HABILITATION A DIRIGER DES RECHERCHES

présentée à

L'Université des Sciences et Technologies de Lille

par

Yann Guérardel

Diversité Structurale des Manno-conjugués Microbiens

et

Mise en Place de Nouveaux Modèles d'Etude de la Sialylation

Soutenue le 24 octobre 2008 devant la commission d'examen :

Président :	Docteur J.C. MICHALSKI
Rapporteurs :	Professeur P. LEROUGE
	Professeur D. POULAIN
	Docteur M. RIVIERE
Membre invité :	Docteur G. STRECKER

Résumé

Les glycannes font montre d'une extraordinaire diversité structurale, et ce à tous les niveaux d'organisation biologique. Les travaux présentés sont centrés sur la notion de spécificité d'espèce de la glycosylation, au travers de l'étude de plusieurs modèles procaryotes et eucaryotes.

Dans une première partie, nous exposons nos travaux portant sur les différentes formes de mannosylation chez les eucaryotes unicellulaires et les bactéries. L'utilisation de modèles aussi divers que les mycobactéries, les levures pathogènes et les eucaryotes unicellulaires parasites, nous a permis de mettre en lumière l'évolution des structures et éventuellement des fonctions associées aux manno-conjugués.

Dans une deuxième partie, nous exposons nos travaux portant sur la comparaison des glycosyltransférases animales. Leur étude est en effet devenue un enjeu majeur pour la compréhension des rôles des glycoconjugués dans l'embryogenèse, la fécondation ou les réactions immunitaires. Nos travaux consistent à mettre en évidence de nouvelles activités glycosyltransférasiques sur des modèles animaux dont le génome, la physiologie et le développement sont suffisamment connus pour qu'ils servent de plateforme à l'étude du rôle de la glycosylation dans les processus physiologiques précités. Les espèces animales principales répondant à ces critères et que nous avons étudiées au cours des dernières années sont *Xenopus laevis*, *X.tropicalis*, *Caenorhabditis elegans* et *Danio rerio*. Dans ce cadre, nous présenterons nos travaux destinés à promouvoir le poisson zèbre (*D. rerio*) comme modèle d'étude cohérent en glycobiologie.

Enfin, dans une dernière partie sont compilés les résultats des développements technologiques portant sur la purification et l'analyse des glycoconjugués nécessaires à nos études.

Je tiens à remercier

Gérard Strecker pour m'avoir accueilli dans son équipe, initié à la glycobiologie structurale et avoir constamment soutenu mes recherches. Il sera toujours pour moi un modèle d'honnêteté scientifique, de dévouement à son art et de modestie.

Jean-Claude Michalski pour avoir soutenu mon intégration au CNRS et m'avoir accordé sa confiance. Il m'a donné les moyens et toute latitude pour développer mes recherches au sein de l'institut.

Les membres de l'équipe pour le superbe travail qu'ils accomplissent quotidiennement et la cohésion dont ils font preuve. En particulier, je désire remercier Emmanuel Maes, mon camarade de tous les instants, et Bernadette Coddeville, qui grâce à leur travail acharné et leur générosité scientifique soutiennent l'édifice de notre groupe. Je remercie également les nouveaux venus, Elisabeth Ellas, Emeline Fabre, Frédéric Krzewinski, Colette Brassard et Eliane Aïssi pour la confiance qu'ils m'ont témoigné en intégrant notre équipe et pour toutes les compétences et la bonne humeur qu'ils y amènent. Je remercie Estelle Garénaux, Lan-Yi Chang et Yoann Rombouts de m'avoir confié une partie de leur éducation et pour l'excellent travail de thèse qu'ils ont effectué. Je remercie enfin Xavier Trivelli pour collaborer activement avec notre groupe et avoir intégré notre famille.

*L'ensemble des membres de l'UGSF qui participent tous à la cohésion et au bon fonctionnement de l'unité. Je remercie personnellement Christophe Mariller, Frédéric Chirat, Tony Lefebvre, Fanny Bonachera, Yves Leroy, Fabrice Allain, Dominique Legrand, René Cacan, Philippe Delannoy, Anne-Harduin Lepers, Jean-Marie Lacroix et Joël Mazurier pour m'avoir soutenu de mille et une façons différentes. Je tiens à témoigner mon attachement à un certain '**esprit C9**' qui se maintient grâce aux efforts des directeurs successifs et à un consensus général, malgré les pressions externes comme internes. Il est fait de partage, de générosité, d'échanges scientifiques et de la reconnaissance que l'intérêt commun prévaut sur les intérêts particuliers. Ce cap est souvent difficile à maintenir mais est garant d'un développement durable collectif de la glycobiologie face à la montée croissante des individualités.*

Mes anciens patrons dans des laboratoires étrangers, Mike Elliot (UK), Kay-Hooi Khoo (Taiwan) and Kitajima sensei (Japon), qui m'ont chacun ouverts de nouveaux horizons et de nouvelles méthodes de travail.

Enfin, mes innombrables collaborateurs extérieurs, parmi lesquels Laurent Kremer, Daniel Poulain, Christine faille, Vincent Gloagen, Christine Ferrier et Hitochi Sawada.

Présentation

Yann Guérardel

7 rue Saint Saëns,
59790 Ronchin,
Né le 1er Nov. 1972
yann.guerardel@univ-lille1.fr

Fonction actuelle

Chargé de Recherche (CR1) au CNRS UMR 8576, USTL, Villeneuve d'Ascq
Depuis Sept 2006, responsable de l'équipe 'Diversité associée aux Glycoconjugués'

Diplômes de l'enseignement supérieur

2002 : **Doctorat de Biochimie** dans le Laboratoire de Glycobiologie Structurale et Fonctionnelle de Lille, CNRS UMR 8576, équipe du **Dr. G. Strecker**
'Variabilité structurale et fonctionnelle des glycoconjugués. *Etudes réalisées chez les mycobactéries, les nématodes, les éponges et les amphibiens*'

1999 : **DEA** des Sciences de la Vie et de la Santé. Université de Sciences et Technologies de Lille (USTL)

1998 : **Maîtrise** de Biochimie. USTL

1996 : Cours de troisième année en Biochimie et Ecotoxicologie marine, dans le cadre du programme ERASMUS. University of Hull, **U.K**.

Expérience

2004- Chargé de Recherche (CR) au CNRS UMR 8576
2006 Professeur invité à l'Université de Nagoya, **JAPON** durant 5 mois.
2003-2004 Stage postdoctoral dans l'équipe du Dr. K.H. Khoo, AS, **TAIWAN**.
2002 Assistant Temporaire d'Enseignement et de Recherche, USTL
1996-1998 Technicien de recherche dans le Laboratoire de Glycobiologie Structurale et Fonctionnelle, CNRS UMR 8576.

Thématiques scientifiques

- Relations structure-fonction des glycoconjugués microbiens
- Spécificité des glycosyltransférases

Compétences scientifiques

- Purification et analyse structurale de glycoconjugués : lipopolysaccharides bactériens, polysaccharides, glycoprotéines et glycolipides d'eucaryotes
- Techniques analytiques et préparatives de chromatographie et d'électrophorèse
- Chromatographie en phase gazeuse, spectrométrie de masse, résonance magnétique nucléaire

Implications dans des contrats de recherche récents

Juil. 2004 **Vaincre la Mucoviscidose**, 'Relations entre profils de glycosylation individuels des mucines bronchiales et infection par *Pseudomonas aeruginosae*'

Dec. 2005 **ANR** (2006-2008), 'Roles of specific mycobacterial cell wall components in tuberculous disease: an *in vivo* analysis using purified cell wall molecules from *mycobacterium marinum* in the zebrafish model'

Janv. 2006 **Contrat de Plan Etat-Région** – 2ème tranche (2005-2006) 'Variabilité intraspécifique de deux pathogènes (*Listeria monocytogenes*, spores de *Bacillus cereus*) et aptitude à coloniser les aliments et les surfaces en environnement défavorable'

Sept. 2007 **ANR** (2007-2009), 'Modélisation de la dynamique des interactions spores de Bacillus/matériau. Rôle de la complexité de surface des spores, application à *Listeria monocytogenes*, pathogène non sporulé'

Communications orales récentes

Juin 2008	Belgian Biophysical Society Summer school on Glycobiology; Bruxelles, conférencier invité
Mai 2008	Groupe Français de Glucides, Ax-les-Thermes
Dec 2007	EMBO Workshop GlycoDev, Lille, conférencier invité
Oct 2007	Franco-Korean Sialobiology meting, Seoul, Corée, conférencier invité
Sept 2007	EuroCarb, Lubeck, Allemagne, conférencier invité
Sept 2007	Du Glycosynthon à la Glycoprotéine, Orléans, conférencier invité

Divers

Anglais courant
Chinois parlé (Mandarin classique) niveau débutant
Japonais parlé et lu niveau débutant

Liste des publications

Publiées

1. Chang LY, Harduin-Lepers A, Kitajima K, Sato C, Huang CJ, Khoo KH Guérardel Y Developmental regulation of oligosialylation in zebrafish. ***Glycoconj. J***. *Sous presse*

2. Chang LY, Mir AM, Thisse C, Guérardel Y, Delannoy P, Thisse B, Harduin-Lepers A Molecular cloning and characterization of the expression pattern of the zebrafish α2,8-sialyltransferases (ST8Sia) in the developing nervous system. ***Glycoconj. J***. *Sous presse*

3. Garenaux E, Yu SY, Florea D, Strecker G, Khoo KH, Guérardel Y. Single step method for purification of sulfated oligosaccharides. ***Glycoconj. J*** *Sous presse*

4. Ojha AK, Baughn AD, Sambandan D, Hsu T, Trivelli X, Guerardel Y, Alahari A, Kremer L, Jacobs WR, Graham F *Mycobacterium tuberculosis* mycolate-rich biofilms harbor drug tolerant bacteria. ***Mol. Mic***. 2008, 69, 164-74.

5. Gloaguen V, Krausz P, Brudieux V, Closs B, Leroy Y, Guerardel Y Structural Patterns of Rhamnogalacturonans Modulating In vitro Human Keratinocyte Differentiation. ***Molecules*** 2008, 13, 1207-1218

6. Mille C, Bobrowicz P, Trinel PA, Li H, Maes E, Guerardel Y, Fradin C, Martínez-Esparza M, Davidson RC, Janbon G, Poulain D, Wildt S. Identification of a new family of genes involved in beta-1,2 mannosylation of glycans in *Pichia pastoris* and *Candida albicans*. **J Biol Chem**. 2008, 283, 9724-36

7. Elass E., Coddeville B., Kremer L., Mortuaire M., Mazurier J., Guérardel Y. Mycobacterial lipomannan induces MAP kinase phosphatase-1 expression in macrophages. **FEBS Lett.**, 2008, 582, 445-450

8. Alahari A., Trivelli X., Guerardel Y., Dover L.G., Besra G.S., Sacchettini J.C., Reynolds R.C., Coxon G.D., Kremer L. Thiacetazone, an Antitubercular Drug that Inhibits Cyclopropanation of Cell Wall Mycolic Acids in Mycobacteria. **PLoS ONE**, 2007, **2** (12), e1343

9. Maes, E, Coddeville, B, Kremer, L, Guérardel, Y. (2007) Polysaccharide structural variability in mycobacteria: identification and characterization of phosphorylated mannan and arabinomannan. ***Glycoconj J***. 24(8):439-48

10. Gutierrez, A.L., Farage, L., Melo, M.N., Mohana-Borges, R.S., Guérardel, Y., Coddeville, B., Wieruszeski, J.M., Mendonca-Previato, L., Previato, J.O. (2007) Characterization of glycoinositolphosphoryl ceramide structure mutant strains of *Cryptococcus neoformans*. ***Glycobiology***. 17, 1-11C.

11. Elass, E., Coddeville, B., Guérardel, Y., Kremer, L., Maes, E., Mazurier, J., Legrand, D. (2007) Identification by surface plasmon resonance of the mycobacterial lipomannan and lipoarabinomannan domains involved in binding to CD14 and LPS-binding protein. ***FEBS Lett.*** 581, 1383-1390

12. Puissegur, MP.,Lay, G., Gilleron, M., Botella, L., Nigou, J., Marrakchi, H., Mari, B., Duteyrat, J.L., Guérardel, Y., Kremer, L., Barbry, P., Puzo, G., Altare, F. (2007) Mycobacterial lipomannan induces granuloma macrophage fusion via a TLR2-dependent, ADAM9- and beta1 integrin-mediated pathway. ***J. Immunol***. 178, 3161-3169

13. Taupin,V., Garenaux, E., Mazet, M., Maes, E., Dense, H., Prensier, G., Vivares, C.P., Guérardel, Y. and Metenier G. (2007) Major O-glycans in the spores of two microsporidian parasites are represented by unbranched manno-oligosaccharides containing α-1,2 linkages. ***Glycobiology***, 17, 56-67

14. Florea, .D, Maes, E., Guérardel, Y., Page, A., Zanetta, J.P., Cogalniceanu, D., and Strecker G. (2006) Structure elucidation of NeuAc, NeuGc and Kdn-containing O-glycans released from Triturus alpestris oviductal mucins : Characterization of the poly LacdiNAc sequence: HSO3(4)(GalNAcbeta1-4GlcNAcbeta1-3)(1-3)GalNAcbeta1-4(GlcNAcbeta1-3) (0-1)GlcNAcbeta1-6GalNAc-ol. ***Glycoconj J.*** 23, 377-399.

15. Treffort, N., Dubreucq, G., Canu, M.H., Guerardel, Y., Falempin, M. and Picquet, F. (2006) Variations in amino acid neurotransmitters in the rat ventral spinal cord after hindlimb unloading. *Neurosci. Lett*. 403, 147-150.

16. Canu, M.H., Treffort, N., Picquet, F., Dubreucq, G., Guerardel, Y., Falempin, M. (2006) Concentration of amino acid neurotransmitters in the somatosensory cortex of the rat after surgical or functional deafferentation. ***Exp. Brain. Res***. 173, 623-628

17. Guerardel, Y., Chang, L.Y., Maes, E., Huang, C.J., Khoo, K.H. (2006) Glycomic survey mapping of zebrafish identifies unique sialylation pattern. ***Glycobiology***. 16, 244-257

18. Elass, E., Aubry, L., Masson, M., Denys, A., Guerardel, Y., Maes, E., Legrand, D., Mazurier, J., Kremer, L. (2005) Mycobacterial lipomannan induces matrix metalloproteinase-9 expression in human macrophagic cells through a Toll-like receptor 1 (TLR1)/TLR2- and CD14-dependent mechanism. ***Infect. Immun***. 73, 7064-7068

19. Teintenier-Lelievre, M., Julien, S., Juliant, S., Guerardel, Y., Duonor-Cerutti, M., Delannoy, P., Harduin-Lepers, A. (2005) Molecular cloning and expression of a human hST8Sia VI (alpha2,8-sialyltransferase) responsible for the synthesis of the diSia motif on O-glycosylproteins. ***Biochem. J.*** 392, 665-674

20. Maes, E., Garenaux, E., Strecker, G., Leroy, Y., Wieruszeski, J.M., Brassart, C., Guerardel, Y. (2005) Major O-glycans from the nest of Vespula germanica contain phospho-ethanolamine. *Carbohydr. Res*. 340, 1852-1858

21. Coppin, A., Varre, J.S., Lienard, L., Dauvillee, D., Guerardel, Y., Soyer-Gobillard, M.O., Buleon, A., Ball, S., Tomavo, S. (2005) Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite Toxoplasma gondii argues for a red alga ancestry. ***J. Mol. Evol***. 60, 257-267

22. Guerardel, Y., Leleu, D., Coppin, A., Lienard, L., Slomianny, C., Strecker, G., Ball, S. and Tomavo S (2005) Amylopectin biogenesis and characterization in the protozoan parasite

Toxoplasma gondii, the intracellular development of which is restricted in the HepG2 cell line. ***Microbes Infect.*** 7, 41-48

23. D. Dao, L. Kremer, Y. Guérardel, A. Molano, W.R. Jacobs Jr., S.A. Porcelli and V. Briken (2004) *Mycobacterium tuberculosis* lipomannan induces apoptosis and Il-12 production of macrophages. ***Infect. Immun***. 72, 2067-2074

24. Quesniaux, V.; Nicolle, D.; Torres, D.; Kremer, L.; Guérardel, Y.; Nigou, J.; Puzo, G. ; Erard, F. and Ryffel, B. (2004) Toll-Like Receptor (TLR)-dependent positive and TLR-independent negative regulation of proinflammatory cytokine production in macrophages by mycobacterial lipomannans. ***J. Immunol.***172, 4425-4434

25. Guérardel, Y.; Czeszak, X; Sumanovski, L.; Karamanos, Y; Popescu, P.; Strecker, G. and Misevic, GN. (2004) Molecular Fingerprinting of carbohydrate structures phenotypes of three porifera proteoglycan-like glyconectins. ***J. Biol. Chem***., 279, 15591-15603

26. Misevic, GN.; Guérardel, Y.; Sumanovski, L.; Slomianny, M.C.; Demarty, M.; Ripoll, C.; Karamanos, Y; Maes, E.; Popescu, P. and Strecker, G. (2004) Molecular recognition between glyconectins as a adhesion self-assembly pathway to multicellularity. ***J. Biol. Chem.,*** 279, 15579-15590

27. Guérardel, Y., Petit, D.; Madigou, B., Guillet, T, Maes, E., Maftah, A., Boujard, D., Strecker, G. and Kol, O. (2003) Identification of the blood group Lewis^a determinant in the oviducal mucins of *Xenopus tropicalis*. ***FEBS letters*** 554, 330-336

28. Guérardel, Y.; Maes, E.; Bricken, V., Chirat, F., Leroy, Y; Locht, C.; Strecker, G. and Kremer, L. (2003) Lipomannan and lipoarabinomannan from a clinical isolate of Mycobacterium kansasii: novel structural features and apoptosis-inducing properties. ***J. Biol. Chem***. 278, 36637-36651

29. Vignal, C.; Guérardel, Y.; Kremer, L.; Masson, M.; Legrand, D.; Mazurier, J. and Elass, E. (2003) Lipomannans, but not lipoarabinomannans, purified from *Mycobacterium chelonae* and *Mycobacterium kansasii* induce TNF-α and IL-8 secretion by a CD14-Toll-Like receptor 2-dependent mechanism. ***J. Immunol***. 171, 2014-2023

30. Guérardel, Y.; Maes, E.; Elass, E., Leroy, Y; Timmerman, P., Besra, G.S; Locht, C.; Strecker, G. and Kremer, L. (2002) Structural study of lipomannan and lipoarabinomannan from *Mycobacterium chelonae*: presence of unusual components with α1,3 mannopyranose side chains. ***J. Biol. Chem***. 277, 30635-30648

31. Kremer, L.; Guérardel, Y.; Gursha, S.; Locht, C. and Besra, G.S. (2002) Temperature induced changes in the cell wall components of *M. thermoresistibile*. ***Microbiology*** 148, 3145-54

32. Lefebvre, T.; Cieniewski, C.; Lemoine, J.; Guérardel, Y.; Leroy, Y.; Zanetta, J.P. and Michalski, J.C. (2001) Identification of GlcNAc specific lectins from rat liver cytosolic and nuclear compartments as heat-shock proteins. ***Biochem. J***. 360, 179-188

33. Guérardel, Y.; Balanzino, L.; Maes, E.; Leroy, Y.; Coddeville, B.; Oriol, R.; and Strecker, G. (2001) The nematode *Caenorhabditis elegans* synthesises unusual O-linked glycans:

identification of glucose-substituted mucin-type O-glycans and short chondroitin-like oligosaccharides. ***Biochem. J***. 357, 167-182

34. Guérardel, Y.; Kol, O.; Maes, E.; Lefebvre, T.; Boilly, B.; Davril, M.; Strecker, G. (2000) O-glycan variability of egg-jelly mucins from *Xenopus laevis* : characterization of four phenotypes that differ by the terminal glycosylation of their mucins. ***Biochem. J***. 352, 449-463

35. Guérardel, Y.; Morelle, W.; Plancke, Y.; Lemoine, J.; and Strecker, G. (1999) Structural analysis of three sulfated oligosaccharides isolated from human milk. ***Carbohydr. Res.*** 320, 230-238

Soumises/en révision

1. Mille C, Trinel PA, Fradin C, Delplace F, Bobrowicz P, Coddeville B, Guerardel Y, Wildt S, Janbon G, Poulain D. *Candida albicans* BMT5 and BMT6 genes are involved in β-mannosylation of phospholipomannan

2. Rivière C, Goossens L, Guerardel Y, Maes E, Garénaux E, Pommery N, Lemoine A, Tellieza A, Pommery J, Désiré O, Delelis A, Hénichart JP. Anti-inflammatory activities of iridoids and polyphenols isolated from *Perichlaena richardii* Baill

3. Tremblay P, Weinbauer MG, Rottier C, Guérardel Y, Nozais C, Ferrier-Pagès C Compartmental and species specific differences in bacterial diversity associated with the tissue and skeleton of three scleractinian coral species: *Galaxea fascicularis*, *Pavona cactus* and *Turbinaria reniformis*

4. Driss V, Legrand F, Hermann E, Loiseau S, Guerardel Y, Kremer L, Adam E, Dombrowicz D, Capron M TLR2-dependent eosinophil interactions with mycobacteria : role of α-defensins

5. Maes E, Mille C, Poulain D, Guérardel Molecular phenotyping of β-mannosyltransferases deficient *Candida albicans* cells by High Resolution-Magic Angle Spinning NMR

Activités de recherches

A- Orientation générale de notre thématique de recherche

Présentation de l'équipe

Depuis la création de l'Unité de Glycobiologie Structurale et Fonctionnelle (UGSF), l'analyse structurale des glycoconjugués apparaît comme l'une de ses spécificités thématiques les plus fortes. A cause de l'hétérogénéité structurale énorme de cette classe de composés, ce type d'étude requiert la maîtrise d'un grand nombre de techniques spécifiques incluant les modifications chimiques et enzymatiques des glycoconjugués et leur analyse par un vaste panel de méthodes chromatographiques et spectroscopiques. Ainsi, l'analyse structurale des glycoconjugués s'est progressivement définie comme une spécialité distincte au sein de la glycobiologie. A ce titre, l'UGSF a grandement bénéficié de son implication dans la mise en place sur le site de l'Université des Sciences et Technologies de Lille de plateformes technologiques incluant les Centres Communs de Spectrométrie de Masse et de Résonance Magnétique Nucléaire. Celles-ci donnent accès à notre personnel scientifique aux équipements les plus modernes et nous permettent une constante évolution technologique. De plus, les renouvellements thématiques et technologiques de l'axe structural ont été récemment assurés grâce à l'intégration de jeunes chercheurs ayant effectué des stages post-doctoraux à l'étranger et au recrutement d'ingénieurs de recherche en RMN qui interagissent fortement avec les équipes impliquées dans l'analyse structurale.

Ces dernières années coïncident également avec une évolution sensible des projets structuraux de l'unité vers une recherche de type intégrée *via* des interactions plus étroites avec des équipes orientées sur des thématiques biologiques, au sein ou en dehors de l'unité. De fait, il apparaît qu'à ce jour la grande majorité de nos travaux fassent l'objet de collaborations étroites avec une ou plusieurs autres équipes. Celles-ci, soit sont le résultat de la mise en place de projets concertés, soit proviennent d'une demande d'une équipe extérieure désirant mettre à profit notre expérience dans l'analyse structurale. L'intérêt croissant suscité par l'approche post-génomique et le perfectionnement rapide des technologies spectroscopiques nous ont récemment permis d'initier de nouveaux développements dans l'analyse des glycomes et glycoprotéomes. Ces nouvelles stratégies, fondées sur l'étude des profils de glycosylation totaux d'un organisme et l'identification des glycoprotéines porteuses, sont rendues possibles non seulement par la miniaturisation des techniques d'analyse mais également par un accès plus efficace à l'information structurale *via* les banques de données. Ainsi, l'approche glycomique apparaît dès à présent comme un complément indispensable aux approches classiques de biologie moléculaire destinées à étudier l'expression des enzymes impliquées dans la synthèse des glycannes. Cette interface promet des développements rapides dans la connaissance de la régulation de la synthèse des glycannes et de leur rôle

physiologique. Ces nouvelles stratégies sont actuellement mises en place au laboratoire, en particulier dans le cadre des études sur la biodiversité des glycoconjugués, et constituent un des axes de développement prioritaires de notre équipe dans les prochaines années.

Suite au départ en retraite en septembre 2006 de son ancien responsable, le Dr Gérard Strecker, j'ai pris en charge la direction de l'équipe 'Biodiversité Associée aux Glycoconjugués'. Au cours de la dernière année, nous avons redéfini les thématiques du groupe à la lueur des besoins et de l'orientation de l'UGSF. Deux axes thématiques complémentaires dans lesquels s'intègre la majorité de nos travaux de recherche se sont dégagés : L'étude des **relations structure-fonction des glycoconjugués microbiens** et l'étude de la **spécificité des glycosyltransférases**. Au sein de ces deux thématiques, les travaux de l'équipe s'organisent selon trois modes de fonctionnement:

1- *Projet initié dans le cadre d'une collaboration étroite avec un ou plusieurs partenaires*. Celui-ci représente le mode principal de fonctionnement de notre recherche et correspond à une évolution des projets structuraux de l'unité vers une recherche de type intégrée *via* des interactions plus étroites avec des équipes orientées sur des thématiques biologiques. L'intégration récente au sein de notre équipe du Pr. Elisabeth Elass, avec laquelle nous collaborons depuis plusieurs années sur les relations structures-fonctions des glycoconjugués microbiens, est une opportunité pour développer nos recherches dans ce domaine et acquérir une indépendance thématique plus poussée.

2- *Projet indépendant pouvant donner lieu à des collaborations ultérieures*. Ces travaux s'intègrent plus spécifiquement dans une perspective de prospection sur des sujets de glycobiologie fondamentale.

3- *Prestation de service*. Ces travaux proviennent d'appels d'offres d'équipes extérieures. En effet, la reconnaissance de l'importance des glycoconjugués dans les processus biologiques incite un nombre croissant de biologistes à nous contacter pour mettre à profit notre expérience dans l'analyse structurale. Bien qu'ils représentent une partie congrue de notre activité, ces travaux permettent de mettre en valeur et de partager notre savoir faire avec la communauté scientifique, essentiellement française, au travers de projets ponctuels.

L'équipe est actuellement constituée de huit personnels statutaires, d'un post-doctorant et de deux doctorants. Au cours de l'année passée, les effectifs de l'équipe se sont enrichis grâce au recrutement d'un Maitre de Conférence (Emeline Fabre) et à l'arrivée du Pr. Elisabeth Elass issu du groupe de J. Mazurier ainsi que du Dr Frédéric Krzewinski de l'ancien groupe de Stéphane Bouquelet. Chaque membre de l'équipe est responsable d'un projet, mais intervient également dans

plusieurs projets. C'est particulièrement le cas des Drs B. Coddeville et E. Maes qui ont la responsabilité des analyses de résonance plasmonique de surface, de spectrométrie de masse et de résonance magnétique nucléaire.

Nom	**Fonction**	**Spécialités scientifiques**	**Responsabilité thématique**
Yann Guérardel	CR CNRS	Structure	Modèles animaux
Elisabeth Elass	Pr USTL	Biologie cellulaire	Fonction des glycoconjugués microbiens
Ossarath Kol	Pr Polytech-Lille	Structure	Glycannes d'amphibiens
Florence Delplace	McF USTL	Structure	Glycolipides de levure
Emeline Fabre	McF USTL	Enzymologie	Mannosyltransférases de levure
Frédéric Krzewinski	McF Polytech-Lille	Enzymologie	Mannosyltransférases de levure
Bernadette Coddeville	IgR USTL	MS, SPR	Lipoglycannes microbiens
Emmanuel Maes	IR CNRS	RMN	Polysaccharides de *B. cereus*
Estelle Garénaux	ATER USTL	Structure	Glycoprotéines de *B. cereus* et *T. gondii*
Yoann Rombouts	Doctorant USTL	Structure	Glycolipides mycobactériens
Lan-Yi Chang	Doctorante USTL-NTU	Structure, microbiologie	Glycoprotéines de poisson zèbre

Présentation des travaux

Dés la thèse, mes travaux ont en partie consisté en l'étude de la structure des glycannes isolés d'espèces animales génétiquement définies. Ils ont pour but de définir précisément leurs profils de glycosylation et de mettre en évidence de nouvelles activités glycosyltransférasiques sur des modèles animaux dont le génome, la physiologie et le développement sont suffisamment connus pour permettre l'étude subséquente de ces activités et ainsi servir de plateforme à l'étude du rôle éventuel de la glycosylation dans divers processus physiologiques. Ces travaux sont regroupés dans l'axe intitulé '**spécificité des glycosyltransférases'.** Les espèces animales répondant à ces critères et que nous avons étudiées au cours des dernières années sont *Xenopus laevis, Xenopus tropicalis, Caenorhabditis elegans, Danio rerio, Halocynthia roretzi* et *Brachiostoma belcheri.* Ces travaux ont été effectués entre 2000 et 2008 au sein de l'UGSF, dans l'équipe du Dr Khoo (Taiwan) et dans

l'équipe du Pr Kitajima (Japon). Nous nous concentrons actuellement plus particulièrement sur le modèle *Danio rerio* (poisson zèbre) qui présente le plus d'intérêts en terme d'étude du rôle de la glycosylation au cours de la fertilisation et de l'embryogenèse. Ainsi, nous essayons actuellement de promouvoir le poisson zèbre comme modèle d'étude cohérent en glycobiologie au travers de la mise en place d'un réseau de collaborateurs et d'une plateforme d'élevage de poissons zèbres disponible sur l'USTL.

En marge de ces travaux, j'ai initié au cours de ma thèse de doctorat une série d'études sur la structure des glycoconjugués mycobactériens en collaboration avec le docteur Laurent Kremer, du Laboratoire de Mécanismes Moléculaires de la Pathogénie Microbienne INSERM U447 à l'Institut de Biologie de Lille. Ces travaux se sont d'abord focalisés sur l'étude structurale des lipoglycannes isolés d'espèces de mycobactéries encore jamais étudiées pour en évaluer la diversité structurale. Dans un deuxième temps, en collaboration avec plusieurs équipes de microbiologistes et d'immunologistes au sein et en dehors de l'UGSF, nous avons tenté d'évaluer l'influence de la structure des lipoglycannes sur leurs propriétés immunomodulatrices au cours de l'infection mycobactérienne. Au cours des années et au gré de nombreuses collaborations, ces travaux nous ont également initié à l'étude de diverses formes de glycosylation microbienne et constituent à présent un axe thématique majeur de notre laboratoire intitulé '**relations structure-fonction des glycoconjugués microbiens'**. D'une part, le renforcement de la collaboration avec le Dr L. Kremer, maintenant dans le Laboratoire de Dynamique des Interactions Membranaires Normales et Pathologiques à Montpellier, nous a permis d'étendre nos travaux à l'étude structurale d'autres types de glycoconjugués mycobactériens, en particulier aux glycolipides (Lipo-oligosaccharides, Tréhalose dimycolates, Glycolipides phénoliques) et aux glycoprotéines, ainsi qu'à la régulation de la biosynthèse des acides mycoliques. D'autre part, la pérennisation de collaborations anciennes mises en place par Gérard Strecker (Dr. D. Poulain, INSERM U799) et la mise en place de nouvelles collaborations (Dr C. Faille, INRA UR638 ; Dr M. Gohar INRA UMR1238 ; Dr Schwarz Philipps-University Marburg ; Dr J. Previato, Universidade Federal do Rio de Janeiro) nous a permis d'étendre nos champs de compétences à d'autres micro-organismes incluant *Corynebacterium diphtheriae, Candida albicans, Bacillus cereus, Toxoplasma gondii* et *Encephalitozoon cuniculi.*

Dans le présent manuscrit, je présenterai le résultat de nos travaux dans deux chapitres principaux reprenant l'organisation des deux axes thématiques **'relations structure-fonction des glycoconjugués microbiens'** et '**spécificité des glycosyltransférases'**. Dans le premier chapitre,

j'exposerai une partie des travaux sur les glycoconjugués microbiens qui ont en commun l'étude des différentes formes de mannosylation chez les eucaryotes unicellulaires et les bactéries en essayant de mettre en lumière l'évolution de leurs structures et éventuellement de leurs fonctions. De fait, je n'exposerai pas en détail nos travaux actuels concernant l'étude des glycolipides de *Mycobacterium marinum* ou des glycoconjugués de *Bacillus cereus* pour ne pas surcharger le présent manuscrit. Une partie sera néanmoins brièvement abordée dans les perspectives d'études. Dans un deuxième chapitre, j'exposerai brièvement les résultats obtenus sur les premiers modèles d'études que nous avons utilisés, les xénopes et *C. elegans*, et m'attarderai sur nos travaux en cours qui concernent le poisson zèbre (*Danio rerio*) et deux modèles annexes de tunicier (*Halocynthia roretzi*) et d'amphioxus (*Branchiostoma belcheri*). Enfin, dans une dernière partie je compilerai les résultats de quelques développements technologiques portant sur la purification et l'analyse des glycoconjugués nécessaires à nos études. D'une manière générale, j'essaierai de mettre l'accent sur les études réalisées au cours des deux dernières années en incluant quelques résultats sous forme d'articles et d'autres encore non publiés.

B- Mannosylation microbienne

1- Contexte

Le mannose est un monosaccharide très largement répandu dans l'ensemble du monde vivant. C'est un constituant majeur d'un grand nombre de glycoconjugués : polysaccharides, glycolipides, glycoprotéines. Le présent chapitre résume brièvement des connaissances de base sur la structure et la biosynthèse des manno-conjugués majoritaires présents chez les organismes unicellulaires : bactéries, archae et protistes. Nous en avons néanmoins exclu une classe pourtant fortement représentée chez les protistes, les glycosylphosphatidylinositol (GPI), du fait que nous n'ayons jamais abordé leur étude.

1.1- La N-glycosylation

1.1.1- Structure des N-Glycannes

La N-glycosylation est une modification post-traductionnelle commune chez les eucaryotes impliquant la liaison covalente d'un oligosaccharide sur un résidu d'asparagine au sein d'une séquence consensus Asn-X-Ser/Thr (X ≠ Proline). Tous les N-glycannes ont en commun un noyau pentasaccharidique, le trimannosyl-di-N-acétylchitobiose ($Man_3GlcNAc_2$). De fait, la N-glycosylation représente chez la majorité des eucaryotes la principale source de manno-glycoconjugués.

Selon la substitution de ce noyau, on distingue classiquement trois types de N-glycannes: le type N-acétyl-lactosaminique ou complexe, le type oligomannosidique et le type hybride (Fig. 1). Pour les N-glycannes de type complexe, l'élongation peut se faire par des antennes N-acétyllactosaminiques de type 1 : Gal(β1-3)GlcNAcβ ou de type 2 : Gal(β1-4)GlcNAcβ. La diversité structurale des N-glycannes de type complexe ou hybride réside en particulier dans les possibilités de substitution à leur extrémité par des sucres dits périphériques ou terminaux. Le plus souvent, la périphérie comprend des monosaccharides d'anomérie α, en positions terminales non-réductrices, sur les antennes ou directement sur le noyau. Chez les mammifères, ces monosaccharides consistent généralement en des unités de Fuc ou NeuAc. Le plus souvent ce sont des acides sialiques liés en α-2,3 ou α-2,6 sur un Gal terminal ou du Fuc en (α1,6) sur le noyau ou en (α1,3) sur un Gal ou une GlcNAc terminale. Chez l'homme, ces sucres périphériques peuvent constituer le support de motifs antigéniques. Enfin, les N-glycannes peuvent posséder des résidus de Gal 3-, 4- ou 6-sulfate ainsi que de la GlcNAc-6-phosphate.

Fig. 1 : Les trois types de structure N-glycanniques. *A : type complexe ou N-lactosaminique; B : type oligomannosidique ou « high-mannose » ; C : type hybride ; D : chaîne lactosaminique de type 1 et 2 (Structures classiques de mammifères).Noir : noyau ; bleu : antennes ; vert : sucres périphériques ou terminaux.*

1.1.2- Voie de biosynthèse conservée

La biosynthèse d'un *N*-glycanne s'effectue en plusieurs étapes : assemblage du précurseur oligosaccharidique, transfert, action de diverses glycosidases puis de glycosyltransférases. Ce processus est localisé à la fois dans le Réticulum Endoplasmique (RE) et dans l'appareil de Golgi. Les étapes décrites ci-dessous correspondent à la biosynthèse des *N*-glycannes telle qu'elle se déroule dans une cellule de mammifère. Dans le cadre de nos études sur les processus de mannosylation, nous nous concentrerons sur les étapes précoces de la biosynthèse.

Le précurseur tétradécasaccharidique $Glc_3Man_9GlcNAc_2$ (Fig. 2) est pré assemblé sur un résidu polyisoprénique particulier, le dolichol, *via* une liaison pyrophosphate.

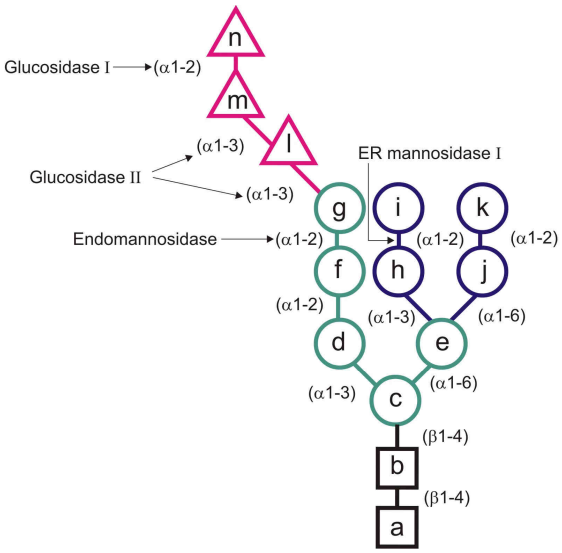


Fig. 2 : Structure de l'oligosaccharide précurseur des N-glycannes. *Ce tétradécasaccharide contient 3 glucoses (triangles), 9 mannoses (cercles) et 2 N-Acétylglucosamines (carrés). Les lettres correspondent à l'ordre d'addition du monosaccharide lors de la biosynthèse du précurseur. Les résidus de mannose en vert ou en bleu diffèrent à la fois par la localisation de leur transfert et par le donneur impliqué (vert : face cytosolique, GDP-Man ; bleu : lumen, DolPMan). L'oligosaccharide entier correspond au précurseur synthétisé dans les cellules de mammifères, de plantes et de champignons. Les sites de clivage de quelques enzymes impliqués dans l'élagage sont indiqués.*

Cet assemblage s'effectue dans le RER au cours d'un processus cyclique, le cycle des dolichols, impliquant l'action séquentielle de différentes glycosyltransférases, produits des gènes ALG (Helenius et Aebi, 2004). Les différentes étapes de cet assemblage sont schématisées dans la Figure 3.

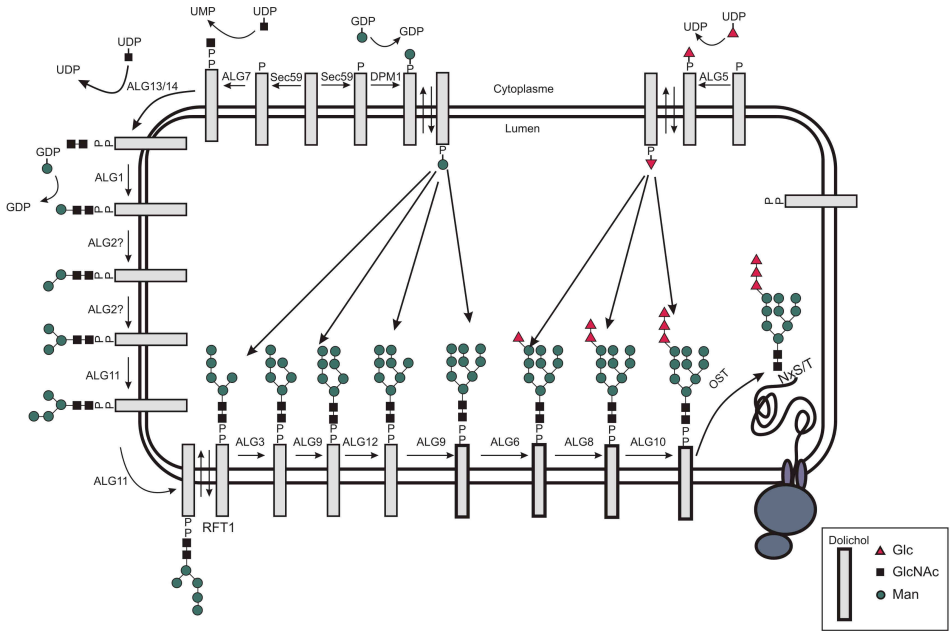


Fig. 3 : *Le cycle des dolichols (d'après Helenius et Aebi, 2004)*

Dans un premier temps, le $GlcNAc_2$-P-P-Dol est formé par addition successive de GlcNAc-1-P sur le P-Dol, puis de GlcNAc sur le GlcNAc-PP-Dol ainsi formé. Les glycosyltransférases impliquées utilisent dans les deux cas le nucléotide sucre UDP-GlcNAc comme donneur. Ces premières réactions se déroulent sur la face cytoplasmique du RE. Les étapes suivantes consistent en l'addition successive de 9 résidus de mannose. Les cinq premiers (c,d,e,f,g) le sont à partir de GDP-Man en une suite de réactions se déroulant sur la face cytoplasmique. Les 4 derniers (h,i,j,k) s'ajoutant, au contraire, sur la face luminale après retournement de l'intermédiaire $Man_5GlcNAc_2$-P-P-Dol vers la face interne de la membrane. Les mannosyltransférases impliquées utilisent le Man-P-Dol, donneur synthétisé face cytoplasmique à partir de GDP-Man qui, étant donné son caractère hydrophobe, permet une réaction de transfert localisée dans la bicouche lipidique interne du RE (Kornfeld and Kornfeld, 1985). De façon similaire, l'addition des trois derniers résidus de glucose se fait à partir de Glc-P-Dol, conduisant au produit final, le $Glc_3Man_9GlcNAc_2$-P-P-Dol. Une fois complet, le précurseur est transféré en bloc sur un résidu asparaginyl d'une protéine nouvellement synthétisée au sein de la séquence-consensus Asn-X-Ser/Thr ($X \neq Pro$) : il s'agit d'un événement co-traductionnel.

Chez les mammifères, l'addition des trois résidus de glucose confère à l'oligosaccharide précurseur une affinité plus grande pour l'oligosaccharyltransférase (OST) qui en fait son substrat préférentiel pour le transfert sur la protéine. Le résidu asparaginyl sur lequel se fait l'addition doit être dans un triplet-consensus: Asn-X-Ser/Thr (X étant un aminoacide différent de la proline). Si en général l'OST affiche une affinité particulière pour le précurseur $Glc_3Man_9GlcNAc_2$-P-P-Dol, celle-ci n'exclue cependant pas le transfert de précurseurs tronqués (Freeze et Aebi, 2005). C'est à ce moment, une fois l'étape de transfert "en bloc" effectuée, que le devenir de la glycoprotéine est défini : transfert vers l'appareil de Golgi, adressage vers le lysosome ou dégradation. C'est précisément dans l'appareil de Golgi que le précurseur oligosaccharidique subira les étapes indispensables de maturation à l'origine des structures oligomannosidiques, complexes ou hybrides. L'accès à l'appareil de Golgi est conditionné par divers signaux portés à la fois par la protéine et le glycanne.

La diversification ultérieure des *N*-glycannes de type oligomannosyl en un large répertoire de N-glycannes hybrides ou complexes résulte de l'action combinée de glycosyltransférases et glycosidases. L'élagage du précurseur N-glycannique commence dans le RE et se poursuit dans les différents saccules golgiens pour ne conserver de la structure initiale que l'heptasaccharide $Man_5GlcNAc_2$.

1.1.3- Diversification des voies de biosynthèse

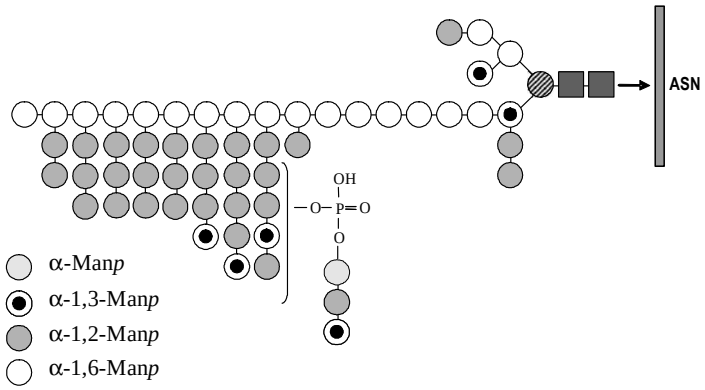
1.1.3.1- Chez les eucaryotes unicellulaires

La majorité des eucaryotes synthétisant des N-glycannes (champignons, plantes, animaux, moisissure rampante et euglènes) respectent le processus de N-glycosylation décrit ci-dessus. La conservation du core pentasaccharidique indique que chaque cellule eucaryote produisant des N-glycannes a conservé l'étape princeps, réticulaire, de la biosynthèse des N-glycannes : le cycle des dolichols. Si cette étape est conservée, le peu de données structurales disponibles concernant la N-glycosylation chez les protistes montre néanmoins que certains de ces organismes présentent des voies de biosynthèse originales, tronquées. Ces travaux indiquent ainsi que le transfert d'un précurseur unique, l'oligosaccharide $Glc_3Man_9GlcNAc_2$, n'a pas été strictement conservé au cours de l'évolution, et qu'il existe ainsi divers précurseurs oligosaccharidiques synthétisés et transférés (Castro *et al.*, 2006). La nature du précurseur synthétisé dépend directement du panel d'enzymes de type ALG présent dans un organisme. L'existence de domaines conservés au sein des enzymes de type ALG facilite leur recherche dans des banques de données. De fait, il est possible de prédire la nature du précurseur oligosaccharidique synthétisé par un organisme en déterminant les enzymes ALG qu'il possède. Ainsi, il apparaît que chez certains eucaryotes, les processus de N-glycosylation aient été modifiés dans les étapes précoces malgré leur niveau important de conservation, c'est le cas chez les eucaryotes unicellulaires.

Les levures

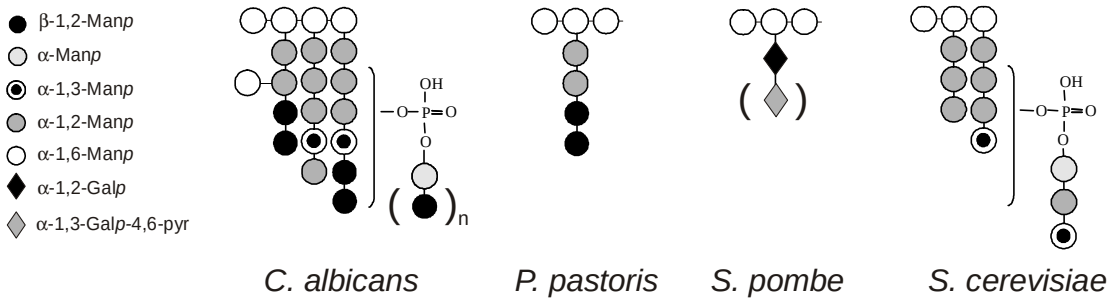
Les levures sont les protistes dont les voies de N-glycosylation ont été de loin les mieux étudiées. La principale modification dans la voie de biosynthèse précoce des N-glycannes chez les levures en comparaison du processus précédemment exposé provient du fait que ces organismes ne raccourcissent pas le $Man_8GlcNAc_2$ comme le font les cellules de mammifères. De fait, les profils de N-glycosylation des levures sont caractérisés par la présence de structures oligomannosylées qui sont à leurs tours allongées par un grand nombre de résidus de mannoses. Ce noyau interne, une fois la glycoprotéine transférée vers l'appareil de Golgi, peut subir une extension par le transfert de mannoses additionnels liés en (α1–6), menant à un noyau externe comprenant de 8 à 15 résidus de mannose. Finalement, chacun des mannoses liés en (α1,6) peut porter une branche en (α1,2) ou (α1,3), résultat de l'action de mannosyltransférases golgiennes spécifiques aux levures qui permettent la synthèse de N-glycannes polymannosylés (Ballou, 1990). De plus, chez de nombreuses espèces de levures, sont observées des chaînes oligomannosidiques supplémentaires liées aux chaînes latérales *via* des liaisons de type phosphodiester. Sur la base des résultats obtenus

par l'étude de *Saccharomyces cerevisiae*, la structure des N-glycannes polymannosylés de levure peut être schématisée comme ci-dessous.



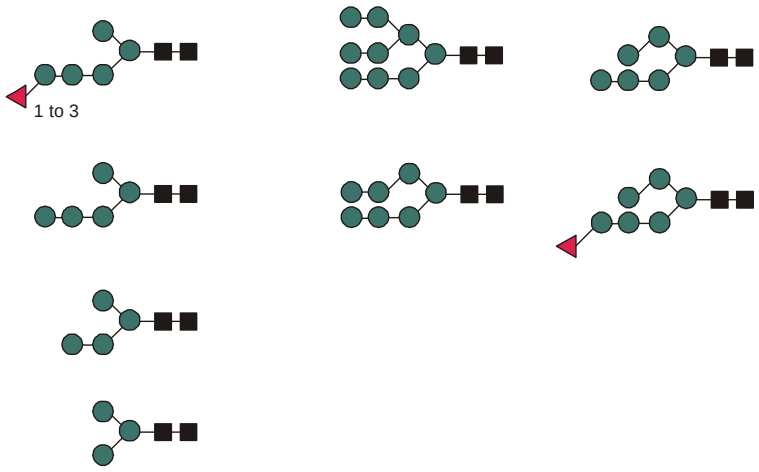
Ainsi, à l'exception des structures $Man_8GlcNAc_2$ et $Man_9GlcNAc_2$ les N-glycannes de levures sont donc largement différents de ceux observés chez les mammifères. De plus, aucun N-glycanne de type complexe n'est observé chez les levures. La charge normalement apportée par la présence de résidus d'acides sialiques ou de groupements sulfates est éventuellement remplacée par des groupements phosphates, voire dans de rares cas, par des acides uroniques (Fukazawa *et al.*, 95) ou des groupements pyruvates (Gemmil *et al.*, 96). Néanmoins, la fonction éventuelle de telles charges est inconnue à ce jour.

Si le modèle de glycosylation général semble être bien conservé chez toutes les levures étudiées, il faut néanmoins noter une diversité structurale relativement importante entre les différentes espèces voire entre différents sérotypes de la même espèce. Ainsi, les N-glycannes vont différer à la fois dans la composition des leurs extrémités (Man, Gal, GlcNAc), dans la taille et degré de branchement de leurs oligomannosides ainsi que dans l'anomérie de leurs résidus glycosidiques. Ci-dessous sont représentés quelques exemples de spéciation structurale des N-glycannes polymannosylés.



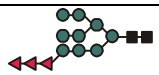
Les autres protistes

En dehors des levures, les données concernant les processus de N-glycosylation des protistes restent très fragmentaires. Néanmoins, quelques études structurales assez anciennes ont mis en évidence la présence de N-glycannes oligomannosylés tronqués chez plusieurs espèces d'eucaryotes unicellulaires tels que *Tretrahymana pyriformis*, *Trypanosoma cruzi* ou *Leishmania major*, comme indiqués ci-dessous :



T. pyriformis Taniguchi *et al.*, 1985 — *T. cruzi* Engel *et al.*, 1985 — *L. major* Funk *et al.*, 1987

Ces résultats suggèrent fortement la modification des processus réticulaires de glycosylation chez ces organismes, due à l'absence d'un certain nombre de glycosyltransférases réticulaires. Des études récentes de Robbins et collaborateurs, sur la base de tests biochimiques et de l'exploration des bases de données génomiques, ont élégamment démontré l'absence de panels différents de glycosyltransférases de type Alg chez la plupart des eucaryotes unicellulaires (Samuelson *et al.*, 2005 ; Banerjee *et al.*, 2007). Ces résultats sont résumés ci-dessous :

Organismes	**Enzymes réticulaires manquantes**	**Forme limite de transfert prédite**
S. cerevisiae ; Dictyostelium	Aucune	
T. cruzi	Alg 6, 8, 10 GlcTfase utilisant Dol-P-Glc	
T. gondii ; T. pyriformis	Alg 3, 9, 12 ManTfase utilisant Dol-P-Man	
G. intestinalis ; P. falciparum	Toutes les Algs, sauf Alg 7	
E. cuniculi	Toutes les Algs et STT3	

Cette diversité de la glycosylation ne résulterait pas de la complexification progressive du processus de glycosylation mais bien de pertes secondaires des différentes enzymes au cours de l'évolution, à partir d'un ancêtre commun possédant le panel complet d'enzymes qui permettent la biosynthèse du précurseur $Glc_3Man_9GlcNAc_2$.

1.1.3.2- Chez les bactéries et *archaea*

Bien qu'aboutissant à la synthèse de structures très différentes, les rares exemples de processus de N-glycosylation décrits chez les bactéries relèvent de nombreuses homologies fonctionnelles. En particulier, un équivalent bactérien de N-glycanne consistant en l'heptasaccharide GalNAc(α1-4)GalNAc(α1-4)[Glc(α1-3)GalNAc(α1-4)GalNAc(α1-4)GalNAc(α1-3)$BacNH_2$ lié sur un résidu d'asparagine a été mis en évidence chez *C. jejuni* [Young *et al.*, 2002]. Bien que non essentiel à sa survie, l'absence de N-glycosylation réduit fortement la capacité de la bactérie à coloniser son hôte (Szymanski *et al.*, 2002). De la même manière que chez les eukaryotes, la biosynthèse des N-glycannes est initiée sur un substrat lipidique (un bactoprénylpyrophosphate inséré dans la membrane en place du dolichylphosphate de la membrane du RE) dans le compartiment cytoplasmique au sein duquel se poursuit l'addition séquentielle des divers monosaccharides à partir de nucléotides sucres. Puis, l'heptasaccharide précurseur traverse la membrane pour s'orienter vers le périplasme et être transféré en bloc sur le résidu d'Asn d'une séquence consensuelle Asp/Glu-X-Asn-Z-Ser/Thr (X et Z différents de Pro) par une protéine équivalente à STT3, la protéine PglB (Linton *et al.*, 2005; Kelly *et al.*, 2006 ; Kowarik, 2006). Au contraire des eucaryotes chez qui la diversification structurale s'effectue précisément dans la lumière du RE et du Golgi, les bactéries ne semblent pas modifier la structure oligosaccharidique après changement de compartiment. Néanmoins, le manque de données structurales ne permet d'apprécier ni la diversité structurale des N-glycannes bactériens ni la spécificité *in vivo* de PglB envers le précurseur oligosaccharidique. Un processus équivalent à celui des bactéries a été observé chez les a*rchaea*, chez lesquels la biosynthèse serait initiée sur des précurseurs lipidiques de type dolichylphosphate et dolichylpyrophosphate (Lechner *et al.*, 1985). Au contraire des bactéries, les données structurales accessibles établissent le transfert de N-glycannes de structures différentes en fonction de l'espèce : Man(β1-4)Gal chez *Haloferax volcanii* (Kuntz et al 1997), ManpNAcA6Thr(β1-4)GlcpNAc3NAcA(β1-3)GlcpNAc(β- chez *Methanococcus voltae* (Voisin *et al.*, 20005). Ces données suggèrent que la spécificité de l'oligosaccharidyl-transferase *d'archaea* soit assez lâche pour transférer des glycannes de structures différentes, ce que confirme

l'observation d'un transfert d'oligosaccharides tronqués chez le mutant *M. Voltae aglA* (Chaban *et al.*, 2006). Les homologies entre eucaryotes, bactéries et *archaea* sont résumées ci dessous :

	Eucaryotes	**Bactéries**	**Archae**
Donneur de sucre	Nucléotides-sucres Dol-P-sucres	Nucléotides-sucres	Nucléotides-sucres
Transporteur lipidique	Dolichyl pyrophosphate	Bactoprényl-pyrophosphate	Dolichyl mono- ou pyrophosphate
Lieu du transfert d'oligosaccharide	Lumière du ER	Face externe de la membrane plasmique	Face externe de la membrane plasmique
Oligosaccharidyl transférase	Multimérique ; STT3 sous-unité catalytique	Monomérique ; PglB	Monomérique ; PglB
Séquence consensuelle	N-X-S/T (X ≠ P)	D/E-X-N-Z-S/T (Z, X ≠ P)	N-X-S/T (X ≠ P); N-X-N/L/V
Modifications du glycopeptide	Dans le RE et Golgi	Aucune	Inconnue

D'un point de vue évolutif, la présence dans les trois domaines du vivant de N-glycannes dont la biosynthèse repose sur des principes communs suggère que cette modification post-traductionnelle soit très ancienne et que le système commun à l'ensemble des eucaryotes dérive d'un ancêtre bactérien ou archéal unique.

1.2- La O-glycosylation

1.2.1- Structure des O-glycannes

La O-glycosylation dans son sens large définit la liaison d'un oligosaccharide sur un polypeptide *via* une liaison O-glycosidique. Il existe différents types de O-glycosylation, certains types étant limités à des espèces, des tissus, ou des polypeptides particuliers. Elle représente une famille extrêmement diversifiée que l'on observe dans l'ensemble du monde vivant. En effet, la O-glycosylation peut s'effectuer sur plusieurs types d'acides aminés (Ser, Thr, Hyp, Tyr) *via* de nombreux monosaccharides différents (Man, GalNAc, GlcNAc, Gal, Fuc, Glc, Ara), chaque combinaison étant à l'origine d'un type particulier de O-glycannes. La plus fréquemment répandue dans le monde animal, implique la liaison d'un résidu de α-D-N-acétylgalactosamine en α sur un acide aminé hydroxylé, la sérine (Ser) ou la thréonine (Thr). Elle définie une famille

structurellement très hétérogène de composés connus sous le nom de 'O-glycannes de type mucine'. Néanmoins cette famille de O-glycannes ne contient jamais de résidus de mannose. Une autre famille de O-glycannes, beaucoup plus restreinte, est composée quant à elle presque uniquement de résidus de mannose et est connue sous le nom de 'O-mannosylation'. Elle est en particulier retrouvée chez les levures ou elle constitue la famille majeure de O-glycannes, mais a été également identifiée chez certaines bactéries et les eucaryotes supérieurs.

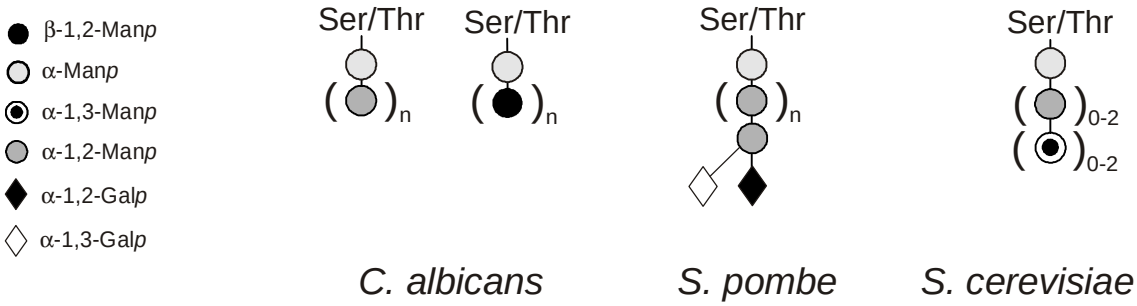
Chez les bactéries

Chez les bactéries la O-mannosylation a été identifiée dans plusieurs espèces de mycobactéries dont *Mycobacterium tuberculosis* et *M. bovis*. En particulier, la protéine du complexe 45/47-kDa (Apa) est substituée par des O-mannosides constitués d'α1,2Man*p* (Dobos *et al.*, 1995 ; Dobos *et al.*, 1996). La présence de ces O-mannanes serait indispensable aux capacités immuno-modulatrices d'Apa et servirait de ligand à la protéine surfactante pulmonaire A (PSP-A) (Romain *et al.*, 1999 ; Horn te al., 1999 ; Ragas *et al.*, 2007). La protéine MPB83 retrouvée elle aussi chez *M. tuberculosis* et *M. bovis* est quant à elle substituée par un O-mannoside constitué d'unités d'α1,3Man*p*, ce qui laisse présager l'existence de processus de mannosylation spécifiques des protéines au sein d'une même espèce de mycobactérie (Michell *et al.*, 2003). Plusieurs autres glycoprotéines mycobatériennes, telles que les lipoprotéines de 19 et 38 kDa de *M. tuberculosis*, sont supposées être O-mannosylées sur la base de reconnaissances spécifiques avec des lectines, mais la structure exacte de leurs domaines mannosylés respectifs est encore inconnue (Garbe *et al.*, 1993 ; Herrmann *et al.*, 1996). En dehors des mycobactéries, quelques rares autres exemples de mannosylation directe sur les protéines bactériennes ont été mis en évidence, telles les N-acétylendoglucosaminidases F2 et F3 de *Flavobacterium meningosepticum* substituées par le glycanne 2-O-Me-Man(1-4)GlcNAcA(1-4)GlcA(1-4)Glc(1-4)2-O-Me-GlcA-1-4)[2-O-Me-Rham(1-2)]Man et les cellulases de *Streptomyces lividans* et *Cellulomonas fimi* substituées par un O-glycanne contenant du mannose et du galactose (Ong *et al.*, 1994 ; Reinhold *et al.*, 1995).

Chez les levures

La O-mannosylation reste la seule forme connue de O-glycosylation chez les levures. Toutes les levures étudiées jusqu'à présent sont capables de synthétiser des structures de type Man(α1-2)Man(α1-2)Man(α1-)Ser/Thr. Néanmoins, comme pour les N-glycannes, les structures synthétisées varient selon l'espèce considérée. *S. cerevisiae* se distingue par l'addition de Man(α1,3) pour donner lieu à de courtes structures mannosylées, éventuellement substituées par

des groupements phosphates. *Candida albicans* et *Pichia pastoris* ne transfèrent que des Man liés en α-1,2, formant des chaînes de tailles variables, tandis que *Candida albicans* semble également être capable de transférer des Man liés en β-1,2 (Fradin *et al.*, résultats non publiés). *Saccharoyces pombe* se caractérise par l'attachement d'un ou deux résidus de Gal, conduisant à l'éventuelle obtention d'oligosaccharides ramifiés. Cette diversité structurale est résumée ci-dessous. Il est intéressant de remarquer un parallèle entre la structure des O-glycannes et celles des chaînes latérales des N-glycannes dans une espèce donnée, ce qui laisse présager des systèmes de biosynthèse au moins en partie commun.



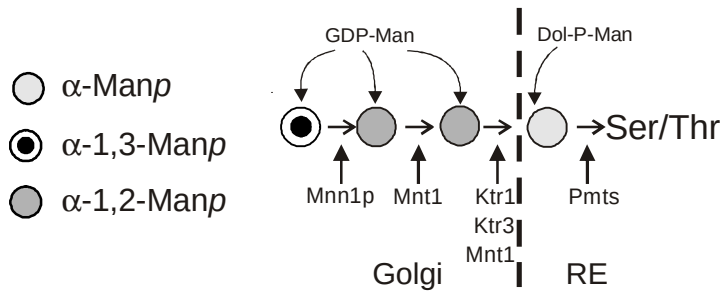
Les O-glycannes liés *via* un résidu de mannose sont également retrouvés chez les eucaryotes supérieurs sous d'autres formes. On notera en particulier, l'α-dystroglycan, une glycoprotéine indispensable à la neurogenèse, qui est substituée par une chaîne glycannique de type : Sia(α2-3)Gal(β1-4)GlcNAc(β1-2)Manα (Shiba et al, 1997).

1.2.2- Biosynthèse des O-glycannes

La biosynthèse des O-mannosides a été décryptée dans un premier temps dans l'organisme chez qui ils ont été découverts, la levure *S. cerevisiae*. Au contraire de la majorité des autres types de O-glycosylation, l'initiation de la O-mannosylation a lieu dans la lumière du RE, par le transfert de la première unité de mannose sur le peptide en cours de translocation. Ce transfert s'effectue sous l'action d'une famille de mannosyltransférases, les Pmts, qui utilisent le Dol-P-Man comme substrat donneur de mannose et non un nucléotide sucre. *S. cerevisiae* comprend actuellement sept membres présentant de fortes homologies et classés dans trois sous familles : la sous famille Pmt1 (Pmt1, pmt5, pmt7), la sous famille Pmt2 (Pmt2, Pmt3, Pmt6) et la sous famille Pmt4. Par comparaison, 5 isoformes de Pmts ont été identifiés chez *C. albicans* et seulement 3 chez *S. pombe*. La multiplicité des Pmts chez les levures, servant toutes au transfert d'un unique résidu de mannose,

est encore une question ouverte et est réminiscente de la multiplicité des polypeptides transférases chez les mammifères (17 à ce jour) ; ce sujet a été récemment abordé en détail par Ernst et collaborateurs (Lengeler *et al.*, 2008). Il semble néanmoins que ces enzymes ne soient activent que sous forme d'hétérodimères présentant des spécificités de substrats accepteurs différentes et qu'au sein d'une même protéines, différents domaines puissent être mannosylées par des Pmt différentes (Girrbach et al, 2005 ; Ecker *et al.*, 2003).

Dans un deuxième temps, les O-mannosides sont allongés dans le Golgi sous l'action séquentielle d'α-1,2-mannosyltransférases de la famille Ktr et d'α-1,3-mannosyltransférases de la famille Mnn1 utilisant de GDP-Man comme substrat donneur (Lussier *et al.*, 1999). Il est à noter que ces mêmes enzymes participent conjointement au remodelage des chaînes latérales N-glycannes dans le Golgi. Les différentes étapes de la biosynthèse des O-mannosides chez *S. cerevisiae* peuvent être résumées comme ci-dessous :

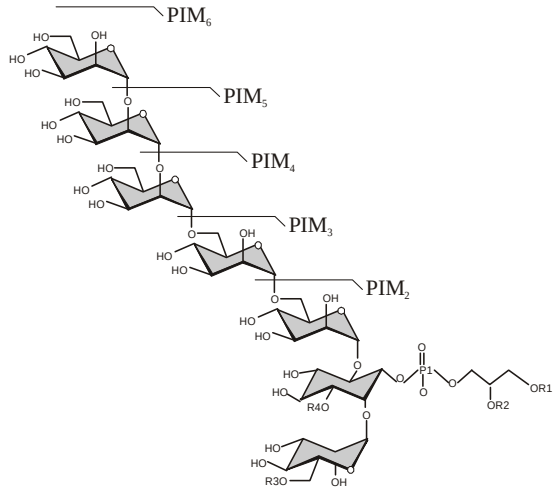


Chez les eucaryotes supérieurs, seuls deux homologues des gènes PMTs ont été identifiés : POMT1 et POMT2 chez les mammifères et leurs homologues *rt* (rotated abdomen) et *tw* (twisted) chez *Drosophila melanogaster* (Jurado *et al.*, 1999 ; Willer *et al.*, 2002). Tout comme les Pmts chez les levures, l'hétérodimère POMT1/POMT2 initie la biosynthèse de la O-mannosylation dans le RE chez les mammifères (Lommel *et al.*, 2008). Aucun équivalent des protéines de mannosylation golgiennes Ktrs, Mnt1 et Mnn1 n'a été observé chez les eucaryotes supérieurs, ce qui expliquerait l'absence d'oligomannosides.

1.3- Les Lipoglycannes

Une dernière classe majeure de manno-glycoconjugués que nous avons étudiée sont les lipoglycannes et leurs dérivés isolés de mycobactéries. Les lipoglycannes sont des constituants majeurs de la paroi des mycobactéries. Ils sont tous basés sur une ancre de type phospo-inositol-glycérol acylée. Ils comprennent trois familles de composés reliés par une filiation biosynthétique

directe: les phosphatidyl-inositolmannosides (PIMs), les lipomannanes (LMs) et les lipoarabinomannanes (LAMs). Les PIMs sont constitués d'une famille homogène de composés présentant de deux (PIM_2) à six (PIM_6) résidus de Man*p*. L'inositol de PIM_2 est mannosylé en positions C-2 et C-6 tandis que l'élongation subséquente de la chaîne mannosylée dans les PIM_3 à PIM_6 s'effectue à partir de la position C-6 de l'inositol (Lee, Y.C. & Ballou, C.E., 1965). La position C-2 de l'inositol étant toujours substituée par un unique mannose, c'est la longueur de la chaîne en C-6 (1 à 5 mannoses) qui détermine la nature du PIM (PIM_2 à PIM_6). De fait, l'archétype de la structure de la partie oligosaccharidique des PIMs peut être résumée comme ci-dessous :



En plus des positions C-1 (position R1) et C-2 (position R2) du glycérol, les positions C-2 (position R4) et C-6 (position R3) de l'inositol et du Man*p* lié en 1,2 sur l'inositol respectivement peuvent également être acylées, donnant naissance à des formes di-, tri- et tétra-acylées des PIMs (Ac_2-, Ac_3- et Ac_4PIMs). Les trois acides gras les plus communément identifiés lors des études citées sont les acides palmitique (C_{16}), stéarique (C_{18}) et tuberculostéarique (C_{19}, acide 10-méthyl-octadécanoïque).

Les LM et LAM sont des lipoglycannes complexes constitués de plusieurs domaines: une ancre phosphatidyl-*myo*-inositol et un domaine polysaccharidique plus ou moins complexe. Il sont issus de l'allongement de PIM_2 et peuvent être considérées comme des formes multi-glycosylées des PIMs. De fait, l'ancre phospho-*myo*-inositol des LM et LAM est strictement identique à celle des PIMs. LM et LAM semblent systématiquement co-exister chez toutes les mycobactéries et ne diffèrent que par la constitution de leur domaine polysaccharidique. Alors que celui du LM n'est constitué que d'un homopolymère de D-mannose (domaine mannane), celui du LAM est constitué d'un domaine mannane et d'un domaine arabinane, ce dernier étant lui même substitué par un

troisième domaine de composition variable appelé "coiffe". Les LM et LAM partagent un même domaine mannane directement lié sur la position C-6 de l'inositol de l'ancre phospho-*myo*-inositol-glycérol. Il est constitué d'un enchaînement linéaire d'unités α-D-Man*p* liées en 6 partiellement substituées en α-1,2 par un résidu de D-Man*p*. Cet arrangement à été décrit à l'identique chez *M. tuberculosis* (Chatterjee *et al.*, 1991), *M. bovis* BCG (Venisse *et al.*, 1995) et *M. smegmatis* (Khoo, *et al.*, 1996). Par contre, la taille et le degré de substitution de ce domaine sont sujets à variations. Au sein d'une même espèce, le polymère de mannose exhibe un polydispersité importante. Ceci a été confirmé grâce à la mesure de la masse du domaine mannane du LM de *M. smegmatis* par spectrométrie de masse qui a montré que sa taille varie de 17 à 35 résidus de mannoses, avec un maximum de 26 unités (Khoo *et al.*, 1996). Dans cette même espèce il a été montré que les domaines mannanes du LAM et du LM étaient de même taille et présentaient une polydispersité équivalente, alors que ceux de *M. tuberculosis* Erdman sont de tailles différentes (Chatterjee & Khoo 1998). La taille du domaine mannane varie également en fonction de l'espèce: en moyenne 26 chez *M. smegmatis*, 20 chez *M. tuberculosis* Erdman et 18 chez *M. bovis* BCG. De même, le degré de substitution des chaînes de mannose semble varier en fonction de l'espèce de 50 à 70 % environ. Le détail des relations entre structure et fonctions des lipoglycannes a fait l'objet de plusieurs revues très complètes (Chatterjee & Khoo, 1998, Gilleron *et al.*, 2008).

2- Travaux

2.1- Simplification des voies de glycosylation chez les eucaryotes unicellulaires

Ces travaux ont été entrepris dans le cadre d'une étude générale concernant la simplification des voies de biosynthèse des protistes. Sur la base des travaux précédemment exposés, il apparaît que de nombreux protistes ont perdu au cours de leur évolution un nombre variable d'enzymes de N-glycosylation, ce qui a modifié leurs potentiels de glycosylation (Samuelson *et al.*, 2005 ; Banerjee *et al.*, 2007). Ces données génétiques ont été confirmées dans quelques cas (*T. pyriformis, T. cruzi*) par des études structurales indépendantes, mais ces dernières restent trop fragmentaires. En particulier, aucune donnée structurale solide concernant les eucaryotes unicellulaires parasites obligatoires tels que *T. gondii* et *P. falciparum* n'est réellement disponible ce qui limite l'interprétation des données génétiques.

Nous avons eu l'opportunité d'aborder cette thématique par l'étude de *T. gondii* et de deux espèces de microsporidies au travers de collaborations étroites avec deux équipes spécialisées dans l'étude de ces organismes. Nous nous sommes attaché à déterminer la structure des N- et des O-glycannes synthétisées par ces modèles d'étude. L'une des difficultés majeures de ces études et de différencier le matériel parasitaire de celui de la cellule hôte dans laquelle il se développe. En effet, malgré des protocoles de purification sophistiqués, il est virtuellement impossible de se prémunir de contaminations éventuelles de la cellule hôte (Robbins *et al.*, 2005 ; Monk *et al.*, 2006). Ce problème est encore aggravé quand le parasite synthétise potentiellement des composés structurellement proches de ceux de l'hôte et/ou qu'il les synthétise dans des quantités restreintes. De fait, ces études ne peuvent s'effectuer que par une comparaison minutieuse des profils de glycosylation des cellules hôtes et de leurs parasites.

Etude des profils de mannosylation des microsporidies

Ces travaux ont été initiés en 2005 dans le cadre d'une collaboration avec le Dr Christian Vivares (UMR CNRS 6023). Ils ont pour objectif d'étudier le rôle éventuel de la glycosylation des protéines membranaires des microsporidies. Une première étape a consisté à définir les profils de glycosylation de deux espèces de microsporidies, *Encephalitozoon cuniculi* et *Antonospora locustae*.

Très peu d'informations sont actuellement disponibles sur la diversité glycannique et les processus de glycosylation chez les microsporidies, un ensemble de plus de 1200 espèces d'eucaryotes unicellulaires, toutes parasites intracellulaires. Plusieurs espèces sont des pathogènes de l'homme et provoquent des infections chez les patients immunodéprimés. La spore des microsporidies est entourée par une paroi épaisse. Elle contient un organe appelé tube polaire qui

joue un rôle essentiel lors de l'invasion de la cellule hôte. La paroi épaisse de la spore est majoritairement composée de chitine. Par contre, ces parasites ne seraient capables de synthétiser ni glycogène, souvent utilisé comme polysaccharide de réserve, ni β1,3 glucane pourtant communément retrouvé dans la paroi des champignons. La présence de glycoconjugués a été cytochimiquement suggérée pour au moins trois types de structures sporales : l'exospore, le capuchon polaire et le tube polaire. En particulier, l'utilisation de lectines a démontré la présence de glycannes O-mannosylés sur des protéines pariétales (PTP1). Ces résultats préliminaires étaient en accord avec l'annotation du génome d'*E. cuniculi* qui suggérait la O-mannosylation comme seul mode de O-glycosylation dans cette espèce.

Dans le but de définir leur rôle éventuel et les enzymes impliquées dans leur biosynthèse, nous avons entrepris d'étudier la structure fine des glycannes majeurs de *E. cuniculi* et *A. locustae*. Nous avons démontré l'absence de N-glycosylation chez ces deux organismes, en accord avec l'absence des homologues microsporidiens pour les enzymes-clefs de la N-glycosylation. Par contre, nous avons mis en évidence et séquencé par l'utilisation combinée de la spectrométrie de masse, de la chromatographie en phase gaz et de la RMN à très haut champ une famille homogène de O-glycannes linéaires constitués de résidus de mannoses liés en α1-2 (Taupin *et al.*, 2007). Ces composés sont très similaires à ceux identifiés dans la levure pathogène *C. albicans*. La définition précise de la structure des O-glycannes endogènes ouvre la voie à l'étude des enzymes de biosynthèse des glycannes et du rôle de la glycosylation chez ces organismes sur le modèle des études actuellement réalisées chez les levures.

Glycobiology vol. 17 no. 1 pp. 56–67, 2006
doi:10.1093/glycob/cwl050
Advance Access publication on September 15, 2006

Major *O*-glycans in the spores of two microsporidian parasites are represented by unbranched manno-oligosaccharides containing α-1,2 linkages

Vanessa Taupin[2], Estelle Garenaux[3], Muriel Mazet[2], Emmanuel Maes[3], Hubert Denise[2], Gérard Prensier[2], Christian P. Vivarès[1,2], Yann Guérardel[3], and Guy Méténier[2]

[2]Equipe Parasitologie Moléculaire et Cellulaire, LBP, CNRS UMR6023, Université Blaise Pascal, 24 Avenue des Landais, 63177 Aubière Cedex, France; and [3]Laboratoire de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq Cedex, France

Received on July 12, 2006; revised on September 8, 2006; accepted on September 11, 2006

Protein glycosylation in microsporidia, a fungi-related group comprising exclusively obligate intracellular parasitic species, is still poorly documented. Here, we have studied glycoconjugate localization and glycan structures in spores of *Encephalitozoon cuniculi* and *Antonospora locustae*, two distantly related microsporidians invading mammalian and insect hosts, respectively. The polar sac-anchoring disc complex or polar cap, an apical element of the sporal invasion apparatus, was strongly periodic acid-thiocarbohydrazide-Ag proteinate-positive. Mannose-binding lectins reacted with the polar cap and recognized several bands (from 20 to 160 kDa) on blots of *E. cuniculi* protein extracts. Physicochemical analyses provided the first determination of major glycostructures in microsporidia. *O*-linked glycans were demonstrated to be linear manno-oligosaccharides containing up to eight α1, 2-linked mannose residues, thus resembling those reported in some fungi such as *Candida albicans*. No *N*-linked glycans were detected. The data are in accordance with gene-based prediction of a minimal O-mannosylation pathway. Further identification of individual mannoproteins should help in the understanding of spore germination mechanism and host–microsporidia interactions.

Keywords: glycan analysis/microsporidia/O-mannosylation/polar cap/ultracytochemistry

Introduction

The identification of glycans and glycoconjugates in eukaryotic parasites is of interest for the knowledge of host–parasite interactions and pathogenic determinants (Guha-Niyogi et al. 2001). It is worth noting that the capacity of classical *O*-linked and *N*-linked glycosylations can be highly reduced in some species exhibiting an obligate intracellular lifestyle. This is the case of the apicomplexan parasite *Plasmodium falciparum* (malaria agent) in which the most abundant glycoconjugate structures are glycosylphosphatidylinositol (GPI) anchors (Gowda et al. 1997). Bioinformatic analysis of the *P. falciparum* genome sequence revealed only four potential enzymes of the N-glycosylation pathway and no enzyme characteristic of the synthesis of complex *O*-linked glycans (Aravind et al. 2003).

Relatively little is known about carbohydrate diversity and glycosylation processes in microsporidia, an assemblage of over 1200 unicellular eukaryotic species that are all obligate intracellular parasites. These organisms are viewed as highly derived fungi having undergone rapid reductive evolution (Keeling 2003; Thomarat et al. 2004). Several species are human pathogens and may cause severe diseases in immune-deficient patients (Weiss 2001). Their development inside host cells comprises a proliferation phase (merogony) followed by a differentiation phase (sporogony) producing small spores that can be released in the environment. Surrounded by a resistant cell wall, the microsporidian spore contains a very long coiled organelle (polar tube) that plays an essential role in the onset of cell invasion. Indeed, the polar tube can be quickly extruded at the apical pole of the spore in order to inject the sporoplasm into a new host cell (Xu and Weiss 2005).

Chitin appears as the unique microsporidial polysaccharide and is associated with the thick inner layer of the spore wall, named the endospore (Bigliardi et al. 1996; Vavra and Larson 1999). Glycoconjugates should be abundant in a periodic acid-Schiff (PAS)-positive apical region of the spore, called the "polar sac-anchoring disc complex" or "polar cap", that comprises a dome-shaped vesicular structure closely associated with a discoid element capping the anterior end of the polar tube (Vavra and Larson 1999). Current biochemical information about microsporidial glycoproteins mainly derives from lectin-binding experiments. Blotting of *Glugea plecoglossi* spore extracts with eight different lectins showed that only concanavalin A (ConA) and wheat germ agglutinin (WGA) react with some protein bands (Kim et al. 1999). In *Encephalitozoon intestinalis*, two proteins of the outer spore wall layer (exospore) were assumed to be N-glycosylated on the only basis of their detection with specific antibodies in the fractions of infected cell lysates that were immobilized on either ConA or WGA-coated agarose beads (Hayman et al. 2001). In contrast, purified *E. hellem* polar tube protein (PTP) 1 should be O-mannosylated because of its lack of reactivity with an antibody to *O*-GlcNAc, its binding to only ConA among 10 different lectins, and elimination of this binding by NaOH treatment (Xu et al. 2004). The occurrence of O-mannosylation is more in agreement with annotation data for the genome sequence of the closely related species *E. cuniculi* (Katinka et al. 2001; Vivarès and Méténier 2004).

[1]To whom correspondence should be addressed; Tel: +33 4 73 40 74 57; Fax: +33 4 73 40 76 70; e-mail: christian.vivares@univ-bpclermont.fr

In this article, we report on the localization and structure of mannoconjugates in spores of the microsporidians *E. cuniculi* and *Antonospora locustae*. Labeling with mannose-specific lectins at the electron microscope (EM) level suggested that the polar cap is a major site for mannose-rich glycoproteins. Analyses of glycan fractions with mass spectrometry (MS) and nuclear magnetic resonance (NMR) techniques led to the first determination of oligosaccharide structures linked to microsporidial proteins.

Results

Our primary choice of *E. cuniculi*, a monokaryotic (mononucleate) microsporidian that infects a wide range of mammals, including humans, was justified by the extreme reduction of its genome (2.9 Mbp) that was fully sequenced (Katinka et al. 2001) and viewed as a model of "minimal genome" in eukaryotic cells (Méténier and Vivarès 2004). For comparison, analyses were also performed in the diplokaryotic (binucleate) species *A. locustae* that invades grasshoppers and locusts and has a 5.3-Mbp genome for which a sequencing project is in progress (http://gmod.mbl.edu/perl/site/antonospora01; *Antonospora locustae* Genome Project, Marine Biological Laboratory at Woods Hole, funded by NSF award number 0135272).

Reactivity of spore structures to PATAg and mannose-binding lectins

An adaptation of the PAS reaction to the EM localization of polysaccharides and glycoconjugates, known as the periodic acid-thiocarbohydrazide-Ag proteinate (PATAg) reaction (Thiéry 1967), was applied to ultrathin sections of epoxy resin-embedded *E. cuniculi* spores. A strongly labeled structure was a cup-shaped organelle, named the polar cap, close to the spore apex and abutting the anterior end of the polar tube (Figure 1A). Another PATAg-positive region was the lamellar polaroplast, a tightly folded membrane system that represents the precursor of the new plasma membrane surrounding the sporoplasm when transferred into a host cell. The reactivity of the spore envelope and polar tube was rather low. Interestingly, a section through a germinated spore (with extruded polar tube) revealed that the reactive material of the polar cap forms a collar-like structure around the aperture required for the passage of the polar tube (Figure 1B). Moreover, a significant labeling was associated with the surface of the extruded polar tube. The carbohydrate richness of the polar cap was also evident when applying the PATAg test to ultrathin frozen sections of spores from both *A. locustae* (Figure 1C) and *E. cuniculi* (data not shown).

The EM localization of potential mannoconjugates in *E. cuniculi* cells was investigated by treatment of cryosections with two biotinylated mannose-binding lectins [ConA and *Galanthus nivalis* agglutinin (GNA)] followed by immunogold detection. ConA labeling was associated with polar tube coils located in the posterior region of the spore (Figure 2A). In the anterior region, both the polar tube straight part and the polar cap were ConA-reactive (Figure 2B). The main GNA-labeled spore structure was the polar cap (Figure 2C), as previously found with PATAg reaction. A few images were obtained for sporoblasts, the precursor cells in which the biogenesis of the extrusion apparatus occurs. Gold particles were clustered in a more or less central cytoplasmic area consisting of small vesicles and extending close to a tubular network characteristic of polar tube formation (Figure 2D). These vesicles possibly derive from the nondictyosomal Golgi apparatus that remains difficult to identify only on the basis of morphological criteria. The same labeling pattern was observed in *A. locustae* cells (data not shown).

Detection of E. cuniculi *glycoproteins on electrophoretic profiles*

The spore proteome of *E. cuniculi* has been recently investigated using two-dimensional (2-D) electrophoresis and MS techniques, leading to sequence characterization of 177 protein spots (Brosson et al. 2006). In a first attempt to detect glycoproteins in 2-D profiles, gels (pI range: 3–10) were either stained with Coomassie blue (Figure 3A) or blotted for further staining with a PAS-derived procedure involving sugar biotinylation (Figure 3B). A low number of reactive spots were observed. A poorly resolved acidic region was centered on 55 kDa and correlated with a group of 10 different proteins. Two of these proteins are known to be specifically secreted during sporogony: PTP1 (Delbac, Peyret, et al. 1998) and exospore protein spore wall protein (SWP) (Bohne et al. 2000). No sequence assignation was available for other glycoprotein spots, mainly seen between 30 and 35 kDa at neutral pI and in a 110-kDa basic region.

The presence of mannose-containing glycoproteins was tested by ConA and GNA blotting after sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of spore protein extracts. Because of their highly resistant cell wall, *E. cuniculi* spores were subjected to two different extraction procedures: one based on the combined action of urea and a disulfide-reducing agent [dithiothreitol (DTT)] and the other involving hot SDS treatment. Although protein patterns of the two kinds of extracts were significantly different, ConA and GNA reacted with some common bands in the two profiles: two bands at 38 and 55 kDa and least two other bands smearing between 100 and 160 kDa (Figure 4). In SDS extracts, a 65-kDa band and a doublet close to 20 kDa were also recognized by the two lectins. Ovalbumin preincubation prevented lectin binding for all bands (data not shown). Thus, in accordance with our previous ultracytochemical data, some microsporidial spore proteins should be mannosylated. Subsequent physicochemical studies were oriented toward the characterization of glycan structures.

N-*glycans are virtually absent in peptide* N-*glycosidase-digested fractions*

Potential *N*- and *O*-linked glycans were purified from total delipidated extracts by sequential enzymatic digestions with peptide *N*-glycosidase F (PNGase F) and peptide *N*-glycosidase A (PNGase A), followed by reductive β-elimination. Separation of released glycans from remaining glycopeptides was achieved between each step by passage through C-18 columns, which was shown to permit collection of *N*- and *O*-linked glycans from a single sample (Dell et al. 1994). Both PNGase F and A, with different substrate specificities depending on the presence of fucose residue on chitobiose core, were used to release both possible types of *N*-glycans. Each of the three glycan fractions was tested for

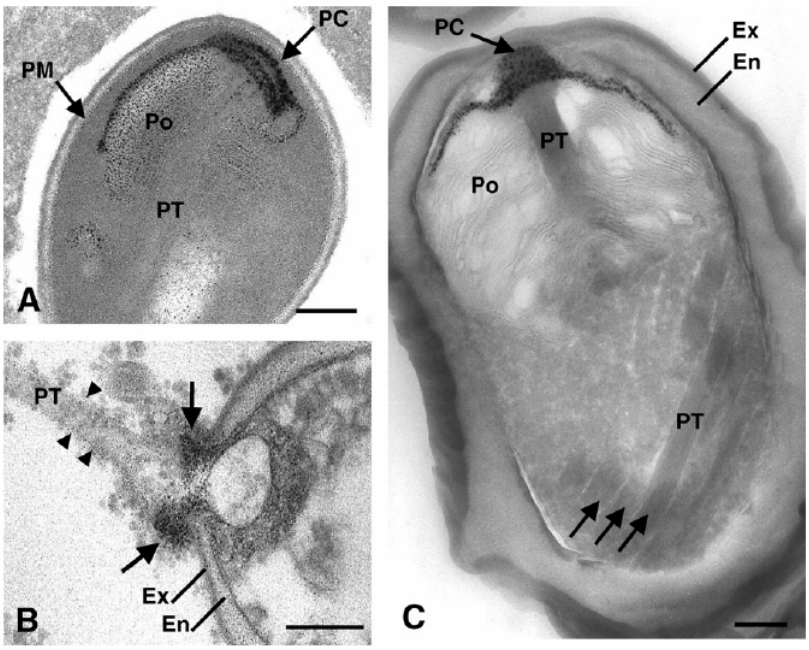


Fig. 1. EM localization of glycoconjugates with PATAg reaction in microsporidian spores. [(**A**) and (**B**)] Epoxy resin sections of *E. cuniculi* spores. (A) Strong reactivity of the polar cap (PC) overlying the lamellar polaroplast (Po) and anterior straight part of the polar tube (PT). Polaroplast and plasma membrane (PM) are moderately labeled. (B) Section across a germinated spore focussing on the apical aperture surrounded by PATAg-positive material (arrows) originating from the PC. Note also the reactivity of the surface of the extruded polar tube (arrowheads). (**C**) Cryosection of *A. locustae* spore. PC labeling is evident. Arrows indicate several coils of the polar tube. Ex, exospore; En, endospore. Bar, 200 nm.

its monosaccharide composition in gas chromatography coupled to mass spectrometry (GC–MS) and MS-analyzed as permethylated derivatives. This strategy repetitively failed to conclusively demonstrate the presence of *N*-linked glycans in several batches of spores from both *E. cuniculi* and *A. locustae*. Composition analysis of PNGase-digested products did not show the presence of GlcNAc and Man residues in significant amounts to be attributed to *N*-linked glycans. Accordingly, MS analyses of these fractions in native forms, or after permethylation, did not show any signal due to potential *N*-glycans. We conclude that either endogenous *N*-glycans were absent or, if present, the quantities were too low to be detected by physicochemical means. As expected, *N*-glycans were found in mammalian Madin–Darby canine kidney (MDCK) cells used for our *E. cuniculi* cultures after application of the same purification protocol and matrix assisted-laser desorption/ionization time of flight (MALDI-TOF) MS.

O-*glycans are linear mannosylated oligosaccharides*

Composition analyses established that mannose is the major monosaccharide component of *O*-linked glycan fractions (at least 70%). In order to facilitate further analysis, *O*-glycan fractions were separated into two subfractions (flow through subfraction and included subfraction) by gel filtration on a Bio-Gel P2 column. These subfractions were then permethylated and subjected to MALDI-TOF and electro-spray (ES)-ion trap MS analyses. In all experiments involving permethylated products, CD_3I was used as primary permethylation reagent to detect natural monosaccharides modified by methyl groups, as found in various organisms (Kocharova et al. 2000; Guerardel et al. 2001). The MALDI-TOF MS spectrum of permethylated flow through *O*-glycan fraction from *E. cuniculi* revealed a set of ions ranging from m/z 947 to m/z 1799 (Figure 5A). Their calculated compositions correspond to sodium adducts of permethylated-reduced tetra-hexosides (Hex_4-ol) to octa-hexosides (Hex_8-ol). The absence of $[M\text{-}3+Na]^+$ or $[M\text{-}6+Na]^+$ clearly demonstrated the absence of natural mono- or di-methylated hexose residues within the detected oligosaccharides. MS analysis of the gel filtration included fraction of *O*-glycans and showed only two $[M+Na]^+$ ions at m/z 520 and 733, attributed to smaller reduced di- and tri-hexosides (data not shown).

The nature of all oligosaccharides was confirmed by collision-induced decay tandem mass spectrometry (CID-MS/MS) analysis of all $[M+Na]^+$ ions detected by MALDI-TOF MS. For ions at m/z 1373 and 1586, fragmentation patterns of all compounds were consistent with the presence of linear-reduced oligomers of hexoses (Figure 5A and B). All fragmentation spectra were dominated by a series of Y-type (m/z at 503, 716, 929, 1143, 1356) and B-type (m/z at 466, 679, 892, 1105, 1319) ions resulting from the cleavage of glycosidic bonds from terminal nonreducing and reducing ends of oligomers, respectively. In addition, two sets of ions tentatively attributed to internal $^{1,3}A$ or $^{2,4}A$ fragment ions (m/z at 756, 969, 1182) and to internal $^{1,5}A$ fragment ions (m/z at 651, 864, 1077, 1291) were observed as minor fragments. The nature of the $^{1,5}A$ fragment ions was confirmed by MS^3 analysis. As expected, MS^3

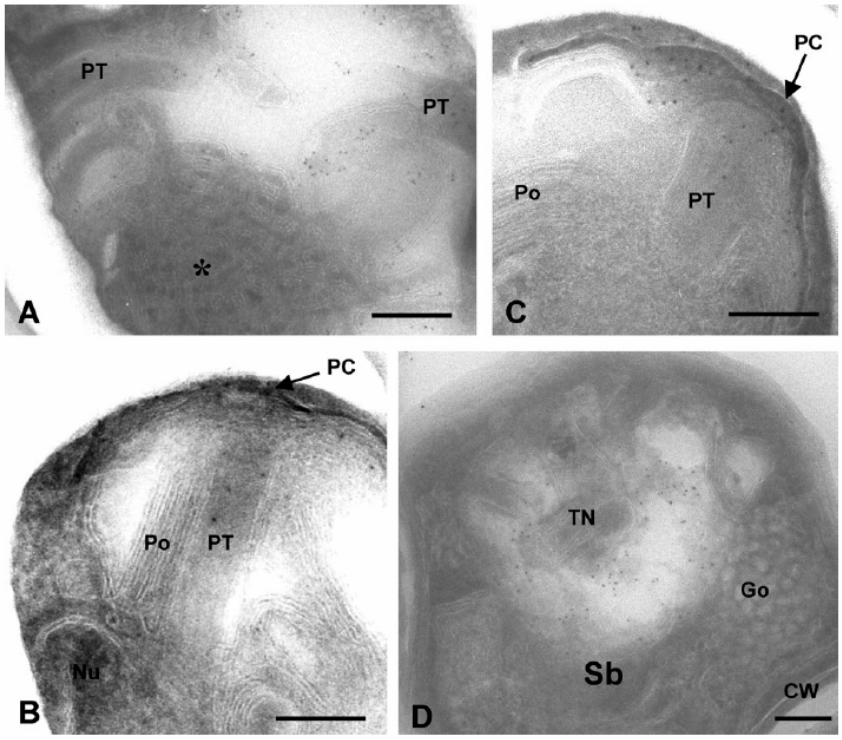


Fig. 2. EM localization of ConA- and GNA-binding sites in cryosectioned *E. cuniculi* cells. [(**A**) and (**B**)]. ConA binding. (A) Gold particles are located on polar tube (PT) coils, as seen in an immature spore. A membranous body likely involved in the formation of the posterior vacuole is unlabeled (*). (B) Anterior spore region showing labeled polar cap (PC). [(**C**) and (**D**)] GNA binding. (C) GNA also binds to PC. (D) Labeling of some vesicular elements in a sporoblast stage (Sb). These elements are located between densely packed Golgi vesicles (Go) and a tubular network (TN) representative of PT precursors. CW, cell wall; Nu, nucleus; Po, polaroplast. Bar, 200 nm.

fragmentation spectra were all characterized by the presence of a set of B-type $[M\text{-}185\text{-}(213)_n + Na]^+$ ions indicating the release of $CH_2OCD_3\text{-}CH_2\text{-}(CHOCD_3)_2\text{-}CH_2O$ fragment and a set of Y-type ions resulting from the release of nonreducing hexose residues (data not shown). It is noteworthy that no hydroxyl group containing secondary fragment ion, indicative of the presence of branched structures, was observed in any of the CID-MS/MS spectra. Taken together, these data strongly suggested the presence of a family of mannosylated linear oligosaccharides with a degree of polymerization (DP) ranging from two to at least eight.

O-mannosyl glycans consist of α-1,2 linked mannose residues

In order to identify the linkage position of Man residues within the detected oligomers, permethylated *O*-glycan fractions from *E. cuniculi* were submitted to methanolysis and acetylation prior to analysis by GC–MS. Total ion chromatogram (TIC) showed three major peaks labeled 1

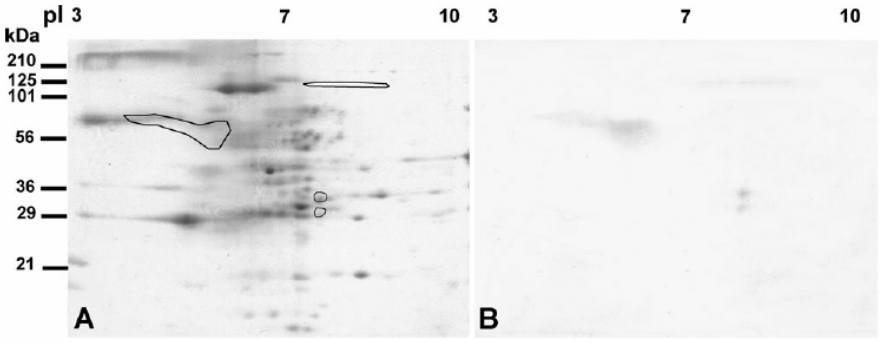


Fig. 3. Two-dimensional gel electrophoresis of *E. cuniculi* spore proteins. (**A**) Coomassie blue-stained gel. (**B**) Corresponding blot stained with a PAS-derived procedure. On the left, a reactive acidic region includes at least PTP1 and SWP1, two known glycoproteins assigned to the polar tube and exospore, respectively. Other reactive spots (30, 35, and 110 kDa) represent still unidentified proteins.

