
- Zaro, B.W., Batt, A.R., Chuh, K.N., Navarro, M.X., Pratt, M.R., **2017**. The Small Molecule 2-Azido-2-deoxy-glucose Is a Metabolic Chemical Reporter of O-GlcNAc Modifications in Mammalian Cells, Revealing an Unexpected Promiscuity of O-GlcNAc Transferase. *ACS Chem. Biol.* 12, 787–794. <https://doi.org/10.1021/acscchembio.6b00877>
- Zhang, H., Li, Z., Wang, Y., Kong, Y., **2021**. O-GlcNAcylation is a key regulator of multiple cellular metabolic pathways. *PeerJ* 9, e11443. <https://doi.org/10.7717/peerj.11443>

Zhao, N., Kamijo, K., Fox, P.D., Oda, H., Morisaki, T., Sato, Y., Kimura, H., Stasevich, T.J., **2019**. A genetically encoded probe for imaging nascent and mature HA-tagged proteins in vivo. *Nat. Commun.* 10, 2947. <https://doi.org/10.1038/s41467-019-10846-1>