59

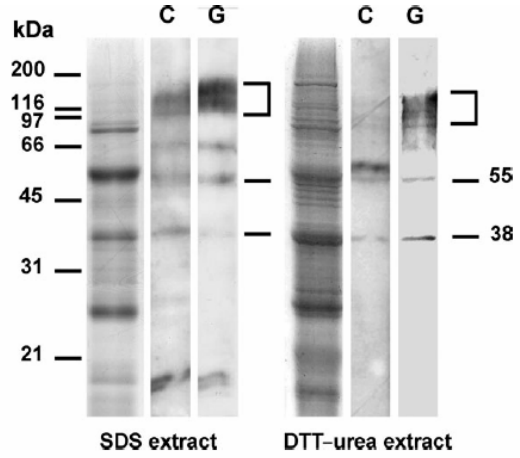


Fig. 4. ConA and GNA lectin overlay. SDS–PAGE and lectin blotting (C, ConA; G, GNA) of *E. cuniculi* proteins extracted with either usual SDS-containing lysis buffer or DTT–urea solution. First lanes show Coomassie blue-stained profiles. Common lectin-binding bands are marked.

to 3 (Figure 6A). By comparison of their retention times with standard molecules, peaks 1 and 2 were identified as 2,3,4,6-Me_4 Man(Me) (terminal nonreducing mannose) and 3,4,6-Me_3-2-Ac Man(Me) (internal 2-linked mannose), respectively. The positions of methyl groups were confirmed by electronic impact (EI)-MS (Figure 6B and C). Similarly, peak 3 was first identified as 1,3,4,5,6-Me_5-2-Ac Hex-ol residue according to its fragmentation pattern (Figure 6D). Subsequent comparison of its retention time with standards produced from acetolysis and reduction of commercial *Saccharomyces cerevisiae* mannan definitely typified it as a 2-linked Man-ol residue. This residue results from the reduction of the reducing mannose following release of *O*-glycans by reductive β-elimination. In accordance with the ES-MS/MS analysis, no di-substituted mannose residue indicative of branched structures was observed in GC–MS analysis. Differential integration of peak areas established that t-Man, 2-Man, and 2-Man-ol were present in a ratio of 0.8/2.9/1, establishing an average DP of five for the high molecular mass fraction. Similar results were obtained from low molecular mass *O*-glycan fraction, which differs from high molecular mass fraction only by a lower average DP of 2.5 (data not shown). It is, therefore, clear that these *O*-mannosyl glycans are made of linear stretches of 2-linked Man residues.

Anomery of mannose residues was determined by 800 MHz ^{1}H NMR analyses. One-dimensional (1-D) ^{1}H (Figure 7A) and 2-D COSY 90 (Figure 7B) NMR spectra revealed two individual anomeric protons at δ 5.224 and 5.045 ppm that correlated with their respective H-2 protons at δ 3.98 and 4.07 ppm. It also revealed a group of at least four H-1 protons (5.286–5.312 ppm), correlating with H-2 protons between 4.08 and 4.10 ppm. Chemical shifts of anomeric protons superior to 5.04 ppm as well as $J_{1,2}$ coupling constant of approximately 1.8 Hz clearly established that all residues were α anomers (Cohen and Ballou 1980; Faille et al. 1992). According to published ^{1}H NMR parameters, signals at δ 5.045 and 5.286–5.312 ppm were attributed to terminal nonreducing α-Man residue and internal α1,2-linked Man residues, whereas signal at δ 5.224 was attributed to the single α-Man residue substituting Man-ol in C-2 position (Hayette et al. 1992; Trimble et al. 2004). The multiplicity of signals associated with internal (α1-2)Man residues results from the polydispersity of the oligomannosides, as revealed by MS analysis, that influences individual NMR parameters (Kobayashi et al. 1994).

Altogether, the data collected from composition, linkage, MS, MS/MS, and NMR analyses established the presence in *E. cuniculi* of a family of *O*-linked glycans exclusively composed of α1,2-linked mannose residues, with a size ranging from two to at least eight residues. Identical strategy conducted in *A. locustae* showed very similar results. As the two species belong to two different genera that have been shown to be distantly related in rRNA phylogenetic analyses (Slamovits et al. 2004), such glycan structures might be a general feature in the microsporidian phylum.

Discussion

As first observed in *Stempellia* spores (Vavra 1972), the application of the PATAg procedure to ultrathin sections of *E. cuniculi* and *A. locustae* spores indicated glycoconjugate richness of the polar cap. This organelle was also labeled with biotinylated ConA and GNA, which supports the presence of mannose-containing glycoconjugates. After extrusion, the polar tube must be firmly attached to the spore apex in order to maintain the integrity of the sporoplasm flowing toward a host cell. A critical role of the polar cap in this attachment is conceivable because of its positioning between apical plasma membrane and polar tube domains and its conversion into a collar-like structure during polar tube extrusion (Lom 1972). Some PATAg-positive material seen at the junction between extruded polar tube and spore body (Figure 1B) likely reflects controlled exocytosis of polar cap contents. An abundance of hydrophilic sugars associated with polar cap proteins may facilitate polar tube exit than sliding through the apical aperture. The IgG response of immunocompetent humans against *Encephalitozoon* spp. was found to be directed against carbohydrate moieties of PTP1 and of some proteins migrated as a smear in SDS–PAGE between 100 and 250 kDa (Peek et al. 2005). Immunofluorescence data supported a localization of the last antigenic glycoproteins in the anchoring region, and a role in the adherence of the spore to the host cell surface has been hypothesized by the authors. The adherence of *E. intestinalis* spores to host cells in vitro involved sulfated glycosaminoglycans of the host cell surface, and inhibition of spore adherence by chondroitin sulfate A caused a significant decrease in the percentage of infected host cells (Hayman et al. 2005). Exospore glycoproteins are the best candidates for interacting with host glycosaminoglycans. A putative contribution of polar cap glycoproteins to host cell–microsporidia interactions should occur only after polar tube extrusion.

The mannose-binding lectins ConA and GNA recognized several bands in SDS–PAGE patterns of *E. cuniculi* spore proteins, and only mannose was clearly identified after GC–MS analysis of *E. cuniculi* and *A. locustae* oligosaccharide fractions subjected to methanolysis. ConA binding was reported for protein bands within a size range of 45–57 kDa

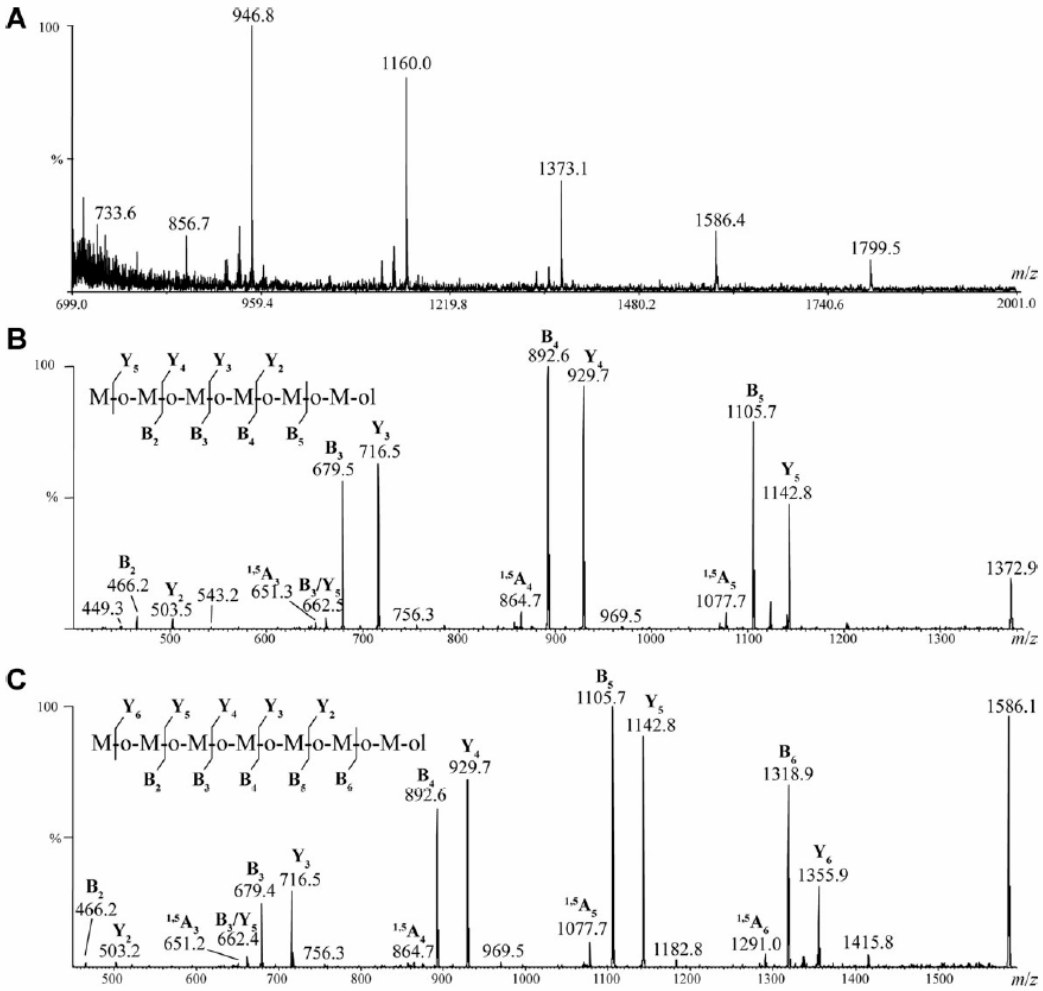


Fig. 5. MS analysis of glycans released by reductive β-elimination. (**A**) MALDI-MS screening of perdeuteromethylated O-glycans from *E. cuniculi*; CID-MS/MS sequencing of (**B**) reduced hexasaccharide at m/z 1373 and (**C**) of reduced heptasaccharide at m/z 1586, labeled according to Domon and Costello (1988). M stands for mannose residue.

in a fish-infecting microsporidian, the most labeled band being at 55 kDa (Kim et al. 1999), as for exospore proteins SWP1 and SWP2 in *E. intestinalis* (Hayman et al. 2001) and PTP1 in *E. hellem* (Xu et al. 2004), *A. locustae*, and *Paranosema grylli* (Polonais et al. 2005). SWP1 and PTP1 anomalously migrate in SDS–PAGE between 50 and 55 kDa and exist in different species, whereas SWP2 has a larger size (150 kDa) and is apparently restricted to *E. intestinalis*. In *E. cuniculi*, the ConA- and GNA-binding bands at 55 kDa may be partially correlated with a large PAS-reactive area that was also close to 55 kDa in 2-D gels and known to include both SWP1 and PTP1. Evidence for binding of PTP1 to GNA was previously obtained by lectin blotting in two insect-infecting microsporidia (Polonais et al. 2005). Whether at least one high molecular weight smeared band may contain a polar cap protein should deserve further investigations. The reactivity with GNA, a lectin that preferentially recognizes terminal α1,3-linked Man residues (Shibuya et al. 1988), may be somewhat surprising because only α1,2-linked Man residues were detected here and no microsporidial gene encoding α1,3-mannosyltransferase was identified. However, in the work of Shibuya et al. (1988), manno-oligosaccharides with an α1,2-linked linkage were found to also inhibit mannan-GNA precipitation, although less efficiently than oligosaccharides with terminal Man(α1,3)Man unit. Moreover, the retardation of glycopeptides on GNA column was strongly dependent on the number and heterogeneity of disaccharide units. The binding of GNA to homogeneous α1,2-linked mannose chains seems therefore likely.

Neither GlcNAc and Man residues nor mass-specific *N*-glycan signal was identified in PNGase-treated glycan fractions of the studied microsporidia. Although this cannot formally exclude the possibility of a very low frequency of N-glycosylation, it should be stressed that a real loss of N-glycosylation is consistent with bioinformatic analyses of gene repertoires in *E. cuniculi* (Katinka et al. 2001) and *A. locustae* (*Antonospora locustae* Genome Project; http://gmod.mbl.edu/perl/site/antonospora01). Strikingly,

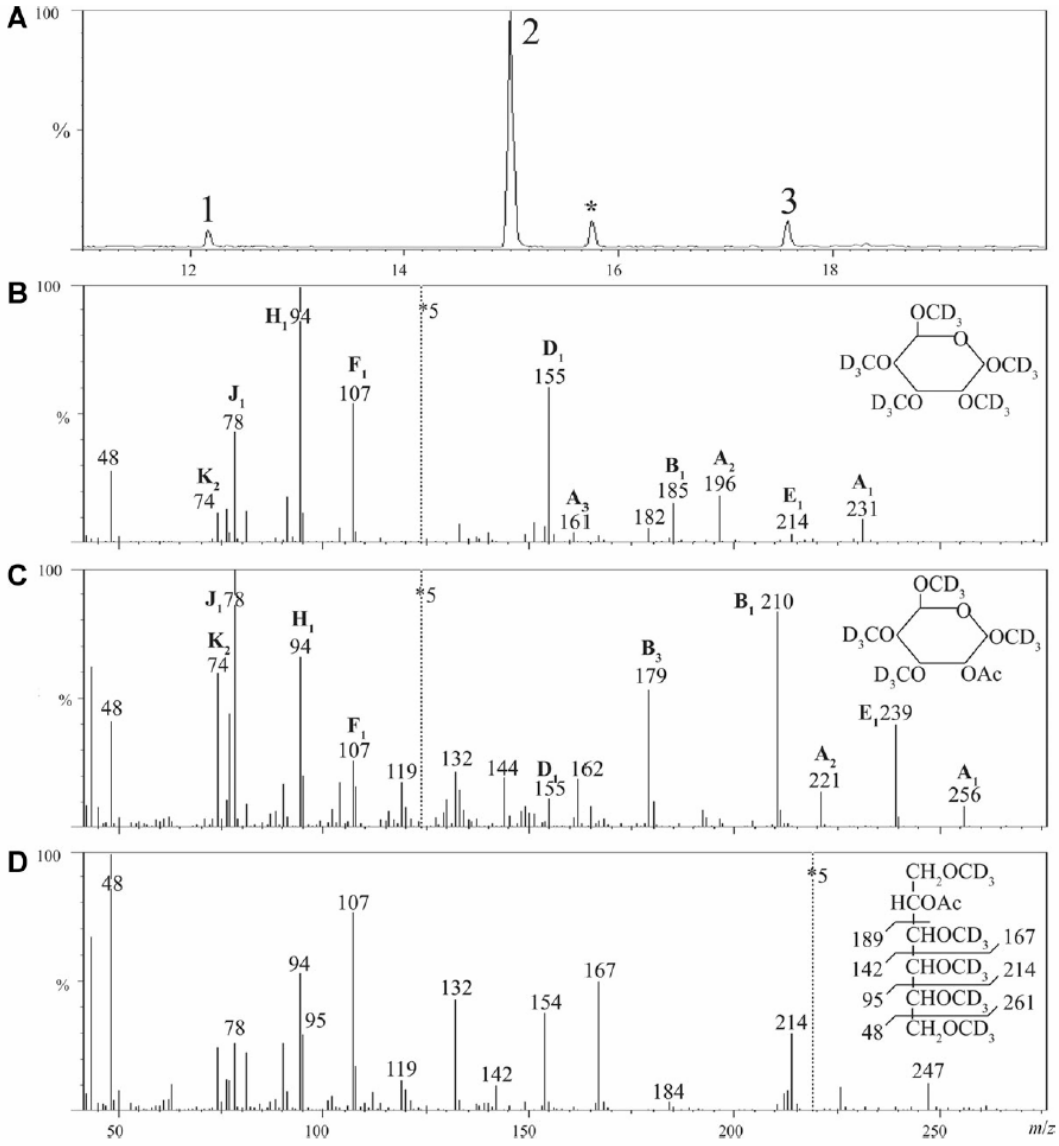


Fig. 6. Linkage analysis of *E. cuniculi* O-linked glycans by GC–MS. (**A**) TIC chromatogram of partially deuteromethylated and acetylated methyl-glycosides; (**B**–**D**), EI-MS spectra of compounds 1 (terminal mannose), 2 (2-linked internal mannose), and 3 (2-linked mannitol), respectively. Each compound was identified according to its retention time and fragmentation pattern.

no genes were found to encode subunits of the oligosaccharyltransferase complex needed for flipping of dolichol (Dol)-PP-GlcNAc$_2$Man$_5$ across the endoplasmic reticulum (ER) membrane and then linking the high-mannose oligosaccharide to asparagine residues on nascent peptides. The lack of key enzymes for protein N-glycosylation in *E. cuniculi* has been also verified through a recent inventory of Alg glycosyltransferases in several eukaryotic organisms; their comparison leading the authors to postulate that various secondary losses of enzymes from a common eukaryotic ancestor may have occurred (Samuelson et al. 2005). The potential *E. cuniculi* proteome also lacks critical factors for correct folding of N-glycosylated proteins in the ER, such as processing glycosidases and calnexin–calreticulin chaperone system (Katinka et al. 2001). In fungal organisms, the major fraction of *N*-linked glycans is incorporated within cell wall phosphomannoprotein complexes, and the importance of these glycans in host–fungal interactions is especially well illustrated by their involvement in the adherence of *Candida albicans* cells to host macrophages (Cutler 2001). It is noteworthy that, unlike typical fungi, microsporidia have no permanent cell wall. Proliferating intracellular stages (meronts) are indeed delimited only by plasma membrane, with cell wall formation occurring during sporogony. The low diversity of glycoprotein spots revealed after 2-D gel electrophoresis might be related to lacking N-glycosylated proteins.

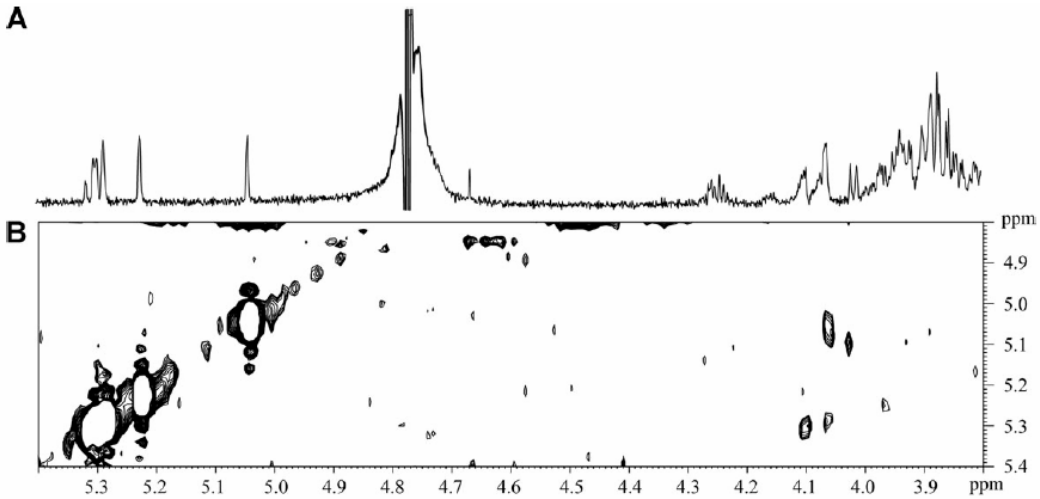


Fig. 7. 800 MHz ^{1}H NMR analysis of *E. cuniculi* O-linked glycans. (**A**) One-dimensional and (**B**) 2-D COSY-90 ^{1}H NMR spectra of glycans.

Detailed analyses of microsporidial *O*-glycans have revealed linear oligosaccharides consisting of α1,2-linked Man residues. These structures are comparable with those identified in several fungi and share a common α1,2-linked mannotriose with α-linkage of the reducing terminal Man residue to an hydroxy amino acid (Gemmill and Trimble 1999; Willer et al. 2003). Further processing varies according to the species under consideration. In *S. cerevisiae*, the mannotriose is capped with one or two α1,3-linked Man residues, whereas in *Schizosaccharomyces pombe*, up to two galactose residues can be attached via α1,2- and α1,3-linkages. *Pichia pastoris* and *C. albicans* have oligosaccharides with only α1,2-linked Man residues. Thus, *E. cuniculi* and *A. locustae* share with the two last yeasts very similar linear *O*-mannosyl glycans, except that maximum glycan length is higher in microsporidia (up to eight Man residues). The immobilization of *E. intestinalis* exospore proteins SWP1 and SWP2 on ConA-agarose columns (Hayman et al. 2001) might be re-interpreted as due to the presence of O-mannosylated chains. However, why the same proteins were found to be also immobilized on WGA-agarose columns is still unclear. The *E. cuniculi* coding sequence ECU08_1340 was initially annotated as having similarity with the 110-kDa subunit of *O*-GlcNAc transferase (OGT) (Katinka et al. 2001). In fact, it is improbable that ECU08_1340 encodes such an enzyme, the partial homology concerning only tetratricopeptide repeats, not the catalytic C-terminal domain of OGT. As the microsporidian spore wall contains chitin but no β glucan, we tentatively suggest that binding of WGA to SWPs may be due to the presence of a GPI structure linked to chitin oligomers.

The finding of *O*-mannosyl glycans fits with biochemical evidence for O-mannosylation of *E. hellem* PTP1 (Xu et al. 2004) and metabolic potentials inferred from *E. cuniculi* genome sequence (Vivarès and Méténier 2004). Starting from Dol-P-Man in the ER rather than from a nucleotide sugar in the Golgi apparatus, the O-mannosylation pathway in fungal organisms (Ernst and Prill 2001) should be minimally represented in microsporidia. Only one gene encodes α1,2 mannosyltransferase (KTR family) in *E. cuniculi* (locus ECU04_1130; UniProtKB entry: Q8SS28_ENCCU) and probably also in *A. locustae* (http://gmod.mbl.edu/perl/site/antonospora01). This excludes the presence of partially redundant enzymes involved in the addition of the second and third α1,2-linked Man residues, as demonstrated in both *S. cerevisiae* (Romero et al. 1999) and *C. albicans* (Munro et al. 2005). The conserved PMT family of protein *O*-mannosyltransferases is divided into PMT1, PMT2, and PMT4 subfamilies and contains up to seven members in *S. cerevisiae*. The unique member of the PTP4 subfamily forms homomeric complexes, whereas members of the PMT1 subfamily interact heterophilically with those of the PMT2 subfamily, the PMT1/PMT2 and PMT4 members differing in protein substrate specificity (Girrbach and Strahl 2003; Willer et al. 2003). In *E. cuniculi*, only two potential PMTs are present. The protein sequence ECU02_1300 (Q825D9) is clearly representative of the PMT2 subfamily (50–52% similarity with *C. albicans* and *Aspergillus fumigatus* PMT2s). The highly divergent character of the other PMT candidate (ECU06_0950; Q8SVA5) does not allow assignation to a known PMT subfamily. A clear homolog of ECU02_1300 has been found in *A. locustae* (ORF 1659; http://gmod.mbl.edu/perl/site/antonospora01). The formation of a PMT1/PMT2-like complex in microsporidia remains debatable.

Further isolation and characterization of individual O-mannosylated proteins should be useful for a better knowledge of the molecular organization of the microsporidian invasion apparatus, including the polar cap, and of possible ligands involved in host–microsporidia cell interactions. Moreover, as microsporidian gene potentialities exist for glypiation (Vivarès and Méténier 2004) and some recently studied endospore proteins are likely GPI-anchored (Brosson et al. 2005; Peuvel-Fanget et al. 2005; Xu et al. 2006), physicochemical analyses are still needed to elucidate the precise structure of these GPI anchors.

Materials and methods

Isolation of microsporidian spores

E. cuniculi GB-M1, a mouse isolate that was used for genome sequencing (Katinka et al. 2001), was produced on MDCK cells as previously described (Delbac, Duffieux, et al. 1998). Parasite cells were released in culture media by natural lysis of heavily infected MDCK cells. Culture media were harvested and then parasite cells were sedimented and repeatedly washed with phosphate-buffered saline (PBS), pH 7.4, through 10 successive centrifugations (20 000g, 3 min) to remove host cell contaminants of low density. The final pellet of *E. cuniculi* cells was represented by late sporogonial stages including a majority of mature spores (more than 70%).

A. locustae spores, arising from infected grasshoppers, were commercially available from M&R Durango Insectary (Bayfield, CO). These spores were also washed in PBS, pH 7.4, prior to protein extraction.

EM cytochemistry

E. cuniculi and *A. locustae* spores were either epoxy resin-embedded or frozen. For resin embedding, spores were fixed for 1 h in 2% glutaralhehyde, 0.05% ruthenium red, and 0.07 M cacodylate buffer, pH 7.4. After washing for 30 min in 0.1 M cacodylate buffer, pH 7.4, they were postfixed for 1 h in 1% OsO_4, dehydrated through a graded series of ethanol, infiltrated in propylene oxide, and embedded in Epikote 812 resin (Agar Scientific, Essex, UK). Ultrathin sections were obtained with a Leica Ultracut S ultramicrotome, then classically stained with uranyl acetate and lead citrate. For ultracryotomy, spores were fixed for 1 h with 4% paraformaldehyde–0.1% glutaraldehyde in 0.1 M cacodylate buffer, pH 7.4. After infusion for 1 h at room temperature in a 25% glycerol–5% dimethylsulfoxide mixture, the samples were rapidly frozen in slush nitrogen. Cryosections (90 nm) were obtained using a dry sectioning device at −110 °C and Ultracut S ultramicrotome fitted with the low-temperature sectioning system FC4. Sectioned material was mounted on collodion-coated nickel grids (150 meshes) and stored at 4 °C in PBS prior to cytochemical protocols.

The localization of glycoconjugates and polysaccharides was investigated using the PATAg procedure (Thiéry 1967). Unstained sections of resin-embedded or frozen cells were submitted to oxidation with 1% periodic acid for 30 min, washed in distilled water, and incubated in 0.2% thiocarbohydrazide for 4 h. After washing with 10% acetic acid and then with distilled water, grids were treated with 1% silver proteinate for 30 min.

For the localization of lectin-binding sites, cryosections were saturated for 1 h with PBS-1% ovalbumin and then incubated for 1 h with a 1:50 dilution of biotinylated ConA or GNA lectins (EY Laboratories, San Mateo, CA). Grids were subsequently reacted for 1 h with goat antibiotin antibody (Sigma, St. Louis, MO) at 1:100, then for 1 h with 5 nm gold-conjugated antigoat IgG (Sigma) at 1:100. Cryosections were finally contrasted and protected with a 0.8% uranyl acetate–1.6% methylcellulose mixture. All specimens were examined under a JEOL 1200EX transmission EM.

Protein extractions

For 2-D electrophoresis, *E. cuniculi* spore proteins were extracted through repeated cycles of freezing–thawing in liquid nitrogen and sonication (15 × 1 min on ice) in the presence of 7 M urea, 2 M thiourea, 100 mM DTT, 4% CHAPS, and 0.2% SDS. For SDS–PAGE followed by lectin blotting, proteins were solubilized in either Laemmli buffer ("SDS extract") or a solution containing 4 M urea and 100 mM DTT ("DTT–urea extract").

For MS and NMR analyses, *E. cuniculi* or *A. locustae* spores (total cell number: 10^9–10^{10}) were disrupted in a lysis buffer containing 1% (v/v) Triton X-100 and 100 mM DTT, by repeated cycles of freezing–thawing and sonication, then incubated in the extraction solution (7 M urea, 2 M thiourea, 100 mM DTT, 1% Triton X-100) under agitation for 3 days. The entire sample was finally dialyzed (6000–8000 cutoff) for further 3 days.

2-D gel electrophoresis and glycoprotein detection

Isoelectrofocalization (IEF) of *E. cuniculi* protein samples (50 μg) was performed along linear immobilized pH gradient strips of 7 cm, pH 3–10 (GE Healthcare, Piscataway, NJ) in rehydration buffer (7 M urea, 2 M thiourea, 2% CHAPS, 2 mM tributyl phosphine, and 0.5% ampholytes), with the IPGPhor apparatus (GE Healthcare). The program of voltage increase was 30 V for 12 h, 400 V for 30 min, 500 V for 30 min, 800 V for 30 min, 1000 V for 1 h, 4000 V for 1 h, and 8000 V for 2 h. After equilibration with 50 mM Tris–HCl (pH 8.8), 6 M urea, 30% glycerol, 2% SDS, 100 mM DTT, and then 135 mM iodoacetamide, strips were deposited on 12% polyacrylamide slab gels. After SDS–PAGE for 1 h at 25 mA, proteins were either stained with Coomassie brilliant blue or electrophoretically transferred onto poly vinylidenedifluoride PVDF membranes (Millipore, Billerica, MA). Glycoprotein detection on these membranes was carried out with the BioRad Immun-Blot® kit, involving successively periodate oxidation of carbohydrate groups, biotinylation, incubation with streptavidin-alkaline phosphatase conjugate, and color development with 5-bromo-4-chloro-3-indolyl phosphate and nitro-blue tetrazolium chloride.

Lectin overlay

SDS and DTT–urea extracts of *E. cuniculi* were analyzed by SDS–PAGE (12% polyacrylamide) and separated proteins were transferred onto PVDF membranes. Blots were saturated in Tris buffer saline (TBS) (50 mM Tris–HCl, pH 7.4, 150 mM NaCl)–5% skimmed milk and washed in TBS and then in lectin reaction buffer (20 mM Tris–HCl, pH 7.4, 0.2 M NaCl, 1 mM $MgCl_2$, 1 mM $MnCl_2$, and 1 mM $CaCl_2$). They were subsequently incubated for 1 h with one of the two following biotin-labeled lectins (EY Laboratories): ConA diluted at 1:1500 and GNA at 1:1000. After washing in TBS, the membranes were reacted with a goat antibiotin antibody (Sigma) diluted at 1:1000 and finally with a peroxidase-conjugated antigoat IgG (Sigma) at 1:10 000. Lectin binding was visualized with a chemoluminescent system (ECL + Western blot detection kit, Amersham). Specificity was tested by preincubation of lectin conjugates with an ovalbumin blot at 4 °C overnight.

Isolation of glycan fractions

E. cuniculi and *A. locustae* extracts were delipidated by chloroform–methanol (2:1). The separated pellet was reduced

with 6 M guanidine chloride, 25 mM DTT in 0.6 M Tris for 8 h at 50 °C, alkylated with 50 mM iodoacetamide overnight, and dialyzed. The product (3 mL) was then digested with 1 mg of trypsin in 50 mM ammonium bicarbonate pH 8.4 at 37 °C for 1 day and with additional 0.2 mg of chymotrypsin at 37 °C under agitation for a night. A C18 Sep-Pak (Waters, Milford, CT) was used to purify the resulting supernatant. Putative *N*-linked glycans were released by digestion with PNGase F and PNGase A. Oligosaccharides were separated from peptides and glycopeptides using a C18 Sep-Pak. Putative *O*-linked glycans were released by alkaline reductive degradation in 1 M $NaBH_4$ and 0.1 M NaOH at 37 °C for 72 h. The reaction was stopped by the addition of Dowex 50 × 8, 25–50 mesh, H^+ form (Bio-Rad) at 4 °C until pH reaches 6.5, and after evaporation to dryness, borate salts were removed by repeated evaporation with methanol. Total material was subjected to cationic exchange chromatography on Dowex 50 × 2, 200–400 mesh, H^+ form (Bio-Rad) to remove residual peptides. The oligosaccharide fraction was then purified on a Bio-Gel P2 column (Bio-Rad) and C18 Sep Pak.

Monosaccharide composition

Monosaccharides were analyzed by GC–MS as perheptafluorobutyryl derivatives (Zanetta et al. 1999). Shortly, *N*- and *O*-oligosaccharides were subjected to methanolysis in 500 μL of 0.5 M HCl in anhydrous methanol at 80 °C for 20 h and incubated in 200 (L of anhydrous acetonitrile (ACN) and heptafluorobutyric acid (HFB) at 180 °C for 10 min. The reagents were evaporated, and the sample was dissolved in ACN prior to GC–MS analysis.

Permethylation and linkage analysis

Permethylation was performed according to the procedure of Ciucanu and Kerek (1984). Briefly, compounds were incubated overnight in a suspension of 200 mg/mL NaOH in dry dimethylsulfoxide (300 μL) and iodomethane (200 μL). The methylated products were extracted in chloroform and washed with water. After methanolysis, they were dried and then peracetylated in 200 μL of acetic anhydride and 50 μL of pyridine overnight at room temperature. The reagents were evaporated, and the sample was dissolved in chloroform before analysis in GC–MS.

MALDI-TOF and ES-MS^n

The molecular masses of *N*- and *O*-oligosaccharides were measured by MALDI-TOF on a Voyager Elite reflectron mass spectrometer (PerSeptive Biosystems, Framingham, MA), equipped with a 337 nm UV laser. Native and permethylated samples were prepared by mixing directly on the target 1 μL of water (native) or ACN (permethylated) diluted oligosaccharide solution and 1 μL of 2.5-dihydroxybenzoic acid matrix solution (10 mg/mL dissolved in ACN–H_2O). For electro-spray ionization multistage mass spectrometry (ES-MS^n), permethylated samples were reconstituted in methanol and analyzed by mass spectrometry on a LCQ DK XP + ion trap (Thermo Finnigan, Waltham, MA) instrument. After mixing with an equal volume of methanol–0.1 M aqueous formic acid, samples were directly infused at 50 nL/min, using the nanoflow probe option for MS and MS^n analyses.

^{1}H NMR spectroscopy

One-dimensional and 2-D ^{1}H NMR spectra were recorded on a Bruker Avance 800 spectrometer (Université des Sciences et Technologies de Lille), equipped with a TXI probe-head. Prior to ^{1}H NMR analyses, oligosaccharides were twice exchanged with 99.97% ^2H_2O and finally solubilized in 250 μL of ^2H_2O in 5-mm Shigemi tube matched for ^2H_2O. The spectrometer operated at 300 K without solvent presaturation. The chemical shifts were expressed relative to residual acetate salts (δ 1.909 ppm). Spectral width was 8012 Hz with 16 k points for a spectral resolution of 0.49 Hz/pt. The 2-D ^{1}H–^{1}H COSY (correlation spectroscopy) spectrum was acquired with *z*-gradient pulse from cosygp pulse program available in Bruker software. Spectral width was 8012 Hz for both dimensions with 4016 points for F2 and 256 points for F1 giving spectral resolution of 1.96 and 31.3 Hz/pt, respectively.

Acknowledgments

We thank Yves Leroy (CNRS UMR 8576) and Philippe Timmerman (CNRS UMR 8576) for technical assistance on carbohydrate composition. We are grateful to Dr Catherine Texier for her help in the interpretation of two-dimensional protein profiles.

Conflict of interest statement

None declared.

Abbreviations

ACN, acetonitrile; CID-MS/MS, collision-induced decay tandem mass spectrometry; ConA, concanavalin A; COSY, correlation spectroscopy; Dol, dolichol; DP, degree of polymerization; DTT, dithiothreitol; EI, electronic impact; EM, electron microscopy; ER, endoplasmic reticulum; ES, electro-spray; GC–MS, gas chromatography coupled to mass spectrometry; GNA, *Galanthus nivalis* agglutinin; GPI, glycosylphosphatidylinositol; HFB, heptafluorobutyric acid; IEF, isoelectrofocalization; MALDI-TOF, matrix assisted-laser desorption/ionization time of flight; MDCK, Madin–Darby canine kidney; NMR, nuclear magnetic resonance; OGT, *O*-GlcNAc transferase; PAS, periodic acid-Schiff; PATAg, periodic acid-thiocarbohydrazide-Ag proteinate; PBS, phosphate-buffered saline; PNGase A, peptide *N*-glycosidase A; PNGase F, peptide *N*-glycosidase F; PTP, polar tube protein; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; SWP, spore wall protein; TBS, Tris buffer saline; TIC, total ion chromatogram; WGA, wheat germ agglutinin; 1-D, one-dimensional; 2-D, two-dimensional.

References

Aravind L, Iyer LM, Wellems TE, Miller LH. 2003. *Plasmodium* biology: genomic gleanings. Cell. 115:771–785.

Bigliardi E, Selmi MG, Lupetti P, Corona S, Gatti S, Scaglia M, Sacchi L. 1996. Microsporidian spore wall: ultrastructural findings on *Encephalitozoon hellem* exospore. J Eukaryot Microbiol. 43:181–186.

Bohne W, Ferguson DJ, Kohler K, Gross U. 2000. Developmental expression of a tandemly repeated, glycine- and serine-rich spore wall protein in the microsporidian pathogen *Encephalitozoon cuniculi*. Infect Immun. 68: 2268–2275.

Brosson D, Kuhn L, Delbac F, Garin J, Vivares C, Texier C.. Forthcoming 2006. Proteomic analysis of the eucaryotic parasite *Encephalitozoon cuniculi* (Microsporidia): a reference map for proteins expressed in late sporogonial stages. Proteomics 6:3625–35.

Brosson D, Kuhn L, Prensier G, Vivarès CP, Texier C. 2005. The putative chitin deacetylase of *Encephalitozoon cuniculi*: a surface protein implicated in microsporidian spore-wall formation. FEMS Microbiol Lett. 247:81–90.

Ciucanu I, Kerek F. 1984. A simple and rapid method for the permethylation of carbohydrates. Carbohydr Res. 131:209–217.

Cohen RE, Ballou CE. 1980. Linkage and sequence analysis of mannose-rich glycoprotein core oligosaccharides by proton nuclear magnetic resonance spectroscopy. Biochemistry. 19:4345–4358.

Cutler JE. 2001. N-glycosylation of yeast, with emphasis on *Candida albicans*. Med Mycol. 39 (Suppl 1):75–86.

Delbac F, Duffieux F, David D, Metenier G, Vivares CP. 1998. Immunocytochemical identification of spore proteins in two microsporidia, with emphasis on extrusion apparatus. J Eukaryot Microbiol. 45: 224–231.

Delbac F, Peyret P, Méténier G, David D, Danchin A, Vivarès CP. 1998. On proteins of the microsporidian invasive apparatus: complete sequence of a polar tube protein of *Encephalitozoon cuniculi*. Mol Microbiol. 29: 825–834.

Dell A, Reason AJ, Khoo KH, Panico M, McDowell RA, Morris HR. 1994. Mass spectrometry of carbohydrate-containing biopolymers. Methods Enzymol. 230:108–132.

Domon B, Costello CE. 1988. Structure elucidation of glycosphingolipids and gangliosides using high-performance tandem mass spectrometry. Biochemistry. 27:1534–1543.

Ernst JF, Prill SK. 2001. O-glycosylation. Med Mycol. 39 (Suppl 1):67–74.

Faille C, Wieruszeski JM, Michalski JC, Poulain D, Strecker G. 1992. Complete 1H- and 13C-resonance assignments for D-mannooligosaccharides of the beta-D-(1–> 2)-linked series released from the phosphopeptidomannan of *Candida albicans* VW.32 (serotype A). Carbohydr Res. 236:17–27.

Gemmill TR, Trimble RB. 1999. Overview of N- and O-linked oligosaccharide structures found in various yeast species. Biochim Biophys Acta. 1426:227–237.

Girrbach V, Strahl S. 2003. Members of the evolutionarily conserved PMT family of protein O-mannosyltransferases form distinct protein complexes among themselves. J Biol Chem. 278:12554–12562.

Gowda DC, Gupta P, Davidson EA. 1997. Glycosylphosphatidylinositol anchors represent the major carbohydrate modification in proteins of intraerythrocytic stage *Plasmodium falciparum*. J Biol Chem. 272: 6428–6439.

Guerardel Y, Balanzino L, Maes E, Leroy Y, Coddeville B, Oriol R, Strecker G. 2001. The nematode *Caenorhabditis elegans* synthesizes unusual O-linked glycans: identification of glucose-substituted mucin-type O-glycans and short chondroitin-like oligosaccharides. Biochem J. 357:167–182.

Guha-Niyogi A, Sullivan DR, Turco SJ. 2001. Glycoconjugate structures of parasitic protozoa. Glycobiology. 11:45R–59R.

Hayette MP, Strecker G, Faille C, Dive D, Camus D, Mackenzie DW, Poulain D. 1992. Presence of human antibodies reacting with *Candida albicans* O-linked oligomannosides revealed by using an enzyme-linked immunosorbent assay and neoglycolipids. J Clin Microbiol. 30: 411–417.

Hayman JR, Hayes SF, Amon J, Nash TE. 2001. Developmental expression of two spore wall proteins during maturation of the microsporidian *Encephalitozoon intestinalis*. Infect Immun. 69:7057–7066.

Hayman JR, Southern TR, Nash TE. 2005. Role of sulfated glycans in adherence of the microsporidian *Encephalitozoon intestinalis* to host cells in vitro. Infect Immun. 73:841–848.

Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P, et al. 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature. 414:450–453.

Keeling PJ. 2003. Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. Fungal Genet Biol. 38:298–309.

Kim JH, Ogawa K, Wakabayashi H. 1999. Lectin-reactive components of the microsporidian *Glugea plecoglossi* and their relation to spore phagocytosis by head kidney macrophages of ayu *Plecoglossus altivelis*. Dis Aquat Organ. 39:59–63.

Kobayashi H, Komido M, Watanabe M, Matsuda K, Suzuki M, Ikeda T, Oyamada H, Shibata N, Suzuki S. 1994. Structure of cell wall mannan of *Candida kefyr IFO 0586*. Infect Immun. 62:4425–4431.

Kocharova NA, Knirel YA, Widmalm G, Jansson PE, Moran AP. 2000. Structure of an atypical O-antigen polysaccharide of *Helicobacter pylori* containing a novel monosaccharide 3-C-methyl-D-mannose. Biochemistry. 39:4755–4760.

Lom J. 1972. On the structure of the extruded microsporidian polar filament. Zeitung Parasitenkund. 38:200–213.

Méténier G, Vivarès CP. 2004. Genomics of microbial parasites: the microsporidial paradigm. In: Hirt RP, Horner DS, editors. Organelles, genomes and eukaryote phylogeny. An evolutionary synthesis in the age of genomics. Boca Raton (FL): CRC Press; p. 207–236.

Munro CA, Bates S, Buurman ET, Hughes HB, MacCallum DM, Bertram G, Atrih A, Ferguson MA, Bain JM, Brand A, et al. 2005. Mnt1p and Mnt2p of *Candida albicans* are partially redundant alpha-1,2-mannosyltransferases that participate in O-linked mannosylation and are required for adhesion and virulence. J Biol Chem. 280:1051–1060.

Peek R, Delbac F, Speijer D, Polonais V, Greve S, Wentink-Bonnema E, Ringrose J, van Gool T. 2005. Carbohydrate moieties of microsporidian polar tube proteins are targeted by immunoglobulin G in immunocompetent individuals. Infect Immun. 73:7906–7913.

Peuvel-Fanget I, Polonais V, Brosson D, Texier C, Kuhn L, Peyret P, Vivares C, Delbac F. 2006. EnP1 and EnP2, two proteins associated with the *Encephalitozoon cuniculi* endospore, the chitin-rich inner layer of the microsporidian spore wall. Int J Parasitol. 36:309–318.

Polonais V, Prensier G, Méténier G, Vivarès CP, Delbac F. 2005. Microsporidian polar tube proteins: highly divergent but closely linked genes encode PTP1 and PTP2 in members of the evolutionarily distant *Antonospora* and *Encephalitozoon* groups. Fungal Genet Biol. 42: 791–803.

Romero PA, Lussier M, Veronneau S, Sdicu AM, Herscovics A, Bussey H. 1999. Mnt2p and Mnt3p of *Saccharomyces cerevisiae* are members of the Mnn1p family of alpha-1,3-mannosyltransferases responsible for adding the terminal mannose residues of O-linked oligosaccharides. Glycobiology. 9:1045–1051.

Samuelson J, Banerjee S, Magnelli P, Cui J, Kelleher DJ, Gilmore R, Robbins PW. 2005. The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. Proc Natl Acad Sci USA. 102:1548–1553.

Shibuya N, Goldstein IJ, Van Damme EJ, Peumans WJ. 1988. Binding properties of a mannose-specific lectin from the snowdrop (*Galanthus nivalis*) bulb. J Biol Chem. 263:728–734.

Slamovits CH, Williams BA, Keeling PJ. 2004. Transfer of *Nosema locustae* (microsporidia) to *Antonospora locustae* n. comb. based on molecular and ultrastructural data. J Eukaryot Microbiol. 51:207–213.

Thiéry JP. 1967. Mise en évidence des polysaccharides sur coupes fines en microscopie électronique. J Microscopie. 6:987–1018.

Thomarat F, Vivarès CP, Gouy M. 2004. Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. J Mol Evol. 59:780–791.

Trimble RB, Lubowski C, Hauer CR III, Stack R, McNaughton L, Gemmill TR, Kumar SA. 2004. Characterization of N- and O-linked glycosylation of recombinant human bile salt-stimulated lipase secreted by *Pichia pastoris*. Glycobiology. 14:265–274.

Vavra J. 1972. Detection of polysaccharides in microsporidian spores by means of the periodic acid-thiosemicarbazide-silver proteinate test. J Microscopie. 14:357–360.

Vavra J, Larson J. 1999. Structure of the microsporidia. In: Wittner M, Weiss LM, editors. The microsporidia and microsporidiosis. Washington (DC): ASM Press; p. 7–84.

Vivarès CP, Méténier G. 2004. The microsporidia genome: living with minimal genes as an intracellular eukaryote. In: Lindsay DS, Weiss LM, editors. Opportunistic infections: toxoplasma, sarcocystis, and

microsporidia. World class parasites. Boston (MA): Kluwer Academic Pub; p. 215–242.

Weiss LM. 2001. Microsporidia: emerging pathogenic protists. Acta Trop. 78:89–102.

Willer T, Valero MC, Tanner W, Cruces J, Strahl S. 2003. O-mannosyl glycans: from yeast to novel associations with human disease. Curr Opin Struct Biol. 13:621–630.

Xu Y, Takvorian PM, Cali A, Orr G, Weiss LM. 2004. Glycosylation of the major polar tube protein of *Encephalitozoon hellem*, a microsporidian parasite that infects humans. Infect Immun. 72:6341–6350.

Xu Y, Takvorian P, Cali A, Wang F, Zhang H, Orr G, Weiss LM. 2006. Identification of a new spore wall protein from *Encephalitozoon cuniculi*. Infect Immun. 74:239–247.

Xu Y, Weiss LM. 2005. The microsporidian polar tube: a highly specialised invasion organelle. Int J Parasitol. 35:941–953.

Zanetta JP, Timmerman P, Leroy Y. 1999. Gas–liquid chromatography of the heptafluorobutyrate derivatives of the O-methyl-glycosides on capillary columns: a method for the quantitative determination of the monosaccharide composition of glycoproteins and glycolipids. Glycobiology. 9: 255–266.

Etude de la biosynthèse des oligomannosides de Toxoplasma gondii

Dans un contexte scientifique très similaire à celui de l'étude de la glycosylation des microsporidies, nous avons engagé une collaboration début 2006 avec le Dr. Ralph Schwarz (Philipps University Marburg, Allemagne) concernant l'étude du profil de glycosylation du parasite *Toxoplasma gondii*. En effet, malgré les nombreuses études dont il fait l'objet et le problème sanitaire majeur dont il est la source, les modes de glycosylation de ce parasite sont méconnus. La difficulté majeure de son étude provient du statut de parasite intracellulaire strict de *T. gondii* qui engendre des contaminations importantes de la part de la cellule hôte lors des analyses biochimiques. Notre stratégie est basée sur une étude multifactorielle de la glycosylation dans le but de suivre la biosynthèse des glycoconjugués et ce à chaque étape du processus de glycosylation. Pour ce faire, nous suivons l'incorporation de différents précurseurs radioactifs (mannose, mannosamine) au sein des glycoprotéines et glycolipides, l'apparition des précurseurs biosynthétiques (dol-P-P-glycannes) des N-glycannes par extraction organique et couplages fluorescents et des produits finaux (N-glycannes) par spectrométrie de masse. L'ensemble de ces expériences sont réalisées en parallèle sur des parasites cultivés dans deux types de lignées cellulaires (vero et HFF) et sur les cellules non infectées (vero et HFF).

Les résultats d'incorporation de précurseurs indiquent l'existence d'une biosynthèse indépendante de N-glycannes par *T. gondii*. De plus, les parasites libres, expulsés après lyses cellulaires sont également capables d'incorporer des précurseurs biosynthétiques. L'analyse par spectrométrie de masse indique que les glycoprotéines de *T. gondii* sont uniquement substituées par des N-glycannes de type oligomannosyl alors que les cellules hôtes contiennent également des types complexes. Ces résultats sont néanmoins en contradiction partielle avec des recherches bio-informatiques qui indiquent qu'une partie des homologues des gènes responsables de la biosynthèse des N-glycannes (ALG3, ALG9 et ALG12) est absente du génome de *T. gondii*. Pour résoudre ce problème nous avons infecté des cellules hôtes déficientes dans la synthèse de Dolichol Phospho-mannose Synthase (lignée B3F7) par *T. gondii* dont les N-glycannes ont été analysés par spectrométrie de masse avant et après dégradation enzymatique. Ces expériences nous ont permis de démontrer que le parasite possède un système de N-glycosylation hybride unique en son genre. En effet, d'une part il synthétise et transfert de manière indépendante sur ses protéines des N-glycannes tronqués de structures anormales. D'autre part, il utilise également les intermédiaires de biosynthèse glycolipidiques de la cellule hôte pour transférer des N-glycannes de type eucaryote identiques à ceux de la cellule hôte sur ses propres protéines.

The dual origin of *Toxoplasma gondii N*-Glycans

Estelle Garénaux [‡], Hosam Shams-Eldin [§], Frederic Chirat [‡], Ulrike Bieker [§], Jörg Schmidt [§], Jean-Claude Michalski [‡], René Cacan [‡], Yann Guérardel [‡,*] and Ralph T. Schwarz [‡, §]

[‡] Unité de Glycobiologie Structurale et Fonctionnelle, UMR 8576 CNRS, Université des Sciences et Technologies de Lille, 59655, Villeneuve d'Ascq cedex, France

[§] Institut für Virologie, AG Parasitologie, Philipps-Universität Marburg, Hans-Meerwein-Strasse 2, 35043 Marburg, Germany.

* To whom correspondence should be addressed.
e. mail: yann.guerardel@univ-lille1.fr;
tel : +33(0)320336347
Fax : +33(0)320436555

Running title: *Toxoplasma gondii* synthesizes truncated *N*-glycans but imports mature *N*-glycans from host cells

Abbreviations

GPI, glycosylphosphatidylinositol, DolPPOS, Dolichol-PP-oligosaccharide, ER, endoplasmic reticulum, OST, oligosaccharyltransferase, Man, mannose, Glc, Glucose, GlcNAc, N-acetyl-glucosamine, MS, mass spectrometry.

Abstract

N-linked glycosylation is the most frequent modification of secreted proteins in eukaryotic cells that plays a crucial role in protein folding and trafficking. Mature *N*-glycans are sequentially processed in endoplasmic reticulum and Golgi apparatus through a pathway highly conserved in most eukaryotes organisms. Here, we demonstrate that the obligate intracellular protozoan parasite *Toxoplasma gondii* independently transfers endogenous truncated as well as other host derived *N*-glycans onto its own proteins. Therefore, we propose that the apicomplexan parasite scavenges *N*-glycosylation intermediates from the host cells to compensate the rapid evolution of its biosynthetic pathway, which is primarily devoted to modification of proteins with glycosylphosphatidylinositols (GPIs) rather than *N*-glycans.

Introduction

In eukaryotes, asparagine-linked glycosylation (ALG) is initiated in the endoplasmic reticulum- (ER)-membrane by the synthesis of lipid-linked oligosaccharides, the Dol-PP-oligosaccharide (DolPPOS) (Fig.1A) (*1*, *2*, *3*, *4*, *5*). The precursor oligosaccharide $Glc_3Man_9GlcNAc_2$-PP-Dol is assembled in a stepwise manner by the sequential actions of specific glycosyltransferases (Alg1–Alg12) and glycosidases. The five first mannose residues are transferred from GDP-Man to the chitobiose core, leading to the original structure $Man_5GlcNAc_2$. After flipping the oligosaccharide to the luminal side of the ER (*6, 7*) four other mannose residues are attached, provided by Dol-P-Man and three glucose residues from Dol-P-Glc. The complete oligosaccharide moiety ($Glc_3Man_9GlcNAc_2$) is then transferred to selected asparagines of the nascent proteins on the consensus motif Asn-X-Ser/Thr (Fig. 1A). This central reaction is catalyzed by the hetero-oligomeric protein oligosaccharyltransferase (OST) complex (*8*), which contains a catalytic subunit, STT3 (*9*, *10*).

Whereas many reports are available concerning *N*-glycosylation in free-living organisms, information regarding obligate intracellular parasites is still fragmentary. The data concerning ALG genes in several eukaryotic organisms suggest that database mining may allow the prediction of DolPPOS structures synthesized by any organism (*11*). This bioinformatic approach has been verified in several unicellular organisms such as *Trypanosoma cruzii, Tetrahymena thermophilia* or

Leishmania major. Interestingly the capacity of classical *O*-linked and *N*-linked glycosylations appears to be greatly reduced in some species exhibiting an obligate intracellular lifestyle, for example, the apicomplexan parasite *Plasmodium falciparum* contains low levels of *N*-glycosylation capability (*12, 13, 14*) and two microsporidian species, *Encephalitozoon cuniculi* and *Antonospora locustae,* seem to have lost the ability to synthesize *N*-linked glycans through the secondary losses of several key enzymes of *N*-glycosylation pathway (*15, 16*). Similar to *P. falciparum*, glycosylphosphatidylinositol (GPI) anchors represent the major carbohydrate modification of proteins in *Toxoplasma gondii* (*17, 18, 14*). *T. gondii* is a coccidian parasite found worldwide, that infects a wide range of warm-blooded vertebrates and has emerged as an important opportunistic pathogen for immunocompromised persons, for example coinfection with HIV often leads to fatal encephalitis (*19, 20, 21*). *T. gondii* shares common features with other apicomplexan parasites and represents a promising model for the study of the biosynthesis and role of glycans in the apicomplexa. The first direct biochemical evidence for *N*-glycosylation of the proliferative-stage, tachyzoite, glycoprotein gp23 provided impetus for the study of *N*-glycosylation in *T. gondii* (*22*). Furthermore, it was recently reported that *N*-glycosylation is essential for successful infection (*23*). It's worthy mentioning that lacking of solid structural data represents a major obstacle to deciphering glycoconjugate metabolism using bioinformatic tools. *T. gondii* clearly expresses a functional STT3p (*24*), however genome analysis indicates that orthologs of ALG3, ALG9 and ALG12 are absent. These correspond to the set of luminal mannosyltransferases that use Dol-P-Man as the sugar donor (Table S1, *13, 25*), *http://www.toxodb.org/toxo/home.jsp*), indicating a profound modification in the *N*-glycan biosynthetic pathway of the parasite. Thus, based on the presence of a specific subset of orthologous ALG genes, *T. gondii* possesses the enzymatic equipment for the biosynthesis of truncated $Glc_3Man_5GlcNAc_2$-PP-Dol, but not complete $Glc_3Man_9GlcNAc_2$-PP-Dol oligosaccharide precursors (Fig. 1B). However, the exact nature of oligosaccharides transferred onto nascent parasite proteins is unknown.

Experimental Procedures

Cell and parasite cultures

T. gondii grown in Human Foreskin Fibroblasts or African green monkey kidney cells (VERO cells ATCC CCL-81), Human foreskin fibroblasts (HFF; ATCC CRL-1635), CHO Pro-5 and CHO B3F7 were cultured in DMEM (Gibco BRL), supplemented with 10% FCS (Gibco), 2 mM glutamine, 100 units/ml penicillin and 0.1 mg/ml streptomycin. Parasites (5×10^7) were added to confluent monolayers cells (T175 cm^2), harvested after 72 h of cultivation and liberated from their host cells using Mixer Mill homogenizer (Retsch). The suspension was run through a 20 ml glass wool column to remove cellular debris. The purity of the tachyzoite suspension was monitored microscopically. Cell lines and parasites were routinely tested for *Mycoplasma* contamination.

To control efficiency of parasites purification, *T. gondii* grown in CHO B3F7 liberated from their host cells were mixed with homogenized Vero-cells and the mixture was purified using glass-wool coluns as described before.

Strains and media

The *S. cerevisiae* and *E. coli* strains used in this work were: *YPH* 499 [Mat a; *ura* 3- 52; *lys* 2-801*amber; ade* 2- 101*ochre; trp* 1- 63; *his* 3- 200; *leu* 2- 1] (Stratagene) and *E. coli* strain *XL*1-*blue* (Stratagene), which was used for subcloning and other standard recombinant DNA procedures. *S. cerevisiae* strains were grown in YPAD medium [1% (w/v) Bacto yeast extract, 2% (w/v) Bactopeptone, 2% (w/v) dextrose, 4 mg l−1 adenine] or SD medium (0.17% Bacto yeast nitrogen base, 0.5% ammonium sulphate and 2% dextrose) containing the nutritional supplements necessary to complement strain auxotrophs or allow selection of transformants. *YPH499–HIS–GAL–ALG7* was maintained on SGR medium (4% galactose, 2% raffinose, 0.17% Bacto yeast nitrogen base, 0.5% ammonium sulphate) in which dextrose is replaced by galactose/raffinose as a source of carbohydrates. *E. coli* strains were grown in LB medium.

Extraction of glycoproteins and preparation of N-glycans

Collected *T. gondii* were homogenized by ultrasonic treatment at 4°C. Lipids were extracted by sequential extraction with 20 volumes of chloroform/methanol (2:1, v/v), then with a chloroform/methanol/water (40:20:3) solution.

Delipidated homogenates were suspended in a solution of 6 M guanidinium chloride and 5 mM Ethylene Diamine Tetraacetic Acid (EDTA) in 0.1 M Tris/HCl, pH 8, and agitated for 4 h at 4°C. Dithiothreitol was then added to a final concentration of 20 mM and incubated for 5 h at 37°C,

followed by the addition of iodoacetamide to a final concentration of 50 mM and further incubated overnight in the dark at room temperature. Reduced/alkylated sample was dialyzed against water at 4°C for 3 days and lyophilized. The recovered protein samples were then sequentially digested by (TPSK)-treated trypsin overnight at 37°C, in 50 mM ammonium bicarbonate buffer, pH 8.4 and by chimotrypsine.

Crude peptides/glycopeptides were loaded onto a Sep-Pak C18 cartridge. Glucan polymers and other hydrophilic contaminants were washed off with 5% aqueous acetic acid and the bound peptides/glycopeptides were eluted with a step gradient of 20, 40 and 60% 1-propanol in water. Eluted fractions were pooled, dried down and then incubated with *N*-glycosidase F (Roche, Basel, Switzerland)) overnight at 37°C in 50 mM ammonium bicarbonate buffer, pH 8.4. Released *N*-glycans were separated from peptides/glycopeptides using the same C18 Sep-Pak procedure. Pooled propan-1-ol fractions were then digested with *N*-glycosidase A from almond (0.5 mU, Calbiochem) in 50 mM ammonium acetate buffer, pH 5, at 37°C overnight. *N*-glycans released were likewise separated from the peptides by the same Sep-pak C18 procedures.

Chemical derivatization

For MALDI-MS analyses, the glycan samples were permethylated using the NaOH/dimethyl sulfoxide slurry method (*26*). The permethyl derivatives were then extracted in chloroform and repeatedly washed with water.

Coupling oligosaccharides with 2-aminopyridine

After hydrolysis, oligosaccharides were reductively aminated with 2-aminopyridine at the reducing end (*27*). Then, to remove excess reagent and purify 2-PA oligosaccharides, two methanol solutions (75% and 85%) are successively added to the reactional mixture and dry under a nitrogen stream. After adding 1ml water, pH is adjusted at 10 by adding NH_4OH (25%). Aqueous phase is washed 10 times with chloroform to eliminate excess 2-AP. Aqueous phase is then transfered into a clean tube and pH neutralized with glacial acetic acid before lyophilization. Finally, derivatized oligosaccharides are purified by SPE onto a Sep-Pack C18 column.

Mannosidase treatment of 2-aminopyridinylated N-glycans

2-PA coupled PNGase F released-oligosaccharides were digested with alpha-Mannosidase from *Aspergillus saitoï (*GKX5009 Glyko®). Each sample was treated with α(1,2) mannosidase in 100 mM sodium acetate, 2mM Zn^{2+} pH 5.0, and incubated overnight at 37°C. The mixture was finally applied on a Sep-Pak C18 to purify 2-aminopyridinylated oligosaccharides.

MS analyses of glycans

For MALDI-time-of-flight (MALDI-TOF) MS glycan profiling, native compounds in water were mixed 1:1 with 2,5-dihydroxybenzoic acid (DHB) matrix (10 mg/mL in MeOH/Water 50:50) spotted on the target plate and dried under vacuum. Data acquisition was performed manually on a Voyager (Applied biosystem) operated in the reflectron mode. Laser shots were accumulated until a satisfactory signal to noise ratio was achieved when combined and smoothed.

Extraction of Lipid Linked Oligosaccharides

Extractions were performed according to (*28*). Briefly, cell pellets were extracted with 2ml of chloroform/methanol/H_2O (3:2:1). After centrifugation, upper phase and proteins were extracted twice with 1.5 mL of theoretical lower phase (MeOH 70mL, MgCl2 4mM 5mL, chloroform 430mL). Then proteins were washed 4 times with 3 mL of theoretical upper phase (MeOH 240mL, H_2O 225mL, $MgCl_2$ 100mM 9.4mL, CH_3Cl_3 15mL). After partial drying of pellets Dol-PP-OS were extracted by 1.5 mL of CH_3Cl_3/MeOH/H_2O (10:10:3; v/v) mixture. Dried Dol-PP-OS were hydrolysed by THF/H_2O/HCl (9:1:0.083; v/v) at 50°C for 2h. Released oligosaccharides are dried under a stream of nitrogen and purified on a carbograph column (Alltech carbograph SPE Column).

Results and discussion

The alg7 gene encodes GlcNAc-1-P transferase that initiates biosynthesis of *N*-Glycan precursors. Therefore, as an initial step toward an evaluation of genetic regulation of the *T. gondii* *N*-glycosylation, the alg7 gene homolog from *T. gondii* was identified. Initially, a 338bp fragment of putative *T. gondii* alg7 was identified by BLAST search of the EST database (*http://www.toxodb.org/toxo/home.jsp*) using the sequences of known orthologs from other organisms. This was subsequently used to screen a *T. gondii* Lambda ZAP cDNA library and a clone containing a 1503 bp open reading frame predicted to encode 54 KDa type III transmembrane protein was isolated (EMBL Data Bank Accession No. AJ436993). The deduced protein sequence shows 36.3% and 43% identity to amino acid sequences of the *S. cerevisiae*- and *Homo sapiens*-homologue proteins, respectively. Furthermore, the putative *T. gondii* alg7 was able to complement a conditional lethal yeast mutant (Fig S1 A and B) strongly indicating that it encodes a functional GlcNAc-1 transferase. The fact that *T. gondii* possesses genes coding for at least two key enzymes responsible for the initiation of DolPPOS precursor biosynthesis (alg7) and the transfer of the final products to nascent proteins (STT3) (*24*), strongly suggests that the parasite exhibits a functional, although possibly truncated, *N*-glycosylation pathway. The presence of *de novo* glycosylation pathway in the parasite was confirmed by observing differences between the DolPPOS profiles of the parasite and those of its host cells. Whereas the profile of DolPPOS oligosaccharide moieties from control African green monkey kidney epithelial (Vero) cells is characterized by prominent $Man_8GlcNAc_2$, $Man_9GlcNAc_2$ and $Glc_{1-3}Man_9GlcNAc_2$ (Fig. S2), *T. gondii* tachyzoites exhibit a simpler pattern containing only $Hex_8GlcNAc_2$ and $Hex_9GlcNAc_2$. The presence of precursors with more than eight hexoses does not correlate with the gene repertoire of the parasite, whose genome is likely to lack a number of genes required for the biosynthesis of *N*-glycan (Fig.1B). To reconcile these apparently contradictory observations, we defined the nature of the final *N*-glycans biosynthetic products using a glycomic approach, releasing glycans from the total pool of proteins from *T. gondii* grown in different host cells then analyzing their distribution using mass spectrometry (MS). Surprisingly, and in contrast to host cells, tachyzoites grown in Vero cells were found to lack complex-type *N*-glycans (Fig. 2). Similar results were obtained with *T. gondii* cultivated in HFF cells (Fig. S3). *T. gondii* exhibited m/z values consistent with the presence of oligomannosylated type *N*-glycans ranging from $Man_3HexNAc_2$ to $Man_9HexNAc_2$ that are typically observed in mammalian cells. Presence of higher mannosylated glycans on parasitic proteins is again in contradiction with bioinformatic predictions, leaving in question origin of these compounds.

To begin to address this discrepancy, we cultivated tachyzoites in a CHO mutant cell line deficient in Dol-P-Man synthase, B3F7 (*29*). These cells synthesize $Glc_{1-3}Man_5GlcNAc_2$-PP-Dol,

instead of the classical $Glc_{1-3}Man_9GlcNAc_2$-PP-Dol lipidic precursors, and transfers truncated $Glc_3Man_5GlcNAc_2$ onto their newly synthesized proteins (Fig. 1C; *30*). Accordingly, *N*-glycosylation profile of host B3F7 cell line is dominated by truncated $Man_3GlcNAc_2$ to $Man_5GlcNAc_2$ as well as complex-type *N*-glycans that are likely to result from the trimming and processing of $Glc_3Man_5GlcNAc_2$ (Fig. S4A and Fig. 3C) (*30*). $Hex_{6-8}GlcNAc_2$ signals were tentatively attributed to glucosylated $Glc_{1-3}Man_5GlcNAc_2$ and which are found on proteins due to their large accumulation in mutant cells. Both *T. gondii* and B3F7 cells are expected to synthesize similar truncated glycans as the result of the deficiency in different activities: Dol-P-Man synthase in B3F7 and luminal mannosyltransferases using Dol-P-Man in *T. gondii*. If the synthesis of $Man_9GlcNAc_2$ observed in *T. gondii* grown in Vero and HFF cells was the result of the complementation of parasite-deficient ALG3, ALG9 and ALG12 enzymatic activities by the host cell, we would then also find this glycan in *T. gondii* grown in B3F7, which also harbors these enzymes. However, *T. gondii* grown in B3F7 mutant cells synthesize oligosaccharides varying from $Hex_3GlcNAc_2$ to $Hex_8GlcNAc_2$ (Fig. S4B and Fig. 3E) but no $Hex_9GlcNAc_2$ (Fig 2B, Fig. S3B). In all infected host cell types analyzed no complex-type was observed to modify parasite proteins. The specific absence of $Hex_9GlcNAc_2$ signals in *T. gondii* grown in B3F7 cells compared with other host cells established that parasite glycosylation capacities are somehow dependant of the host. However, the parasite does not use its host enzymatic repertoire to complement its own incomplete repertoire. Based on the known mammalian glycosyltransferase repertoire, the $Hex_{5-8}HexNAc_2$ signals observed in Vero and HFF cells are classical branched mature *N*-glycans derived from the trimming of $Glc_3Man_9GlcNAc_2$, while in B3F7 cells, they largely correspond to truncated $Glc_{0-3}Man_5GlcNAc_2$. It is noteworthy that the mutant (B3F7) does not have a complete deficiency in the DolPMan synthase activity as minute, but measurable, amounts of $Man_9GlcNAc_2$ were also detected.

Finally, to assess the relative contributions of the parasite and host cells to *T. gondii N*-glycosylation, we looked for the precise origin of the protozoan *N*-glycans in various host cells. The exquisite specificity of *Aspergillus saitoi* exo-mannosidase toward (α1-2) linkage allows the differentiation of the Golgi-degradation products (Fig. 3G) from the linear ER related compounds (Fig 3H) and the glucose containing truncated glycans (Fig. 3I). The efficiency of this enzymatic treatment was evaluated by MS using a standard mixture of high mannose-type chains ranging in size from $Man_6GlcNAc_2$ to $Man_9GlcNAc_2$. This demonstrated that glycans ranging from $Man_7GlcNAc_2$ to $Man_9HexNAc_2$ were completely degraded to branched $Man_5HexNAc_2$, with a small amount degraded to $Man_6GlcNAc_2$ (Fig S5). On the other hand, treatment of *N*-glycans isolated from tachyzoites grown in B3F7 cells demonstrated complete degradation of $Man_4GlcNAc_2$

and $Man_5GlcNAc_2$ to generate $Man_3GlcNAc_2$. This indicated that these two oligosaccharides are ER-derived compounds (Fig. 3E and 3F). Furthermore, resistance of $Hex_6GlcNAc_2$, $Hex_7GlcNAc_2$ and $Hex_8GlcNAc_2$ to mannosidase activity, as well as absence of a branched $Man_5GlcNAc_2$ degradation product establishes that these compounds are $Glc_{1\text{-}3}Man_5GlcNAc_2$ truncated glycans. A similar strategy was employed to confirm that truncated *N*-glycans are the major type of glycosylation in B3F7 cells (Fig. 3C and 3D). It also confirmed the presence of small amounts of branched glycans, not observed in *T. gondii,* due to residual DolPMan synthase activity in B3F7 cells, as already suggested by the presence of $Man_9GlcNAc_2$ signals. Taken together, these data show that *T. gondii* transfers truncated *N*-glycans onto newly synthesized proteins. However, since both parasite and B3F7 mutant cells may synthesize identical *N*-glycans, therefore the final observed products may originate either from DolPPOS *de novo* synthesized by the parasite or scavenged from the host cells.

As expected, mannosidase treatment of Vero cell-derived *N*-glycans eliminated $Man_{7\text{-}9}GlcNAc_2$ confirming that the all mature *N*-glycans are indeed Golgi-modified products (Fig. S6), as observed with the standard *N*-glycans, their complete degradation generated a prominent branched $Man_5GlcNAc_2$ and minor $Man_6GlcNAc_2$. In contrast, $Hex_{7\text{-}9}GlcNAc_2$ isolated from *T. gondii* grown in Vero cells was partially resistant to digestion by mannosidase, indicating the presence of truncated $Glc_{1\text{-}3}Man_5GlcNAc_2$ glycans (Fig. 3A and 3B). However, concomitant marked increase of branched $Man_5GlcNAc_2$ and disappearance of $Hex_9GlcNAc_2$ clearly indicate the presence of branched $Man_{6\text{-}9}GlcNAc_2$, as well. To confirm that the $Hex_9GlcNAc_2$ signal originated from *T. gondii* rather then host-cell contaminant, parasites released from B3F7 cells were mixed with homogenized Vero cells and, after purification as described before, the *N*-glycan profile was analyzed. This analysis revealed an absence of $Hex_9GlcNAc_2$, thereby indicating that any contamination from host cells is below the level of detection (Fig. S7).

Analyses of *N*-glycans extracted from tachyzoites grown in different host cells clearly demonstrated that the parasite is able to synthesize, as well as transfer, truncated, immature oligomannosylated *N*-glycans to parasite proteins (*23*), in a host independent manner. These results are in agreement with *in silico* prediction of *T. gondii* glycosyltransferase pathway as well as, the presence of key enzymes of the dolichol pathway in the parasite (*11, 25*). However, *T. gondii* also transfers classical *N*-glycans identical to those of its host cells, as indicated by the presence of Golgi-derived oligomannosylated-types, when tachyzoites were grown in Vero and HFF, but not in DolPMan synthase mutant B3F7 cells. These glycans are not synthesized by the parasite through functional complementation of the incomplete *T. gondii* glycosyltransferase pathway by host cell enzymes, since parasites grown in B3F7 cells (which still maintain this activity) do not synthesize

mature *N*-glycans. Therefore, we hypothesize that *T. gondii* scavenges DolPPOS from its host cell in a non-specific manner and subsequently transfers its oligosaccharidic moieties onto its own newly-synthesized proteins, as evidenced by the presence of functional STT3p (*24*). Once transferred, the destiny of these glycans remains unclear; however they seem to undergo very limited processing in contrast to mammalian cells. Indeed, in contrast to B3F7 mutant cells that extensively process the $Glc_{1-3}Man_5GlcNAc_2$ precursor glycans toward $Man_4GlcNAc_2$ and complex types, *T. gondii* grown in these cells transfer mainly unprocessed $Glc_{1-3}Man_5GlcNAc_2$ onto nascent proteins. This is in good agreement with the apparent lack of most proteins involved in the quality control of newly synthesized proteins in *T. gondii* (*31*).

To our knowledge, this represents the first documented example of the scavenging of glycan intermediates from the host cell by an intracellular parasite. In a similar way, *T. gondii* has been shown to mobilize selected host lipids to fullfil its high metabolic requirements during proliferation (*32*, *33*, *34*). These exchanges could be facilitated by the close association of the parasitophorous vacuole membrane with the host endoplasmic reticulum, where the early steps of *N*-glycan biosynthesis take place. From an evolutionary perspective, it is noteworthy that the absence of Dol-P-Man dependent mannosyltransferase activities in the *N*-glycosylation pathway is counterbalanced by the synthesis of very large quantities of GPI-type glycosylation (*35*). Considering the postulated common evolutionary origin of ALG enzymes and PIG enzymes involved in the biosynthesis of phosphatidyl-inositol glycans (*36*), one may postulate that Dol-P-Man dependent mannosyltransferases could have rapidly evolved toward the exclusive synthesis of GPI in *T. gondii*. The presence of such enzymes in apicomplexan protozoa is in agreement with this hypothesis (*37*). The opportunity for the parasite to use *N*-glycan precursors synthesized by the host could then have lifted the selective pressure to keep an intact reticular *N*-glycosylation pathway, acting as a powerful drive for the evolution of glycosylation machinery in intracellular obligatory parasites.

Acknowledgments

This research was supported by CNRS (Centre National de la Recherche Scientifique), Deutsche Forschungsgemeinschaft, Bonn and PROCOPE of DAAD/Egide to R.T. Schwarz, and Y. Guérardel. We thank Dr. P. Denny, Durham University for careful reading of this manuscript. We also acknowledge the expert help of Ms. Melanie Sauer.

Supporting Information Available :

Seven supporting figures (Fig. S1 to S7) and their corresponding legends.

One supporting table (Table S1).

References

1. Wirth, D.F. Lodish, H.F., and Robbins, P.W. (1979) Requirements for the insertion of the Sindbis envelope glycoproteins into the endoplasmic reticulum membrane, *J. Cell Biol. 81(1)*, 154-162.

2. Schwarz, R.T., and Datema, R. (1982) The lipid pathway of protein glycosylation and its inhibitors: the biological significance of protein-bound carbohydrates, *Adv. Carbohydr. Chem. Biochem. 40*, 287-379.

3. Kornfeld, R., and Kornfeld, S. (1985) Assembly of asparagine-linked oligosaccharides, *Annu. Rev. Biochem. 54*, 631-664.

4. Lehle, L., Eiden A., Lehnert, K., Haselbeck, A., and Kopetzki, E. (1995) Glycoprotein biosynthesis in *Saccharomyces cerevisiae*: ngd29, an N-glycosylation mutant allelic to och1 having a defect in the initiation of outer chain formation, *FEBS Lett 370(1-2)*, 41-45.

5. Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum, *Annu. Rev. Biochem. 73*, 1019-1049.

6. Snider, M.D., Sultzman, L.A., and Robbins P.W. (1980) Transmembrane location of oligosaccharide-lipid synthesis in microsomal vesicles, *Cell 21(2)*:385-392.

7. Snider, M.D., and Robbins, PW., (1982) Transmembrane organization of protein glycosylation. Mature oligosaccharide-lipid is located on the luminal side of microsomes from Chinese hamster ovary cells, *J. Biol. Chem.. 257(12)*, 6796-6801.

8. Kelleher, D.J., Kreibich, G., and Gilmore, R. (1992) Oligosaccharyltransferase activity is associated with a protein complex composed of ribophorins I and II and a 48 kd protein, *Cell 69(1)*, 55-65.

9. Zufferey, R., Knauer, R., Burda, P., Stagljar, I., te Heesen, S., Lehle, L., and Aebi, M. (1995) STT3, a highly conserved protein required for yeast oligosaccharyl transferase activity in vivo, *EMBO J. 14(20)*, 4949-4960.

10. Parodi, A.J. (2000) Protein glucosylation and its role in protein folding, *Annu. Rev. Biochem. 69*, 69-93.

11. Samuelson, J., Banerjee, S., Magnelli, P., Cui, J., Kelleher, D.J., Gilmore, R., and Robbins, P.W. (2005) The diversity of dolichol-linked precursors to AsN-linked glycans likely results from secondary loss of sets of glycosyltransferases, *Proc. Natl. Acad. Sci. U S A. 102(5)*, 1548-1553.

12. Dieckmann-Schuppert, A., Bender, S., Odenthal-Schnittler, M., Bause, E., and Schwarz, R.T. (1992) Apparent lack of N-glycosylation in the asexual intraerythrocytic stage of *Plasmodium falciparum*, *Eur. J. Biochem. 205(2)*, 815-825.

13. Dieckmann-Schuppert, A., Bause, E., Schwarz, R.T. (1994) Glycosylation reactions in *Plasmodium falciparum*, *Toxoplasma gondii*, and *Trypanosoma brucei* probed by the use of synthetic peptides, *Biochim. Biophys. Acta. 1199(1)*, 37-44.

14. Gowda, D.C., Gupta, P., and Davidson, E.A. (1997) Glycosylphosphatidylinositol anchors represent the major carbohydrate modification in proteins of intraerythrocytic stage *Plasmodium falciparum*, *J. Biol. Chem. 272(10)*, 6428-6439.

15. Katinka, M.D., Duprat, S., Cornillot, E., Metenier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., and Vivares, C.P. (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*, *Nature 414(6862)*, 450-453.

16. Taupin, V., Garenaux, E., Mazet, M., Maes, E., Denise, H., Prensier, G., Vivares, C.P., Guerardel, Y., and Metenier, G. (2007) Major O-glycans in the spores of two microsporidian parasites are represented by unbranched manno-oligosaccharides containing alpha-1,2 linkages, *Glycobiology 17(1)*, 56-67.

17. Striepen, B., Zinecker, C.F., Damm, J.B.L., Melgers, P.A.T., Gewig, G.J., Koolen, M., Vliegenthart, J.F.G, Dubremetz, J.F. and Schwarz RT (1997) Molecular structure of the low molecular weight antigen of *Toxoplasma gondii*: a glucose α1–4-N-acetylgalactosamine makes free glycosyl-phosphatidylinositols highly immunogenic, *J. Mol. Biol. 266*, 797–813.

18. Zinecker, C.F., Striepen, B., Geyer, H., Geyer, R., Dubremetz, J.F., and Schwarz, R.T. (2001) Two glycoforms are present in the GPI-membrane anchor of the surface antigen 1 (P30) of *T. gondii*, *Mol. Biochem. Parasitol. 116(2)*, 127-135.

19. Neuen-Jacob, E., Figge, C., Arendt, G., Wendtland, B., Jacob, B., and Wechsler, W. (1993) Neuropathological studies in the brains of AIDS patients with opportunistic diseases, *Int. J. Legal Med. 105(6)*, 339-350.

20. Klaren, V.N., and Kijlstra, A. (2002) Toxoplasmosis, an overview with emphasis on ocular involvement, *Ocul Immunol Inflamm 10(1)*, 1-26.

21. Kravetz, J.D., and Federman, D.G. (2005) Toxoplasmosis in pregnancy, *Am. J. Med. 118(3)*, 212-216.

22. Odenthal-Schnittler, M., Tomavo, S., Becker, D., Dubremetz, J.F., and Schwarz, R.T. (1993) Evidence for N-linked glycosylation in *T. gondii*, *Biochem. J. 291*, 713-721.

23. Fauquenoy, S., Morelle, W., Hovasse, A., Bednarczyk, A., Slomianny, C., Schaeffer, C., Van Dorsselaer, A., and Tomavo, S. (2008) Proteomics and glycomics analyses of N-glycosylated structures involved in *Toxoplasma gondii*-host cell interactions, *Mol. Cell. Proteomics 7(5)*, 891-910.

24. Shams-Eldin, H., Blaschke, T., Anhlan, D., Niehus, S., Muller, J., Azzouz, N., and Schwarz, R.T. (2005) High-level expression of the *T. gondii* STT3 gene is required for suppression of the yeast STT3 gene mutation, *Mol Biochem Parasitol. 143(1)*, 6-11.

25. Luk, F.C., Johnson, T.M., and Beckers, C.J. (2008) N-linked glycosylation of proteins in the protozoan parasite *Toxoplasma gondii*, *Mol. Biochem. Parasitol. 157(2)*, 169-78.

26. Ciucanu, I., and Kerek, F. (1984) A simple and rapid method for the permethylation of carbohydrates, Carbohyd. Res. 131, 209-217.

27. Hase, S. (1994) High-performance liquid chromatography of pyridylaminated saccharides, *Meth. Enzymol 230*, 225-237.

28. Villers, C., Cacan, R., Mir, A.M., Labiau, O., and Verbert, A. (1994) Release of oligomannoside-type glycans as a marker of the degradation of newly synthesized glycoproteins, *Biochem J. 298 (Pt 1)*, 135-142.

29. Stoll, J. (1986) The isolation and characterization of mammalian glycoprotein mutants, *PhD Thesis, School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, MD, USA.*

30. Duvet, S., Chirat, F., Mir, A.M., Verbert, A., Dubuisson, J., and Cacan, R. (2000) Reciprocal relationship between alpha1,2 mannosidase processing and reglucosylation in the rough endoplasmic reticulum of Man-P-Dol deficient cells, *Eur. J. Biochem. 267*, 1146-1152.

31. Banerjee, S., Vishwanath, P., Cui, J., Kelleher, D.J., Gilmore, R., Robbins, P.W., and Samuelson, J. (2007) The evolution of N-glycan-dependent endoplasmic reticulum quality control factors for glycoprotein folding and degradation, *Proc. Natl. Acad. Sci. 104(28)*,11676-11681.

32. Coppens, I., Sinai, A.P., Joiner, K.A. (2000) *Toxoplasma gondii* exploits host low-density lipoprotein receptor-mediated endocytosis for cholesterol acquisition, *J. Cell Biol. 149(1)*, 167-180.

33. Charron, A.J. and Sibley, L.D. (2002) Host cells: mobilizable lipid resources for the intracellular parasite *Toxoplasma gondii*, *J. Cell Sci. 115(15)*, 3049-3059.

34. Coppens, I., and Joiner, K.A. (2001) Parasite host cell interactions in toxoplasmosis: new avenues for intervention?, *Expert Rev. Mol. Med. 2001*, 1-20.

35. Tomavo, S., Dubremetz, J.F., and Schwarz, R.T. (1992) Biosynthesis of glycolipid precursors for glycosylphosphatidylinositol membrane anchors in a *Toxoplasma gondii* cell-free system, *J. Biol. Chem. 267(30)*, 21446-21458.

36. Oriol, R., Martinez-Duncker, I., Chantret, I., Mollicone, R., and Codogno, P. (2002) Common origin and evolution of glycosyltransferases using Dol-P-monosaccharides as donor substrate, *Mol. Biol. Evol. 19(9)*,1451-1463.

37. Delorenzi, M., Sexton, A., Shams-Eldin, H., Schwarz, R.T., Speed, T., and Schofield, L. (2002) Genes for glycosylphosphatidylinositol toxin biosynthesis in *Plasmodium falciparum*, *Infect. Immun. 70(8)*, 4510-4522.

Figure 1 Inventory of ALG glycosyltransferases genes and predicted dolichol-linked *N*-glycans in (A) classical mammalian cell, (B) *T. gondii* and (C) B3F7 mutant cell line

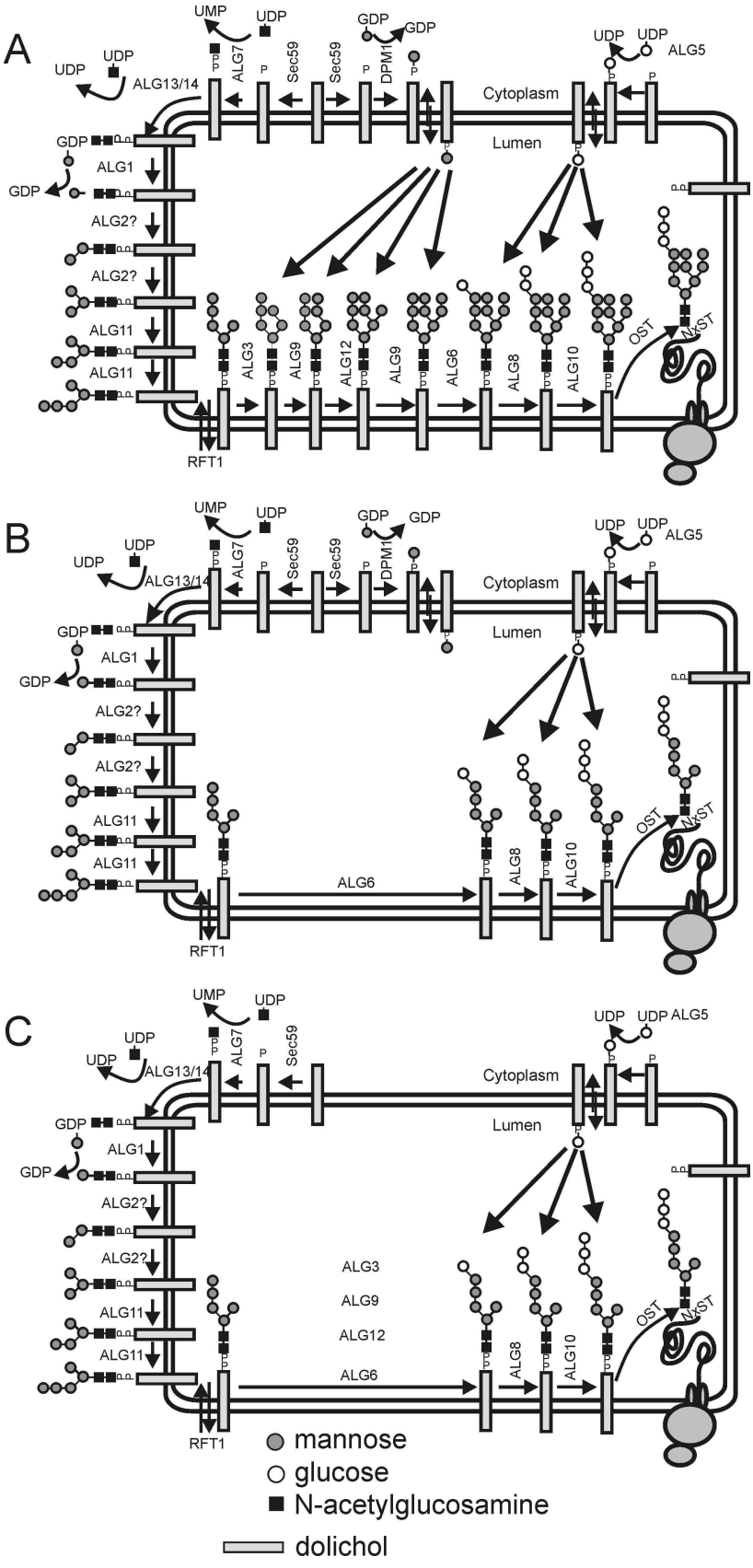
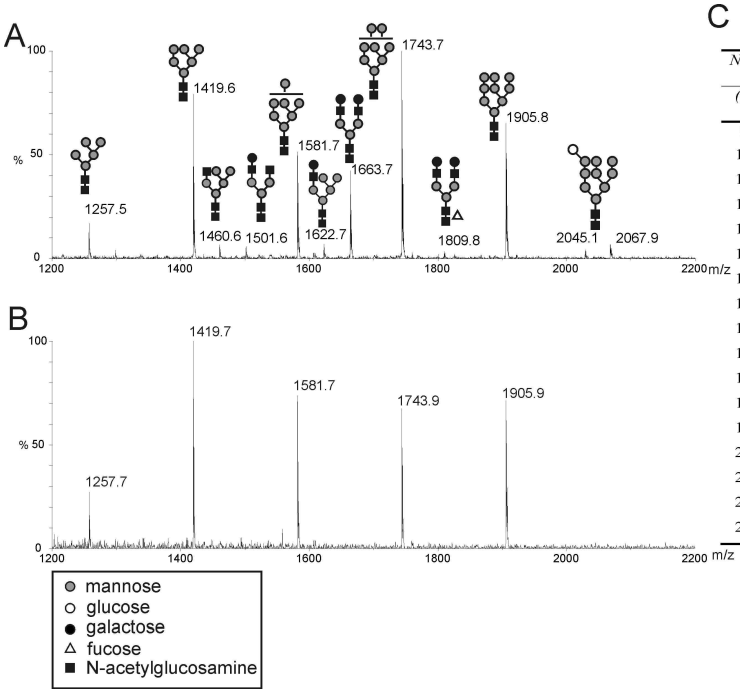


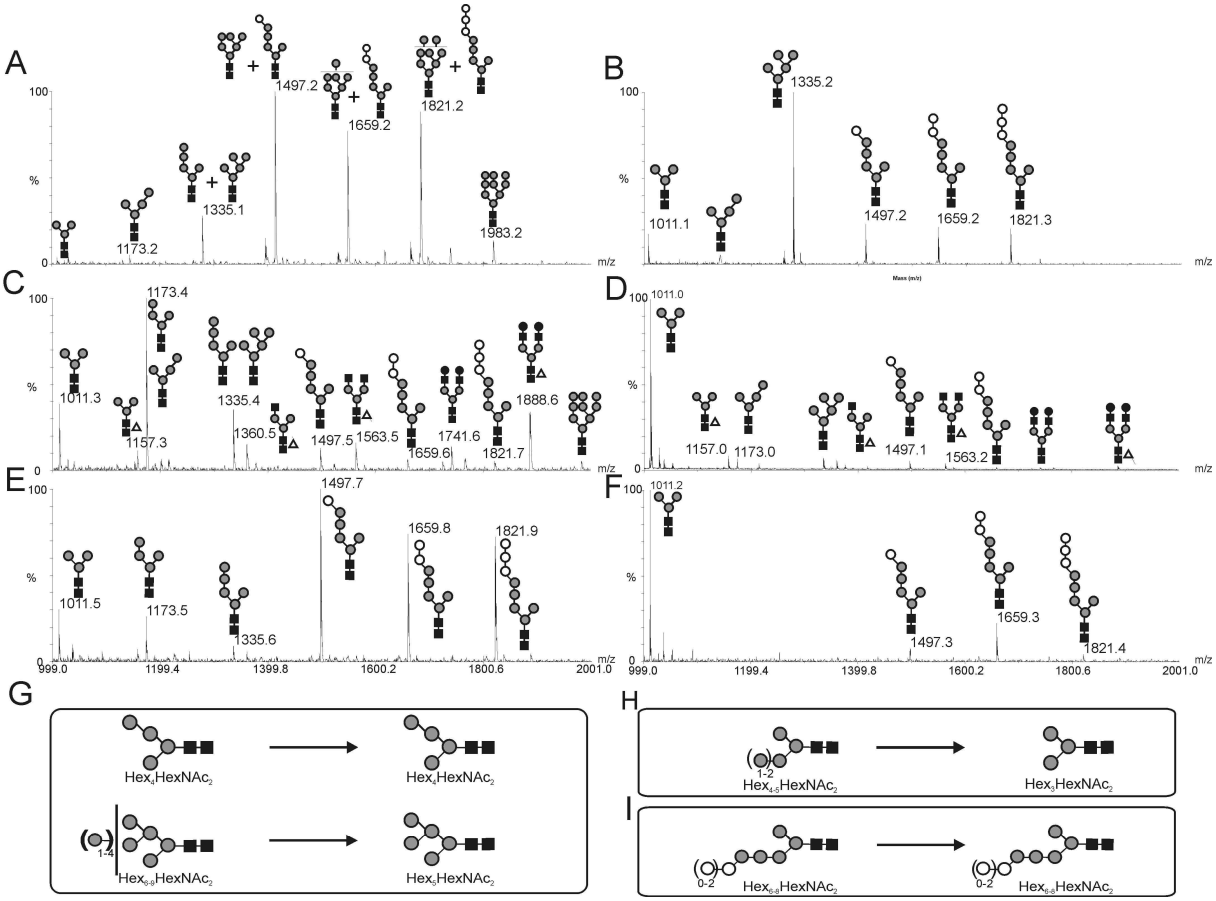
Figure 2 MALDI-TOF/MS comparison of native PNGase F released oligosaccharides from (A) Vero cells and (B) *T. gondii* grown in Vero cells. (C) Summary of native and 2-AP derivatives compounds observed by MS in vero cells and *T. gondii* grown in Vero cells.



C

Native	*2-PA*	*Assignment*	*Sample*	
(m/z)	*(m/z)*	*Composition*	*Vero*	*T. gondii*
933	1011	$Man_3GlcNAc_2$	√	√
1079	1157	(Fuc) $Man_3GlcNAc_2$	√	
1095	1173	Hex_1 $Man_3GlcNAc_2$	√	√
1257	1335	Hex_2 $Man_3GlcNAc_2$	√	√
1419	1497	Hex_3 $Man_3GlcNAc_2$	√	√
1460	1538	GlcNAc $Man_5GlcNAc_2$	√	
1501	1579	$GlcNAc_2$ $Man_3GlcNAc_2$	√	
1581	1659	Hex_4 $Man_3GlcNAc_2$	√	√
1622	1700	Gal GlcNAc $Man_3GlcNAc_2$	√	
1663	1741	$Gal_2GlcNAc_2Man_3GlcNAc_2$	√	
1743	1821	Hex_5 $Man_3GlcNAc_2$	√	√
1809	1887	Fuc $Gal_2GlcNAc_2Man_3GlcNAc_2$	√	
1905	1983	Hex_6 $Man_3GlcNAc_2$	√	√
2012	2090	Fuc $Gal_2GlcNAc_3Man_3GlcNAc_2$	√	
2028	2106	$Gal_3GlcNAc_3Man_3GlcNAc_2$	√	
2067	2145	Hex_7 $Man_3GlcNAc_2$	√	√
2174	2252	Fuc $Gal_3GlcNAc_3Man_3GlcNAc_2$	√	

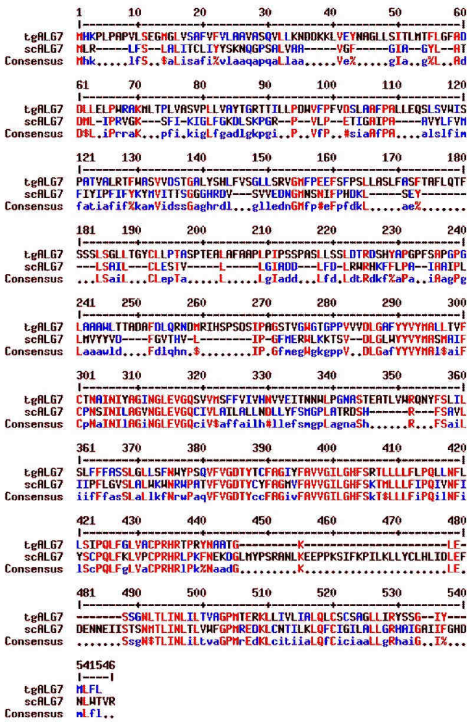
Figure 3 MS profiles of oligosaccharides released from proteins of (A,B) *T. gondii* grown in vero cells, (C,D) B3F7 cells, (E,F) *T. gondii* grown in B3F7 cells, (A,C,E) before and (B,D,F) after *Aspergillus saitoï* exo-mannosidase digestion. Susceptibility to *A. saitoï* exo-mannosidase digestion of (G) branched Golgi derived oligomannosyl-types, (H) ER derived *N*-glycans and (I) ER derived linear truncated *N*-glycans.



Supporting Figures

Figure S1 Functional complementation of the conditional lethal *S. cerevisiae* strain YPH499-HIS-GAL-ALG7 by the *T. gondii* ALG7. The conditional lethal mutant was transformed with plasmids carrying either the Human ALG7 (HsALG7) or the *T. gondii* ALG7 (*Tg*ALG7). The transformed cells were then streaked onto plates containing minimal medium lacking histidine and containing either galactose (SGR) (2) or glucose (SD) (3) and incubated at 30°C.

A



B

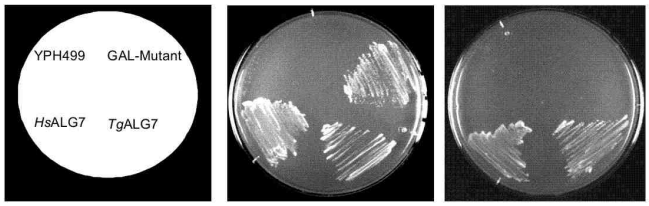
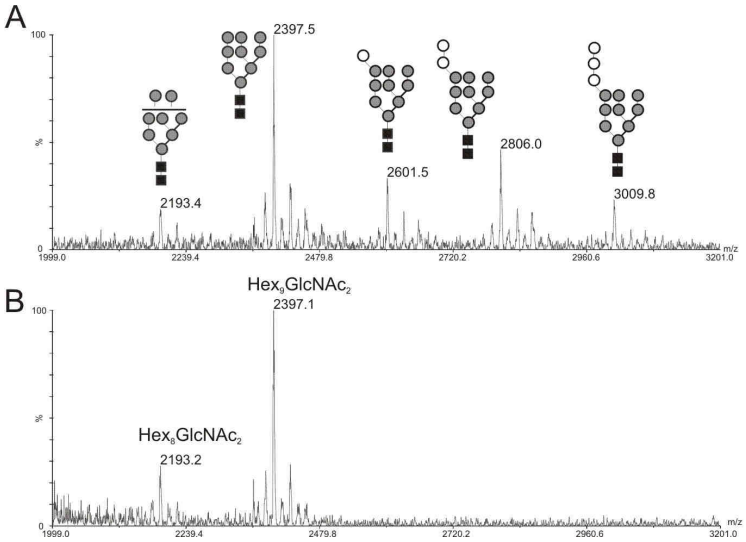
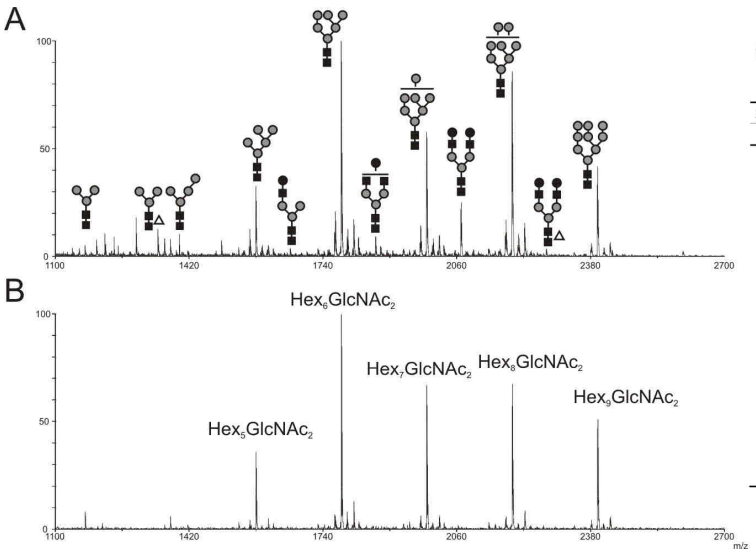


Figure S2 MALDI-TOF/MS comparison of *DolPPOS* acid hydrolysed released oligosaccharides from (A) Vero cells and (B) *T. gondii* grown in Vero cells. (C) Summary of permethylated compounds observed by MS in Vero cells and *T. gondii* grown in Vero cells.



Signal	*Composition*	*Sample*	
m/z		*Vero*	*T. gondii*
1783	$Hex_6GlcNac_2$	√	
1987	$Hex_7GlcNac_2$	√	
2191	$Hex_8GlcNac_2$	√	√
2395	$Hex_9GlcNac_2$	√	√
2599	$Hex_{10}GlcNac_2$	√	
2802	$Hex_{11}GlcNac_2$	√	
3006	$Hex_{12}GlcNac_2$	√	

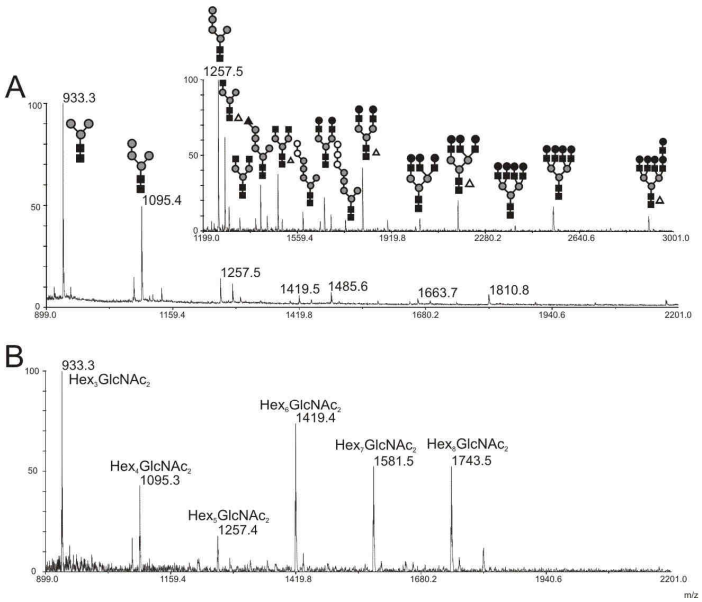
Figure S3 MALDI-TOF/MS comparison of native PNGase F released oligosaccharides from (A) HFF cells and (B) *T. gondii* grown in HFF cells. (C) Summary of native and 2-AP derivatives compounds observed by MS in HFF cells and *T. gondii* grown in HFF cells.



C

pMeth	*Assignement*	*Sample*	
m/z	*Composition*	*HFF*	*T. gondii*
1171	$Man_3GlcNAc_2$, Na^+	√	√
1345	Fuc $Man_3GlcNAc_2$, Na^+	√	
1375	Hex_1 $Man_3GlcNAc_2$, Na^+	√	√
1579	Hex_2 $Man_3GlcNAc_2$, Na^+	√	√
1620	Gal GlcNAc $Man_3GlcNAc_2$, Na^+	√	
1783	Hex_3 $Man_3GlcNAc_2$, Na^+	√	√
1865	$GalGlcNAc_2Man_3GlcNAc_2$, Na^+	√	
1987	Hex_4 $Man_3GlcNAc_2$, Na^+	√	√
2069	$Gal_2GlcNAc_2Man_3GlcNAc_2$, Na^+	√	
2191	Hex_5 $Man_3GlcNAc_2$, Na^+	√	√
2243	Fuc $Gal_2GlcNAc_2Man_3GlcNAc_2$, Na^+	√	
2395	Hex_6 $Man_3GlcNAc_2$, Na^+	√	√
2488	Fuc $Gal_2GlcNAc_3Man_3GlcNAc_2$, Na^+	√	
2518	$Gal_3GlcNAc_3Man_3GlcNAc_2$, Na^+	√	
2599	Hex_7 $Man_3GlcNAc_2$, Na^+	√	√
2692	Fuc $Gal_3GlcNAc_3Man_3GlcNAc_2$, Na^+	√	

Figure S4 MALDI-TOF/MS comparison of native PNGase F released oligosaccharides from (A) B3F7 cells and (B) *T. gondii* grown in B3F7 cells. (C) Summary of native and 2-AP derivatives compounds observed by MS in B3F7 cells and *T. gondii* grown in B3F7 cells.



Native (m/z)	2-PA (m/z)	Composition Assignment	*Sample* B3F7	*T. gondii*
933	1011	$Man_3GlcNAc_2$, Na^+	√	√
1079	1157	Fuc $Man_3GlcNAc_2$, Na^+	√	
1095	1173	Hex_1 $Man_3GlcNAc_2$, Na^+	√	√
1257	1335	Hex_2 $Man_3GlcNAc_2$, Na^+	√	√
1282	1360	Fuc GlcNAc $Man_3GlcNAc_2$, Na^+	√	
1339	1417	$GlcNAc_2$ $Man_3GlcNAc_2$, Na^+	√	
1419	1497	Hex_3 $Man_3GlcNAc_2$, Na^+	√	√
1486	1564	Fuc $GlcNAc_2$ $Man_3GlcNAc_2$, Na^+	√	
1501	1579	Gal $GlcNAc_2$ Man_3 $GlcNAc_2$, Na^+	√	
1581	1659	Hex_4 $Man_3GlcNAc_2$, Na^+	√	√
1648	1726	Fuc Hex $GlcNAc_2$ $Man_3GlcNAc_2$, Na^+	√	
1664	1742	$Gal_2GlcNAc_2Man_3GlcNAc_2$, Na^+	√	
1689	1767	Fuc $GlcNAc_3$ $Man_3GlcNAc_2$, Na^+	√	
1743	1821	Hex_5 $Man_3GlcNAc_2$, Na^+	√	√
1809	1887	Fuc $Gal_2GlcNAc_2Man_3GlcNAc_2$, Na^+	√	
1905	1983	Hex_6 $Man_3GlcNAc_2$, Na^+	√	
2012	2090	Fuc Gal_2 $GlcNAc_3$ $Man_3GlcNAc_2$, Na^+	√	
2028	2106	$Gal_3GlcNAc_3Man_3GlcNAc_2$, Na^+	√	
2174	2252	Fuc $LacNAc_3$ Man_3 $GlcNAc_2$, Na^+	√	
2394	2472	$LacNAc_4$ Man_3 $GlcNAc_2$, Na^+	√	
2541	2619	Fuc $LacNAc_4$ Man_3 $GlcNAc_2$, Na^+	√	
2906	2984	Fuc $LacNAc_5$ Man_3 $GlcNAc_2$, Na^+	√	

Figure S5 Susceptibility of standard branched olimannosyl-types *N*-Glycans to *A. saitoi* exo-mannosidase digestion. MS profiles of 2-AP derivatized olimannosyl-types *N*-Glycans (A) before and (B) after mannosidase digestion.

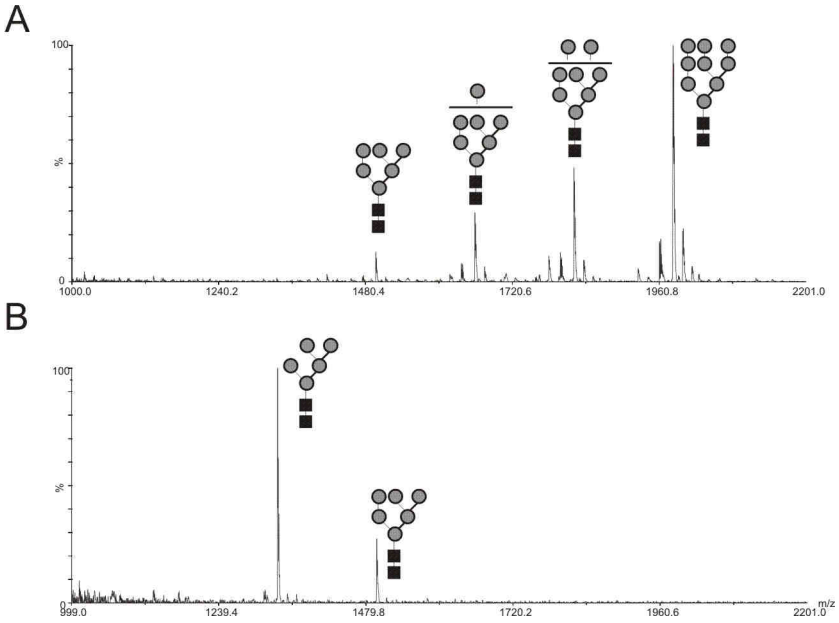


Figure S6 Susceptibility of Vero cells derived *N*-Glycans to *A. saitoi* exo-mannosidase digestion. MS profiles of 2-AP derivatized olimannosyl-types *N*-Glycans (A) before and (B) after mannosidase digestion. * correspond to complex types *N*-glycans.

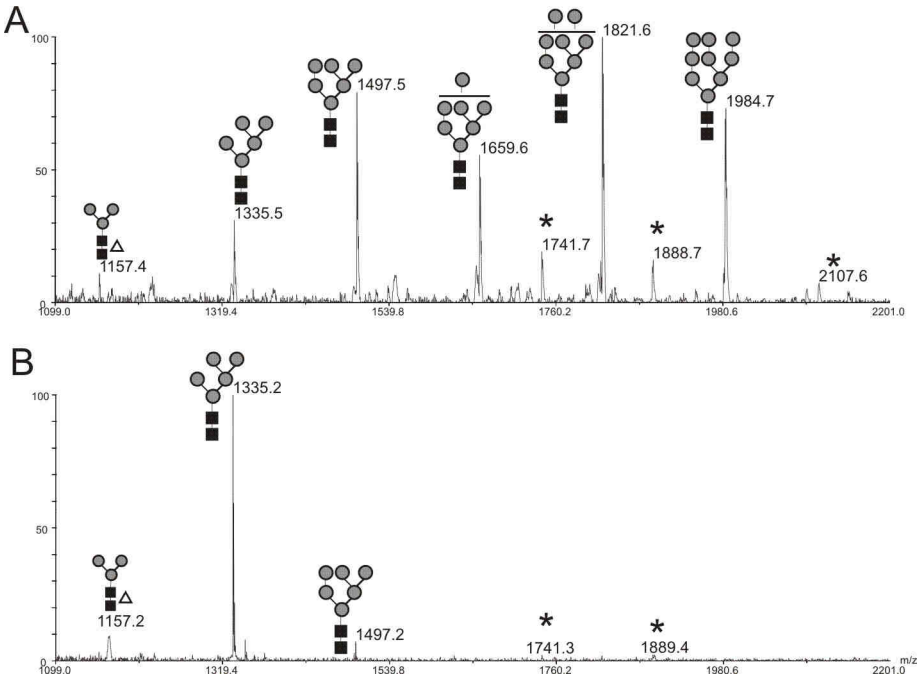


Figure S7 MALDI-TOF/MS of native PNGase F released oligosaccharides from *T. gondii* grown in B3F7 mixed with VERO cells debris after purification on a glass wool column. * correspond to glucose polymers.

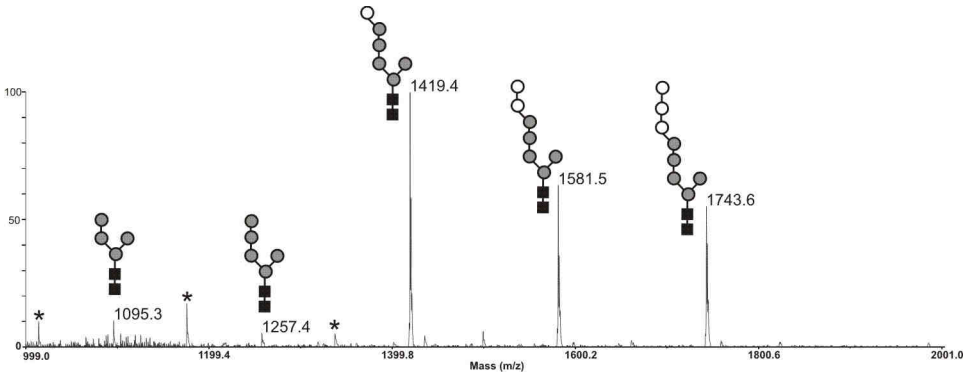


Table S1 Genes predicted to be involved in protein *N*-glycosylation in *T. gondii*.

Gene	Function	Statistical summary on the proteins				References
		Item	Yeast	Human	*T. gondii*	
ALG7	UDP-N-acetyl-glucosamine-1-P transferase, transfers Glc-Nac-P from UDP-GlcNac to Dol-P in the ER	Molecular weight	50.37	46.09	53.95	Eckert *et al.* (1998) Hartog and Bishop,1987
		Residue numbers	448	408	500	
		Charge	6	3	-6	
		Isoelectric point	8.40	7.92	5.90	
DPM1	catalyzes the formation of Dol-P-Man from Dol-P and GDP-Man/ER membrane	Molecular weight	30.36	29.63	26.93	Orlean *et al.* (1988); Mazhari-Tabrizi *et al.* (1996) ; Maeda *et al.* (2000) ; Colussi PA, *et al.* (1997)
		Residue numbers	267	260	237	
		Charge	1	10	4	
		Isoelectric point	7.97	10.40	9.76	
ALG1	β-1 ,4-mannosyl-transferase; essential for viability	Molecular weight	51.93	52.52	63.40	Huffaker & Robbins, 1983; Couto *et al.*,1984; Albright & Robbins, 1990
		Residue numbers	449	464	574	
		Charge	9	-2	10	
		Isoelectric point	9.31	7.19	8.18	
ALG2	α-1,3-mannosyl- and α-1,6-mannosyl-Transferases that catalyze two consecutive steps	Molecular weight	58.05	47.09	56.37	Jackson *et al.*, 1993
		Residue numbers	503	416	506	
		Charge	-7	-2	1	
		Isoelectric point	6.22	7.02	7.65	
ALGII	α-1,2-mannosyl transferase, catalyzes sequential addition of the two	Molecular weight	63.15	55678.66	58.92	Nasr *et al.*, 1996
		Residue numbers	548	492	540	
		Charge	13	9	13.00	

		Isoelectric point	9.70	8.48	8.72	
ALG3	Dolichol-P-Man dependent α-(1-3) mannosyl-transferase	Molecular weight	52.86	50.12	**missing**	Huffaker & *Robbins, 1983; Sharma et al.,* 1990; Verostek *et al,* 1993; Aebi *et al.,*1996; Kimura *et al.,* 1997
		Residue numbers	458	438		
		Charge	11	19		
		Isoelectric point	9.29	10.21		
ALG9	α-1,2-mannosyl-transferase that catalyzes the transfer of mannose from Dol-P-Man to lipid-linked oligosaccharides	Molecular weight	63.78	70.78	**missing**	Burda *et al.,* 1996
		Residue numbers	555	618		
		Charge	9	12		
		Isoelectric point	8.74	8.84		
ALG12	α-1,6-mannosyl-transferase that catalyzes the addition of the alpha-1,6 mannose to dolichol-linked Man7GlcNAc2,	Molecular weight	62.67	54.65	**missing**	Grubenmann *et al.,* 2002
		Residue numbers	551	488		
		Charge	-2	17		
		Isoelectric point	7.09	10.43		
ALG6	α-1,3-glucosyl-transferase, mutations in human ortholog are associated with disease	Molecular weight	62.78	58.34	61.14	Reiss *et al.,* 1996
		Residue numbers	544	509	564	
		Charge	15	12	10	
		Isoelectric point	9.87	8.39	8.61	
ALG5	UDP-glucose: dolichol-phosphate glucosyltransferas	Molecular weight	38.35	36.94	**missing**	Huffaker & Robbins, 1983; te Heesen *et al.,* 1994; Lennon *et al.,* 1995
		Residue numbers	334	324		
		Charge	6	10		
		Isoelectric point	9.17	10.05		
ALG8	α-1,3-glucosyl-	Molecular weight	67.38	60.08	81.47	Stagljar *et al.,* 1994

		Residue numbers	577	526	737	
		Charge	24	18	15	
		Isoelectric point	10.31	9.81	8.85	
*ALG*10	Dolichyl-phosphoglucose -dependent α-1,2 glucosyl-transferase	Molecular weight	61.74	55.60	66.66	Burda and Aebi , 1998
		Residue numbers	525	473	603	
		Charge	18	20	14	
		Isoelectric point	10.19	10.04	8.82	
Stt3	subunit of the oligosaccharyl -transferase that forms a subcomplex with Ost3p and Ost4p and is directly involved in catalysis.	Molecular weight	81.52	80.52	83	Yoshida *et al.,* 1995; Zufferey *et al.*,1995; Hong *et al.*, 1996 Shams-Eldin *et al.*, 2005
		Residue numbers	718	705	753	
		Charge	2	4	17	
		Isoelectric point	8.10	8.07	10.00	

2.2- Biosynthèse des β-Mannanes chez *C. albicans*

2.2.1- Contexte

Ces travaux ont été initiés par le Dr. G. Strecker en 1990 en collaboration étroite avec le Dr. Daniel Poulain (INSERM Unité 799). Ils concernent l'étude structurale des glycoconjugués de la paroi de la levure pathogène *Candida albicans*.

Candida albicans est une levure ascomycète saprophyte de la flore digestive normale de l'humain. Cependant, du fait de l'augmentation des traitements immuno-suppresseurs, des procédures chirurgicales invasives et du vieillissement de la population des pays occidentaux, cette levure est responsable d'infections sévères, souvent disséminées, entraînant une mortalité élevée. Le contact entre *C. albicans* et les cellules hôtes se fait au niveau de la paroi du champignon. Comme pour toutes les levures, les parois de ces cellules sont des structures complexes formées d'un réseau d'exo-polysaccharides mêlés de nombreuses protéines incluses ou ancrées dans ces structures. Plus spécifiquement, la paroi de *C. albicans* est principalement composée de glucanes (polymères de résidus glucoses liés en β-1,3 et β-1,6), de mannanes (polymères de résidus de mannoses liés α-1,2 et α-1,3) et de mannoprotéines. En plus de ces polymères d'α-Man, des β-1,2 oligomannosides sont associés soit aux mannoprotéines pariétales soit au phospholipomannane (PLM) (Shibata *et al.*, 1985 ; Kobayashi *et al.*, 1992 ; Trinel *et al.*, 2002). Absents chez *Saccharomyces cerevisiae*, les β-1,2 oligomannosides ont aussi été identifiés chez d'autres microorganismes, dont certains potentiellement pathogènes, tels que des salmonelles (Lindber *et al.*, 1988), le parasite *Leishmania* (Ralton *et al.*, 2003), ou la levure *Pichia pastoris* (Vinogradov *et al.*, 2000). De nombreuses études ont montré que la présence de ces résidus était associée aux mécanismes de reconnaissance hôte-pathogène, ainsi qu'à la pathogénicité de *C. albicans.*

Dès les premiers jours de l'existence, les levures font partie intégrante de la flore humaine puis deviennent ensuite une composante régulière de l'alimentation. Les relations entre le système immunitaire et ces micro-organismes s'équilibrent par le biais de l'immunité innée puis d'un relais par l'immunité adaptative dont une des composantes est une réponse humorale dirigée contre des polymères de mannoses de la paroi des levures qui peut être mise en évidence chez tous les individus. Il existe au moins deux circonstances dans lesquelles les perturbations des relations hôte-levures aboutissent ou sont liées à des manifestations pathologiques. Il s'agit des candidoses et de la maladie de Crohn. Dans le premier cas, des perturbations de l'homéostasie de l'hôte vont aboutir à

l'invasion tissulaire par la levure *C. albicans* qui est un endosaprophyte pathogène opportuniste. Dans le second cas, des anticorps anti *S. cerevisiae* -levure essentiellement alimentaire- vont, pour des raisons encore obscures, être des marqueurs prédictifs d'une pathologie inflammatoire sévère du tube digestif. L'équipe du Dr Daniel Poulain a contribué à explorer les mécanismes physiopathologiques au travers desquels les structures oligo-mannosylées synthétisées par *C. albicans* ou *S. cerevisiae* peuvent contribuer à la pathogénie. Ces études ont permis d'établir que la régulation d'une partie importante des relations hôte-levure était conditionnée par des «codes oligomannosidiques » synthétisés différemment par ces dernières en fonction des conditions imposées ou permises par l'hôte. Dans ce cadre, la collaboration avec l'équipe du Dr Gérard Strecker a été essentielle à l'établissement de la spécificité structurale des domaines oligomannosylés au sein des différents glycoconjugués pariétaux de la levure. En particulier, ces analyses structurales ont permis de mettre en évidence la présence de β-mannanes spécifiques de *C. albicans* au sein de plusieurs de ces composés et d'identifier une nouvelle classe de glycolipides à longues chaînes de β1-2 mannosides (Trinel *et al.*, 1997 ; 1999 ; 2002). De par leur conformation hélicoïdale particulière (Nitz *et al.*, 2002), les β mannosides présentent des activités biologiques distinctes des α mannosides trouvés chez *S. cerevisiae.* Alors que ces derniers interagissent avec des lectines de type-C, les β mannosides interagissent spécifiquement avec la galectine 3. De plus, les β mannosides sont des antigènes pour la réponse immunitaire adaptative et élicitent la synthèse d'anticorps protecteurs contre les infections par *C. albicans*. Enfin, l'administration orale de β mannosides synthétiques protège de la colonisation du tractus intestinal de la souris nouveau-né par *C. albicans*, alors que celle d'α mannosides ne présente aucun effet.

2.2.2- Résultats

Identification des β-mannosyltransférases de *C. albicans*

Nos travaux en collaboration avec l'équipe du docteur D. Poulain concernent la biosynthèse des β mannosides dans *C. albicans*. Ainsi, nous avons suivi par des moyens physicochimiques les modifications de β mannosylation dans les parois des *C. albicans* déficients dans la biosynthèse de diverses activités β-mannosyltransférasiques. Une nouvelle famille de β-mannosyltransférases, les BMT, a en effet été identifiée dans *Pichia pastoris* et leurs homologues ont été individuellement inactivés chez *C. albicans* par l'équipe du Dr Poulain. Nous avons ensuite procédé à leur phénotypage en analysant par RMN la présence des β mannosides sur les glycolipides et les

domaines acide-stable et acide-labile des N-glycannes de leurs phospho-peptido-mannanes. Ainsi, nous avons participé à identifier les activités β-mannosyltransférasiques déficientes en fonction de la longueur des chaînes oligomannosidiques observées chez chaque mutant. Par exemple, cette stratégie nous a permis de démontrer que trois enzymes distinctes CaBmt2, CaBmt3 et CaBmt4 étaient responsables de l'ajout séquentiel des trois premiers résidus de mannose en béta sur le phospho-mannane des N-glycannes (Mille *et al.*, 2008). Les enzymes responsables de la biosynthèse du domaine acido-stable ont également été identifiées de la même manière. Parallèlement, l'analyse structurale des phospholipomannanes isolés des souches mutées dans la synthèse des β-mannosyltransférases nous a permis de déterminer la spécificité de deux autres membres de la famille des BMTs, les BMT5 et BMT6, pour les glycolipides (Mille *et al., soumis*). Il est à noter que tous les membres de cette nouvelle famille d'enzymes font preuve d'une remarquable spécificité de substrats malgré de très fortes homologies de séquences. De fait, les BMTs apparaissent comme un modèle extrêmement prometteur pour l'étude des relations entre structure et activité des glycosyltransférases. Ce point sera abordé plus avant dans les projets.

Application de la RMN HR-MAS à l'étude des profils de glycosylation des levures

Suite à la définition de la spécificité fine de chaque β-mannosyltransférase par l'analyse détaillée de la structure des mannanes purifiés de chaque souche mutante, les profils de glycosylation totaux de ces souches ont été déterminés par Résonance Magnétique Nucléaire HR-MAS (High Resolution-Magic Angle Spinning). Cette technique nouvellement introduite, et mise en place par l'équipe de Guy Lippens à Lille, permet d'observer les signaux RMN des molécules majeures présentes à la surface de cellules vivantes en effectuant des acquisitions sur les cellules totales. Elle a récemment été utilisée pour l'analyse des lipoglycannes de la paroi des mycobactéries (Li *et al.*, 2005). Nous avons adapté cette technologie à l'étude des mannanes présents à la surface des levures, ce qui nous a permis de déterminer le phénotype glycannique de souches mutantes sans *a priori* structuraux. Pour ce faire, les culots cellulaires de différents sérotypes contrôles de *C. albicans* et des souches mutantes déficientes dans la synthèse des β-mannosyltransférases impliquées dans la synthèse des mannanes ont été insérés dans des rotors en oxyde de zirconium et analysés en RMN HR-MAS ^{1}H homonucléaire mono et bidimensionnelles et ^{1}H-X hétéronucléaires bidimensionnelles. Les spectres ont ensuite été comparés à ceux de mannanes purifiés pour déterminer les modifications structurales des mannanes exprimés à la surface cellulaire dans les différentes souches mutantes. La comparaison de la région anomérique des spectres HR-MAS RMN

^{1}H-^{13}C hétéronucléaires du mannane purifié et de la souche standard démontre la présence de β-mannoses à la surface des cellules vivantes. L'anomérie des différents signaux est attribuée grâce à la mesure des constantes de couplages directes ($^{1}J_{H,C}$ 170 Hz pour les α -Man et 160 Hz pour les β-Man). En accord avec nos observations sur les molécules purifiées, nous avons démontré que l'inactivation des gènes de chaque β-mannosyltransférase induisait des changements spécifiques dans l'expression des épitopes β-mannosylés à la surface des cellules. De plus, la comparaison des profils glycanniques de surface a mis en évidence la présence de structures β-mannosylées additionnelles différentes de celles observées sur les phospho-peptido mannanes et phospholipomannanes purifiés. En particulier, nous observons un ensemble de paramètres RMN attribués à un résidu de phospho β-Man encore jamais décrit sur les manno-glycoconjugués fongiques. De fait, la RMN HR-MAS s'est révélée être une technique d'analyse extrêmement puissante permettant non seulement de suivre l'évolution de la structure de motifs glycanniques pariétaux connus, mais également de mettre en évidence des motifs inconnus que les techniques classiques d'analyse après purification ne permettent pas d'observer.

Notre implication dans l'étude des activités associées aux BMTs et l'application de la RMN HR-MAS est décrite dans les deux publications suivantes : Mille *et al.*, 2008; Maes *et al.*, soumis.

Identification of a New Family of Genes Involved in β-1,2-Mannosylation of Glycans in *Pichia pastoris* and *Candida albicans*[*]

Received for publication, October 25, 2007, and in revised form, January 30, 2008 Published, JBC Papers in Press, January 30, 2008, DOI 10.1074/jbc.M708825200

Céline Mille[‡1], Piotr Bobrowicz[§1], Pierre-André Trinel[‡], Huijuan Li[§], Emmanuel Maes[¶], Yann Guerardel[¶], Chantal Fradin[‡], María Martínez-Esparza[‖], Robert C. Davidson[§], Guilhem Janbon[**], Daniel Poulain[‡2], and Stefan Wildt[§]

*From the [‡]Unité de Physiopathologie des Candidoses, INSERM U799, Université de Lille 2 EA 2684, Lille 59045, France, [§]Strain Development, GlycoFi, Inc./Merck & Co., Inc., Lebanon, New Hampshire 03766, [¶]Unité de Glycobiologie Structurale et Fonctionnelle, UMR CNRS 8576, Villeneuve d'Ascq 59655, France, the [‖]Department of Biochemistry, Molecular Biology (B) and Immunology, Medical School, University of Murcia, Murcia 30100, Spain, and **Unité de Mycologie Moléculaire, Institut Pasteur, Paris 75015, France*

Structural studies of cell wall components of the pathogenic yeast *Candida albicans* have demonstrated the presence of β-1,2-linked oligomannosides in phosphopeptidomannan and phospholipomannan. During *C. albicans* infection, β-1,2-oligomannosides play an important role in host/pathogen interactions by acting as adhesins and by interfering with the host immune response. Despite the importance of β-1,2-oligomannosides, the genes responsible for their synthesis have not been identified. The main reason is that the reference species *Saccharomyces cerevisiae* does not synthesize β-linked mannoses. On the other hand, the presence of β-1,2-oligomannosides has been reported in the cell wall of the more genetically tractable *C. albicans* relative, *P. pastoris*. Here we present the identification, cloning, and characterization of a novel family of fungal genes involved in β-mannose transfer. Employing *in silico* analysis, we identified a family of four related new genes in *P. pastoris* and subsequently nine homologs in *C. albicans*. Biochemical, immunological, and structural analyses following deletion of four genes in *P. pastoris* and deletion of four genes acting specifically on *C. albicans* mannan demonstrated the involvement of these new genes in β-1,2-oligomannoside synthesis. Phenotypic characterization of the strains deleted in β-mannosyltransferase genes (*BMTs*) allowed us to describe the stepwise activity of Bmtps and acceptor specificity. For *C. albicans*, despite structural similarities between mannan and phospholipomannan, phospholipomannan β-mannosylation was not affected by any of the *CaBMT1–4* deletions. Surprisingly, depletion in mannan major β-1,2-oligomannoside epitopes had little impact on cell wall surface β-1,2-oligomannoside antigenic expression.

A number of yeast species have adapted to colonize human tissue, and the adverse effects caused by these pathogenic yeasts have gained importance over recent decades. In particular, the endosaprophytic yeast *Candida albicans* is able to invade human tissues in immunosuppressed patients leading to frequent nosocomial systemic infections with a high mortality rate (1). Current research has shown that successful pathogenic adaptation requires the ability to utilize elaborate sensing and regulation pathways, impacting expression of a wide range of virulence factors (2). Among these virulence attributes critical for survival under changing environmental conditions is the yeast cell wall containing large amounts of carbohydrates and carbohydrates covalently linked to a noncarbohydrate moiety classified as glycoconjugates, either glycoproteins or glycolipids. In yeasts and other eukaryotes, glycoproteins typically contain *N*- and/or *O*-linked glycans (3, 4). Glycans serve diverse and important functions, such as intracellular trafficking of glycoproteins or glycolipids, protein folding, anchoring of macromolecules to the cell membrane, signal transduction, cell/cell interactions, as well as determination of glycoprotein half-life (5). In pathogenic microbes, glycans trigger host innate and/or adaptive immunity responses.

Despite similarities in the early steps of processing, the mature structure of glycans differs substantially between yeasts and mammals. For example, mammalian *N*-linked glycans are typically of the complex type, whereas typical fungal *N*-glycans are categorized as high mannose (3). Depending on the species, fungal high mannose glycans contain distinctive modifications, such as the addition of mannosyl phosphate (6) and β-linked mannose. Suzuki and co-workers (7) was the first to demonstrate and confirm that *C. albicans*, in contrast to *Saccharomyces cerevisiae*, harbors β-1,2-linked oligomannosides (β-Mans)[3]. The expression pattern of β-Mans in *C. albicans* is quite complex. β-Mans have been found to be associated with the acid-labile and acid-stable part of the cell wall phosphopeptidomannan (PPM) (8)

[*] This work was supported in part by the European Programme Interreg IIIA. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: Inserm U799, Lille 59045, France. Tel.: 33-3-20-62-34-20; Fax: 33-3-20-62-34-16; E-mail: dpoulain@univ-lille2.fr.

[3] The abbreviations used are: β-Man, β-1,2-linked oligomannoside; PPM, phosphopeptidomannan; PLM, phospholipomannan; mAb, monoclonal antibody; ORF, open reading frame; ANTS, 8-aminonaphthalene-1,3,6-trisulfonate; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; MS, mass spectrometry; HSQC, heteronuclear single quantum coherence; ROESY, rotating frame nuclear Overhauser enhancement spectroscopy; FACE, fluorophore-assisted carbohydrate electrophoresis.

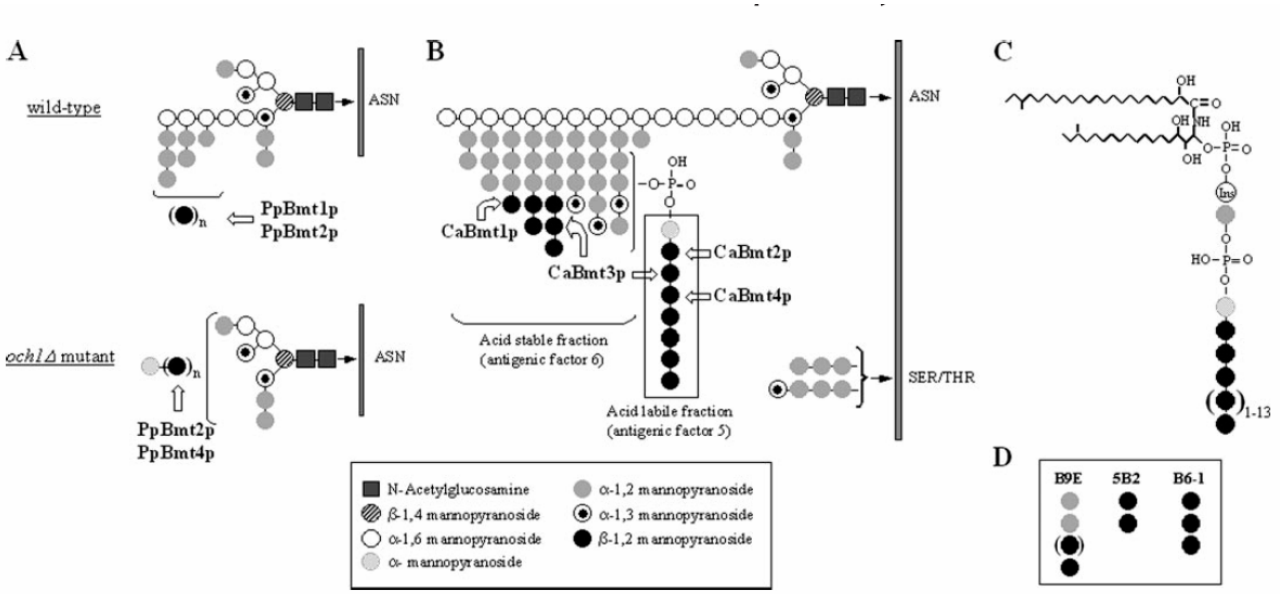


FIGURE 1. **β-1,2-Mannosylation pathways.** Schematic structure of *P. pastoris N*-linked glycans and *C. albicans* serotype A glycoconjugates in which the presence of β-1,2-linked oligomannosides has been demonstrated chemically. *A, N*-linked glycans from *P. pastoris* wild-type strain and *och1*Δ mutant. *B, C. albicans* cell wall phosphopeptidomannan. *C, C. albicans* cell wall phospholipomannan, a glycosphingolipid product of the sphingolipid biosynthesis pathway. *D,* schematic representation of the main epitopes recognized by mAbs B9E, 5B2, and B6-1. The *arrow* indicates the function of the PpBmt and CaBmt proteins identified in this work by showing where the enzymes act and which hexose is added.

where they correspond, respectively, to antigenic factors 5 and 6 (Fig. 1*B*). They are also associated with phospholipomannan (PLM), a surface cell wall glycolipid derived from the mannose-inositol-phosphoceramide biosynthetic pathway (9) (Fig. 1*C*). Although not yet confirmed by chemical/structural studies, association of β-Mans with several cell wall non-mannan mannoproteins has been suggested repeatedly by the use of anti-β-Man-specific antibodies (10, 11). The mammalian immune system specifically reacts with synthetic β-Mans that display a unique spatial conformation (12). β-Mans have been established to be potent antigens for the adaptive immune response and to elicit specific infection-protective antibodies (13). In contrast to α-linked oligomannosides, which interact with C-type lectins, β-Mans specifically bind to galectin-3 (14, 15). PLM from *C. albicans* displays exclusively β-Mans (9) and has been shown to stimulate a variety of host signaling pathways (16, 17). Finally, it has been demonstrated that oral administration of synthetic β-Mans inhibited *C. albicans* colonization of the intestinal track in newborn mice, whereas synthetic α-Man did not (18). This experimental evidence strongly suggests that β-Mans contribute in a very specific manner to the virulence of *C. albicans*.

To further analyze the importance of β-Man residues in *C. albicans* virulence, the construction of mutants partially or fully depleted of these residues is necessary. However, the genes encoding *C. albicans* β-mannosyltransferases (*BMTs*) have not been identified. Furthermore, *S. cerevisiae* lacks β-Man residues and thus is a poor model organism to elucidate mechanisms of β-Man transfer. To date, attempts to decipher the role of β-Mans in virulence and cell wall structure of *C. albicans* have relied on the deletion of genes homologous to *S. cerevisiae* genes involved in PLM and PPM biosynthesis pathways upstream of actual β-mannose transfer (10, 19–21).

The presence of β-Mans is not restricted to *C. albicans* and has been demonstrated chemically in bacteria (22), *Leishmania* (23, 24), as well as in nonpathogenic yeasts such as *Pichia pastoris* (25).

Here we describe the identification and characterization of a novel family of genes involved in the transfer of β-Mans in *P. pastoris* and the pathogenic yeast *C. albicans*. Based on the available genomic sequence of *P. pastoris* and *C. albicans*, we identified four novel genes in *P. pastoris* and nine previously uncharacterized family members in *C. albicans*. Interestingly, homologs of these genes are not present in *S. cerevisiae*. A series of deletion strains of *P. pastoris* and *C. albicans* revealed different immunochemical phenotypes allowing us to gain insights into the function of these genes, including the impact of deletion on the global cell wall surface expression of β-Man epitopes. Moreover, systematic biochemical and structural analyses of these mutants have allowed us to define their role in different aspects of glycosylation, specifically core mannosylation of *N*-linked glycans in *P. pastoris* and β-mannose transfer onto mannan in *P. pastoris* and *C. albicans*.

EXPERIMENTAL PROCEDURES

Strains and Culture Conditions—*Escherichia coli* strains TOP10 or DH5α were used for recombinant DNA work. The yeast strains used in this study are listed in Table 1. Protein expression in *P. pastoris* cells was carried out as reported previously (26). *C. albicans* was grown in YPD-Arg-His-Urd medium (1% yeast extract, 2% peptone, 2% dextrose, 20 mg/liter arginine, 20 mg/liter histidine, 20 mg/liter uridine) at 37 °C for 16 h. All procedures for manipulating DNA were performed as described previously (27).

Deletion of BMT Genes—The *PpBMT* deletion alleles were generated by the PCR overlap method (28–30). In the first PCR 5′- and 3′-flanking regions of the *PpBMT* genes (500–1000 bp

in size) and the NAT or G418 resistance marker (31, 32) were amplified. Then, in a second reaction all three first round templates were used to generate an overlap product that contained all three fragments as a single linear allele. The final PCR product was then directly employed for transformation. Transformants were selected on YPD medium containing 200 μg/ml of G418 or 100 μg/ml of nourseothricin. In each case the proper integration of the mutant allele was confirmed by PCR. Each *C. albicans* gene, listed in Table 2, was deleted sequentially from strain BWP17 by PCR-based gene targeting (33). Two plasmids were used to release two selectable markers, *CaARG4* or *CaHIS1*, by NotI digest. Each marker was amplified by PCR using primers, including the first and the last 100 bp of gene-specific sequences. The generated disruption cassettes were used to transform BWP17 by the lithium acetate method as described previously (34). First, the selectable marker *ARG4* was used to generate independent heterozygous strains on synthetic dextrose plates supplemented with 20 mg/liter histidine and 20 mg/liter uridine. Correct insertion of the marker was verified by PCR. A second round of transformation was performed to delete the second allele with the *HIS1* marker by the same method. Homologous integration was verified by PCR.

Reintroduction of CaBMTs into the Null Strains—Each *CaBMT* open reading frame with ~500-bp upstream and downstream nucleotides was amplified by PCR using AccuPrime *Pfx* DNA polymerase (Invitrogen). The amplified fragment was cloned into the pCRII-TOPO vector (Invitrogen). Then the *CaBMT* gene was excised by digestion with SacI and NotI and ligated into SacI- and NotI-digested CIp10 (35) to obtain the reintegration vector. The *Cabmt*Δ null strains were transformed with StuI- or NcoI-digested reintegration vector. Transformants were screened by PCR to check the reintroduction of the *CaBMT* gene at the *RPS10* locus.

Whole Cell Protein Extraction and Western Analyses—*P. pastoris* cells were grown for 1 day in YPD medium at room temperature, and whole cell proteins were extracted using Y-PER (yeast protein extraction reagent, Pierce). Extracts were adjusted to the same protein concentration and analyzed by SDS-PAGE on 5–20% acrylamide gels. Membranes were probed with monoclonal antibody (mAb) 5B2 diluted 1:2000 (11).

Reporter Protein Purification and Release of N-Linked Glycans—The Kringle 3 domain was used as a model protein and was purified using His_6 tag as reported previously (26). The glycans were released and separated from glycoproteins by modification of a method reported previously (36). After the proteins were reduced and carboxymethylated, and the membranes blocked, the wells were washed three times with water. The protein was deglycosylated by the addition of 30 μl of 10 mM NH_4HCO_3, pH 8.3, containing 1 milliunit of *N*-glycanase (Glyko). After 16 h at 37 °C, the solution containing the glycans was removed by centrifugation and evaporated to dryness.

Complementation of P. pastoris OCH1 Gene—A 2.86-kb portion of the *P. pastoris OCH1* locus was amplified by PCR using primers PB198 5′-GAAGGATCCTGATATAGACCTGCGACACCATC-3′ and PB199 5′-ATTCGGATCCTGTCAATGGGAAGAGATGTCTTGTGCACA-3′. The resulting amplified fragment was subcloned into the *P. pastoris*/*E. coli* shuttle vector, pRCD252, constructed in our laboratory, which contains the *HPH* gene obtained from plasmid pAG32 (31), encoding for hygromycin resistance and the *P. pastoris TRP2* gene (integration locus). The resulting plasmid, pPB294, was linearized with a unique XbaI site located in the *TRP2* gene and transformed into *P. pastoris* by electroporation. Transformants were selected on YPD medium containing 300 mg/ml hygromycin.

Phenotypic Characterization of C. albicans Strains—To determine the growth rate, strains were grown in YPD-Arg-His-Urd medium at 37 °C for 9 h. Every 2 h, absorbance at 620 nm was measured. Hyphae formation was induced in 5% horse serum at 37 °C for 3 h. Chlamydospore formation was analyzed in medium containing 17 g/liter corn meal agar (Difco), 0.33% Tween 80, and 2% agar. Cultures were grown at 28 or 37 °C for 2–5 days. To determine the sensitivity to chemical agents, 5 μl of serial 1:10 dilutions of an overnight culture were spotted onto YPD-Arg-His-Urd agar plates containing 25 or 50 mM $CaCl_2$; 10, 20, or 50 μg/ml calcofluor white; 1 or 1.5 M NaCl; 0.001, 0.01, or 0.05% SDS. Plates were incubated at 37 °C for 1 day. For sensitivity assays to antifungal agents, including amphotericin B, 5-fluorocytosine, itraconazole, fluconazole, voriconazole, and caspofungin, strains were grown in Sabouraud's dextrose broth at 37 °C for 24 h. Two hundred microliters of 1:10 dilution in Shadomy medium were inoculated onto antifungigram plates and incubated at 30 °C for 24 h. The growth was determined by measuring the absorbance at 620 nm.

Phosphopeptidomannan Extraction and Western Analyses—PPM from cells grown in YPD-Arg-His-Urd medium was extracted as described previously (37). Briefly, cell pellets were suspended in 0.02 M citrate buffer and autoclaved at 125 °C for 90 min. Suspensions were harvested, and Fehling's solution was added to the supernatant to precipitate PPM. The PPM was then washed in methanol:acetic acid (8:1) and dried in a Speed-Vac concentrator. Sugar concentrations were estimated by the sulfuric phenol colorimetric method (38). PPM was analyzed by SDS-PAGE as described above. Membranes were probed with mAbs 5B2 diluted 1:1000 (11), B6.1 diluted 1:1000 (13), or B9E diluted 1:750 (39).

Extraction of Sphingolipids—Cells grown at 37 °C in YPD-Arg-His-Urd medium were broken with a French press (Aminco) at 20,000 psi, dialyzed, and lyophilized. Extraction was carried out as described previously (9), using chloroform:methanol:water extractions, butanol:water partitions, and chromatography on phenyl-Sepharose.

Analysis of the β-Mans Released from PPM and PLM—Previously extracted PPM or PLM were hydrolyzed in 0.1 N HCl for 1 h at 100 °C to release their β-Man-containing acid-labile moiety and neutralized with 0.2 N NaOH. Hydrolysates were then dried directly for PLM fractions or after an ultrafiltration step on Centricon YM-30 filters (Amicon) for PPM fractions and tagged with 0.15 M 8-aminonaphthalene-1,3,6-trisulfonate (ANTS) and 1 M sodium cyanoborohydride for 16 h at 37 °C as described previously (40). The dried samples were resuspended in glycerol:water (1:4) buffer. Electrophoresis of ANTS-labeled oligomannosides was performed on 25 or 35% (w/v) acrylamide separating gels, depending on the degree of polymerization of the oligomannosides, and a

5% acrylamide stacking gel at 600 V. Electrophoresis buffers were the same as for SDS-PAGE except that they did not contain SDS. Acid-hydrolyzed dextran and oligosaccharides were also tagged with ANTS and used as carbohydrate standards. Gels were dried, and images were acquired with the Gel Doc 2000 image analysis apparatus from Bio-Rad equipped with a 365 nm UV transilluminator.

NMR Analyses—For a better selectivity of mutant analyses, NMR experiments were performed separately on acid-resistant domains and on total mixtures of acid-labile oligomannosides isolated from PPM from mutant and wild-type strains. After mild acid hydrolysis, acid-labile and acid-resistant domains of mannans were separated by gel filtration and analyzed separately. All anomeric ^{1}H NMR signals were then readily attributed to α- and β-linked mannose residues according to published data (37, 41). 400-MHz ^{1}H NMR experiments were performed on a Bruker Avance® spectrometer equipped with a 5-mm TXI probehead (^{1}H resonating at 400.33 and ^{13}C at 100.25 MHz) (Centre Commun de Resonance Magnétique Nucléaire of USTL, Villeneuve d'Ascq, France). Samples were exchanged three times with $^{2}H_2O$ (99.97% deuterium atoms, Euriso-Top, Saclay, France) and intermediate lyophilization and then dissolved in 250 μl of $^{2}H_2O$. All experiments are performed with 5-mm Shigemi® tubes matched for $^{2}H_2O$. Experiments were recorded at 300 K for acid-labile fractions and 318 K for acid-stable fractions. Chemical shifts were expressed in ppm downfield from the signal of the methyl group of acetone (^{1}H resonating at 2.225 ppm and ^{13}C at 31.55 ppm for all fractions). For one-dimensional spectra, a spectral width of 4006 Hz was collected as 16,000 complex data points. Sine-shifted bell was used prior to Fourier transformation. Spectra were base-line corrected with a fourth order polynomial function.

Two-dimensional ^{1}H-^{13}C heteronuclear (HSQC) spectra were recorded without sample spinning, and data were acquired in the phase-sensitive mode using the time-proportional phase increment method.

ROESY experiment was acquired using 400-ms mixing time and acquired in States mode according to Bax and Davis (42). All parameters (90 hard pulses, soft pulses, delays, and pulse powers) were optimized for each experiment.

Flow Cytometry—Surface expression of β-Mans was determined as described previously (43). Briefly, cells grown on YPD-Arg-His-Urd medium at 37 °C for 16 h were washed with phosphate-buffered saline and then incubated with factor 5 and 6 sera (Iatron), diluted 1:200. After washing, cells were incubated with specific secondary fluorescein isothiocyanate-labeled antibody, diluted 1:100. Then cells were fixed with paraformaldehyde and analyzed by fluorescence-activated cell sorter.

Data acquisition was performed using an EPICS XLMCL4 (Beckman Coulter, High Wycombe, UK) equipped with an argon ion laser with an excitation power of 15 milliwatts at 488 nm. Fluorescence intensity of 5000 cells was analyzed using WINMDI software (available on line) and represented on a logarithmic scale.

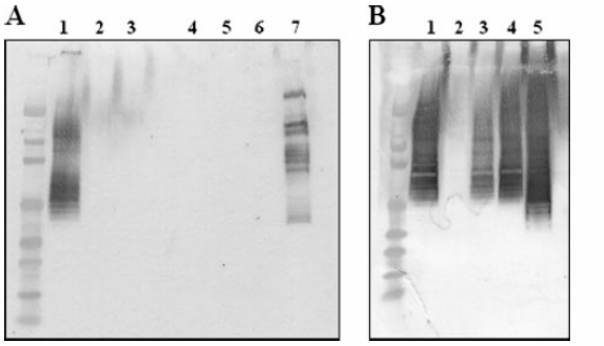


FIGURE 2. **Analysis of β-1,2-oligomannoside expression in *Ppbmt*Δ mutants.** Western blots of whole cell extracts from *P. pastoris* strains stained with mAb 5B2. *A, Ppbmt*Δ strains in *och1*Δ background (see the structure of glycans on the Fig. 1*A, och1*Δ mutant): *lane 1*, BK64 (wild-type, control strain); *lane 2*, BK3-1 (*och1*Δ in BK64); *lane 3*, YAS137 (*pno1*Δ-*mnn4*Δ in BK3-1); *lane 4*, PBP129 (*Ppbmt1*Δ in YAS137); *lane 5*, PBP130 (*Ppbmt2*Δ in YAS137); *lane 6*, PBP126 (*Ppbmt3*Δ in YAS137); *lane 7*, PBP135 (*Ppbmt4*Δ in YAS137). *B, Ppbmt*Δ strains complemented with the *OCH1* gene (see the structure of glycans on the Fig. 1*A, wild-type*): *lane 1*, PBP282 (control strain); *lane 2*, PBP283 (*Ppbmt1*Δ); *lane 3*, PBP284 (*Ppbmt2*Δ); *lane 4*, PBP286 (*Ppbmt3*Δ); *lane 5*, PBP287 (*Ppbmt4*Δ).

RESULTS

P. pastoris Glycoproteins Express β-1,2-Oligomannoside Epitopes—Previous studies reported the existence of β-Mans in *P. pastoris*. Vinogradov *et al.* (25) employed NMR spectroscopy to show the presence of β-Mans on the outer chain of *P. pastoris* mannan. To confirm these results, a Western blot of whole cell extracts stained with mAb 5B2 reacting with β-1,2-linked mannobiose as a minimal epitope was performed and revealed the presence of β-Mans in the parental *P. pastoris* strain BK64 (Fig. 2*A*, *lane 1*), but not in BK3-1 (*och1*Δ mutant) or YAS137 (*pno1*Δ, *mnn4L1*Δ in BK3-1 background)[4] strains depleted in their outer chain (Fig. 2*A*, *lanes 2* and *3*).

Identification of a Family of Putative P. pastoris β-Mannosyltransferase Genes—We speculated that the protein(s) involved in β-linked mannose transfer might share sequence similarity or some specific structural motifs with other mannosyltransferases. Therefore, we performed a BLAST search of a *P. pastoris* partial genomic sequence (Integrated Genomics, Inc.) for genes similar to the known fungal Golgi-residing mannosyl- and mannosylphosphate transferases. Using the C-terminal part of *S. cerevisiae* Mnn4 protein containing lysine-glutamic acid rich repeats (KKKKEEEE) (44) as a probe, a sequence with similar lysine-glutamic acid-rich repeats was identified. A detailed analysis of the corresponding contiguous DNA sequence revealed the presence of an open reading frame (ORF) encoding a hypothetical protein of 644 amino acids with a putative N-terminal transmembrane domain, suggesting a possible Golgi localization. In subsequent searches we identified, in the genome of *P. pastoris*, three more ORFs encoding proteins that share significant sequence similarity (Fig. 3). Interestingly, two of these ORFs are located adjacent to each other. Based on the results of the subsequent experiments, we named these new genes *PpBMT1*, *PpBMT2*, *PpBMT3*, and *PpBMT4* (for *P. pastoris* β-mannosyl transfer 1–4).

[4] T. A. Stadheim, P. Bobrowicz, H. Li, R. C. Davidson, T. U. Gerngross, and S. Wildt, unpublished data.

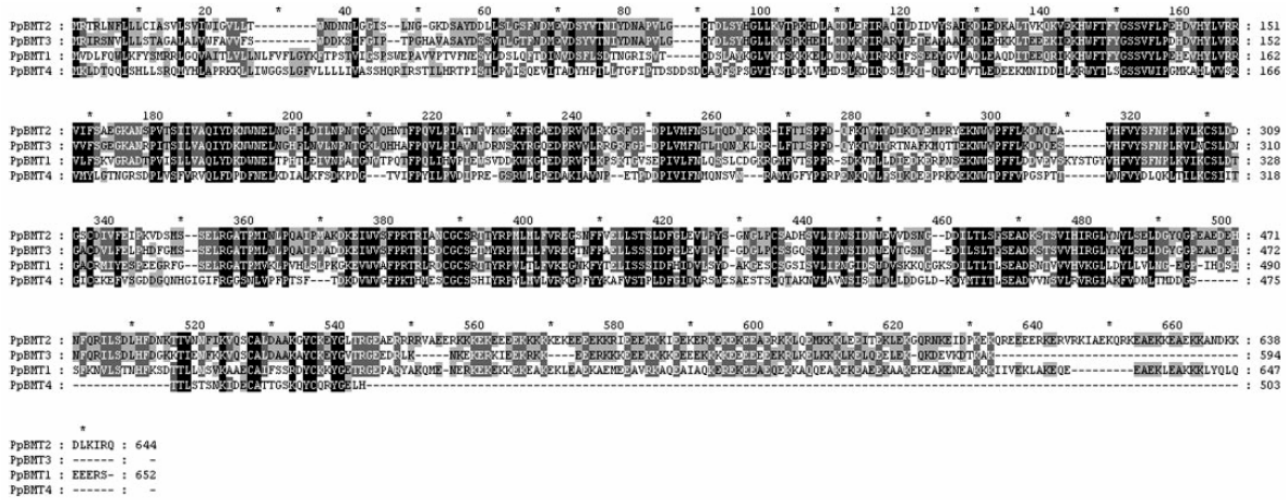


FIGURE 3. **Conserved regions between the four *P. pastoris* β-mannosyltransferase protein sequences.** The alignment was done with ClustalW and edited with Genedoc.

TABLE 1
Strains used in this study

Strain	Parental strain	Genotype	Refs.
BK64	GS115 (Invitrogen)	*his4; K3-ZEO*^{R}	26
BK3-1	BK64	*och1Δ::HIS4; K3-ZEO*^{R}	26
YAS137	BK3-1	*och1Δ::HIS4; pno1Δ::HYG*^{R}*; mnn4L1Δ::NAT*^{R}*; K3-ZEO*^{R}	[a]
PBP126	YAS137	*och1Δ::HIS4; pno1Δ::HYG*^{R}*; mnn4L1Δ::NAT*^{R}*; Ppbmt3Δ::G418*^{R}*; K3-ZEO*^{R}	This study
PBP129	YAS137	*och1Δ::HIS4; pno1Δ::HYG*^{R}*; mnn4L1Δ::NAT*^{R}*; Ppbmt1Δ::G418*^{R}*; K3-ZEO*^{R}	This study
PBP130	YAS137	*och1Δ::HIS4; pno1Δ::HYG*^{R}*; mnn4L1Δ::NAT*^{R}*; Ppbmt2Δ::G418*^{R}*; K3-ZEO*^{R}	This study
PBP135	YAS137	*och1Δ::HIS4; pno1Δ::HYG*^{R}*; mnn4L1Δ::NAT*^{R}*; Ppbmt4Δ::G418*^{R}*; K3-ZEO*^{R}	This study
PBP282	BK3-1	*och1Δ::HIS4; OCH1-HYG*^{R}*; K3-ZEO*^{R}	This study
PBP283	PBP245	*och1Δ::HIS4; Ppbmt1Δ::NAT*^{R}*; OCH1-HYG*^{R}*; K3-ZEO*^{R}	This study
PBP284	PBP277	*och1Δ::HIS4; Ppbmt2Δ::NAT*^{R}*; OCH1-HYG*^{R}*; K3-ZEO*^{R}	This study
PBP286	PBP128	*och1Δ::HIS4; Ppbmt3Δ::G418*^{R}*; OCH1-HYG*^{R}*; K3-ZEO*^{R}	This study
PBP287	PBP247	*och1Δ::HIS4; Ppbmt4Δ::NAT*^{R}*; OCH1-HYG*^{R}*; K3-ZEO*^{R}	This study
BWP17	RM1000	*arg4::hisG/arg4::hisG;his1::hisG/his1::hisG;ura3Δ::limm434/ura3Δ::limm434*	62
AL84	BWP17	*bmt4Δ::ARG4/bmt4Δ::HIS1*	This study
AL86	BWP17	*bmt2Δ::ARG4/bmt2Δ::HIS1*	This study
AL90	BWP17	*bmt3Δ::ARG4/bmt3Δ::HIS1*	This study
AL91	BWP17	*bmt1Δ::ARG4/bmt1Δ::HIS1*	This study
AL93	AL91	*bmt1Δ::ARG4/bmt1Δ::HIS1, RPS10*::CIp10-*BMT1-URA3*	This study
AL94	AL86	*bmt2Δ::ARG4/bmt2Δ::HIS1, RPS10*::CIp10-*BMT2-URA3*	This study
AL95	AL90	*bmt3Δ::ARG4/bmt3Δ::HIS1, RPS10*::CIp10-*BMT3-URA3*	This study
AL96	AL84	*bmt4Δ::ARG4/bmt4Δ::HIS1, RPS10*::CIp10-*BMT4-URA3*	This study

[a] T. A. Stadheim, P. Bobrowicz, H. Li, R. C. Davidson, T. U. Gerngross, and S. Wildt, unpublished data.

*Analysis of β-1,2-Oligomannoside Expression in Ppbmt*Δ *Mutants*—To characterize the function of these novel genes, *PpBMT* deletion strains were generated in *P. pastoris* BK3-1 and YAS137 mutants (lacking outer chain, see Fig. 1*A*) using fusion PCR-generated mutant alleles (Table 1). Disruptions were confirmed by PCR (data not shown). Both host strains express the Kringle 3 domain of human plasminogen (K3) used as an *N*-glycosylated reporter protein (26).

Similar to the YAS137 parental strain (Fig. 2*A*, *lane 3*), Western blot analyses of whole cell extracts from *Ppbmt1-3*Δ mutants generated in this background (Fig. 2*A*, *lanes 4–6*) revealed that they did not express β-Man epitopes recognizable by mAb 5B2. However, strikingly, these epitopes were detected in the *Ppbmt4*Δ mutant (Fig. 2*A*, *lane 7*) suggesting that β-Man epitopes are accessible following *PpBMT4* deletion. Similar results were obtained for deletion mutants constructed in the BK3-1 strain (data not shown). Whole cell analysis was further substantiated by Western blot analyses of the secreted K3 protein (data not shown).

To further characterize the core glycan structure in *Ppbmt*Δ mutants, *N*-glycans were released from secreted Kringle 3 protein by treatment with peptide:*N*-glycosidase (26) and analyzed by MALDI-TOF mass spectrometry with or without preliminary mannosidase digestion (Fig. 4). As shown previously by Davidson *et al.* (30)with BK3-1, in reference strain YAS137 (*och1*Δ background), core glycans displayed extensive and heterogeneous mannosylation (Fig. 4*A*). β-Mannosylation of a fraction of these glycans blocked the action of α-1,2 mannosidase (peak with 10 mannoses; Fig. 4*B*). By contrast, disruption of the *PpBMT2* gene strongly reduced the degree of mannosylation of the core glycans (Fig. 4*C*) and eliminated glycans resistant to α-1,2 mannosidase treatment (loss of peak with 10 mannoses in Fig. 4*D*) and yielded the $Man_5GlcNAc_2$ structure as expected for core *N*-glycans with terminal α-1,2 mannose extensions. In *Ppbmt4*Δ strain, an intermediate mannosylation of the core was observed (Fig. 4*E*), and after α-1,2 mannosidase treatment, recalcitrant peaks were still present, but their sizes were smaller than those of the parental strain (Fig. 4*F*). No

reduction in size or distribution of the recalcitrant peaks was observed in the *Ppbmt1*Δ and *Ppbmt3*Δ strains (data not shown). These observations suggest that PpBmt2p initiates β-mannosylation of core *N*-linked glycans and that PpBmt4p is responsible for addition of a hexose to the extended β-Man chain (Fig. 1*A*), which obscures the mAb 5B2 epitope, as suggested by the results on Fig. 2*A*, *lane 7*.

To test the effect of *PpBMT* deletions on β-Man epitopes on the outer chain, *Ppbmt1–4*Δ was transformed with the wild-type *P. pastoris OCH1* gene. Western blot analyses of whole cell extracts with mAb 5B2 (Fig. 2*B*) revealed that *Ppbmt1*Δ did not express β-Man epitopes (*lane 2*), in contrast to the parental control strain (Fig. 2*B*, *lane 1*). This suggests that PpBmt1p is involved in β-mannosylation of the outer chain (Fig. 1*A*). A slight reduction in the signal was observed in the *Ppbmt2*Δ strain (Fig. 2*B*, *lane 3*), whereas *PpBMT3* deletion had no discernible effect on mAb 5B2 reactivity (*lane 4*). Interestingly, *PpBMT4* deletion resulted in a subtle increase in staining, consistent with results obtained in the *och1*Δ mutant background.

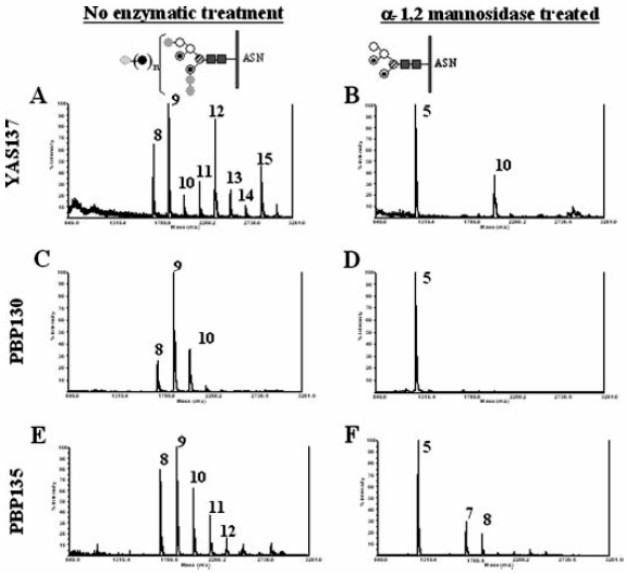


FIGURE 4. **Analysis of *N*-glycans released from K3 reporter protein produced in *Ppbmt*Δ mutants in a YAS137 strain background.** *N*-Glycans were released from K3 by peptide:*N*-glycosidase treatment and directly analyzed by MALDI-TOF mass spectrometry (*No enzymatic treatment panel*). In addition, a portion of the released glycans was digested with α-1,2-mannosidase prior to MALDI-TOF MS analysis (*α-1,2 mannosidase-treated panel*). *Numbers* next to the peaks correspond to the number of mannose residues present in *N*-glycans, calculated on the basis of the molecular mass of the peak (*e.g.* the peak designated with the number 8 corresponds to a $Man_8GlcNAc_2$ *N*-glycan). *A* and *B*, YAS137 (control strain); *C* and *D*, PBP130 (*Ppbmt2*Δ); *E* and *F*, PBP135 (*Ppbmt4*Δ).

The putative functions of PpBmt1–4p, as deduced from this phenotypic analysis, are schematized in Fig. 1*A*.

Identification of a Related Family of BMT Genes in C. albicans—To determine whether these four new *P. pastoris* genes have homologs in other organisms, we performed BLAST searches against public genome data bases. Homologs were found in the fungal species *C. albicans, Candida glabrata, Debaryomyces hansenii, Saccharomyces castellii, Saccharomyces kluyveri, Aspergillus fumigatus,* and *Aspergillus terreus*. This suggests that these *BMT* genes form a new family of fungus-specific genes. By homology with the PpBmt2p sequence, nine proteins with unknown function were revealed in *C. albicans*, a fungal species known to display large amounts of β-Mans on various glycoconjugates (8, 11). Table 2 summarizes some characteristics of the newly identified putative proteins. The CaBmtp predicted proteins display strong homology to each other (Fig. 5). However, none of these sequences contains the conserved glycosyltransferase domain (pfam Gly_transf_sug; PF04488).

To determine the function of these *C. albicans* proteins, alleles of each corresponding gene were disrupted using a PCR-based method described previously (33). This generated a set of strains carrying disruptions of single or both alleles of each gene in the parental strain BWP17 (Table 1).

Preliminary characterization of the nine deletion strains revealed phenotypic alterations of PPM, PLM, and mannoprotein β-mannosylation (data not shown). We first focused our analyses on four deletion strains selected for displaying phenotypic alterations of PPM β-mannosylation, as far as the distribution of β-Man within this molecule is clearly known (Fig. 1*B*) and accessible to structural analyses. We named the corresponding genes *CaBMT1–4*.

Deletion of CaBMTs Had No Effect on Growth and Morphogenesis in Vitro—When the growth of each *CaBMT* deletion strain was monitored in rich medium, no differences were observed. Similarly, all strains had the same ability to produce hyphae in serum or chlamydospores on RAT medium. None of the strains displayed increased sensitivity to $CaCl_2$, calcofluor white, NaCl, or SDS. No differences in growth were observed on media containing increasing concentrations of antifungal agents. Results from Alcian blue binding assays demonstrated no effect on glycan phosphorylation in mutant strains (data not shown).

CaBMT1–4 Deletions Affect β-Mannosylation of Either the Acid-stable or the Acid-labile Fractions of Phosphopeptidomannan—PPMs were extracted from selected strains and first analyzed by Western blot. Staining with mAb B9E, specific

TABLE 2
Identification of *C. albicans* genes

Accession no.	Gene name	Gene name (*Candida* data base)	Protein length	Description	Refs.
orf19.6782	*BMT1*	*IFQ2*	685	Unknown function	This study
orf19.54	*BMT2*	*IFQ3*	655	Unknown function	This study
orf19.3282	*BMT3*	*IPF300*	550	Unknown function	This study
orf19.5612	*BMT4*	*IPF15015*	790	Unknown function	This study
orf19.1464	*BMT5*	*IFQ4*	613	Unknown function	Mille *et al.*, unpublished
orf19.5602	*BMT6*	*IPF19772*	647	Unknown function	Mille *et al.*, unpublished
orf19.342	*BMT7*	*IPF1520*	652	Unknown function	Mille *et al.*, unpublished
orf19.860	*BMT8*	*IPF7647*	695	Unknown function	Mille *et al.*, unpublished
orf19.4673	*BMT9*	*IFQ1*	783	Unknown function	Mille *et al.*, unpublished

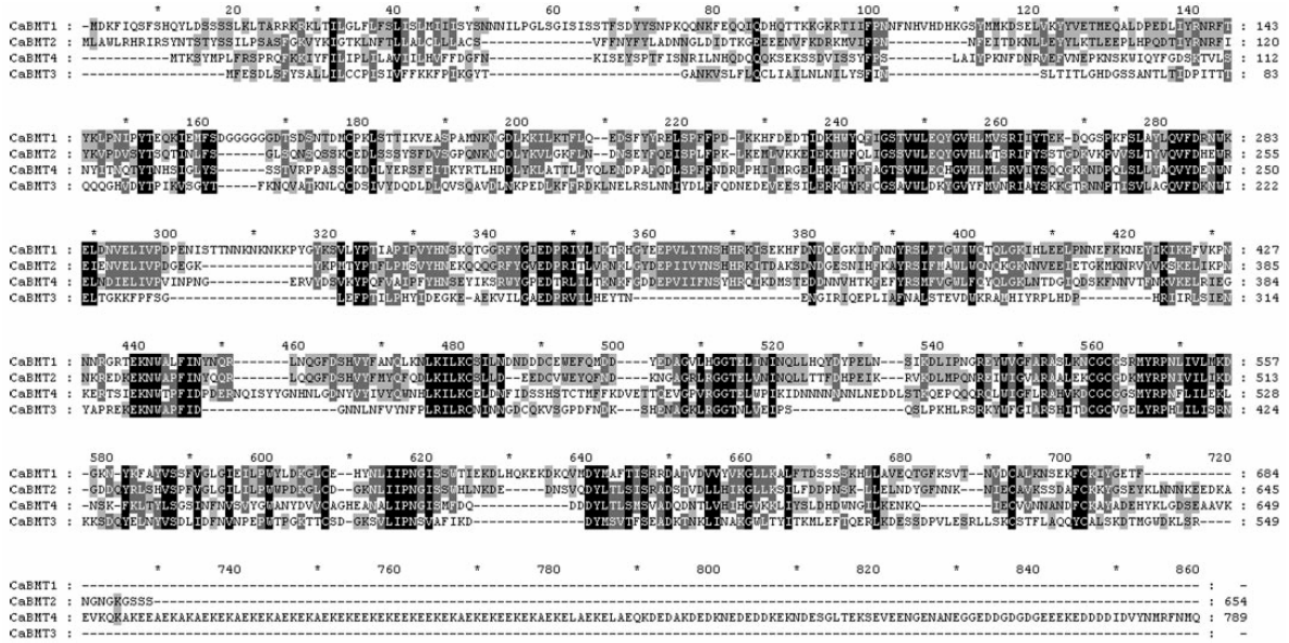


FIGURE 5. **Conserved regions between the four *C. albicans* β-mannosyltransferase protein sequences.** The alignment was done with ClustalW and edited with Genedoc.

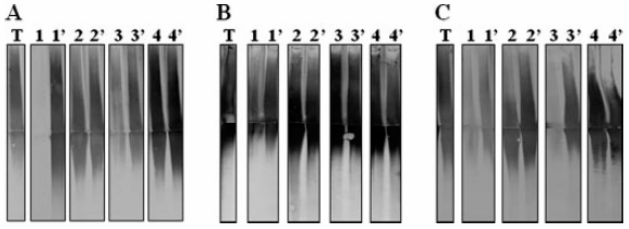


FIGURE 6. **Analysis of phosphopeptidomannan β-mannosylation in *Cabmt*Δ mutants.** Western blots of PPMs (see the structure of PPM on Fig. 1*B*) extracted from deletion strains stained with the mAbs as follows: *A*, B9E (specific for β-oligomannoside present on the nonreducing end of α-1,2 chains of the acid-stable fraction of serotype A); *B*, 5B2 (specific for β-1,2-mannobiose); and *C*, B6.1 (specific for β-1,2-mannotriose). *Lane T*, BWP17 (parental strain); *lane 1*, AL91 (*Cabmt1*Δ); *lane 1′*, AL93 (*Cabmt1*Δ*/BMT1*); *lane 2*, AL86 (*Cabmt2*Δ); *lane 2′*, AL94 (*Cabmt2*Δ*/BMT2*); *lane 3*, AL90 (*Cabmt3*Δ); *lane 3′*, AL95 (*Cabmt3*Δ*/BMT3*); *lane 4*, AL84 (*Cabmt4*Δ); *lane 4′*, AL96 (*Cabmt4*Δ*/BMT4*).

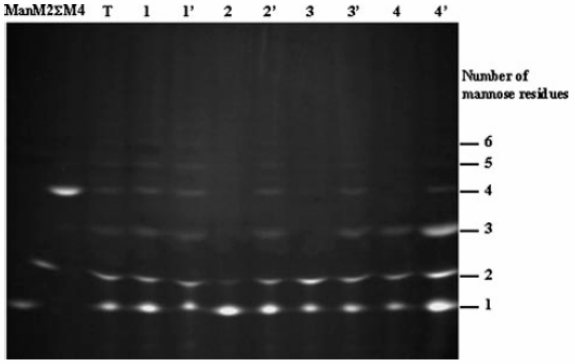


FIGURE 7. **Analysis of phosphopeptidomannan acid-labile fraction.** Oligomannosides released by acid hydrolysis from PPMs previously extracted were analyzed by FACE. Different carbohydrate standards were used to evaluate the monomer number in the oligomannoside chains: *Man*, mannose; *M2*Σ, synthetic di-mannoside; and *M4*, tetra-mannoside. *Lane T*, BWP17 (parental strain); *lane 1*, AL91 (*Cabmt1*Δ); *lane 1′*, AL93 (*Cabmt1*Δ*/BMT1*); *lane 2*, AL86 (*Cabmt2*Δ); *lane 2′*, AL94 (*Cabmt2*Δ*/BMT2*); *lane 3*, AL90 (*Cabmt3*Δ); *lane 3′*, AL95 (*Cabmt3*Δ*/BMT3*); *lane 4*, AL84 (*Cabmt4*Δ); *lane 4′*, AL96 (*Cabmt4*Δ*/BMT4*).

for β-Mans at the nonreducing end of α-1,2 chains of the acid-stable fraction of *C. albicans* serotype A strains (antigenic factor 6, see Fig. 1*B*), revealed a complete loss of reactivity for *Cabmt1*Δ and *Cabmt3*Δ strains (Fig. 6*A*, *lanes 1* and *3*). Reactivity to mAbs 5B2 and B6.1, specific for β-1,2-mannotriose, was reduced (Fig. 6, *B* and *C*) but was completely abolished (data not shown) after acid hydrolysis, which removes the acid-labile fraction of PPM (see Fig. 1*B*). This suggests that *CaBMT1* and *CaBMT3* deletions affect β-mannosylation of the acid-stable part of PPM. In contrast, PPMs from *Cabmt2*Δ and *Cabmt4*Δ strains displayed wild-type reactivity with mAb B9E (Fig. 6*A*, *lanes 2* and *4*) but a strongly reduced reactivity with mAb 5B2 suggesting that β-mannosylation of the acid-labile part of PPM is affected in these mutants.

CaBmt2p, CaBmt3p, and CaBmt4p Are Involved in β-Mannosylation of the PPM Acid-labile Fraction—The oligomannosides released from PPMs by mild acid hydrolysis were analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE) (Fig. 7). In comparison with the parental strain (*lane T*), deletion of *CaBMT1* had no effect on β-mannosylation of the PPM acid-labile fraction (Fig. 7, *lane 1*). In contrast, FACE analysis of the acid-labile fraction from *Cabmt2*Δ, *Cabmt3*Δ, and *Cabmt4*Δ (Fig. 7, *lanes 2–4*) revealed the accumulation of a single mannose, a mannobiose, and a mannotriose, respectively, suggesting that these three *CaBMTs* are successively involved in β-mannosylation of the PPM acid-labile fraction. *Cabmt3*Δ also displayed a slight expression of β-Man with a degree of polymerization >2, which may arise from redundant activity of other CaBmtps (Fig. 7, *lane 3*).

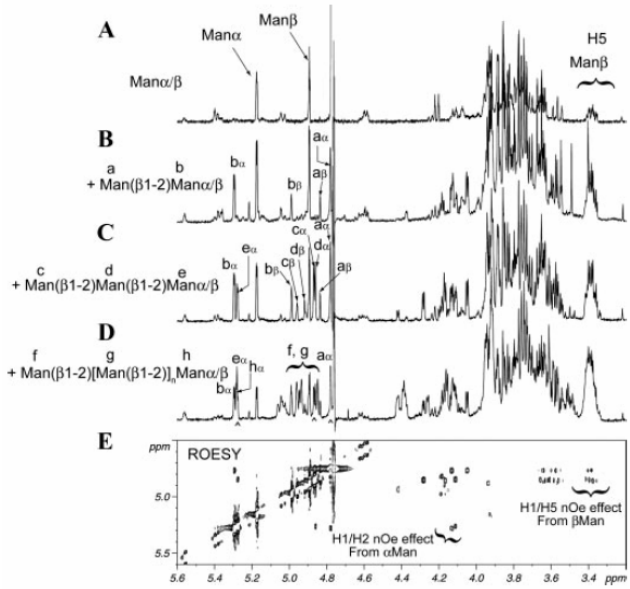


FIGURE 8. **NMR analysis of acid-labile domain.** NMR structural analysis of purified acid-labile oligomannosides after mild acid hydrolysis from wild-type strain and mutants. One-dimensional ^{1}H NMR spectrum of oligomannoside fraction from the following: *A*, AL86 (*Cabmt2*Δ); *B*, AL90 (*Cabmt3*Δ); *C*, AL84 (*Cabmt4*Δ); and *D*, BWP17 (parental strain). *E*, ^{1}H-ROESY spectrum of oligomannoside fraction from AL84 (*Cabmt4*Δ). The detailed structure of β-Man chains shows from which part of the molecules the signals labeled in the various spectra originated. Each anomer *a–h* is observed as two distinct signals according to the α or β forms of the reducing mannose residue of the oligomannosylated chain.

These results were further confirmed by NMR analysis of the acid-labile fractions from all strains composed of different sets of free oligosaccharides with a reducing mannose residue. As shown in Fig. 8*D*, the ^{1}H NMR spectrum of the acid-labile fraction from the serotype A wild-type strain displayed a complex pattern of H-1 β-Man signals between δ 4.75 and 5.10, as well as several signals around δ 5.30 attributed to the reducing H-1 α-Man residues of oligomannosides. The α-anomer of free mannose residues is also easily identified at δ 5.17. These parameters are in agreement with the presence of a heterogeneous family of acid-labile β-Mans as described previously in a *C. albicans* serotype A strain (37).

Analysis of *Cabmt2*Δ (Fig. 8*A*) showed a much simpler ^{1}H NMR spectrum dominated by two signals at δ 5.173 ($^3J_{1,2}$ 1.7 Hz) and 4.890 ($^3J_{1,2} < 0.8$ Hz), which were tentatively attributed to the H-1 signals of α- and β-anomers of free mannose in accordance with published data. This was confirmed by observing their respective spin systems on COSY and ROESY spectra (data not shown). No signal from substituted α-Man residues or internal β-Mans was identified, indicating the total absence of oligosaccharides in the acid-labile fraction of *Cabmt2*Δ mannan. These results confirmed that acid hydrolysis of *Cabmt2*Δ PPM released single mannose residues exclusively.

In addition to free reducing α- and β-Man H-1 signals as observed in ^{1}H NMR spectra of *Cabmt2*Δ, the *Cabmt3*Δ spectra exhibited two intense anomeric signals at δ 5.29 (bα) and 4.78 (aα) (Fig. 8*B*), which correlated with their respective H-2 at δ 4.13 and 4.05 on a COSY spectrum (data not shown). The strong downfield shifts of bα H-1 and H-2 signals at 5.29 and 4.13 ppm compared with free α-Man indicated that this residue of α-Man is substituted in the C-2 position. The H-1 signal of β-Man residue (aα) of the α-form of mannobiosyl was identified based on the H1/H5 internal nuclear Overhauser effect observed from the signal at δ 4.78 (data not shown). Furthermore, ROESY analysis showed a cross-peak between the aα H-1 signal at δ 4.78 ppm and the H-2 signal at δ 4.13, confirming that the β-Man residue was linked to the C-2 position of the α-Man residue (bα). The H-1 of α-Man (bβ) and of the β-Man (aβ) residues of the β form of mannobiosyl was also identified as minor signals at δ 4.99 and δ 4.82, respectively. Thus, NMR data confirmed that the acid-labile fraction of *Cabmt3*Δ mannan contains a mixture of free mannose and Man(β1–2)Man disaccharide.

Compared with *Cabmt3*Δ, the ^{1}H NMR spectrum of *Cabmt4*Δ (Fig. 8*C*) showed three additional major signals at δ 5.28 (eα), 4.87 (cα), and 4.86 (dα), which were attributed to the H-1 of 1,2-linked reducing α-Man, terminal nonreducing β-Man, and 1,2-linked internal β-Man residues of the α anomer of mannotriosyl, respectively. Attribution was based on published data and observation of internal H1/H2 and H1/H5 nuclear Overhauser effects on ROESY spectra (Fig. 8*E*). Two additional minor signals at δ 4.96 (cβ) and 4.92 (dβ) were attributed to H-1 of terminal nonreducing β-Man and 1,2-linked internal β-Man residues of the β anomer of mannotriosyl, respectively. Indeed, as demonstrated previously, H-1 and H-2 of internal and terminal nonreducing mannose residues of mannotriosyl showed splitting because of the α and β forms of the reducing mannose residue (37). These data demonstrate that the acid-labile domain of *Cabmt4*Δ mannan contains a simple mixture of free mannose, Man(β1–2)Man and Man(β1–2)Man(β1–2)Man.

CaBmt1p and CaBmt3p Are Involved in β-Mannosylation of the PPM Acid-stable Fraction—One-dimensional ^{1}H NMR analysis of the acid-stable fraction of mannan from the serotype A wild-type strain showed a complex pattern of H-1 protons of α- and β-linked Man residues (Fig. 9*A*). Nondecoupled ^{13}C-^{1}H HSQC heteronuclear analysis (data not shown) allowed us to distinguish β-Man from α-Man residues because of their low $^1J_{CH}$ constant around 157 ppm, compared with 171 ppm for α-Man. As shown on the ^{13}C-^{1}H HSQC heteronuclear spectrum (Fig. 9*E*), four signals labeled *a–d* were attributed to β-Man H-1 depending on their location within the β-mannan chain. It has been reported previously that serotype B strains lack β-mannosylation on the acid-stable fraction of mannans (7). In accordance with these data, the NMR spectrum of a serotype B strain did not show any signal originating from β-Man residues when the same analysis was performed (Fig. 9*B*).

NMR analysis of the acid-stable fraction of *Cabmt1*Δ (Fig. 9*C*) showed a total absence of signals attributed to β-Man residues, identical to the serotype B strain (Fig. 9*B*). In contrast, NMR analysis of the acid-stable fraction of *Cabmt3*Δ (Fig. 9*D*) showed a single β-Man H-1 signal at 4.79 ppm attributed to Man(β1–2)Man(α1-terminal) motif. Presence of a single signal out of four was confirmed by nondecoupled and decoupled ^{13}C-^{1}H HSQC analysis (data not shown). These data directly confirmed that the *Cabmt1*Δ strain does not contain any

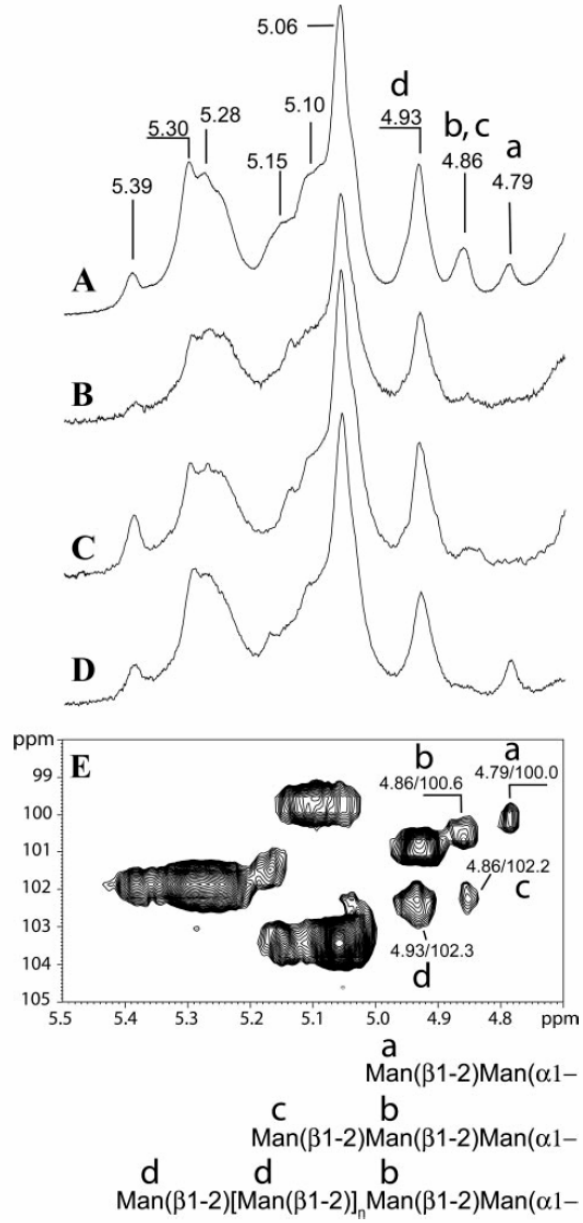


FIGURE 9. **NMR analysis of acid-stable domain.** NMR structural analysis of purified acid-stable domain of mannan after mild acid hydrolysis from wild-type strains and mutants. ^{1}H NMR spectrum of acid-stable domain from the following: *A, C. albicans* serotype A strain; *B, C. albicans* serotype B strain; *C,* AL91 (*Cabmt1*Δ); and *D,* AL90 (*Cabmt3*Δ). *E,* ^{13}C-^{1}H HSQC heteronuclear spectrum of AL90 (*Cabmt3*Δ). *E, a–d,* signals have been identified as β-Man residues based on their low $^{1}J_{CH}$ constant around 157 ppm, whereas all other signals were identified as α-Man, based on their $^{1}J_{CH}$ constant around 171 ppm, as observed by the nondecoupled ^{13}C-^{1}H HSQC heteronuclear experiment. The detailed structure of β-Man chains shows from which part of the molecules the signals labeled in the various spectra originated.

β-Man residues on the acid-resistant domain of mannan, whereas *Cabmt3*Δ retains a single β-linked mannose residue.

Control of Bmtp Activity by Construction of Revertant Strains—Although all of the deletions had an impact of β-mannosylation, it was necessary to confirm specific activity by complementation of one gene copy. Wild copies of *CaBMT* genes were

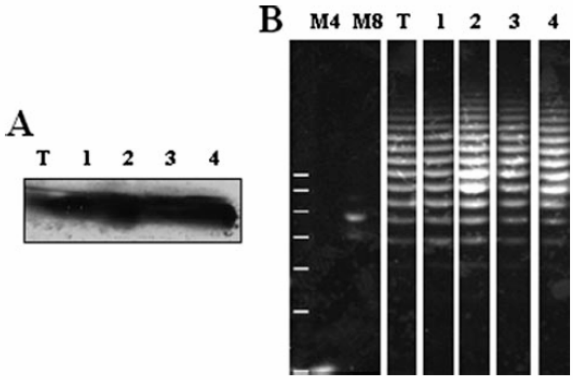


FIGURE 10. **Analysis of phospholipomannan β-mannosylation in *Cabmt*Δ mutants.** *A,* Western blot analysis of PLM stained with mAb 5B2 (specific for β-1,2-mannobiose); *B,* FACE analysis of oligomannosides released from PLM by acid hydrolysis. For the FACE analysis, *M4* and *M8* (tetra- and octo-mannoside) were used to evaluate the monomer number in the oligomannoside chains. *Lane T,* BWP17 (parental strain); *lane 1,* AL91 (*Cabmt1*Δ); *lane 2,* AL86 (*Cabmt2*Δ); *lane 3,* AL90 (*Cabmt3*Δ); *lane 4,* AL84 (*Cabmt4*Δ).

reintegrated into the *RPS10* locus of the corresponding *Cabmt*Δ null strains. Complementation was confirmed by PCR (data not shown). Phenotypic analyses either by Western blot for *Cabmt1*Δ*/BMT1* and *Cabmt3*Δ*/BMT3* (Fig. 6, *lanes 1′* and *3′*) or by FACE for *Cabmt2*Δ*/BMT2, Cabmt3*Δ*/BMT3,* and *Cabmt4*Δ*/BMT4* (Fig. 7, *lanes 2′*, *3′*, and *4′*) showed complete restoration of phenotype assigning definitely a function for these representative members of the CaBmtp family.

CaBMT1–4 Deletions Do Not Affect PLM β-Mannosylation—Considering homology of some β-Man acceptor sites (namely -Man-P-Man) between PPM and PLM (Fig. 1, *B* and *C*), we investigated whether deletion of *CaBMT1–4* had an impact on PLM. No effect on PLM β-mannosylation was observed in Western blot analysis of whole cell extracts (Fig. 10*A*). These results were confirmed by the more analytical FACE method performed on oligomannosides released from PLM in acidic conditions. Altogether these data demonstrate substrate specificity for CaBmt1–4p between PPM and PLM.

Single Deletions Have a Slight Effect on β-1,2-Oligomannoside Surface Expression—Flow cytometry assays were performed using factor 5 serum recognizing homopolymers of β-Mans in the acid-labile fraction of *C. albicans* and factor 6 serum recognizing β-Man residues at the nonreducing end of the α-1,2 lateral chain specific to the PPM acid-stable fraction of *C. albicans* serotype A (Fig. 1*B*). Surface expression of β-Mans of deletion strains detected with polyclonal antibodies was similar to that of the parental strain (Fig. 11) except for *Cabmt1*Δ, which displayed a significant reduction in antigenic factor 6 expression (Fig. 11*A*), and for *Cabmt3*Δ whose reactivity slightly increased with serum factor 5 (Fig. 11*C*). In contrast to results obtained with purified PPM, deletions did not result in dramatic changes suggesting that, besides PLM which was not affected, other cell wall molecules contribute to compensate cell wall expression of β-Mans.

DISCUSSION

The significant increase in opportunistic fungal infections observed over the last 2 decades represents an important med-

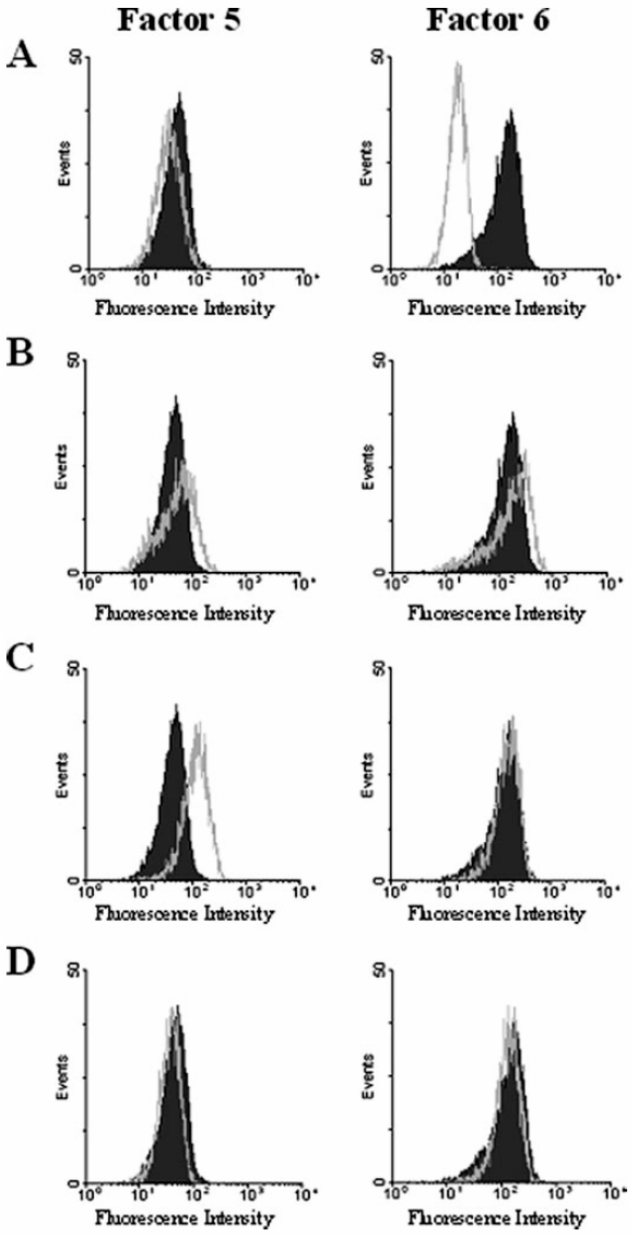


FIGURE 11. **Flow cytometric analysis of β-linked oligomannoside surface expression.** β-1,2-Oligomannoside expression of deletion strains (*gray line*) revealed with factor 5 and 6 sera was compared with that of the parental strain (*black area*). *A*, AL91 (*Cabmt1*Δ); *B*, AL86 (*Cabmt2*Δ); *C*, AL90 (*Cabmt3*Δ); and *D*, AL84 (*Cabmt4*Δ).

ical challenge. In particular, *C. albicans* is one of the most frequent causes of nosocomial bloodstream infections and is associated with high mortality rates (45). Despite increased investments in antifungal therapies, only limited progress has been achieved in controlling nosocomial candidiasis (1). *C. albicans* exists as a harmless, commensal micro-organism in the majority of immunocompetent people. However, the increase in number of patients with compromised immunity has contributed to the observed rise in opportunistic fungal infections. Several traits, including differential expression of adhesins (46–48) and lytic enzymes (49), morphological changes (50), stress response (51), and changes in the glyoxylate cycle (52) have been linked to pathogenicity of *C. albicans*. However, the way in which *C. albicans* coordinates these factors in its continuous adaptation to host conditions is still poorly understood.

Among factors contributing to *C. albicans* virulence, which still pose unresolved questions, are β-Mans. β-Mans contribute to at least two important pathogenic mechanisms as follows: adhesion to the host cells (14) and modulation of the immune response (53). At the cell wall level, β-Mans are known to be associated with PPM (7, 8) and PLM (9) as well as several studies which suggest their association with cell wall mannoproteins. Failure to identify genes involved in the biosynthesis of β-Mans has hindered the implementation of genetic approaches to study their function, *i.e.* due to their notable absence in most model fungal species.

In this study, we describe the identification and characterization of a new family of genes responsible for the synthesis of β-Mans in the following two yeast species: *P. pastoris* and *C. albicans*. First, employing glycan analyses by MALDI-TOF mass spectrometry and Western blot analyses of glycoproteins with mAb 5B2, raised against *C. albicans* cell wall associated β-Mans, we were able to confirm the presence of β-Mans in *P. pastoris*. The ability of *P. pastoris* cells to modify glycoproteins with immunogenic β-Mans could be a potential obstacle in the successful implementation of this organism as a production platform for therapeutic glycoproteins (54). To overcome this problem, we attempted to identify the genes responsible for the transfer of β-1,2-mannose in *P. pastoris*. Making the assumption that putative β-1,2-mannosyltransferases might share some structural motifs or domains with other yeast Golgi glycosyltransferases, we employed BLAST to systematically search for novel genes with such sequence similarities. Golgi mannosyltransferases are typically type II membrane proteins that add α-1,2-, α-1,3-, and α-1,6-linked mannoses to protein- and lipid-linked glycans (55). In several yeasts, including *S. cerevisiae*, *C. albicans*, and *P. pastoris*, glycans also contain mannosyl phosphate, and the proteins responsible for mannosyl phosphate transfer are also type II membrane proteins such as Mnn4p (44). A BLAST search using the *S. cerevisiae* Mnn4p tail as a probe resulted in the identification of a new gene encoding a putative type II membrane protein with C-terminal repeats of lysines and glutamic acids but with no homology to known genes. Subsequently, three similar genes were identified suggesting the presence of a novel gene family with four members in *P. pastoris*. The results of subsequent BLAST searches against public data bases revealed that the new family of genes is present only in a limited number of fungal species. None of the commonly studied fungal model organisms, including *S. cerevisiae* and *Aspergillus nidulans*, harbored homologs of these genes. However, we were able to identify nine homologs in *C. albicans* consistent with the prominent expression of β-Man residues by *C. albicans*, as established by serological classification (56) and subsequently by biochemical and cytological studies (8, 9, 11, 57).

Analyses of a library of deletion strains revealed that these novel genes are indeed involved in the transfer of β-Man residues in *P. pastoris*. Western blot, mannosidase digests,

and MALDI-TOF MS analyses were performed to characterize the *N*-glycans produced by the mutant strains. The analyses were performed in strains with wild-type glycans and in an *och1*Δ mutant background with only core *N*-glycans present. All the results point to the conclusion that this new family of genes is responsible for the addition of β-Man to *N*-glycans. In the wild-type strain background, the deletion of *PpBMT1* eliminated the interaction between whole cell extracts and mAb 5B2, suggesting that the PpBmt1 protein is necessary for the synthesis of a structure recognized by this mAb. *PpBMT2* deletion slightly reduced the interaction with mAb 5B2. No changes in β-Man were observed in the *PpBMT3* deletion strain. Interestingly, deletion of *PpBMT4* in the wild-type background enhanced the signal on Western blot probed with mAb 5B2. This suggests that β-mannosylation catalyzed by PpBmt4p allows further steps that mask β-Man epitopes recognized by mAb 5B2.

In the *och1*Δ strains, deletion of *PpBMT2* eliminates neutral glycans resistant to α-mannosidase treatment as shown by MALDI-TOF MS analysis. In contrast to the wild-type strain, no effect of a *PpBMT1* deletion was observed in the *och1*Δ strain background. This might indicate that PpBmt1p is responsible for β-mannose transfer only on outer chain *N*-glycans. Again, we were unable to detect any changes in *N*-glycans associated with the *PpBMT3* deletion. The deletion of *PpBMT4* in *och1*Δ strains seemed to expose the β-Man structures recognized by mAb 5B2 and even more dramatically in this case because of the complete lack of signal in the *och1*Δ mutant parental strain.

However, MALDI-TOF analysis of the *N*-glycans from the *PpBMT4* deletion strain showed a size reduction of glycans resistant to α-mannosidase treatment. This result, in agreement with Western blot analysis of whole cell extracts, confirms that PpBmt4p extends glycans containing β-Man with hexose. Future NMR studies of glycans from wild-type *P. pastoris* as well as *Ppbmt*Δ mutants will be necessary to determine the structure of β-Man containing glycans and the exact function of *PpBMT* genes, particularly *PpBMT4*, in their synthesis.

Similarly, we investigated whether the nine homologs identified in the genome of *C. albicans* were also involved in the transfer of β-Man residues. First, we created a library of *C. albicans BMT* (*CaBMT*) deletion strains. *In vitro*, under the conditions tested, none of the deletion strains displayed alterations in either growth phenotype or sensitivity to chemical or antifungal compounds. However, all of the strains displayed an altered pattern of β-Man epitopes in PPM. Based on previously established chemical structures of PPMs, we designed and conducted a series of experiments involving purification of these molecules. Western blot of PPMs with anti-β-Man antibodies of different specificity, FACE analysis of oligomannosides, and confirmation of their structure by NMR led to the following conclusions, summarized in Fig. 1*B*. CaBmt1p and CaBmt2p are responsible for the addition of the first β-mannose on PPM acid-stable and acid-labile fractions, respectively, and therefore act on α-linked mannose residues as acceptors. CaBmt3p and CaBmt4p are active on β-mannoses as acceptors, because they are involved in the elongation of β-Man chains of the PPM acid-labile fraction. Analysis of the *CaBMT3* deletion also suggested its involvement in the addition of the second β-mannose to the PPM acid-stable domain. None of these deletions had an impact on PLM β-mannosylation demonstrating the substrate specificity of Bmtps. Finally, flow cytometry analysis revealed that modifications of PPM β-mannosylation had a limited effect on β-Man cell wall surface expression.

A number of studies have implicated a role for mannan in *C. albicans* biology and pathogenicity. The importance of mannan in virulence has been demonstrated by inactivation of genes involved in transfer of α-mannosyls on *O*-glycans (58) or the *N*-glycan outer chain (19). In parallel, experimental studies either *in vitro* (17, 53) or *in vivo* (13, 18), as well as clinical studies involving purified (14, 53) or synthetic β-mannosides (18), or purified molecules selectively containing these residues (16, 17), have shown that β-Mans also contribute to *C. albicans* virulence. Moreover, a separate study has revealed that the spatial conformation of β-Man (12) differs from the more ubiquitous α-Mans, most of which are shared with *S. cerevisiae*. These differences affect β-Man recognition by the adaptative (13) and innate immune system (15, 59). Furthermore, an interesting observation is that members of the *BMT* gene family are present in only a limited number of known fungal species, but among these is the prominent fungal pathogen *A. fumigatus* (60). Despite these findings, recently published papers have contradicted this notion and led to the conclusion that β-Man may actually have little effect on *C. albicans* virulence (20, 21). However, the strategy followed in these papers did not specifically target β-Man, but rather was based on the inactivation of *C. albicans* genes, homologous to *S. cerevisiae* genes, which prevented downstream association of β-Man to PPM (20, 21).

The results presented here demonstrate a multiplicity and specificity of Bmtps for their substrates, as well as a relatively limited impact of individual deletions on overall β-Man expression. It is striking that deletion of *CaBMT2* leading to the complete absence of PPM acid-labile fraction (identified as antigenic factor 5 (61)) had no impact on direct agglutination by the same factor. Even the more analytic flow cytometry method did not find any difference between the mutant and parental strain using factor 5. This implies that the complexity of the process of β-mannosylation in *C. albicans* has been underestimated.

This study was designed as a principal study to define the functions of a new gene family. As such it focused on *C. albicans* PPM as a reference molecule whose structure elucidation has required many studies. Besides definition of Bmt1–4p functions, it was shown that disruption of these *BMTs* acting on PPM had no impact on PLM β-mannosylation. Further studies will focus on this using *CaBMTs*.

Considering this specificity of Bmtps in early biosynthetic steps, pleiotropic effects of *OCH1* deletion are probably also of importance to be considered. The unknown mechanisms of *N*-mannan core elaboration in the *C. albicans och1*Δ mutant noticed by Bates *et al.* (19) present striking homologies with what was observed on *P. pastoris och1*Δ regarding β-Mans.

However, for mature mannoglycoconjugates, considering previous immunological studies on β-Man expression in the *C. albicans* cell wall, and the present results obtained by flow cytometry, the hidden part of the iceberg is probably represented by the numerous cell wall mannoproteins harboring altogether massive amounts of β-Man epitopes. As far as the function of a cell wall molecule and its recognition by the host are dependent on glycosylation, it can be anticipated that a role for β-Mans in virulence may yet be revealed.

With this is mind, the identification of genes responsible for β-mannose transfer creates the opportunity to directly study the β-Man biosynthetic pathway in *C. albicans* and its relation to pathogenicity. More generally detailed analysis of *bmt*Δ mutants in different living organisms should allow a more complete understanding of the extent and function of this post-translational modification and how to manage it for engineering therapeutics and industrial purposes.

Acknowledgments—We gratefully acknowledge Pr. J. Cutler (Research Institute for Children, New Orleans, LA), Pr. J. Pontón (Universidad del País Vasco, Bilbao, Spain), and Pr. A. J. P. Brown (Department of Molecular and Cell Biology, Aberdeen, UK) for providing mAbs B6.1, B9E, and pCIp10, respectively. We are indebted to Annick Masset and Beata Bobrowicz for excellent technical assistance and to Nadine François for the morphogenesis studies. We also thank Dr. T. U. Gerngross for continuous support and Drs. T. Jouault, B. Sendid, and T. A. Stadheim for discussion and help in preparing the manuscript.

REFERENCES

1. Sims, C. R., Ostrosky-Zeichner, L., and Rex, J. H. (2005) *Arch. Med. Res.* **36,** 660–671
2. Calderone, R. A., and Fonzi, W. A. (2001) *Trends Microbiol.* **9,** 327–335
3. Dean, N. (1999) *Biochim. Biophys. Acta* **1426,** 309–322
4. Strahl-Bolsinger, S., Gentzsch, M., and Tanner, W. (1999) *Biochim. Biophys. Acta* **1426,** 297–307
5. Helenius, A., and Aebi, M. (2001) *Science* **291,** 2364–2369
6. Jigami, Y., and Odani, T. (1999) *Biochim. Biophys. Acta* **1426,** 335–345
7. Shibata, N., Ichikawa, T., Tojo, M., Takahashi, M., Ito, N., Okubo, Y., and Suzuki, S. (1985) *Arch. Biochem. Biophys.* **243,** 338–348
8. Kobayashi, H., Shibata, N., and Suzuki, S. (1992) *Infect. Immun.* **60,** 2106–2109
9. Trinel, P. A., Maes, E., Zanetta, J. P., Delplace, F., Coddeville, B., Jouault, T., Strecker, G., and Poulain, D. (2002) *J. Biol. Chem.* **277,** 37260–37271
10. Mille, C., Janbon, G., Delplace, F., Ibata-Ombetta, S., Gaillardin, C., Strecker, G., Jouault, T., Trinel, P. A., and Poulain, D. (2004) *J. Biol. Chem.* **279,** 47952–47960
11. Trinel, P. A., Faille, C., Jacquinot, P. M., Cailliez, J. C., and Poulain, D. (1992) *Infect. Immun.* **60,** 3845–3851
12. Nitz, M., Ling, C. C., Otter, A., Cutler, J. E., and Bundle, D. R. (2002) *J. Biol. Chem.* **277,** 3440–3446
13. Han, Y., Kanbe, T., Cherniak, R., and Cutler, J. E. (1997) *Infect. Immun.* **65,** 4100–4107
14. Fradin, C., Poulain, D., and Jouault, T. (2000) *Infect. Immun.* **68,** 4391–4398
15. Kohatsu, L., Hsu, D. K., Jegalian, A. G., Liu, F. T., and Baum, L. G. (2006) *J. Immunol.* **177,** 4718–4726
16. Jouault, T., Ibata-Ombetta, S., Takeuchi, O., Trinel, P. A., Sacchetti, P., Lefebvre, P., Akira, S., and Poulain, D. (2003) *J. Infect. Dis.* **188,** 165–172
17. Ibata-Ombetta, S., Idziorek, T., Trinel, P. A., Poulain, D., and Jouault, T. (2003) *J. Biol. Chem.* **278,** 13086–13093
18. Dromer, F., Chevalier, R., Sendid, B., Improvisi, L., Jouault, T., Robert, R., Mallet, J. M., and Poulain, D. (2002) *Antimicrob. Agents Chemother.* **46,** 3869–3876
19. Bates, S., Hughes, H. B., Munro, C. A., Thomas, W. P., MacCallum, D. M., Bertram, G., Atrih, A., Ferguson, M. A., Brown, A. J., Odds, F. C., and Gow, N. A. (2006) *J. Biol. Chem.* **281,** 90–98
20. Hobson, R. P., Munro, C. A., Bates, S., MacCallum, D. M., Cutler, J. E., Heinsbroek, S. E., Brown, G. D., Odds, F. C., and Gow, N. A. (2004) *J. Biol. Chem.* **279,** 39628–39635
21. Singleton, D. R., Masuoka, J., and Hazen, K. C. (2005) *Eukaryot. Cell* **4,** 639–648
22. Lindberg, B., Leontein, K., Lindquist, U., Svenson, S. B., Wrangsell, G., Dell, A., and Rogers, M. (1988) *Carbohydr. Res.* **174,** 313–322
23. Ralton, J. E., Naderer, T., Piraino, H. L., Bashtannyk, T. A., Callaghan, J. M., and McConville, M. J. (2003) *J. Biol. Chem.* **278,** 40757–40763
24. Sernee, M. F., Ralton, J. E., Dinev, Z., Khairallah, G. N., O'Hair, R. A., Williams, S. J., and McConville, M. J. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 9458–9463
25. Vinogradov, E., Petersen, B. O., and Duus, J. O. (2000) *Carbohydr. Res.* **325,** 216–221
26. Choi, B. K., Bobrowicz, P., Davidson, R. C., Hamilton, S. R., Kung, D. H., Li, H., Miele, R. G., Nett, J. H., Wildt, S., and Gerngross, T. U. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 5022–5027
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *in Molecular Cloning: A Laboratory Manual* (Ford, N. N., and Ferguson, M., eds) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
28. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989) *Gene* (*Amst.*) **77,** 61–68
29. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* (*Amst.*) **77,** 51–59
30. Davidson, R. C., Nett, J. H., Renfer, E., Li, H., Stadheim, T. A., Miller, B. J., Miele, R. G., Hamilton, S. R., Choi, B. K., Mitchell, T. I., and Wildt, S. (2004) *Glycobiology* **14,** 399–407
31. Goldstein, A. L., and McCusker, J. H. (1999) *Yeast* **15,** 1541–1553
32. Goldstein, A. L., Pan, X., and McCusker, J. H. (1999) *Yeast* **15,** 507–511
33. Gola, S., Martin, R., Walther, A., Dunkler, A., and Wendland, J. (2003) *Yeast* **20,** 1339–1347
34. Sanglard, D., Ischer, F., Monod, M., and Bille, J. (1996) *Antimicrob. Agents Chemother.* **40,** 2300–2305
35. Murad, A. M., Lee, P. R., Broadbent, I. D., Barelle, C. J., and Brown, A. J. (2000) *Yeast* **16,** 325–327
36. Papac, D. I., Briggs, J. B., Chin, E. T., and Jones, A. J. (1998) *Glycobiology* **8,** 445–454
37. Faille, C., Wieruszeski, J. M., Michalski, J. C., Poulain, D., and Strecker, G. (1992) *Carbohydr. Res.* **236,** 17–27
38. Dubois, M., Gilles, K., Hamilton, J. K., Rebers, P. A., and Smith, F. (1951) *Nature* **168,** 167
39. Ponton, J., Marot-Leblond, A., Ezkurra, P. A., Barturen, B., Robert, R., and Senet, J. M. (1993) *Infect. Immun.* **61,** 4842–4847
40. Goins, T. L., and Cutler, J. E. (2000) *J. Clin. Microbiol.* **38,** 2862–2869
41. Shibata, N., Hisamichi, K., Kikuchi, T., Kobayashi, H., Okawa, Y., and Suzuki, S. (1992) *Biochemistry* **31,** 5680–5686
42. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* **63,** 207–213
43. Martinez-Esparza, M., Sarazin, A., Jouy, N., Poulain, D., and Jouault, T. (2006) *J. Immunol. Methods* **314,** 90–102
44. Odani, T., Shimma, Y., Tanaka, A., and Jigami, Y. (1996) *Glycobiology* **6,** 805–810
45. Edmond, M. B., Wallace, S. E., McClish, D. K., Pfaller, M. A., Jones, R. N., and Wenzel, R. P. (1999) *Clin. Infect. Dis.* **29,** 239–244
46. Gale, C. A., Bendel, C. M., McClellan, M., Hauser, M., Becker, J. M., Berman, J., and Hostetter, M. K. (1998) *Science* **279,** 1355–1358
47. Sundstrom, P. (1999) *Curr. Opin. Microbiol.* **2,** 353–357
48. Hoyer, L. L. (2001) *Trends Microbiol.* **9,** 176–180
49. Naglik, J., Albrecht, A., Bader, O., and Hube, B. (2004) *Cell. Microbiol.* **6,** 915–926
50. Whiteway, M. (2000) *Curr. Opin. Microbiol.* **3,** 582–588
51. Enjalbert, B., Smith, D. A., Cornell, M. J., Alam, I., Nicholls, S., Brown, A. J., and Quinn, J. (2006) *Mol. Biol. Cell* **17,** 1018–1032
52. Lorenz, M. C., and Fink, G. R. (2001) *Nature* **412,** 83–86
53. Jouault, T., Fradin, C., Trinel, P. A., and Poulain, D. (2000) *Infect. Immun.* **68,** 965–968

54. Wildt, S., and Gerngross, T. U. (2005) *Nat. Rev. Microbiol.* **3,** 119–128
55. Oriol, R., Martinez-Duncker, I., Chantret, I., Mollicone, R., and Codogno, P. (2002) *Mol. Biol. Evol.* **19,** 1451–1463
56. Suzuki, S. (1995) *Yakugaku Zasshi* **115,** 280–294
57. Molinari, A., Gomez, M. J., Crateri, P., Torosantucci, A., Cassone, A., and Arancia, G. (1993) *Eur. J. Cell Biol.* **60,** 146–153
58. Rouabhia, M., Schaller, M., Corbucci, C., Vecchiarelli, A., Prill, S. K., Giasson, L., and Ernst, J. F. (2005) *Infect. Immun.* **73,** 4571–4580
59. Poulain, D., and Jouault, T. (2004) *Curr. Opin. Microbiol.* **7,** 342–349
60. Latge, J. P. (1999) *Clin. Microbiol. Rev.* **12,** 310–350
61. Shibata, N., Arai, M., Haga, E., Kikuchi, T., Najima, M., Satoh, T., Kobayashi, H., and Suzuki, S. (1992) *Infect. Immun.* **60,** 4100–4110
62. Wilson, R. B., Davis, D., and Mitchell, A. P. (1999) *J. Bacteriol.* **181,** 1868–1874

Molecular phenotyping of β-mannosyltransferases deficient *Candida albicans* cells by High Resolution-Magic Angle Spinning NMR

Emmanuel Maes[1], Céline Mille[2], Daniel Poulain[2] and Yann Guérardel[1*]

[1]Unité de Glycobiologie Structurale et Fonctionnelle, UMR CNRS 8576, IFR 147, Université des Sciences et Technologies de Lille 1, 59655 Villeneuve d'Ascq, France.

[2]Unité de Physiopathologie des Candidoses, Institut National de la Santé et de la Recherche Médicale (Inserm) U799, Université Lille 2, 59045, Lille, France.

* To whom correspondence should be addressed
e. mail: yann.guerardel@univ-lille1.fr;

tel : +33(0)320336347

Fax : +33(0)320436555

Key words: *Candida albicans*, High Resolution Magic Angle NMR, β-mannosylation, mannosyltransferase

Running title: HR-MAS analysis of *Candida albicans*

Abbreviations: HR-MAS NMR, High Resolution-Magic Angle Spinning Nuclear Magnetic Resonance; Man, Mannose; Man*p*, mannopyranoside, Glc*p*NAc, N-AcetylGlucopyranoside, PPM, PhosphoPeptidoMannan; HSQC, Heteronuclear Single Quantum Coherence

Abstract

The yeast *Candida albicans* is an opportunistic pathogen, part of the normal human microbial flora that causes infections in immunocompromised individuals with a high morbidity and mortality levels. Recognition of yeasts by host cells is directly mediated by cell wall components of the yeast, including a wide range of abundantly expressed glycoconjugates. Of particular interest in *C. albicans* are the β-mannosylated epitopes that show a complex expression pattern on *N*-glycan moiety of phosphopeptidomannans and are absent in the non-pathogenic species *Saccharomyces cerevisiae*. Being known as potent antigens for the adaptive immune response and elicitors of specific infection-protective antibodies, the exact delineation of β-mannosides regulation and expression pathways has lately become a major milestone toward the comprehension of host-pathogen interplay. Using the newly developed HR-MAS NMR methodology, we demonstrate here the possibility of assessing the general profiles of cell surface exposed glycosylates from intact living yeast cells without any prior, potentially destructive, purification step. This technique permitted to directly observe structural modifications of surface expressed phosphodiester-linked β-mannosides on a series of deletion strains in β-mannosyltransferases and phospho-mannosyltransferases compared with their parental strains.

Introduction

The cell surfaces of yeast are covered with a dense glyco-shield whose biogenesis is tightly controlled by environmental stimuli. Definition of structure to activity relationships classically requires selective extraction procedures of large quantities of material to be compatible with physicochemical analyses. The application of these procedures in order to unravel basic *N*- and *O*-mannosylation processes by using different temperature sensitive mutants of the eukaryotic yeast model *Saccharomyces cerevisiae* took several decades to achieve a workable structural model partially applicable to other yeast model species (1). Based on sequence homologies of mannosyltransferases and the deciphering of its mannan structure, the pathogenic yeast *Candida albicans* became a popular model in both biochemistry and immunology considering that oligomannose sequences act as adhesins and may even coordinate innate and adaptative immunity responses (2, 3). However, progresses in genetics are still impaired by the length, and to a lesser extend by the lack of standardization, of cell wall molecules preparation and analysis procedures. Moreover application of harsh and selective extraction methods may give a poor physicochemical view of the cell wall molecules actually accessible to the host tissues and immune system.

We therefore adapted the newly developed HR-MAS technology to the analysis of yeast cell surface as a mean for connecting genetic and immunologic background knowledge to structure and ultra-structure of *C. albicans* cell wall. We took advantage of our recent involvement in the definition of *C. albicans* mannan phenotypes resulting from selective disruption of 4 members of genes encoding for a mannosyltransferase family (4). This family, recently discovered, comprises 9 genes which transfer β-Man residues to mannan moiety of several cell wall glycoconjugates, including the phosphopeptidomannans (PPM). Indeed, β-oligomannosides are one of the most salient features of *C. albicans* cell wall glycoconjugates, compared with non the pathogenic species *S. cerevisiae* (5). They have been first identified by Suzuki and co-workers and then extensively characterized both for their distribution on different *Candida* molecules and their contribution to *C. albicans* virulence (6, 7). We therefore used this model of β-mannosylation previously dissected with conventional methods, to demonstrate the usefulness of HR-MAS technology as a new, powerful tool to directly assess the molecular phenotypes of yeast glyco-shield.

Materials and Methods

Strains and Culture Conditions

The *Candida albicans* strains used in this study are listed in Table 2. Strains were grown in YPD medium (1% yeast extract, 2% peptone, 2% dextrose) at 37°C for 16h. Then the cultures were harvested before HR-MAS NMR experiments.

Phosphopeptidomannan Extraction

Phosphopeptidomannan from cells grown in YPD medium was extracted as described previously (8). Briefly, cell pellets were suspended in 20 mM citrate buffer and autoclaved at 125°C for 90 min. Suspensions were harvested and Fehling's solution was added to the supernatant to precipitate PPM. The PPM was then washed in methanol:acetic acid (8:1) and dried in a Speed Vac concentrator. Sugar concentrations were estimated by the sulfuric phenol colorimetric method (9).

NMR experiments

Liquid NMR experiments were performed on a 9.4 T Avance Bruker® spectrometer where ^{1}H, ^{13}C and ^{31}P resonate at 400.33, 100.67 and 162.06 MHz respectively. The experiments were acquired with a Broad Band Inverse self-shielded *z*-gradient probehead. Spectra were recorded at 318K in D_2O after two chemical exchanges with $^{2}H_2O$ (Euriso-top, Gif-sur-Yvette, France).

For ^{1}H-^{13}C HSQC spectra, the spectral widths were 4006 Hz for ^{1}H with 4k acquisition data points and 12080 Hz for ^{13}C during 256 scans or transients and for ^{1}H-^{31}P HSQC spectra the spectral widths were 4006 Hz with 4k acquisition data points and 32411 Hz for ^{1}H and ^{31}P during 128 scans or transients respectively. Durations and power levels were optimized for each experiment. The other experiments (i.e., 1D spectra ^{1}H, ^{31}P and 2D homonuclear spectra) were acquired using classical pulse programs taken in the Bruker® pulse program library.

Spectra were recorded without sample spinning. The chemical shifts were expressed in ppm downfield from the signals of internal acetone, ^{1}H 2.225 ppm ^{13}C 31.55 ppm. HR-MAS NMR experiments were achieved on an 18.8 T Avance II Bruker® spectrometer where ^{1}H, ^{13}C and ^{31}P resonate at 800.13, 201.2 and 323.9 MHz respectively. The experiments were acquired with a $^{1}H/^{13}C/^{31}P/^{2}H$ probe with uniaxial gradients. Before analysis cell-pellets were twice washed with deuterium oxide (Euriso-top, Gif-sur-Yvette, France) in order to remove all exchangeable protons. Four mm ZrO_2 rotor (CortecNet, Paris, France) was filled with 50µL of cell pellets including 0.5 µL of acetone as internal standard and finally centrifuged at 3000 rpm. All spectra were recorded at 293

K, rotor spinning rate was 8 kHz. All experiments came from Bruker library pulse program, delays and powers are optimized for each. For ^{1}H-^{13}C HSQC, the spectral widths were 12820 Hz (^{1}H) with 1024 points for fid resolution and 29994 Hz (^{13}C) during 400 scans or transients giving 12.5 Hz/pt and 75.0 Hz/pt respectively.

Results and Discussion

Detailed structural analysis of mannan fragments generated by acetolysis and mild hydrolysis previously established that the polymannosylated mannan core of *C. albicans* is an extremely heterogeneous molecule composed by a linear α-1,6 mannan substituted by a wide range of α-1,2-, α-1,3-linked mannose residues containing oligomannans and β-1,2-linked mannose residues in serotype A. Furthermore, in both serotypes A and B, the mannan core is substituted by a phosphodiester linked α- and β-Man residues containing oligomannans commonly referred as acid-labile domain (Fig. 1). Structure of mannan moieties of phosphopeptidomannan (PPM) from a wide range of yeast species, including serotypes A and B of *Candida albicans* have already been extensively studied by multidimensional NMR experiments (10, 11, 12, 13). Shear heterogeneity of polysaccharide structure generates a wide variability in the ^{13}C and ^{1}H NMR parameters of individual residues depending on their exact location within the polysaccharide chain because of the known susceptibility of chemical shifts to direct electronic environment. Whereas the micro-heterogeneity of NMR parameters is dealt with when analysing purified oligosaccharidic sequences, it may be a problem when trying to assess the general glycosylation pattern of intact mannan because of signal overlaps which impair its exhaustive sequencing. Still, based on NMR parameters of linked mannose residues established from analysis of acetolysis and hydrolysis fragments (10, 13, 4) the presence of β-Man residues was readily assessed by ^{1}H-^{13}C NMR-HSQC spectroscopy owing to the clear observation in the anomeric region of five signals at 4.79/99.9, 4.85/100.1, 4.86/102.3, 4.94/102.4 and 5.04/102.5 ppm (Fig. 2A). A sixth discrete β-Man signal was also observed at 4.90/100.1 ppm (signal **f**) that will be discussed later. β-anomeric configuration of all residues was unambiguously demonstrated by non-decoupled ^{1}H-^{13}C HSQC heteronuclear analysis owing to $^{1}J_{CH}$ values of about 160 Hz (data not shown). Although all five main signals are presumably the sum of numerous sub-signals from individual residues H-1 in slightly different environments, they each are distinct enough to be associated with a clear substitution pattern. So, in accordance with published data, the five signals were respectively labelled **a** to **e** and been assigned to constant locations of β-

Man residues within non phosphorylated β-oligomannosylated sequences of different sizes (Table 1A). Accordingly, all five signals are also observed on the ^{1}H-^{13}C NMR-HSQC spectrum of the total acid resistant moiety of PPM that results form the mild acid hydrolysis of total serotype A mannan and release of the phosphodiesterified β-oligomannosides (Fig. 2B). Then, recently Shibata and collaborators proposed that phosphodiester link influenced the H-1 chemical shifts of the first β-Man residues of acid labile oligomannosidic chains by a constant value of Δδ +0.08 (13) and thus deshielded the H-1 values of Manβ1-2Manα1-phosphate and [Manβ1-2]$_n$Manβ1-2Manα1-phosphate (residues B) to 4.83-4.86 and 4.89-4.90 ppm respectively, but left the H-1 parameters of other β-Man residues essentially unchanged. This assumption was confirmed by the data of the present study which also showed a slight modification of the C-1 chemical shift (Δδ +0.3) of Manβ1-2Manα1-phosphate residue (Table 1B). Thus, it appears that H-1/C-1 parameters of Manβ1-2Manα1-phosphate switch from signal **a** to signal **b** and thus cannot be easily distinguished from Manβ1-2Manβ1-2Manα1-2Manα by liquid NMR (Table 1B). H-1/C-1 signal of [Manβ1-2]$_n$Manβ1-2Manα1-phosphate can be observed on ^{1}H-^{13}C NMR-HSQC spectrum at 4.90/100.1 ppm (signal **f**) but partly overlaps with intense Manα1-6Manα- H-1 signal and thus is also difficult to observe (Fig. 2B). Then, H-1/C-1 parameters of other β-Man residues associated to phosphodiesterified β-oligomannosides cannot be differentiated from those associated to acid-stable moiety. In summary, although liquid ^{1}H-^{13}C HSQC NMR analysis of intact mannan isolated from PPM permits to easily observe β-mannosylated epitopes owing to their very distinct H-1/C-1 parameters compared with those of α-mannosylated epitopes, it does not permit to distinguish the β-Man residues signals associated to acid-stable from those associated to acid-labile fraction due to a series of signal overlaps.

As a step forward to study the expression of β-mannosylated epitopes at the surface of intact yeast cells by NMR, we assessed the usefulness of High Resolution-Magic Angle Spinning (HR-MAS) NMR by analysing a wide range of wild type and *Candida albicans* deletion strains. This recently developed technology, which enables the acquisition of NMR spectra directly from living intact cells, was previously shown to be effective to observe and analyse the structure of complex molecules present either at the cell surface or inside the cells as long as they are mobile enough and in sufficient quantity. In particular, it was effectively used to analyse the structure of cell wall components, including mycobacterial polysaccharides (14) and complex lipids (15), lipooligosaccharides from *Campylobacter* cells (16), as well as intracellular metabolites (17) or periplasmic glucans (18). First, we analysed a solution of intact PPM purified from serotype A *C. albicans* (BWP17 strain) by HR-MAS NMR in order to standardize experimental procedure. ^{1}H-^{13}C

HSQC experiments showed that HR-MAS NMR exhibited a comparable resolution to liquid NMR which permitted to readily discriminate all β-Man associated signals (Fig. Sup 1). Furthermore, all α- and β-Man residues on purified PPM analysed by HR-MAS NMR exhibited identical parameters than those observed by liquid NMR, which validated the use of this technology to analyse PPM structure based on previously established chemical shifts. Then, in order to analyse glycan phenotype of whole cells, spectra were acquired on 50 µL of intact cells from different strains. As shown on Fig. 3A, anomeric part of ^{1}H-^{13}C HSQC HR-MAS NMR spectrum of living serotype A *C. albicans* (BWP17 strain) cells shows most of the anomeric signals of the mannose residues associated to PPM, including both α- and β-Man type residues. All major α-Man residues signals that have been identified on purified PPM are clearly observed on ^{1}H-^{13}C HSQC HR-MAS NMR spectrum with the noteworthy exception of Manβ1-2Manα1-2 Manα at δ 5.16/101.3 (^{1}H/^{13}C) that completely disappeared. Furthermore, anomeric signal for Manα1-2 residue at δ 5.11/99.6 (^{1}H/^{13}C) was barely visible on intact cell compared with purified PPM. Then, very intense signals at δ 5.45/97.4 and 5.56/95.4 (^{1}H/^{13}C) assigned to Manα1-phosphate and [Manβ1-2]$_{n}$Manα1-phosphate of the PPM acid-labile fraction were also clearly observed. The nature of α anomeric configuration was further confirmed by observation of their $^{1}J_{H\text{-}C}$ coupling constants at 171 Hz by non-decoupled ^{1}H-^{13}C HSQC and linkage to phosphorus atom also confirmed by ^{1}H-^{31}P HSQC experiments (data not shown). Surprisingly, a third signal in the region of phosphorylated α-Man residues was identified at δ 5.56/96.0 ($^{1}J_{H\text{-}C}$ 164 Hz) exclusively on intact cells by HR-MAS-NMR whereas it never appeared on purified PPM analysed either by liquid NMR or by HR-MAS NMR, which strongly suggest it is not associated to the purified PPM but to another cell wall related molecule. Of particular interest for the present study, are signals previously associated to β-Man residues that were easily identified, with the notable exception of **a**, at identical chemical shifts than those observed for purified PPM analysed by liquid and HR-MAS NMR. Signal **a** (Manβ1-2Manα1-2Manα-) being the only β-Man signal exclusively assigned to PPM acid-stable fraction, its specific and reproducible absence on whole cell analyses suggests that β-Man residues associated to acid-stable moiety is somehow not detected by HR-MAS NMR analysis of intact *C. albicans* and thus that only phosphodiesterified β-oligomannosides are observed. The transparency of cell wall associated acid-stable β-mannosides to HR-MAS NMR analyses was demonstrated by comparing data obtained from a serotype A (*Ca*A-*mnn4Δ* strain) and a serotype B (*Ca*B-*mnn4Δ* strain) both devoid of acid-labile mannoside domain. Indeed, whereas *Ca*A-*mnn4Δ* strain differs from *Ca*B-*mnn4Δ* strain by the presence of acid-stable β-mannosides, they both show identical HSQC spectra,

characterized by the total absence of β-Man signals **a-e** associated to acid-stable moiety (Fig. 3B and 3C). This is most probably the result of the specific feature of HR-MAS NMR requiring that the investigated samples possess some degrees of mobility. Indeed, constrain molecules exhibiting low transversal relaxation time (T2) will not be detected. In agreement with the absence of acid-labile mannoside domain, no signal associated to Manα1-phosphate and (Manβ1-2)$_n$Manα1-phosphate at δ 5.45/97.4 and 5.56/95.4 can be observed in *Ca*B-*mnn4Δ* and *Ca*A-*mnn4Δ* strains. In summary, absence of detection of acid-stable β-mannosides *in vivo*, whereas they are easily observed in purified PPM, gives a clear indication that these epitopes have a rigid conformation at the cell surface compared to the phosphodiesterified β-oligomannosides that are very mobile. Altogether, data established that HR-MAS NMR enables the *in situ* observation of cell wall PPM directly from intact living yeasts without prior purification. Also, the difference in respective mobility of both fractions gives the opportunity to specifically target the phosphodiesterified β-oligomannosides at the cell surface without interference from acid-stable β-mannosides.

So, HR-MAS NMR analysis of intact yeasts was used to determine the *in vivo* phenotypes of β-mannosyltansferase mutants *Ca*A-*bmt2Δ, Ca*A-*bmt3Δ* and *Ca*A-*bmt4Δ* by assessing the structure of the phosphodiesterified β-oligomannosides expressed at their cell surfaces. Detailed structural analysis of the PPMs purified from these mutants by a combination of mass spectrometry, fluorophore-assisted carbohydrate electrophoresis (FACE) and liquid NMR strongly suggested that the CaBmt2, CaBmt3 and CaBmt4 proteins were involved in the sequential transfer of the first, the second and the third β-Man residues, respectively, onto the acid-labile mannoside of *C. albicans* PPM (Fig. 1) (4). However, the possibility that methodology based on multiple purification and chemical degradation steps generate artefacts or overlook eventual alternative substrates can never be completely ruled out, which justifies the use of a non-destructive type of analysis on un-processed samples. Anomeric region of the ^{1}H-^{13}C HSQC HR-MAS NMR spectrum of intact *Ca*A-*bmt2Δ* cells shows most α-Man signals previously identified on parental BWP17 strain as well as on *Ca*B-*mnn4Δ* and *Ca*A-*mnn4Δ* strains, but none of the β-Man signals (Fig. 3D). However, *Ca*A-*bmt2Δ* differs from *Ca*B-*mnn4Δ* and *Ca*A-*mnn4Δ* strains by the observation of Manα1-phosphate H-1 signal at δ 5.45/97.4 in agreement with the presence of a PPM acid-labile domain, the Manβ1-2Manα1-phosphate H-1 signal still being absent. Altogether, these data unambiguously demonstrates that *Ca*A-*bmt2Δ* expresses at its cells surface PPMs substituted by phosphodiester side chains that are constituted by a single, non substituted, α-Man residue. Compared to *Ca*A-*bmt2Δ* , ^{1}H-^{13}C HSQC NMR spectrum of *Ca*A-*bmt3Δ* showed a single additional β-Man residue H-

1 signal (**b**: Manβ1-2Manα1-phosphate) as well as a single additional α-Man residue H-1 signal at δ 5.56/95.4 (Manβ1-2Manα1-phosphate) (Fig. 3E). The assignment of this last signal as an internal 2-substituted Man was confirmed by the deshielding of its C-2 value to 78.8 ppm (data not shown). Then, spectrum of *Ca*A-*bmt4Δ* was characterised by the simultaneous presence of β-Man associated signals **b**, **c** and **f** (Fig. 3F). In association with the observation of both terminal and internal Manα1-phosphate H-1 signals and in accordance with previously established NMR parameters (Table 1), these signals established the presence of Manα1-phosphate, Manβ1-2Manα1-phosphate and Manβ1-2Manβ1-2Manα1-phosphate side chains at the surface of intact *Ca*A-*bmt4Δ* cells.

Altogether, presented data established the validity of HR-MAS NMR technology for analysing glycosylation phenotypes of intact yeast cells. We have shown a perfect correlation between *in vitro* and *in vivo* analyses of β-mannoside structures by using a number of *C. albicans* serotypes strains as well as strains deficient in extremely specific mannosyltransferases. *In vivo* analysis has the enormous advantage over classical approach of combining low quantities requirements (50 µL of packed cells) and no need for any purification procedure, making it useful for quick, although structurally informative, phenotyping of mutant strains. A possible drawback of this approach is the selective signal suppression phenomenon associated to low mobility. However, in the case of *C. albicans*, we took advantage of it to specifically target mobile molecules by getting rid of overlapping non informative signals. We do believe that HR-MAS technology has an enormous potential for analysing the regulation of expression a wide range of yeast parietal molecules and could be extremely useful to the yeast community.

Acknowledgements

We thank Xavier Trivelli for technical help for NMR experiments. The 800 MHz spectrometer was funded by European Union (FEDER), Ministère Français de l'Enseignement Supérieur et de la Recherche, région Nord-Pas de Calais, Université des Sciences et Technologies de Lille, Institut Fédératif de Recherche 147 and Centre National de la Recherche Scientifique.

References

1. Ballou CE. 1990 Isolation, characterization, and properties of *Saccharomyces cerevisiae* mnn mutants with nonconditional protein glycosylation defects. Methods Enzymol. **185**:440-70.

2. Jouault T, El Abed-El Behi M, Martínez-Esparza M, Breuilh L, Trinel PA, Chamaillard M, Trottein F, Poulain D. 2006 Specific recognition of *Candida albicans* by macrophages requires galectin-3 to discriminate *Saccharomyces cerevisiae* and needs association with TLR2 for signaling. J Immunol. **177**:4679-4687.

3. Poulain D. and Jouault T. 2004 *Candida albicans* cell wall glycans, host receptors and responses: elements for a decisive crosstalk. Curr Opin Microbiol. **7**:342-349.

4. Mille, C., Bobrowicz, P., Trinel, P.A., Li, H., Maes, E., Guerardel, Y., Fradin, C., Martínez-Esparza, M., Davidson, R.C., Janbon, G., Poulain, D., and Wildt, S. 2008 Identification of a new family of genes involved in beta-1,2 mannosylation of glycans in *Pichia pastoris* and *Candida albicans*. J Biol Chem. **283**:9724-9236

5. Martínez-Esparza M, Sarazin A, Jouy N, Poulain D, and Jouault T. 2006 Comparative analysis of cell wall surface glycan expression in Candida albicans and Saccharomyces cerevisiae yeasts by flow cytometry. J Immunol Methods. **314**:90-102.

6. Shibata, N., Ichikawa, T., Tojo, M., Takahashi, M., Ito, N., Okubo, Y., and Suzuki, S. 1985 Immunochemical study on the mannans of *Candida albicans* NIH A-207, NIH B-792, and J-1012 strains prepared by fractional precipitation with cetyltrimethylammonium bromide. Arch. Biochem. Biophys. **243**:338-348

7. Trinel PA, Maes E, Zanetta JP, Delplace F, Coddeville B, Jouault T, Strecker G, and Poulain D. 2002 *Candida albicans* phospholipomannan, a new member of the fungal mannose inositol phosphoceramide family. J Biol Chem. **277**:37260-71.

8. Faille, C., Wieruszeski, J. M., Michalski, J. C., Poulain, D., and Strecker, G. 1992 Complete 1H- and 13C-resonance assignments for D-mannooligosaccharides of the beta-D-(1-->2)-linked series released from the phosphopeptidomannan of *Candida albicans* VW.32 (serotype A). Carbohydr Res **236:**17-27.

9. Dubois, M., Gilles, K., Hamilton, J. K., Rebers, P. A., and Smith, F. 1951 A colorimetric method for the determination of sugars. Nature **168**:167.

10. Shibata, N., Hisamichi, K., Kobayashi, H., and Suzuki, S. 1993 Complete assignment of ^{1}H and ^{13}C nuclear magnetic resonance chemical shifts of beta-1,2-linked mannooligosaccharides isolated from the phosphomannan of the pathogenic yeast *Candida albicans* NIH B-792 strain. Arch Biochem Biophys. **302**:113-117.

11. Kobayashi, H., Takaku, M., Nishidate, Y., Takahashi, S., Takikawa, M., Shibata, N., and Suzuki, S. 1992 Structure of the D-mannan of the pathogenic yeast, *Candida stellatoidea* ATCC 20408 (type II) strain, in comparison with that of *C. stellatoidea* ATCC 36232 (type I) strain. Carbohydr Res. **231**:105-116.

12. Kobayashi, H., Takahashi, S., Shibata, N., Miyauchi, M., Ishida, M., Sato, J., Maeda, K., and Suzuki, S. 1994 Structural modification of cell wall mannans of *Candida albicans* serotype A strains grown in yeast extract-Sabouraud liquid medium under acidic conditions. Infect Immun. **62**:968-973.

13. Shibata, N., Suzuki, A., Kobayashi, and H., Okawa, Y. 2007 Chemical structure of the cell-wall mannan of *Candida albicans* serotype A and its difference in yeast and hyphal forms. Biochem J. **404**:365-372.

14. Lee, R.E., Li, W., Chatterjee, D., and Lee, R.E. 2005 Rapid structural characterization of the arabinogalactan and lipoarabinomannan in live mycobacterial cells using 2D and 3D HR-MAS NMR: structural changes in the arabinan due to ethambutol treatment and gene mutation are observed. *Glycobiology* **15**:139-151.

15. Alahari, A., Trivelli, X., Guérardel, Y., Dover, L.G., Besra, G.S., Sacchettini, J.C., Reynolds, R.C., Coxon, G.D., and Kremer, L. 2007 Thiacetazone, an antitubercular drug that inhibits cyclopropanation of cell wall mycolic acids in mycobacteria. PLoS ONE. **2**:e1343.

16. Szymanski, C.M., Michael, F.S., Jarrell, H.C., Li, J., Gilbert, M., Larocque, S., Vinogradov, E., and Brisson, J.R. 2003 Detection of conserved N-linked glycans and phase-variable lipooligosaccharides and capsules from campylobacter cells by mass spectrometry and high resolution magic angle spinning NMR spectroscopy. J. Biol. Chem. **278**:24509-24520.

17. Hanoulle, X., Wieruszeski, J.M., Rousselot-Pailley, P., Landrieu, I., Locht, C., Lippens, G., Baulard, A.R. 2006 Selective intracellular accumulation of the major metabolite issued from the activation of the prodrug ethionamide in mycobacteria. J Antimicrob Chemother. **58**:768-772.

18. Wieruszeski, J.M., Bohin, A., Bohin, J.P., and Lippens, G. 2001 *In vivo* detection of the cyclic osmoregulated periplasmic glucan of *Ralstonia solanacearum* by high-resolution magic angle spinning NMR. J Magn Reson. **151**, 118-123.

19. Wilson, R. B., Davis, D., and Mitchell, A. P. 1999 Rapid hypothesis testing with *Candida albicans* through gene disruption with short homology regions. J Bacteriol, **181**, 1868-1874.

20. Hobson, R. P., Munro, C. A., Bates, S., MacCallum, D. M., Cutler, J. E., Heinsbroek, S. E., Brown, G. D., Odds, F. C., and Gow, N. A. 2004 Loss of cell wall mannosylphosphate in *Candida albicans* does not influence macrophage recognition. J Biol Chem **279**:39628-39635.

21. Singleton, D. R., Masuoka, J., and Hazen, K. C. 2005 Surface hydrophobicity changes of two *Candida albicans* serotype B mnn4delta mutants. Eukaryot Cell **4**:639-648.

Table 1. Estimation of ^{1}H-^{13}C chemical shifts (δ) of β-mannosides associated residues substituting **A**, acid-stable and **B**, acid-labile domains of *C. albicans* PPM, according to 4, 10, 13 and personal data. **a** to **f** correspond to the signals actually observed on liquid ^{1}H-^{13}C HSQC NMR and ^{1}H-^{13}C HSQC HR-MAS NMR spectra (Fig. 2 and 3).

A

Sugar residues		Range of ^{1}H and ^{13}C chemical shifts									
E D C B A		E		D		C		B	
Mβ1-2 Mα1-2/3 Mα1-	^{1}H							**a**	4.76-4.79
	^{13}C								99.6-100.0
Mβ1-2 Mβ1-2 Mα1-2/3 Mα1-	^{1}H					**c**	4.80-4.86	**b**	4.83-4.86
	^{13}C						101.8-102.2		99.9-100.6
Mβ1-2 Mβ1-2 Mβ1-2 Mα1-2/3 Mα1-	^{1}H		**d**	4.89-4.94	**d**	4.89-4.94	**b**	4.83-4.86	
	^{13}C			102.2-102.4		102.2-102.4		100.1-100.6	
Mβ1-2 Mβ1-2 Mβ1-2 Mβ1-2 Mα1-2/3 Mα1-	^{1}H	**d**	4.89-4.94	**e**	4.96-5.04	**d**	4.89-4.94	**b**	4.83-4.86
	^{13}C		102.2-102.4		102.4-102.6		102.2-102.4		100.2-100.6

B

Sugar residues		Range of ^{1}H and ^{13}C chemical shifts				
E D C B A		D		C		B		A
Mα1-phosphate	^{1}H							5.45
	^{13}C							97.4
Mβ1-2 Mα1-phosphate	^{1}H				**b**	4.83-4.86		5.56
	^{13}C					99.9-100.6		95.4
Mβ1-2 Mβ1-2 Mα1-phosphate	^{1}H		**c**	4.80-4.86	**f**	4.89-4.90		5.56
	^{13}C			101.8-102.2		100.1		95.4
Mβ1-2 Mβ1-2 Mβ1-2 Mα1-phosphate	^{1}H	**d**	4.89-4.94	**d**	4.89-4.94	**f**	4.89-4.90		5.56
	^{13}C		102.2-102.4		102.2-102.4		100.1		95.4

Table 2. Strains used in this study.

Strain	Genotype	References
BWP17	Serotype A strain, controle strain	Wilson *et al*, 1999 (19)
*Ca*A-*mnn4Δ*	Serotype A strain, *mnn4Δ*/*mnn4Δ*	Hobson *et al*, 2004 (20)
*Ca*B-*mnn4Δ*	Serotype B strain, *mnn4Δ*/*mnn4Δ*	Singleton *et al*, 2005 (21)
*Ca*A-*bmt2Δ*	Serotype A strain, *bmt2Δ*/*bmt2Δ*	Mille *et al*, 2008 (4)
*Ca*A-*bmt3Δ*	Serotype A strain, *bmt3Δ*/*bmt3Δ*	Mille *et al*, 2008 (4)
*Ca*A-*bmt4Δ*	Serotype A strain, *bmt4Δ*/*bmt4Δ*	Mille *et al*, 2008 (4)

Figure 1. Structural model of serotype A *C. albicans* phosphopeptidomannan (PPM) N-glycan moiety. *In vitro* analysis of PPM purified from *C. albicans* deletion mutants strongly suggested that the acid-labile β-mannosides is initiated by the sequential activities of three β-mannosyltransferases names CaBmt2-4p onto the Manα1-Phosphate of the phosphodiester-linked side chain of PPM (4).

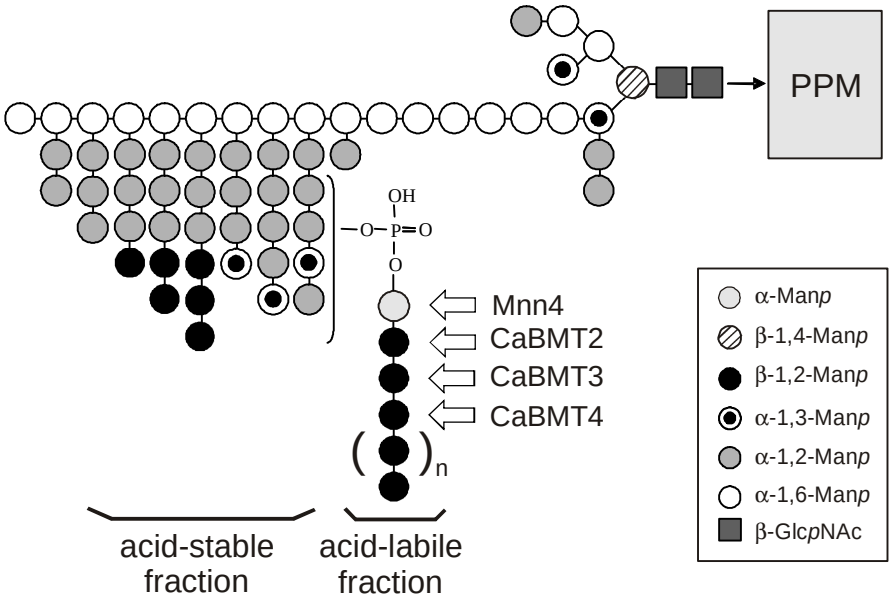


Figure 2 Liquid NMR analysis of purified PPM. ^{1}H-^{13}C HSQC NMR spectra of **A**, native PPM and **B**, acid-stable domain of PPM isolated from serotype A *C. albicans* (BWP17) enable the identification of the β-mannoside moiety owing to six NMR signals **a**-**f** (Table 1) but do not permit to easily discriminate between acid-stable and acid-labile domains. ^{1}H-^{31}P HSQC NMR analyses (**C**) confirm the presence of Manα1-phosphate and Manβ1-2Manα1-phosphate residues in native PPM, and their absence in acid-stable domain of PPM (**D**).

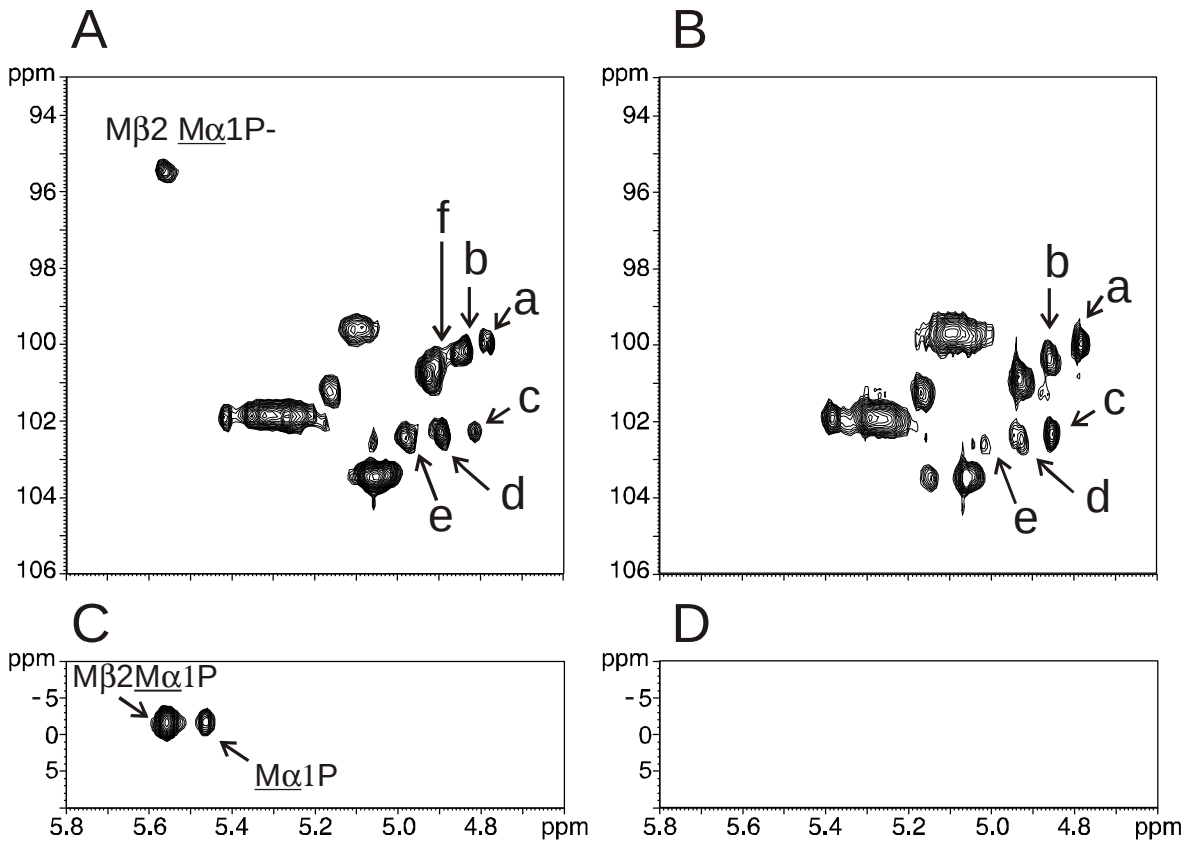
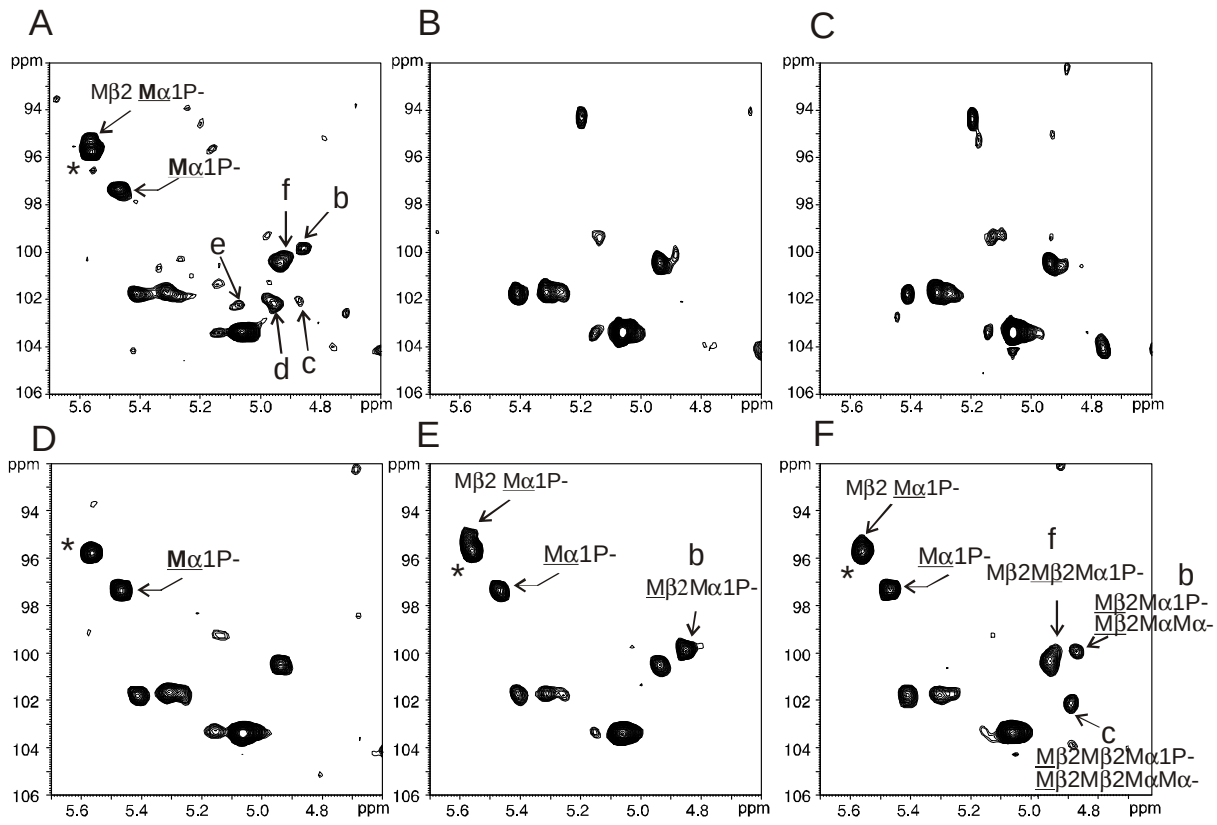
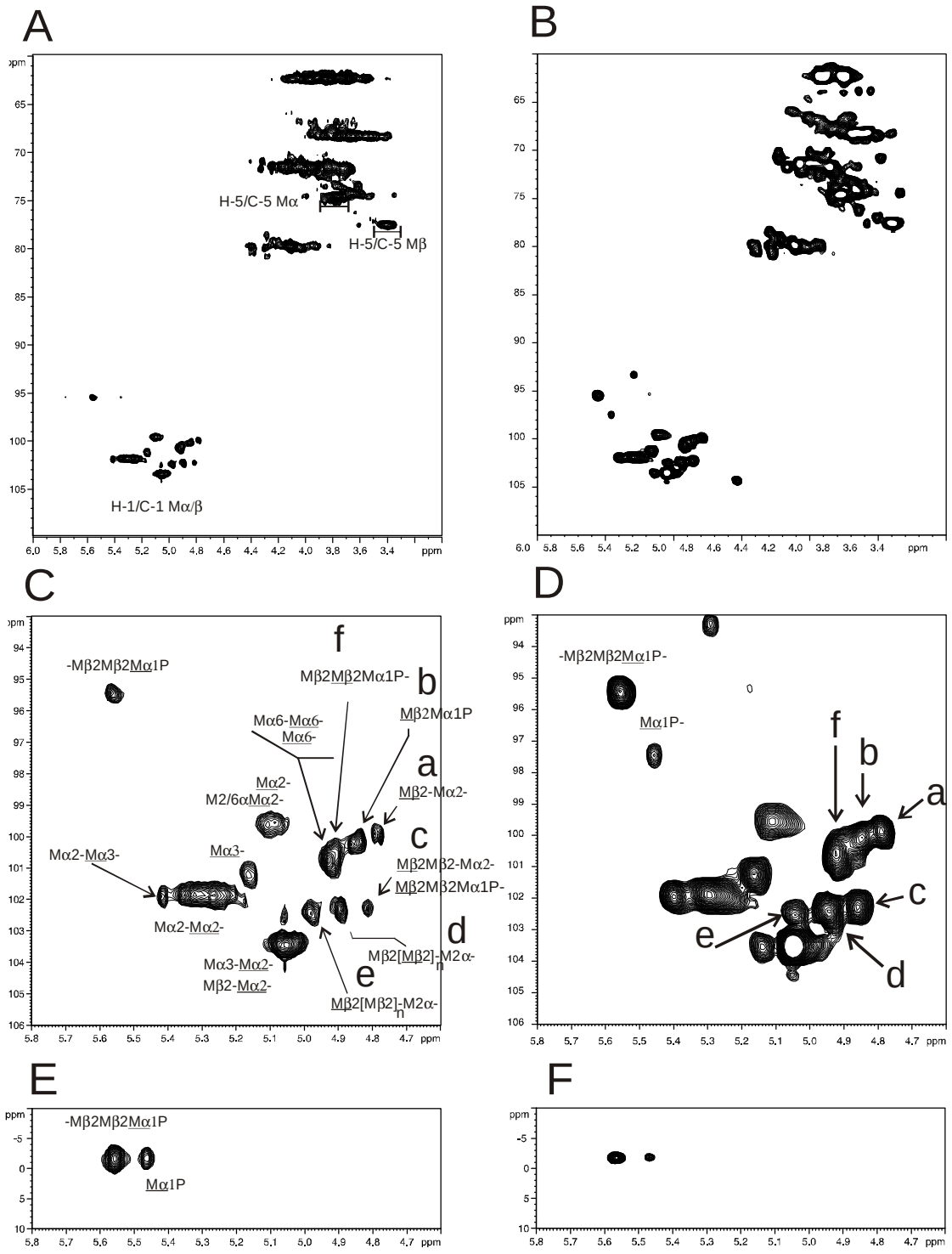


Figure 3 HR-MAS NMR analysis of six intact living *C. albicans* cells. ^{1}H-^{13}C HSQC HR-MAS NMR experiments of intact cells permitted to establish the *in vivo* structure of phosphodiester-linked side chains of the cell wall associated PPMs of **A**, serotype A (BWP17); **B**, serotype A devoid of acid-labile domain (*Ca*A-*mnn4Δ*); **C**, serotype B devoid of acid-labile domain (*Ca*B-*mnn4Δ*); **D**, *Ca*A-*bmt2Δ* ; **E**, *Ca*A-*bmt3Δ* and **F**, *Ca*A-*bmt4Δ* strains. * unidentified signal exclusively observed on ^{1}H-^{13}C HSQC HR-MAS NMR spectra of total cells.



SupFig. 1 Comparison of NMR parameters of purified mannan analysed by liquid and HR-MAS NMR. Details of **A** and **C**, ^{1}H-^{13}C HSQC liquid NMR; **B** and **D**, ^{1}H-^{13}C HSQC HR-MAS NMR; **E**, ^{1}H-^{31}P HSQC liquid NMR and **F**, ^{1}H-^{31}P HSQC HR-MAS NMR spectra of native PPM isolated from serotype A *C. albicans* demonstrating that PPMs exhibit identical NMR parameters in liquid and HR-MAS NMR. **a** to **f** signals are all assigned to H-1 β-Man residues as described in Table 1.



2.3- Diversité structurale des lipoglycannes mycobactériens

2.3.1- Contexte

Mes travaux concernant la diversité structurale des lipoglycannes ont été initiés lors de ma thèse de doctorat en collaboration avec le Dr Laurent Kremer. Ils portaient initialement sur l'élucidation de la structure des lipoglycannes. En effet, les lipoglycannes sont des molécules intrinsèquement hétérogènes. En plus de la polydispersité dont ils font preuve au sein d'une même espèce, ils présentent un certain degré de variabilité structurale inter-spécifique. L'étude de leur propriétés biologiques a non seulement révélé qu'ils étaient impliqués dans de nombreux phénomènes associés aux infections mycobactériennes, mais également que leurs activités étaient très dépendantes de leur origine. De fait, un enjeu de l'étude des LAMs était de déterminer les domaines structuraux potentiellement impliqués dans chacune de ces activités et ainsi d'aider à la compréhension des mécanismes moléculaires à la base de celles-ci. Les différences d'activité observées entre LAMs ont été en partie reliées à la nature de la coiffe, ainsi qu'au statut d'acylation de l'ancre lipidique (voir Gilleron *et al.*, 2008). Néanmoins, plusieurs cas ont été observés dans lesquels aucune corrélation précise n'a pu être dressée entre activité et structure. C'est le cas des pouvoirs chimiotactiques différents des ManLAMs de *M. bovis* BCG et *M. tuberculosis* ou des pouvoirs d'adhésions différents des ManLAMs de souches distinctes de *M. tuberculosis*.

L'étude des relations entre structure et fonctions des lipoarabinomannannes était rendue délicate à cause de deux facteurs principaux: l'hétérogénéité des préparations de LAM et le nombre alors réduit de modèles disponibles. Comme l'ont démontré les nombreuses études portant sur la structure des LAMs, l'hétérogénéité des préparations de LAM est principalement le fait de la polydispersité de leurs domaines saccharidiques et de la multiplicité des états d'acylation de leur ancre lipidique. De fait, le développement de protocoles expérimentaux permettant de réduire l'hétérogénéité du LAM s'est imposé comme une stratégie valide pour l'étude de la fonctionnalité de cette molécule. Ainsi, le fractionnement du LAM en fonction de leur localisation présumée a mis en évidence des sous-familles de molécules de structures plus homogènes, qui éventuellement induisaient des activités biologiques distinctes (Nigou, *et al*, 1997; Gilleron, *et al*, 1997; Gilleron, *et al.*, 2000 ; Gilleron *et al.*, 2006) De même, leur séparation en fonction du degré d'acylation a permis d'évaluer l'importance de la partie lipidique des LAMs pour certaines de leurs activités (Sidobre *et al.*, 2000; Sidobre *et al.* 2001; Riviere *et al.*, 2004).

Une deuxième stratégie d'étude, complémentaire de la première, consistait à rechercher des LAMs présentant des déterminants structuraux différents de ceux utilisés jusqu'à présent. En effet, seules deux classes majeures de LAMs, qui diffèrent par la nature de leurs coiffes, avaient été

décrites: les ManLAMs chez *M. tuberculosis, M. bovis, M. Avium* et *M. leprae* et les PILAMs chez *M. smegmatis* et une espèce indéfinie de mycobactéries à croissance rapide. Malgré l'existence d'une variabilité structurale au sein des familles, le faible nombre de modèles disponibles ne permet que très peu d'études comparatives. Ce type d'approche s'avère souvent être très riche en enseignements et est couramment utilisée dans l'étude des relations structure/fonction des molécules biologiques.

Plus de 70 espèces de mycobactéries ont été décrites à ce jour. Considérant l'extrême variabilité des structure glycanniques dans le monde bactérien, comme le démontre celle exhibée par les glycolipides synthétisés par les mycobactéries, nous sommes partis du postulat que la structure du LAM avait de fortes probabilités d'être beaucoup plus hétérogène que ne le suggéraient les études effectuées sur un nombre réduit d'espèces. Nos travaux se sont donc inscrits dans cette optique avec pour but de mettre en évidence l'existence de nouveaux déterminants structuraux dans les LAMs de mycobactéries encore non étudiées. A terme, la purification de nouveaux LAMs devait fournir de nouveaux outils pour décrypter les fonctions de cette famille de molécules, et éventuellement mettre en évidence de nouvelles fonctions biologiques.

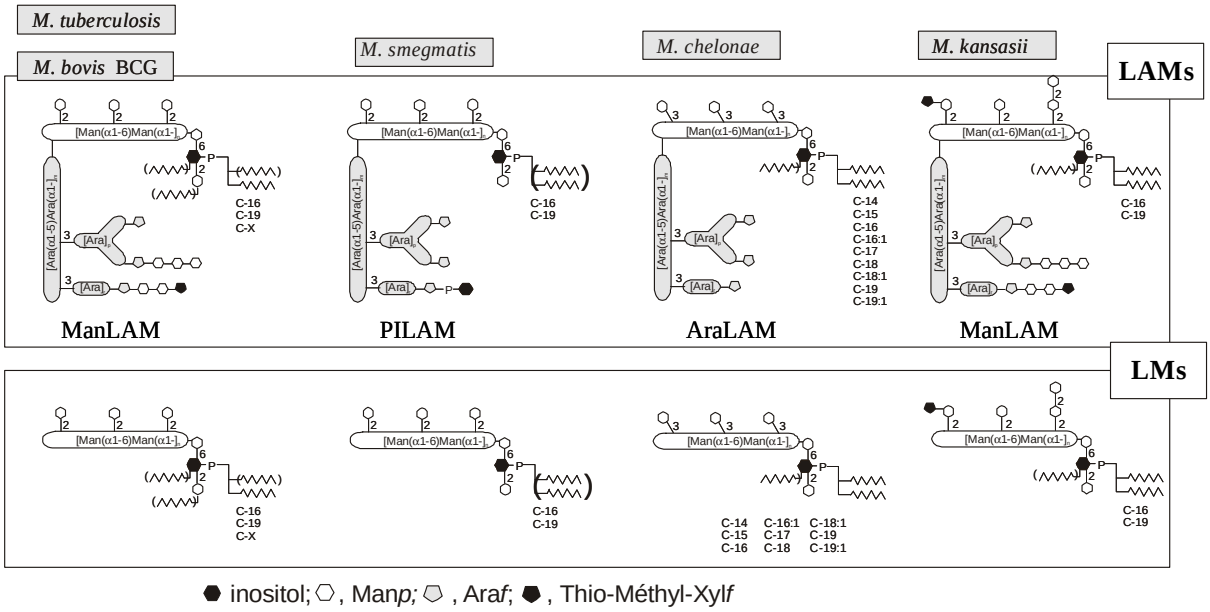
2.3.2- Résultats

Diversité structurale des lipoglycannes

Nos premiers travaux ont consisté en l'étude structurale des lipomannannes et liporabinomannannes extraits de deux espèces de mycobactéries: *M. chelonae* et *M. kansasii* (Guérardel *et al.*, 2002 ; Guérardel *et al.*, 2003). *M. chelonae* est une mycobactérie non tuberculeuse, essentiellement impliquée dans des infections cutanées, qui présente l'intérêt d'avoir une croissance rapide. Le fait qu'elle puisse être transformée en faisait de plus un modèle d'étude attractif. *M. kansasii* est quant à elle considérée comme l'une des mycobactéries atypiques la plus pathogène et nous est apparue comme un modèle intéressant par sa relevance clinique potentielle. La structure de ces composés a été réalisée au laboratoire grâce à des méthodologies d'étude classiques, directement adaptées de la littérature. Ces travaux ont mis en évidence un ensemble de variations structurales tant au niveau du domaine lipidique que des différents domaines polysaccharidiques.

Ces travaux ont permis de montrer que les lipomannannes et lipoarabinomannannes de ces deux espèces présentaient de nombreuses caractéristiques communes avec ceux des espèces déjà étudiées. En particulier, l'organisation en trois domaines de cette famille de molécule -ancre acylglycérol phosphatididyl *myo*inositol, domaine mannosyle et domaine arabinosyle- semble conservée dans l'ensemble des espèces mycobactériennes. Par contre, une liberté structurale limitée

présiderait à la synthèse de chaque domaine. Le résultat le plus surprenant était l'identification d'un nouveau facteur de variabilité au sein des domaines mannannes des LMs et des LAMs. En effet, d'après les études antérieures effectuées, ce domaine semblait être la partie la plus conservée des lipoglycannes. Les seules variations mises en évidence étaient d'ordre quantitative et non qualitative, et concernaient la taille et le degré de substitution en α-1,2 de la chaîne linéaire α-1,6-Man par des mannoses terminaux. Au contraire, l'étude de *M. chelonae* et *M. kansasii* a montré que la nature des liaisons était également variable, et que ce domaine pouvait être substitué en petite quantité par d'autres monosaccharides tel que le thio-méthyl-pentose. Il est intéressant de constater que ce même monosaccharide soufré avait déjà été mis en évidence dans le LAM de *M. tuberculosis*, mais était localisé sur le domaine arabinane (Treumann *et al.*, 2002). Sa structure exacte a été déterminée récemment comme étant un 5-déoxy-5méthylthio-*xylo*furanose (Joe *et al.*, 2006). Enfin, ces travaux ont démontré que le domaine mannanne des lipoglycannes de *M. kansasii* puisse être substitué non seulement par des unités de mannoses terminales, mais également par des di- et trimères d'α-1,2-Man, motifs normalement cantonnés à la coiffe oligomannosidique. La diversité glycanniques des lipoglycannes étudiés est résumée ci-dessous :



Le domaine arabinane, quant à lui, présente une structure qui semble invariable, quelle que soit l'espèce. Il a été proposé que les arabinosyltransférases responsables de la synthèse des domaines arabinanes du lipoarabinomannanne et de l'arabinogalactane soient, au moins en partie, les mêmes. Si tel est les cas, au vu du rôle structurel essentiel que jouent les arabinogalactanes chez

toutes les mycobactéries, on peut supposer que la pression de sélection empêche toute déviation de l'activité des glycosyltransférases impliquées dans la synthèse des domaines arabinanes, ce qui expliquerait leur constance chez les mycobactéries. Par contre, chez *M. chelonae* le domaine arabinane du LAM ne semble substitué par aucun type de coiffe, ce qui permet de définir une nouvelle classe de LAM, pour l'instant spécifique à cette espèce. Le LAM de *M. kansasii* est coiffé par des motifs oligomannosidiques et appartient donc à la famille des ManLAMs.

Enfin, les ancres lipidiques des lipoglycannes de ces deux espèces ne variaient que par leurs compositions en acides gras. Alors que *M. kansasii* présentait des acyl-glycerols très semblables à ceux de *M. tuberculosis* et *M. bovis* BCG -essentiellement substitués par des acides tuberculostéariques et palmitiques- les lipoglycannes de *M. chelonae* étaient substitués par un mélange complexe d'acides gras jamais observé dans ces molécules auparavant.

Activités biologiques associées aux lipoglycannes

Dans un deuxième temps, les propriétés immunologiques des différents lipoglycannes que nous avions isolés et caractérisés ont été comparées avec celles des lipoglycannes de souches déjà étudiées. Ces travaux ont été rendus possibles grâce à une collaboration étroite avec le Pr E. Elass qui s'est fortement impliquée à la fois dans la mise en place et dans la réalisation de ces études. De fait, elle a été la principale instigatrice de ces travaux qui ont permis de placer nos études structurales dans un contexte biologique cohérent.

L'activité anti-inflammatoire des LAM à coiffe mannosyle de *M. tuberculosis* avait été largement décrite dans la littérature, montrant que ces molécules favorisent l'adhérence et l'internalisation du bacille par les macrophages et permettent au pathogène de développer des stratégies visant à favoriser sa survie et sa multiplication au sein des granulomes. En revanche, lorsque nous avons commencé ces recherches, peu d'études étaient consacrées aux LMs. En 2003, nous avons ainsi démontré que les LMs isolés de diverses espèces opportunistes, mais pas les LAM correspondants, étaient de puissants inducteurs de cytokines inflammatoires dans les macrophages humains et que cette activité était dépendante de TLR-2 et du corécepteur CD14 (Vignal *et al.*, 2003). La 'LPS-binding protein' (LBP), une glycoprotéine sérique jouant un rôle importante dans la désagrégation micellaire des lipides bactériens et leur présentation aux cellules immunitaires, potentialise l'effet inducteur des LMs. Nos travaux ont été confirmés par de nombreux auteurs et notamment chez des souris dépourvues de TLR2.

Afin de pouvoir mieux définir les mécanismes moléculaires et cellulaires d'action des lipomannanes, nos principaux objectifs ont porté sur:

1-L'identification des motifs structuraux des LMs et des LAMs reconnus par les différents récepteurs ou co-récepteurs du système immunitaire:

En effet, il est actuellement admis que l'interaction des glycolipides bactériens avec les cellules est fondée sur l'existence de complexes pluri-moléculaires faisant intervenir à la fois, des récepteurs impliquées dans l'activation des voies de signalisation intracellulaires comme les TLR et, des molécules co-réceptrices membranaires ou solubles comme le CD14 ou la LBP. Dans le cas des TLRs, la présentation de monomères de glycolipides paraît importante pour potentialiser l'effet de ces molécules. Ceci ayant été particulièrement bien décrit pour l'interaction des lipopolysaccharides avec TLR4. En revanche, les lectines à mannose telles que DC-SIGN ou le récepteur à mannose, reconnaissent avec une meilleure affinité les structures multimoléculaires agrégées. Etant donné que nous avons observé une augmentation de l'effet inducteur des LMs en présence de la LBP et du CD14, nous avons étudié les interactions moléculaires de ces molécules avec ses co-récepteurs et identifié les motifs reconnus.

En utilisant, différents LMs et LAMs ainsi que diverses fractions glucidiques de mannanes et arabinomannes purifiées à partir de l'enveloppe de diverses mycobactéries, nous avons démontré par résonance plasmonique de surface (Biacore), l'existence d'interactions moléculaires spécifiques de ces glycolipides avec le CD14 et la LBP. Les paramétres de fixation (constantes d'affinité et constantes de vitesse d'association et de dissociation) ont été caractérisés. L'affinité des LM et LAM est plus élevée pour le CD14 que pour la LBP. L'ancre phosphatidyl-*myo*-inositol dimannosylé et acylé est le plus petit domaine des LMs et LAMs capable d'interagir avec ces récepteurs, suggérant que ces interactions exigent la présence de la partie lipidique mais pas des domaines mannanes et arabinanes. Etant donné que les LMs et LAMs forment en solution aqueuse des agrégats micellaires de taille variable, ces résultats indiquent que la LBP et le CD14 peuvent intervenir dans la présentation de ces glycolipides aux récepteurs en facilitant la désagrégation de micelles formées par les LMs et LAMs en solution aqueuse (Elass *et al.*, 2007).

2- L'analyse des activités inflammatoires des LMs et des LAMs dans les macrophages:

- Effet des LMs sur la production de métalloprotéinases matricielles

Sachant que l'une des conséquences d'une infection à mycobactéries est la formation de granulomes suivie par une nécrose tissulaire permettant la dissimination dans l'organisme du pathogène, il nous a paru intéressant d'étudier l'effet des LMs et LAMs sur l'induction de métalloprotéases matricielles par les macrophages, en particulier sur la synthèse de la MMP9, une collagénase de type IV. Ces enzymes sont une famille d'endopeptidases zinc et calcium dépendante qui jouent un rôle important dans le remodelage de la matrice extracellulaire au cours des processus physiologiques et pathologiques**.** En particulier, le taux de MMP9 sécrété par les monocytes-macrophages, est augmenté lors d'infection microbienne. Une production excessive de cette enzyme peut conduire en effet conduire à des réactions inflammatoires amplifiées conduisant à la destruction des tissus.

Nos résultats indiquent que les LMs de *M. chelonae, M. kansasii* et *M. bovis BCG*, augmentent fortement l'expression et la sécrétion de la MMP9 humaine chez les macrophages alors que l'activité inductrice des LAMs correspondants est faible. Cette induction est dépendante des récepteurs TLR2 associés à TLR1 et au CD14. En présence de la LBP, l'effet inducteur des LMs est augmenté. Ces travaux suggèrent que cette surexpression de la MMP9 pourrait participer aux processus de dégradations tissulaires observés lors d'infection mycobactérienne et favoriser ainsi la dissémination du pathogène dans l'organisme (Elass *et al.* 2005).

- Régulation par les LMs des voies d'activation intracellulaires des MAPK (Mitogen-Activated Protein Kinases)

Les mécanismes intracellulaires induits par les TLR sont complexes et dépendants de l'hétérodimérisation de ces récepteurs. En effet, suivant leur coopération avec divers co-récepteurs, les TLR peuvent activer plusieurs voies intracellulaires conduisant en particulier à la stimulation du facteur NF-*kappa* B mais également des MAPK. La durée et l'intensité d'activation des MAPK sont contrôlées par l'action de phosphatases, notamment MKP-1. Ce mécanisme de régulation pourrait être un des mécanismes par lequel le pathogène contrôle le système de défense de l'organisme. Etant donné qu'aucune étude ne décrivait l'effet des LMs sur l'induction des MAPK et leur régulation, nous avons analysé cette activité.

Nous avons démontré que les LMs, purifiés de différentes espèces et reconnus par TLR2, induisent de façon transitoire l'activation des voies MAPK (ERK, p38 and JNK MAP kinases) dans les macrophages humains. Cette stimulation est suivie par l'induction de l'expression de la MAP kinase phosphatase 1 (MKP-1) qui régule la durée et l'intensité d'activation des MAPK, en les

déphosphorylant. Les LAMs à coiffe mannosyle de *M. kansasii* ou de *M.bovis BCG* sont dépourvus de cette activité. Des inhibiteurs spécifiques des voies ERK and p38 MAPK diminuent l'effet inducteur des LMs sur la transcription de *MKP-1*. Ces travaux indiquent que par ce mécanisme, les LMs pourraient contrôler l'activité des macrophages durant le développement de l'infection (Elass *et al.* 2008).

En conclusion, ces travaux ont contribué à éclaircir les mécanismes moléculaires et cellulaires d'action des LMs dans les macrophages et visent à mieux comprendre le rôle de ces molécules dans la pathogénicité des mycobactéries. Les compétences que nous avons acquises dans ce domaine nous permettront d'élargir nos recherches à d'autres molécules isolées de l'enveloppe de mycobactéries et caractérisées au niveau structural au sein du groupe, notamment les lipooligosaccharides et les glycoprotéines, isolées de *M. marinum*. Ce travail nous a permis également d'ouvrir une nouvelle collaboration avec le Dr Chambers (TB Research Group, Veterinary Laboratories Agency Weybridge, New Haw, Addlestone, Surrey KT15 3NB, UK) concernant l'effet d'une glycoprotéine mannosylée la MBP83, fortement exprimée et secrétée par *M. bovis* et *M. tuberculosis*, sur l'expression de la métalloprotéinase MMP9 et sa reconnaissance par TLR2.

3- Projets

3.1- Bases moléculaires de la spécificité des β-mannosyltrasférases de *C. albicans*

Contexte

Alors que l'étude des voies de biosynthèse des β-glucanes et des mannoprotéines est d'ores et déjà avancée, l'identification de gènes participant à formation des β-1,2 oligomannosides du PLM ou du phospho-peptidomannane (PPM) n'est que très récente et encore partielle. Notre projet a donc pour but le clonage des gènes identifiés, l'identification de nouveaux gènes également impliqués, et la caractérisation de l'activité des enzymes codées par ces gènes. En particulier, notre intérêt se porte sur la compréhension de la spécificité de ces enzymes vis-à-vis de leurs substrats (ou « accepteurs »).

Sur la base d'homologies de séquence avec la partie C-terminale de Mnn4 (une phosphomannosyltransferase de *S. cerevisiae*), l'équipe du Dr. Daniel Poulain (INSERM U799, Lille) a récemment identifié chez *P. pastoris* 4 gènes potentiellement impliqués dans le transfert en β-1,2 de résidus mannose, la famille Bmt (Mille *et al.*, 2008). Les séquences de ces protéines ont par la suite permis l'identification de séquences homologues chez les espèces fongiques *C. albicans*, *C. glabrata*, *Debaryomyces hasenii*, *Saccharomyces castelli*, *S. kluyveri*, *Aspergillus fumigatus* et *A. terreus*, suggérant la découverte d'une nouvelle famille de gènes spécifique des champignons. Par homologie avec la séquence de PpBmt2p, 9 protéines de fonction inconnue, et fortement homologues entre elles, ont été mises en évidence chez *C. albicans*. L'étude de la structure des glycoconjugués isolés de souches mutantes déficientes dans la biosynthèse de chacune de ces enzymes a permis de déterminer la spécificité fine de six de ces β-1,2 mannosyltransférases. Elles sont spécifiques non seulement du motif accepteur, mais également de la molécule porteuse, PPM ou PLM.

A ce jour, la structure tri-dimensionnelle d'une seule mannosyltransferase fongique a été résolue (Lobsanov *et al.*, 2004) : il s'agit d'une α-1,2 mannosyltransférase impliquée dans la biosynthèse des glycoprotéines pariétales de *S. cerevisiae*, et appartenant à la famille 15 des glycosyltransférases. Celle-ci possède un repliement mixte α/β similaire à celui de la glycosyltransférase-A qui utilise le GDP-mannose comme donneur et un α-mannose comme accepteur, l'anomérie du donneur étant conservée en fin de transfert. Les séquences codant pour les β-1,2 mannosyltransférases de *C. albicans* et de *P. pastoris* ne présentent, quant à elles, aucune homologie avec les domaines structuraux répertoriés à ce jour (pas de réponse significative lors de blasts contre la base de données nr ou d'analyse pfam). Ainsi l'originalité de ces protéines réside

tant dans la spécificité de la réaction catalysée (liaison synthétisée et accepteur employé), que dans leur structure primaire (et probablement tridimensionnelle).

Objectifs

L'objectif de ce projet est de mieux comprendre les voies de biosynthèse de ces structures originales. Nous allons nous focaliser sur les trois points suivants :

- Caractérisation biochimique des β-1,2 mannosyltransférases. Les étapes préliminaires consisteront à cloner les gènes chez *P. pastoris*, exprimer les enzymes sous forme soluble directement dans le milieu de culture, puis de les purifier. Enfin, des études enzymologiques devraient nous permettre de mieux comprendre le mode de synthèse de ces β-1,2 oligomannosides.
- Compréhension de la séquence d'action des différentes enzymes pour l'élaboration de structures complexes telles que les PPM, sachant qu'une seule et même β-1,2 mannosyltransférase est probablement capable d'employer différents accepteurs et donc d'intervenir sur différents sites de la chaîne glycannique.
- A plus long terme, la résolution de la structure tridimensionnelle de ces enzymes, par cristallographie, pourra venir compléter les caractérisations biochimiques. L'obtention de ces nouvelles structures sera d'un intérêt majeur pour le design de nouveaux agents antifongiques potentiels.

Programme de recherche

Cloner les gènes chez Pichia pastoris

Le clonage des gènes codant pour les β-1,2 mannosyltransférases de *C. albicans*, ainsi que la caractérisation des enzymes correspondantes seront réalisés en relation étroite avec les laboratoires du Dr. Daniel Poulain (INSERM U799, Lille) et du Dr Stefan Wildt (GlycoFi, Inc./Merck & Co.).

Les gènes CaBmt1 à 9 pourront être individuellement insérés en multi-copie dans le génome de *P. pastoris*, et surexprimés à l'aide du promoteur fort et inductible au méthanol, AOX1. La sécrétion des protéines d'intérêt dans le milieu de culture sera rendue possible par la fusion N-terminale d'un signal de sécrétion (« facteur α », Invitrogen). Une fusion N- ou C-terminale avec une étiquette telle que 6xHis, permettra la purification ultérieure des enzymes, par chromatographie d'affinité. Ces dernières seront en principe rendues solubles par la troncature de leur fragment transmembranaire. En effet, l'utilisation du logiciel TopPred (http://mobyle.pasteur.fr/cgi-

bin/MobylePortal/portal.py) nous a permis de prédire l'existence systématique, pour les 9 protéines, d'un segment transmembranaire d'une vingtaine d'acides aminés, situé côté N-terminal.

La production à l'échelle pilote de ces différentes enzymes sera éventuellement assurée dans le cadre d'une collaboration avec Dr. Stephan Wildt (Strain Development, GlycoFi, Inc/Merck and Co., Inc, Lebanon, NH 03766, Etats-unis).

Etudier la spécificité d'accepteurs

L'activité individuelle de chaque enzyme purifiée sera déterminée à partir d'accepteurs oligosaccharidiques synthétiques et naturels. Les différentes souches mutantes de *C. albicans*, construites au cours de l'étude de Mille *et al.* (2008), et déficientes pour les différents gènes de β-1,2 mannosyltransférases, pourront être utilisées pour l'obtention des accepteurs. Ainsi, les accepteurs endogènes putatifs de chaque enzyme impliquée dans l'addition de β-Man sur le PPM (Bmt1-4) seront générés à partir des souches mutantes correspondantes. Les PPM seront purifiés et des oligomannosyles phosphorylés et non phosphorylés seront libérés par hydrolyse alcaline et acide respectivement, puis purifiés par chromatographie de gel filtration et HPLC préparative. De même, des molécules tronquées de PLM seront purifiées à partir des cellules déficientes dans la synthèse de Bmt5 et 6. La structure de ces fractions de mannanes sera précisément déterminée par RMN. Par ailleurs, les produits synthétisés à partir de l'une ou l'autre des différentes β-1,2 mannosyltransférases objets de notre étude pourront eux-mêmes servir d'accepteurs potentiels pour les autres enzymes. Les paramètres cinétiques de chaque enzyme (K_M et V_{max}), obtenus pour les différents accepteurs, seront comparés afin d'appréhender les "préférences" de chacune. L'originalité de ce projet repose en partie sur la capacité de notre groupe à purifier et analyser des accepteurs endogènes putatifs des enzymes étudiées à partir de différentes souches de *C. albicans*.

L'ensemble de ces résultats pourra ainsi permettre de mieux comprendre la séquence d'action de ces β-1,2 mannosyltransférases pour la synthèse d'oligosaccharides complexes.

Etudier les relations structure-fonctions

Dans une dernière étape, nous allons tenter de déterminer les motifs structuraux essentiels contrôlant la spécificité fine de chaque β-mannosyltransférase que nous avons précédemment déterminée. Pour ce faire, nous étudierons la structure tri-dimensionnelle du domaine catalytique d'un des membres de la famille des Bmts par co-cristallisation de l'enzyme avec son substrat accepteur oligosaccharidique préférentiel. Ces études s'effectueront en collaboration avec le Dr Julie Bouckaert de l'Université de Bruxelles, dont l'équipe possède une expertise reconnue dans

l'étude de la structure de lectines bactériennes dont certaines sont spécifiques de manno-glycoconjugués. Sur la base de leurs fortes homologies structurales nous modéliserons la structure des autres membres de la famille Bmts au sein du laboratoire.

3.2- Etude du rôle de la glycosylation des protéines chez *Mycobacterium marinum*

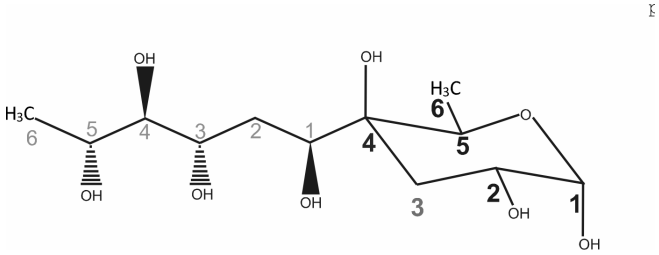
Contexte

Ces travaux s'intègrent dans un projet plus large, basé sur l'utilisation du modèle émergeant *Mycobacterium marinum/Danio rerio*. Une partie de ce projet est actuellement financée par l'Agence Nationale de la Recherche en collaboration avec les Dr. L. Kremer (CNRS UMR 5539) et P. Herbomel (CNRS URA2578). *M. marinum* (*M. ma*) est un pathogène naturel des poissons qui induit une infection aux caractéristiques physiopathologiques semblables à celles rencontrées chez l'homme infecté par *M. tuberculosis* (*M. tb*). Outre les similitudes entre *M. ma* et *M. tb* du point de vue de la pathogenèse, ces deux organismes sont très proches génétiquement, d'où l'intérêt porté à *M. ma* en tant que modèle pour l'étude de la physiopathologie tuberculeuse. De plus, *Danio rerio* ("poisson zèbre"), un hôte naturel de *M. ma*, représente un modèle de choix car son embryon est optiquement transparent, ce qui permet de suivre les interactions hôte-pathogènes *in vivo* et ce en temps réel. Nos travaux en cours sur ce sujet consistent à comparer les profil lipidiques et glycolipidiqueses de souches de *M. ma* induisant soit une infection aiguë soit une infection chronique granulomateuse dans le poisson. Après identification, ces composants sont purifiés afin premièrement de réaliser leur analyse structurale fine, et deuxièmement de les coupler à des billes de latex fluorescentes. Ces billes chargées en lipides et glycolipides spécifiques sont alors injectées dans l'embryon du poisson. La participation de ces molécules dans divers phénomènes tels que la phagocytose, le recrutement cellulaire, la formation de granulomes, et l'incidence des réponses innée/adaptative sont ainsi directement visualisées et évaluées *in vivo*.

Suite à l'analyse systématique après marquage métabolique des glycolipides de sept souches de *M. ma*, en collaboration avec le Dr. Laurent Kremer, nous avons mis en évidence chez une souche induisant une pathologie chronique (Mma7) la présence d'une accumulation d'un intermédiaire de biosynthèse glycolipidique. L'analyse chromatographique suggère un défaut de biosynthèse du lipo-oligosaccharide IV (LOS IV) entraînant l'accumulation de LOS III dans la paroi de la mycobactérie. La structure de ces composés chez *M. ma* n'étant que partiellement

connue, nous avons initié la purification des différents LOS à partir de la souche Mma7 et d'une souche présentant un profil lipidique normal dans le but de définir la nature du défaut de glycosylation dans la souche Mma7.

L'étude structurale des LOS purifiés des deux souches MmaM et Mma7, a tout d'abord confirmé que la souche Mma7 présentait bien un défaut de biosynthèse du LOS IV. De plus nous avons établit que les structures des LOSII, II et IV étaient toutes basées sur un corps oligosaccharidique commun de séquence C4-(tetrahydroxy-1,3,4,5-hexyl)-3,6-dideoxy-α-Gal*p*-(1-4)-β-Xyl*p*-(1-4)-α-Rha*p*3Me-(1-3)-β-Glc*p*-(1-3)-β-Glc*p*-(1-4)-α-Glc*p*-(1-1)-α-Glc*p* qui contient en position terminale non réductrice un monosaccharide très rare, le caryophyllose, dont la structure est représentée ci-dessous :



Dans le cadre de ce projet, nous couplons également tous les composés purifiés à des billes de latex fluorescentes qui sont injectées dans l'embryon transparent du poisson par l'équipe du Dr. Herbomel. Ces expériences permettent de suivre en temps réel et *in vivo* le devenir de chaque composé et sa participation dans divers phénomènes comme la phagocytose, le recrutement cellulaire, la formation de granulomes et l'incidence des réponses innée/adaptative. Les billes fluorescentes recouvertes de LM, LAM et LOS ont ainsi été injectées dans le ventricule postérieur du cerveau pour évaluer la capacité des composés purifiés à induire le recrutement des macrophages dans le ventricule injecté (chimiotactisme). De ces premières expériences il apparaît que le nombre de macrophages recrutés dans le ventricule des embryons injectés avec les billes couplées aux lipoglycanes est environ 1.5 à 2 fois plus élevé qu'après injection de billes contrôle et celles couplées aux lipo-oligosaccharides l'est de 3 à 4 fois plus. En parallèle, nous déterminons *in vivo* les capacités immuno-modulatrices de ces composés par injection dans la circulation sanguine et quantification des cytokines pro-inflammatoires zTNFa et zTNFalp présentes chez le poisson zèbre (équipe du Dr. L. Kremer). Ces test sont également réalisés *in vitro* sur des cultures de cellules humaines. Les premiers résultats obtenus sur cellules Thp1 différenciées montrent une induction spécifique de cytokines pro-inflammatoires par les billes couplées aux LOS.

Nous nous proposons d'étendre le cadre du projet qui vient d'être décrit à l'étude des relations entre structure et fonction des glycoprotéines en utilisant le même modèle expérimental pathogène-hôte *Mycobacterium marinum-Danio rerio*.

Objectifs

La motivation de ce projet de recherche est de mieux comprendre l'influence de la glycosylation des protéines mycobactériennes sur la pathologie tuberculeuse par une approche de biochimie structurale. Nos trois principaux objectifs sont :

- D'étudier systématiquement les glycoprotéines exprimées par *M. marinum* potentiellement impliquées dans la pathologie.
- De déterminer l'importance relative des gènes impliqués dans la glycosylation des protéines.
- D'etudier le rôle de glycoprotéines excrétées sur la modulation de la réponse immunitaire.

Programme de recherche

Analyse des profils de glycosylation.

Les glycoprotéines mannosylées pariétales et excrétées seront purifiées en bloc par chromatographie d'affinité sur colonne de lectine ConA. Sur la fraction retenue, nous mettrons en évidence les manno-protéines par blot avec lectines de spécificités différentes (ConA et GNA). Nous vérifierons que les protéines retenues par chromatographie d'affinité sont réellement mannosylées en inhibant la fixation des lectines par des inhibiteurs glycanniques et en démannosylant enzymatiquement. Les parties glycanniques seront libérées à partir du mélange de glycoprotéines par méthodes chimiques et analysées en mélange par SM, RMN, HPLC et GC/MS. En parallèle, les protéines mannosylées seront séquencées par spectrométrie de masse et identifiées par comparaison avec les banques de données génétiques. Sur la fraction non retenue par la colonne d'affinité, nous évaluerons la possibilité de la présence d'autres types de glycosylation que la mannosylation. Ceci sera réalisé sur la base d'études directes après libération chimique et l'utilisation d'un large panel de lectines.

L'ensemble de ces analyses nous fournira une vision globale de l'état de glycosylation des protéines dans une souche témoin de *M. marinum* et permettra de déterminer :

- la présence ou non des glycoprotéines déjà connues dans d'autres souches de mycobactéries
- la nature des profils de glycosylation globaux et individuels des glycoprotéines
- l'identité d'éventuelles nouvelles glycoprotéines
- la topologie partielle des glycoprotéines

Comparer les profils de glycoprotéines dans différentes souches

Sur la base de l'identification des glycoprotéines dans la souche de référence, nous comparerons les profils de glycoprotéines, totales et excrétées, exprimées dans six souches de *M. marinum* induisant des pathologies chroniques ou aigües.

- Ces analyses nous permettront de déterminer l'implication éventuelle de glycoprotéines dans le type de pathologie induite par *M. marinum*

Purifier les glycoprotéines d'intérêt

En fonction des résultats acquis nous sélectionnerons une ou plusieurs glycoprotéines dans le but d'en étudier la structure fine et de l'utiliser dans des tests fonctionnels. Nous sélectionnerons les glycoprotéines nouvelles, qui présentent des profils de glycosylation particuliers ou dont l'expression varie en fonction de la pathologie. Les glycoprotéines choisies seront purifiées par chromatographies successives (affinité, échange d'ions et interactions hydrophobes) et leur structure fine analysée par MS/MS.

Déterminer l'importance des gènes impliqués dans la glycosylation des protéines.

Sur la base de l'identification d'un gène codant pour une O-mannosyltransférase chez *M. tuberculosis*, nous rechercherons les homologues des mannosyltransférases potentiellement impliqués dans la biosynthèse des mannoprotéines chez *M. marinum*. Nous génèrerons des souches mutantes de *M. marinum* dans lesquelles les enzymes mises en évidence seront soit surexprimées soit inhibées. Le phénotype de ces souches sera déterminé en terme de croissance et de pouvoir infectieux. L'ensemble de ces travaux sera effectué par l'équipe du Dr Laurent Kremer (CNRS UMR 5235, Université Montpellier II).

Nous analyserons les profils glycanniques et glycoprotéiques des différentes souches mutantes selon la même méthodologie précédemment décrites.

- Ces analyses nous permettront de déterminer l'influence individuelle des enzymes de O-mannosylation sur la glycosylation des protéines de *M. marinum*.

Etudier le rôle de glycoprotéines sur la modulation de la réponse immunitaire.

Nous étudierons le rôle des glycoprotéines excrétées, en mélange et purifiées, dans les réponses cellulaire et immunitaire de l'hôte. Nous profiterons de l'avantage unique que confère le modèle *M. marinum/Danio rerio* pour évaluer l'induction des réponses pro- et anti-inflammatoires *in vivo*. Les composés seront micro-injectés dans les embryons de *Danio rerio* et la modulation du système immunitaire sera suivie au cours de l'infection par dosage des ARN codant pour un panel de cytokines. En parallèle, nous utiliserons une lignée macrophagique humaine (THP-1) pour déterminer plus précisément la signalisation intracellulaire. Les activités des glycoprotéines issues des différentes souches sauvages et mutantes seront comparées dans les différents systèmes expérimentaux. L'action de glycosidases spécifiques sera également évaluée sur la fonction des différents composés.

- Ces analyses ont pour objectif de corréler les activités biologiques des glycoprotéines avec leur degré de glycosylation.

C- Développement de modèles animaux d'étude de la glycosylation

1- Contexte

Notre équipe de recherche est impliquée depuis de nombreuses années dans l'étude de la diversité structurale des glycannes, en particulier des O-glycannes, dans le règne animal. Etudiée de façon extensive chez les mammifères, la structure des O-glycannes demeure largement inconnue dans le reste du règne animal. Quelques études ponctuelles chez les insectes, les parasites et les oiseaux mettent en évidence l'existence de schémas structuraux très différents de ceux existant chez les mammifères, générant une extraordinaire diversité structurale. A ce jour, les deux seules études systématiques de la structure des O-glycannes chez différentes espèces concernent la classe des amphibiens par le Dr Strecker, et la classe des poissons par l'équipe du Dr Inoue, au Japon.

De manière surprenante, la structure du noyau des O-glycannes est remarquablement conservée tout au long de l'évolution, à quelques rares exceptions. A titre d'exemple, tous les O-glycannes des mucines salivaires de martinet d'Asie (genre *Collocalia*) contiennent exclusivement les noyaux de types 2 ou 5 (Wieruszeski *et al.*, 1987; Strecker *et al.*, 1992). De même les O-glycannes isolés des œufs de nombreuses espèces de *Salmonidae* s'articulent autours des noyaux de types 1 et 5 tandis que la grande majorité des O-glycannes d'amphibiens possèdent les noyaux de types 1, 2 et 3 (Inoue *et al.*, 1997). De fait, la remarquable variabilité structurale des O-glycannes tient essentiellement à l'élongation des chaînes glycanniques, qui peut prendre les formes les plus diverses en fonction de l'organisme. Ainsi, les O-glycannes isolés de la mucine salivaire du martinet d'asie présentent un squelette constitué du disaccharide Gal(β1-4)Galβ, séquence inconnue chez l'homme. Les divergences observées dans la structure des O-glycannes des modèles animaux les plus éloignés des mammifères rendent les notions de squelette et de périphérie difficiles à appliquer.

Des études systématiques effectuées chez diverses espèces d'oiseaux, de poissons, d'insectes et d'amphibiens, ont permis de mettre en évidence deux faits majeurs. Premièrement, la diversité structurale des glycannes semble sans limite. Deuxièmement, la structure des glycannes est spécifique à chaque espèce. La spécificité d'espèce, bien sûr, n'exclut pas qu'une même structure puisse être commune à plusieurs espèces, car certaines activités glycosyltransférasiques sont ubiquitairement retrouvées dans le règne animal. Ceci est particulièrement vrai pour les structures les plus courtes, souvent réduites au noyau des O-glycannes. Néanmoins, chaque espèce étudiée possède bien un profil de glycosylation unique. Malgré le nombre considérable d'espèces animales, le potentiel de diversité de la glycosylation semble être compatible avec une telle hypothèse. Au delà de l'aspect purement phénoménologique, ces observations suscitaient de nombreuses questions quant aux origines d'une telle diversité et aux implications biologiques qui en découlent.

Les glycosyltransférases impliquées dans la synthèse des glycannes sont des enzymes dont la spécificité est habituellement très stricte. Ainsi, l'observation directe d'un motif glycannique quelconque suppose l'existence d'au moins une glycosyltransférase à l'origine de sa synthèse. Suivant ce principe, les études effectuées au laboratoire sur 22 espèces d'amphibiens suggèrent l'existence d'une vingtaine d'activités glycosyltransférasiques jamais décrites auparavant et donc d'au moins autant de glycosyltransférases nouvelles. Par exemple, l'analyse des O-glycannes de trois espèces d'ambystomes démontre l'existence de quatre activités fucosyltransférasiques inédites permettant le transfert de fucose en α-1,4 et α-1,5 sur un résidu de Kdn et en α-1,2 et α-1,3 sur du fucose (Fig. 4).

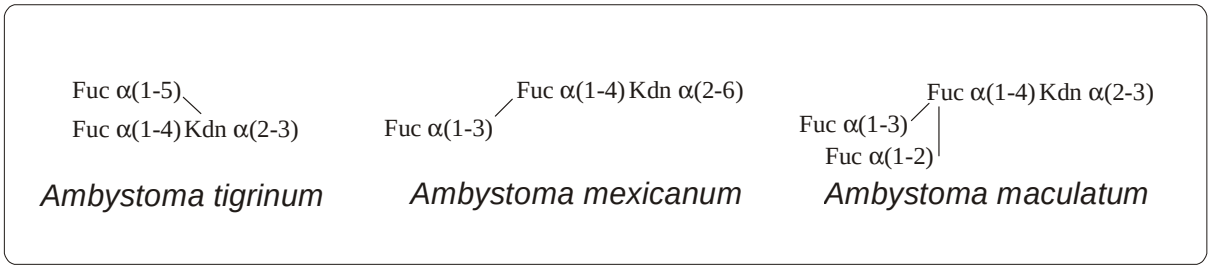


Fig.4 : Epitopes fucosylés observés chez les ambystomes (http://glycobase.univ-lille1.fr/base/)

Le fait que ces activités ségréguent différemment dans chaque espèce suggère qu'elles sont bien le fait de quatre enzymes différentes. En effet, *A. maculatum* exhibe les activités α-1,2 et α-1,3 fucosyltransférasiques tandis que *A. mexicanum* n'exhibe que l'activité α-1,3 fucosyltransférasique. De même, l'activité α-1,4 est commune aux trois espèces, tandis que l'activité α-1,5 n'est présente que chez *A. tigrinum*. Enfin, ces activités sont hautement spécifiques au substrat. Ainsi, seul le fucose présent dans la séquence Fuc(α1-4)Kdn peut être à son tour fucosylé, tandis que les fucoses présents dans les motifs Fuc(α1-2)Gal et Fuc(α1-3)GlcNAc ne le sont jamais. Sur la base d'homologies de séquences, des études phylogénétiques ont permis de classifier les fucosyltransférases des vertébrés en trois groupes: α1,2-, α1,3- et α1,4-fucosyltransférases ayant pour substrat le di-N-acétylchitobiose GlcNAc(β1-4)GlcNAc ou le motif Gal(β1-3/4)GlcNAc (Oriol *et al.*, 1999). Ainsi, les activités fucosyltransférasiques très particulières observées chez les ambystomes ne s'intègrent *a priori* dans aucun modèle d'évolution proposé et l'origine des enzymes impliquées reste inconnue. Bien que très spécifique, l'activité des glycosyltransférases semble très sensible aux mutations. Ainsi, il a été montré qu'une mutation unique ($Trp^{111} \rightarrow Arg$) dans le domaine hyper-variable de la FUT-III était suffisant à modifier son substrat accepteur préférentiel du disaccharide de type 1 au type 2 (Dupuy, F. *et al.*, 1999). De fait, les fucosyltransférases non

identifiées d'ambystome pourraient aussi bien dériver de l'ancêtre commun des autres fucosyltransférases de vertébrés que d'avoir une origine totalement différente.

L'étude et la comparaison des glycosyltransférases animales sont devenues un enjeu majeur pour la compréhension à la fois des mécanismes moléculaires à la base de leurs activités, mais aussi de leurs rôles dans l'embryogenèse, la fécondation ou les réactions immunitaires. De ce point de vue, il apparaît que l'existence d'activités glycosyltransférasiques imprévisibles est la limitation la plus flagrante à l'utilisation de modèles basés uniquement sur l'homologie de séquence des glycosyltransférases pour prédire le potentiel de glycosylation d'un organisme ou d'un tissu. De même, l'utilisation d'accepteurs exogènes standards pour tester l'activité d'une glycosyltransférase isolée d'un modèle biologique quelconque peut se révéler trompeuse, le substrat endogène pouvant être imprévisible. A ce titre, le cas de la fucosylation du Kdn ou du fucose chez les ambystomes est éloquent. Il est à noter également que l'existence d'activités "exotiques" ne se limite pas aux modèles très éloignés de l'homme, comme l'a démontré l'identification du motif Fuc(β1-6)GalNAc dans la mucine sous-maxillaire bovine (Martensson, S. *et al.*, 1998). *A contrario*, les études systématiques des glycannes d'amphibiens ont eu le mérite de mettre en évidence l'existence d'une énorme diversité dans les activités glycosyltransférasiques liées à la synthèse des O-glycannes, mais le manque d'outils moléculaires et de données génétiques sur les modèles utilisés rendent utopique la caractérisation de toute glycosyltransférase d'intérêt à partir de ceux ci. En conclusion, il apparaît que l'étude fine de toute nouvelle glycosyltransférase nécessite une connaissance préalable de ses substrats endogènes potentiels et que les études effectuées sur le modèle humain à partir de bases structurales solides ne sont pas systématiquement applicables aux autres modèles animaux.

2- Travaux

2.1- Profils glycanniques d'animaux modèles

Le modèle xenope

Initialement, dans le cadre de la thématique principale du Dr Strecker concernant la diversité structurale des O-glycannes de gangues oviducales d'amphibiens, j'ai entrepris l'étude des modèles *Xenopus laevis* et *Xenopus tropicalis*. En effet, *Xenopus laevis* est un modèle historique d'étude en biologie du développement qui est largement utilisé depuis des décennies. L'espèce *Xenopus tropicalis* s'est quant à elle imposée comme l'un des modèles les plus prometteurs en biologie du développement. En effet, bien que *Xenopus laevis* soit le modèle amphibien le plus utilisé jusqu'à présent, et par là même celui dont la génétique est la mieux connue, son étude se révèle complexe du fait sa nature pseudotétraploïde. Dernièrement, pour pallier cet inconvénient majeur, *Xenopus tropicalis*, une espèce diploïde voisine de *X. laevis*, a été développée dans de nombreux laboratoires pour remplacer ce dernier. La mise en place d'outils génétiques spécifiques à cette espèce laisse espérer un développement rapide de *X. tropicalis* comme modèle d'étude.

Les O-glycannes des deux espèces d'amphibiens ont été libérés de préparations de mucines de gangues oviducales par β-élimination en milieu réducteur, isolés par HPLC et analysés par une combinaison de techniques physico-chimiques -RMN, spectrométrie de masse MALDI, chromatographie en phase gazeuse- et l'utilisation de lectines spécifiques. Chez *X. laevis*, ces analyses ont été effectuées en parallèle sur six spécimens issus du même élevage et ont nécessité le séquençage complet de 76 oligosaccharides. Ces analyses ont permis de mettre en évidence pour la première fois chez une espèce animale un polymorphisme intra-spécifique portant sur la structure des O-glycannes. Celui ci résulte de l'expression différentielle de deux activités glycosyltransférasiques distinctes -α1,4-galactosyl- et α1,3-fucosyltransférasiques- impliquées dans la synthèse de la partie terminale non-réductrice des O-glycannes. L'étude de la répartition des structures glycanniques sur les différentes mucines oviducales a révélé que le polymorphisme glycannique touche également les glycoprotéines au sein d'un même spécimen.

Chez *X. tropicalis*, nous avons identifiés douze O-glycannes majeurs, dont neuf avaient déjà été identifiés chez d'autres espèces d'amphibiens (Fig. 5). Les composés les plus simples avaient également déjà été mis en évidence dans les mucines de diverses autres espèces animales, dont l'homme. Par contre, trois O-glycannes substitués par des déterminants de type Le^a présentaient des structures encore jamais mises en évidence. De fait, la glycosylation de *X. tropicalis* se caractérise par plusieurs traits qui lui sont spécifiques. En particulier, la présence du déterminant Le^a le distingue de toutes les autres espèces d'amphibiens qui ont été étudiées jusqu'à présent.

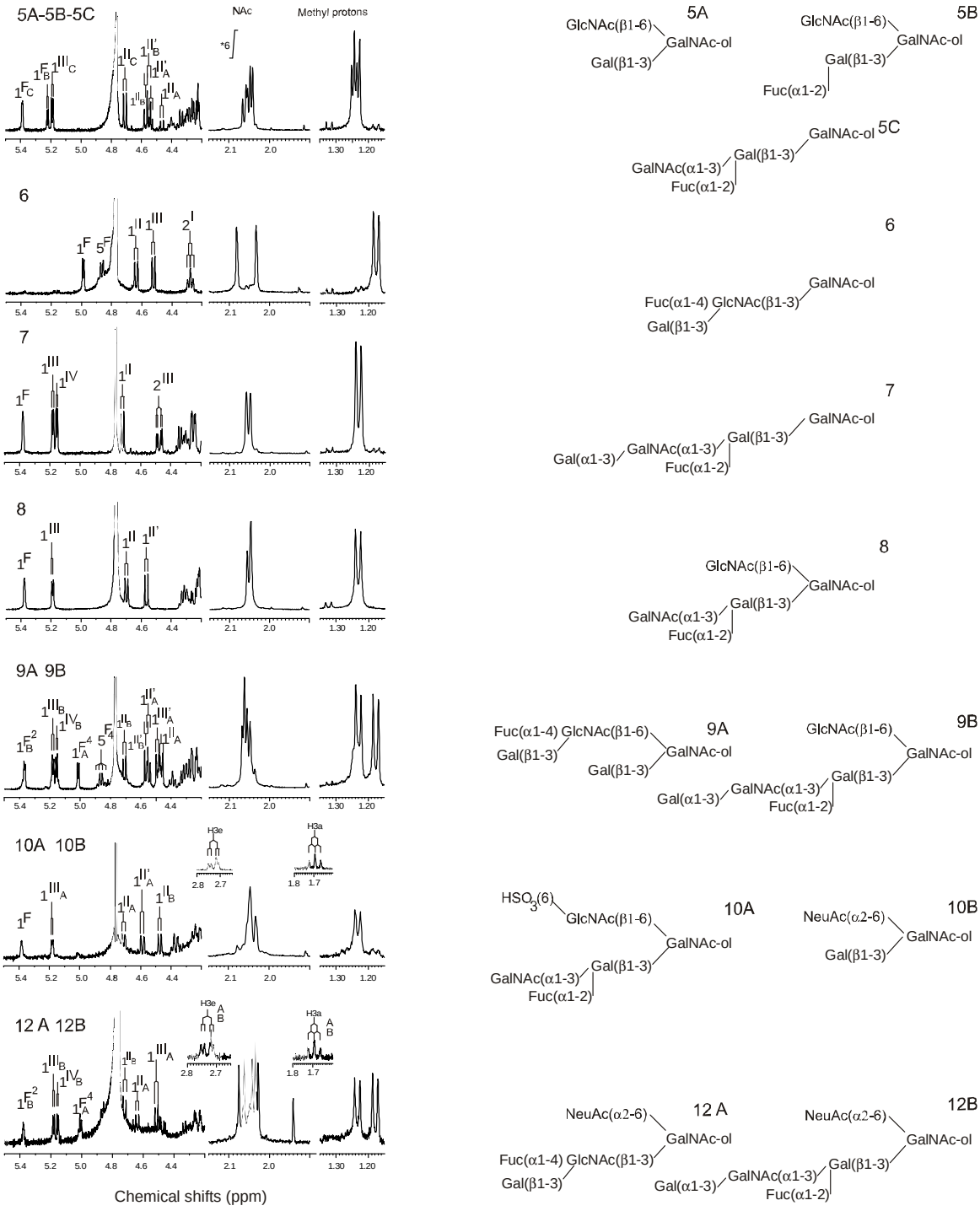


Fig. 5: Résumé des structures des O-glycannes isolés des gangues oviducales de *X. tropicalis*

D'une manière générale, le déterminant Le^a n'a été observé que très rarement dans les glycoprotéines animales, alors qu'il est couramment trouvé dans les glycolipides. A notre connaissance, les seuls exemples connus sont les O-glycannes de type mucine qui substituent les

immunoglobulines A sécrétées dans le lait humain (Pierce-Crétel *et al.*, 1989). Par contre, le déterminant Le^a est très commun dans les glycoprotéines de plante (Palma *et al.*, 2001). De plus, le composé Gal(β1,3)[Fuc(α1,4)]GlcNAc(β1,6)[Gal(β1,3)]GalNAc-ol (glycanne **9A**) représente l'unique exemple connu dans le règne animal de chaîne de type I directement liée au GalNAc terminal. En effet, jusqu'à présent, dans les O-glycannes, les glycolipides et les oligosaccharides libres du lait, la chaîne de type I avait exclusivement été observée en position terminale non réductrice, liée le plus souvent sur un enchaînement de type II, ou tout du moins sur un résidu de galactose en position C-3. Dans tous les O-glycannes décrits, la position C-6 est communément substituée par un disaccharide de type II, mais jamais par un type I.

Au sein du règne animal, seuls les deux gènes FUT3 et FUT5, qui font partie du cluster de gènes de type Lewis (FUT3, -5 et -6), sont connus pour coder des α3/4-fucosyltransférases permettant de synthétiser l'épitope Le^a (Costache, M. *et al.*, 1997). Les α3/4-fucosyltransférases se différencieraient des α3-fucosyltransférases grâce à une séquence consensus très courte HHWD à la place de HHRD/E dans leur domaine hypervariable (Table 1).

Substrat accepteur	Nom de l'enzyme (espèces)	Séquence conservée
Type-1 et -2	FUT3 (Homme/Chimpanzé/macaque)	**HHWD**
"	FUT5 (Homme/Chimpanzé)	**HHWD**
Type-2	FUT6 (Homme/Chimpanzé)	**HHRE**
"	FUT7 (Homme/Souris)	**HHRE**
"	FUT4 (Homme/Souris)	**HHRD**

Table 1 : Nature des enzymes responsables de la synthèse des épitopes Le^x et Le^a (Costache *et al.*, 2002)

Ces gènes n'ont jusqu'à présent été identifiés que chez les homoïdés (humains et chimpanzés), tandis que l'activité α3/4-fucosyltransférases n'a été mise en évidence que chez les homoïdés et les singes du vieux monde (Cercopitécoides), mais jamais chez les singes du nouveau monde (Platyrrhines). Un orthologue de FUT3 dont le produit présente une activité α4-fucosyltransférasique, *FUT3 rh*, a néanmoins été cloné d'un représentant des Cercopitécoides (Dupuy, F. *et al.*, 2002). Ces études suggéraient que l'activité α4-fucosyltransférasique est apparue deux fois de façon indépendante au cours de l'évolution des primates, à partir d'un même gène

ancestral codant une α3-fucosyltransférase. Néanmoins, aucune donnée n'était jusqu'à présent disponible quant à l'existence de telles activités ailleurs que chez les primates.

Suite à la découverte de nouvelles activités fucosyltransférasiques chez les xénopes, une collaboration a été mise en place avec le groupe du Pr A. Maftah de l'Unité de Génétique Moléculaire Animale (UMR-INRA 1061), dont le but était d'étudier les enzymes responsables de ces activités. Les premiers travaux effectués par des techniques de biologie moléculaires ont permis de mettre en évidence l'existence chez *X. laevis* de trois gènes codant pour des α1,3-fucosyltransférases sur la base des séquences consensus HHRD/E qui étaient potentiellement responsables du polymorphisme observé, mais aucune fucosyltransérase présentant la séquence HHWD typique des α1,3/4-fucosyltransférases. Par contre, un unique gène présentant de fortes homologies avec les fucosyltransférases de type Lewis a été observé dans le génome de *X. tropicalis*. En place de la séquence permettant de distinguer les α1,3/4- des α1,3-fucosyltransférases il présentait un motif inédit HSRD, ce qui suggère fortement que ce gène code pour un nouveau type d' α1,4-fucosyltransférase chez *X. tropicalis*.

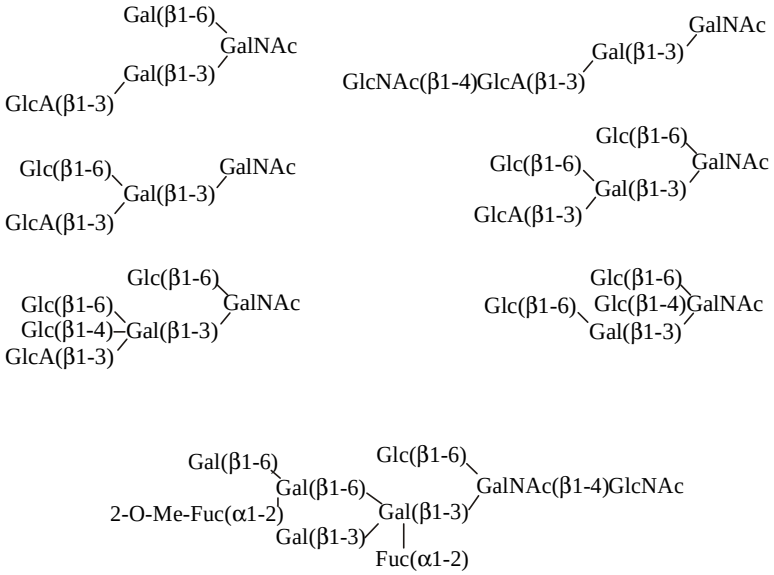
Le modèle Caenorhabditis elegans

A l'heure actuelle, *Caenorhabditis elegans* est l'animal le plus simple et le mieux caractérisé: son anatomie et son développement ont été totalement déterminés, et le séquençage de son génome est complet depuis plusieurs années. L'analyse du génome de *C. elegans* a révélé l'existence de nombreux gènes homologues à des glycosyltransférases et des transporteurs de nucléotides-sucres, dont certains codent des glycosyltransférases actives. Des gènes codant pour des polypeptides GalNAc-transférases (Hagen & Nehrke, 1998), GlcNAc-transférases (Chen *et al.*, 1999; Warren *et al.*, 2001) et fucosyltransférases (DeBose-Boyd *et al.*, 1998 ; Zheng *et al.*, 2002 ; Zheng *et al.*, 2008) actives ont été caractérisés. En effet, vingt deux gènes homologues à des α2-Fuc-transférases, cinq à des α3- et une à une α6- ont été identifiés dans son génome (Shachter *et al.*, 2004 ; Oriol *et al.*, 1999 ; Paschinger *et al.*, 2004 ; Paschinger *at al.*, 2005). De plus, son génome ne contient pas moins de neuf gènes homologues à la Gal-transférase des vertébrés responsable de la synthèse du corps 1 (Galβ1-3GalNAc) et dix huit à la GlcNAc-transférase responsable de la synthèse du corps 2 (Galβ1-3[GlcNAcβ1-6]GalNAc) (Ju *et al.*, 2002 ; Schachter *et al.*, 2004 ; Ju *et al.*, 2006). De fait certains auteurs estimaient qu'à l'exception de la sialylation, le génome de *C. elegans* coderait un répertoire de gènes capables d'élaborer un profil de glycosylation homologue à celui des mammifères (Dennis *et al.*, 1999). *C. elegans* apparaissait donc comme un modèle idéal

pour l'étude du rôle des glycoconjugués dans nombre de processus biologiques reliés au développement.

Ainsi, l'étude de la glycosylation de *C. elegans* s'intégrait parfaitement dans les objectifs de nos travaux, c'est à dire de donner un support structural à la détermination des activités glycosyltransférasiques isolées de modèles génétiquement bien définis. Lorsque nous avons entrepris cette étude, aucune structure de N- ou de O-glycanne n'avait encore été établie chez *C. elegans*. Par contre les structures de six glycosphingolipides appartenant à la série arthro-, trois neutres et trois acides, avaient été établies à partir de lysats totaux de *C. elegans* (Gerdt *et al.*, 1997; Gerdt *et al.*, 1999). Ces structures présentaient des homologies importantes avec le nématode parasite *Ascaris suum*, ce qui confortait la position de *C. elegans* en tant que modèle d'étude des nématodes parasites.

Pour notre part, nous nous sommes concentré sur l'étude de la structure des O-glycannes. D'une manière générale, nos travaux ont démontré que le nématode *C. elegans*, bien qu'il possède de nombreux gènes homologues à des gènes de mammifères codant des glycosyltransférases, et que les activités de plusieurs des produits de ces gènes sont similaires à celles de glycosyltransférases humaines connues, synthétise *in vivo* des O-glycannes de structures imprévisibles. Les structures de ces composés sont décrites ci-dessous :



En particulier, la plupart de ces structures sont caractérisées par la présence inhabituelle de résidus de β-Glc. A ce jour, ces composés représentent les seuls exemples connus de O-glycannes de types mucine incorporant du glucose. A ce titre, seule l'utilisation combinée de la RMN et de la spectrométrie de masse a pu permettre de mettre en évidence cette originalité. Ainsi, les quelques

structures que nous avons décrites démontrent l'existence de cinq nouvelles activités glycosyltransférasiques très particulières et imprévisibles: deux activités β-1,6-glucosyltransférasiques, deux activités β-1,3-glucosyltransférasiques et une activité N-acétylgalactosamine β-1,6-galactosyltransférasique.

Ces prédictions ont été en partie confirmées par une étude indépendante, dans laquelle les auteurs avaient mis en évidence l'existence chez *C. elegans* de six gènes apparentés (*gly-1* et *gly-15* à *-19*) dont les séquences présentaient des homologies importantes avec une famille de β-1,6-N-acétylglucosaminyltransférases de mammifères (Warren, C.E *et al.*, 2001). Cette famille est composée de plusieurs membres qui présentent des activités variées: la β6-GlcNAc-T(L) synthétise uniquement le noyau de type 2, la β6-GlcNAc-T(M) les noyaux de types 2, 4 et l'antigène I, tandis que la I β6-GlcNAc-T synthétise uniquement l'antigène I. Sur la base de comparaisons de séquences, ces gènes de *C. elegans* apparaissaient comme les meilleurs candidats pour coder des β6-GlcNAc-T actifs sur les O-glycannes. Néanmoins, il n'était pas possible de déterminer si ils codaient pour des homologues de la β6-GlcNAc-T(L), de la β6-GlcNAc-T(M) ou de la I β6-GlcNAc-T de mammifère. Sur la base de nos travaux, les auteurs ont réévalué les activités endogènes probables de ces enzymes. En particulier, en accord avec la présence du motif Gal(β1-3)[Glc(β1-6)]GalNAc dans les O-glycannes de *C. elegans*, ils ont démontré que *gly-1* transférait préférentiellement du glucose à partir de l'UDP-glucose sur un noyau de type 1 plutôt que de la N-acétylglucosamine (Warren *et al.*, 2002). Cet exemple démontre directement que l'activité endogène d'une telle glycosyltransférase inconnue peut être attribuée uniquement sur la base des substrats donneurs et accepteurs endogènes préalablement déterminés par une étude structurale. Ceci est d'autant plus vrai que des gènes présentant des homologies importantes peuvent coder des glycosyltransférases d'activités différentes, comme le démontre l'exemple cité.

Une deuxième étude a caractérisé une UDPGalNAc:GlcNAcβ-R β1,4-N-acétylgalactosaminyltransférase de *C. elegans* (Kawar *et al.*, 2002). Les deux substrats accepteurs exogènes préférentiels testés de cette enzyme sont GlcNAcβ-S-pNP et GlcNAc(β1-6)Gal. De fait les auteurs suggèrent que cette enzyme puisse être impliqué dans la synthèse du noyau très inhabituel R-GalNAc(β1-4)GlcNAc-Ser/Thr (LacdiNAc terminal) présent dans un des O-glycannes que nous avons décrits. Ceci est d'autant plus probable qu'à ce jour aucune autre séquence LacdiNAc n'a été décrite ni dans les O-glycannes, ni dans les N-glycannes de *C. elegans* (Paschinger *et al.*, 2008).

Enfin, hors du champ de l'étude des glycosyltransférases et de leurs activités, l'établissement du profil de O-glycosylation de *C. elegans* a également donné une base structurale

solide à l'établissement des phénotypes glycanniques de souches mutantes de nématodes. C'est en particulier le cas du mutant *sfr-3* chez qui il a été démontré que la résistance aux infections par *Microbacterium nematophilum* et à l'attachement du biofilm des bactéries *Yersinia pseudotuberculosis* et *Yersinia pestis* serait en partie le résultat de la perte sélective des O-glycannes chargés que nous avions décrits (Cipollo *et al.*, 2004).

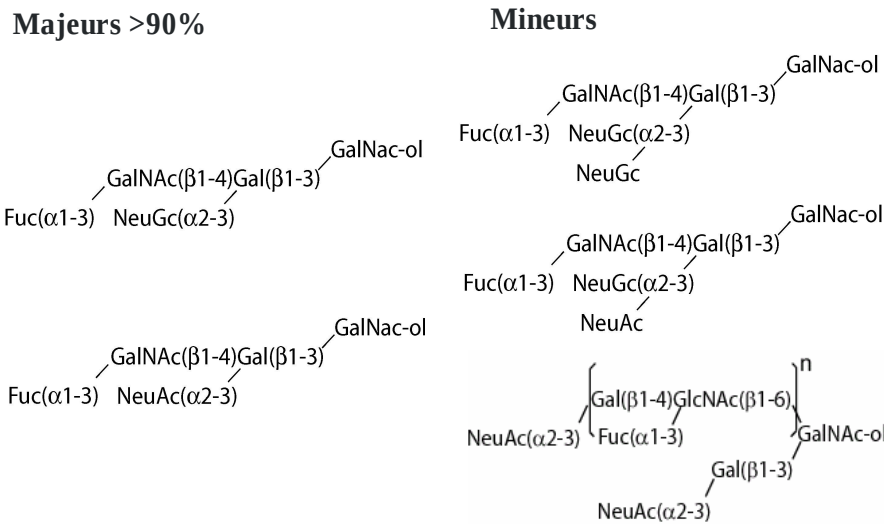
2.2- Développement de nouveaux modèles animaux d'étude de la sialylation

2.2.1- Poisson zèbre

Les travaux que nous avons réalisés sur les profils de glycosylation des xénopes et *de C. elegans* ont pleinement remplis leurs objectifs qui étaient de mettre en évidence de nouvelles activités glycosyltransférasiques et de servir de plateformes structurales solides à l'étude de la biosynthèse et des fonctions des glycoconjugués dans ces différents modèles. Néanmoins, au-delà de l'identification d'un gène codant pour un nouveau type de fucosyltransférase chez *X. tropicalis*, ces travaux n'ont pas été pérennisés au sein du laboratoire et n'ont donc pas dépassés pour nous le stade de l'analyse structurale. En particulier, les travaux sur *C. elegans* faisaient l'objet d'une collaboration avec le Dr R. Oriol et devaient se poursuivre par le clonage et l'étude des fucosyltransférases dans ce modèle, sur la base des prédictions bio-informatique (Oriol *et al.*, 1999). Malheureusement, ces travaux n'ont pu être réalisés suite au départ du Dr Balanzino qui gérait la plateforme d'élevage de *C. elegans*. Cependant, ces travaux ont été repris et sont en court de réalisation dans d'autres équipes (Zheng *et al.*, 2002 ; Zheng *et al.*, 2008).

Ainsi, nous recherchions un nouveau d'étude modèle du rôle de la glycosylation compatible avec les recherches en biologie du développement. C'est au cours de mon tage post-doctoral chez le Dr Khoo (Academia Sinica, Taipei) que j'ai eu l'opportunité d'aborder l'étude du poisson zèbre. En effet, le poisson zèbre (*Danio rerio*) est communément utilisé en tant que modèle d'étude du développement chez les vertébrés en raison de son développement embryonnaire rapide, de sa transparence optique et de la facilité avec laquelle il peut être manipulé *in vivo*. Ces qualités le prédisposent à devenir un bon modèle d'étude du rôle de la glycosylation chez les vertébrés. Néanmoins, l'étude de l'implication éventuelle de la glycosylation dans l'embryogenèse était jusqu'à présent rendue difficile par l'absence totale de données quant aux profils de glycosylation exprimés par cet organisme. Ainsi, en prélude à l'étude des relations entre structures et fonctions des glycoconjugués, nous avons entrepris de décrire l'évolution des profils de glycosylation au cours du développement du poisson zèbre. Dans un premier temps, les glycolipides, N-glycannes et

O-glycannes ont été séquentiellement purifiés d'extraits totaux de poissons zèbre et étudiés par une combinaison de spectrométrie de masse, de digestion enzymatique, de dérivation chimique et de résonance magnétique nucléaire (RMN). Nous avons ainsi mis en évidence l'existence de profils de glycosylation très particuliers dominés par des composés hautement sialylés (Guérardel *et al.*, 2006). La structure des O-glycannes identifiée est résumée ci-dessous :



Dans un deuxième temps, le suivi de la glycosylation au cours du développement a révélé que la polysialylation des glycoconjugués était différentiellement régulée en fonction de leur nature. En effet, alors que l'on observe une chute brutale de la polysialylation des glycoprotéines au cours des stades précoces du développement, la polysialylation des glycolipides augmente (Chang *et al.*, 2008). De plus, les séquences polysialylées présentent des différences structurales entre ces différents types de glycoconjugués : alors que les acides polysialiques des N- et O-glycannes sont majoritairement constitués d'acide N-glycolyl neuraminique, ceux des glycolipides sont constitués d'un mélange équimolaire d'acide N-glycolyl neuraminique et N-acétyl neuraminique. Ces observations démontrent l'existence de phénomènes extrêmement subtils de régulation de la glycosylation au cours du développement précoce et suggère un rôle important de ces composés.

Parallèlement à ces travaux, le Dr. Anne Harduin-Lepers au sein de notre institut a révélé l'existence d'une trentaine d'orthologues de gènes codant potentiellement des sialyltransférases au sein du génome de *Danio rerio* (Harduin-Lepers *et al.*, 2005). Sur la base de ces résultats, j'ai mis en place une collaboration internationale sur le thème du rôle de la sialylation au cours de l'embryogenèse du poisson zèbre. Outre notre équipe, elle comprend les équipes du Dr. A. Harduin-Lepers de l'UGSF, l'équipe du Dr K. Kitajima de l'Université de Nagoya (Japon) et l'équipe du Dr K.H. Khoo de l'Academia Sinica de Taipei (Taiwan). Dans le but de dynamiser ces travaux et de

me former à l'étude physicochimique des acides polysialiques, j'ai effectué un séjour de cinq mois (février à juin 2006) dans le laboratoire du Dr. Ken Kitajima, au cours duquel j'ai finalisé l'étude de la polysialylation du poisson zèbre. Dans le même cadre, le Dr K.H. Khoo et moi-même encadrons en cotutelle une étudiante en thèse taiwanaise qui étudie la régulation des polysialyltransférases chez le poisson zèbre en collaboration avec le Dr. Anne Harduin-Lepers. Dans le cadre de cette collaboration, nous avons suivi les profils d'expression de toutes les polysialyltransférases au cours du développement précoce pour déterminer l'implication individuelle des polysialyltransférases dans la biosynthèse de chaque classe de glycoconjugués (Chang *et al.*, 2008).

Glycomic survey mapping of zebrafish identifies unique sialylation pattern

Yann Guérardel[1,2,3], Lan-Yi Chang[2,4], Emmanuel Maes[3], Chang-Jen Huang[2,4], and Kay-Hooi Khoo[1,2,4]

[2]Institute of Biological Chemistry, Academia Sinica, Nankang, Taipei 11529, Taiwan; [3]Unité de Glycobiologie Structurale et Fonctionnelle, USTL, 59655 Villeneuve d'Ascq, France; and [4]Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan

Received on October 9, 2005; accepted on November 14, 2005

Functional genomics and proteomics studies of the developmental glycobiology of zebrafish are greatly hampered by the current lack of knowledge on its glycosylation profile. To furnish the requisite structural basis for a more insightful functional delineation and genetic manipulation, we have initiated a survey mapping of the possible expression of stage-specific glycoconjugates in zebrafish. High-sensitivity mass spectrometry (MS) analysis in conjunction with the usual array of enzymatic and chemical derivatization was employed as the principal method for rapid differential mapping of the glycolipids and sequentially liberated *N*- and *O*-glycans from the total extracts. We demonstrated that all developmental stages of the zebrafish under investigation, from fertilized eggs to hatched embryos, synthesize oligomannosyl types of *N*-glycans, as well as complex types with additionally β4-galactosylated, Neu5Ac/Neu5Gc monosialylated Lewis x termini. A combination of collision-induced dissociation (CID)-MS/MS and nuclear magnetic resonance (NMR) analyses led to the identification of an abundant and unusual mucin-type O-glycosylation, based on a novel sequence Fucα1-3GalNAcβ1-4(Neu5Ac/Neu5Gcα2-3)Galβ1-3GalNAc. This core structure may be further oligosialylated, but exclusively in the earlier development stages. Similarly, MS and MS/MS analyses of the extracted glycolipid fraction revealed the presence of a heterogeneous family of oligosialylated lactosylceramide compounds. In contrast to the *O*-glycans, these glycolipids only appear in the later development stages, suggesting a complex pattern of regulation for sialyltransferase activities during zebrafish embryogenesis.

Key words: glycomics/mass spectrometry/sialylation/structure analysis/zebrafish

Introduction

The zebrafish, *Danio rerio*, has emerged in recent years as an excellent model system to study the genetic underpinnings of vertebrate development. Large-scale genetic screens have identified thousands of mutant variants that allow *in vivo* dissection of developmental processes at single cell and molecular resolution. Among other advantages, it has a very short generation time which enables fast generation of transgenic lines. Its embryos are optically transparent, develop externally, and can absorb mutagens directly from surrounding water. The popularity of this model system has led to rapid accumulation of a large body of genetic data which provides a unique opportunity to follow the functional involvement of glycoconjugates in a vertebrate model throughout its complete embryogenesis. Such studies are increasingly accessible owing to recent development of tools designed to identify and classify enzymes involved in complex carbohydrate biogenesis based on sequence and folding similarities (Coutinho and Henrissat, 1999).

Strategies based on sequence similarities appear to be very successful in higher vertebrate models, including human, of which the glycosylation potentials are already well known. However, they show major shortcomings in the case of phylogenetically distant models because of the existence of unknown forms of glycosylation deriving from unpredictable enzymatic activities. In such cases, the definition of the fine specificity of glycosylation-related enzymes is rendered very delicate by the extreme variability of their activities. Such variability has been well illustrated by the modulation of FUT-3 substrate acceptor through the mutation of a single amino acid in its hypervariable stem (Dupuy *et al.*, 1999). The presence of unpredictable glycan structures has been identified in many animal models including those commonly used for developmental studies (Guerardel *et al.*, 2000, 2001, 2003; Haslam and Dell, 2003), which amply demonstrated the pertinence of integrated strategies combining genetic approach with direct observation of endogenous forms of glycosylation. Only then can the fine specificity of identified enzymes be truly assessed using endogenous substrates rather than commonly found glycan motifs.

In zebrafish, structural and functional studies of glycoconjugates have so far been focused only on very specific forms of glycosylation. In particular, the involvement of chitin oligosaccharides during zebrafish embryogenesis was strongly suggested by the inhibition of their biosynthesis (Bakkers *et al.*, 1997; Semino *et al.*, 1998; Semino and Allende, 2000). More recently, the existence of a specific receptor for chitin tetrasaccharide that would link its activity to the Raf, MEK, and ERK pathway in zebrafish cells was demonstrated (Snaar-Jagalska *et al.*, 2003). Other glycoconjugates such as the glycosaminoglycans (GAG) and polysialic acids were likewise shown to express a wide range of functions during zebrafish embryogenesis, including central nervous system (Bernhardt and Schachner, 2000; Marx *et al.*, 2001; Becker and Becker, 2002) and muscle development (Bink *et al.*, 2003). However, definitive structures of

[1]To whom correspondence should be addressed; e-mail: kkhoo@gate.sinica.edu.tw/yann.guerardel@univ-lille1.fr

glycolipids, N- or O-linked glycans have hitherto not been reported which precluded any functional studies on these essential components.

As a prelude to decipher the influence of glycosylation during zebrafish embryogenesis, we have initiated a systematic profiling of glycoconjugates at different developmental stages. The first aim was to define the structures of major glycoprotein-derived glycans and of glycolipids expressed by this organism, from which the endogenous activity of glycosylation-related enzymes, including glycosyltransferases, can be inferred. Mass spectrometry (MS)-based glycomics mapping were followed by more detailed analysis for the novel structural features thus identified. In parallel, comparative analyses were extended to the extracted glycomes from other developmental stages to critically evaluate their differential expression, especially in relation to the unique sialylation pattern.

Results

Our overall glycomic survey mapping strategy involved sequential extraction of glycolipids and glycoproteins and the subsequent sequential release of *N*- and *O*-glycans from the proteolytic-digested peptides/glycopeptides mixtures, for matrix-assisted laser-desorption ionization-MS (MALDI-MS) and MS/MS analyses. The released glycans were permethylated to allow more informative MS/MS sequencing, but native glycans were also analyzed where sample amount permitted. Such approach, in general, gives a good representative profile of the glycome but does not optimize for the yield of any particular class of glycoconjugates. It provides the first picture, uncovers any novel structural features, and facilitates subsequent more detailed investigations. Typically, the fertilized eggs at five distinctive developmental stages, 0.5, 8, 24, 45, and 48 h, were analyzed to allow a fair assessment of possible developmental regulation from first cell stage to hatching. Any significant differences were noted while common features were reported without distinguishing the origin of sample stage.

Identification of the major N-*glycans*

MALDI-MS profiling of the permethyl derivatives of *N*-glycans released from the total zebrafish embryo extracts afforded five major peaks at *m/z* 1579, 1783, 1988, 2193, and 2397, corresponding respectively, to sodiated molecular ions, $[M + Na]^+$, of the composition $Hex_{5-9}HexNAc_2$ (Figure 1A). Further collision-induced dissociation (CID)-MS/MS analysis and treatment with α-mannosidase (data not shown) demonstrated that these major signals are indeed the common high–mannose-type structures. In addition, several signals of lower intensity were visibly present among which a prominent cluster at *m/z* 3551, 3581, and 3611 could be tentatively assigned as $[M + Na]^+$ of $Neu5Ac_2Fuc_2Hex_7HexNAc_4$, $Neu5Ac_1Neu5Gc_1Fuc_2Hex_7HexNAc_4$, and $Neu5Gc_2Fuc_2Hex_7HexNAc_4$, respectively. These unusual compositions were shown by MALDI-MS/MS analyses to be biantennary complex-type structures with monosialylated Hex_2(Fuc)HexNAc sequence on both antennae (Figure 2). Both parent ions at *m/z* 3551 and 3611 afforded similar consecutive losses of terminal sialic acid residues and monosialylated antennae. Importantly, after losing both sialylated antennae, a common fragment ion at *m/z* 1143 was produced which corresponds to the sodiated trimannosyl core, $Man_3GlcNAc_2$, containing two free OH groups and thus confirming their biantennary nature.

For the $Neu5Ac_2$-containing parent (*m/z* 3551), a primary sodiated b ion, $Neu5Ac_1Hex_2$(Fuc)HexNAc, was detected at *m/z* 1225, accompanied by a $Neu5Ac^+$ oxonium ion at *m/z* 376 and a sodiated c ion, Neu5Ac-Hex_2-OH, at *m/z* 824 (Figure 2B). In comparison, the $Neu5Gc_2$-containing parent (*m/z* 3611) afforded a sodiated b ion, $Neu5Gc_1Hex_2$(Fuc)HexNAc, at *m/z* 1255, a $Neu5Gc^+$ ion at *m/z* 406, and a sodiated c ion, Neu5Gc-Hex_2-OH, at *m/z* 854 (Figure 2C). Further loss of the distinguishing Neu5Ac/Neu5Gc residue from the respective primary b and c ions degenerated the mass difference and yielded the common secondary fragment ions at *m/z* 850 and 449, corresponding to $(HO)_1Hex_2$-$(Fuc)_1$HexNAc and $(HO)_1Hex_2$-OH, respectively. The mass difference of 60 u between the two parents could thus be unambiguously attributed to a Neu5Ac and Neu5Gc difference (30 u) on each of the two monosialylated antennae. It could be further deduced that the third molecular ion signal (*m/z* 3581) in between the $Neu5Ac_2$- and $Neu5Gc_2$-containing parents corresponds to a similar biantennary complex-type structure carrying a Neu5Ac- and a Neu5Gc-sialylated antennae.

Higher in mass (Figure 1A) and of even lower abundance was another cluster of molecular ion signals which could be assigned as trisialylated triantennary complex-type structures with similar monosialylated terminal sequence carrying the Neu5Gc/Neu5Ac heterogeneity. Thus the signal at *m/z* 4830 corresponds to species with all three Hex_2(Fuc)HexNAc antennae sialylated by Neu5Gc, whereas the one at *m/z* 4740 carries only Neu5Ac-Hex_2(Fuc)HexNAc antenna. Supporting data were obtained when after desialylation by neuraminidase, a peak at *m/z* 3656 was detected which corresponds to $[M + Na]^+$ of a triantennary complex-type structure with three Hex_2-$(Fuc)_1$HexNAc antennae. The corresponding desialylated biantennary structure was observed as a major sodiated molecular ion at *m/z* 2827 (Figure 1B). Interestingly, after an overnight (>12 h) digestion, Neu5Ac desialylation appeared to be more complete than removal of Neu5Gc. Additional molecular ions corresponding to incompletely digested mono-Neu5Gc sialylated bi- and triantennary structures were detected at *m/z* 3218 and 4048, respectively (Figure 1B), but not their mono-Neu5Ac-sialylated counterparts which could only be observed if the neuraminidase digestion was kept to a shorter period (data not shown).

In accordance with MS/MS sequencing of the sialylated counterparts described above, MALDI-MS/MS on the desialylated biantennary structure yielded prominent nonreducing terminal primary fragment ions at *m/z* 463 and 864 (Figure 2A), corresponding to sodiated c ion, Hex_2-OH, and b ion, Hex_2-(Fuc)HexNAc, respectively. The Fuc substitution could be deduced as 3-linked to the HexNAc based on the detection of the secondary ions produced through the elimination of the Fuc (minus 206 mass units from the parent and other major primary fragment ions), whereas the

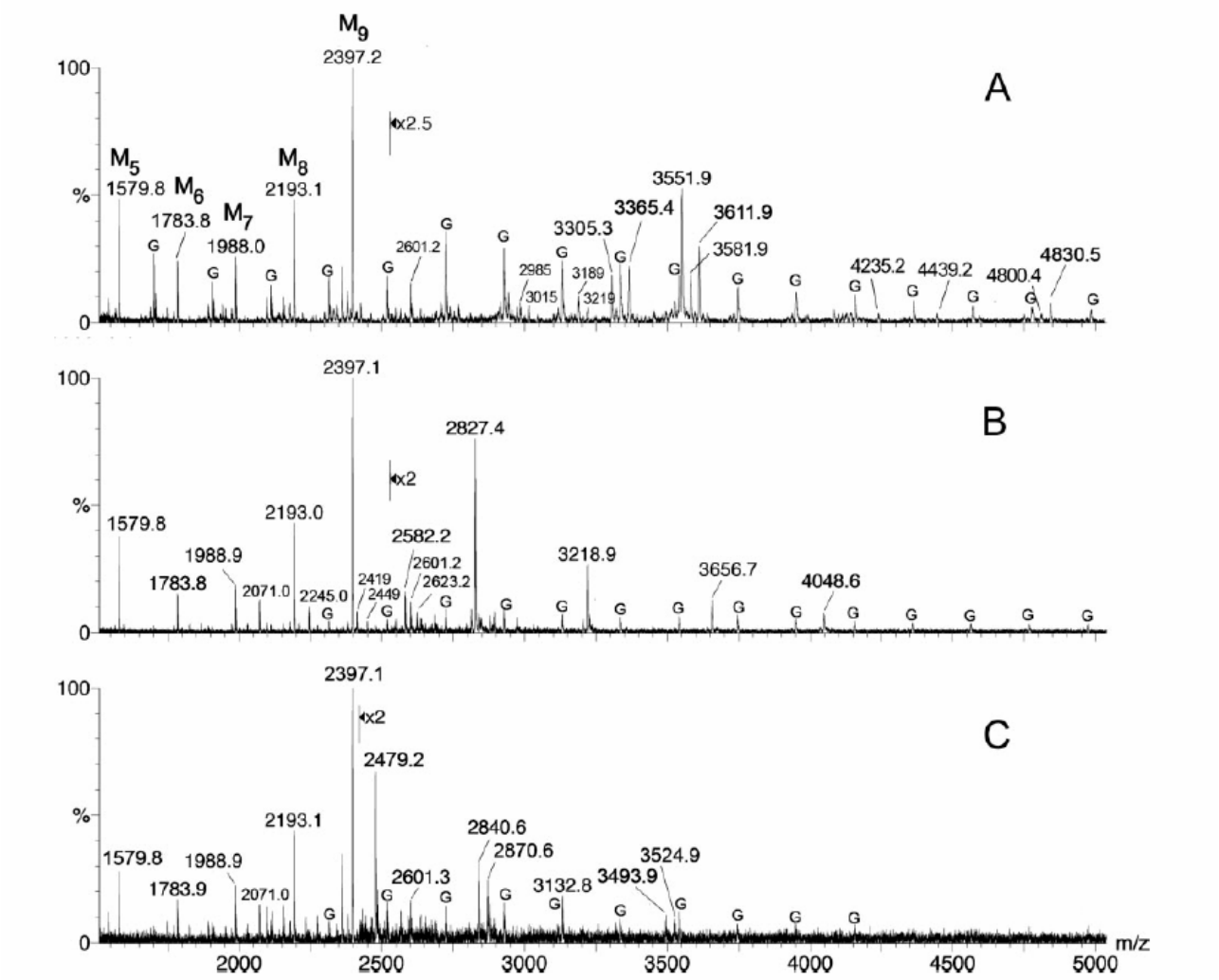


Fig. 1. MALDI-MS profiles of the permethylated *N*-glycans from zebrafish embryos before (**A**) and after neuraminidase (**B**) or after aqueous hydrofluoric acid (**C**) treatment. High–mannose-type structures were labeled M5–M9 in (**A**), representing $Man_{5-9}GlcNAc_2$ structures. Signal at *m/z* 2601 most likely corresponds to $Glc_1Man_9GlcNAc_2$. Glucose oligomer contaminants were labeled G. In panel **B**, additional minor signals corresponding to nonsialylated biantennary structures with incomplete fucosylation and/or galactosylation were also detected (*m/z* 2071, 2245, 2419, 2449, 2623), the smallest of which at *m/z* 2071 could be assigned as nonfucosylated biantennary *N*-glycan with simple Gal-GlcNAc termini. In panel **C**, complete defucosylation and desialylation produced the bi- (*m/z* 2479) and triantennary (*m/z* 3132) structures carrying $Hex_2HexNAc$ termini, accompanied by their mono-Neu5Ac/Neu5Gc sialylated counterparts at *m/z* 2840/2870 and 3492/3524, respectively. Under the conditions employed, aqueous hydrofluoric acid would remove α2,3,4-Fuc almost completely but sialic acid only partially. For simplicity, other minor products corresponding to a combination of incomplete galactosylation and desialylation were not labeled.

elimination of ±Neu5Ac/Neu5Gc-Hex_2 was not observed. Further confirmation was obtained when the desialylated structures were digested with β4-galactosidase. MALDI-MS and MS/MS analyses demonstrated that one Hex was removed from each of the nonreducing termini, whereas prior defucosylation with aqueous hydrofluoric acid afforded bi- and triantennary structures with Hex_2-HexNAc termini (Figure 1C) which could then be completely degalactosylated by β4-galactosidase (data not shown). The failure to remove the internal Gal attached to a fucosylated GlcNAc is consistent with the well-known selectivity of the β4-galactosidase acting on a Gal-(Fuc)GlcNAc unit. Linkage analysis on the isolated sialylated structures (see *Materials and methods*) further showed that the amount of terminal Gal relative to 2-linked Man or 3,6-linked Man did not change significantly before and after desialylation. Moreover, mono-substituted Gal residue was not detected. Instead, 3,4-linked Gal was quantitatively converted to 4-linked Gal after desialylation, therefore indicating that the sialic acid was attached to the 3-position of an internal 4-linked Gal and not to the terminal Gal.

Taken together, the data unambiguously defined the common monosialylated terminal sequence on each antennae of the major bi- and triantennary complex-type

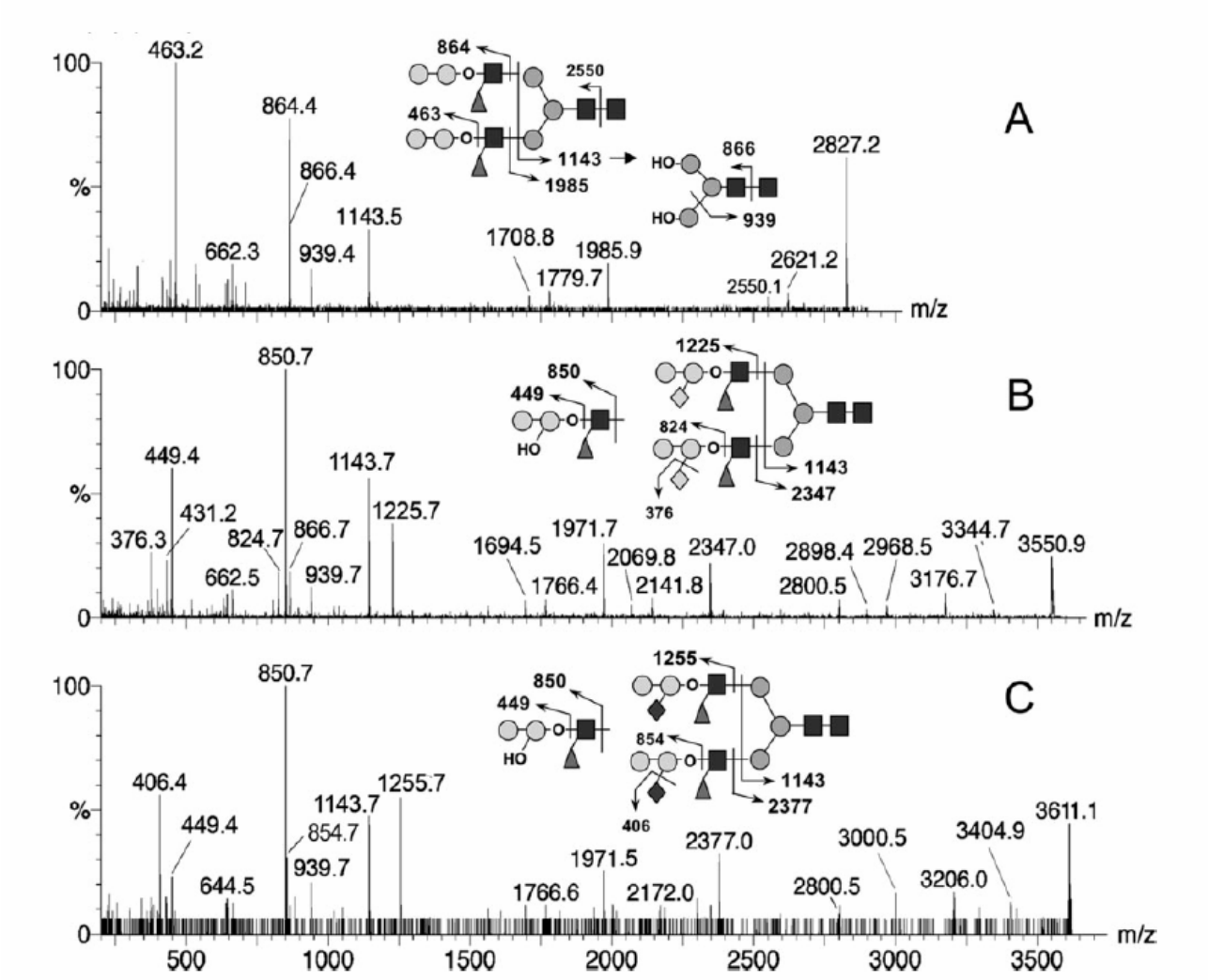


Fig. 2. MALDI-MS/MS sequencing of the biantennary complex-type *N*-glycans from zebrafish embryos. (**A**) MS/MS on the desialylated parent at *m/z* 2827; (**B**) MS/MS on the Neu5Ac-disialylated parent ion at *m/z* 3551; (**C**) MS/MS on the Neu5Gc-disialylated parent ion at *m/z* 3611. Major fragmentation pattern is as indicated schematically. Cleavage ions are mostly b and y ions except when indicated with an oxygen atom which correspond to c ions. All three gave common trimannosyl core ions at *m/z* 1143 as shown in panel A which yielded additional common fragment ions (*m/z* 939, 866, 662). Elimination of Fuc (minus 206 u) from parent and primary fragment ions are commonly observed, whereas loss of terminal Neu5Ac yielded signals at *m/z* 3176, 2800 (from 3550), and 1971 (from 2347) in panel **B**; loss of terminal Neu5Gc gave signals at *m/z* 3206, 2800 (from 3611), and 1971 (from 2377) in panel **C**. Symbols used are square, HexNAc; circle, Hex; diamond, Neu5Gc (dark) and Neu5Ac (light); triangle, Fuc. OH denotes exposed hydroxyl group because of cleavage on the permethyl derivatives.

N-glycan structures as Galβ1-4(Neu5Gc/Neu5Acα2-3)Galβ1-4(Fucα1-3)GlcNAc, namely an internal Lewis x unit which was further galactosylated and sialylated with either Neu5Gc or Neu5Ac. A very small amount of incomplete sialylation could be detected (Figure 1A) as monosialylated bi- (*m/z* 3189/3219) and disialylated triantennary structures (*m/z* 4439), as well as species that lack both sialic acid and Gal residues on one of the antennae (*m/z* 2985/3015; 4235). Curiously, the major disialylated biantennary structures were found to occur also as minor species lacking the reducing terminal GlcNAc, giving sodiated molecular ion signals at *m/z* 3305/3335/3365 (Figure 1A). CID MS/MS analysis of its desialylated counterpart (*m/z* 2582; Figure 1B) firmly established that the same antennary sequence is carried on the implicated $Hex_3HexNAc_1$ core in place of the usual $Man_3GlcNAc_2$ for *N*-glycans (data not shown).

Identification of the major O*-glycans*

MALDI-MS analysis of the permethylated *O*-glycans, released from de-N-glycosylated peptides as oligoglycosyl alditols through reductive elimination, afforded two predominant molecular ion signals at *m/z* 1315 and 1345. Other barely detectable weak signals at higher *m/z* values became more apparent only after enrichment by stepwise elution on an anion exchange column (Dowex 1 × 2 anionic resin). Thus, an early eluting fraction was found to contain only the two major signals, whereas a higher salt-eluted

fraction carried additional minor signals at higher mass range (Figure 3A), including the pair at *m/z* 1706/1736 which was related to the *m/z* 1315/1345 pair by a Neu5Gc increment. Assignment of the corresponding compositions and sequences was afforded by MALDI CID-MS/MS analyses, as schematically shown in Figure 3B–D.

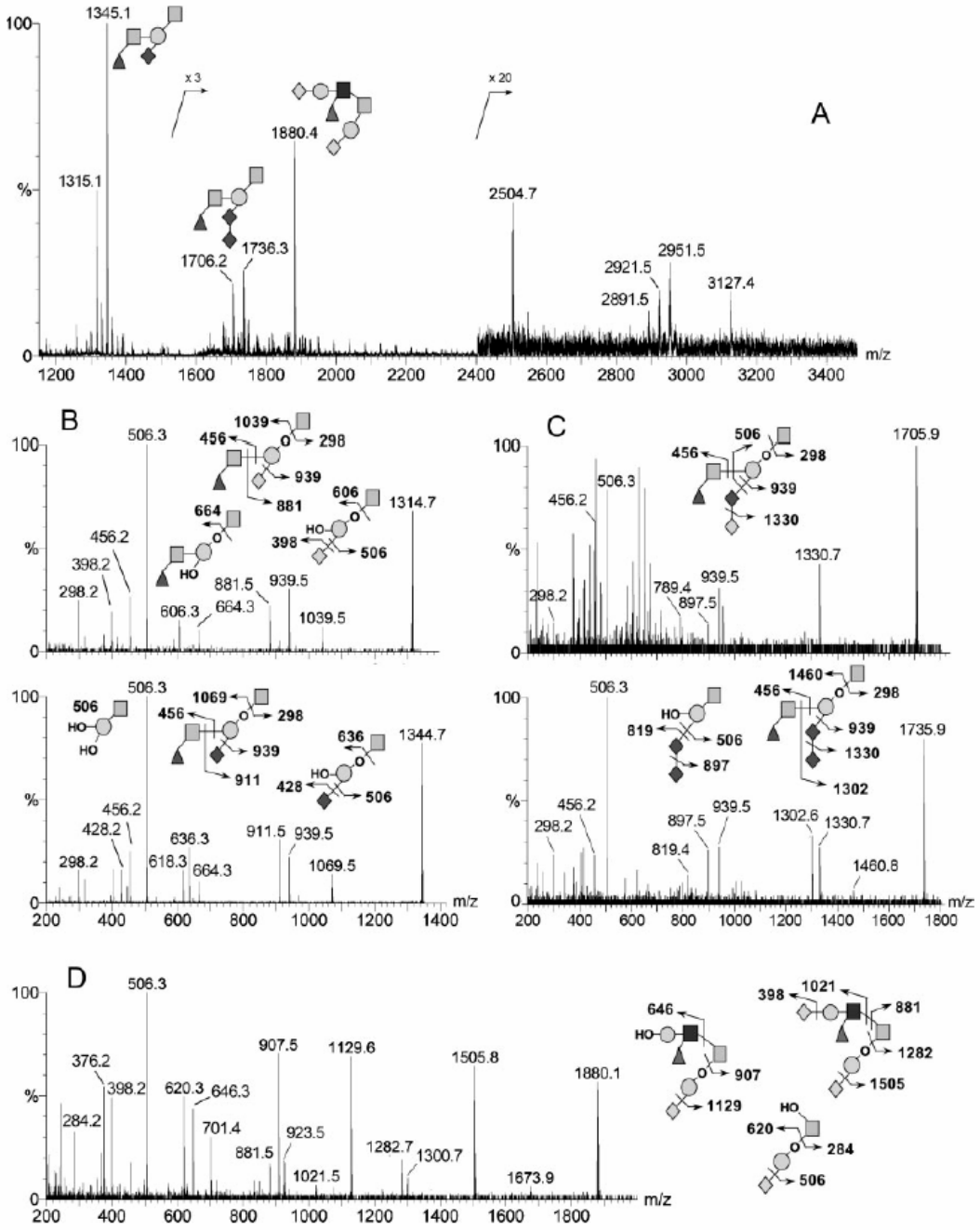


Fig. 3. MALDI-MS profile of permethylated *O*-glycans from zebrafish embryos (**A**) and MS/MS analyses on the major signals detected (**B–D**). MS profile shown in panel A was that on higher salt-eluted fraction from anion exchange chromatography. Nonenriched and earlier eluting fraction both afforded only the two major peaks at *m/z* 1315/1345 (not shown). For clarity, fragmentation pattern and the origin of major fragment ions are schematically indicated on structures corresponding to both parent and primary fragment ions. As in Figure 2, cleavage ions are mostly b and y ions except when indicated with an oxygen atom which correspond to c and z ions. OH denotes exposed hydroxyl group because of cleavage on the permethyl derivatives. All reducing end HexNAc of the released *O*-glycans is HexNAcitol by implication. The MS/MS spectrum of *m/z* 1705 (**C**, upper panel) contained several abundant low mass ions clearly not derived from the parent. Symbols used are square, HexNAc; circle, Hex; diamond, Neu5Gc (dark) and Neu5Ac (light); triangle, Fuc.

Notably for the parent ions at *m/z* 1315 and 1335 (Figure 3B), a core 1-type structure was indicated by the common z_1 ion at *m/z* 298, consistent with a mono-substituted reducing end HexNAcitol. The presence of a nonreducing terminal Neu5Ac/Neu5Gc was supported by the respective b_1 ions at *m/z* 398 (Neu5Ac) and 428 (Neu5Gc), as well as their facile loss from either the parent ions or the c ions at *m/z* 1039/1069 which degenerated the 30 u mass difference, giving rise to the common ions at *m/z* 939 and 664, respectively. Another distinctive set of b and y ion pair at *m/z* 456 and 881/911 defined a nonreducing terminal Fuc-HexNAc moiety, thus completing a rather unique sequence of (Fuc-HexNAc)(Neu5Ac/Neu5Gc)Hex-HexNAcitol. Localization of the terminal Neu5Ac/Neu5Gc to an internal Hex was supported by the common y ion at *m/z* 506 which corresponds to $(OH)_2$Hex-HexNAcitol, and the ions at *m/z* 606/636 corresponding to $(OH)_1$(Neu5Ac/Neu5Gc)$_1$Hex-OH.

Both the parent ions at *m/z* 1706 and 1736 likewise afforded a prominent fragment ion at *m/z* 506, which together with the z_1 ion at *m/z* 298, clearly indicated a similar core 1-type structure branched at the internal Hex (Figure 3C). For the *m/z* 1736 parent which gave a better quality MS/MS data, direct loss of either a single terminal Neu5Gc or a dimeric Neu5Gc-Neu5Gc yielded the y ions at *m/z* 1330 and 939, respectively. This was accompanied by the corresponding b ion for Neu5Gc-Neu5Gc at *m/z* 819 which firmly established the presence of a Neu5Gc-disialylated sequence. Importantly, the characteristic b ion at *m/z* 456 defined a similar nonreducing terminal Fuc-HexNAc moiety, complemented by the pairing y ion at *m/z* 1302 which corresponds to $(OH)_1$(Neu5Gc-Neu5Gc)Hex-HexNAcitol. Further loss of a terminal Neu5Gc from the latter yielded the ion at *m/z* 897, whereas losing both Neu5Gc residues gave the aforementioned $(OH)_2$Hex-HexNAcitol at *m/z* 506. The MS/MS spectrum of the other parent at *m/z* 1706 was of inferior quality and did not afford a full range of fragment ions. Nevertheless, apart from the common ions at *m/z* 456 and 506 which established its basic core structure, direct loss of either a terminal Neu5Ac (*m/z* 1330) or a dimeric Neu5Ac-Neu5Gc (*m/z* 939), but not a terminal Neu5Gc, indicated that its internal Hex was extended by a Neu5Gc and terminating with a Neu5Ac, contrasting with the other disialylated structure which carries a Neu5Gc-Neu5Gc extension. The detection of the corresponding b ion at *m/z* 789 for Neu5Ac$_1$Neu5Gc$_1$ and not *m/z* 819 for Neu5Gc$_2$ further supported the assignment. Thus, the disialylated structures may be considered as carrying an extra Neu5Ac or Neu5Gc extension from a Neu5Gc monosialylated structure but not a Neu5Ac monosialylated structure, all of which share the same basic Fuc-HexNAc-Hex-HexNAcitol backbone.

Monosaccharide composition analysis of the purified monosialylated fraction revealed an almost equal relative amount of Gal, Fuc, GalNAc, GalNAcitol, and sialic acid, fully consistent with the assigned sequence. To elucidate the complete primary structures of the two oligoglycosyl alditols, the monosialylated fraction was analyzed as a mixture by 400 MHz ^{1}H-nuclear magnetic resonance (^{1}H-NMR) spectroscopy. Chemical shifts of the protons of individual constituents obtained from sequential ^{1}H-^{1}H homonuclear correlation spectroscopy-90 (COSY-90) and total correlation spectroscopy (TOCSY) experiments were compiled in Table I. Spin systems of monosaccharides confirmed the results obtained from composition analysis. In particular, as shown on TOCSY spectrum (Figure 4A), the internal HexNAc residue was unambiguously identified as a GalNAc owing to its spin system, reminiscent of the galacto configuration. The presence of both Neu5Ac and Neu5Gc was easily assessed owing to the observation of intense *N*-acetyl and *N*-glycolyl protons as singlets at δ2.045 and 4.122 p.p.m., respectively. Differential integration of these signals afforded an estimated Neu5Gc/Neu5Ac ratio of about 8:2. However, chemical shifts of their respective H-3ax, H-3eq, and H-5 could not be distinguished from one another.

Table I. ^{1}H-NMR chemical shifts of monosialylated *O*-glycans

	Chemical shifts of ^{1}H (p.p.m.)						
Residue	1	2	3	4	5	6	NAc/NGc
GalNAc-ol I	3.79/3.76	4.38	4.064	3.531	4.148	3.64	–
Gal(β1-3) II	4.571	3.416	4.182	4.117	3.78	nd	–
GalNAc(β1-4) III	4.791	4.052	3.705	3.964	nd	nd	–
Fuc(α1-3) F	5.088	3.702	3.903	3.812	4.141	1.206	–
Neu5Gc(α2-3)	–	–	1.951/2.705	3.873	3.655	–	4.122
Neu5Ac(α2-3)	–	–	1.951/2.705	3.873	3.655	–	2.045

nd, not determined.

As expected, the chemical shifts of H-2 and H-5 of GalNAc-ol are clearly indicative of mono substitution of GalNAc-ol by Gal in β1-3 linkage (Kamerling and Vliegenthart, 1992), consistent with a core 1-type structure deduced by CID/MS-MS analyses. The chemical shift of GalNAc H-1 at δ4.791 p.p.m. in conjunction with its coupling constant $J_{1,2}$ of about 8.5 Hz established the GalNAc III residue as β-anomer. In accordance with previous work (Herkt *et al.*, 1985; Coppin *et al.*, 2002), the chemical shifts of Gal, GalNAc, and Neu5Ac/Neu5Gc residues are indicative of the sequence GalNAcβ1-4(Neu5Acα2-3)Galβ1-, known as the Cad determinant. In particular, the downfield position of H-3ax of Neu5Ac or Neu5Gc to δ1.951 p.p.m. is very specific to this motif (Mourad *et al.*, 2001). However, the chemical shift values of GalNAc protons showed significant discrepancies from those of Cad determinant, in particular, its H-2 signal that was noticeably deshielded by ~0.15 to δ3.903 p.p.m., suggesting that the GalNAc residue was further substituted. Substitution of GalNAc residue was assessed by the observation of GalNAc ^{13}C chemical shifts owing to heteronuclear ^{13}C-^{1}H heteronuclear multiple quantum coherence (HMQC) experiment (data not shown). In particular, the GalNAc C-3 signal was strongly deshielded to δ79.1 p.p.m. compared with C-4 signal at δ70.0 p.p.m., which indicated that only C-3 position bore further substitution. Accordingly, nuclear Overhauser effect spectroscopy (NOESY) experiment showed an intense NOE contact between Fuc H-1 at δ5.088 p.p.m. and GalNAc III H-3 at δ3.705 p.p.m. (Figure 4B), confirming the terminal Fucα1-3GalNAc sequence. It also corroborated

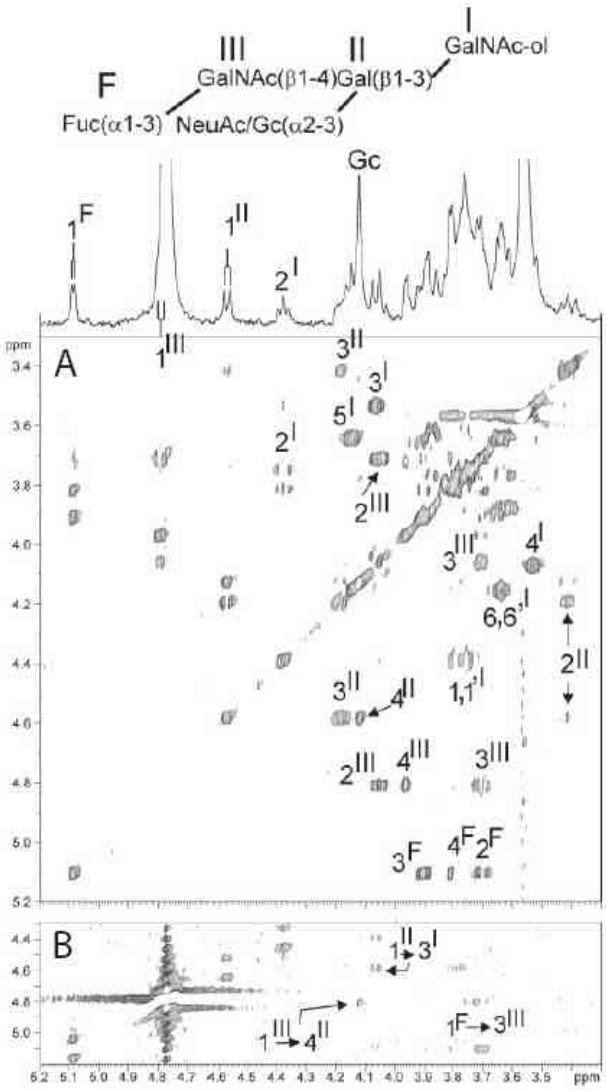


Fig. 4. Four hundred megahertz NMR analysis of monosialylated *O*-glycans purified from the total *O*-glycans by anion exchange chromatography. (**A**) Two-dimensional ^{1}H-1H TOCSY spectrum (δ^1H: 3.3–5.2 and 3.3–5.2) that allows the observation of the spin systems of residues II, III, and F. The presence of Neu5Gc and Neu5Ac are clearly established owing to intense peaks of *N*-glycolyl protons δ4.122 p.p.m. and *N*-acetyl protons at δ2.045 p.p.m. (not shown). (**B**) Two-dimensional ^{1}H-1H NOESY spectrum (δ^1H: 3.3–5.2 and 4.3–5.2) showing the connectivities between residues F, III, II, and I.

the assigned GalNAc1-4Gal1-3GalNAc-ol backbone sequence owing to clear NOE contacts between GalNAc H-1 and Gal H-4 and between Gal H-1 and GalNAc-ol H-3 at δ4.571 and 4.064 p.p.m., respectively.

In summary, the monosaccharide composition, MS/MS and NMR data collectively and unambiguously defined the two major monosialyated *O*-glycans as: Fucα1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-3GalNAc-itol and Fucα1-3GalNAcβ1-4(Neu5Gcα2-3)Galβ1-3GalNAc-itol. At a much lower level, the latter, but not the former, can be further extended by an additional Neu5Gc or Neu5Ac. At even higher mass, another cluster of extremely weak signals at *m/z* 2891, 2921, and 2951 apparently also exhibited the Neu5Gc/Neu5Ac heterogeneity of 30 u apart (Figure 3A), but no CID-MS/MS could be successfully obtained to establish their sequence. In contrast, the signals at *m/z* 1880, 2504, and 3127 did not afford the Neu5Gc/Neu5Ac sialylation pattern and apparently constitute a different series of *O*-glycan structures.

MALDI-MS/MS on the parent ion at *m/z* 1880 (Figure 3D) afforded the same fragment ion at *m/z* 506, corresponding to a disubstituted Hex-HexNAcitol core. However, instead of z_1 ion at *m/z* 298, *m/z* 284 was detected and therefore indicating that branching is at the reducing end HexNAcitol. Terminal Neu5Ac substitution was evident from the protonated and sodiated b_1 ions at *m/z* 376 and 398, respectively, whereas consecutive losses of two terminal Neu5Ac (*m/z* 1505 and 1129) from the parent suggested a disialylated structure with two different monosialylated termini. The c and z ion pair at *m/z* 620 and 1282, together with the other y_1 ion at *m/z* 881 (see the schematic drawing on Figure 3D) is consistent with a Neu5Ac-Hex motif attached to the 3-arm of a 3,6-branched HexNAcitol and a Neu5Ac$_1$-(Fuc$_1$Hex$_1$HexNAc$_1$)- motif on the 6-arm. The latter was supported by observing the corresponding b ion at *m/z* 1021 and after losing the terminal Neu5Ac at *m/z* 646. In the absence of other data owing to lack of sample material, the exact structure for the fucosylated motif could not be established although facile elimination of a 3-linked Fuc (minus 206 u) from the parent and several other ions was strongly indicative of a Lewis x epitope. This is also consistent with it being extended by another one or two such -[Hex(Fuc)HexNAc]- repeats on the 6-arm to give the higher mass molecular ions detected at *m/z* 2504 and 3127 (Figure 3A), corresponding to a sialylated poly-Lewis x sequence.

As in the case with the *N*-glycans, the two major monosialylated *O*-glycans characterized were consistently detected across all five developmental stages of the zebrafish embryos. However, larger *O*-glycans, either based on the same core 1-type structure but disialylated on the Neu5Gc appendage or based on a distinct branched core 2-type structure with Neu5Ac monosialylation on both arms, were in general of very low abundance and only readily detectable in earlier stages (0.5 and 8 h). Similar enrichment on anion exchange column failed to yield any signal for samples derived from the 45 and 48 h embryonic stages while there was some batch-to-batch variation for detecting their presence in the 24 h samples. Thus although their low abundance and heterogeneity in structures precluded firm conclusion with respect to their exact structures and developmentally regulated expression, oligosialylation on *O*-glycans appeared to be preferentially associated with early development, before the completion of morphological differentiation.

MS profiling of the glycolipids

In contrast to oligosialylation on the *N*- and *O*-glycans, MALDI-MS profiling of the permethyl derivatives of glycolipids extracted from various developmental stages of zebrafish embryos revealed that potential molecular ion signals corresponding to oligosialylated glycosphingolipids were only present at the later stages (Figure 5). Some minor

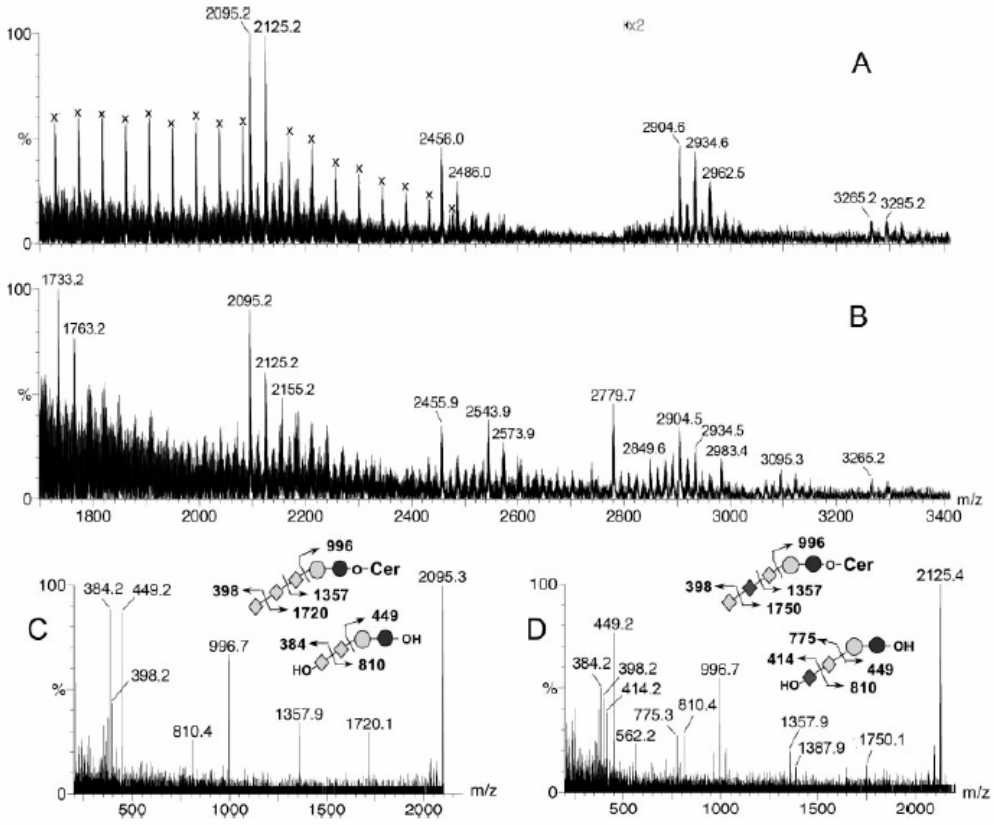


Fig. 5. MALDI-MS and MS/MS analyses of permethylated glycolipids from zebrafish embryos. Oligosialylated glycolipids related to the series at *m/z* 2095/2125 were only apparent in the profiles of samples from 24 h (**A**) and 45–48 h (**B**) after the fertilization. MS profiles of earlier stages (not shown) afforded other weak signals that may be attributed to glycan-containing species which were not further characterized in this study because of low abundance. MS profiles of the 45 and 48 h samples are almost identical, and only the latter was shown in panel **B**. Signals marked with x in panel A are contaminants. The two major sodiated molecular ions at *m/z* 2095 and 2125 from panel A were selected for MS/MS analyses, and the derived trisialylated sequences were schematically shown in panels **C** and **D**, respectively, along with the major fragmentation pattern observed. All fragment ions are of b and y ions. Signals at *m/z* 2543/2573 in panel **B** are a Hex-HexNAc increment from the trisialylated LacCer at *m/z* 2095/2125. Signals at *m/z* 2779 and 2983 are related by a Hex, but their exact sequences were not established. Symbols used are circle, Glc (dark) and Gal (light); diamond, Neu5Gc (dark) and Neu5Ac (light); Cer, ceramide; OH, exposed hydroxyl group because of primary cleavage.

variations between the profile of 24 h (Figure 5A) and those of 45–48 h (Figure 5B) were noted, but the salient features were mostly conserved. Among the weak signals, a most prominent cluster occurred at *m/z* 2095 and 2125. MALDI-MS/MS sequencing of the parent ion at *m/z* 2095 (Figure 5C) revealed a consecutive loss of Neu5Ac residues. Only the first Neu5Ac loss corresponds to loss of a fully methylated terminal residue, whereas subsequent losses of the other two Neu5Ac correspond to further cleavage of internal Neu5Ac residues, as distinguished by their distinctive residual mass values. In addition to the linear stretch of a $Neu5Ac_3$ sequence thus established, the fragment ions at *m/z* 449 and 810 could be assigned as sodiated $(OH)_1$Hex-Hex-OH and $(OH)_1$Neu5Ac-Hex-Hex-OH, respectively, consistent with a direct attachment of the $Neu5Ac_3$ terminal sequence to a lactosylceramide (LacCer). Based on the *m/z* values of the sodiated molecular ion and the deduced glycosyl sequence, the ceramide moiety could be calculated as corresponding to a d18:1 base with a C16:0 fatty acyl chain or their equivalent permutation thereof.

For the parent ion at *m/z* 2125, MS/MS sequencing (Figure 5D) demonstrated that one of the internal Neu5Ac in the $Neu5Ac_3$-LacCer was replaced by a Neu5Gc. Significantly, only loss of terminal Neu5Ac and not Neu5Gc was observed, giving the y ion at *m/z* 1750. This could be followed by further loss of an internal Neu5Gc (*m/z* 1357) and then an internal Neu5Ac to yield the sodiated $(OH)_1$LacCer ion at *m/z* 997. Additional sodiated ions at *m/z* 449 and 810 indicated an internal -Neu5Ac-Hex_2- sequence, therefore suggesting a unique sequence of Neu5Ac-Neu5Gc-Neu5Ac-LacCer. Alternative arrangement of the $Neu5Ac_2Neu5Gc_1$ sequence on the LacCer was, however, not ruled out nor stage and/or batch variations be investigated due to its low abundance. At one Neu5Ac residue smaller, the molecular ion signals at *m/z* 1733 and 1763 could be tentatively assigned as $Neu5Ac_2$-LacCer and $Neu5Ac_1Neu5Gc_1$LacCer, respectively, whereas signals at *m/z* 2456 and 2486 are consistent with $Neu5Ac_4$-LacCer and $Neu5Ac_3Neu5Gc_1$-LacCer, respectively. Other weak signals that may be tentatively assigned include the clusters

at *m/z* 2904/2934 which correspond to a Hex-HexNAc increment from the tetrasialylated LacCer, and signals at *m/z* 3265/3295 which carried an additional sialic acid. In all cases though, the low abundance amid contaminant noise peaks, compounded by additional heterogeneity which may be contributed by Neu5Ac/Neu5Gc differences and/or the lipid moiety, precluded firm definition of the entire spectra of the glycolipids.

Discussion

A glycomic analysis aims, in general. to define the glycosylation potential of a biological source under investigation. As applied to lower organism such as the zebrafish for which current knowledge on its glycobiology is scarce, a MS approach offers several distinctive advantages. Of prime consideration, the detection and tentative compositional assignment, including *de novo* sequencing, is not dependent on standard references and hence more conducive to identification of novel structures than any other methods. In this context, MALDI-MS mapping coupled with facile CID MS/MS sequencing on the permethyl derivatives is by far the most informative and sensitive analytical strategy although not without its limitations. As a first attempt, we have successfully derived an overall picture of the zebrafish glycome, as presented on both the glycoproteins and the glycolipids, but have excluded analysis on the GAG or the chitin oligosaccharides. Our collective results show that a most striking feature is the diverse oligosialylation pattern which appears to be developmentally regulated.

The high–mannose-type *N*-glycans are the only nonsialylated population of the glycome that occurs at any abundance. Otherwise, both the complex-type *N*-glycans and the *O*-glycans are each predominantly represented by a single unique terminal sequence, monosialylated with either Neu5Ac or Neu5Gc. For the *N*-glycans, the Galβ1-4(Neu5Gc/Neu5Acα2-3)Galβ1-4(Fucα1-3)GlcNAc nonreducing terminal sequence constitutes the antenna of the major bi- and triantennary structures, along with some minor degrees of incomplete sialylation and/or galactosylation on the internal Lewis x epitope. For the *O*-glycans, a core 1-type structure was identified which carries a Fucα1-3GalNAcβ1-4(Neu5Gc/Neu5Acα2-3)Galβ1-3GalNAc sequence. In both cases, α2-3-sialylation was found on an internal β-Gal. The zebrafish sequences can be distinguished from those of more commonly found mammalian type by virtue of either an additional β4-Gal extension on a sialyl Lewis x or an α3-fucosylated β4-GalNAc extension on a sialylated core 1 *O*-glycan. On the other hand, they bear much similarity to other characterized fish glycans (reviewed in Inoue and Inoue, 1997).

The Galβ1-4Galβ1-4GlcNAc motif was first identified on the complex-type free sialoglycans released from the glycophosphoproteins of unfertilized eggs of *Tribolodon hakonensis* and *Oryzia latipes* (Inoue *et al.*, 1989; Iwasaki *et al.*, 1992). α2-3-Neu5Ac monosialylation was found to occur on either the terminal or the internal Gal, giving monosialylated antenna for the predominantly bi- and triantennary structures. Such epitope was also identified on the bulky multiantennary *N*-glycans isolated from cortical alveolus glycoproteins (hyosophorins) of fertilized fish eggs which carry species-specific, highly branched poly-*N*-acetyllactosaminoglycans (Taguchi *et al.*, 1993, 1994, 1995, 1996). Of all the structural variants determined, only those presented by the hyosophorins of medaka fish, *O. latipes*, also contain the fucosylated version of the α2-3-Neu5Ac-sialylated Galβ1-4Galβ1-4GlcNAc motif, identical to that currently identified on the zebrafish *N*-glycans. Interestingly, the Fucα1-3GalNAcβ1 epitope as found on the zebrafish *O*-glycans has also been identified on the *N*-glycans of hysophorin of flounder (Seko *et al.*, 1989), but not sialylated and is attached to the same -3Galβ1-4Galβ1-4GlcNAc sequence. Thus, a Galβ1-4Galβ1-4GlcNAc unit, occurring either at the terminal or as an internal unit, with and without further α2-3Neu5Ac sialylation on the β4-Gal and/or α3-fucosylation on the β4-GlcNAc appears to be a shared feature among the *N*-glycans from several fish eggs characterized to date. A slightly different version with an additional α4-Gal capping instead of sialylation or fucosylation has also been recently identified on the pigeon serum immunoglobulin G (Suzuki *et al.*, 2003), suggesting a possible wider occurrence on nonmammalian vertebrates.

The presence of the Lewis x-type α3-fucosylation is in agreement with the characterization of two zebrafish α,1-3 fucosyltransferases capable of synthesizing Lewis x from lacto-*N*-neotetraose *in vitro* (Kageyama *et al.*, 1999). A recent communication (Natsuka *et al.*, 2005) has further reported the identification of such Lewis x carrying *N*-glycans which appeared from the segmentation period (18 h) onward. The occurrence of the sialylated version of the same structure or other *N*-glycans were, however, not investigated (Natsuka *et al.*, 2005; Takemoto *et al.*, 2005), which probably biased the overall glycomic representation. Our MS-based profiling did not reveal a significant increase in the relative abundance of nonsialylated complex-type *N*-glycans following segmentation. It should, however, be noted that the discrepancy may also arise from different starting materials used because we have attempted to start from total delipidated extracts, whereas Takemoto *et al.* (2005) have treated the lyophilized embryos, free of corion and yolk, directly with hydrazinolysis for *N*-glycan release. Interestingly, they have also identified a significant proportion of biantennary *N*-glycans with the reducing end GlcNAc missing, namely with a trimannosyl $GlcNAc_1$ core instead of the expected di-*N*-acetylchitobiose, and attributed the findings to elevated endo-β-*N*-acetylglucosaminidase activity (Natsuka *et al.*, 2005). Although the activity of peptide:*N*-glycanase or glycoamidase has been convincingly demonstrated in the early embryos of medaka fish (Seko *et al.*, 1991) and elsewhere in other animals, an endoglycosidase F or a chitobiase-type activity has not been previously identified in fish. It is nevertheless conceivable that mammalian-type stepwise action of the lysosomal aspartylglucosaminidase and chitobiase (Michalski *et al.*, 1977; Strecker *et al.*, 1988) could lead to intracellular generation of such free sialylated complex-type *N*-glycans in zebrafish embryos which were not completely removed from our glycoprotein sample preparation because of the use of a low molecular weight cutoff dialysis (3500 Da), coupled with subsequent omission of C18 Sep-Pak step after tryptic digestion to increase the yield of oligosialylated *N*-glycans. Chemical degradation during permethylation

was deemed unlikely as similar structures were also observed with peracetylation. Furthermore, none of the more abundant high–mannose-type *N*-glycans was found to exhibit similar degradation which seems to be restricted to the complex-type subset. The origin of these atypical "*N*-glycans" remains a moot point and may also represent a novel linkage or form of glycosylation merit further investigations.

Despite similarity to previously characterized glycan structures of fish eggs, the zebrafish glycans are unique and novel in many additional aspects. First, the hyosophorin *N*-glycans are bulky, and those of *O. latipes* have been shown to be exclusively pentaantennary, whereas the much simpler bi- and triantennary *N*-glycans from the fish glycophosphoproteins characterized to date do not have fucosylation. Second, whereas sialylated core 1 *O*-glycans with terminal Fucα1-3GalNAc epitope have also been identified in fish eggs (Inoue and Inoue, 1997), none carries the exact sequence as defined here. Oligo- or polysialylation, when occurs, extends from the C6 of the reducing end GalNAc or nonreducing terminal GalNAc. In contrast, the zebrafish Neu5Ac-Neu5Gc or Neu5Gc-Neu5Gc disialyl unit extends from an internal Gal of the *O*-glycans. Further work is needed to confirm the tentatively defined structures for the disialylated core 2-type *O*-glycans which appear to carry one to several units of Lewis x on its 6-arm.

Notably then, we have shown that Neu5Ac and Neu5Gc sialylation were not evenly distributed. Most Neu5Ac is located at the nonreducing terminal position and, mostly, if not, exclusively as monosialylated motif. Neu5Gc occurs as both terminal and internal residues. For the *O*-glycans, our MS/MS data indicated that only the species sialylated with Neu5Gc can be further sialylated with another Neu5Ac or Neu5Gc residue. Likewise, although both Neu5Ac and Neu5Gc monosialylated antennae could be detected for the *N*-glycans, a preliminary oligosialylation analysis with the more sensitive 1,2-diamino-4,5-methylenedioxybenzene high-pressure liquid chromatography (DMB-HPLC) fluorescent detection method indicated that only a Neu5Gcα2-8Neu5Gc-DMB derivative could be additionally derived from the *N*-glycans (Guerardel, unpublished data). These data strongly suggest the occurrence of a certain form of donor and acceptor substrate selectivity in the differential transfer of Neu5Ac and Neu5Gc on glycoprotein-type glycans of zebrafish or a strict intra-/extracellular compartmentalization of sialyltransferase activities.

The concentration of CMP-Neu5Gc in the cytosol has been suggested to play the most important role in regulating the level of Neu5Gc sialylation because neither the CMP-sialic acid antiporter nor the sialyltransferases examined so far seem to exhibit a preference for CMP-Neu5Ac or CMP-Neu5Gc (Higa and Paulson, 1985; Lepers *et al.*, 1989, 1990; Schauer and Kamerling, 1997). In contrast, different donor substrate specificities have been observed for enzymes involved in the elongation of oligo-/polysialylated chains. Thus, whereas rainbow trout polysialyltransferase (polyST) can use both CMP-Neu5Ac and CMP-Neu5Gc as activated sialyl donors, chick brain polyST was shown to not recognize CMP-Neu5Gc (Kitazume *et al.*, 1994; Sevigny *et al.*, 1998). However, to our knowledge, nothing is presently known on the possible specificity of these enzymes toward their acceptor substrates for Neu5Ac/Neu5Gc composition that may explain the absence of polysialyl elongation from Neu5Ac residues in zebrafish *O*-glycans.

Contrary to the glycoproteins, synthesis of oligosialyl sequences in glycolipids did not seem to be affected by the same biosynthetic restrictions. The major sialylated glycolipids detected conform to a family of lactosylceramides extended by up to four sialic acids which can be further elongated by a Hex-HexNAc unit to form either sialylated ganglio-tetraglycosylceramides or sialylated (neo)lactotetraglycosylceramides, with up to five sialic acids. Both series of glycolipids have previously been identified in other fishes (Ando and Yu, 1979; DeGasperi *et al.*, 1987; Nakamura *et al.*, 1997). However, to our knowledge, tetrasialylated lactosylceramides have not been observed previously in any model system. Such a compound, that would be named GQ3 according to used nomenclature, does not fit into accepted ganglioside synthesis pathway model in which GT3 is the biosynthetic precursor of the so called c-series (including GT2, GT1c, GQ1c, and GP1c) and is not further elongated by sialic acids (Freischutz *et al.*, 1995). The sialylated moieties of all observed glycolipids are made up by heterogeneous mixtures of Neu5Ac and Neu5Gc residues in all possible combinations. The presence of polymerized Neu5Ac sequences distinguishes their sialylation pattern from those of *N*- and *O*-glycans. Furthermore, homogeneously Neu5Ac-sialylated glycolipids are the major forms compared with Neu5Gc-containing glycolipids. Altogether, these data demonstrate that although the glycoprotein glycans and glycolipids are both highly sialylated, the biosynthesis of their respective oligosialylated moieties are differently regulated, and the sialylation pattern changes as the embryos develop.

In particular, the disialylated *O*-glycans were exclusively observed in the very first stages of development, before 24 h after the fertilization, whereas, surprisingly, the pattern of oligosialylation in glycolipids seems to follow the opposite trend with the oligosialylated glycolipids being exclusively observed in later developmental stages. The general low yield of the glycolipids relative to the major *N*- and *O*-glycans prevented more definitive structural characterization. We nevertheless could detect glycolipids of even higher degree of sialylation which collectively represent a complete shift in the glycolipid profile from the very early stage that contained a putative range of very different, highly heterogeneous neutral glycolipids. These findings strongly suggest the existence of a very complex, regulated expression pattern of sialylation according to the class of glycoconjugates and developmental stages. By furnishing the structural data pertaining to the glycome of *D. rerio*, our studies reported here provide a solid basis for further functional investigations into the specificity of glyco-related enzymes and, by extension, the role of glycosylation during development. In particular, we believe that *D. rerio* is a very promising model for the study of the fine regulation of sialylation events. Preliminary screening of gene data banks revealed that zebrafish genome not only contains orthologs of all identified human polyST-coding genes, but also several potentially new members of this family (Harduin-Lepers *et al.*, 2005). The structures reported also give the opportunity to identify other novel glycosyltransferase activities

such as the fucosyltransferase activity involved in the synthesis of Fuc(α1-3)GalNAc(β1-motif.

Materials and methods

Sample collection

Zebrafish (*D. rerio*) were maintained at 28°C on a 14-h light/10-h dark cycle. Embryos were incubated at 28°C, and different developmental stages were determined according to the description in the Zebrafish Book (Westerfield, 1995).

Extraction and preparation of glycoconjugates

Embryos were suspended in 200 µL of water and homogenized by sonication on ice. The resulting material was dried and then sequentially extracted three times by chloroform/methanol (2:1, v/v) and three times by chloroform/methanol/water (1:2:0.8, v/v/v). Supernatants from the latter extractions were pooled, dried, and subjected to a mild saponification in 0.1 M sodium hydroxide in methanol at 37°C for 3 h and then evaporated to dryness (Schnaar, 1994). Sample was reconstituted in methanol/water (1:1, v/v) and applied to a C18 Sep-Pak cartridge (Waters, Milford, MA) equilibrated in the same solvent system. After washing with five volumes of methanol/water (1:1, v/v), glycosylceramides were eluted by five volumes of methanol and five volumes of chloroform/methanol (2:1, v/v).

Delipidated pellet from chloroform/methanol/water extraction was resuspended in a solution of 6 M guanidinium chloride and 5 mM ethylenediaminetetraacetic acid (EDTA) in 0.1 M Tris/HCl, pH 8, and agitated for 4 h at 4°C. Dithiothreitol was then added to a final concentration of 20 mM and incubated for 5 h at 37°C, followed by the addition of iodoacetamide to a final concentration of 50 mM and further incubated overnight in the dark at room temperature. Reduced/alkylated sample was dialyzed against water at 4°C for 3 days and lyophilized. The recovered protein samples were then sequentially digested by (L-(tosylamido-2-phenyl) ethyl chloromethyl ketone)-treated trypsin for 5 h and chymotrypsin overnight at 37°C, in 50 mM ammonium bicarbonate buffer, pH 8.4. Crude peptide fraction was separated from hydrophilic components on a C18 Sep-Pak cartridge equilibrated in 5% acetic acid by extensive washing in the same solvent and eluted with a step gradient of 20, 40, and 60% propan-1-ol in 5% acetic acid. Pooled propan-1-ol fraction was dried and subjected to *N*-glycosidase F (Roche, Basel, Switzerland) digestion in 50 mM ammonium bicarbonate buffer, pH 8.4, overnight at 37°C. Alternatively, both the chloroform/methanol/water extraction and the C18 Sep-Pak purification step following tryptic digestion may be omitted to increase the yield of the sialylated *N*-glycans. Omission of the latter step would, however, increase the content of contaminant Hex polymers and possibly other free glycans.

The released *N*-glycans were separated from peptides using the same C18 Sep-Pak procedure as described above. To liberate *O*-glycans, retained peptide fraction from C18 Sep-Pak was submitted to alkaline-reductive elimination in 100 mM NaOH containing 1.0 M sodium borohydride at 37°C for 72 h. The reaction was stopped by addition of Dowex 50 × 8 cation-exchange resin (25–50 mesh, H^+ form) at 4°C until pH 6.5 and, after evaporation to dryness, boric acid was distilled as methyl ester in the presence of methanol. Total material was then submitted to cation-exchange chromatography on a Dowex 50 × 2 column (200–400 mesh, H^+ form) to remove residual peptides.

Chromatographic separation of glycans

The released *N*-glycans were either analyzed directly or after separation into neutral and sialylated fractions on a weak anion exchanger, DEAE Sephadex A-25 column (Amersham, Piscataway, NJ). Samples were dissolved in 20 mM Tris/HCl, pH 8, for loading onto a column equilibrated in the same buffer. Nonbinding neutral glycans were recovered in the washed through fractions, whereas sialylated glycans were eluted in a single fraction by a 0.8 M NaCl solution in 20 mM Tris/HCl, pH 8. Both fractions were desalted by passage through a Bio-Gel P2 column (Bio-rad, Hercules, CA) equilibrated in water.

To remove the contaminating neutral *N*-glycans and to enrich for the sialylated components, the *O*-glycans were dissolved in water and fractionated on a strong anion exchanger Dowex 1 × 2 (200–400 mesh, $HCOO^-$ form) column preequilibrated in water. Neutral glycans were washed off by water, whereas mono- and oligosialylated compounds were recovered by a stepwise elution at 0.1 and 2 M pyridine acetate, pH 5.5, respectively. High salt fractions were desalted by passage through a Bio-Gel P2 column equilibrated in water.

Exoglycosidase digestions

The *N*-glycans were digested with 20 mU of neuraminidase from *Arthrobacter ureafaciens* (Roche) in 100 µL of 50 mM sodium acetate buffer, pH 5.5, at 37°C for 16–18 h. Desialylated *N*-glycans were further treated with 3 mU of β1-4 galactosidase from *Streptococcus pneumoniae* (Calbiochem, Merck, Darmstadt, Germany) in 100 µL of 50 mM sodium acetate buffer, pH 5.5, at 37°C for 12 h, before and after chemical defucosylation by 48% aqueous hydrofluoric acid at 4°C for 48 h.

Chemical derivatization

Monosaccharide compositions were determined by gas chromatography (GC)-MS analysis as either per-heptafluorobutyryl (Zanetta *et al.*, 1999) or alditol acetate derivatives. For alditol acetates analysis, glycan samples were hydrolyzed in 4 M trifluoroacetic acid (TFA) for 4 h at 100°C and then reduced with sodium borohydride in 0.05 M NaOH for 4 h. Reduction was stopped by dropwise addition of acetic acid until pH 6 was reached, and borate salts were codistilled by repetitive evaporation in dry methanol. Peracetylation was performed in acetic anhydride at 100°C for 2 h. To determine the chemical nature of sialic acids, intact sialic acids were liberated directly by mild hydrolysis in 0.01 N TFA at 50°C and reacted with a volume of DMB reagent at 50°C for 2 h 30 min (Hara *et al.*, 1987). The monomeric DMB-sialic acid derivatives were separated isocratically on a C18 reverse phase (RP) HPLC column (250 × 4.6 mm, 5 micron, Vydac, Hesperia, CA) by

a solvent mixture of acetonitrile/methanol/water (7:9:84) and identified by referring to the elution positions of standard Neu5Ac and Neu5Gc derivatives.

For MALDI-MS analyses, the glycan samples were permethylated using the NaOH/dimethyl sulfoxide slurry method (Ciucanu and Kerek, 1984), as described by Dell *et al.* (1994). The permethyl derivatives were then extracted in chloroform and repeatedly washed with water. GC-MS linkage analysis was performed as described previously (Suzuki *et al.*, 2003).

MS analyses of glycans and glycolipids

For MALDI-time-of-flight (MALDI-TOF) MS glycan profiling, the permethyl derivatives in acetonitrile were mixed 1:1 with 2,5-dihydroxybenzoic acid (DHB) matrix (10 mg/mL in acetonitrile), spotted on the target plate, air-dried, and recrystallized on-plate with ethanol whenever necessary. Data acquisition was performed manually on a benchtop M@LDI LR system (Micromass) operated in the reflectron mode. For DHB matrix, the coarse laser energy control was set at high and fine adjusted using the % slider according to sample amount and spectra quality. Laser shots (5 Hz, 10 shots/spectrum) were accumulated until a satisfactory signal to noise ratio was achieved when combined and smoothed. Glycan mass profiling was also performed on a dedicated Q-Tof Ultima MALDI instrument (Micromass, Manchester, UK), in which case the permethylated samples in acetonitrile were mixed 1:1 with α-cyano-4-hydrocinnamic acid matrix (in acetonitrile : 0.1% TFA, 99:1 v:v) for spotting onto the target plate. The nitrogen UV laser (337 nm wavelength) was operated at a repetition rate of 10 Hz under full power (300 μJ/pulse). MS survey data were manually acquired, and the decision to switch over to CID MS/MS acquisition mode for a particular parent ion was made on the fly upon the examination of the summed spectra. Argon was used as the collision gas with a collision energy manually adjusted (between 50 and 200 V) to achieve optimum degree of fragmentation for the parent ions under investigation.

NMR analyses

Before NMR spectroscopic analysis, sample was repeatedly exchanged in $^{2}H_2O$ (99.97% purity, Euriso-top, CEA Saclay, France) with intermediate freeze drying and then dissolved in 250 μL of Me_2SO-d_6 (Euriso-top). Chemical shifts were expressed in p.p.m. downfield from the signal of the methyl group of Me_2SO-d_6 ($\delta^{1}H$/tetramethyl-silyl ester [TMS] = 2.52 p.p.m., $\delta^{13}C$/TMS = 40.98 p.p.m. at 343 K). The sample was analyzed in 200 × 5 mm BMS-005-B Shigemi tubes on a Bruker ASX-400 spectrometer (Centre d'Analyses RMN, Villeneuve d'Ascq, France) ^{1}H, 400.33 MHz; ^{13}C, 100.66 MHz equipped with a double resonance (^{1}H/X) Broad Band Inverse z-gradient probe head. All NMR data were recorded without sample spinning.

The one-dimensional proton ^{1}H spectrum was measured using a 90° tipping angle for the pulse and 1.5 s as a recycle delay between each of 32 acquisitions of 2.4 s. The spectral width of 4006 Hz was collected in 16384 complex data points. Two-dimensional homonuclear (^{1}H-1H) spectra (COSY and TOCSY) were measured using standard Bruker pulse programs. Rotating frame Overhauser enhancement spectroscopy (ROESY) spectra were acquired with 400 ms mixing times and acquired in the States mode, according to Bax and Davis (1985). Moreover, the two-dimensional TOCSY spectrum was recorded using a MLEV-17 mixing sequence of 120 ms. The spin lock field strength corresponded to a 90° pulse width of 62 μs. The spectral width was 2402 Hz in both dimensions. About 256 spectra of 4096 data points with 32 scans per t1 increment were recorded giving a spectral resolution of 0.6 Hz/pt in F2 and ~9.4 Hz/pt in F1. Heteronuclear HMQC ^{1}H-13C spectrum was obtained using with standard Bruker inv4tp pulse sequence. The spectral width was 2403 Hz in F2 and 12080 Hz in ^{13}C dimension, giving a spectral resolution of 0.6 Hz/pt and 47.2 Hz/pt, respectively.

Acknowledgments

This work was supported by Academia Sinica, Taiwan (to K.H.K. and C.J.H.), and an Academia Sinica Postdoctoral Fellowship (to Y.G.). Mass spectrometry analyses were performed at the Core Facilities for Proteomics Research located at the Institute of Biological Chemistry, Academia Sinica, supported by a Taiwan National Science Council grant (NSC 93-3112-B-001-010-Y) and the Academia Sinica.

Abbreviations

CID, collision-induced dissociation; DMB, methylenedioxybenzene; Hex, hexose; HexNAc, *N*-acetyl hexosamine; HexNAcitol, reduced *N*-acetyl hexosaminitol; LacCer, lactosylceramide; MALDI, matrix-assisted laser-desorption ionization; MS, mass spectrometry; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; polyST, polysialyltransferase; TFA, trifluoroacetic acid; TMS, tetramethyl-silyl ester; TOCSY, total correlation spectroscopy.

References

Ando, S. and Yu, R.K. (1979) Isolation and characterization of two isomers of brain tetrasialogangliosides. *J. Biol. Chem.*, **254**, 12224–12229.

Bakkers, J., Semino, C.E., Stroband, H., Kijne, J.W., Robbins, P.W., and Spaink, H.P. (1997) An important developmental role for oligosaccharides during early embryogenesis of cyprinid fish. *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 7982–7986.

Bax, A.D. and Davis, J. (1985) Practical aspect of two-dimentional transvers NOE spectroscopy. *J. Magn. Reson.*, **63**, 207–213.

Becker, C.G. and Becker, T. (2002) Repellent guidance of regenerating optic axons by chondroitin sulfate glycosaminoglycans in zebrafish. *J. Neurosci.*, **22**, 842–853.

Bernhardt, R.R. and Schachner, M. (2000) Chondroitin sulfates affect the formation of the segmental motor nerves in zebrafish embryos. *Dev. Biol.*, **221**, 206–219.

Bink, R.J., Habuchi, H., Lele, Z., Dolk, E., Joore, J., Rauch, G.J., Geisler, R., Wilson, S.W., den Hertog, J., Kimata, K., and Zivkovic, D. (2003) Heparan sulfate 6-O-sulfotransferase is essential for muscle development in zebrafish. *J. Biol. Chem.*, **278**, 31118–31127.

Ciucanu, I. and Kerek, F. (1984) A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res.*, **131**, 209–217.

Coppin, A., Maes, E., and Strecker, G. (2002) Species-specificity of amphibia carbohydrate chains: the Bufo viridis case study. *Carbohydr. Res.*, **337**, 121–132.

Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active enzymes. Available at: http://afmb.cnrs-mrs.fr/CAZY/. Accessed December 12, 2005.

DeGasperi, R., Koerner, T.A., Quarles, R.H., Ilyas, A.A., Ishikawa, Y., Li, S.C., and Li, Y.T. (1987) Isolation and characterization of gangliosides with hybrid neolacto-ganglio-type sugar chains. *J. Biol. Chem.*, **262**, 17149–17155.

Dell, A., Reason, A.J., Khoo, K.H., Panico, M., McDowell, R.A., and Morris, H.R. (1994) Mass spectrometry of carbohydrate-containing biopolymers. *Methods Enzymol.*, **230**, 108–132.

Dupuy, F., Petit, J.M., Mollicone, R., Oriol, R., Julien, R., and Maftah, A. (1999) A single amino acid in the hypervariable stem domain of vertebrate alpha1,3/1,4-fucosyltransferases determines the type 1/type 2 transfer. Characterization of acceptor substrate specificity of the Lewis enzyme by site-directed mutagenesis. *J. Biol. Chem.*, **274**, 12257–12262.

Freischutz, B., Saito, M., Rahmann, H., and Yu, R.K. (1995) Characterization of sialyltransferase-IV activity and its involvement in the c-pathway of brain ganglioside metabolism. *J. Neurochem.*, **64**, 385–393.

Guerardel, Y., Kol, O., Maes, E., Lefebvre, T., Boilly, B., Davril, M., and Strecker, G. (2000) O-glycan variability of egg-jelly mucins from *Xenopus laevis*: characterization of four phenotypes that differ by the terminal glycosylation of their mucins. *Biochem. J.*, **352**, 449–463.

Guerardel, Y., Balanzino, L., Maes, E., Leroy, Y., Coddeville, B., Oriol, R., and Strecker, G. (2001) The nematode Caenorhabditis elegans synthesizes unusual O-linked glycans: identification of glucose-substituted mucin-type O-glycans and short chondroitin-like oligosaccharides. *Biochem. J.*, **357**, 167–182.

Guerardel, Y., Petit, D., Madigou, T., Guillet, B., Maes, E., Maftah, A., Boujard, D., Strecker, G., and Kol, O. (2003) Identification of the blood group Lewis (a) determinant in the oviducal mucins of *Xenopus tropicalis*. *FEBS. Lett.*, **554**, 330–336.

Hara, S., Takemori, Y., Yamaguchi, M., Nakamura, M., and Ohkura, Y. (1987) Fluorometric high-performance liquid chromatography of N-acetyl- and N-glycolylneuraminic acids and its application to their microdetermination in human and animal sera, glycoproteins, and glycolipids. *Anal. Biochem.*, **164**, 138–145.

Harduin-Lepers, A., Mollicone, R., Delannoy, P., and Oriol, R. (2005) The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology*, **15**, 805–817.

Haslam, S.M. and Dell, A. (2003) Hallmarks of Caenorhabditis elegans N-glycosylation: complexity and controversy. *Biochimie*, **85**, 25–32.

Herkt, F., Parente, J.P., Leroy, Y., Fournet, B., Blanchard, D., Cartron, J.P., van Halbeek, H., and Vliegenthart, J.F. (1985) Structure determination of oligosaccharides isolated from Cad erythrocyte membranes by permethylation analysis and 500-MHZ 1H-NMR spectroscopy. *Eur. J. Biochem.*, **146**, 125–129.

Higa, H.H. and Paulson, J.C. (1985) Sialylation of glycoprotein oligosaccharides with N-acetyl-, N-glycolyl-, and N-O-diacetylneuraminic acids. *J. Biol. Chem.*, **260**, 8838–8849.

Inoue, S. and Inoue, Y. (1997) Fish glycoproteins. In Montreuil, J., Vliegenthart, J.F.G., and Schachter, H. (eds), *Glycoproteins II*. Elsevier Sceince B.V., Amsterdam, The Netherlands, pp. 143–161.

Inoue, S., Iwasaki, M., Ishii, K., Kitajima, K., and Inoue, Y. (1989) Isolation and structures of glycoprotein-derived free sialooligosaccharides from the unfertilized eggs of Tribolodon hakonensis, a dace. Intracellular accumulation of a novel class of biantennary disialooligosaccharides. *J. Biol. Chem.*, **264**, 18520–18526.

Iwasaki, M., Seko, A., Kitajima, K., Inoue, Y., and Inoue, S. (1992) Fish egg glycophosphoproteins have species-specific N-linked glycan units previously found in a storage pool of free glycan chains. *J. Biol. Chem.*, **267**, 24287–24296.

Kageyama, N., Natsuka, S., and Hase, S. (1999) Molecular cloning and characterization of two zebrafish alpha (1,3) fucosyltransferase genes developmentally regulated in embryogenesis. *J. Biochem. (Tokyo)*, **125**, 838–845.

Kamerling, J.P. and Vliegenthart, J.F.G. (1992) High-resolution ^{1}H-nuclear magnetic resonance spectroscopy of oligosaccharide-alditols released from mucin-type O-glycoproteins. In Berliner, L.J. and Reuben, J. (eds), *Carbohydrates and Nucleic Acids*. Plenum Press, New York, pp. 1–194.

Kitazume, S., Kitajima, K., Inoue, S., Troy, F.A., 2nd, Cho, J.W., Lennarz, W.J., and Inoue, Y. (1994) Identification of polysialic acid-containing glycoprotein in the jelly coat of sea urchin eggs. Occurrence of a novel type of polysialic acid structure. *J. Biol. Chem.*, **269**, 22712–22718.

Lepers, A., Shaw, L., Cacan, R., Schauer, R., Montreuil, J., and Verbert, A. (1989) Transport of CMP-N-glycoloylneuraminic acid into mouse liver Golgi vesicles. *FEBS. Lett.*, **250**, 245–250.

Lepers, A., Shaw, L., Schneckenburger, P., Cacan, R., Verbert, A., and Schauer, R. (1990) A study on the regulation of N-glycoloylneuraminic acid biosynthesis and utilization in rat and mouse liver. *Eur. J. Biochem.*, **193**, 715–723.

Marx, M., Rutishauser, U., and Bastmeyer, M. (2001) Dual function of polysialic acid during zebrafish central nervous system development. *Development*, **128**, 4949–4958.

Michalski, J.C., Strecker, G., and Fournet, B. (1977) Structures of sialyl-oligosaccharides excreted in the urine of a patient with mucolipidosis I. *FEBS. Lett.*, **79**, 101–104.

Mourad, R., Morelle, W., Neveu, A., and Strecker, G. (2001) Diversity of O-linked glycosylation patterns between species. Characterization of 25 carbohydrate chains from oviducal mucins of Rana ridibunda. *Eur. J. Biochem.*, **268**, 1990–2003.

Nakamura, K., Tamai, Y., and Kasama, T. (1997) Gangliosides of dogfish (*Squalus acanthias*) brain. *Neurochem. Int.*, **30**, 593–604.

Natsuka, S., Takemoto, T., Moriguchi, K., Aoki, T., Ito, N., and Hase, S. (2005) Lewis X-oligosaccharides appeared at segmentation period of zebrafish embryo. *Glycoconj. J.*, **22**, 338. Abstract for Glyco XVIII.

Schauer, R. and Kamerling, J.P. (1997) Chemistry, biochemistry and biology of sialic acids. In Montreuil, J., Vliegenthart, J.F.G., and Schachter, H., eds. *Glycoproteins II*. Elsevier Science B.V., Amsterdam, The Netherlands, pp. 243–402.

Schnaar, R.L. (1994) Isolation of glycosphingolipids. *Methods Enzymol.*, **230**, 348–370.

Seko, A., Kitajima, K., Iwasaki, M., Inoue, S., and Inoue, Y. (1989) Structural studies of fertilization-associated carbohydrate-rich glycoproteins (hyosophorin) isolated from the fertilized and unfertilized eggs of flounder, Paralichthys olivaceus. Presence of a novel penta-antennary N-linked glycan chain in the tandem repeating glycopeptide unit of hyosophorin. *J. Biol. Chem.*, **264**, 15922–15929.

Seko, A., Kitajima, K., Inoue, S., and Inoue, Y. (1991) Identification of free glycan chain liberated by de-N-glycosylation of the cortical alveolar glycopolyprotein (hyosophorin) during early embryogenesis of the Medaka fish, *Oryzias latipes*. *Biochem. Biophys. Res. Commun.*, **180**, 1165–1171.

Semino, C.E. and Allende, M.L. (2000) Chitin oligosaccharides as candidate patterning agents in zebrafish embryogenesis. *Int. J. Dev. Biol.*, **44**, 183–193.

Semino, C.E., Allende, M.L., Bakkers, J., Spaink, H.P., and Robbins, P.P. (1998) Expression of Rhizobium chitin oligosaccharide fucosyltransferase in zebrafish embryos disrupts normal development. *Ann. N. Y. Acad. Sci.*, **842**, 49–54.

Sevigny, M.B., Ye, J., Kitazume-Kawaguchi, S., and Troy, F.A., 2nd. (1998) Developmental expression and characterization of the alpha2,8-polysialyltransferase activity in embryonic chick brain. *Glycobiology*, **8**, 857–867.

Snaar-Jagalska, B.E., Krens, S.F., Robina, I., Wang, L.X., and Spaink, H.P. (2003) Specific activation of ERK pathways by chitin oligosaccharides in embryonic zebrafish cell lines. *Glycobiology*, **13**, 725–732.

Strecker, G., Michalski, J.C., and Montreuil, J. (1988) Lysosomal catabolic pathway of N-glycosylprotein glycans. *Biochimie*, **70**, 1505–1510.

Suzuki, N., Khoo, K.H., Chen, C.M., Chen, H.C., and Lee, Y.C. (2003) N-glycan structures of pigeon IgG: a major serum glycoprotein containing Galalpha1-4 Gal termini. *J. Biol. Chem.*, **278**, 46293–46306.

Taguchi, T., Seko, A., Kitajima, K., Inoue, S., Iwamatsu, T., Khoo, K.H., Morris, H.R., Dell, A., and Inoue, Y. (1993) Structural studies of a novel type of tetraantennary sialoglycan unit in a carbohydrate-rich glycopeptide isolated from the fertilized eggs of Indian Medaka fish, Oryzias melastigma. *J. Biol. Chem.*, **268**, 2353–2362.

Taguchi, T., Seko, A., Kitajima, K., Muto, Y., Inoue, S., Khoo, K.H., Morris, H.R., Dell, A., and Inoue, Y. (1994) Structural studies of a

novel type of pentaantennary large glycan unit in the fertilization-associated carbohydrate-rich glycopeptide isolated from the fertilized eggs of *Oryzias latipes*. *J. Biol. Chem.*, **269**, 8762–8771.

Taguchi, T., Kitajima, K., Muto, Y., Inoue, S., Khoo, K.H., Morris, H.R., Dell, A., Wallace, R.A., Selman, K., and Inoue, Y. (1995) A precise structural analysis of a fertilization-associated carbohydrate-rich glycopeptide isolated from the fertilized eggs of euryhaline killi fish (*Fundulus heteroclitus*). Novel penta-antennary N-glycan chains with a bisecting N-acetylglucosaminyl residue. *Glycobiology*, **5**, 611–624.

Taguchi, T., Iwasaki, M., Muto, Y., Kitajima, K., Inoue, S., Khoo, K.H., Morris, H.R., Dell, A., and Inoue, Y. (1996) Occurrence and structural analysis of highly sulfated multiantennary N-linked glycan chains derived from a fertilization-associated carbohydrate-rich glycoprotein in unfertilized eggs of *Tribolodon hakonensis*. *Eur. J. Biochem.*, **238**, 357–367.

Takemoto, T., Natsuka, S., Nakakita, S., and Hase, S. (2005) Expression of complex-type N-glycans in developmental periods of zebrafish embryo. *Glycoconj. J.*, **22**, 21–26.

Westerfield, M. (ed.) (1995) *The Zebrafish Book*, 3rd edn. University of Oregan Press, Eugene, OR.

Zanetta, J.P., Timmerman, P., and Leroy, Y. (1999) Gas-liquid chromatography of the heptafluorobutyrate derivatives of the O-methyl-glycosides on capillary columns: a method for the quantitative determination of the monosaccharide composition of glycoproteins and glycolipids. *Glycobiology*, **9**, 255–266.

Developmental regulation of oligosialylation in zebrafish

Lan-Yi Chang · Anne Harduin-Lepers · Ken Kitajima · Chihiro Sato · Chang-Jen Huang · Kay-Hooi Khoo · Yann Guérardel

Received: 21 April 2008 / Revised: 9 June 2008 / Accepted: 10 June 2008

Abstract Zebrafish appears as a relevant model for the functional study of glycoconjugates along vertebrate's development. Indeed, as a prelude to such studies, we have previously identified a vast array of potentially stage-specific glycoconjugates which structures are reminiscent of glycosylation pathways common to all vertebrates. In the present study, we have focused on the identification and regulation of major protein and lipids associated α2,8-linked oligosialic acids motifs in the early development of zebrafish. By a combination of partial hydrolysis, anion exchange HPLC-FD and mass spectrometry, we demonstrated that glycoproteins and glycolipids differed by the extent and the nature of their substituting oligosialylated sequences. Furthermore, relative quantifications showed that α2,8-linked sialylation was differentially regulated in both families of glycoconjugates along development. Accordingly, we established that α2,8-sialyltransferase mRNA levels was directly correlated with changes of α2,8-sialylation status of glycolipids, but independent of those observed on major glycoproteins that appear to originate from the mother.

Keywords α2,8-sialylation · α2,8-sialyltransferases · Neu5Ac/Neu5Gc · Zebrafish · Early development · Glycolipids · Glycoproteins

Introduction

Sialic acids constitute a vast family of heterogeneous monosaccharides that are distributed throughout most of the living organisms, from bacteria to vertebrates. About 40 different naturally occurring members have been described so far [1]. Most of them occur in bound forms to glycoproteins, glycolipids and free oligosaccharides as terminal monosaccharide units. Internal α2,8-linked sialic acids in the diasialosyl motif have been early recognized as a common constituent of gangliosides and then of glycoproteins [2–6]. α2,8-polySia chain with a degree of polymerization (DP) >7 was first identified in bacteria soluble colominic acid prepared from culture medium [7]. Not until 1980 was it identified in animal kingdom associated with a major sialoglycoprotein, the so called polysialoglycoprotein (PSGP), isolated from the eggs of rainbow trout [8]. Although ubiquitously identified in all fish species examined so far, α2,8-polySia chains exhibited a remarkable species specificity associated with differential sequences, nature of constituent sialic acids and degree of *O*-acetyl and *O*-lactyl substitutions [9]. Subsequently, α2,8-

Electronic supplementary material The online version of this article (doi:10.1007/s10719-008-9161-5) contains supplementary material, which is available to authorized users.

L.-Y. Chang · A. Harduin-Lepers · Y. Guérardel (✉)
Unité de Glycobiologie Structurale et Fonctionnelle,
UMR CNRS 8576,
Université des Sciences et Technologie de Lille 1,
59655 Villeneuve d'Ascq, France
e-mail: yann.guerardel@univ-lille1.fr

L.-Y. Chang · C.-J. Huang · K.-H. Khoo
Institute of Biochemical Sciences, National Taiwan University,
Taipei 106, Taiwan

L.-Y. Chang · C.-J. Huang · K.-H. Khoo
Institute of Biological Chemistry, Academia Sinica, Nankang,
Taipei 11529, Taiwan

K. Kitajima · C. Sato
Laboratory of Animal Cell Function,
Bioscience and Biotechnology Center,
Department of Bioengineering Sciences,
Nagoya University, Japan

polySia chain were shown to be associated with multiple glycoproteins in eukaryotic organisms, including neural cell adhesion molecule (N-CAM) [10], sodium channels [11], CD-36 from human milk [12] and neuropilin [13].

The cellular sialic acid content is metabolically regulated mainly by sialyltransferases and sialidases. α2,8-sialyltransferases (ST8Sia) are biosynthetic enzymes catalyzing the transfer of sialic acid residues either to sialoglycoproteins or sialoglycolipids (reviewed in [14–16]), while sialidases or neuraminidases are glycohydrolytic enzymes that remove sialic acid residues during sialoglycoconjugate degradation [17]. Sialylated glycoconjugates are known to be tightly regulated developmentally. In particular, gangliosides are exclusively expressed in mice embryos from 7 days, when neural crest appears, in agreement with their major localisation in central nervous system [18]. Then along brain development, gangliosides pattern show drastic modifications, including a decrease of GM3 and GD3 concomitant to an increase of GD1a and GT1b from mid-embryonic development onward [19]. Similarly in rat, ganglioside expression pattern was shown to shift from simple b-series to complex gangliosides along brain development [20]. Polysialylated N-CAM exhibits also an exquisite spatio-temporal regulation in the developing brain. Indeed, shortly after its appearance in mouse brain at 8 days, N-CAM become polysialylated with a peak expression of α2,8-polySia during perinatal phase and a complete clearing within the three weeks of post-natal brain development, with the exception of sites of neuronal plasticity [21, 22]. In fish, polysialylation profile of PSGP is regulated along oocytogenesis in ovary. Indeed, it was shown to shift from a diSia-PSGP in earlier stages of oogenesis to a α2,8-polySia-PSGP in later stages, based on the temporal regulated expression of the α2,8-sialyltransferases ST8Sia II and ST8Sia IV [23, 24]. However, with the exception of N-CAM, little is known about the regulation and functions of polySia sequences of various glycoconjugates along embryogenesis in vertebrates.

The zebrafish, *Danio rerio*, has emerged in recent years as an excellent model system to study the genetic underpinnings of vertebrate development. Recent profiling of zebrafish embryos glycosylation established that this organism synthesizes a wide variety of sialylated glycoconjugates, the expression of which is potentially regulated along development [25]. Collected data has poised zebrafish as an ideal model to study the role of sialylation in embryogenesis. The present report focuses on the identification and regulation of α2,8-linked sialylated glycoconjugates in the early stages of zebrafish development through a concerted approach combining structural, biochemical and molecular biology analyses. In a first step, we established with different structures of oligoSia sequences associated to different types of glycoconjugates. Then, we evaluated the implication of different glycogenes, including α2,8-sialyltransferases (ST8Sia) and sialidases in the early developmental regulation of the expression of α2,8-linked sialic acids.

Materials and methods

Materials

The molecular biology kits RNeasy Midi and Plasmid extraction were obtained from Qiagen (Chatsworth, CA, U. S.A.), the TOPO TA cloning kit was from Invitrogen (Cergy Pontoise, France), the NucleoSpin® RNA II kit was from Macherey-Nagel (Düren, Germany). The oligonucleotides were synthesized and purified by Eurogentec (Seraing, Belgium), Sybr Green Brilliant Q-PCR master mix, eight-well strip tubes and the MX-4000 Quantitative PCR System were from Stratagene (La Jolla, CA, USA). The first strand cDNA synthesis kit was from Amersham Pharmacia Biotech (Little Chalfont, U.K.). The cDNA Kidney library was kindly provided by L. Zon (ZFIN, Oregon). The experion ARN Std Sens Analysis kit was from Biorad (Marnes-la Coquette, France). Taq polymerase, 4-methylumbelliferone (4-MU) and 2'-(4methylumbelliferyl)-α-d-*N*-acetylneuraminic acid (4-MU-Neu5Ac) were from Sigma (St Louis, MO, USA).

Sample collection

Zebrafish (*D. rerio*) were maintained at 28°C on a 14 h-light/10 h-dark cycle. Embryos were incubated at 28°C and different developmental stages were determined according to the description in the Zebrafish Book [26].

Extraction and preparation of glycoconjugates

Embryos were suspended in 200 µl of water and homogenised by sonication on ice. The resulting material was dried and then sequentially extracted three times by chloroform/methanol (2:1, *v*/*v*) and chloroform/methanol (1:2, *v*/*v*). Supernatants from the extractions were pooled, dried and subjected to a mild saponification in 0.1 M sodium hydroxide in methanol at 37°C for 3 h, and then evaporated to dryness. Sample was reconstituted in methanol/water (1:1, *v*/*v*) and applied to a C18 Sep-Pak cartridge (Waters) equilibrated in the same solvent system. After washing with five volumes of methanol/water (1:1, *v*/*v*), glycosylceramides were eluted by five volumes of methanol and five volumes of chloroform/methanol (2:1, *v*/*v*).

Delipidated pellet from chloroform/methanol/water extraction was re-suspended in a solution of 6 M guanidinium chloride and 5 mM EDTA in 0.1 M Tris/HCl, pH 8, and agitated for 4 h at 4°C. Dithiothreitol was then added to a final concentration of 20 mM and incubated for 5 h at 37°C, followed by addition of iodoacetamide to a final concentration of 50 mM and further incubated overnight in the dark at room temperature. Reduced/alkylated sample was dialysed against water at 4°C for 3 days and lyophilized. The recovered protein samples were then sequentially digested by TPCK treated trypsin for 5 h and chymotrypsin overnight at 37°C, in 50 mM ammonium bicarbonate buffer, pH 8.4. Crude peptide fraction was separated from hydrophilic components on a C18 Sep-Pak cartridge (Waters) equilibrated in 5% acetic acid by extensive washing in the same solvent and eluted with a step gradient of 20, 40 and 60% propan-1-ol in 5% acetic acid. Pooled propan-1-ol fraction was dried and subjected to *N*-glycosidase F (Roche) digestion in 50 mM ammonium bicarbonate buffer pH 8.4, overnight at 37°C. The released *N*-glycans were separated from peptides using the same C18 Sep-Pak procedure as described above. To liberate *O*-glycans, retained peptide fraction from C18 Sep-Pak was submitted to alkaline reductive elimination in 100 mM NaOH containing 1.0 M sodium borohydride at 37°C for 72 h. The reaction was stopped by addition of Dowex 50× 8 cation-exchange resin (25–50 mesh, H^+ form) at 4°C until pH 6.5 and, after evaporation to dryness, boric acid was distilled as methyl ester in the presence of methanol. Total material was then submitted to cation-exchange chromatography on a Dowex 50×2 column (200–400 mesh, H^+ form) to remove residual peptides.

Chemical derivatization and MS analyses

Monosaccharide compositions were determined by gas chromatography (GC)-mass spectrometry (MS) analysis as alditol acetate derivatives. Briefly, glycan samples were hydrolysed in 4 M trifluoroacetic acid (TFA) for 4 h at 100°C and then reduced with sodium borohydride in 0.05 M NaOH for 4 h. Reduction was stopped by drop wise addition of acetic acid until pH 6 was reached and borate salts were co-distilled by repetitive evaporation in dry methanol. Peracetylation was performed in acetic anhydride at 100°C for 2 h.

For MALDI-MS analyses, the glycan samples were permethylated using the NaOH/dimethyl sulfoxide (Cuicanu and Kerek 1984), and then extracted in chloroform and repeatedly washed with water. MALDI-MS and MS/MS data were acquired on either a Q-TOF Ultima MALDI instrument (Micromass) or a MALDI-TOF/TOF system, the ABI 4700 Proteomic Analyzer, exactly as described [27].

Analysis of oligo-sialylated sequences

In order to minimize internal fragmentation of polysialylated sequences, sialylated glycan samples were directly coupled to 1,2-diamino-4,5-methylenedioxybenzene (DMB) without prior mild hydrolysis [28]. Samples were incubated for 2.5 h at 50°C in 50 µl of a DMB reagent solution (2.7 mM DMB, 9 mM sodium hydrosulfite, and 0.5 mM β mercaptoethanol in 20 mM TFA). 10 µl of 1 M NaOH was then added and the reaction mixtures further incubated in the dark at room temperature for 1 h. Samples were stored at 4°C before analysis.

DMB-derivatized sialic acid oligomers were separated on a HPLC apparatus fitted with either an anion exchanger column, mono-Q (Amersham-Biosciences), or a CarboPac PA-100 column (Dionex). For mono-Q column, the sample was loaded and eluted with a flow rate of 0.5 ml/min with 20 mM Tris–HCl (pH 8.0), followed by a NaCl gradient (0–10 min, 0 M; 10–60 min, 0 to 0.6 M; 60–65 min, 0.80 M) in 20 mM Tris–HCl (pH 8.0). CarboPac column was eluted at 1 ml/min with a concentration gradient of 2 to 32% of 1 M NaNO3 in water. In both systems, elution was monitored by an on line fluorescence detector set at wavelengths of 373 nm for excitation and 448 nm for emission. Periodate oxidation and C7/C9 analyses for oligosialyl linkage determination were performed essentially as described by Sato *et al.* (1998) [33]. Briefly, samples were dissolved in a mixture of 25 µl of 40 mM sodium acetate buffer (pH 5.5) and 2 µl of 0.25 M sodium metaperiodate and left at 0°C for 45 min in the dark. Five microliters of 5% glycerol was then added and allowed to react for another 40 min at 0°C, followed by 32 µl of 0.2 M sodium borohydride in 0.2 M sodium borate buffer (pH 8.0) and left overnight at 0°C. Finally, TFA was added to a final concentration of 1 M and incubated at 80°C for 1 h before subjected to DMB derivatization. To determine the chemical nature of sialic acids, intact sialic acids were liberated directly by mild hydrolysis in 0.01 N TFA at 50°C and reacted with a volume of DMB reagent at 50°C for 2 h 30 min. The monomeric DMB-sialic acid derivatives were separated isocratically on a C18 reverse phase HPLC column (250×4.6 mm, 5 µm, Vydac) by a solvent mixture of acetonitrile/methanol/ water (7:9:84) and identified by referring to the elution positions of standard Neu5Ac and Neu5Gc derivatives. For additional MS analysis, the DMB-derivatives were separated instead with a gradient of acetonitrile/methanol/water (7:9:84) mixture in water (0–10 min, 10%; 10–40 min, 10 to 100%). Fluorescence-detected peaks were individually collected on ice and immediately freeze dried. Samples were then reconstituted in 10 µl of water and analysed by nanoESI-MS and MS/MS on an LCQ DK XP+ ion trap (Thermo Finnigan).

RNA extraction and cDNA synthesis

D. rerio unfertilized eggs and embryos (0, 6, 14, 24 and 36 hpf) kindly provided by the Thisses' Lab were sampled (200 embryos) and snap frozen in liquid nitrogen. Total cellular RNA was extracted from embryos at various developmental stages using the RNeasy Midi kit according to the manufacturer's instructions. Total RNA purity was established by calculating the ratio of the absorbance readings at 260 and 280 nm and quantified using the NanoDrop® ND-100 spectrophotometer (NanoDrop Technologies, Wilmington, DE, U.S.A.). The integrity and purity of the extracted RNA was also analyzed by means of gel electrophoresis on a bioanalyzer (Experion, Biorad). Total RNA (1.2 µg) was reverse transcribed using the first strand cDNA synthesis kit in 33 µl following the manufacturer's instruction. RNA samples were tested for genomic DNA contamination by PCR amplification of the zebrafish β-actin (GenBank accession number AF025305; [29], using oligonucleotide primers designed in two distinct exons (Sup. Table 1) and aerosol contamination by including no template controls (NTC). Another set of developmental stages cDNAs were kindly prepared by H. Ahmed and G. Vasta according to [30].

Real-time PCR of ST8Sia genes during development in zebrafish

Primers used for quantitative Q-PCR (Sup. Table 1) were designed in the coding region of previously identified zebrafish ST8Sia genes [14] using the Primer Premier version 31.1 software (Primer Premier, Biosoft International, Palo Alto, USA). Each primer pair was carefully selected so to give rise to an amplified DNA fragment of about 300 bp and such that their Tm values were very close (around 51°C). The suitability of the primers for their uniqueness to amplify a single PCR product was assured by regular end-point PCR (Denaturation step at 94°C for 2 min followed by 38 cycles at 95°C 1 min; 50°C 1 min; 72°C 1 min and an elongation step at 72°C for 10 min) using cDNA kidney library provided by L. Zon. The amplified products were subsequently run on an agarose gel, subcloned in TOPO TA cloning vector and finally, fully sequenced (Genoscreen, Lille). The TOPO plasmids containing the amplified regions of the targeted genes were amplified, purified and quantified by nanodrop and used for the establishment of a standard curve for absolute quantification. Efficiency of target amplification for each primer set (ST8Sia I, ST8Sia II, ST8Sia III, ST8Sia IV, ST8Sia V, ST8Sia VI, Sup. Table 1) was optimized by real-time PCR performed in a Stratagene MX4000 by trialing several final primer concentrations. Each 25 µl Q-PCR master mix contained 12.5 µl 2X Master Mix (Brilliant® SYBR® Green Q-PCR Master Mix (Stratagene, CA)), 150 nM of each primer, and 5 µl of diluted cDNA (equivalent to 100 ng total RNA) extracted from 0, 6, 14, 24 hpf embryos and the real-time quantitative PCR were the thermal cycling program consisted of 10 min at 95°C followed by 45 cycles of 30 s at 95°C, 1 min at 50°C and 30 s at 72°C and this was followed by a melting step consisting of heating from 50°C to 95°C at an increment of 1°C per 30 s to check the specificity of the amplified product. PCR for all the samples were carried out in triplicate in eight-well strip tubes and data were expressed as means +/− SD. The reactions were quantified by selecting the amplification cycle when the PCR product of interest was detected (threshold cycle, Ct). Calibration curves were generated by ten-fold serial dilution of Hind III linearized TOPO plasmids containing the amplified regions of the targeted genes (from 2×10 5 copies to 2×10 1 copies). The same PCR master mix and thermocycler conditions as described above were used and plasmid standard curve equations were used to calculate the absolute copy number of each gene. The amplification efficiencies of each calibrator were found to be between 95.9% and 100.5%. We used absolute quantification relying on the serial diluted DNA fragment with known concentration, called calibrators, which were amplified from cDNA of 24-hpf embryos with the same primers.

Sialidases assays

D. rerio embryos (0, 8, 24 and 48 hpf) and unfertilized eggs were sampled (500 embryos) and snap frozen in liquid nitrogen. The eggs were homogenized in 500 µl of water, then different amount of total cell lysate corresponding to 1 to 60 eggs, were mixed with 0.2 mM 4-MU-Neu5Ac in 50 mM sodium acetate buffer (pH 3, 4, 5, 6 and 7) in a final volume of 250 µl. Protein concentration used were determined using the micro BCA TM protein assay reagent kit (Thermo Scientific Pierce, Rockford, USA). The incubation was performed at 37°C. At 0.5, 1, 2 and 4 h, 30 µl of the reaction mixture was taken back, and the reaction was quenched by adding 120 µl 0.5 M Na_2CO_3. The released 4-methyl-umbeliferone (MUN) was measured and quantified by fluorescence detector at 360 nm for excitation, 460 nm for excitation. Sialidase activity was calculated according to a MUN standard curve.

Results

Glycans from embryos contain oligosialic acid chains

Oligosialylation on N-glycans

We previously described in zebrafish embryos a family of unusual di- and tri-antennary sialylated *N*-glycans along

with major ubiquitous oligomannosylated *N*-glycans [25]. They are characterized by the presence of Neu5Ac/Neu5Gc monosialylated Lewis x motifs further substituted by a β4-Gal residue. A 10% sialic acid content of the total *N*-glycan fraction (Fig. 1) nonetheless implied that a much larger proportion of sialylated *N*-glycans than that detectable by MS might be present. In fact, after purification of sialylated *N*-glycans by anion exchange chromatography, the proportion of sialic acids in sialylated *N*-glycans increased sharply up to 23% of total monosaccharides, which represents an average of three to four sialic acid residues per *N*-glycan. As demonstrated by reverse phase (RP)-HPLC analysis of sialic acid-DMB derivatives, the Neu5Gc:Neu5Ac ratio, which ranges between 2:1 and 4:1 depending on the sample batch, was also somewhat inconsistent with a prevalence of Neu5Ac over Neu5Gc implicated by MALDI-MS profiling of the *N*-glycans. These discrepancies between the MS and sialic acid composition data indicated that some additional oligosialylated *N*-glycans may be refractory to MALDI-MS detection.

To gain a better picture of the sialylation, the well established DMB-tagging and HPLC analytical method [31] was further employed to identify possible presence of oligo- or polysialyl motifs. We first conducted structural analyses of glycoproteins associated oligosialylation on 1 hpf embryos, then established that oligosialylation was qualitatively identical in other developmental stages. The *N*-glycan sample was incubated in the acidic derivatization reaction mixtures without prior acidic liberation to minimize internal fragmentation of polysialic acid chain. The resulting tagged products were then separated on anion exchange HPLC columns (MonoQ and CarboPac PA-100) according to their degree of polymerization (DP) and detected with a fluorescent detector (FD). Under the experimental conditions employed, a monoQ column permits a ready detection of polymeric sialic acid chains

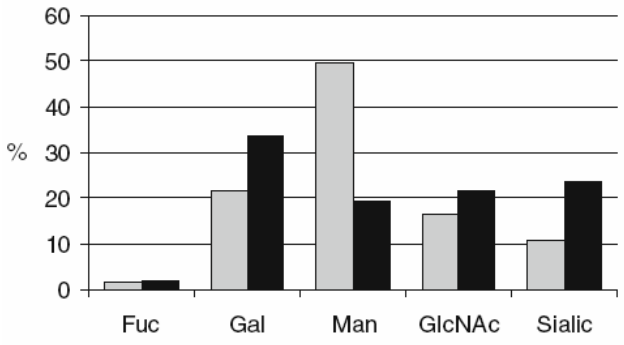


Fig. 1 Relative monosaccharide compositions of *N*-glycans. Monosaccharide compositions of total (in *grey*) and acidic (in *black*) *N*-glycans liberated from 1 hpf embryos were analysed by gas-chromatography. Results are expressed in percentage of total monosaccharides

from DP 2 up to DP 50 (data not shown), whereas the CarboPac PA-100 column also allows detection of Neu5Ac monomer. Since the FD response per mol of $(Neu5Ac)_n$-DMB remains constant for low DPs [32], integration of peak areas therefore provides a good estimation of the relative abundance of various Sia_n units. On MonoQ column, total *N*-glycans fraction was found to yield at most six peaks with retention times corresponding to $Sia[(\alpha2–8)Sia]_n$-DMB standards of DP 2 to DP 7 (Fig. 2a). Identical result was obtained with acidic *N*-glycans obtained after purification by anion exchange chromatography (data not shown). A sharp drop in the relative intensities of peaks occurred from DP 2 onwards. On PA-100 column, it could be estimated that DP 1 and DP 2 constitute 59% and 38%, respectively of the total content with higher oligosialyl chains contributing to less than 3% in total (data not shown). As expected, the observed peaks are sensitive to the action of exoneuraminidase (data not shown). Closer examination of the chromatograms showed that standard $[-8)Neu5Ac(\alpha2-]_n$ and $[-8)Neu5Gc(\alpha2-]_n$ exhibited slightly different retention times, in particular for DP 2 and DP 3 (Fig. 3a). Accordingly, chromatographic behaviours of DMB-tagged oligoSia from *N*-glycans suggest that Sia_2 and Sia_3 are exclusively composed of Neu5Gc residues. As shown in Fig. 3c, standard $Neu5Gc_2$ peak co-migrates with DP 2, whereas $Neu5Ac_2$ peak exhibits a clear time shift compared to DP 2 (Fig. 3b).

To ascertain the identity of the major dimeric peak, the same DMB derivative mixtures were subjected to RP-HPLC in order to purify DP 2. The elution position of dimeric sialic acid-DMB was inferred from standard $Neu5Ac_{1–3}$-DMB mixtures. A single major dimeric peak was detected at a retention time similar to that of standard $Neu5Ac_2$-DMB (data not shown) and was collected for MS and MS/MS analyses. As shown in Fig. 4a, ESI-MS analysis of a standard $Neu5Ac_2$-DMB afforded three molecular ion signals in positive ion mode, corresponding to $[M+H]^+$, $[M+Na]^+$ and $[M–H+2Na]^+$ at *m/z* 717, 739 and 761, respectively. Further CID-MS/MS on the mono-sodiated parent ion (Fig. 4b) yielded a major y ion at *m/z* 448 due to facile loss of the non-reducing terminal Neu5Ac residue. In contrast, similar ESI-MS analysis on the collected dimeric peak from the sample gave the corresponding molecular ions at *m/z* 749, 771 and 793 (Fig. 4c), which differ from those afforded by $Neu5Ac_2$-DMB dimer by 32 mass units and are consistent with a $Neu5Gc_2$-DMB composition. This is supported by CID-MS/MS on the candidate mono-sodiated parent ion at *m/z* 771 (Fig. 4d), which afforded a major y ion at *m/z* 464, corresponding to loss of a non-reducing terminal Neu5Gc. Further confirmation was then sought by referring to the CID MS/MS spectrum of an authentic $Neu5Gc_2$-DMB standard which was found to co-elute with $Neu5Ac_2$-DMB

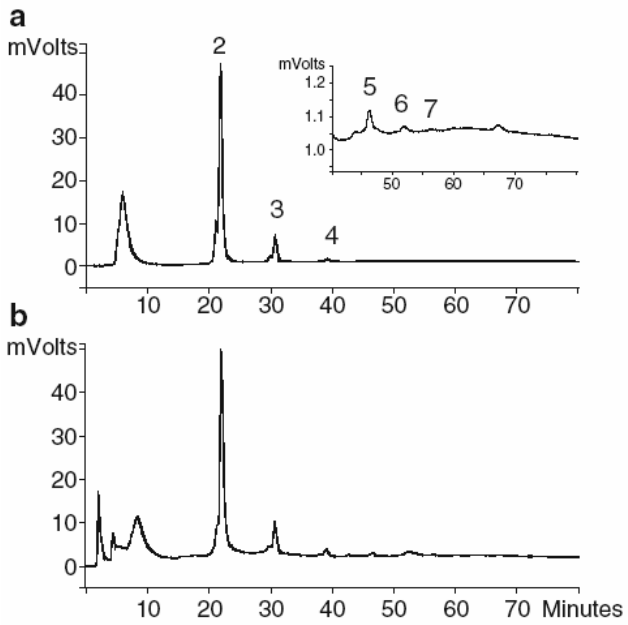


Fig. 2 Profiles of oligosialylated sequences on *O*- and *N*-glycans. OligoSia sequences were released from glycans isolated from 1hpf embryos, tagged with DMB and separated by HPLC-FD on an anion exchange column. OligoSia profiles from (a) *N*-glycans including the *inset in upper panel*, and (b) *O*-glycans. Peaks are labeled according to the DP values as established by comparison with authentic standards

standard at the same retention time under the HPLC conditions employed. In contrast, a putative mono-sodiated Neu5Ac1Neu5Gc$_1$-DMB peak at *m/z* 755 did not afford either a loss of Neu5Ac or Neu5Gc and was subsequently shown to be a prominent ESI-MS contaminant peak commonly observed when sample amount was low. Thus, our innovative MS and MS/MS approaches have provided unambiguous evidence for the presence of a Neu5Gc-Neu5Gc dimer, and not a Neu5Ac$_2$ or Neu5Ac$_1$Neu5Gc$_1$ dimer, as a major oligosialyl motif on the *N*-glycans.

Exclusive presence of Neu5Gc in oligosialylated sequences was further assessed by mild periodate oxidation followed by hydrolysis and DMB-labelling. Applied to the sialylated *N*-glycans, it cleaves the non-substituted side chains of Neu5Ac/Neu5Gc at the C_7–C_8 bond which are identified by RP-HPLC as DMB-labelling C7/C9 analogues [33]. This demonstrated that all 4 expected products, namely C_9(Neu5Gc)-DMB, C_7(Neu5Gc)-DMB, C_9(Neu5Ac)-DMB and C_7(Neu5Ac)-DMB could be detected at increasing retention time (Fig. 5a), and quantified as representing 28, 51, 2 and 18% of the total sialic acid content, respectively on an *N*-glycan sample with a Neu5Gc to Neu5Ac ratio of 4:1. Assuming the mild periodate oxidation of sialic acid has proceeded to completion and not hampered by any undetected non-saccharide substitution on the side chain, the recovery of intact C_9 Neu5Ac/Neu5Gc-DMB derivatives (about 30% of total) is normally indicative of an internal α2,8-linked sialyl motif since terminal or α2–9 linked sialic acids would be cleaved to C7 analogues. Strikingly, there was 14 fold more α2,8-linked Neu5Gc than Neu5Ac which suggests that where oligosialylation may occur, it preferentially extends from Neu5Gc and not from Neu5Ac. This conclusion is consistent with the observation that Neu5Gc dimer and trimers are the major oligosialyl motifs. Moreover, after subtracting the proportion of terminal Neu5Gc that was α2–8 linked to internal Neu5Gc/Neu5Ac in the dimer, the amount of the remaining C_7-Neu5Gc (21%) was roughly the same as that of C_7-Neu5Ac (18%). This figure is in good agreement with the MALDI-MS analysis which detected complex type *N*-glycans with antenna monosialylated by approximately equal amount of terminal Neu5Gc and Neu5Ac [25].

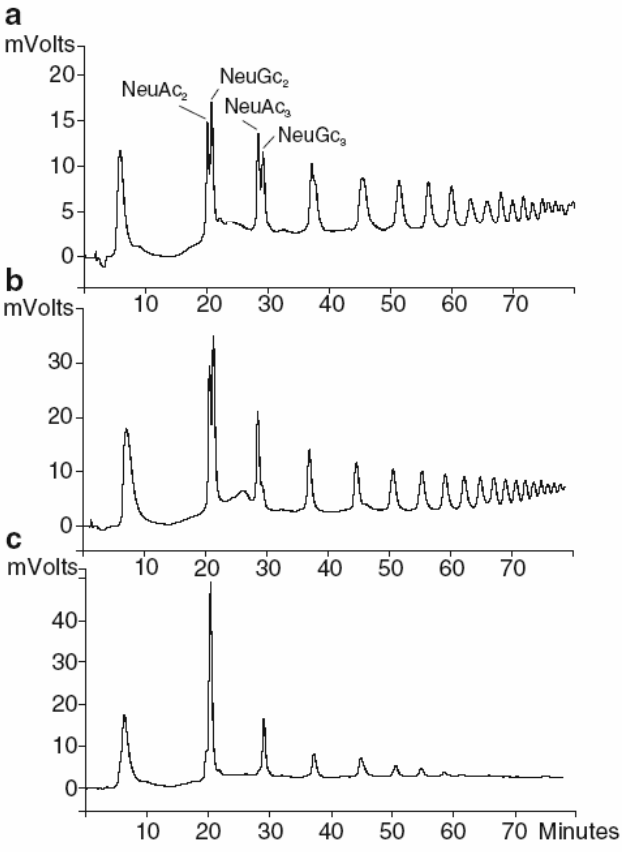


Fig. 3 Identification of oligosialylation on *N*-glycans by anion exchange DMB/HPLC-FD. Chromatographic profiles of co-injected DMB-derivatized **a** [-8)Neu5Ac(α2-]$_n$ and [-8)Neu5Gc(α2-]$_n$ standards, **b** [-8)Neu5Ac(α2-]$_n$ standard and oligo-Sia from zebrafish 1-hpf embryos *N*-glycans, **c** [8)Neu5Gc(α2-]$_n$ standard and oligo-Sia from zebrafish 1 hpf embryos *N*-glycans, showing that the diSia (DP=2) peak from zebrafish *N*-glycans co-migrate exclusively with Neu5Gc (α2–8)Neu5Gc

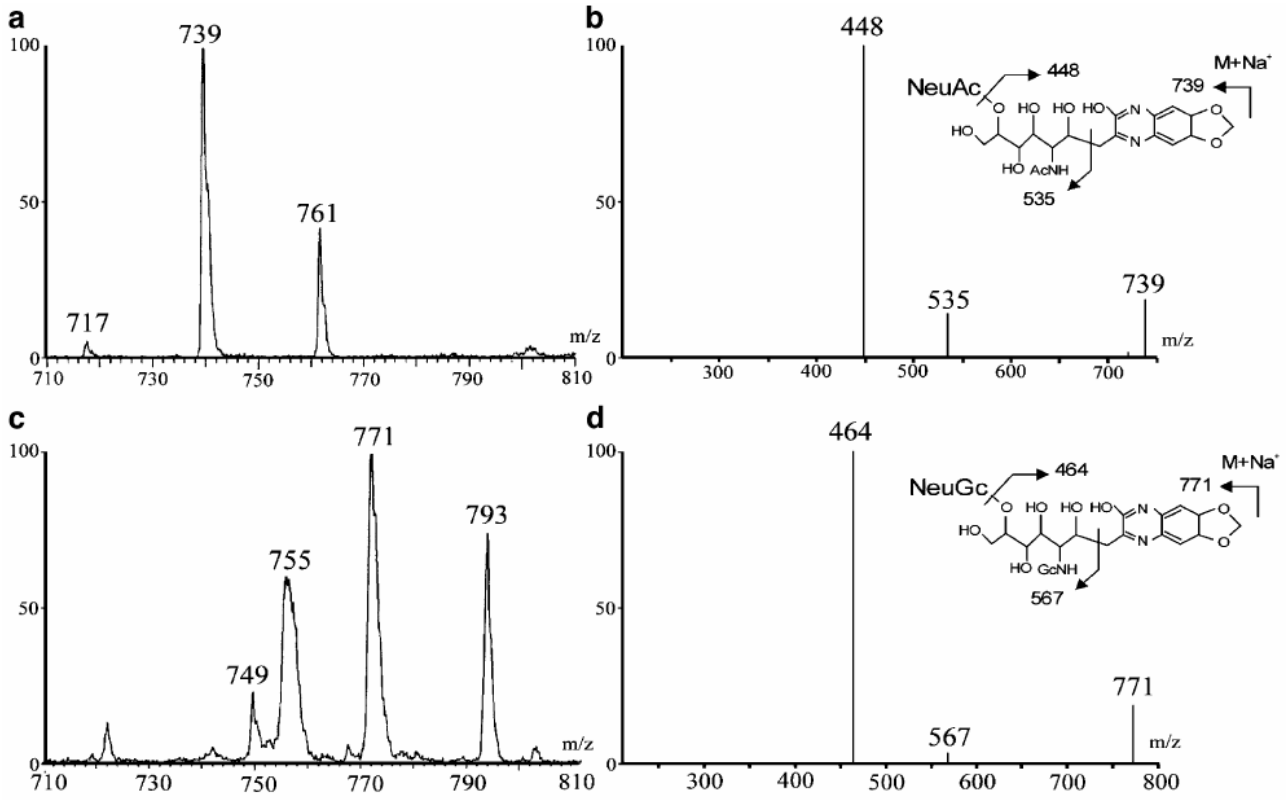


Fig. 4 Structural determination of the disialylated sequences on *N*-glycans by CID-MS/MS of DMB-derivatives. **a** ESI-MS profile of a Neu5Ac2-DMB standard isolated by C18 HPLC fractionation. The most abundant molecular ion at *m/z* 739 corresponds to a monosodiated species which was selected for CID MS/MS sequencing as shown in (**b**). **c** ESI-MS analysis of the putative dimeric sialic acid-DMB peak afforded by zebrafish embryonic *N*-glycans and similarly isolated by C18 HPLC, followed by CID MS/MS on the most abundant molecular ion at *m/z* 771 **d** which established its identity as monosodiated Neu5Gc2-DMB derivative. In both MS/MS, loss of 204 u corresponds to loss of the common DMB moiety through cleavage at C3–C4 of the derivatized, reducing end Neu5Ac/Neu5Gc, as shown schematically

Oligosialylation on O-glycans

Profiling of the glycosylation pattern of zebrafish embryos also demonstrated the presence of prominent sialylated O-linked glycans [25]. In contrast to *N*-glycans, disialylated motifs on *O*-glycans could be directly identified by MS analysis owing to lower molecular mass of *O*-glycans compared with *N*-glycans. In particular, we identified Neu5Gc–Neu5Gc as well as Neu5Ac–Neu5Gc motifs, but could not observe Neu5Gc–Neu5Ac and Neu5Ac–Neu5Ac, suggesting again the existence of exquisite specificity in the synthesis of α2,8-sialylated epitopes. The extent of α2,8-sialylation on *O*-glycans was evaluated using an identical experimental approach identical to that for *N*-glycans and showed very similar results. *O*-glycans are substituted by oligosialylated motifs including up to seven residues, as determined by DMB/HPLC-FD (Fig. 2b). As observed for *N*-glycans, slight shifts in the retentions times compared with Neu5Ac[(α2–8)Neu5Ac]n-DMB suggest the prevalence of Neu5Gc containing oligoSia over Neu5Ac (Sup. Fig. 1). Separation of periodate oxidised compound by RP-HPLC confirmed also the absence of internal α2,8-linked Neu5Ac residues in the molecule, as observed on *N*-glycans (Fig. 5b). However, *O*-glycans differed from *N*-glycans by a lower C_9(Neu5Gc) to C_7(Neu5Gc) ratio which suggests that the proportion of oligosialylation is lower in *O*-glycans.

Collectively, the data presented show that both *O*- and *N*-glycans are substituted by Neu5Gc containing oligosialylated sequences which exhibit similar overall patterns. Although Neu5Ac has been identified along Neu5Gc in *O*- and *N*-glycans, it seems to be restricted to monosialylated compounds or in non-reducing terminal position of oligosialylated sequences.

Oligosialylation on glycolipids

Direct MALDI-MS-mapping of the acidic glycolipids demonstrated the presence of oligosialylated glycolipids substituted by up to five sialic acids (Fig. 6b). The major tri-sialylated components were previously shown to be substituted by a mixture of Neu5Ac and Neu5Gc residues

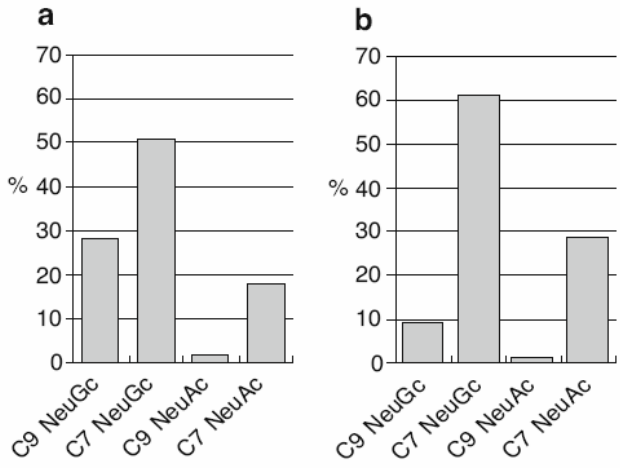


Fig. 5 Relative quantification of internal and external non-reducing sialic acids in glycoproteins oligosialylated motifs. Periodate oxidized Neu5Ac/Neu5Gc-DMB monomers from zebrafish embryonic glycans were resolved on RP-HPLC to distinguish the respective C7/C9 products by referring to the elution positions of authentic standards. Chromatographic profiles of **a** *N*-glycans and **b** *O*-glycans from 1 hpf embryos. Results are expressed in percentage of total oxidized derivatives and are representative of two independent experiments

in all possible combinations [25]. In contrast to *N*- and *O*-glycans, the presence of polymerized Neu5Ac indicates that no restriction seems to prevail in the synthesis of oligosialylated motifs in glycolipids. Accordingly, characterization of sialic acids by DMB/RP-HPLC demonstrated the prevalence of a molar ratio of 1.5:1 for Neu5Ac: Neu5Gc, indicating that sialylation patterns of glycolipids differ from those of glycoproteins in which Neu5Gc prevails over Neu5Ac. In agreement with direct observation of sialylated glycolipids by MS, DMB/HPLC-FD analysis shows the presence of oligosialylated motifs up to DP 6 (data not shown).

The sialic acid content changes along development

The presence of α2,8-sialylation on glycoproteins and glycolipids was assessed along the development timeline from 0 to 48 hpf. For *O*-glycans, presence of di-sialylated glycans could be directly assessed by MS profiling. MALDI-TOF analysis of permethylated glycans after separation of mono- and disialylated compounds shows that previously identified di-sialylated *O*-glycans can be observed exclusively in the earlier stages of developments (0 and 8 hpf) as signals at *m/z* 1706 (Fucα1–3GalNAcβ1–4(Neu5Ac–Neu5Gcα2–3)Galβ1–3GalNAc-itol) and at *m/z* 1736 (Fucα1–3GalNAcβ1–4(Neu5Gc–Neu5Gcα2–3)Galβ1–3GalNAc-itol), but never in the later stages (Fig. 7). Although mass spectrometry does not provide quantitative information, it suggests a disappearance of oligosialylation on *O*-glycans along development. DMB/HPLC-FD analysis provided a semi-quantitative comparison of the oligosialylation content of total *O*-glycan fractions purified from identical numbers of embryos at each development stage. Each fraction presented a very similar pattern of $(Sia)_n$-DMB with $1<n<6–7$, and thus did not show qualitative variation in the extent of α2,8-sialylation. However, a clear decrease in the quantity of each oligomer was observed as shown by integration of chromatographic peaks for di-, tri and tetrasialyl components (Fig. 8a), confirming the rapid decrease of oligosialylated *O*-glycans during embryonic development. Indeed, after 8 hpf, quantity of disialylated motif dropped by more than 60% and less than 5% of the initial di-sialylated

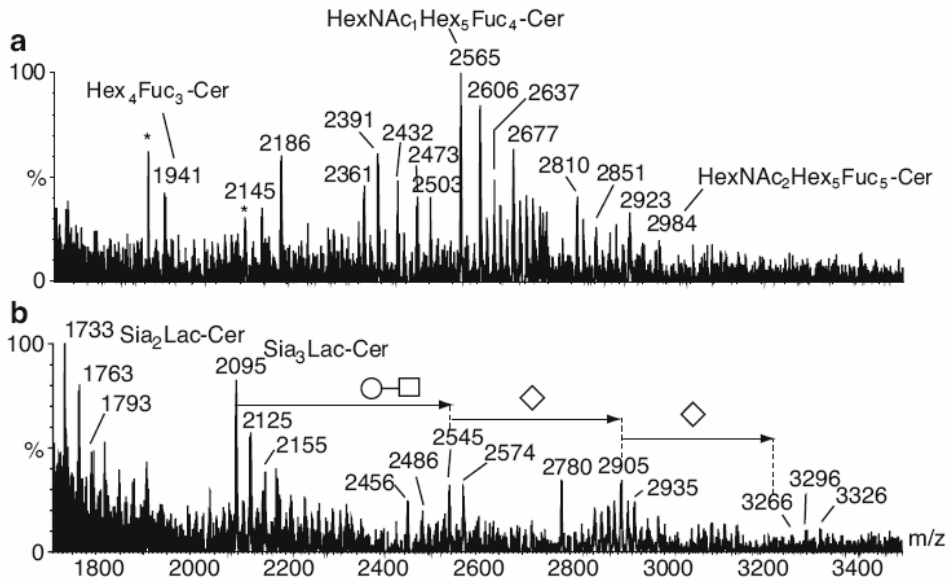


Fig. 6 MALDI-MS analyses of permethylated glycolipids from zebrafish embryos. Glycolipid profiles of high molecular mass glycolipids from **a** 1 hpf and **b** 48 hpf embryos. No sialylated glycolipids were observed at 1 h psf, whereas a complex mixture of oligosialylated glycolipids containing from 2 to 5 sialic acids were detected at the later stages. Symbols used: *circle* Hex, *square* HexNAc, *diamond* sialic acid, *Cer* ceramide

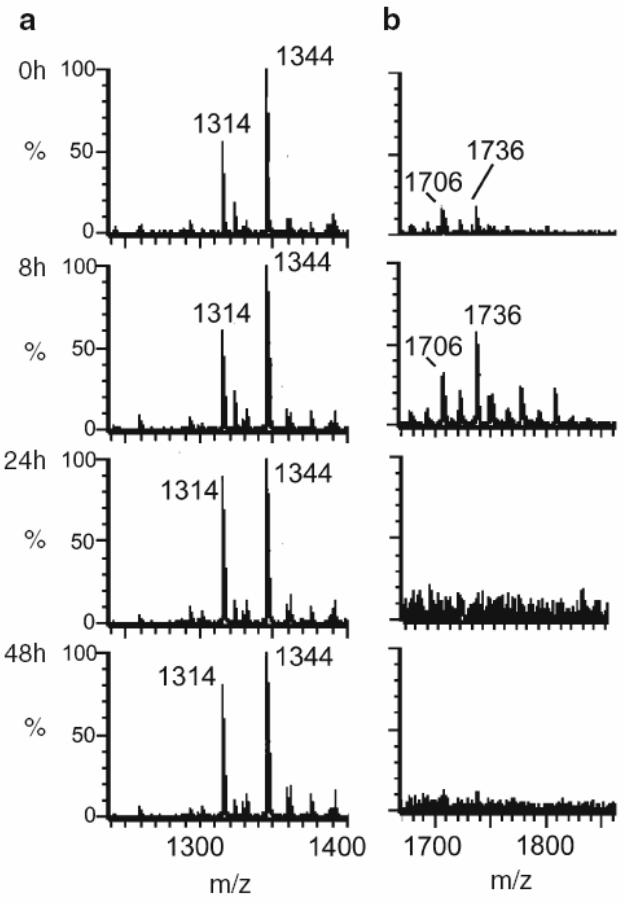


Fig. 7 MALDI-MS profiling of permethylated *O*-glycans from zebrafish embryos. The presence of **a** mono-sialylated *O*-glycans Fucα1–3GalNAcβ1–4(Neu5Acα2–3)Galβ1–3GalNAc-itol at *m/z* 1,314 and Fucα1–3GalNAcβ1–4(Neu5Gcα2–3)Galβ1–3GalNAc-itol at *m/z* 1344 and **b** di-sialylated *O*-glycans Fucα1–3GalNAcβ1–4 (Neu5Ac-Neu5Gcα2–3)Galβ1–3GalNAc-itol at *m/z* 1,706 and Fucα1–3GalNAcβ1–4(Neu5Gc-Neu5Gcα2–3)Galβ1–3GalNAc-itol at *m/z* 1,736 was checked by MALDI-MS analyses of the four embryonic stages from 0 to 48 hpf

motif could be observed prior to hatching (48 hpf). Similarly, the content of tri-sialylated motifs was reduced by more than 20 fold in 48 h. Identical methodology was applied to *N*-glycans and showed a similar trend of rapid clearance of oligosialylation along development (Fig. 8b).

In contrast to glycoprotein glycosylation, several lines of evidence demonstrated that oligosialylation in glycolipids increases along development. First, direct observation of sialylated glycolipids by MS was only possible in later stages of development (24 and 48 hpf) as previously reported. Indeed, MS profiling of total glycolipid from 0 and 8 hpf embryos exclusively showed a complex pattern of neutral fucosylated glycolipids (Table 1), but no sialylated compounds (Fig. 6). Repeated attempts to purify acidic compounds from early stages embryos failed to provide any evidence for their presence. These results were confirmed by comparing endoceramidase treated total glycolipid fractions from embryos at 0 and 48 hpf. Both samples show overall similar profiles characterized by complex mixtures of

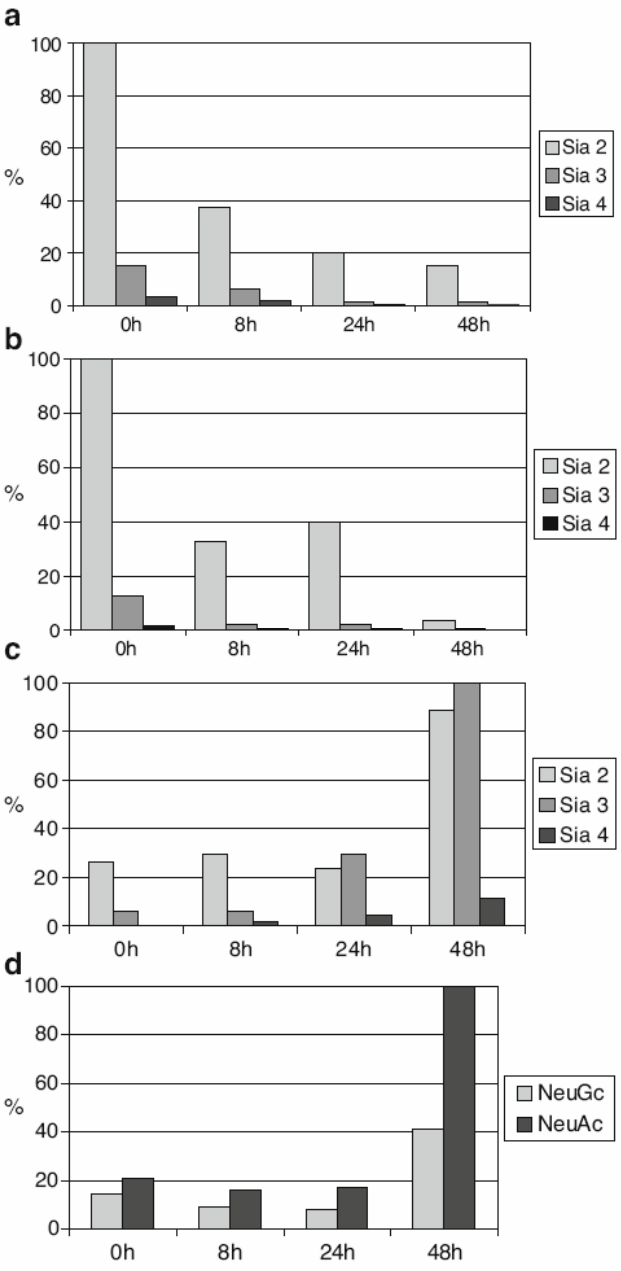


Fig. 8 Relative quantification of oligosialylation along embryos development. Proportions of Sia 2, Sia 3 and Sia 4 associated to **a** *O*-glycans, **b** *N*-glycans and **c** glycolipids were compared from 0 to 48 hpf by anion exchange DMB/HPLC-FD. Relative quantities of Neu5Ac and Neu5Gc from glycolipids were also followed along development by RP-HPLC after total release of sialic acids (**d**). Results are representative of three independent experiments

Table 1 Monosaccharide composition of neutral glycolipids calculated from MALDI-MS analysis of permethyl derivatives

m/z $[M+Na]^+$	Composition		
	HexNAc	Hex	dHex
1,941	0	4	3
2,145	0	5	3
2,186	1	4	3
2,361	1	4	4
2,391	1	5	3
2,432	2	4	3
2,473	3	3	3
2,503	3	4	2
2,565	1	5	4
2,606	2	4	4
2,637	2	5	4
2,677	3	4	3
2,810	2	5	4
2,851	3	4	4
2,923	4	4	3
2,984	2	5	5

identical neutral glycans (Sup. Fig. 2). Later stage sample shows the presence of additional major signals at *m/z* 838.6, 868.6, 1083.8 and 1113.8 attributed to $Neu5Ac_1Hex_2$, $Neu5Gc_1Hex_2$, $Neu5Ac_1Hex_2HN_1$ and $Neu5Gc_1Hex_2HN_1$, respectively. Furthermore, careful analysis of MS spectra reveals the presence of additional minor signals exclusively in later stage at *m/z* 1199.9, 1229.9 and 1259.9 attributed to $Neu5Ac_2Hex_2$, $Neu5Ac_1Neu5Gc_1Hex_2$ and $Neu5Gc_2Hex_2$ as well as *m/z* 1145.0, 1475.1 and 1505.1 attributed to $Neu5Ac_2Hex_2HN_1$, $Neu5Ac_1Neu5Gc_1Hex_2HN_1$ and $Neu5Gc_2Hex_2HN_1$. The chemical nature of oligosialylated motifs in endoceramidase treated glycolipids was confirmed by MS/MS sequencing of their permethylated derivatives. Indeed, fragmentation of molecular ions at *m/z* 1199, 1229 and 1259 showed the presence of $Neu5Ac_2$, $Neu5Ac_1Neu5Gc_1$ and $Neu5Gc_2$ sequences owing to the presence of B/Y ion pairs at *m/z* 760/463, 790/463 and 820/463, respectively (Sup. Fig. 3). Furthermore, presence of both Y ions at *m/z* 825 and 855 established that compound $Neu5Ac_1Neu5Gc_1Hex_2$ at *m/z* 1229 is a mixture of the two isobaric structures Neu5Gc–Neu5Ac–Hex–Hex and Neu5Ac–Neu5Gc–Hex–Hex. In agreement with MS analysis, profiling by DMB/HPLC-FD demonstrated that oligosialylation increased along development, with a sharp increase between 24 and 48 hpf (Fig. 8c). Indeed, as compared with embryos at 0 hpf, quantity of diSia increased by 3.5 fold and triSia by 17 fold in 48 hpf embryos. Accordingly, the total amount of sialic acid at 48 hpf is 4 to 6 higher than in other stage (Fig. 8d).

These data clearly demonstrate that the amount of oligosialylation in embryos varies along development. Surprisingly, comparative profiling of glycoconjugates established that the overall content of sialylation decreases for glycoproteins along embryogenesis, but increases for glycolipids.

Changes of α2,8-sialylation status of glycolipids is directly correlated with α2,8-sialyltransferase mRNA levels, whereas sialidase activities might be responsible glycoproteins

To investigate whether ST8Sia genes may participate in the di-, oligo- and polysialylation of major glycoconjugates during the zebrafish development, we examined the mRNA levels of the α2,8-sialyltransferase genes (ST8Sia I, ST8Sia II, ST8Sia III, ST8Sia IV, ST8Sia V and ST8Sia VI) in different developmental stages by real time PCR. Six human ST8Sia orthologues were previously identified in the zebrafish genome [14] and were molecularly cloned from various zebrafish organs by RT-PCR and sequenced. Since β-actin showed variable level of expression among the various developmental stages with the same amount of total cDNA (data not shown), we have chosen absolute quantification to quantify ST8Sia. Except for ST8Sia II, all the ST8Sia genes showed very low level of expression at 0 hpf (Fig. 9b). They all increase along embryogenesis from 0 to 36 hpf but at very different rates (Fig. 9a). Indeed, expression levels of ST8Sia I and ST8Sia VI sharply increase in the first 14 h of development whereas other ST8Sia exhibited either more modest or delayed increase (Fig. 9a). Then, ST8Sia I and ST8Sia III reach maximum sustained expression levels at 14 hpf whereas the others show a gradual increase in expression level up to 36 hpf. Interestingly, ST8Sia I and ST8Sia V genes, the human orthologs of which are known to be involved in glycolipids sialylation, are both dramatically up regulated along development by a factor of 300 and 165, respectively. This result is in agreement with the observation of an increase of sialylation associated to glycolipids. On the contrary, the modest increased expression levels of ST8Sia II and ST8Sia IV, which are responsible for the biosynthesis of polysialic acid, and of ST8Sia III also involved in the α2,8-sialylation of glycoproteins are still slightly inconsistent with the decreased oligosialylation status observed on glycoproteins. More surprisingly is the 4000 fold increase of ST8Sia VI gene expression that is not directly correlated with an increase of Sia2 motifs synthesis on glycoproteins.

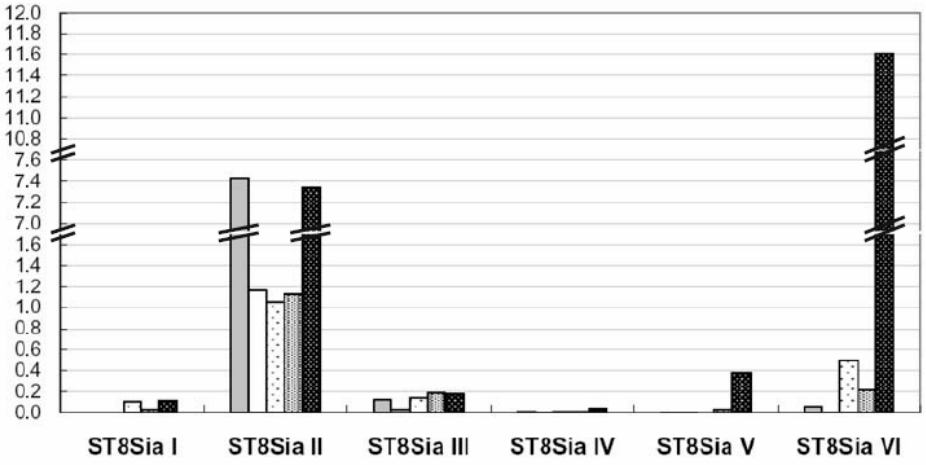
In order to check whether zebrafish sialidases might be involved in these α2,8-sialoglycoprotein metabolism, we performed sialidase assays using 4-methyl-umbelliferyl-Neu5Ac (4MU-Neu5Ac) as a substrate at various pH for various embryonic developmental stages (0, 8, 24, and 48 hpf) and unfertilized eggs. Our preliminary data showed

Fig. 9 Absolute quantification of Dre ST8Sia genes expression by real-time PCR. **a** "Fold of increase" represents the relative expression quantity of each ST8Sia in different developmental stages compared with expression at 0 hpf. **b** The different expression levels of all the ST8Sia genes were analyzed with cDNA from 0 hpf *empty square*, 14 hpf ⬚, 24 hpf ▨, 36 hpf ■ eggs and ovary *filled square* by quantitative real-time PCR. The absolute amount of transcripts was quantified according to the same DNA fragment amplified and cloned in the plasmid. Values are the mean of triplicate points

a

	ST8Sia I	ST8Sia II	ST8Sia III	ST8Sia IV	ST8Sia V	ST8Sia VI
unf	0.15	6.38	3.91	1.82	2.79	21.65
0h	1.00	1.00	1.00	1.00	1.00	1.00
14h	266.78	0.91	4.51	1.19	0.46	173.39
24h	92.78	0.97	5.77	1.55	13.00	75.18
36h	303.74	6.31	5.53	4.58	165.56	4059.40

b



the existence of intense sialidase activities associated to embryos. Survey of sialidase activities showed dramatic differences of intensities depending on the pH and along development (Sup. Fig. 4). The fact that the evolution of total sialidase activities measured at different pH follows different trends strongly suggest the presence of different enzymes, as recently demonstrated by the identification of several genes coding for potential sialidases in zebrafish [34]. Indeed, the sialidase activities detected at pH 5 sharply decreased along the zebrafish development whereas sialidase activities detected at pH 4 increased slightly along the zebrafish development. These data demonstrate a tight regulation of sialidase activities along development and suggest a possible involvement of sialidases in the α2,8-sialylation status of sialoglycoproteins, which may explain the discrepancies observed between the regulation of ST8Sia genes involved in the glycoproteins biosynthesis and the actual synthesis of oligosialylated motifs on glycoproteins.

Discussion

A previous study of the glycosylation profiles of zebrafish embryos demonstrated that this organism synthesizes a vast panel of unusual sialylated glycoconjugates. Of particular interest are the β4-galactosylated sialyl Lewis x, Galβ1–4(Siaα2–3)Galβ1–4(Fucα1–3)GlcNAc, motif observed on all complex *N*-glycans and Fucα1-3GalNAcβ1–4(Siaα2–3)Galβ observed on *O*-glycans. The Galβ1–4Galβ1–4GlcNAc motif now appears to be a common feature of several fish glycoproteins, as described in *Oryzias melastigma*, *Tribolodon hakonensis* and *Oryzia latipes* [35, 36], but the α2–3Neu5Ac sialylated Galβ1–4Galβ1–4GlcNAc motif observed in *D. rerio* was additionally identified only in *O. latipes*. However, its Neu5Gc substituted equivalent is specific to *D. rerio* so far. Similarly, the Fucα1–3GalNAcβ1–4(Siaα2–3)Galβ motif has been exclusively identified on the *O*-glycans from *D. rerio* although a wide range of other *O*-glycan structures have been described in fishes [37].

The present study extended further our knowledge of the fine structures of sialylated glycans synthesized by zebrafish along embryonic development by focusing on the α2,8-sialylation pattern (quality and quantity) of glycoproteins and glycolipids. Our data clearly established that both glycoproteins and glycolipids were α2,8-sialylated. Surprisingly, fine structural analysis demonstrated that the major glycolipids and glycoproteins presented different patterns of oligosialylation. Whilst the α2,8-sialylated glycolipids may be substituted by mixtures of oligo (Neu5Ac), oligo(Neu5Gc) and hybrid type oligo(Neu5Ac, Neu5Gc) sequences, glycoproteins are mainly substituted by oligo(Neu5Gc). Indeed, although Neu5Ac and Neu5Gc

are present in similar proportions on mono-sialylated motifs of *N*-glycans [25], we have now demonstrated that only Neu5Gc is further elongated by other sialic acids to form oligo(Neu5Gc) sequences. Identical biosynthetic restrictions seems to prevail during the extension of oligosialic acids associated with *O*-glycans that prevent formation of oligo(Neu5Ac). These differences between glycoproteins and glycolipids were also reflected in their respective contents of sialic acids: two to four times more Neu5Gc than Neu5Ac in glycoproteins and about twice as more Neu5Ac than Neu5Gc for glycolipids. As previously reported, polysialic acid structures of fish egg glycoproteins are exquisitely species specific, differing by their extent, composition, acetylation and sequences [9]. However, to our knowledge, zebrafish represents the only documented example of the preferential use of Neu5Gc over Neu5Ac for oligosialic acid elongation found on glycoproteins. This suggests an α2,8-sialyltransferase activity dedicated to α2,8-sialylation of glycoproteins, such as ST8Sia II, ST8Sia III, ST8Sia IV or ST8Sia VI, that would exhibit a preference for CMP-Neu5Gc over CMP-Neu5Ac. In mammals, ST8Sia II and ST8Sia IV are known to be involved primarily in the polysialylation of the *N*-glycans of N-CAM, ST8Sia III catalyzes the transfer of one to seven sialic acid residues onto glycoproteins or glycolipids whereas ST8Sia VI catalyzes the transfer of a unique sialic acid residue on the *O*-glycans of glycoproteins [16, 38]. In this context, it is also interesting to note that the zebrafish ST8Sia IV shows very low capability to transfer Neu5Ac from CMP-Neu5Ac onto N-CAM compared to the zebrafish ST8Sia II [39]. This last observation might reflect a preference of the zebrafish ST8Sia IV for CMP-Neu5Gc over CMP-Neu5Ac. Also, it is noteworthy that the major glycoprotein associated α2,8-sialyl motifs in zebrafish exhibit a significantly lower DP compared to those previously identified on the polysialylated glycoprotein (PSGP) of other fish species. Indeed, whereas salmonids eggs contain polysialyl units with chain length up to 20 residues [8, 37], the major sialylated glycoproteins of zebrafish eggs are mainly substituted by diSia (DP=2) motif as well as minute amounts of oligoSia (3<DP<6).

In addition, expression of oligosialylation on glycolipids and glycoproteins is differentially regulated along embryonic development. Indeed, the extent of α2,8-sialylation on major glycoproteins sharply decreases whereas that of glycolipid increases along development. Surveys of sialylation by MS analyses of intact and endoceramidase treated glycolipids, as well as the quantification of sialic acids and oligosialic acids, all indicated that significant sialylation specifically and reproducibly appears around 24 hpf, which strongly suggests that glycolipid associated α2,8-sialylation is triggered during early development. Accordingly, we failed to detect significant glycolipid associated sialylation in mature ovaries before spawning (data not shown). It is noteworthy that the complex neutral glycolipids observed in all development stages are apparently not the substrates for sialylation events occurring in later developmental stages. MS analyses demonstrated that the sialylated glycolipids of later stages comprised mainly Neu5Ac/Neu5Gc substituted $(Sia)_{1-4}$LacCer, GM2 and GD2 with no sialylated equivalents of the fucosylated neutral glycolipids (Table 1), suggesting that these two families of compounds are independently synthesized.

The appearance of glycolipids associated α2,8-sialylation in later stages of embryonic development positively correlates with the up-regulation of genes coding for α2,8-sialyltransferases ST8Sia I, ST8Sia III and ST8Sia V (Fig. 9). The human recombinant enzymes have been shown to be involved in the biosynthesis of gangliosides GD3, GT3, GD1a, GT1b and GQ1c (reviewed in [16]) [40–46]. The onset expression of these genes starts around 10 hpf and is essentially located in the developing brain (Chang *et al.* 2008, this issue). Up-regulation of oligosialylation along zebrafish development is in agreement with previous observations made on *Xenopus laevis* showing by *in vitro* assays that ST8Sia I (SAT-2) and ST8Sia V (SAT-4) activities were dramatically increased along the early development, with a maximum activity at day 4 [47]. However, whereas quantities of both neutral glycolipids and gangliosides sharply increase in *X. laevis*, only gangliosides appear to be up-regulated in *D. rerio* [48]. In contrast to glycolipids, evolution of glycoproteins associated α2,8-sialylation does not correlate with the temporal expression and the general increase of mRNAs of ST8Sia II, ST8Sia IV and ST8Sia VI from 10 hpf along the early stages of zebrafish development, therefore implicating a different kind of regulation. These latter ST8Sia genes were found to be expressed in the ovaries (Fig. 9), suggesting that the α2,8-sialylated glycoproteins of the zebrafish embryo detected at very early developmental stages well before the onset zygotic expression of these ST8Sia (around 10 hpf), might be inherited from the mother. We hypothesize that these inherited α2,8-sialylated glycoproteins could be subsequently degraded in the embryos by endogenous sialidases.

As a first step towards substantiating this hypothesis, we have essayed the sialidase activities with synthetic substrates at various pH (3, 4, 5, 6, 7) in unfertilized eggs. Since the higher and lower activities were obtained at pH 5 and 4 respectively, we then assayed the sialidase activities at pH 4 and 5 for the various developmental stages of interest. Our preliminary data showed that, sialidase activities found at pH 4 increased along zebrafish development, whereas sialidase activities found at pH 5 decreased (Sup. Fig. 4). The correlation between the detected sialidases activities and glycoproteins oligosialylation along

embryonic development therefore remain equivocal and the implicated zebrafish sialidases still need to be clearly identified. A recent study of Manzoni *et al.* [34] reported the identification of seven zebrafish sialidase genes homologous to three of the four known human genes (NEU1, NEU2, NEU3, NEU4) The zebrafish neu3.1, neu3.3 and neu4 were shown to be active towards gangliosides at very low pH (2–3) and the corresponding genes were found to be expressed differentially along the embryonic development. On the other hand, an additional NEU3 orthologous gene named neu3.2 has been described, which started to be expressed at 24 hpf. The corresponding enzyme appears to be soluble in the cytosol and active at higher pH (5.5). However, fine enzymatic characterization of all the zebrafish sialidases still awaits studies.

Such an up-regulation of the ST8Sia II gene expression along embryonic development of zebrafish has been previously reported in the context of an increased synthesis of polysialic acid chains (PSA) on the *N*-glycans of the neural cell adhesion molecule (N-CAM), which reaches a maximum around 27–40 hpf [49]. Our present study focused instead on the global α2,8-sialylation status of the different classes of glycoconjugates along embryonic zebrafish development and has identified a rapid decrease of glycoproteins associated α2,8-sialylation content. This might be explained by the large quantities of PSGP synthesized in the cortical alveoli during oogenesis in fishes [24] compared to the natural low abundance and restricted localization of PSA-N-CAM. It is most probable that the N-CAM polysialylation pattern cannot be discriminated from the one of PSGP or other polysialylated glycoproteins by a global approach.

As a premise to the identification of other oligosialylated glycoproteins in zebrafish fertilized eggs, we have assessed the oligosialylation patterns of glycoproteins in different compartments of the fertilized oocyte: embryo tissue, chorion and perivitelline space. Semi-quantification analysis by DMB-HPLC revealed that about 94% of the total oligoSia in 1 hpf fertilized oocyte was associated with soluble glycoconjugates in the perivitelline space, 5% in the chorion and less than 1% in the embryonic tissue (data not shown). Surprisingly, the extent of oligosialylation distributed within each of these fractions was very different. The perivitelline space associated components show a very short DP distribution dominated by DP 2, reminiscent of the one observed on total glycoprotein fraction, whereas the chorion and embryonic tissues exhibit more evenly distributed patterns with up to DP 10 (Sup. Fig. 5). Each fraction was further differentiated by their sialic acids content, as embryo associated glycoproteins was almost exclusively composed of Neu5Gc (Neu5Gc/Neu5Ac 14:1) whereas perivitelline space associated glycoproteins of Neu5Gc and Neu5Ac in a 4:1 ratio (data not shown).

The presence of large quantities of oligosialylated soluble glycoproteins in perivitelline space of zebrafish is in good agreement with previous reports of the of polysialylated peptides originating from a fertilization induced proteolysis of cortical alveoli PSGP in several other species of fish, including trouts, salmonids and medaka fish [9, 24, 50]. Accordingly, we observed large quantities of protein associated oligoSia chains in zebrafish mature ovaries that exhibit a distribution pattern and a composition identical to that of the soluble glycoproteins of fertilized oocytes (data not shown). However, despite the postulated conservation of this phenomenon among fishes and the large quantity of excreted PSGP, the destiny and the function of these compounds are still largely unknown. Altogether, our data established that the vast majority of oligosialylation observed in fertilized oocytes are synthesized prior to embryogenesis in the mother ovary and then degraded along embryonic development for a yet unknown purpose.

The observation of a protein associated oligosialylation in embryonic tissue as early as 1 hpf, thus long before that N-CAM associated PSA appears, strongly suggests the presence of so far undescribed oligosialylated glycoproteins in oocytes. Diversity of polysialylated components in fish embryos increases the possible endogenous substrates for zebrafish α2,8-sialyltransferases. Indeed, ST8Sia II and IV were shown to be involved in the synthesis of polysialic acid chains on both N-CAM *N*-glycans in zebrafish and on PSGP *O*-glycans in rainbow trout, which suggest that these enzymes might have multiple substrates acceptors in various animal species [49, 51]. Altogether, the results presented here establish the structural bases for the investigation of the fine enzymatic specificities of the different ST8Sia identified from zebrafish genome [14]. These studies are actually in progress by using the endogenous molecules as acceptor substrates. Further work is also needed to identify the protein carriers of the oligosialylated motifs and to assess their localization within the fertilized embryo. The collected data will enable a better evaluation of the importance of sialyltransferases as well as sialidases in the regulation of synthesis of sialylated motifs along zebrafish embryogenesis.

Acknowledgement This research was supported by CNRS (Centre National de la Recherche Scientifique), EGIDE, the Ministère des Affaires Etrangères and PPF Bioinformatique de Lille, and Academia Sinica, Taiwan. We acknowledge Dr. Estelle Garénaux and Dr. Bernadette Coddeville for their help on mass spectrometry analyses of derivatized oligosialic acids and Yi-Cyun CHEN for help in preparation of zebrafish embryos. We acknowledge also the technical assistance of Anne-Marie Mir, Beatrice Catieaux and Dr. Sandrine Duvet, and Pr. Yu-Teh Li for helpful discussion. *D. rerio* embryos were provided by Dr. Bernard and Christine Thisse (IGBMC, Strasbourg) and a set of staged cDNA were kindly prepared by Hafiz Hamed and Gerardo Vasta (Baltimore, USA) that were used for quantification of sialyltransferase mRNA levels.

Abbreviations

Sialyltransferases nomenclature is according to Tsuji *et al.* [52], gangliosides nomenclature is according to Svennerholm [53]. hpf|hours post-fertilization;diSia|disialyl motif;triSia|trisalylmotif;DHB|2,5-dihydroxybenzoic acid;DP|degree of polymerisation;GC|gas chromatography;Hex|hexose;HexNAc|*N*-acetyl hexosamine;HexNAcitol|reduced *N*-acetyl hexosaminitol;LacCer|lactosylceramide;MALDI|matrix-assisted laser-desorption ionization;MS|mass spectrometry;TFA|trifluoroacetic acid;TOF|time-of-flight

References

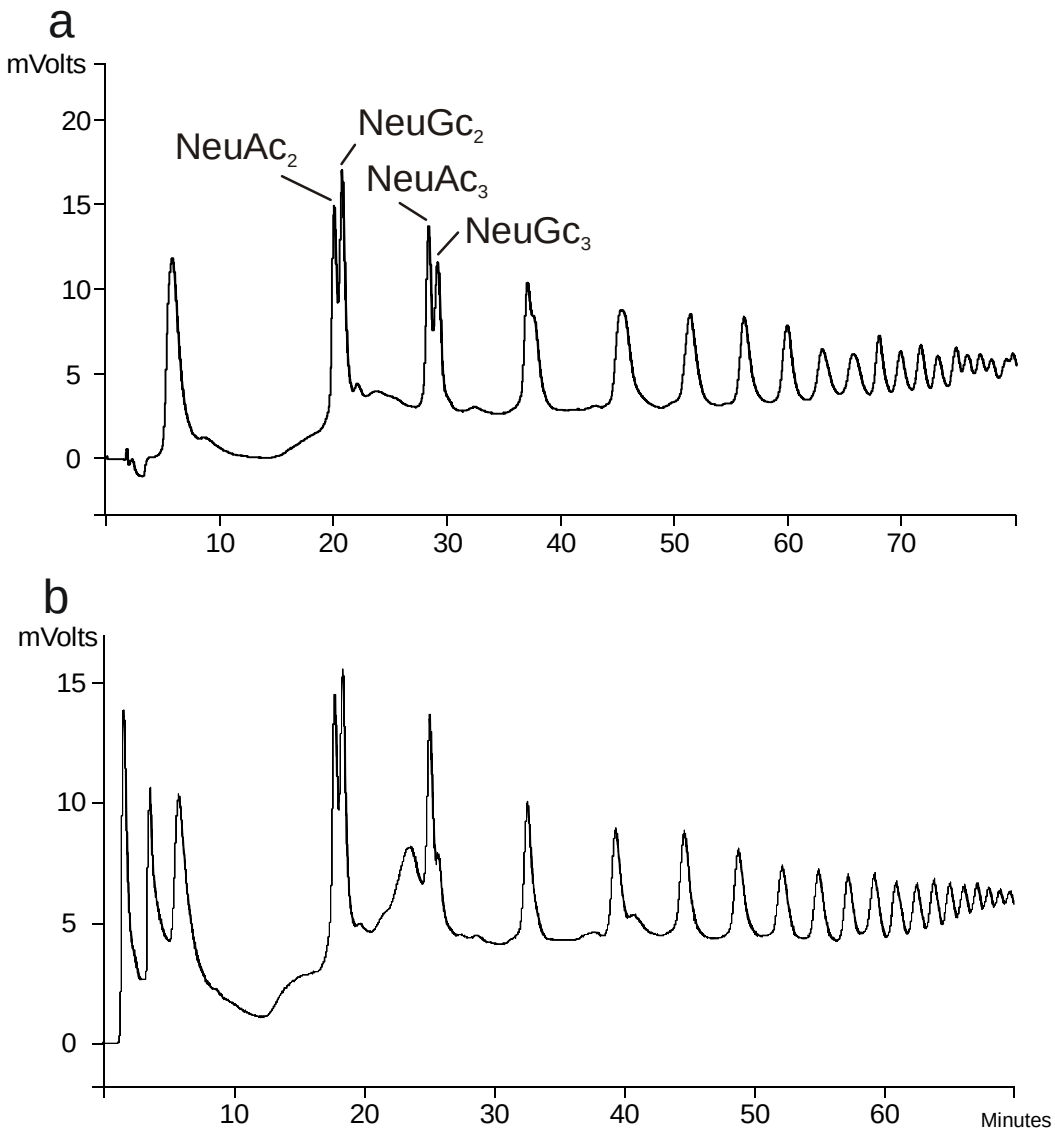
1. Schauer, R., Kamerling, J.: Chemistry, biochemistry and biology of sialic acids. In: Montreuil, J., Vliegenhart, J., Schachter, H. (eds.) Glycoproteins II, pp. 243–372. Elsevier, Amsterdam (1997)
2. Finne, J., Krusius, T., Rauvala, H.: Occurrence of disialosyl groups in glycoproteins. Biochem. Biophys. Res. Commun. **74**, 405–410 (1977). doi:10.1016/0006-291X(77)90318-7
3. Handa, N., Handa, S.: The chemistry of lipids of posthemolytic residue or stroma of erythrocytes. XIV. Chemical structure of glycolipid of cat erythrocyte stroma. Jpn. J. Exp. Med. **35**, 331–341 (1965)
4. Handa, S., Burton, R.M.: Lipids of retina. I. Analysis of gangliosides in beef retina by thin layer chromatography. Lipids **4**, 205–208 (1969). doi:10.1007/BF02532630
5. Krusius, T., Finne, J., Karkkainen, J., Jarnefelt, J.: Neutral and acidic glycopeptides in adult and developing rat brain. Biochim. Biophys. Acta. **365**, 80–92 (1974)
6. Kuhn, R., Wiegandt, H.: Die konstitution der ganglio-n-tetraose und des gangliosids gl. Chem. Ber. **96**, 866–880 (1963). doi:10.1002/cber.19630960329
7. McGuire, E.J., Binkley, S.B.: The structure and chemistry of colominic acid. Biochemistry **3**, 247–251 (1964). doi:10.1021/bi00890a017
8. Inoue, S., Iwasaki, M.: Characterization of a new type of glycoprotein saccharides containing polysialosyl sequence. Biochem. Biophys. Res. Commun. **93**, 162–165 (1980). doi:10.1016/S0006-291X(80)80260-9
9. Sato, C., Kitajima, K., Tazawa, I., Inoue, Y., Inoue, S., Troy 2nd, F.A.: Structural diversity in the alpha 2®8-linked polysialic acid chains in salmonid fish egg glycoproteins. Occurrence of poly (Neu5Ac), poly(Neu5Gc), poly(Neu5Ac, Neu5Gc), poly(KDN), and their partially acetylated forms. J. Biol. Chem. **268**, 23675–23684 (1993)
10. Finne, J.: Occurrence of unique polysialosyl carbohydrate units in glycoproteins of developing brain. J. Biol. Chem. **257**, 11966–11970 (1982)
11. James, W.M., Agnew, W.S.: Multiple oligosaccharide chains in the voltage-sensitive Na channel from electrophorus electricus: evidence for alpha-2,8-linked polysialic acid. Biochem. Biophys. Res. Commun. **148**, 817–826 (1987). doi:10.1016/0006-291X(87)90949-1
12. Yabe, U., Sato, C., Matsuda, T., Kitajima, K.: Polysialic acid in human milk. CD36 is a new member of mammalian polysialic acid-containing glycoprotein. J. Biol. Chem. **278**, 13875–13880 (2003). doi:10.1074/jbc.M300458200
13. Curreli, S., Arany, Z., Gerardy-Schahn, R., Mann, D., Stamatos, N.M.: Polysialylated neuropilin-2 is expressed on the surface of human dendritic cells and modulates dendritic cell-T lymphocyte interactions. J. Biol. Chem. **282**, 30346–30356 (2007). doi:10.1074/jbc.M702965200
14. Harduin-Lepers, A., Mollicone, R., Delannoy, P., Oriol, R.: The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. Glycobiology **15**, 805–817 (2005). doi:10.1093/glycob/cwi063
15. Harduin-Lepers, A., Recchi, M.A., Delannoy, P.: 1994, the year of sialyltransferases. Glycobiology **5**, 741–758 (1995). doi:10.1093/glycob/5.8.741
16. Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M.A., Samyn-Petit, B., Julien, S., Delannoy, P.: The human sialyltransferase family. Biochimie **83**, 727–737 (2001). doi:10.1016/S0300-9084(01)01301-3
17. Saito, M., Yu, R.: Biochemistry and function of sialidases. In: Rosenberg, A. (ed.) Biology of Sialic Acids, pp. 261–313. Plenum, New York (1995)
18. Hakomori, S.I.: Structure and function of glycosphingolipids and sphingolipids: Recollections and future trends, Biochim Biophys Acta (2007).
19. Ngamukote, S., Yanagisawa, M., Ariga, T., Ando, S., Yu, R.K.: Developmental changes of glycosphingolipids and expression of glycogenes in mouse brains. J. Neurochem. **103**, 2327–2341 (2007). doi:10.1111/j.1471-4159.2007.04910.x
20. Yu, R.K., Macala, L.J., Taki, T., Weinfield, H.M., Yu, F.S.: Developmental changes in ganglioside composition and synthesis in embryonic rat brain. J. Neurochem. **50**, 1825–1829 (1988). doi:10.1111/j.1471-4159.1988.tb02484.x
21. Hildebrandt, H., Muhlenhoff, M., Weinhold, B., Gerardy-Schahn, R.: Dissecting polysialic acid and NCAM functions in brain development. J. Neurochem. **103**(Suppl 1), 56–64 (2007). doi:10.1111/j.1471-4159.2007.04716.x
22. Probstmeier, R., Bilz, A., Schneider-Schaulies, J.: Expression of the neural cell adhesion molecule and polysialic acid during early mouse embryogenesis. J. Neurosci. Res. **37**, 324–335 (1994). doi:10.1002/jnr.490370305
23. Asahina, S., Sato, C., Kitajima, K.: Developmental expression of a sialyltransferase responsible for sialylation of cortical alveolus glycoprotein during oogenesis in rainbow trout (*Oncorhynchus mykiss*). J. Biochem. **136**, 189–198 (2004). doi:10.1093/jb/mvh106
24. Kitazume, S., Kitajima, K., Inoue, S., Inoue, Y., Troy 2nd, F.A.: Developmental expression of trout egg polysialoglycoproteins and the prerequisite alpha 2,6-, and alpha 2,8-sialyl and alpha 2,8-polysialyltransferase activities required for their synthesis during oogenesis. J. Biol. Chem. **269**, 10330–10340 (1994)
25. Guérardel, Y., Chang, L.Y., Maes, E., Huang, C.J., Khoo, K.H.: Glycomic survey mapping of zebrafish identifies unique sialylation pattern. Glycobiology **16**, 244–257 (2006). doi:10.1093/glycob/cwj062
26. Westerfield, M.: The zebrafish book. A guide for laboratory use of zebrafish (*Danio rerio*), p. 385. University of Oregon Press, Eugene (1995)
27. Yu, S.Y., Wu, S.W., Khoo, K.H.: Distinctive characteristics of MALDI-Q/TOF and TOF/TOF tandem mass spectrometry for sequencing of permethylated complex type *N*-glycans. Glycoconj. J. **23**, 355–369 (2006). doi:10.1007/s10719-006-8492-3
28. Sato, C., Inoue, S., Matsuda, T., Kitajima, K.: Fluorescent–assisted detection of oligosialyl units in glycoconjugates. Anal. Biochem. **266**, 102–109 (1999). doi:10.1006/abio.1998.2921
29. Chen, W.Y., John, J.A., Lin, C.H., Chang, C.Y.: Molecular cloning and developmental expression of zinc finger transcription factor MTF-1 gene in zebrafish, *Danio rerio*. Biochem. Biophys. Res. Commun. **291**, 798–805 (2002). doi:10.1006/bbrc.2002.6517
30. Ahmed, H., Du, S.J., O'Leary, N., Vasta, G.R.: Biochemical and molecular characterization of galectins from zebrafish (*Danio rerio*): notochord-specific expression of a prototype galectin during early embryogenesis. Glycobiology **14**, 219–232 (2004). doi:10.1093/glycob/cwh032
31. Inoue, S., Inoue, Y.: A challenge to the ultrasensitive chemical method for the analysis of oligo- and polysialic acids at a

nanogram level of colominic acid and a milligram level of brain tissues. Biochimie **83**, 605–613 (2001). doi:10.1016/S0300-9084(01)01307-4

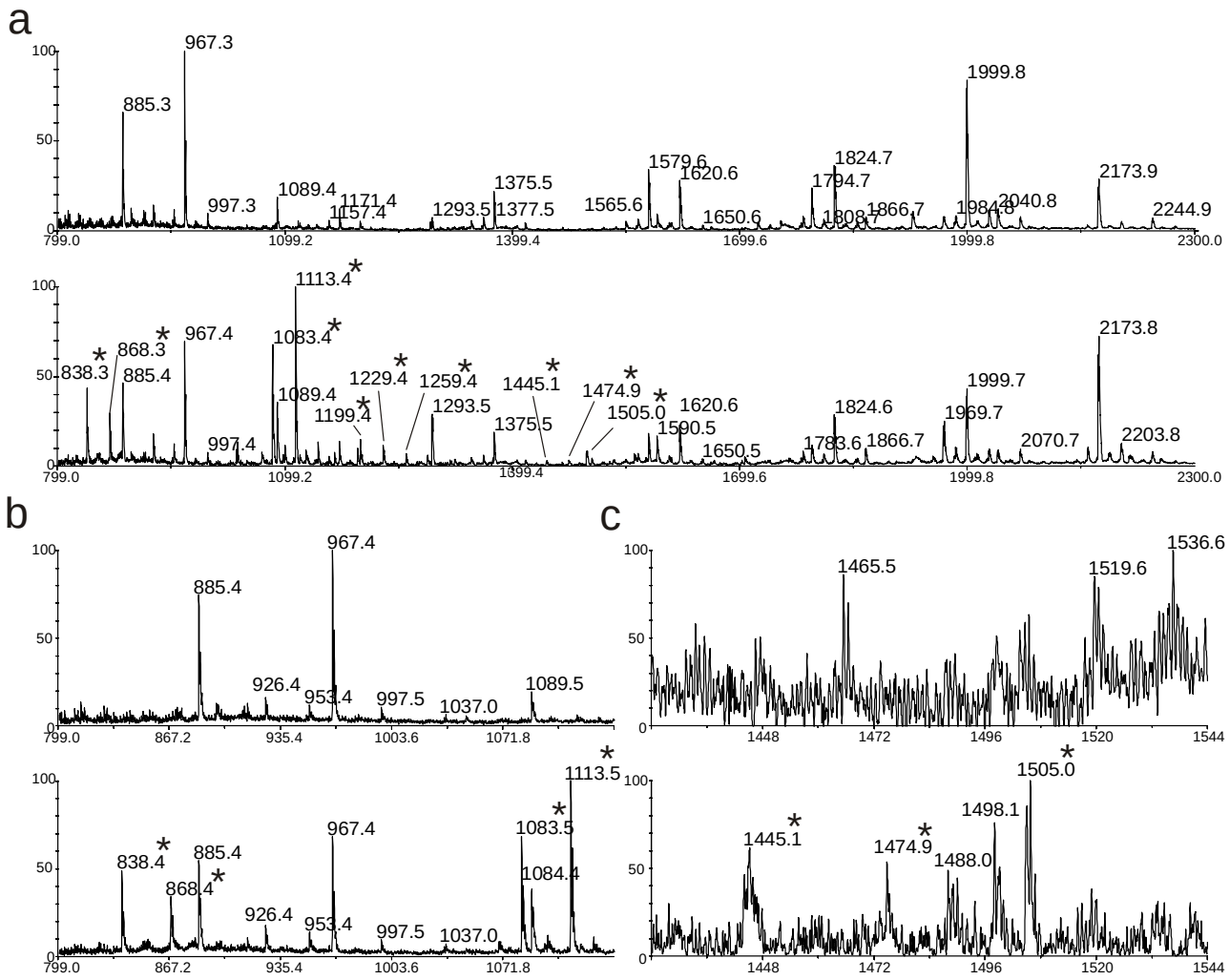
32. Inoue, S., Lin, S.L., Lee, Y.C., Inoue, Y.: An ultrasensitive chemical method for polysialic acid analysis. Glycobiology **11**, 759–767 (2001). doi:10.1093/glycob/11.9.759
33. Sato, C., Inoue, S., Matsuda, T., Kitajima, K.: Development of a highly sensitive chemical method for detecting alpha2®8-linked oligo/polysialic acid residues in glycoproteins blotted on the membrane. Anal. Biochem. **261**, 191–197 (1998). doi:10.1006/abio.1998.2718
34. Manzoni, M., Colombi, P., Papini, N., Rubaga, L., Tiso, N., Preti, A., *et al.*: Molecular cloning and biochemical characterization of sialidases from zebrafish (*Danio rerio*). Biochem. J. **408**, 395–406 (2007). doi:10.1042/BJ20070627
35. Inoue, S., Iwasaki, M., Ishii, K., Kitajima, K., Inoue, Y.: Isolation and structures of glycoprotein-derived free sialooligosaccharides from the unfertilized eggs of *Tribolodon hakonensis*, a dace. Intracellular accumulation of a novel class of biantennary disialooligosaccharides. J. Biol. Chem. **264**, 18520–18526 (1989)
36. Iwasaki, M., Seko, A., Kitajima, K., Inoue, Y., Inoue, S.: Fish egg glycophosphoproteins have species-specific N-linked glycan units previously found in a storage pool of free glycan chains. J. Biol. Chem. **267**, 24287–24296 (1992)
37. Inoue, S., Inoue, Y.: Fish glycoproteins. In: Montreuil, J., Vliegenthart, J., Schachter, H. (eds.) Glycoproteins II, pp. 143–161. Elsevier, Amsterdam (1997)
38. Teintenier-Lelievre, M., Julien, S., Juliant, S., Guerardel, Y., Duonor-Cerutti, M., Delannoy, P., *et al.*: Molecular cloning and expression of a human hST8Sia VI (α2,8-sialyltransferase) responsible for the synthesis of the diSia motif on *O*-glycosylproteins. Biochem. J. **392**, 665–674 (2005). doi:10.1042/BJ20051120
39. Marx, M., Rivera-Milla, E., Stummeyer, K., Gerardy-Schahn, R., Bastmeyer, M.: Divergent evolution of the vertebrate polysialyltransferase Stx and Pst genes revealed by fish-to-mammal comparison. Dev. Biol. **306**, 560–571 (2007). doi:10.1016/j.ydbio.2007.03.032
40. Haraguchi, M., Yamashiro, S., Yamamoto, A., Furukawa, K., Takamiya, K., Lloyd, K.O., *et al.*: Isolation of GD3 synthase gene by expression cloning of GM3 α2,8-sialyltransferase cDNA using anti-GD2 monoclonal antibody. Proc. Natl. Acad. Sci. USA. **91**, 10455–10459 (1994). doi:10.1073/pnas.91.22.10455
41. Nara, K., Watanabe, Y., Maruyama, K., Kasahara, K., Nagai, Y., Sanai, Y.: Expression cloning of a CMP-NeuAc: NeuAcα2–3Galb1–4Glcb1–1¢Cer α2,8-sialyltransferase (GD3 synthase) from human melanoma cells. Proc. Natl. Acad. Sci. USA. **91**, 7952–7956 (1994). doi:10.1073/pnas.91.17.7952
42. Sasaki, K., Kurata, K., Kojima, N., Kurosawa, N., Ohta, S., Hanai, N., *et al.*: Expression cloning of a GM3-specific α2,8-sialyltransferase (GD3 synthase). J. Biol. Chem. **269**, 15950–15956 (1994)
43. Daniotti, J.L., Rosales Fritz, V., Kunda, P., Nishi, T., Maccioni, H.J.: Cloning, characterization and developmental expression of alpha2,8 sialyltransferase (GD3 synthase, ST8Sia I) gene in chick brain and retina. Int. J. Dev. Neurosci. **15**, 767–776 (1997). doi:10.1016/S0736-5748(97)00027-0
44. Lee, Y.C., Kim, Y.J., Lee, K.Y., Kim, K.S., Kim, B.U., Kim, H.N., *et al.*: Cloning and expression of cDNA for a human Sia alpha 2,3Gal beta 1, 4GlcNA:alpha 2,8-sialyltransferase (hST8Sia III). Arch. Biochem. Biophys. **360**, 41–46 (1998). doi:10.1006/abbi.1998.0909
45. Kim, Y.J., Kim, K.S., Do, S., Kim, C.H., Kim, S.K., Lee, Y.C.: Molecular cloning and expression of human alpha2,8-sialyltransferase (hST8Sia V). Biochem. Biophys. Res. Commun. **235**, 327–330 (1997). doi:10.1006/bbrc.1997.6725
46. Kono, M., Yoshida, Y., Kojima, N., Tsuji, S.: Molecular cloning and expression of a fifth type of α2,8-sialyltransferase (ST8Sia V). Its substrate specificity is similar to that of SAT-V/III, which synthesize GD1c, GT1a, GQ1b and GT3. J. Biol. Chem. **271**, 29366–29371 (1996). doi:10.1074/jbc.271.46.29366
47. Gornati, R., Basu, S., Bernardini, G., Rizzo, A.M., Rossi, F., Berra, B.: Activities of glycolipid glycosyltransferases and sialidases during the early development of *Xenopus laevis*. Mol. Cell. Biochem. **166**, 117–124 (1997). doi:10.1023/A:1006891414663
48. Gornati, R., Rizzo, A.M., Tong, X.W., Berra, B., Bernardini, G.: Glycolipid patterns during Xenopus embryo development. Cell Biol. Int. **19**, 183–189 (1995). doi:10.1006/cbir.1995.1061
49. Marx, M., Rutishauser, U., Bastmeyer, M.: Dual function of polysialic acid during zebrafish central nervous system development. Development **128**, 4949–4958 (2001)
50. Taguchi, T., Seko, A., Kitajima, K., Inoue, S., Iwamatsu, T., Khoo, K.H., *et al.*: Structural studies of a novel type of tetraantennary sialoglycan unit in a carbohydrate-rich glycopeptide isolated from the fertilized eggs of Indian Medaka fish, *Oryzias melastigma*. J. Biol. Chem. **268**, 2353–2362 (1993)
51. Asahina, S., Sato, C., Matsuno, M., Matsuda, T., Colley, K., Kitajima, K.: Involvement of the α2,8-polysialyltransferases II/STX and IV/PST in the biosynthesis of polysialic acid chains on the O-linked glycoproteins in rainbow trout ovary. J. Biochem. **140**, 687–701 (2006). doi:10.1093/jb/mvj200
52. Tsuji, S., Datta, A.K., Paulson, J.C.: Systematic nomenclature for sialyltransferases. Glycobiology **6**, v–vii (1996). doi:10.1093/glycob/6.7.647
53. Svennerholm, L.: The gangliosides. J. Lipid Res. **5**, 145–155 (1964)

Supplemental material

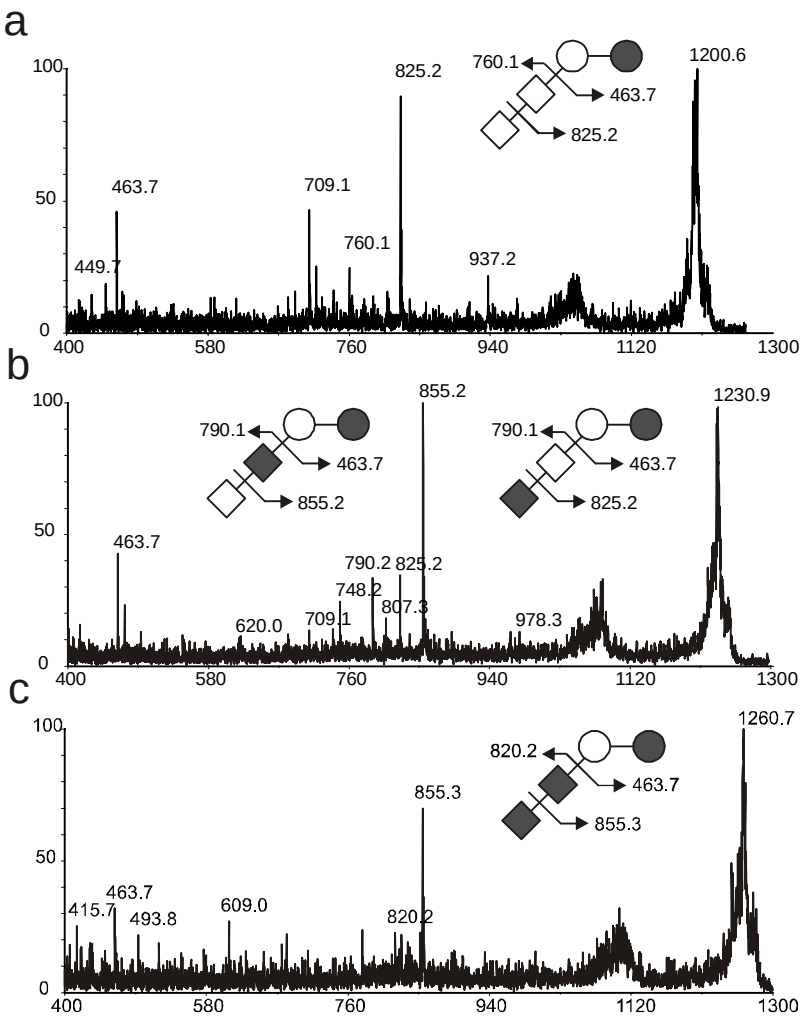
Sup. Figure 1: Identification of oligosialylation on *O*-glycans by anion exchange DMB/HPLC-FD. Chromatographic profiles of co-injected DMB-derivatized (a) [-8)Neu5Ac(α2-]$_n$ and [-8)Neu5Gc(α2-]$_n$ standards, (b) [-8)Neu5Ac(α2-]$_n$ standard and oligo-Sia from zebrafish 1hpf embryos *O*-glycans, showing that the diSia (DP=2) peak from zebrafish *O*-glycans does not co-migrate with Neu5Ac(α2-8)Neu5Ac.



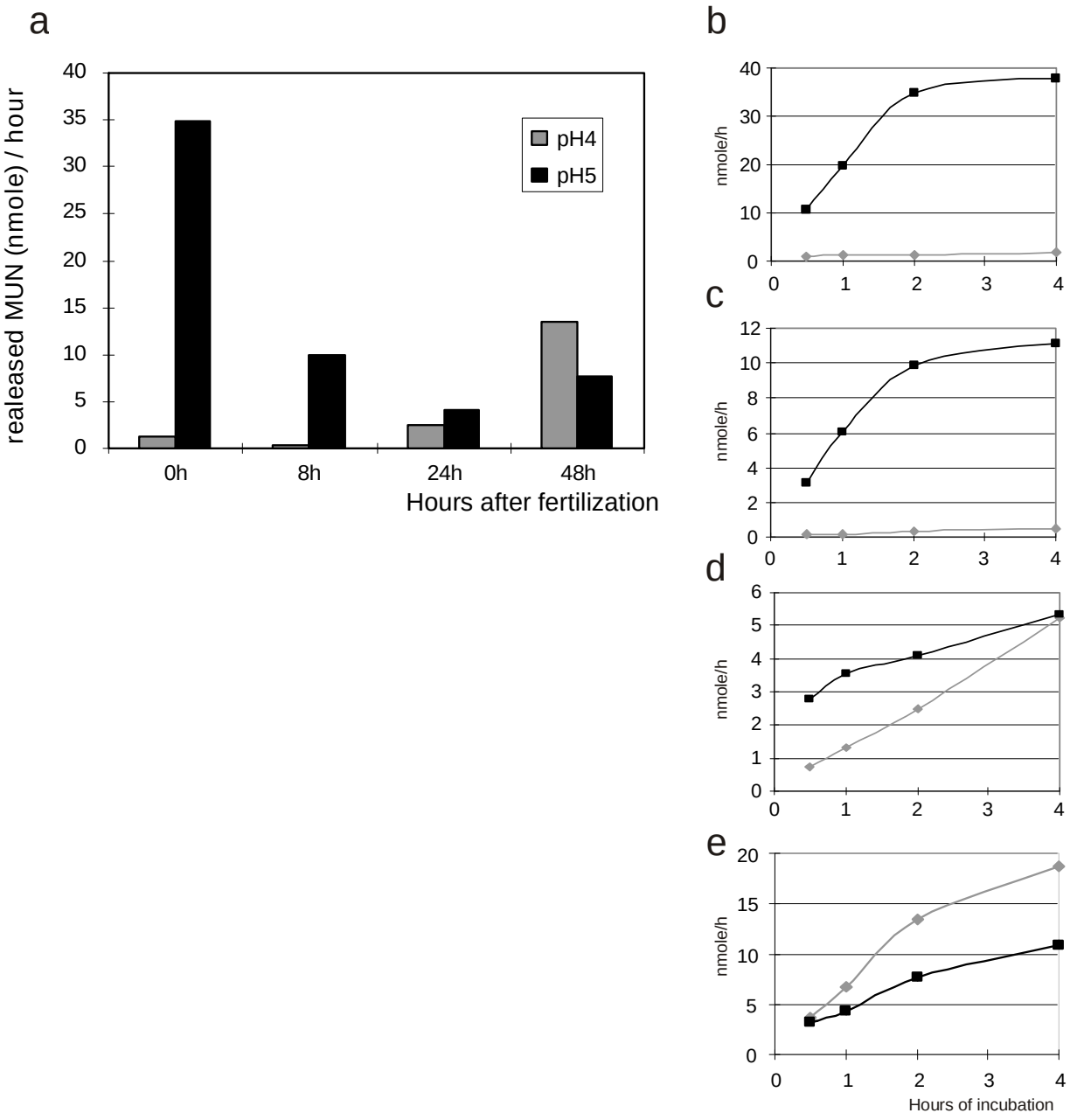
Sup. Figure 2: Comparison of glycolipid profiles of embryos along development. MALDI-MS profiles permethylated derivatives of endoceramidase treated glycolipids from (upper panels) 1hpf embryos and (lower panels) 48hpf embryos. (a) Spectra m/z 800 to 2300, (b) close up of spectra m/z 800 to 1150 and (c) close up of spectra m/z 1425 to 1545. Values labelled with an asterisk where attributed to sialylated components exclusively observed in 48 hpf embryos. m/z 838.6, $Neu5Ac_1Hex_2$; 868.6, $Neu5Gc_1Hex_2$; 1083.8, $Neu5Ac_1Hex_2HN_1$; 1113.8, $Neu5Gc_1Hex_2HN_1$; 1199.9, $Neu5Ac_2Hex_2$; 1229.9, $Neu5Ac_1Neu5Gc_1Hex_2$; 1259.9, $Neu5Gc_2Hex_2$; 1145.0, $Neu5Ac_2Hex_2HN_1$; 1475.1, $Neu5Ac_1Neu5Gc_1Hex_2HN_1$; 1505.1, $Neu5Gc_2Hex_2HN_1$.



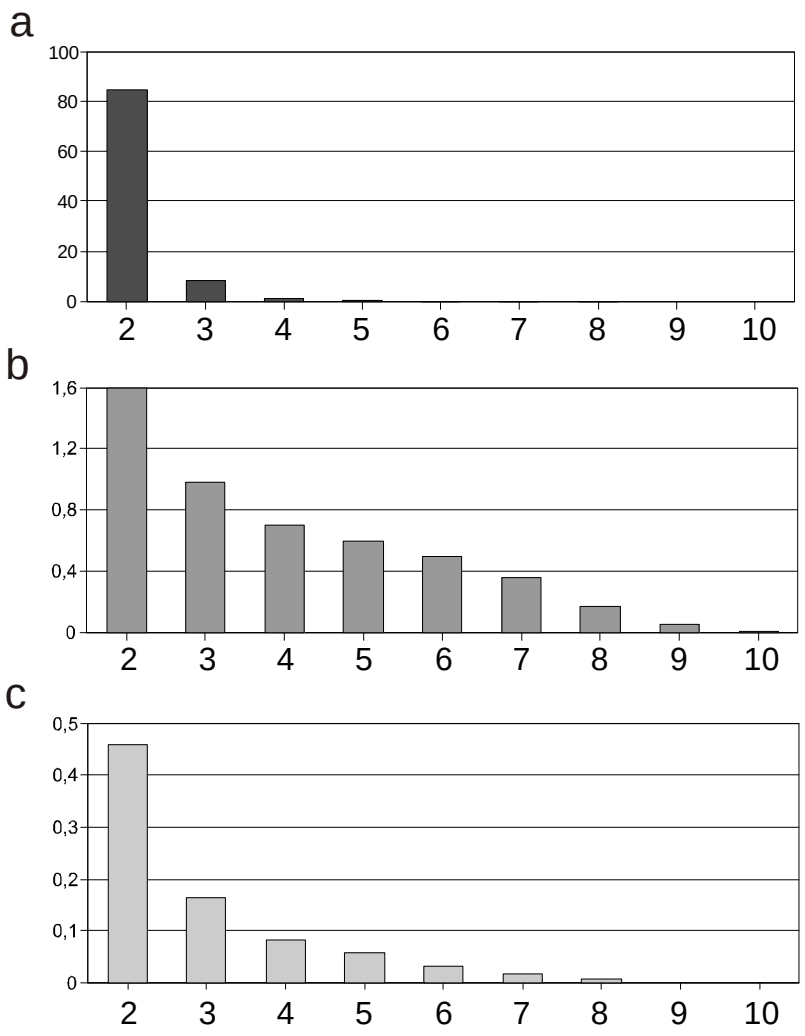
Sup. Figure 3: MALDI-MS/MS sequencing of endoceramidase treated glycolipids from zebrafish embryos. MS/MS analyses of permethylated derivatives of (a) $Neu5Ac_2Hex_2$ at m/z 1200, (b) $Neu5Ac_1Neu5Gc_1Hex_2$ at m/z 1230 and (c) $Neu5Gc_2Hex_2$ at m/z 1260. Fragmentation pattern of parent ion at m/z 1230 established the presence of two isomers differing by the respective positions of Neu5Ac and Neu5Gc within the diSia sequence. Sequences of these compounds are in agreement with their assignment as derivatives of di-sialylated lactose. Symbol used: dark circle, Glc; white circle, Gal; dark diamond, Neu5Gc; white diamond, Neu5Ac.



Sup. Figure 4: Evolution of sialidase activities along embryogenesis. Sialidase activities from total embryos at (b) 0 hpf; (c) 8 hpf; (d) 24 hpf; (e) 48 hpf were assessed at pH 4 and pH 5 for incubation times ranging from 0.5 to 4 hours, by measuring the release of 4-methyl-umbeliferone (MUN). Values of activities at each stage and pH for 2 hours incubation are summarized in (a). Black squares, pH 5; grey diamonds, pH 4.



Sup. Figure 5: Profiles of glycoprotein associated oligoSia from different locations in the fertilized oocyte. Proteins associated oligoSia were purified from (a) perivitelline space, (b) chorion and (c) embryonic tissue, then analyzed by anion exchange DMB/HPLC-FD. Intensity of each DP is plotted in percentage, 100% being the sum of oligoSia from all fractions



Sup. Table 1: Primer nucleotide sequence and expected amplicon size. Accession number in GeneBank for the identified sialyltransferases and actin sequences are indicated.

ST8Sia	Primer sequence	Accession number (GenBank)	PCR Product size (bp)
ST8Sia I	Forward 5'-TTGCGGTTACTAAGGAGA Reverse 5'-ACGAAAGATTTGCGGGAC	AJ715535	346
ST8Sia II	Forward 5'-GACTCGCACGACTTTGTT Reverse 5'-TGGTTGGTCAGCCAGTAA	AY055462	335
ST8Sia III	Forward 5'-AACAACCTGCTGACCATCC Reverse 5'-ATGATACGGCAGCTCCTT	AJ715543	354
ST8Sia IV	Forward 5'-TCTTGACTTGGGAGTTGG Reverse 5'-TCTGACCGCAATCCTACA	AJ715545	366
ST8Sia V	Forward 5'-AAATAAGGAGGAGACGGATAA Reverse 5'-AAAGTCAGAAGCGTCAAT	AJ715546	291
ST8Sia VI	Forward 5'-TGTCTATGATGGCGAAAG Reverse 5'-TGACCGTATGAATGAAGG	AJ715551	333
β-actin	Forward 5'-GTTGGTATGGGACAGAAAGA Reverse 5'-GGCGTAACCCTCGTAGAT	AF025305	378

Conclusion

En conclusion, ces travaux ont permis d'établir le poisson zèbre comme un modèle cohérent d'étude du rôle de la sialylation dans le développement embryonnaire des vertébrés, ce qui ouvre la voie à l'étude des relations entre structure et fonction des sialoconjugués. Au delà des données structurales, nos résultats suggèrent fortement que les glycoprotéines sialylées majeures dont nous observons les profils de glycosylation sont synthétisées par la mère. Celles-ci, après avoir été excrétées dans l'espace périvitellin suite à la fécondation des oocytes sont dégradées par des sialydases solubles et libèrent ainsi des quantités importantes d'acides sialiques libres dont la destinée et le rôle sont encore inconnus. Par contre, les glycolipides sialylés sont synthétisés plus tardivement par l'embryon au cours de son développement, ce qui induit un changement total du profil de glycosylation des glycolipides au cours de l'embryogenèse. Il apparaît également que les modes de sialylations des glycoprotéines et des glycolipides soient régulés différemment, comme le démontrent les différences dans les profils d'oligosialylation et les ratios NeuAc/NeuGc.

Dans le but de mieux cerner le rôle de la sialylation dans l'embryogenèse, l'expression de la CMP-Neu5Ac hydroxylase (CMAH) des embryons de poisson zèbre -l'enzyme responsable de la synthèse du Neu5Gc- a été bloquée par l'utilisation de morpholinos par l'équipe du Dr Khoo. Les embryons ainsi modifiés présentent un phénotype morphologique très clair caractérisé par un excroissance cranienne et un coude de la queue de l'embryon comme montré ci-dessous (Fig. 6).

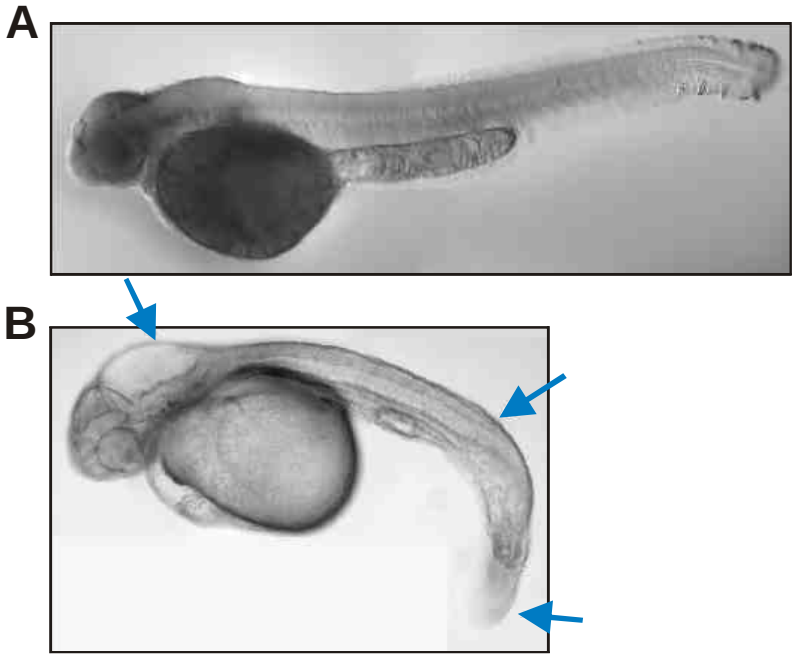


Fig. 6 : A, embryon de poisson zèbre contrôle, 48 hpf ; **B**, embryon traité aux morpholinos, 48 hpf

Il est noter que ce phénotype et en bon accord avec le fait que la CMAH soit exprimée majoritairement dans les deux zones qui présentent des anomalies de développement visibles (Fig. 7).

Fig. 7 : Hybridation *in situ* sur un embryon de poisson zèbre contrôle 48hpf avec une sonde 3'UTR du gène codant pour la CMAH

Ainsi, l'inactivation de cette enzyme laissait présager des modifications dans les profils d'expression du Neu5Ac et du Neu5Gc au sein des différents glycoconjugués. En particulier, si l'hypothèse selon laquelle les glycolipides sialylés sont bien synthétisés par l'embryon et non par la mère, les glycolipides sialylés devraient être substitués de façon homogène par du Neu5Ac et non par un mélange de Neu5Ac et de Neu5Gc comme dans les embryons contrôles. Ainsi, nous avons déterminé la structure des glycannes et leurs profils de sialylation associés aux différents stades de développement embryonnaire entre 1 et 48h post-fertilisation. De façon très surprenante, nous n'avons observé jusqu'à présent aucune différence entre les profils de sialylation des embryons contrôles et mutés, y compris sur les glycolipides. En effet, les embryons mutés incorporent autant de Neu5Gc dans leurs glycolipides que les embryons contrôles, ce qui suggère que la biosynthèse de CMP-Neu5Gc ne soit pas perturbée par l'inactivation de l'expression de la CMP-Neu5Ac hydroxylase et donc que cette enzyme soit redondante. Cette hypothèse est néanmoins fort peu probable car à notre connaissance, l'expression de la CMAH n'est pas redondante, y compris chez le poisson. Une deuxième hypothèse est que malgré l'absence de biosynthèse du CMP-Neu5Gc, l'embryon puisse incorporer du Neu5Gc d'origine exogène dans ses propres glycoconjugués. L'incorporation de Neu5Gc d'origine exogène ainsi que sa réexpression sur des glycoconjugués membranaires a en effet été démontrée et serait à l'origine de la synthèse de glycoconjugués contenant du Neu5Gc chez l'homme, bien que la CMAH humaine soit inactive (Bardor *et al.*, 2005). Sur ce modèle, il est possible d'imaginer que l'acide sialique libéré des glycoprotéines excrétées dans l'espace périvitellin puisse être réincorporé par l'embryon pour servir de source d'acides sialiques au cours de l'embryogenèse. Sur la base des travaux de Inoue effectués sur d'autres

espèces de poissons (salmonidés et 'medaka fish') et de nos études sur le poisson zèbre, le processus pourrait être représenté comme dans la Fig. 8.

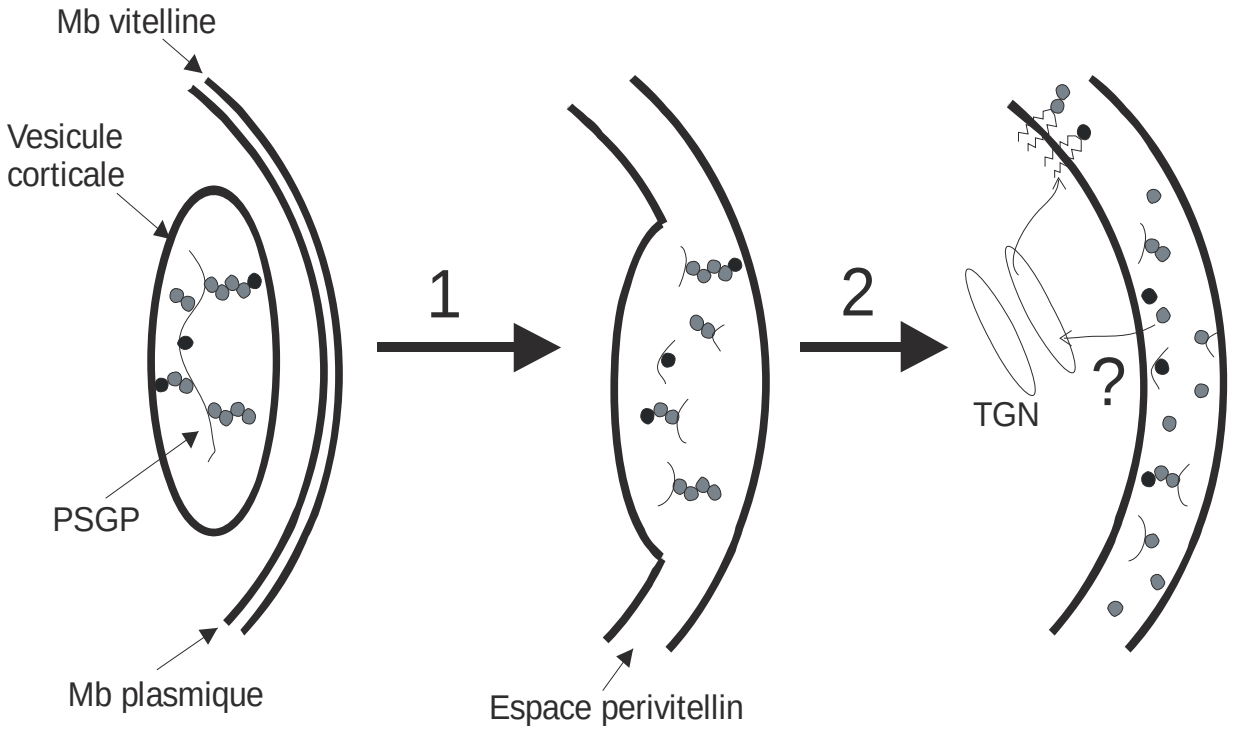


Fig. 8 : **1**, Suite à la fécondation, les glycoprotéines polysialylées (PSGP) sont excrétées des vésicules corticales et la membrane vitelline se soulève pour former l'espace perivitellin. Très rapidement (<10min), les PSGP sont protéolysées pour former des peptides polysialylés (L-PSGP) (Inoue and Inoue, 1997). **2**, dans une deuxième étape, le domaine polysialylé des L-PSGP est dégradé, probablement sous l'action d'une sialidase soluble ou membranaire, pour libérer de l'acide sialique libre (Chang *et al.*, 2008). Nous émettons l'hypothèse que celui-ci puisse être ré-endocyté par l'embryon pour servir à la constitution de néo-sialoconjugués. Rond gris, Neu5Gc ; rond noir, Neu5Ac.

Pour prouver cette hypothèse nous avons récemment mis en évidence l'existence d'une activité sialidasique soluble spécifique à l'espace périvitellin. Pour cela, nous avons séparé la membrane vitelline, l'espace périvitellin et l'embryon à partir d'oocytes fécondés (1 hpf). La mesure de l'activité sialidasique spécifique de chaque fraction montre que l'activité de la fraction périvitelline soluble compte pour 50% de l'activité sialidasique totale de l'embryon (Fig. 9).

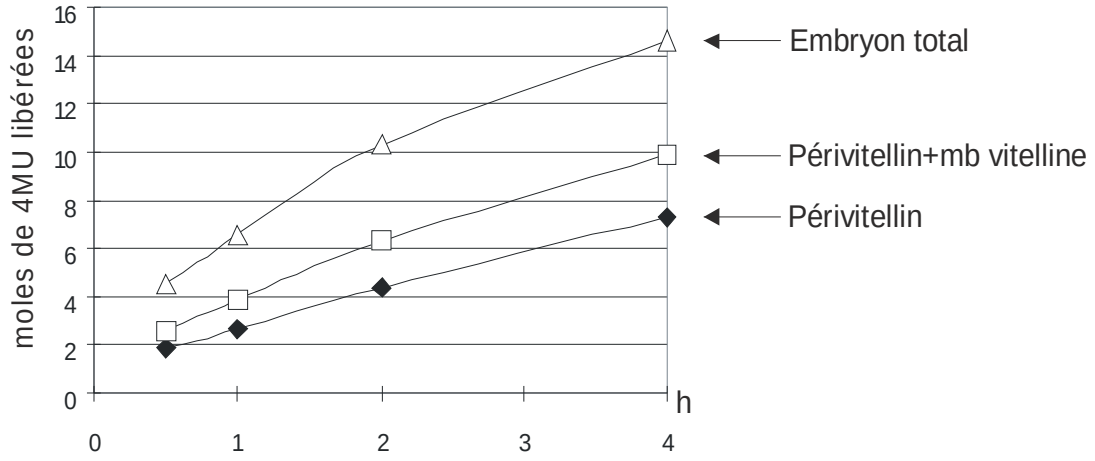


Fig. 9 : Libération d'acide sialique induite par différentes fractions d'embryons de poisson zèbre (1hpf).

L'observation d'une telle activité associée à l'espace périvitellin confirme les résultats obtenus par quantification directe des motifs oligosialylés associés aux glycoprotéines solubles au cours de l'embryogenèse, c'est-à-dire que les motifs oligosialylés sont dégradés. Néanmoins, elle ne démontre pas que les acides sialiques libérés dans l'espace périvitellin puissent être recyclés par l'embryon.

Dans le cadre de la collaboration avec le Dr. Harduin-Lepers, nous avons également abordé l'étude des polysialyltransférases potentiellement impliquées dans la biosynthèse des motifs oligosialylés que nous avions identifiés. En particulier, nous nous sommes focalisés sur un homologue de la ST8Sia VI humaine qui a été précédemment décrite comme une di-sialyltransférase, c'est-à-dire l'enzyme transférant un unique résidu d'acide sialique sur un motif monosialylé (Teintenier-Lelièvre *et al.*, 2005). L'objectif de cette étude est d'identifier l'enzyme responsable de la biosynthèse des O-glycannes oligosialylés. En effet, l'analyse structurale des N- et O-glycannes suggère très fortement l'existence d'une activité di-sialyltransférasique utilisant uniquement le Neu5Gc et non le Neu5Ac comme substrat accepteur. Jusqu'à présent une telle spécificité entre le Neu5Gc et le Neu5Ac n'a jamais été mise en évidence chez les sialyltrasférases, ni envers le substrat accepteur, ni envers le substrat donneur. De fait, la mise en évidence d'une enzyme utilisant préférentiellement le Neu5Gc constituerait une avancée importante dans la compréhension du mécanisme d'action des sialyltransférases, voire de la fonction du Neu5Gc. Ainsi, *Dre*ST8Sia VI a été cloné et exprimé dans des cellules COS-7 sous forme soluble (Fig. 10A). L'enzyme est elle-même N-glycosylée comme le prouve l'augmentation de mobilité en PAGE induite par la digestion avec la PNGAse F (Fig. 10B).

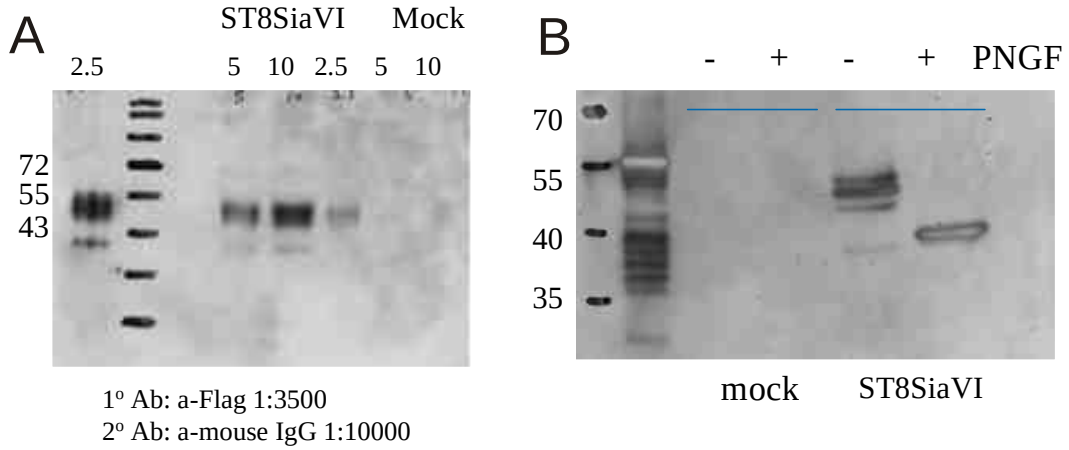


Fig. 10 : A, expression de *Dre*ST8Sia VI dans des cellules COS-7 ;**B**, déglycosylation de *Dre*ST8Sia VI par la PNGAse.

Bien qu'exprimée sous forme soluble, nous n'avons malheureusement pas encore réussi à démontrer une activité di-sialyltransférasique satisfaisante. Des expériences de mesure de transfert ont été réalisée sur de nombreux substrats exogènes ou endogènes, synthétiques ou naturels, mais les résultats obtenus ne sont pour l'instant pas reproductibles. Il est possible que nous n'ayons pas employé les substrats adéquats, ne connaissant pas la séquence de biosynthèse des O-glycannes sialylés.

Perspectives

Tout d'abord, dans le but de pérenniser nos études au sein de l'institut, nous avons participé à la mise en place d'une plateforme d'élevage de poissons zèbres grâce à un financement de l'université. Celui ci nous permettra un accès direct à notre modèle d'étude, élément indispensable à la continuation de nos travaux.

Les perspectives à ces travaux sont multiples :

1- *Analyse des profils glycolipidiques et leur évolution au cours de l'embryogenèse*. En effet, lors de notre première approche structurale, nous nous sommes focalisés sur les glycolipides sialylés car ils apparaissaient spécifiques des stades tardifs de l'embryogenèse, et donc étaient synthétisés par l'embryon. Néanmoins, des résultats préliminaires concernant les profils de glycolipides des oocytes non fécondés et des ovaires matures ont mis en évidence la présence d'un très grand nombre de glycolipides présentant des structures inédites. En particulier nous avons

observés des glycolipides polyfucosylés présentant des motifs Fuc-Fuc. L'analyse de la régulation de la biosynthèse et du catabolisme de ces composés peut se révéler très instructive dans le cadre de la régulation des interactions primordiales sucres-sucres qui régulent la compaction de l'embryon (Eggens *et al.*, 1989; Kojima *et al.*, 1994). Ces analyses seront réalisées au sein du groupe.

2- *Analyse du processus d'endocytose des acides sialiques*. Nous essaierons de déterminer l'existence et le rôle éventuel de ce processus dans l'embryogenèse du poisson zèbre. Nous nous proposons d'analyser l'incorporation par l'embryon de différents précurseurs d'acides sialiques dans ses propres glycoconjugués, soit par microinjection dans l'espace périplasmique, soit par incubation de l'embryon débarrassé de son enveloppe vitelline. Très récemment, Bertozzi et collaborateurs ont démontré qu'un analogue précurseur de l'acide sialique (le N-azidoacetylmannosamine) était métaboliquement incorporé par l'embryon vivant et pouvait servir de sonde pour observer l'apparition de glycoconjugués sialylés à sa surface (Laughlin *et al.*, 2008). Néanmoins la signification de l'incorporation d'acides sialiques dans un contexte physiologique est totalement inconnue. Ces expériences seront réalisées au laboratoire en coopération avec le groupe du Dr Khoo, à Taiwan.

3- *Détermination de la spécificité des polysialyltransférases de poisson zèbre*. En collaboration avec le Dr Harduin-Lepers, nous tenterons de déterminer les spécificités des différentes enzymes de poisson zèbre clonées et purifiées au sein de son groupe. A l'heure actuelle, deux de ces enzymes *Dre*ST8Sia VI, homologue à la ST8Sia VI humaine, et *Dre*ST8Sia VII dont le gène codant n'est homologue à aucune famille connue.

4- *Détermination de la nature des glycoprotéines sialylées*. Les glycoprotéines majeures du poisson zèbre seront identifiées par une approche glyco-protéomique. Les profils de glycosylation individuels seront déterminés par spectrométrie de masse. Une fois les protéines identifiées, leurs profils d'expressions seront analysés au cours de l'embryogenèse. Ces analyses sont en court au sein du laboratoire du Dr Khoo, à Taiwan.

2.2.2- La sialylation chez les invertébrés

En marge des études effectuées sur la sialylation chez le poisson zèbre, nous avons initié une étude prospective sur les profils de sialylation chez les invertébrés marins. En effet, une recherche systématique des gènes codant pour des sialyltransférases dans le règne animal réalisée par le Dr Anne Harduin-Lepers (Harduin-Lepers *et al.*, 2005) avait suggéré l'existence chez les urocordés et les céphalocordés de gènes codant pour des ancêtres des sialyltransférases de vertébrés. Chez les invertébrés, les gènes codant pour des sialyltransférases putatives peuvent être regroupées en seulement quatre familles ancestrales, alors que ceux des vertébrés peuvent être classés en quatre familles et vingt sous familles distinctes (Fig. 11).

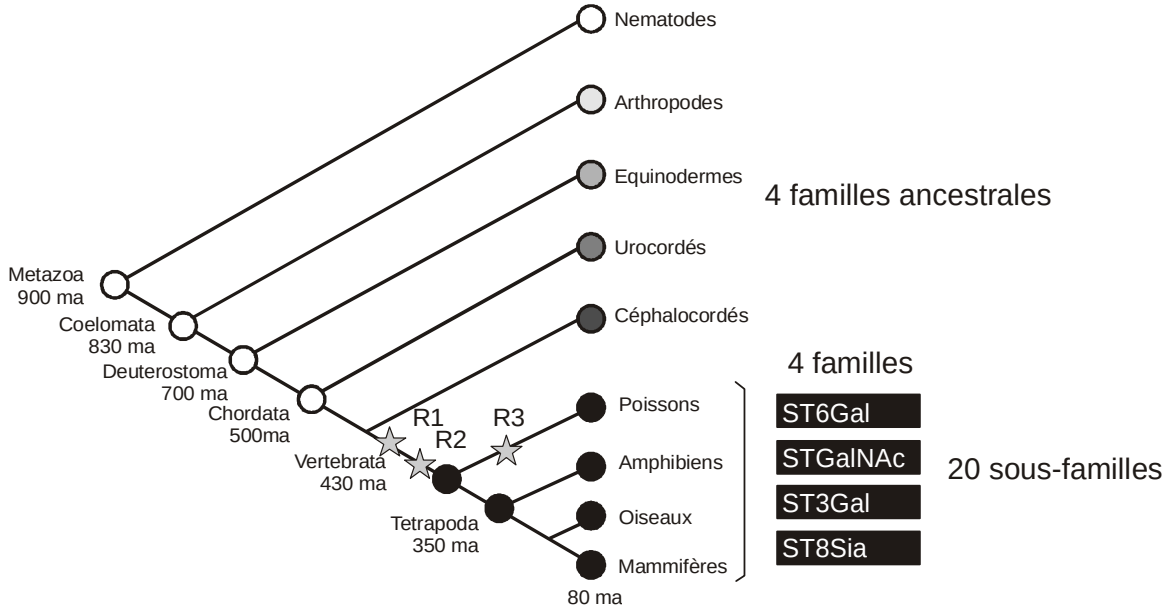


Fig. 11 : L'apparition soudaine de vingt sous familles chez les vertébrés résulterait de trois cycles de duplication des génomes complets (R1, R2 et R3) survenus lors de l'émergence de ces derniers. D'après Harduin-Lepers *et al.*, 2005 (ma : millions d'années)

Dans le cadre de l'étude de l'évolution des activités sialyltransférasiques et de la mise en place de nouveaux modèles dédiés à l'étude des sialo-glycoconjugués, nous avons initié une série d'études sur les profils de glycosylation de représentants des deux *phyla* les plus proches des vertébrés : les céphalocordés et les urocordés. Ils avaient pour objectif initial de déterminer les substrats endogènes d'éventuels sialyltransférases. Ces travaux ont été initiés en en collaboration avec les Dr. H. Sawada (Université de Nagoya) et K. Kubokawa (Université de Tokyo) lors de mon séjour dans le laboratoire du Dr Kitajima. Nous avons donc isolé les N-glycannes, les O-glycannes

et les glycolipides à partir d'ovaires de *Branchiostoma belcheri* (céphalocordé) et de gonades de *Halocynthia roretzi* (urocordés) en vue de déterminer leurs structures et leurs compositions en acides sialiques. Les résultats que nous avons obtenus sont brièvement exposés ci dessous.

Profils de sialylation de Branchiostoma belcheri

Les céphalocordés sont des animaux marins d'organisation relativement rudimentaire assez semblables en apparence aux poissons. Ils comprenent environ 25 espèces, dont la plus utilisée en tant que modèle d'embryologie est *Branchiostoma floridea*. Il est souvent considéré comme le meilleur modèle disponible pour représenter le dernier invertébré à servir d'ancêtre commun aux vertébrés. L'espèce que nous avons étudiée est *B. belcheri,* proche cousin asiatique de *B. floridea*. Par analogie avec les poissons chez lesquels les ovaires sont très riches en sialo-conjugués, nous avons commencé nos études sur ce tissu. Les ovaires entiers de *B. belcheri* ont été grossièrement fractionnés en protéines solubles, protéines insolubles, glycannes libres et glycolipides. L'analyse de leur composition en acides sialiques par HPLC après couplage à un fluorophore a révélée la présence de Neu5Ac, de Neu5Gc, de Kdn (acide 3-deoxy-D-glycero-D-galacto-nonulosonique) et d'un composé supplémentaire inconnu (dénommé ici X) dont le temps de rétention ne correspondait à aucun standard d'acide sialique (Fig. 12A). Neu5Ac et X étaient présents en quantités très importantes dans les deux fractions protéiques (environ 1µmole par g. de tissu frais). Les quantités observées sont du même ordre que celles observées dans les ovaires de poissons. Après libération et purification des O-glycannes et des N-glycannes des fractions protéiques, il est apparu que la grande majorité des acides sialiques était associée aux O-glycannes (Fig.12C.).

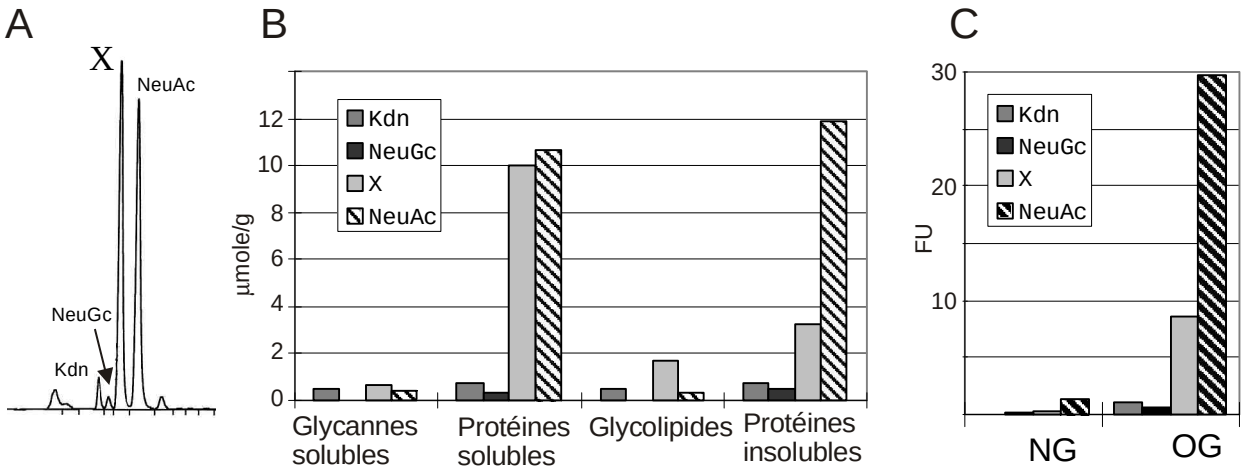


Fig. 12 : Dosage des acides sialiques d'ovaire de *B. belcheri* par HPLC. **A**, exemple de profil HPLC ; **B**, dosage des les différentes fractions ; **C**, comparaison des quantités d'acides sialiques dans les O- et les N-glycannes totaux isolés de protéines insolubles

Le fait que le composé X soit identifié dans des quantités comparables dans les glycannes purifiés et dans les fractions totales suggérait qu'il ne s'agissait pas d'une contamination, mais bien d'un composé spécifiquement associé aux glycannes. Ce composé a ensuite été partiellement identifié par l'analyse de son dérivé fluorescent en LC-CID MS/MS. Le spectre de fragmentation du pic précedemment identifié comme du Neu5Ac sur la base de son temps de rétention en HPLC a tout d'abord confirmé la nature de ce dérivé (Fig. 13A). Le dérivé du composé X présentait quant à lui une masse $[M+Na]^+$ à m/z 421. La valeur de signal, ainsi que celle du fragment correspondant à la coupure ente les carbones C-3 et C-4 à m/z 217, ne correspondent *a priori* à aucun acide sialique classique, mais sont compatibles. Ces valeurs sont par contre compatibles avec un Kdn mono-méthylé (Fig. 13B). Bien que la méthylation soit une modification commune des acides sialiques (Neu5Ac8Me, Neu5,9$Ac_2$8Me, Neu5Gc8Me, Neu9Ac5Gc8Me, Neu2,7an5Gc8Me…), elle n'avait à notre connaissance encore jamais été observée associée au Kdn. La position de groupement méthyl doit néanmoins encore être déterminée par des méthodes complémentaires.

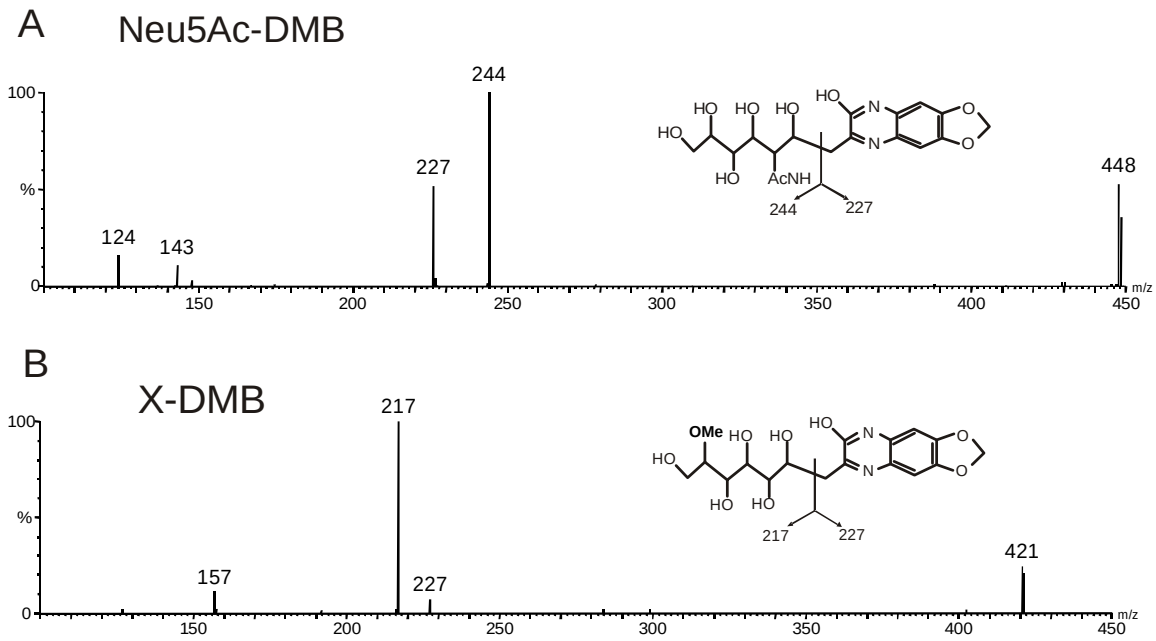


Fig. 13 : Analyse en LC-CID-MS/MS des acides sialiques majeurs dérivés au DMB, des O-glycannes insolubles isolés de *B. belcheri*. Spectres de fragmentation **A**, de Neu5Ac-DMB ; **B**, de X-DMB attribué sur cette base comme du KdnMe-DMB.

Les O-glycannes sialylés majeurs ont ensuite été perméthylés puis analysés par MALDI-MS et MALDI-MS/MS. Le spectre MS total exhibe deux pics majeurs à m/z 1174,7 et 1256,7 pouvant correspondre aux adduits $[M+Na]^+$ de glycannes de compositions respectives $HexNAc_1Hex_1Kdn_2$ et $HexNAc_1Hex_1NeuAc_2$ (données non montrées). Il est à noter qu'une fois perméthylé, il est

impossible de distinguer le Kdn du KdnMe. Néanmoins, considérant le fait que les quantités de KdnMe sont au moins dix fois supérieures à celles du Kdn, il est fort probable que le signal observé à m/z 1174,7 présente la composition suivante : $HexNAc_1Hex_1KdnMe_2$. Les profils de fragmentation suggèrent que ces deux composés présentent des structures équivalentes qui ne différent que par la nature de leurs acides sialiques, Neu5Ac ou KdnMe (Fig. 14). Pour le signal à m/z 1256.7, la perte consécutive de deux Neu5Ac terminaux (M-375 à m/z 881.5 et M-2x375 à m/z 506.3) démontre que les deux acides sialiques sont présents en positions terminales non réductrices et ne forment donc pas de motifs di-sialylés. Enfin, le couple de fragments C/Z à m/z 620/659 démontre qu'un des résidus d'acide sialique substitue le résidu d'hexose. Ainsi, le profil de fragmentation établit la séquence de ce glycanne comme étant Neu5Ac-Hex-[Neu5Ac]-HexNAac-ol. Bien que cette analyse ne permette pas de déterminer formellement la position de substitution des branches sur le résidu d'HexNAc-ol, le fait que le résidu d'hexose soit libéré *via* une coupure de type Z suggère très fortement qu'il soit lié sur le C-6 et non sur le C-3 (le détail des fragmentations spécifiques des oligosaccharides méthylés sera donné dans la dernière partie du présent manuscrit). Basé sur les voies de biosynthèse canoniques des O-glycannes identifiées chez les vertébrés, ce profil de fragmentation est cohérent avec la structure Neu5Ac(α2,3/6)-Gal(β-1,6)[Neu5Ac(α2-3)]-GalNAc-ol. Néanmoins, une telle structure devra être confirmée par d'autres méthodologies. Le profil de fragmentation du composé à m/z 1174.7 est très similaire ; les valeurs des fragments sialylés ne différent que de 41 mu, ce qui correspond à la différence de masse calculée entre le Neu5Ac et le KdnMe permétylés. Il est également notable que la liaison du KdnMe perméthylé semble plus résistante à la fragmentation par CID-MS que ne l'est celle du Neu5Ac. Ainsi, ces analyses suggèrent que ce composé présente la structure suivante : KdnMe(α2,3/6)-Gal(β-1,6)[KdnMe(α2-3)]-GalNac-ol.

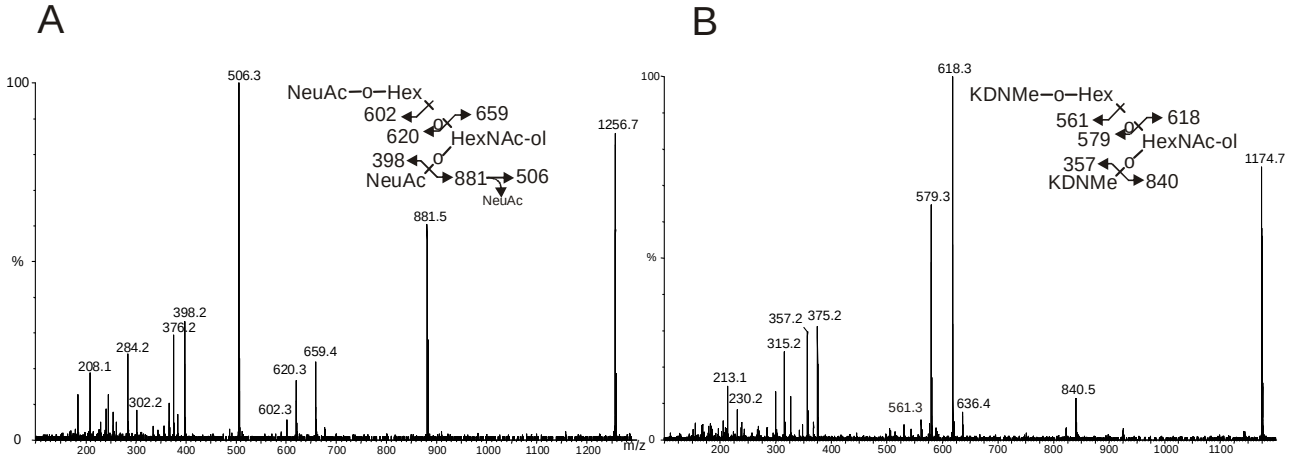


Fig. 14 Profils de fragmentation par MALDI-Q-TOF-MS/MS des deux O-glycannes majeurs isolés des ovaires de *B. belcheri*.

En conclusion, les analyses préliminaires effectuées sur les glycannes de *B. belcheri* ont clairement démontré l'existence de plusieurs activités sialyltransférasiques dans cet organisme. L'analyse plus détaillée du génome d'une espèce apparentée de céphalocordés, *B. floridae*, a révélé l'existence de gènes présentant des homologies avec les familles des ST8Sia et ST3Gal dont les produits pourraient être à l'origine de la biosynthèse des composés que nous avons identifiés (Harduin-Lepers, résultats non publiés). Ainsi, nous pensons que l'analyse de la spécificité fine de ces enzymes ancestrales est possible chez les céphalocordés sur la base de la connaissance de leurs substrats endogènes et qu'elle peut apporter des informations précieuses sur l'évolution des sialyltransférases.

Profils de sialylation de Halocynthia roretzi

Halocynthia roretzi est un représentant des urocordés, appelés également tuniciers. Les tuniciers ont typiquement des corps en forme d'outre recouverts d'une tunique cellulosique. La plupart des espèces sont des organismes sessiles qui se nourrissent en filtrant de grandes quantités d'eau. Les tuniciers sont depuis peu considérés comme les parents les plus proches des vertébrés, et de fait remplaceraient les céphalocordés qui ont longtemps occupés cette position. A priori, les tuniciers adultes ne présentent que très peu de ressemblances avec les vertébrés, mais leurs larves sont libres et possèdent une notocorde, un pharynx et une queue post-anale. Ce type de queue est considéré comme le précurseur de queue cartilagineuse ou osseuse des vertébrés. A ce propos, on note cependant des différences marquées entre les trois classes de tuniciers : la queue larvaire est inexistante chez les Thaliacea, perdue lors de la métamorphose des Ascidiacea -dont fait partie *H. roretzi*-, mais persiste chez les Appendicularia. En raison de leur position stratégique nouvelle en tant qu'ancêtres des vertébrés, ces petits animaux ont dernièrement provoqué un intérêt très large de la part des biologistes du développement.

De manière identique aux analyses effectuées sur *B. belcheri,* nous avons quantifié les acides sialiques dans les fractions protéiques, lipidique et glycannique libre de gonades de *H. roretzi*. Au contraire des céphalocordés, il est à noter que la grande majorité des ascidies, y compris *H. roretzi*, sont hermaphrodites, chaque individu ayant des gonades des deux sexes. En comparaison avec les quantités très importantes détectées chez *B. belcheri*, seules des quantités négligeables d'acides sialiques -de l'ordre de une à deux nmole/g de tissu- constituées de Neu5Ac et Kdn, ont été détectées dans les fractions protéiques et lipidiques. La fraction glycannique libre contenait quant à elle des quantités plus importantes de ces mêmes acides sialiques estimées à 10 nmole/g de tissu. Ainsi, les gonades de *H. roretzi* contiendraient 100 à 1000 fois moins d'acide sialique que les

ovaires de poisson et de céphalocordés, de fait, il est difficile de déterminer si l'acide sialique détecté est effectivement associé à des glycoconjugués endogènes ou à une contamination bactériennes. De fait, nous avons dans un premier temps réalisé une série d'analyses préliminaires concernant la structure des N- et O-glycannes pour lever cette ambiguïté. L'analyse par MALDI-MS des N- et O-glycannes totaux purifiés de gonades a révélé la présence de familles extrêmement hétérogènes de molécules presque uniquement composées de résidus de HexNAc et de deHex L'analyse de composition par GC et RMN a démontré ultérieurement qu'elles étaient composées de résidus de Fuc, GlcNAc et GalNAc pour former des motifs LacdiNAc (GalNAcβ1-4GlcNAc) fucosylés. (Fig. 15 A et B).

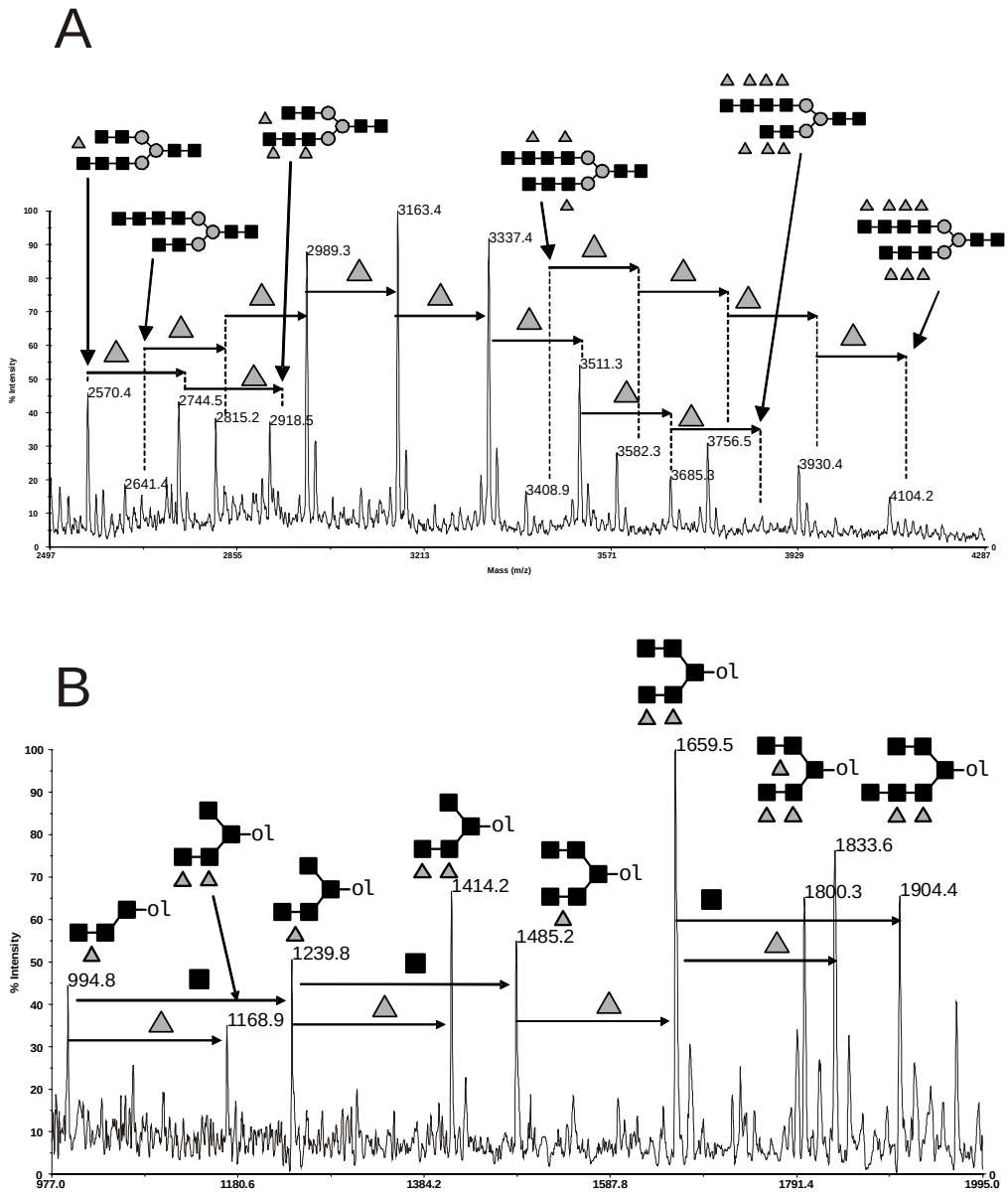


Fig. 15 : Profils en MALDI-MS des **A**, N-glycannes et **B**, O-glycannes perméthylés isolés des gonades de *H. roretzi*. Carré noir, HexNAc ; triangle, deHex ; rond, Hex (probablement gris, Man et blanc, Gal). Les séquences représentées sont proposées sur la base des compositions calculées en monosaccharides, mais n'ont pas été déterminées à ce stade de l'analyse.

Par contre, aucune molécule substituée par de l'acide sialique n'a été mis en évidence lors de ces analyses. De même, des analyses effectuées sur les glycolipides et les glycannes libres n'ont révélé aucune trace d'acide sialique. De fait, l'utilisation de *H. roretzi* comme modèle d'étude de la sialylation paraissait d'ores et déjà largement compromise. Néanmoins, la présence de composés hyperfucosylés s'est avérée être une piste très prometteuse pour identifier le ligand potentiel d'un récepteur de surface du spermatozoïde. En effet, il a récemment été démontré que la fixation du spermatozoïde sur l'œuf de *H. roretzi* était sous le contrôle d'une protéine de surface du spermatozoïde dont le gène codant est très homologue à celui d'une α-L-fucosidase (Matsumoto *et al.*, 2002). L'activité de cette α-L -fucosidase présente un pH optimal extrêmement bas (pH 4.0) dont la valeur n'est pas compatible avec les conditions physiologiques (eau de mer pH 8.2). De plus, l'interaction entre les spermatocytes et la membrane vitelline a été bloquée par les substrats synthétiques de la fucosidase (p-nitrophényle, 4-methylumbelliferylet et de α ou β fucosides) qui ont vraisemblablement agit comme inhibiteurs concurrentiels de la fucosidase pour ses substrats et/ou ses ligands physiologiques. Ainsi, cette fucosidase ne présenterait à pH physiologique qu'une activité lectinique, responsable de l'attachement su spermatozoïde à la membrane vitelline de l'œuf. De fait, notre collaborateur, le Dr Sawada (Université de Nagoya), s'est montré très intéressé par la caractérisation des glycannes fucosylés dont nous avions démontré l'existence, ces derniers étant probablement les ligands du récepteur du spermatozoïde.

Ayant pour objectif d'identifier les ligands glycanniques de la fucosidase à la surface de l'œuf de *H. roretzi*, nous avons donc dans une deuxième étape procédé à l'analyse des glycannes isolés d'enveloppe vitellines purifiées. Ce matériel biologique n'étant disponible qu'en quantités réduites nous avons tout d'abord analysé de manière exhaustive la structure des glycannes isolés des gonades, disponibles en grande quantité. Ensuite, nous avons comparé les profils de glycosylation des oeufs entiers et des membranes vitellines avec ceux des gonades. Pour ce faire, dans le cas des gonades et des oeufs, les N- et les O-glycannes ont été séparément isolés par hydrazynolyse et β-élimination, respectivement à partir des mêmes lots d'échantillons séparés en deux. Par contre, pour les enveloppes vitellines, N- et O-glycannes ont été successivement libérés par action de la PNGAse puis β-élimination. Les profils O-glycanniques (Table 2A) et N-glycanniques (Table 2B) de chaque échantillon ont ensuite été comparés par l'analyse MALDI-MS des glycannes natifs et perméthylés.

A

Natif (m/z)	Permeth (m/z)	HexNAc	Fuc	Hex	Gonades	Oeufs	Mb Vitelline
	1476,5	2	4	1			
652,1	820,6	3					
798,2	994,7	3	1				
944,3	1168,8	3	2				
855,2	1065,8	4					
1001,3	1239,9	4	1				
1147,3	1414,1	4	2				
1294,3		4	3				
1204,3	1485,1	5	1				
1350,3	1659,2	5	2				
1496,9	1833,7	5	3				
1643,9	2007,7	5	4				
1553,9	1904,7	6	2				
1700,0	2078,7	6	3				
1846,2	2251,7	6	4				
1992,2	2426,8	6	5				
2154,0		6	5	1			
2357,4		7	5	1			
2560,0		8	5	1			

B

Permeth (m/z)	HexNAc	Fuc	Hex	Corps	Gonades	Oeufs	Mb Vitellin
2153,1	4			M3GN2			
2327,2	4	1		M3GN2			
2501,2	4	2		M3GN2			
2675,40	4	3		M3GN2			
2851,5	4	4		M3GN2			
2398,1	5			M3GN2			
2572,3	5	1		M3GN2			
2746,4	5	2		M3GN2			
2920,5	5	3		M3GN2			
3094,5	5	4		M3GN2			
3268,6	5	5		M3GN2			
2643,30							
2817,4	6	1		M3GN2			
2991,5	6	2		M3GN2			
3165,6	6	3		M3GN2			
3339,7	6	4		M3GN2			
3513,7	6	5		M3GN2			
3687,7	6	6		M3GN2			
3861,7	6	7		M3GN2			
3410,6	7	3		M3GN2			
3584,6	7	4		M3GN2			
3758,6	7	5		M3GN2			
3932,7	7	6		M3GN2			
4106,7	7	7		M3GN2			
1784			3	M3GN2			
1988			4	M3GN2			
2192			5	M3GN2			
2396			6	M3GN2			

Table 2. Comparaison des profils de A, O- et B, N-glycosylation des gonades, oeufs et membranes vitellines de *H. roretzi*.

Ces analyses ont mis en évidence des différences majeures dans les profils de glycosylation des gonades, œufs et membranes vitellines. Les gonades et les œufs contiennent de nombreux O-glycannes communs de m/z compris entre 798 et 1496 sous formes natives (table 2A). Ils ont été attribués à des molécules contenant uniquement entre 3 et 5 résidus de HexNAc et 1 à 3 résidus de deHex. Par contre, l'ensemble des O -glycannes de masses supérieures comprises entre m/z 1643 et 2560 n'a été identifié que dans les gonades. Cette différence de composition a également été vérifiée par l'analyse des O-glycannes par TLC (données non montrées). Enfin, nous n'avons pu observer de signal caractéristique de O-glycanne dans les membranes vitellines, ce qui suggère fortement l'absence de tout O -glycanne.

La comparaison des N -glycannes issus des différents tissus montre également une grande spécificité de localisation (table 2B). Alors que de très nombreuses molécules sont observées dans les gonades, la diversité observée dans les œufs et les membranes vitellines est beaucoup plus restreinte. En effet, seuls les N -glycannes oligomannosylés sont observés dans les œufs, tandis que seule une sous famille de N-glycannes complexes, en moyenne de masses plus faibles que celles observées dans les gonades, est présente dans les membranes vitellines. Il est également très surprenant de constater l'absence totale de N-glycannes oligomannosylés dans la membrane vitelline (Fig. 16). Bien que les résultats obtenus sur les membranes vitellines soient sans ambiguïté

possible, nous suspectons fortement que l'absence de N-glycannes dans les oeufs soit le résultat d'un problème expérimental lors de leur libération par voie chimique. En effet, des analyses fines des résultats de spectrométrie de masse ont révélé la présence d'un grand nombre de signaux artefactuels probablement due à une hydratation trop importante de l'hydrazine. Cette expérience devra donc être réitérée.

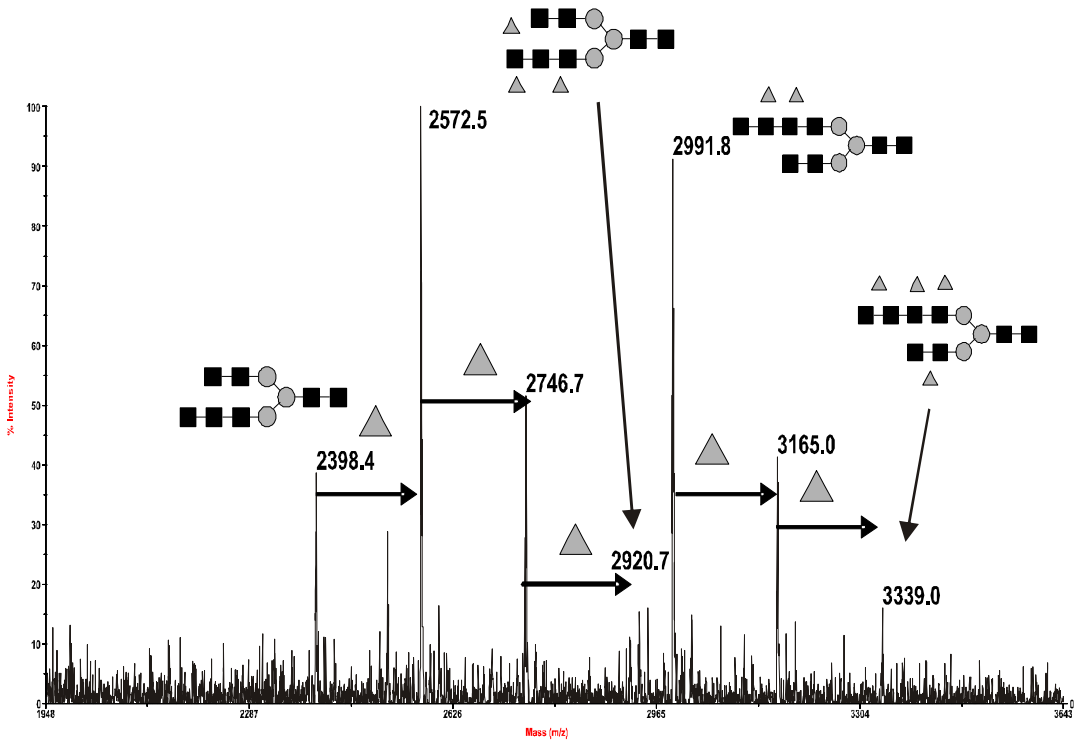


Fig. 16 : Profil de N-glycosylation (oligosaccharides perméthylés) de l'enveloppe vitelline de *H. roretzi*. Les séquences représentées sont proposées sur la base des compositions calculées en monosaccharides, mais n'ont pas été déterminées à ce stade de l'analyse.

Les études en cours ont pour objectif de déterminer la structure fine de l'ensemble des O- et N-glycannes identifiés dans les différents tissus de *H. roretzi*. Des premières expériences de séquençage des O-glycannes par MALDI-MS-MS ont révélé que chaque signal observé par MALDI-MS était en fait un mélange complexe de structures isobariques qui diffèrent par leurs profils de branchement et la position de leurs résidus de fucose. A titre d'exemple, quelques séquences de ces structures isobariques sont montrées Table 3 et 4.

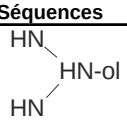
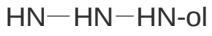
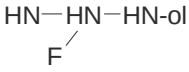
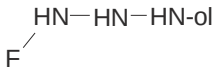
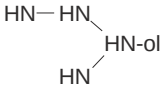
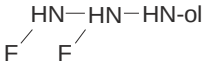
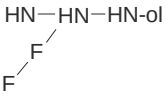
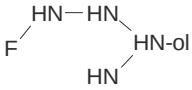
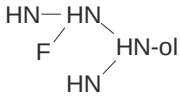
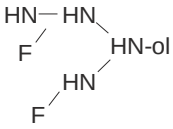
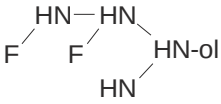
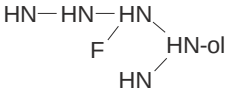
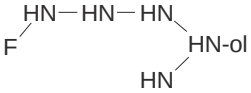
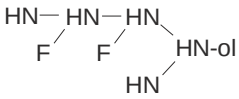
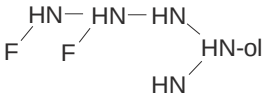
Natif	Permethylé	composition-ol	Séquences	
652,1	820,6	HN3		
798,2	994,7	HN3F		
855,2	1065,8	HN4		
944,3	1168,8	HN3F2		
1001,3	1239,9	HN4F		
1147,3	1414,1	HN4F2		
1204,3	1485,1	HN5F		
1350,3	1659,2	HN5F2		

Table 3 : Séquences des O-glycannes majeurs communs aux gonades et aux oeufs de *H. roretzi* déduites de l'analyse des O-glycannes en mélange par MALDI-MS/MS

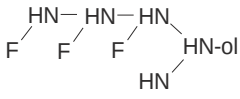
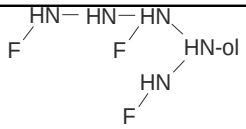
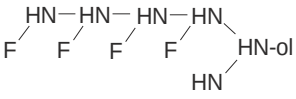
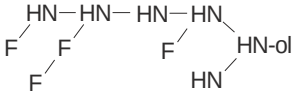
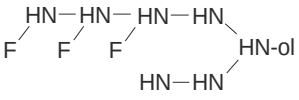
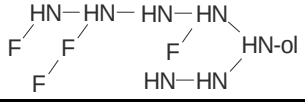
Natif+Na	Perméthylés	Composition-ol		
1496,9	1833,7	HN5F3		
1643,9	2007,7	HN5F4		
1700	2078,7	HN6F3		
1846,2	2251,7*	HN6F4		

Table 4 : Séquences des O-glycannes majeurs spécifiques aux gonades déduites de l'analyse des O-glycannes en mélange par MALDI-MS/MS

Ainsi, dans une dernière étape, les structures isobariques des O-glycannes isolés de gonades ont été partiellement résolues par purification HPLC multidimensionnelle sur colonnes en phase normale et inverse. L'analyse de ces composés purifiés est en cours, mais a d'ores et déjà permis de déterminer la structure exacte de plusieurs O-glycannes purifiés grâce à l'utilisation d'une combinaison d'analyses RMN, de chromatographie en phase gaz et de dégradation chimique. A titre d'exemple un des deux isobares majeurs du composé Fuc_1HN_3-ol (m/z 798) a été isolé et identifié comme étant le GalNAc(β1-4)[Fuc(α1-3)]GlcNAc(β1-3)GalNAc-ol (Fig. 17), tandis que son équivalent di-fucosylé (m/z 944) a été identifié comme étant le Fuc(α1-3)GalNAc(β1-4)[Fuc(α1-3)]GlcNAc(β1-3)GalNAc-ol.

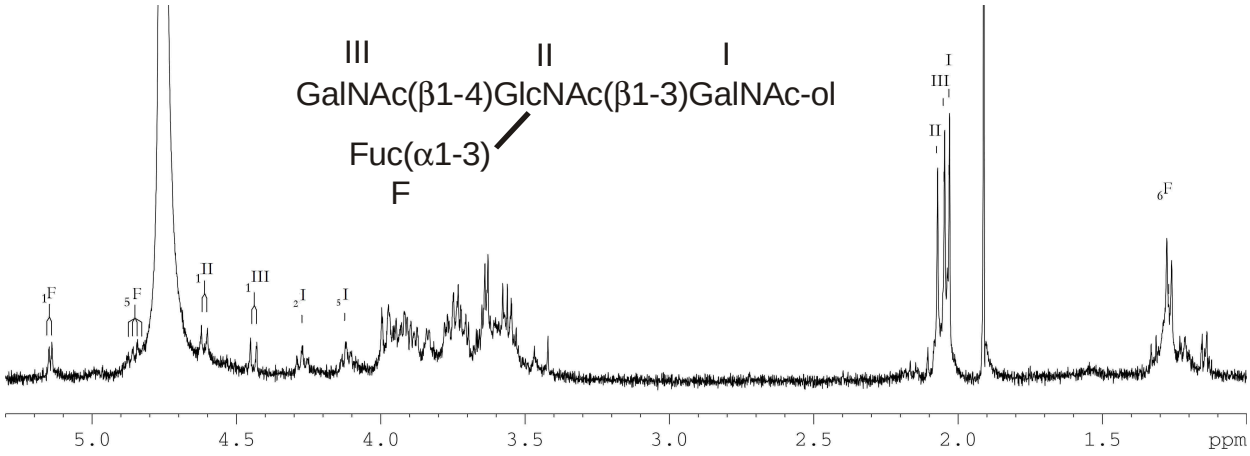


Fig. 17 : Spectre ^{1}H RMN d'un O-glycanne purifié des gonades de *H. roretzi* par HPLC bi-dimensionelle.

Ainsi l'analyse exhaustive de deux O-glycannes purifiés suggère fortement que les enchaînements d'HexNAc observés dans tous les glycannes de *H. roretzi* soient constitués de motifs LacdiNAc plus ou moins fucosylés. L'ensemble des données récoltées a démontré que les structures de tous les N- et O-glycannes était basées sur le même principe. Les gonades présentent la plus vaste diversité structurale, tandis que les oeufs et les membranes vitellines ne contiennent que des sous familles des glycannes observées dans les gonades. En particulier, les glycoprotéines des membranes vitellines sont spécifiquement substituées par une famille relativement homogène de N-glycannes polyfucosylés, mais par aucun N-glycanne oligomannosylé ni O-glycanne. De fait, ces molécules apparaissent comme les ligands probables du récepteur identifié à la surface du spermatozoïde. Dans une prochaine étape, les molécules que nous avons purifiées seront utilisées dans des tests d'inhibition de la fixation du spermatozoïde sur l'œuf de *H. roretzi* pour prouver

l'implication de ces N-glycannes dans l'interaction spermatozoïde-oeuf. Ces tests seront effectués par le Dr Sawada.

En conclusion, les analyses effectuées sur deux espèces de céphalocordés et d'urocordés ont démontré que l'espèce *B. belcheri* pouvait se révéler comme un modèle intéressant pour l'étude de l'évolution des activités sialyltransférasiques. De fait, nous envisageons de pérenniser ces travaux en collaboration avec le Dr Harduin-Lepers dans le but d'étudier la spécificité des sialyltransférases de ce modèle. Par contre, malgré l'identification de gènes codant pour des homologues de sialyltransférases chez les urocordés, en particulier un membre apparenté à la famille des ST3Gal, la présence d'activité sialyltransférasique n'a pas été confirmée par nos travaux sur le modèle *H. roretzi*. Plusieurs hypothèses peuvent être avancées pour résoudre cette apparente contradiction. **1**, les études génétiques sont basées sur l'étude du génome de *Ciona intestinalis*, une espèce différente bien que proche, de *H. roretzi*. Nous envisageons d'initier une étude comparative sur ce modèle animal. **2**, sur la base des données disponibles sur les poissons, et en accord avec celles collectées sur *B. belcheri*, nous avons restreint nos études aux tissus reproductifs et dérivés. De fait, il est possible que la sialylation soit tissu spécifique et qu'il faille la rechercher dans d'autres tissus tels que le pharynx ou l'estomac. **3**, les activités endogènes des sialyltransférases peuvent également être limitées à un type de glycanne très particulier, présent en quantité limité, ce qui rend leur détection difficile. Un tel cas a été observé chez *Drosophila melanogaster* dont le génome code pour une α2-6-sialyltransferase fonctionnelle, mais dont le substrat endogène n'a pas été identifié par une approche classique d'analyse des profils de glycosylation (Koles *et al.*, 2004 ; North *et al.*, 2006). Seule l'analyse détaillée des glycannes après purification multidimensionnelle a permis finalement de mettre en évidence la présence de N-glycannes sialylés mineurs, dont les structures ne correspondent d'ailleurs pas avec les prévisions basées sur l'analyse de l'activité *in vivo* de DSiaT. **4**, la présence d'homologues de gènes codant pour des sialyltransférases n'est pas suffisante à la biosynthèse de glycoconjugués sialylés en l'absence de la machinerie complète de sialylation. C'est le cas *d'Arabidopsis thaliana* dont le génome code plusieurs homologues de sialyltransférases de mammifère mais qui est incapable de synthétiser le moindre sialo-conjugué (Séveno *et al.*, 2004).

D- Développements méthodologiques

L'analyse structurale des glycoconjugués présente une forte composante méthodologique et de fait est très dépendante de l'instrumentation. Néanmoins, malgré les améliorations constantes des techniques d'analyse, la glycobiologie (ou glycomique en l'occurrence), au contraire d'autres disciplines telles que la génomique ou la protéomique, n'a toujours pas vu émerger de méthodologie standard permettant la détermination structurale automatisée des sucres. En cela, elle rejoint l'analyse de toutes les autres modifications post-traductionnelles, mais les limitations méthodologiques sont d'autant plus fragrantes dans le cas des sucres, à cause de leur hétérogénéité intrinsèque. De fait, l'analyse des sucres est restée une discipline très traditionnelle dans sa conception malgré le matériel sophistiqué qui est employé pour la mener à bien. Les améliorations considérables survenues ces dernières années ont ainsi grandement étendu nos capacités d'analyse en terme de complexité de glycannes ou de sensibilité, mais n'ont pas radicalement transformé la discipline. En particulier, les principales améliorations technologiques ont été accomplies dans le secteur de la spectrométrie de masse. Au cours de la dernière décennie, sont apparus un grand nombre d'instruments et de configurations qui ont décuplé les capacités d'analyse par spectrométrie de masse. Par comparaison, les progrès accomplis en instrumentation RMN semblent beaucoup plus restreints et limités à l'augmentation del'intensité des champs magnétiques accessibles. Néanmoins, l'amélioration de la stabilité des aimants et l'avènement de nouveaux supraconducteurs, couplés à l'utilisation de micro tubes ou de cryo-sondes ont permis ces dernières années un gain de sensibilité très appréciable.

La principale limitation technologique à laquelle nous sommes actuellement confrontée ne concerne pas le manque de sensibilité, comme souvent avancé par les constructeurs d'instruments, mais la complexité des mélanges. En effet, l'amélioration de la sensiblité et des capacités de discrimination des techniques analytiques n'a rendu que plus flagrante la reconnaissance de l'hétérogénéité des structures glycanniques et donc la difficulté à les séparer pour les analyser ou les utiliser dans des tests biologiques. En effet, des molécules considérées il y a une dizaine d'année comme pures apparaissent maintenant comme des mélanges complexes. Cet état de fait s'est confirmé à de nombreuses reprises au sein du groupe lorsque nous ré-évaluons la structure des molécules précedemment purifiées par le Dr Strecker et considérées comme pures. Ainsi, il est clair que les capacités de séparation des glycannes n'ont pas suivi celles liées à leur analyse. De fait, deux voies disctinctes peuvent être suivies pour contrebalancer ce problème ; soit améliorer notre capacité à résoudre l'hétérogénéité glycannique, soit être capable d'analyser la structure des glycannes en mélange. Chaque stratégie présente ses propres limites et la résolution de ce problème ne peut venir que de l'adoption d'une voie médianne prenant en compte les améliorations sur les

deux fronts. En effet, la complexité presque sans limite de certaines préparations rend illusoire la séparation individuelle de chaque structure isobarique. C'est par exemple le cas des O-glycannes associés aux mucines trachéobronchiales humaines dont nous estimons le nombre à au moins 2000, basée sur nos analyses en spectrométrie de masse. Par contre, elle rend également très difficile la résolution de leur structure en mélange.

Notre équipe de recherche ne s'implique que très peu dans le développement de nouvelles méthodologies, préférant utiliser des méthodes d'analyse robustes déjà éprouvées pour arriver à nos fins. Néanmoins, nous essayons d'améliorer des protocoles disponibles pour répondre à des demandes particulières de nos propres thématiques de recherche. Cela a été le cas récemment dans le but d'améliorer la purification d'oligosaccharides sulfatés présents au sein de mélanges complexes dans le cadre de la préparation et de l'analyse de O-glycannes sulfatés d'amphibiens et d'ovomucine de poule. Cette méthodologie sera présentée dans le présent chapitre. En parallèle à l'amélioration de protocoles de purification, nous présenterons dans la suite de ce chapitre nos résultats concernant l'analyse de données de spectrométrie de masse en vue d'optimiser l'interprétation structurale de glycannes en mélange. Ce manuscrit, écrit essentiellement pour notre usage, reprend des données classiques de spectrométrie de masse et propose leur utilisation possible dans le cadre de l'analyse de molécules en mélange.

Glycoconj J
DOI 10.1007/s10719-008-9139-3

A single step method for purification of sulfated oligosaccharides

Estelle Garenaux · Shin-Yi Yu · Doina Florea · Gérard Strecker · Kay-Hooi Khoo · Yann Guérardel

Received: 31 January 2008 / Revised: 22 April 2008 / Accepted: 23 April 2008

Abstract Purifying and analysing sulfated oligosaccharides by mass spectrometry often constitutes a challenge. We present here a single step method to isolate sulfated compounds from a complex mixture of neutral and acidic oligosaccharide–alditols. The strategy relies on the exclusion of sulfated molecules from strong cation exchange resin. By testing a wide range of reduced mucin type O-glycans isolated from different biological sources, we demonstrate that this method permits, without prior chemical modification, the specific purification of sulfate-containing oligosaccharides present in any quantity from very complex mixtures of molecules.

Keywords Sulfated oligosaccharides · Purification · Mass spectrometry · Mucins

Abbreviations

dHex	deoxyHexose
DHB	2,5-dihydroxybenzoic acid
ESI	ElectroSpray Ionization
HPLC	High Performance Liquid Chromatography
Kdn	2-keto-3-deoxynononic acid
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MS	Mass Spectrometry
Man	Mannose
S	Sulfate group
TLC	Thin Layer Chromatography

Electronic supplementary material The online version of this article (doi:10.1007/s10719-008-9139-3) contains supplementary material, which is available to authorized users.

E. Garenaux · D. Florea · G. Strecker · Y. Guérardel (✉)
Unité de Glycobiologie Structurale et Fonctionnelle,
CNRS UMR 8576, USTL,
59655 Villeneuve d'Ascq, France
e-mail: yann.guerardel@univ-lille1.fr

S.-Y. Yu · K.-H. Khoo
Institute of Biological Chemistry, Academia Sinica,
Nankang,
Taipei 11529, Taiwan

S.-Y. Yu · K.-H. Khoo
Institute of Biochemical Sciences, National Taiwan University,
Taipei 106, Taiwan

Introduction

Sulfation is a common modification of protein glycans. It is present at a very high prevalence on high molecular weight glycosaminoglycans bearing proteoglycans, whose physico-chemical and biological functions are directly dependent on their number and positions. It is also incorporated in numerous smaller protein glycans, where it may play a crucial role in biological functions of glycoproteins. In humans, it mainly substitutes lactosamine-containing glycans either as GlcNAc-6-sulfate or Gal-3-sulfate [1 3]. Sulfate groups are observed on N-glycans substituted proteins such as thyroglobulin [4], carbonic anhydrase VI [5] and Tamm-Horsfall [6], as well as on O-glycans of ovarian cystadenoma glycoproteins [7]. Unusual sulfations may also occur on very specific epitopes such as HNK1 (SO_4–3-GlcAβ1-4Galβ1-4G1cNAc-R) on myelin-associated glycoproteins [8]. Several glycoconjugates have been shown to play important biological roles mediated by specific recognition of their sulfate moiety by receptors. This is the case with luteinizing hormone receptor on endothelial cells that is highly specific for the SO_4–4-GalNAcβ1-4G1cNAc motifs

present on its N-glycans and is responsible for the rapid clearance of the luteinizing hormone from circulation [9]. Similarly, the oligosaccharide ligand for L-selectin was shown to be the 6′sulfo-sialyl Lex on a mucin-like molecule, termed GIyCAM-1 (for glycosylated cell adhesion molecule-1), during the homing of leukocytes to sites of inflammation [10].

Other than proteoglycans, mucin type glycoproteins are the main carriers of sulfated glycans in humans. Indeed, sulfate groups have been observed in most excreted mucin-containing mucus, including salivary [11], tracheo bronchial [12] and intestinal mucus [13]. Mucins are the major protein component of the mucus protecting the epithelia. They are highly glycosylated macromolecules characterized by the presence of a dense and highly diverse O-glycosylation, linked to the protein core through serines and threonines, that may represent up to 80% of the total weight of mucins. O-glycans contribute to both the physical and chemical protections of sensitive epithelia, such as bronchial epithelia, by providing the rheologic properties of mucus. Furthermore, the extraordinary structural diversity of O-glycosylation generates a wide range of carbohydrate based epitopes that are believed to supply attachment sites for microorganisms, which can be trapped and expelled during the continuous renewal of the mucus layer. By providing competing receptors for cell-surface glycoconjugates, mucins may trap bacteria and make them less successful in their attempts to colonize the epithelium. Thus, the array of oligosaccharides expressed on the mucins of an individual may play a key role in governing the susceptibility to infection [14].

The sulfation status of mucin type O-glycosylation, among other modifications of glycosylation patterns, may be modified in pathological conditions. Such modifications may occur during numerous pathologies affecting airways such as cystic fibrosis or chronic bronchitis. Modifications of sulfation and sialylation patterns were reported for airway mucins prepared from patients suffering from cystic fibrosis or from severe chronic bronchitis [15]. It was suggested that a strong inflammatory reaction, generates an increase in mucin sialylation and an hyperexpression of sialyl-Lewis x that may be responsible for the specific lung colonization by *P. aeruginosa* of cystic fibrosis patients. Indeed, the use of purified sulfated O-glycans coupled to synthetic polymers gave good lines of evidence to directly correlate the increased sulfation of CF mucins and the increased affinity of *P. aeruginosa* to CF mucus [16]. Accordingly, the flagellar cap protein, FliD, from PAO1 strain demonstrated a clear association with sulfosialyl-Lewis x, as well as with other Lewis x derivatives [17].

Structural analysis of sulfated compounds is impaired by the difficulty to purify them from complex mixtures of mucin-type O-glycans. Indeed, anion exchange chromatography cannot easily discriminate sulfated compounds from other acidic compounds such as sialic acid and uronic acid containing O-glycans, which renders subsequent analyses by chromatographic or spectroscopic methods difficult. An elegant method for separating sialylated from sulfated mucin type O-glycans in mixture based on on-column carboxy-methylation of sialic acids was previously proposed [18]. The conversion of sialic acids to methyl esters after collection of neutral glycans permitted to elute them from anion exchanger independently of the sulfated compounds. Although very efficient, this method involves the chemical modification of sialic acids which are then lost for possible later use in biological tests. Furthermore, the necessity to use fairly large amounts of harmful iodomethane and the high concentration of pyridinium acetate for elution of sulfated oligosaccharides is not compatible with possible large scale purification of sulfated glycans from natural sources. Here we propose an alternative procedure aimed to purifying intact sulfated oligosaccharides in batch from complex mixtures of neutral and acidic O-linked glycans. It is based on the charge repulsiveness between sulfate groups of carbohydrates and strong cation exchanger gel and enables an easy one step purification of all sulfate containing glycans.

Material and methods

Preparation of mucins

Amphibian egg-jelly coats were extracted from intact eggs into Dulbecco's phosphate buffered saline (Sigma) containing 10 mM EDTA, 1 mM PMSF and 0.5% 2-mercaptoethanol at 4°C ovemight. The mixture was centrifuged and the supernatant was then dialysed for 72 h against water and finally freeze dried. Tracheobronchial mucins were isolated from a CF patient and were a kind gift of Prof. Philippe Roussel. They were prepared according to published methods [2].

Isolation of oligosaccharide–alditols

As starting materials we used mucins from egg-jelly of *Rana temporaria* (1 g), *Pleurodeles waltl* (200 mg) and human tracheobronchial mucins (less than 100 μg). Samples were submitted to reductive β-elimination for 72 h at 37°C in 100 mM NaOH containing 1 M $NaBH_4$. The reaction was stopped by the addition of DOWEX 50× 8 (25–50 mesh, H^+ form; Sigma-Aldrich) at 4°C until pH 6.5 is reached. After filtration on glass wool and evaporation to dryness, boric acid was eliminated by repetitive distillation as its methyl ester in the presence of methanol. The material was submitted to a first cationic exchange chromatography on DOWEX 50×2 (200–400 mesh, H^+

form) in order to remove residual peptides. Total eluate from the first DOWEX 50×2 column (five column volumes) was concentrated and applied to a second Dowex 50×2 column (200–400 mesh, H^+ form) to purify sulfated oligosaccharides.

Thin layer chromatography

Elution from cation exchange column was assessed by TLC chromatography. An aliquot of each fraction was run on silica gel 60 thin layer chromatography (Merck) in *n*-butanol/ethanol/Acetic acid/Pyridine/H_2O (10:100:3:10:30). Samples were detected following staining with orcinol reagent and charring.

Permethylation of oligosaccharide–alditols

Glycan samples were permethylated using the NaOH/dimethyl sulfoxide slurry method [19], with modifications on the subsequent extraction method for sulfate-containing oligosaccharides. Briefly, the reaction mixture was neutralized with 4N TFA before solid phase extraction on a Sep-pack C18 column. Permethylated oligosaccharides were eluted with ACN/ H2O 0.1% TFA (25:75).

Mass spectrometry analysis

Native and permethylated samples were analysed on a Voyager-DE STR MALDI-Tof Mass Spectrometer, (PerSeptive Biosystems, Framingham, MA, USA), equipped with a 337 nm UV laser. Samples were spotted by mixing directly on the target 1 µl of oligosaccharide solution and 1 µL of 2,5-dihydroxybenzoic acid matrix solution (10 mg/mL dissolved in MeOH/H_2O 0.1% trifluoroacetic acid). For tandem mass spectrometry, native and permethylated oligosaccharides were mixed 1:1 with α-cyano-4-hydrocinnamic acid matrix (in acetonitrile 0.1% trifluoroacetic acid, *v:v*) and analysed on a Q-Tof Ultima MALDI instrument (Micromass). The nitrogen UV laser (337 nm wavelength) was operated at a repetition rate of 10 Hz under full power (300 µJ/pulse). MS survey data were manually acquired and the decision to switch over to CID MS/MS acquisition mode for a particular parent ion was made on-the-fly upon examination of the summed spectra. Argon was used as the collision gas with a collision energy manually adjusted (between 50~200 V) to achieve optimum degree of fragmentation for the parent ions under investigation.

Sulfate content analysis

The sulfate content was measured by HPAEC according to [13]. Briefly, sulfate was released by hydrolysis with 1 M HCl (500 µl) for 5 h at 100°C and HCl was evaporated under a stream of nitrogen. The dry residue was dissolved in 200 µl of Milli-Q quality water (Millipore Corp., Milford, MA). 25 µL of this solution were directly injected onto a Dionex BioLC system equipped with an IonPac AS4A column (250×4 mm), an anion micromembrane suppressor, and a CDM 2 conductivity detector. The column was eluted with 0.04 M NaOH at a flow rate of 2 mL/min, and the separated anions were measured by conductivity detection with 30-microsiemens sensitivity. The chromatograms were analysed and integrated with Chromeleon software, version 6.40 (Dionex Corporation). A standard curve was constructed with K_2SO_4 solutions (6.25, 12.5, 25, and 35 µg/mL) to measure sulfate released from oligosaccharide fractions.

Results and discussion

Purification of sulfated compounds

O-linked oligosaccharide alditols were released from several biological sources by reductive β-elimination. For each sample, borates were co-evaporated with methanol and samples were passed on a first Dowex 50×2 cation exchange column (Sigma-Aldrich) to eliminate residual peptides. Columns were eluted with five column volumes of water and total eluates containing all sulfated and neutral oligosaccharides from each sample were freeze-dried. This step was not used to separate glycans, but to remove peptides that are bound on the column. Care was taken to immediately neutralize samples with diluted ammonia after cation exchange chromatography to protect eventual sialic acids from acid hydrolysis. It is noteworthy that eliminating residual peptides on a first column of Dowex 50×2 resin substantially improved the subsequent purification of sulfated oligosaccharides on a second cation exchanger. So, this step was systematically included in the workflow (data not shown).

Three samples were used in the course of the present study: two egg-jelly mucus from the amphibians *R. temporaria* and *P. waltl* and a sample of human tracheo-bronchial mucus. Oligosaccharides from amphibians have been extensively characterized so their glycosylation profiles are already well known and can be used to evaluate the effectiveness of the separation methodology. Total oligosaccharide samples devoid of peptides were directly loaded onto a second 50×2 (200–400 mesh H^+ form) without prior desalting and eluted with water. Column size was adapted depending on the total quantity of carbohydrate for each sample that ranged from about 200 mg for *R. temporaria* to less than 100 µg for tracheo bronchial O-glycans. During the course of the study, an average of 10 mL of resin per mg of carbohydrates was used to obtain

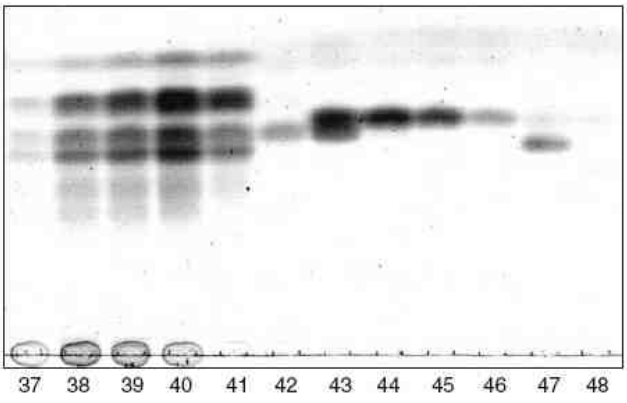


Fig. 1 Fractionation on a strong cation exchange Dowex 50×2 (200–400 mesh H^+ form) chromatography column of total O-glycans released by reductive β-elimination from egg jelly coat of *R. temporaria* eggs, monitored by TLC

a satisfactory resolution of separation between sulfated and non-sulfated glycans. Furthermore, irrespective of column size, columns with highest length to diameter ratios were used in order to maximize the separation efficiency.

O-glycans from *R. temporaria*

About two hundred mg of O-linked oligosaccharides were released and purified from one gram of dried egg jelly coat of the amphibian *R. temporaria*. This species synthesises a very complex mixture of neutral and acidic O-glycans (http://glycobase.univ-lille1.fr/). Acidic properties are conferred by the presence of either sulfate groups, sialic acid (Kdn) or glucuronic acid (GlcA) substituents. Some acidic glycans may be substituted by two different acidic groups. The sheer complexity of this sample appeared as ideal for assessing the selectivity of separation on cation exchange chromatography of different acidic compounds.

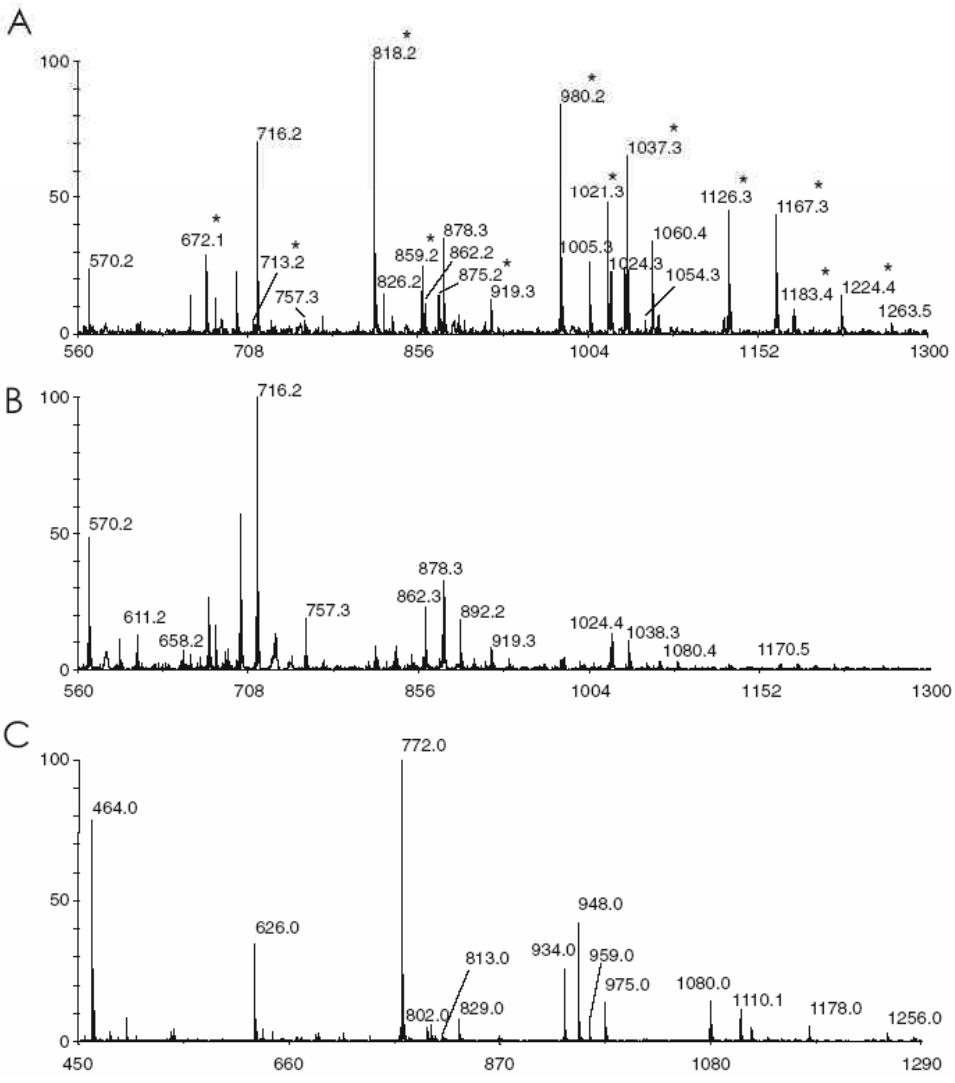
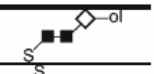
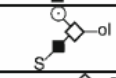
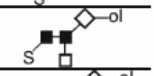
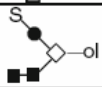
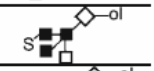
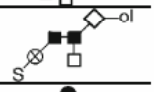
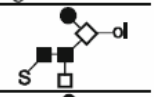
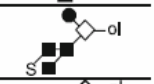
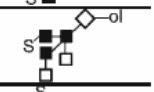
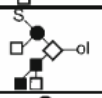
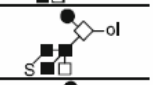
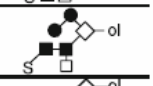
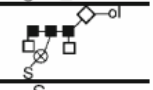
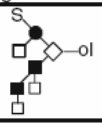


Fig. 2 Mass spectrometry analysis of glycans from *R. temporaria* released by reductive β-elimination. Positive reflectron mode MALDI-TOF spectrum of native **a** excluded fraction and **b** included fraction on a strong cation exchange chromatography. Asterisks labelled *m/z* values correspond to $[M+2Na–H]^+$ molecular ions of sulfated oligosaccharides, the others to $[M+Na]^+$ molecular ions of desulfated neutral oligosaccharides (in **a**) or genuine neutral oligosaccharides (in **b**). **c** negative reflectron mode MALDI-TOF spectrum of native excluded fraction; all *m/z* values correspond to $[M–H]^-$ molecular ions of sulfated oligosaccharides. For predicted compositions, see Tables 1 and 2

Table 1 Assignment of major oligosaccharides released from *Rana temporaria* mucins observed by MALDI-TOF MS in excluded fraction, in positive mode as (*a*) $[M+2Na-H]^+$ adducts and (*b*) negative modes as $[M-H]^-$ adducts. Individual structures were inferred from previous studies [20, 21, 22, unpublished results] and confirmed by ES-MS/MS fragmentation

(A) Excluded fraction.

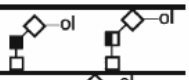
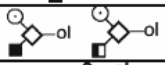
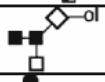
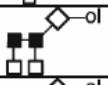
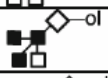
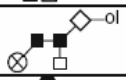
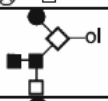
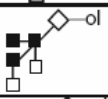
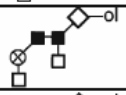
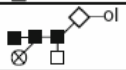
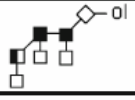
m/z a	m/z b	Calculated compositions	Structures
672.1	626.0	S Hex_2 HexNAc-ol	
713.2	667.0	S HexNAc Hex HexNAc-ol	
760.2	714.0	S Kdn Hex HexNAc-ol	
818.2	772.0	S dHex Hex_2 HexNAc-ol	
848.2	802.0	S HexA Hex_2 HexNAc-ol	
859.2	813.0	S dHex HexNAc Hex HexNAc-ol	
875.2	829.0	S HexNAc Hex_2 HexNAc-ol	
980.2	934.0	S dHex Hex_3 HexNAc-ol	
	948.0	S dHex HexA Hex_2-HexNAc-ol	
1021.2	975.0	S deHex HexNAcHex_2 HexNAc-ol	
1037.3		S HexNAc Hex_3 HexNAc-ol	
1126.3	1080.0	S $dHex_2$ Hex_3 HexNAc-ol	
1167.3	1121.1	S $dHex_2$ HexNAc Hex_2 HexNAc-ol	
1183.4		S dHex HexNAc Hex_3 HexNAc-ol	
1224.4	1178.0	S dHex $HexNAc_2$ Hex_2 HexNAc-ol	
1302.4	1256.0	S $dHex_2$ HexA Hex_3 HexNAc-ol	
1313.5		S $dHex_3$ HexNAc Hex_2 HexNAc-ol	

Structures are depicted according to the nomenclature of Kamerling and Vliegenthart [25]: ◇—ol, GalNAc-ol; ◆, β-GalNAc;  , *empty diamond*; α-GalNAc; *filled circle*, β-GlcNAc; *left-half-filled circle*, α-GlcNAc; *filled square*, β-Gal; *left-half-filled square*, α-Gal; *empty circle with center dot*, α-Kdn; *empty square*, α-Fuc; *empty circle with x*, β-GlcA

Total glycan sample was loaded on a Dowex 50×2 column (200–400 mesh, H + form; 150×2.5 cm) and was eluted with water at 0.2 ml/min; 2 ml fractions were collected. Considering the large quantity of carbohydrates, the elution of carbohydrate was monitored by orcinol-sulfuric colorimetric reaction after migration of all fractions on Thin Layer Chromatography (TLC). This visual method allowed an easy discrimination of different oligosaccharidic fractions according to their migration. As shown on Fig. 1, a carbohydrate rich fraction was first excluded from the column in tubes 37 to 42, which corresponds to the dead volume of the column. This excluded material exhibited a very heterogeneous pattern of migration on TLC, establishing that it is composed of a complex mixture of

Table 2 Assignment of major oligosaccharides released from *Rana temporaria* mucins observed by MALDI-TOF MS in included fraction in positive mode as $[M+Na]^+$ adducts. Individual structures were inferred from previous studies [20, 21, 22, unpublished results] and confirmed by ES-MS/MS fragmentation

(B) Included fraction

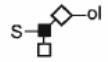
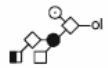
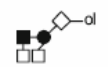
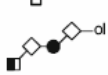
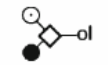
m/z	Calculated compositions	Structures
554	dHex Hex HexNAc-ol	
570	Hex_2 HexNAc-ol	
611	Hex HexNAc HexNAc-ol	
658	Kdn Hex HexNAc-ol	
716	dHex Hex_2 HexNAc-ol	
757	dHex HexNAc Hex HexNAc-ol	
862	$dHex_2$ Hex_2 HexNAc-ol	
878	dHex Hex_3 HexNAc-ol	
892	dHex HexA Hex_2 HexNAc-ol	
919	dHex Hex_2 HexNAc HexNAc-ol	
960	dHex Hex $HexNAc_2$ HexNAc-ol	
1024	$dHex_2Hex_3$ HexNAc-ol	
1038	$dHex_2HexAHex_2$ HexNAc-ol	
1054	dHex HexA Hex_3 HexNAc-ol	
1170	$dHex_3$ Hex_3 HexNAc-ol	

Structures are depicted according to the nomenclature of Kamerling and Vliegenthart [25]: ◇—ol, GalNAc-ol; ◆, β-GalNAc;  , *empty diamond*; α-GalNAc; *filled circle*, β-GlcNAc; *left-half-filled circle*, α-GlcNAc; *filled square*, β-Gal; *left-half-filled square*, α-Gal; *empty circle with center dot*, α-Kdn; *empty square*, α-Fuc; *empty circle with x*, β-GlcA

oligosaccharides of different sizes. Then, retarded carbohydrates were eluted from tube 43 onward. Orcinol positive fractions (tubes 43 to 76) were all pooled as a single included fraction for analysis. Excluded and included material show very distinct chromatographic patterns establishing that both fractions essentially contain different oligosaccharides.

The oligosaccharide content of each fraction was screened by MALDI-TOF MS in both positive and negative modes. Monosaccharide compositions of individual signals were calculated according to the observed *m/z* values and the nature of each compound was inferred from previously described oligosaccharides isolated from *R. temporaria* egg jelly coat (20, 21, 22, unpublished results) and confirmed by ESI-MS/MS fragmentation of most oligosaccharides (data not shown). As expected, MS profiles in positive mode of excluded and included fractions are different (Fig. 2). The included fraction is characterized by a complex pattern of $[M+Na]^+$ signals attributed to neutral and acidic oligosaccharides as presented on Tables 1 and 2. Identification of these oligosaccharides established that the acidic glycans are substituted either by Kdn or by GlcA, but never by sulfate group. In contrast, the profile of excluded fraction was characterized by the presence of $[M+2Na–H]^+$ signals attributed to the sulfate containing oligosaccharides previously described in *R. temporaria* jelly coat. This assignment was confirmed by acquiring MS spectrum in negative mode, which produced a series of $[M–H]^-$ ions of composition identical to their $[M+2Na–H]^+$ counterparts. It is noteworthy that none of the signals attributed to sulfated glycans could be observed in the included fraction. Positive mode MS spectrum of excluded fraction additionally shows

Table 3 Elution profile of oligosaccharide–alditols from *Pleurodeles waltl* mucin after cationic exchange column assessed by positive mode MALDI-TOF MS of each fraction 26 to 45 as (*a*) $[M+2Na–H]^+$ and (*b*) $[M+Na]^+$ adducts

	a	b	*26*	*27*	*28*	*29*	*30*	*31*	*32*	*33*	*34*	*35*	*36*	*37*	*38*	*39*	*40*	*41*	*42*	*43*	*44*	*45*
	656		+	+	+	+	+	+	+	+	+	+	+									
	1370		+	+	+	+	+	+	+	+	+											
	1516		+	+	+	+	+	+	+	+												
		1356										+	+	+	+	+	+	+	+	+		
		1210										+	+	+	+	+	+	+	+	+		
		1153															+	+				
		1106								+	+	+	+	+	+	+	+	+	+	+		
		960									+	+	+	+	+	+	+	+	+	+	+	
		1007																+	+	+		
		903													+	+	+	+	+	+	+	
		804														+	+					
		814												+	+		+	+	+	+		
		757													+	+	+	+	+	+		
		699																			+	+

Individual structures were inferred from previous studies [23, 24] and confirmed by ES-MS/MS fragmentation

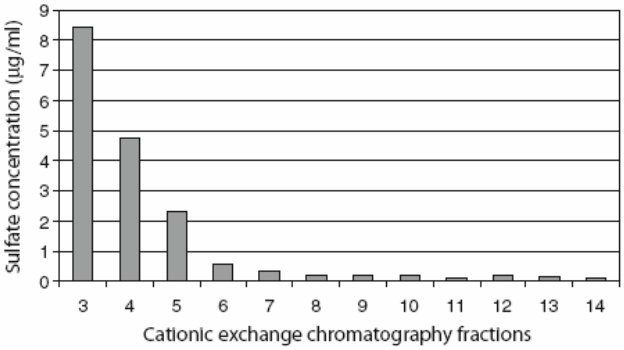


Fig. 3 Sulfate content of oligosaccharide fractions released from human tracheobronchial mucin after separation on strong cation exchange chromatography

numerous $[M+Na]^+$ ions attributed to non-sulfated oligosaccharides that have already been observed in the included fraction. As will be demonstrated in the next section, these ions resulted from a laser-induced desulfation process. Indeed, most of the $[M+2Na–H]^+$ major signals were accompanied by a signal at 102 mass units lower, corresponding to loss of a sodium sulfite moiety, *e.g.* *m/z* 672/570, 818/716, 859/757 and 980/878.

These data demonstrated that strong cation exchange chromatography enables a complete exclusion of sulfate containing oligosaccharides from the column. We hypothesise that repulsion forces between sulfate groups substituting the oligosaccharides and sulfonic acid functional groups attached to the styrene divinylbenzene resin prevent sulfated glycans to penetrate copolymer lattice and exclude them from the column, in contrast to other glycans, which are retarded in the polymer. Optimal separation between excluded and included fractions are obtained when using highest porosity (2% crosslinkage) and lowest particle size (200–400 mesh) resin, which maximizes the retardation of included fraction. So, Dowex 50×2 200×400 mesh (Sigma-Aldrich) and AG 50×2 200×400 mesh (Bio-Rad) appeared as the most suitable resins. The presence of non sulfated GlcA and Kdn containing oligosaccharides in the included fraction established that carboxylated functions are not excluded from the resin, which permits an easy separation between acidic oligosaccharides exclusively based on the presence of sulfate groups.

O-glycans from *P. waltl*

In a second step, we assessed the usefulness of the described method for purifying minute amounts of sulfated oligosaccharides from very complex mixture of non-sulfated glycans. Contrarily to the amphibian *R. temporaria* that synthesises large amounts of sulfated oligosaccharides, oviductal mucins of *P. waltl* are almost exclusively substituted by neutral and sialylated oligosaccharides, as previously established [23, 24]. However, careful re-evaluation of the oligosaccharidic content of *P. waltl* permitted the isolation, by multidimensional HPLC, of single sulfated glycans as a very minor component, establishing the presence of small amounts of sulfate substitutions (Supplementary Fig. 1; unpublished data).

Total O-glycans released by reductive β-elimination from 200 mg of *P. waltl* were separated on a Dowex 50×

Table 4 Assignment of major human tracheobronchial oligosaccharides observed in excluded fraction 3 as $[M–H]^-$adduct by negative mode MALDI-TOF MS and included fraction 12 as $[M+Na]^+$ adduct by positive mode MALDI-TOF MS

m/z	Calculated compositions
Excluded fraction 3 as $[M–H]^-$adduct by negative mode MALDI-TOF MS	
813	S dHex HexNAc Hex HexNAc-ol
829	S HexNAc Hex_2 HexNAc-ol
975	S dHex HexNAcHex_2 HexNAc-ol
1040	NeuAc HexNAc Hex_2 HexNAc-ol
1121	S $dHex_2$ HexNAc Hex_2 HexNAc-ol
1137	S dHex HexNAc Hex_3 HexNAc-ol
1178	S dHex $HexNAc_2$ Hex_2 HexNAc-ol
1186	NeuAc dHex HexNAc Hex_2 HexNAc-ol
1283	S $dHex_2$ HexNAc Hex_3 HexNAc-ol
1340	S dHex $HexNAc_2$ Hex_3 HexNAc-ol
1486	S $dHex_2$ $HexNAc_2$ Hex_3 HexNAc-ol
1632	S $dHex_3HexNAc_2Hex_3$ HexNAc-ol
1648	S $dHex_2HexNAc_2Hex_4$ HexNAc-ol
1794	S $dHex_3HexNAc_2Hex_4$ HexNAc-ol
1851	S $dHex_2HexNAc_3Hex_4$ HexNAc-ol
1997	S $dHex_3HexNAc_3Hex_4$ HexNAc-ol
Included fraction 12 as $[M+Na]^+$ adduct by positive mode MALDI-TOF MS	
757	dHex HexNAc Hex HexNAc-ol
773	HexNAc Hex_2 HexNAc-ol
862	$dHex_2$ Hex_2 HexNAc-ol
878	dHex Hex_3 HexNAc-ol
919	dHex HexNAc Hex_2 HexNAc-ol
935	HexNAc Hex_3 HexNAc-ol
976	$HexNAc_2$ Hex_2 HexNAc-ol
1024	$dHex_2$ Hex_3 HexNAc-ol
1065	$dHex_2$ HexNAc Hex_2 HexNAc-ol
1081	dHex HexNAc Hex_3 HexNAc-ol
1122	dHex $HexNAc_2$ Hex_2 HexNAc-ol
1138	$HexNAc_2$ Hex_3 HexNAc-ol
1211	$dHex_3$ HexNAc Hex_2 HexNAc-ol
1227	$dHex_2$ HexNAc Hex_3 HexNAc-ol
1243	dHex HexNAc Hex_4 HexNAc-ol
1284	dHex $HexNAc_2$ Hex_3 HexNAc-ol
1373	$dHex_3$ HexNAc Hex_3 HexNAc-ol
1389	$dHex_2$ HexNAc Hex_4 HexNAc-ol
1430	$dHex_2$ $HexNAc_2$ Hex_3 HexNAc-ol
1446	dHex $HexNAc_2$ Hex_4 HexNAc-ol
1576	$dHex_3$ $HexNAc_2$ Hex_3 HexNAc-ol
1592	$dHex_2$ $HexNAc_2$ Hex_4 HexNAc-ol
1649	dHex $HexNAc_3$ Hex_4 HexNAc-ol

2 column (200–400 mesh, H^+ form; 50×2 cm) under similar conditions to those above and collected as 2 mL fractions. Elution profile of oligosaccharides was directly assessed by MALDI-TOF MS analysis of each eluted fraction. As shown in Table 3, three oligosaccharides observed as signals at *m/z* 656, 1370 and 1516 are clearly excluded from the cation exchange column in fractions 26 to 34, whereas most other compounds are eluted from fraction 34 to fraction 45. Signals at *m/z* 656, 1370 and 1516 observed in the excluded fractions were attributed to $[M+2Na–H]^+$ adducts of $SO_3deHex_1Hex_1HexNAc_1$-ol, $SO_3deHex_2Hex_2HexNAc_3$-ol and $SO_3deHex_3Hex_2HexNAc_3$-ol, respectively. ^1H NMR analysis of the pooled excluded fraction (data not shown) established the exact

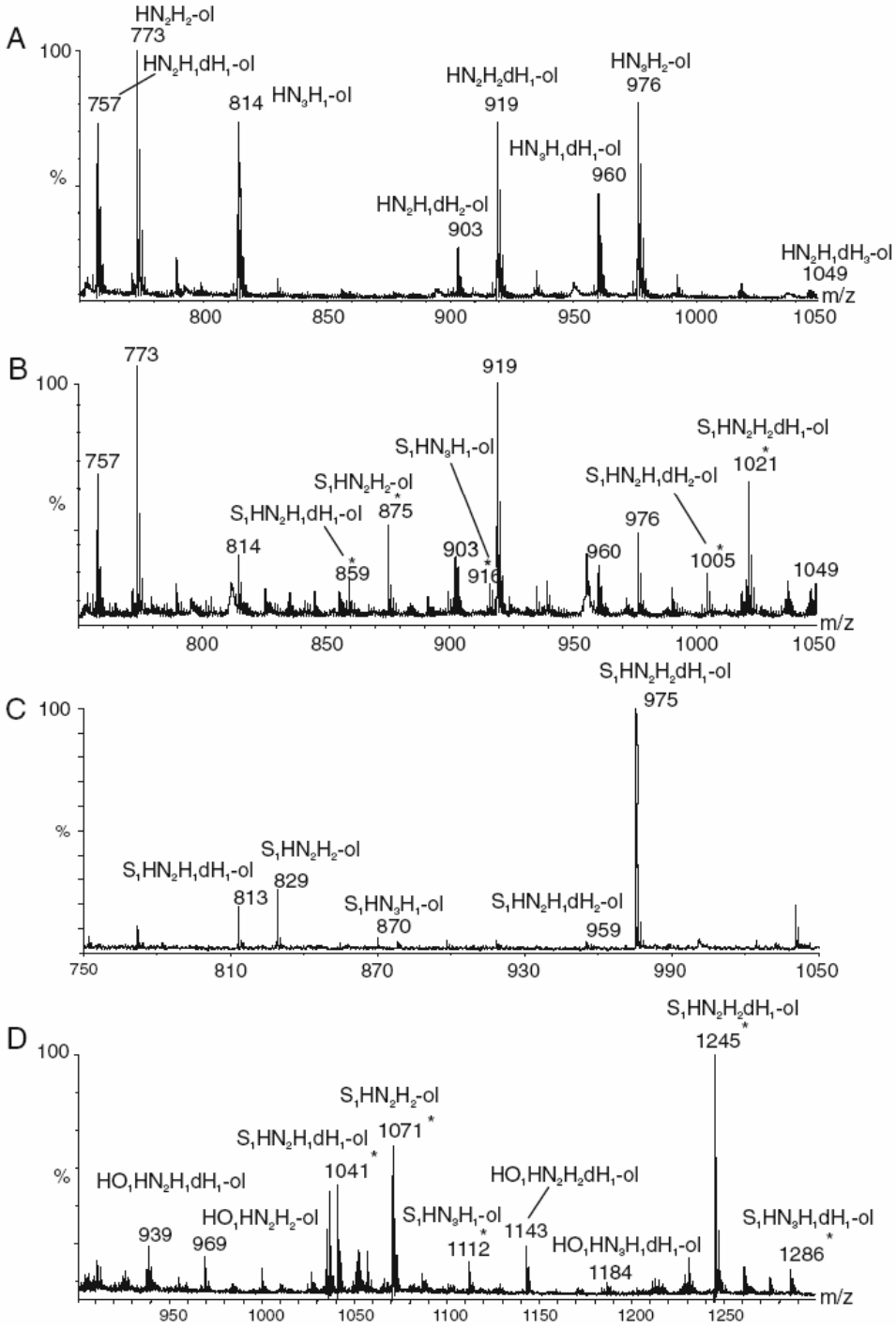


Fig. 4 Details of mass spectrometry analyses of O-glycans released from human bronchial mucins. Positive-mode MALDI-TOF spectra of **a** the native neutral fraction, **b** the native sulfated fraction and **d** the permethylated sulfated fraction. **c** Negative mode MALDI-TOF spectra of the native sulfated fraction. Asterisks labelled *m/z* values correspond to disodiated ions of sulfated oligosaccharides. For predicted compositions, see Table 4

structure of its major component as a core 1 based sulfated trisaccharide Fuc(α1–2)[SO_3(4)]Gal(β1–3)GalNAc-ol, in agreement with the presence of a signal at *m/z* 656 in MS analysis. Signals at *m/z* 1370 and 1516 were identified as very minor components whose structures were tentatively assigned to sulfated trisaccharide extended from GalNAc-ol C6 with Fuc(α1–2)GalNAc(β1–4)GlcNAc(β1- and Fuc (α1–2) GalNAc(β1–4)[Fuc(α1–3)]GlcNAc(β1-branches, respectively. It is noteworthy that all sulfated oligosaccharides are excluded in the same fractions, irrespective of their molecular weights, which demonstrates that their elution volume is exclusively dependent on the presence of sulfate group. Signals observed in included fractions were all attributed to $[M+Na]^+$ adducts of neutral and sialylated oligosaccharides whose structures were assigned based on previous work (Strecker 1992 a and b; unpublished work) and in accordance with their calculated compositions. In contrast to the sulfated compounds, the elution volumes of included compounds are dependent on their molecular weight, which demonstrates that the cation exchanger resin acts as a molecular sieve for neutral and sialylated oligosaccharides. These data established that the cation exchange chromatography is useful for purifying minor sulfated oligosaccharides among a complex mixture of non-sulfated compounds.

Human tracheobronchial mucins

In a last step, we validated the method by analysing minute amounts of oligosaccharides isolated from human tracheobronchial mucins. Less than 100 μg of reduced O-glycans released from purified tracheobronchial mucins by reductive β-elimination were loaded on 150 μL of DOWEX 50× 2 resin (200–400 mesh, H^+ form) packed in a 200 μL automatic pipette cone. Column was eluted with water and 25 μL fractions were collected. Elution of sulfated oligosaccharides was monitored by assessing sulfate concentration on a Dionex BioLC system fitted with an IonPac AS4A column and a conductivity detector. As expected,

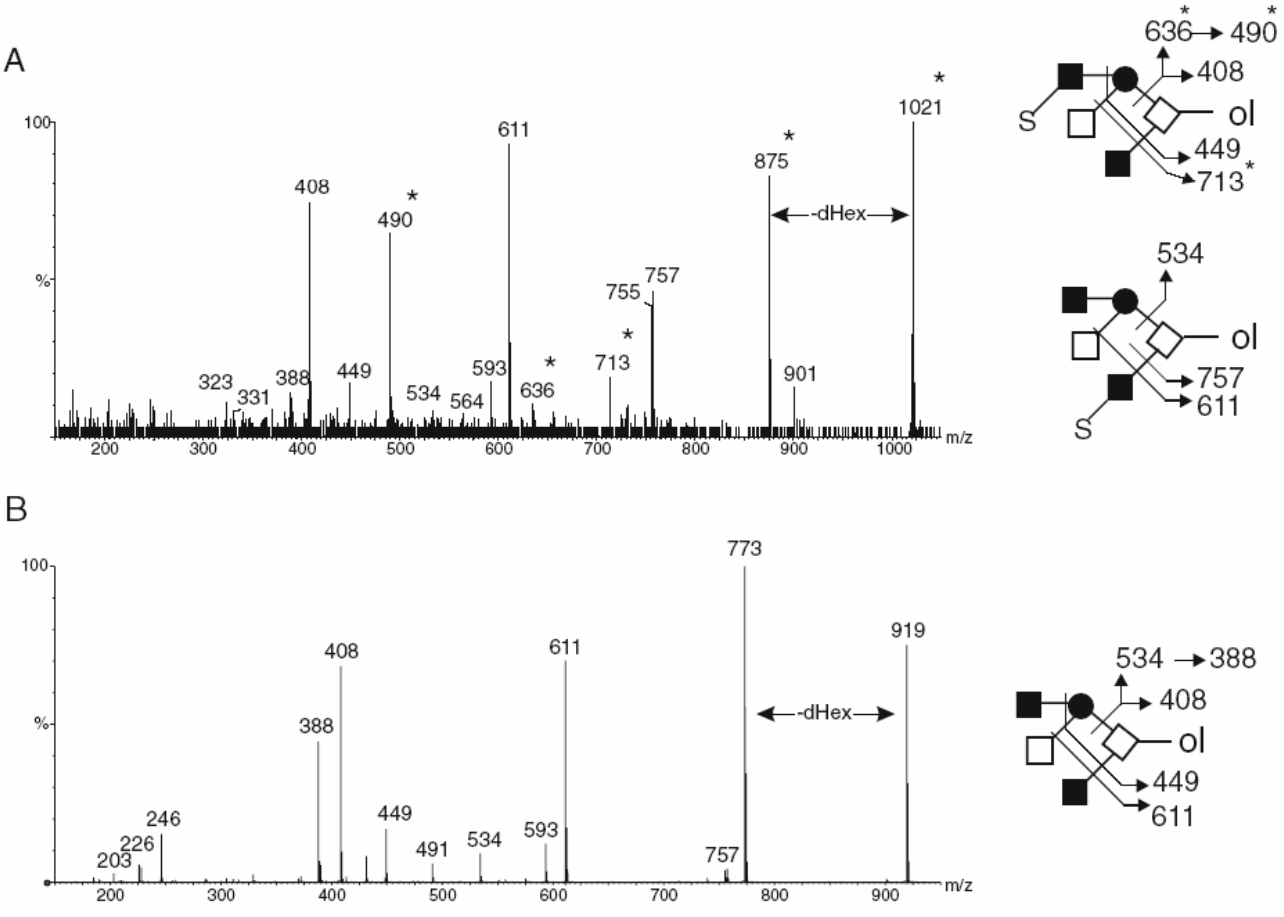


Fig. 5 Collision induced MALDI-Q-TOF fragmentation spectra of **a** $[M+2Na–H]^+$ parent ion at *m/z* 1021 of sulfated oligosaccharide $SO_3Fuc_1Gal_2GlcNAc_1GalNAc$-ol and **b** $[M+Na]^+$ parent ion at *m/z* 919 of on-target desulfated oligosaccharide $Fuc_1Gal_2GlcNAc_1Gal$ NAc-ol isolated from human bronchial mucin. Symbols are depicted in legend from Tables 1 and 2. *Asterisks* labelled *m/z* values correspond to $[M+2Na–H]^+$ molecular ions of the sulfated oligosaccharides. Fragmentation pattern of sulfated oligosaccharides demonstrate the presence of two different isobaric structures in mixture that only differ by the position of sulfation either on the upper or lower branch

sulfate concentration is maximal in fractions 3, 4 and 5, which correspond to the dead volume of the column (Fig. 3). As observed by MALDI-TOF MS, sulfate containing fractions are composed of a very complex mixture of almost exclusively sulfated oligosaccharides (Table 4). No neutral oligosaccharide was detected in sulfate containing fractions, but two minor sialylated oligosaccharides were observed (Table 4). Despite the low bed volume of resin, sulfated and non sulfated molecules were unexpectedly well resolved. Oligosaccharides were detected up to fraction 12, which is completely devoid of sulfated oligosaccharides and exclusively composed of neutral compounds (Table 4).

Analysis of sulfated glycans by MALDI-MS/MS

As mentioned above, monosulfated oligosaccharides crystallized in DHB matrix are observed as $[M+2Na-H]^+$ molecular ions in positive-mode MALDI-TOF MS spectra. As an example, the spectrum of a sulfated fraction of human bronchial O-glycans show five molecular ion signals at *m/z* 859, 875, 916, 1005 and 1021, corresponding to $[M+2Na-H]^+$ of sulfated oligosaccharides (Fig. 4b). Assignment was confirmed by observing their $[M-H]^-$ molecular ions at *m/z* 813, 829, 870, 959 and 975 in negative-mode (Fig. 4c). Furthermore, a series of monosodiated ions at *m/z* 773, 814, 903, 919, 960 and 976 were clearly observed in positive-mode (Fig. 4b), which correspond to the neutral equivalents of all sulfated oligosaccharides having loss a sodium sulfite moiety. All these signals were also observed in the neutral fraction of O-glycans (Fig. 4a), which suggests that they originate either from a contamination of the sulfated fraction by neutral oligosaccharides or from a partial desulfation of oligosaccharides during MALDI-TOF analysis. Positive-mode MALDI-TOF analysis of the permethylated derivatives of

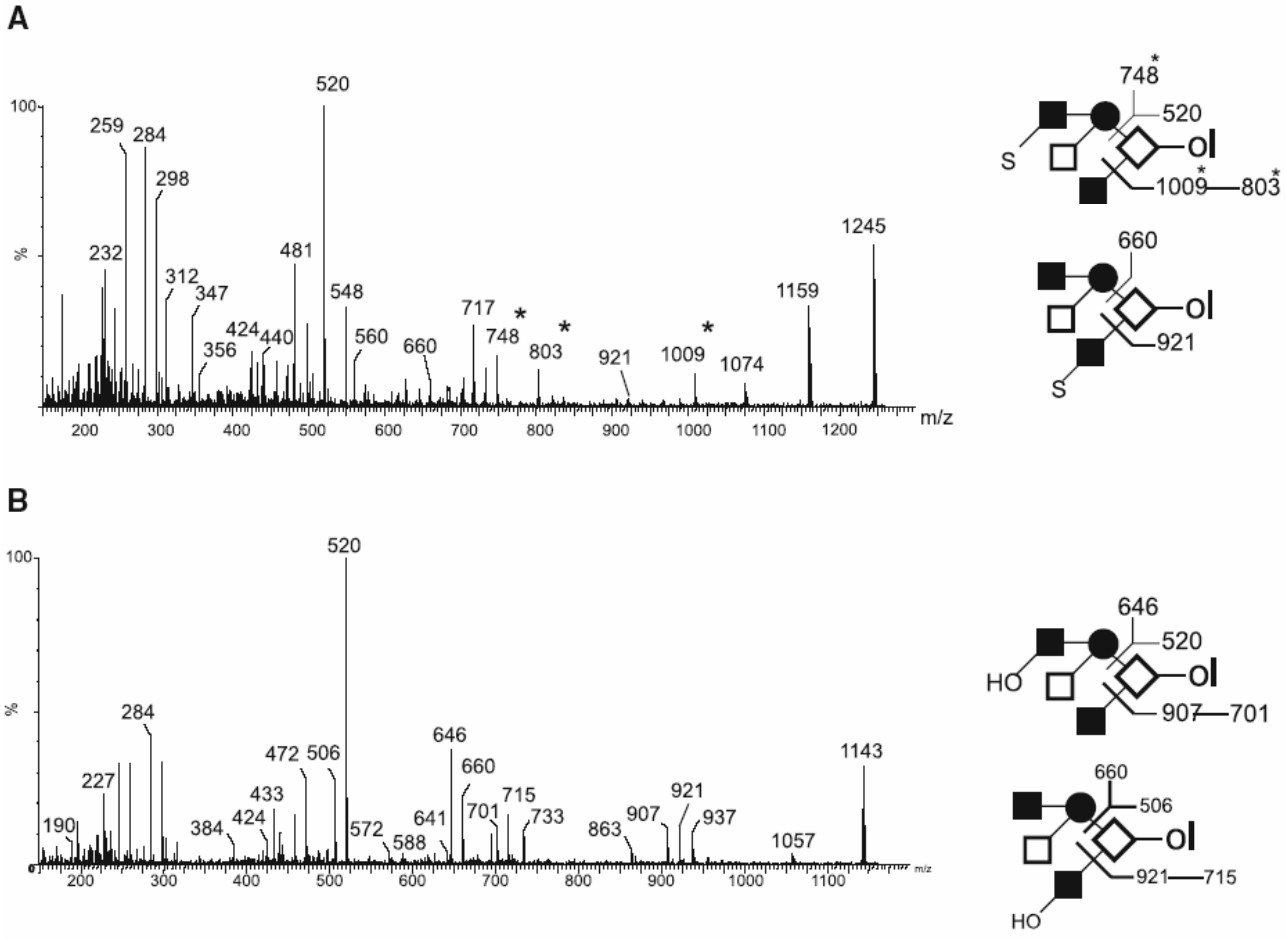


Fig. 6 Collision induced MALDI-Q-TOF fragmentation spectra of **a** $[M+2Na-H]^+$ parent ion at *m/z* 1245 of sulfated oligosaccharide $SO_3Fuc_1Gal_2GlcNAc_1GalNAc$-ol permethylated derivative and **b** $[M+Na]^+$ parent ion at *m/z* 1143 of on-target desulfated oligosaccharide $Fuc_1Gal_2GlcNAc_1GalNAc$-ol permethylated derivative isolated from human bronchial mucin. Fragmentation patterns of both parent ions demonstrate the presence of two different isobaric structures in mixture that only differ by the position of sulfation either on the upper or lower branch

sulfated fraction showed an identical set of sulfated oligosaccharides as $[M+2Na–H]^+$ at *m/z* 1041, 1071, 1112, 1245 and 1286 (Fig. 4d), each accompanied by signals corresponding to loss of sodium sulphite (*m/z* 939, 969, 1010, 1143 and 1184), as described above. These are the neutral equivalents of permethylated sulfated oligosaccharides in which a $–CH_3$ group has been replaced by a hydroxyl group, originating from an on-target desulfation process and not from genuine unsulfated molecules.

The detection of different sets of di- and mono-sodiated adducts enables an easy distinction of sulfated and non sulfated components in positive-mode MALDI-TOF MS spectra. This principle can be extended to MS/MS analyses of sulfated oligosaccharides that generate both mono- and disodiated ions in a single MS/MS spectrum, depending on the presence of sulfate group in the fragment ions. Indeed, as shown in the MS/MS spectrum of sulfated oligosaccharides (Fig. 5a and 6a), sulfated fragment ions, marked by an asterisk, are all characterized by being disodiated. As an example, fragmentation of parent ion at *m/z* 1021 (Fig. 5a) established the presence of a sulfo-Lewis x containing core 2 O-glycans $SO_3Fuc_1Gal_2GlcNAc_1GalNAc$-ol. Although many non-sulfated fragment ions are observed on the spectra, we believe that they do not originate from a secondary fragmentation induced desulfation, as demonstrated by the total absence of ion at *m/z* 773 potentially resulting from desulfation of primary fragment ion at *m/z* 875. This particular ion is however clearly observed in the MS/MS spectrum of the parent ion at *m/z* 919 that originates from the primary on-target desulfation of the sulfated oligosaccharide. A comparison of the MS/MS spectra of the permethylated derivatives of $SO_3Fuc_1Gal_2GlcNAc_1GalNAc$-ol at *m/z* 1245 with its desulfation product at *m/z* 1143 confirmed that all sulfate containing disodiated fragment ions are replaced by monosodiated fragment ions with a free hydroxyl group. Fragmentation of permethylated derivatives also confirmed the absence of desulfation during collision induced fragmentation.

Conclusion

The present data established the use of strong cation exchange resin as a reliable method for specifically purifying sulfated oligosaccharides from a very complex mixture of neutral, sialylated and sulfated components. The purification method, which is exclusively based on the presence of sulfate groups, is selective enough to extract very minor sulfated oligosaccharides. Although not exemplified in the present report, this method also permits to purify disulfated glycans which are co-eluted with monosulfated glycans (unpublished data). Furthermore, this method may be adapted for any quantity of purified glycans from the microgram scale up to at least hundreds of milligram. Indeed, we demonstrated that it may be used for large scale purification simply by adapting bed resin size, collection and detection procedures. It does not require any chemical modification of oligosaccharides of interest and thus can be used in the context of analysis of biologically relevant molecules. Furthermore, contrary to classical protocols of purification of acidic oligosaccharides based on anion exchange chromatography, sulfated molecules are eluted under salt free conditions, which permits an omission of a tricky desalting step.

References

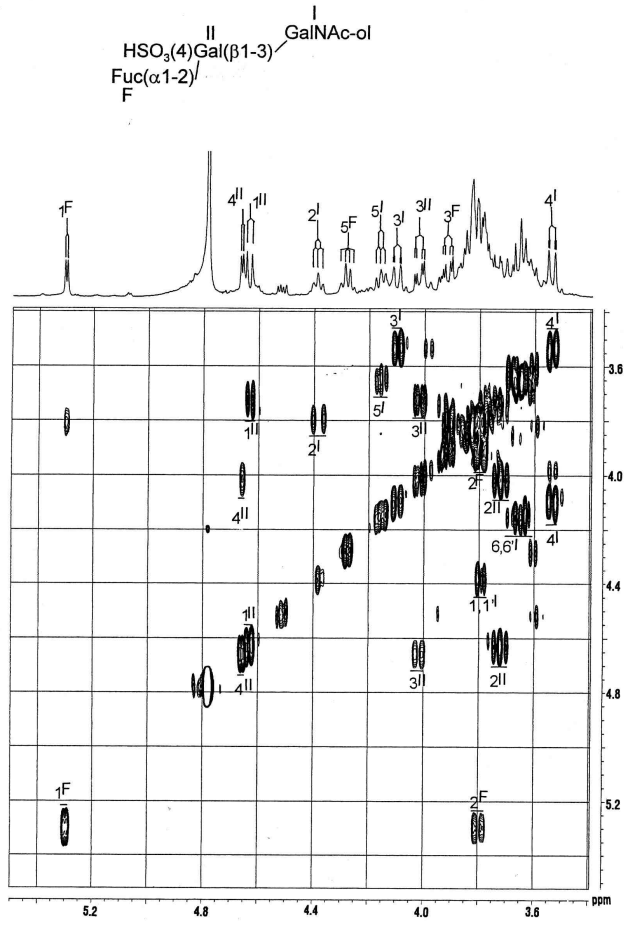
1. Degroote, S., Ducourouble, M.P., Roussel, P., Lamblin, G.: Sequential biosynthesis of sulfated and/or sialylated Lewis x determinants by transferases of the human bronchial mucosa. Glycobiology **9**, 1199–1211 (1999)
2. Degroote, S., Maes, E., Humbert, P., Delmotte, P., Lamblin, G., Roussel, P.: Sulfated oligosaccharides isolated from the respiratory mucins of a secretor patient suffering from chronic bronchitis. Biochimie **85**, 369–379 (2003)
3. Lamblin, G., Degroote, S., Perini, J-M., Delmotte, P., Scharfman, A., Davril, M., Lo-Guidice, J-M., Houdret, N., Dumur, V., Klein, A., Roussel, P.: Human airway mucin glycosylation: a combinatory of carbohydrate determinants which vary in cystic fibrosis. Glycoconj. J. **18**, 661–684 (2001)
4. Spiro, R.G., Bhoyroo, V.D.: Occurrence of sulfate in the asparagine-linked complex carbohydrate units of thyroglobulin-identification and localization of galactose-3-sulfate and *N*-acetylglucosamine-6-sulfate residues in the human and calf proteins. J. Biol. Chem. **263**, 14351–14358 (1988)
5. Hooper, L.V., Beranek, M.C., Manzella, S.M., Baenziger, J.U.: Differential expression of GalNAc-4-sulfotransferase and GalNAc-transferase results in distinct glycoforms of carbonic anhydrase VI in parotid and submaxillary glands. J. Biol. Chem. **270**, 5985–5993 (1995)
6. Hård, K., Van Zadelhoff, C., Moonen, P., Kamerhing, J.P., Vliegenthart, J.F.C.: The Asn-linked carbohydrate chains of human Tamm-Horssfal glycoprotein of one male-novel sulfated and novel *N*-acetylgalactosamine containing N-linked carbohydrate chains. Eur. J. Biochem. **209**, 895–915 (1992)
7. Yuen, C.-T., Lawson, A.M., Chai, W., Larkin, M., Stoll, M.S., Stuart, A.C., Sullivan, F.X., Ahem, T.J., Feizi, T.: Novel sulfated ligands for the cell adhesion molecule E-selectin revealed by the neoglycolipid technology among O-linked ohigosaccharides on an ovarian cystadenoma glycoprotein. Biochemistry **31**, 9126–9131 (1992)
8. Burger, D., Simon, M., Perruisseau, G., Steck, A.J.: The epitope (s) recognized by HNK-1 antibody and IgM paraprotein in neuropathy is present on several N-linked oligosaccharide structures on human P0 and myelin-associated glycoprotein. J. Neurochem. **54**, 1569–1575 (1990)
9. Fiete, D., Srivastava, V., Hindsgaul, O., Baenziger, J.U.: A hepatic reticuloendothelial cell receptor specific for $S0_4$–4GalNAcb1,4GlcNAcb that mediates rapid clearance of lutropin. Cell **67**, 1103–1110 (1991)
10. Hemmerich, S., Leffler, H., Rosen, S.D.: Structure of the O-glycans in GlyCAM-1, an endothelial-derived ligand for L-selectin. J. Biol. Chem. **270**, 12035–12047 (1995)

11. Thomsson, K.A., Schulz, B.L., Packer, N.H., Karlsson, N.G.: MUC5B glycosylation in human saliva reflects blood group and secretor status. Glycobiology **15**, 791–804 (2005)
12. Lo-Guidice, J.M., Wieruszeski, J.M., Lemoine, J., Verbert, A., Roussel, P., Lamblin, G.: Sialylation and sulfation of the carbohydrate chains in respiratory mucins from a patient with cystic fibrosis. J. Biol. Chem. **269**, 18794–18813 (1994)
13. Robbe, C., Capon, C., Maes, E., Rousset, M., Zweibaum, A., Zanetta, J.P., Michalski, J.C.: Evidence of regio-specific glycosylation in human intestinal mucins: presence of an acidic gradient along the intestinal tract. J. Biol. Chem. **278**, 46337–46348 (2003)
14. Thornton, D.J., Sheehan, J.K.: From mucins to mucus: toward a more coherent understanding of this essential barrier. Proc. Am. Thorac. Soc. **1**, 54–61 (2004)
15. Davril, M., Degroote, S., Humbert, P., Galabert, C., Dumur, V., Lafitte, J.J., Lamblin, G., Roussel, P.: The sialylation of bronchial mucins secreted by patients suffering from cystic fibrosis or from chronic bronchitis is related to the severity of airway infection. Glycobiology **9**, 311–321 (1999)
16. Scharfman, A., Delmotte, P., Beau, J., Lamblin, G., Roussel, P., Mazurier, J.: Sialyl-Le(x) and sulfo-sialyl-Le(x) determinants are receptors for *P. aeruginosa*. Glycoconj J. **17**, 735–740 (2000)
17. Scharfman, A., Arora, S.K., Delmotte, P., Van Brussel, E., Mazurier, J., Ramphal, R., Roussel, P.: Recognition of Lewis x derivatives present on mucins by flagellar components of *Pseudomonas aeruginosa*. Infect. Immun. **69**, 5243–5248 (2001)
18. Karlsson, N.G., Karlsson, H., Hansson, G.C.: Strategy for the investigation of O-linked oligosaccharides from mucins based on the separation into neutral, sialic acid- and sulfate-containing species. Glycoconj. J. **12**, 69–76 (1995)
19. Ciucanu, I., Kerek, F.: A simple and rapid method for the permethylation of carbohydrates. Carbohyd. Res. **131**, 209–217 (1984)
20. Maes, E., Florea, D., Delplace, F., Lemoine, J., Plancke, Y., Strecker, G.: Structural analysis of the oligosaccharide–alditols released by reductive b-elimination from oviducal mucins of *Rana temporaria*. Glycoconj. J. **14**, 127–146 (1997)
21. Florea, D., Maes, E., Strecker, G.: Primary structure of seven sulfated oligosaccharide–alditols released by reductive b-elimination from oviducal mucins of *Rana temporaria*. Characterization of the sequence HSO_3(3)GlcA(b1–3)Gal. Carbohydr. Res. **302**, 179–189 (1997)
22. Coppin, A., Maes, E., Morelle, W., Strecker, G.: Structural analysis of 13 neutral oligosaccharide–alditols released by reductive b-elimination from oviducal mucins of *Rana temporaria*. Eur J Biochem. **266**, 94–104 (1999)
23. Strecker, G., Wieruszeski, J.M., Michalski, J.C., Alonso, C., Boilly, B., Montreuil, J.: Characterization of Le(x), Le(y) and A Le(y) antigen determinants in KDN-containing O-linked glycan chains from *Pleurodeles waltl* jelly coat eggs. FEBS Lett. **298**, 39–43 (1992)
24. Strecker, G., Wieruszeski, J.M., Michalski, J.C., Montreuil, J.: ^{1}H- and ^{13}C-n.m.r. spectroscopy of 2-oxo-3-deoxy-D-glycero-D-galactononulosonic acid-containing oligosaccharide–alditols bearing Lewis X, Lewis Y and A-Lewis Y determinants isolated from the jelly coat of *Pleurodeles waltl* eggs. Biochem. J. **287**, 905–909 (1992)
25. Kamerling, J.P., Vliegenthart, J.F.G.: High-resolution ^{1}H-nuclear magnetic resonance spectroscopy of oligosaccharide–alditols released from mucin-type O-glycoproteins. Biol. Magn. Res. **10**, 1–288 (1992)

Supplementary data

SupFig. 1: NMR analysis of major sulphated *O*-glycan from egg jelly coat of *Pleurodeles waltl*. **a**) ^{1}H-^{1}H COSY-90 NMR and signal attributions, **b**) ^{1}H-NMR chemical shifts of sulphated trisaccharide.

a



b

	GalNAc-ol I	Gal II	Fuc F
H-1	3,79	4,613	5,280
H-1'	3,79	_	_
H-2	4,389	3,725	3,800
H-3	4,099	4,019	3,914
H-4	3,537	4,660	3,84
H-5	4,158	n.d.	4,276
H-6	3,65	n.d.	1,222
H-6'	3,65	_	_
Nac	2,28	_	_

Analysis of complex glycan mixtures by CID-MS/MS

Yann Guérardel, Kay Hooi Khoo

Fig. 1

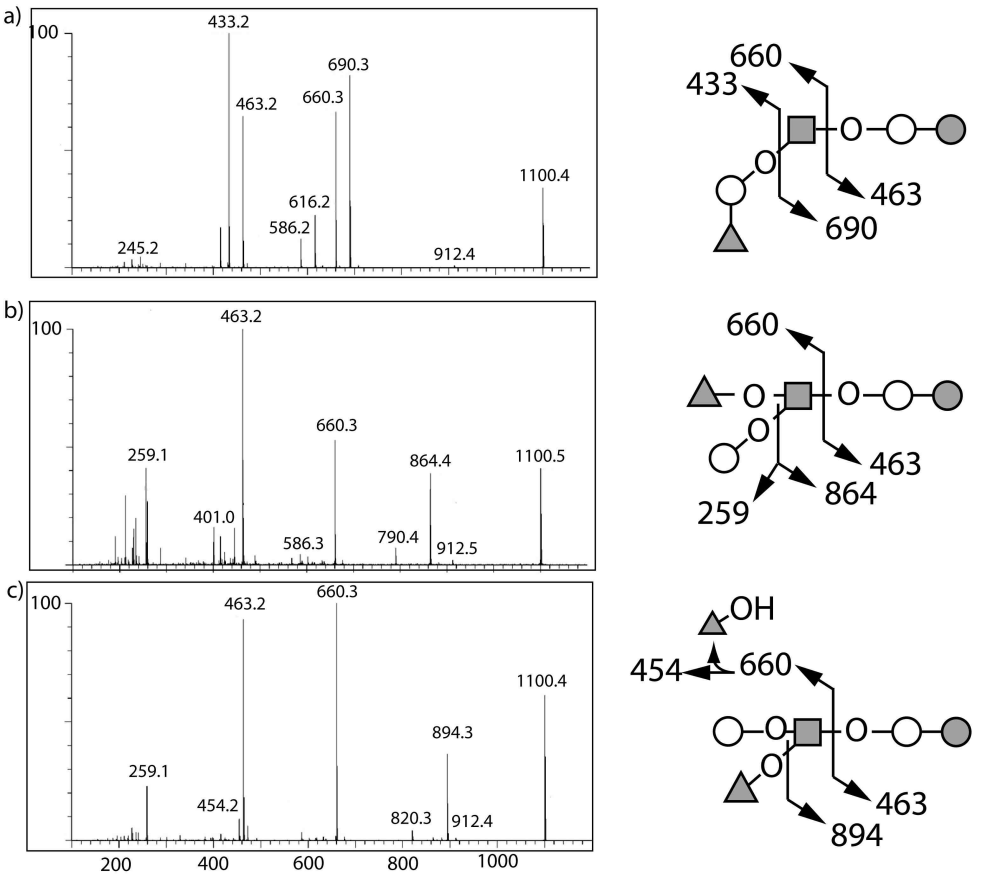
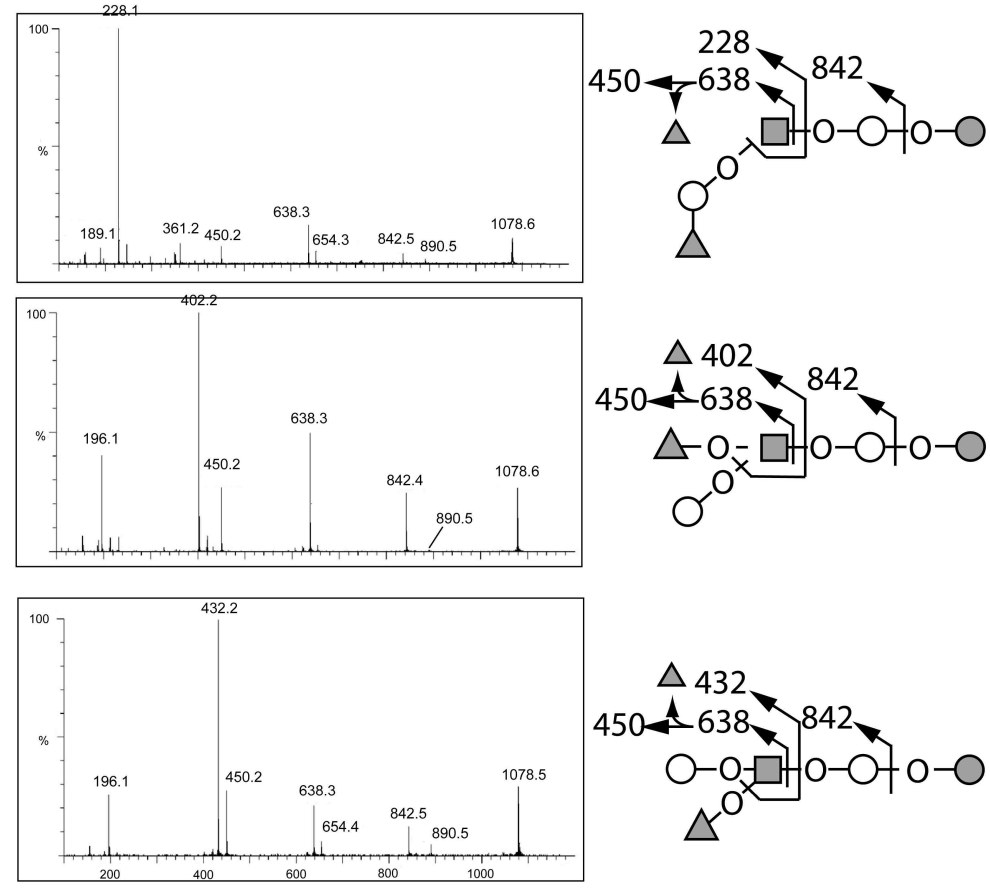


Fig. 2

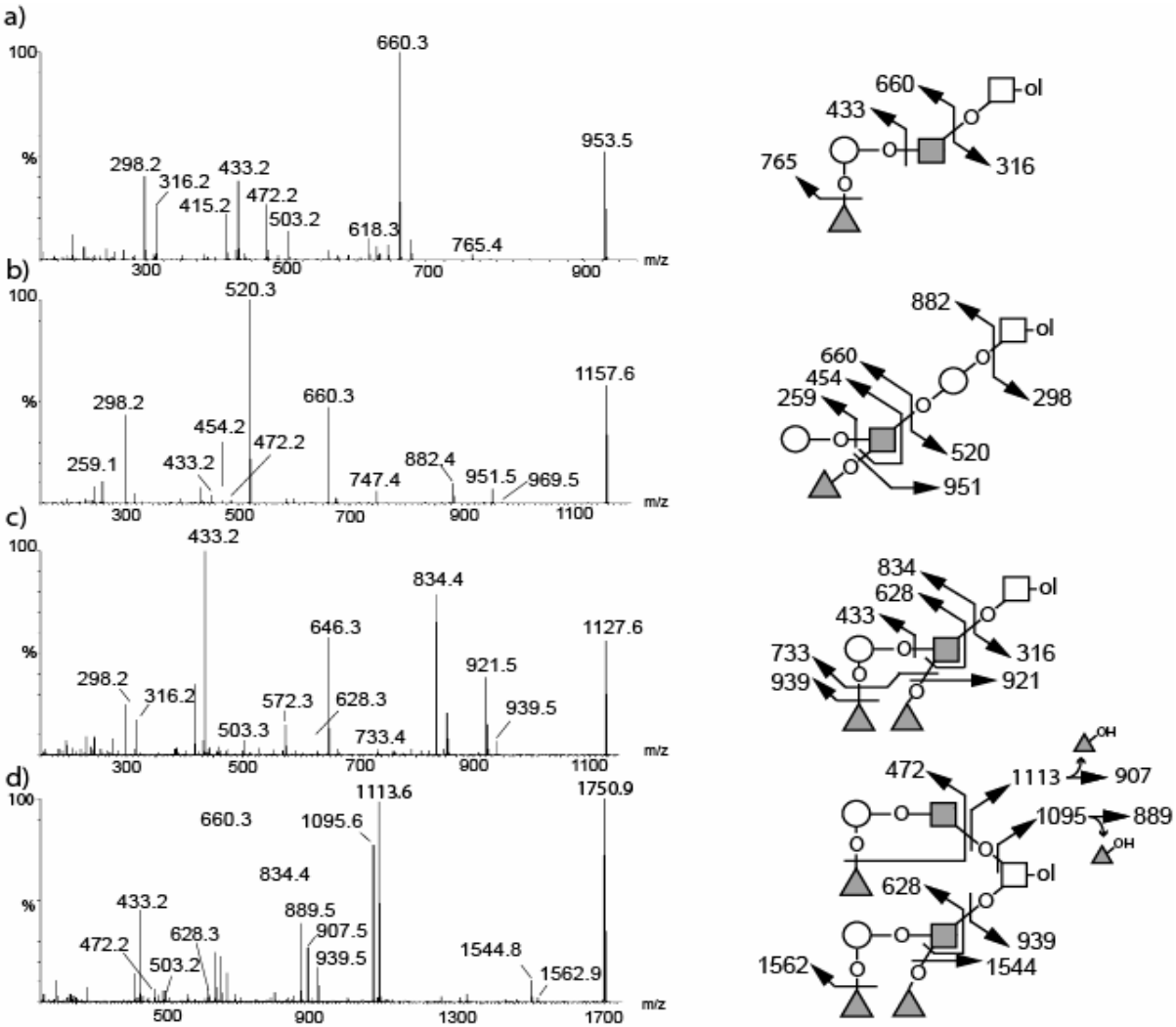


Analysis of complex glycan mixtures by CID-MS/MS

Fragmentation patterns of purified complex glycans

Susceptibility of the X(1-3)GlcNAc linkage. As established by previous studies, the predominant fragmentation pathway observed in per-methylated glycans involves cleavage of the GlcNAc bonds to generate Y- and B-type ions. This cleavage has been widely used for oligosaccharide sequencing in early FAB-MS experiments. It was thought to originate from the so-called A-type cleavage of a $[M+H]^+$ ion to produce an oxonium-type fragment ions. It is presently not known whether or not the fragmentation pattern that is observed during MALDI-QTOF analysis of $[M+Na]^+$ adducts originates from a similar process. As shown on Fig. 1, this fragmentation pathway leads to formation of intense Y- and B-types ions at m/z 660 and 463, respectively, for all three isobars of Lacto-N-Fucopentaose (LNF). These compounds differ according to both their lactosamine type (type 1 for LNF-I and -II, type 2 for LNF-III) and the fucose position (Gal for LNF-I, GlcNAc for LNF-II and -III) which define H-1, Le^a and Le^x motifs for LNF-I, -II and -III, respectively. In addition to these common ions, a set of sequence-specific ions can be observed in each spectrum. Examination of the fragmentation patterns established that these ions all originate from the release of the C-3 substitution of the GlcNAc residue by generating C- and Z-type ions. In this respect, fragmentation patterns of LNFII and LNFIII differ owing to the presence of either a $[M-Gal+Na]^+$ Z-ion at m/z 864 (M-236) or a $[M-Fuc+Na]^+$ Z-ion at m/z 894 (M-206), respectively. Each ion comes together with a related ion of lesser intensity of m/z -74 mu at m/z 790 and 820, respectively. These have been attributed as secondary fragments resulting from the loss of GlcNAc C-5 and C-6 through a $^{0,4}X$-type fragmentation. Contrarily to the C-3 substitution, no Z-type fragment ion resulting from the loss of the C-4 substitution (Fuc or Gal) was observed for either compound. However, C-type ion at m/z 259 resulting from the cleavage of Gal glycosidic bond was observed independently of its linkage position (C-3 or C-4) for both LNFII and LNFIII. Similar rules applied for the fragmentation of LNFI. Again, along Y- and B-types ions at m/z 660 and 463, LNFI showed a specific Z- and C-type couple of ions at m/z 690 and 433, that resulted from the release of terminal Fuc(α1-2)Gal disaccharide linked in C-3 position to the GlcNAc residue. Overall, data suggest that these experimental conditions permit to distinguish type 1 lactosamine disaccharide (Galβ1-3GlcNAc) from type 2 lactosamine (Galβ1-4GlcNAc) owing to the presence of an intense Z-ion resulting from the primary cleavage of the Gal-GlcNAc bond. It should be noted that the C-ion that originates from a cleavage at an identical location is not specific of a single linkage, for it is commonly observed for both lactosamine types in variable intensities and thus

Fig. 3



is not a reliable indicator (see following results). Similarly, the chemical lability of the Fuc bond results in its non specific cleavage and release as a Y ion irrespective of the nature of the linkage. As shown Fig. 1, this ion is observed as a minor ion at m/z 912 for the three isobars of Lacto-N-Fucopentaose.

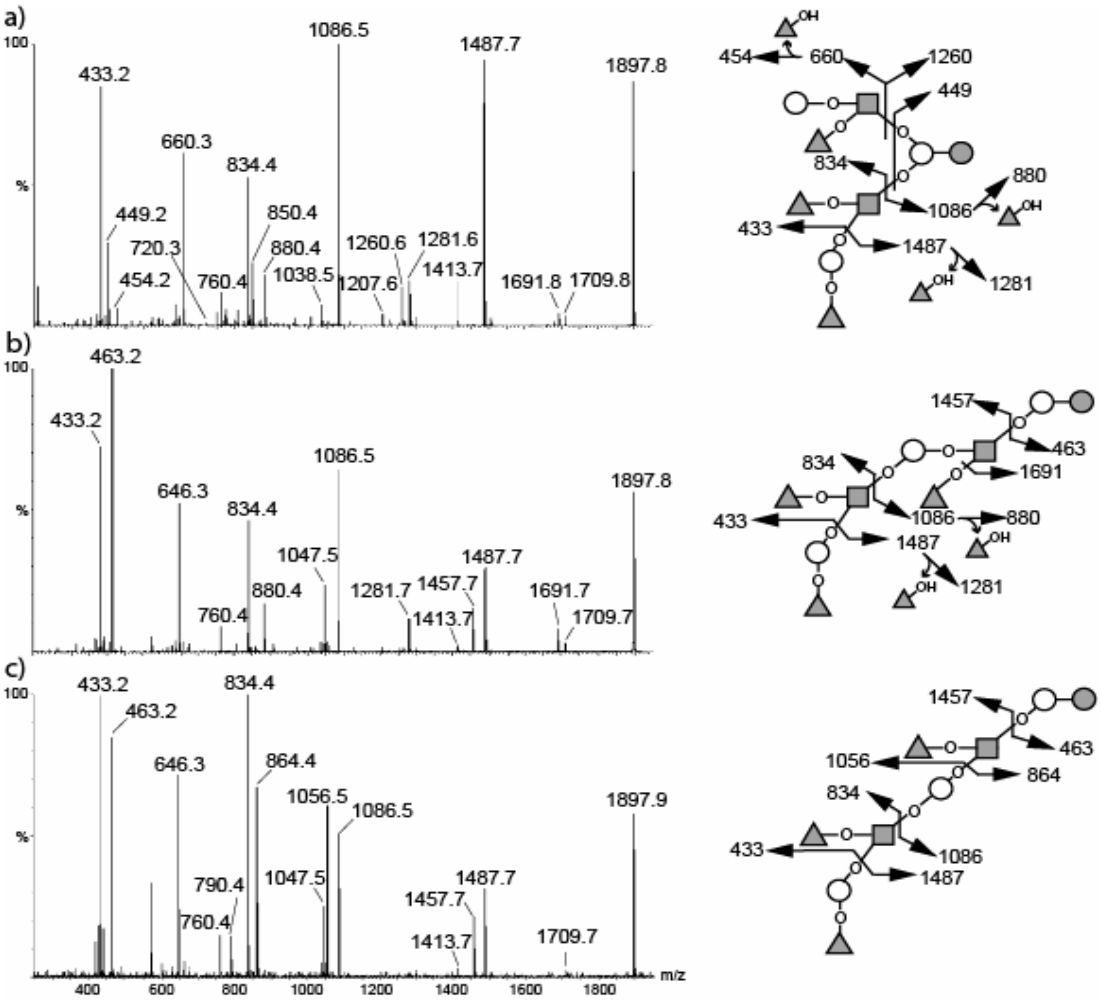
It is noteworthy that specific elimination of the C-3 substitution has already been observed for permethylated oligosaccharides in FAB-CID-MS/MS, but it exclusively occurred as secondary cleavage from the oxonium fragment, and never as primary fragment as observed with MALDI-CID-MS/MS (ref). Such fragmentation patterns could be obtained by CID-MS/MS from protonated adducts of LNFs on a ES-QTOF instrument. In this configuration, different isobars LNF-I to -III were distinguished owing to the presence of intense secondary fragments at m/z 228, 402 and 432 respectively generated from Y-ion at m/z 638 by the elimination of C-3 position (Fig. 2). On another hand, fragmentation of sodiated adducts on the same instrument generated fragmentation spectra almost identical to the ones obtained with MALDI-QTOF, demonstrating that the choice of parent ion adduct and not the instrument configuration conditioned the type of fragmentation (data not shown).

Identification of H/Lewis motifs. From the existence of a specific elimination process of the C-3 substitution of GlcNAc (see above), it is possible to easily and accurately predict the nature of fucosylated lactosamine motifs within a glycan, and so to distinguish between H-1, H-2, Le^x, Le^a, Le^x and Le^y determinants, solely on the base of CID-MS/MS profile. We will first exemplify on simple glycans the divers fragmentation pathways that can be used toward this purpose, and then demonstrate that the same principles apply for complex glycans.

Fig. 3 shows CID-MS/MS fragmentation patterns of four simple O-glycans, among a large panel of others tested, containing variations of fucosylated type-2 lactosamine motifs. As established above, fragmentation patterns of all compounds are dominated by B-ions at m/z 660 or 834 and by Y-ions at m/z 316, 520, 939 or 1113, that result from the cleavage of GlcNac linkages. All lactosamine disaccharides can be typified as type 2 according to the total absence of specific Z-ion at m/z $[M-236+Na]^+$ (loss of terminal Gal) or $[M-410+Na]^+$ (loss of terminal Fuc-Gal) that would result from the cleavage of Gal-GlcNac bond. However, as already pointed out, their non-specific C-ion counterparts at m/z 259 or 433 are systematically observed in variable relative intensities. The H-2 and Le^x isobar motifs are easily distinguished by the way they lose their respective fucose residue. Indeed, fragmentation pattern of H-2 containing glycan exhibits a single $[M-Fuc+Na]^+$ Y-ion at m/z 765 (M-188 mu), while Le^x containing glycan shows two $[M-Fuc+Na]^+$ Y and Z-ions at m/z 969 (M-188 mu) and 951 (M-206 mu), respectively (Fig. 3a and 3b). The generation of a non-specific

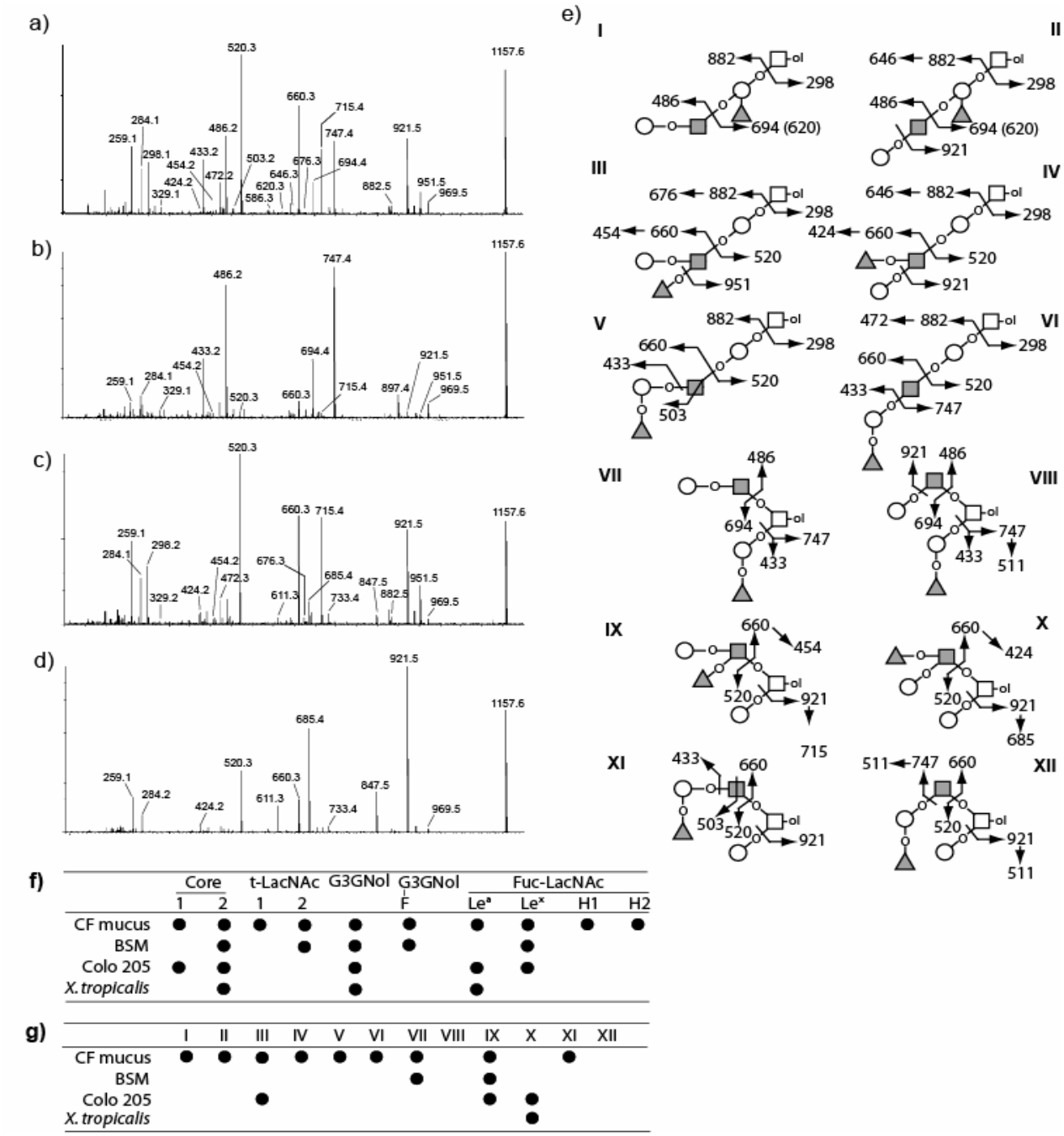
Y-ion at m/z M-188 is also observed in α(1,4)-linked fucose residue containing glycans (Fig. 1b). From these observations, we deduced that fucose linkage may be cleaved in two different ways. First, α(1,2)-, α(1,3)- and α(1,4)-linked fucose residues are released through a non specific cleavage that generates a low intensity Y-type ion at m/z M-188. On top of that, α(1,3)-linked fucose residues are released owing to a specific elimination that leads to the generation of a Z-type at m/z M-206. It is noteworthy that whenever both ions are observed, irrespective of their absolute intensities, Z-ion systematically exhibited a much higher relative intensity than Y-ion. Based on these assumptions, Le^y containing O-glycan (Fig. 3c) was characterized by the presence of both Z-ion at m/z 921 (M-206 mu) generated by the specific elimination of Fuc(α1,3) residue and Y-ion at m/z 939 (M-188 mu) generated by the loss of presumably one or the other Fuc(α1-3) and Fuc(α1-2) residues. Then, a third Le^y specific ion was observed at m/z 733 (M-206-188) which was attributed to the simultaneous loss of both Fuc(α1-3) and Fuc(α1-2) residues. Loss of fucose may also be observed as secondary fragment ions, although these usually display lower relative intensity. In accordance with previous data, one can observe from the primary B-ion fragment (mono fucosylated LacNAc) at m/z 660 either a single Y-type secondary fragment at m/z 472 (660 mu-188 mu) for the H-2 containing glycan (Fig. 3a), or both Y- and Z-type secondary fragments at m/z 472 and 454 (660 mu-206 mu) for the Le^x structure (Fig. 3b). Similarly for Le^y structure, from the primary fragments at m/z 834, 1095 and 1113 (Fig. 3c and 3d) both Y- and Z-type secondary fragments indicative of the presence of a Fuc(α1-3) residue are observed. Contrarily, these secondary fragmentation are not observed from the H-2 containing primary fragments. Finally, presence of the Gal(β1-4)GlcNac motif in a glycan may be further assessed owing to the presence of a very specific ion at m/z 503, attributed to a ^{3,5}A-ion resulting from internal ring fragmentation of GlcNAc residue. Indeed, this ion was exclusively observed in H-2 and Le^y containing glycans but never in Le^x (Fig. 3). Similarly, considering that this ion includes only the C-4, C-5 and C-6 of GlcNAc residue, it is never observed in type 1 LacNAc structures (Fig. 1a and 1b). This observation was further confirmed from analyses of a large panel of standard glycans (data not shown).

Fig. 4



In the prospect of studying complex mixture of isobaric glycans, we will now demonstrate that the above observations enable to distinguish between individual isobar glycans that present distinct combinations of Lewis determinants. As showed Fig. 4, three trifucosyllacto-N-hexaose isobars isolated from milk, arbitrarily labeled I, II and III, displayed very distinctive fragmentation patterns. Compounds II and III were identified as linear molecules owing to the presence of Y- and B-ions at m/z 463 and 1457, that resulted from the release of the intact tri-fucosylated di-LacNAc unit (Fig. 4b and 4c) The observation of Y- and B-ions at 1086 and 834 established the presence of a terminal di-fucosylated LacNAc (Le^b or Le^y) and consequently of an internal mono-fucosylated LacNAc (H-1, H-2, Le^a or Le^x). The presence of the last motif can also be tentatively observed owing to the intense secondary fragment at m/z 646. On the contrary to compounds II and III, no ion indicative of a linear di-LacNAc core was observed for compound I (Fig. 4a). Instead, two sets of Y-ions at m/z 660/834 and B-ions at m/z 1260/1086 established the presence of both terminal di-fucosylated and terminal mono-fucosylated LacNAc units, which suggested the occurrence of a branched glycan. Secondary fragment at m/z 449, corresponding to the release of the Lac core, presented a discrepancy of 28 mu with its fully permethylated counterpart, which directly demonstrates that the Lac core is di-substituted. Additionally, this was further confirmed by the presence of an $^{O,4}A$-ion at m/z 720 establishing that the Lac core is substituted by the mono-fucosylated LacNAc unit in C-6 position. In compound I spectrum, Z-ion at m/z 1691 (M-206) established the occurrence of Fuc(α1-3) residue and Y-ion at m/z 1487 (M-410) of terminal Fuc(α1-2)Gal(α1-3) motif. These data are compatible with a combination of either Le^x and Le^b determinants or H-1 and Le^y determinants. However, it can be observed that fragments ions at m/z 1086 and 660 generate Z-type secondary ions at m/z 880 and 454, through specific elimination of a Fuc(α1-3) residue, which is strongly in favor of the presence of a Le^x determinant. Accordingly, elimination of Fuc(α1-3) could not be observed from fragment ions at m/z 834 and 1260, which confirms the identity of the other branch as Le^b determinant. In addition, we repetitively observed in all Le^b containing glycans a complementary ion at m/z 760, that we tentatively attributed as a $^{O,4}X$-secondary ion originating from the 834 B-ion. For a reason unknown of us, this ion was never observed in any of the Le^y containing glycan that we studied, and so was subsequently used as a marker ion for the presence of Le^b. Through a similar reasoning, sequence of compounds II and III can be completely deciphered. In both compounds, combination of ions at m/z 1487, 834 and 760 clearly indicated the presence of a terminal Le^b motif.

Fig. 5



In compound II, $[M-Fuc+Na]^+$ Z-ion at 1691 established the presence of a Fuc(α1-3) residue, which suggests the existence of an internal Le^x motif. This residue was directly located on the internal LacNac unit owing to the Z-type secondary ion at m/z 880, generated by the elimination of Fuc(α1-3) from Y-ion at m/z 1086. In accordance with the presence of an internal Le^x motif, no cleavage of its Gal-GlcNAc bond occurred. On the contrary for compound III, two intense C- and Y-ion at m/z 1056 and 864 established the nature of the internal LacNAc as type-1. Furthermore, they permitted to locate the Fuc residue on the GlcNac residue, which establishes the presence of internal Le^a motif. Accordingly with the absence of Fuc(α1-3) residue in the molecule and contrarily to compound II, no Z-ion resulting from the loss of fucose can be observed neither from the molecular ion, nor from the Z-ion at m/z 1086.

Study of glycans in mixture

Proteomic and glycomic-type studies are generally not compatible with extensive purification of glycan moiety following its release. Indeed, low quantities of samples generally available prevent the use of any multi-dimension chromatographic purification process that is requires for isolation of individual glycan isobar. So, to be fully effective, mass-spectrometry based method of study of carbohydrate structure should be usable for complex mixture of glycans, and be able to distinguish, at least partially, between closely related isobars. In order to demonstrate that the CID-MS/MS of permethylated glycans enables such study, we will resolve the different isobars present in a single ion from three different sources. To do so, O-glycans have been released and purified from bovine submaxillary glands mucins (BSM), purified human cystic fibrosis (CF) mucus and total extract from colo 205 carcinogenous cell culture. As expected, after permethylation, profiling of total O-glycans by MALDI-MS showed an extensive structural variability within each sample (data not shown). As a mean to study isobaric distribution of glycans, we targeted the $[M+Na]^+$ ion at m/z 1157 (1 HexNAc-ol, 1 HexNAc, 2 Hex and 1 deHex) which was fragmented in CID-MS/MS for each sample. From the known biosynthetic pathways of mucin O-glycans, we deduced the presence of twelve possible isobars which fragmentation patterns were inferred according to the principles aforementioned (Fig. 5e). To our knowledge, only nine out of the twelve possible isobars have been observed so far in mammals (compounds **I** to **VII**, **IX** and **XI**), type 1 LacNAc chain being never observed directly attached to C-6 position of reducing GalNAc residue (as in compounds **VIII**, **X** and **XII**). CID-MS/MS fragmentation pattern models were also directly assessed for five compounds out of the twelve possible by using pure isobars purified from several animal species, including all known core 2 (compounds **VII**, **IX**, **X**, **XI** and **XII**) and a core 1

containing glycans (compound **III**). Models did not show any discrepancy from actual spectra, which contributed to validate the method.

As will be demonstrated, analysis of complex mixtures of glycans differs significantly from similar analysis of pure glycans due to overlaps in generated fragment ions. It is a multi-step analysis which prime concern is to identify precise motifs within the mixture, i.e. what type of core, LacNAc disaccharides and fucosylated motifs are present. From a functional point of view, the expression of such motifs on glycoconjugates clearly appears as the most relevant information to seek. In a second step, precise arrangement of motifs within each glycans will be investigate, and by doing so will permit to definitively establish the presence or the absence of each isobar. This step relies more heavily on secondary fragmentations for it enables to distinguish more efficiently isobars.

In the chosen example, CID MS/MS spectra of the three O-glycans samples showed completely different patterns, which established that they contained distinct sets of isobaric structures (Fig. 5a, 5b and 5c). From the general aspects of spectra, higher complexity of fragment ions pattern suggests that CF mucus contains a higher number of compounds than two other samples. In a first step we will examine the presence of individual structural motifs present in each mixture, which results are summarize on Fig. 5 f.

CF and colo 205 samples exhibited both type 1 and 2 core structures according to the presence of secondary fragment ion at m/z 284 and Z/C-type ion pair at m/z 298/882. Contrarily, BSM exclusively showed type 2 core specific fragment at m/z 284, establishing the total absence of type 1 core containing glycans and so restricting isobaric variability to compounds **VII** to **XII**. Terminal Gal(β1-3) as well as Gal(β1-3)GalNac-ol motif were observed in all three samples as deduced from the presence of Z-type ions at m/z 921 and Y-type ion at m/z 520. Terminal Gal(β1-3) residue originates either from terminal type 1 LacNAc or Gal(β1-3)GalNac-ol motif. Their fucosylated counterparts, terminal Fuc(α1-2)Gal(β1-3) (Z ion at m/z 747) as well as Fuc(α1-2)Gal(β1-3)GalNAc-ol (Z ion at m/z 694) showed a more restricted distribution and were only detected in CF and BSM. Accordingly, owing to Z/C ion pair at m/z 694 and 486, the presence of un-fucosylated terminal LacNAc unit was attested only in CF and BSM samples, colo 205 showing none. Nature of LacNAc unit could also be attributed either to type 2 according to ^{3,5}A-ions at m/z 329 in both CF and BSM, or to type 1 according to secondary ion at m/z 646, exclusively in CF. Although fucosylated LacNAc trisaccharides were observed in all three samples according to Y/B ion pair at m/z 520/660, their isomers (H1, H2, Le^x and Le^a) were differently distributed. Presence of Le^x was easily assessed in all three samples owing to the elimination of 1,3-linked Fuc residue at

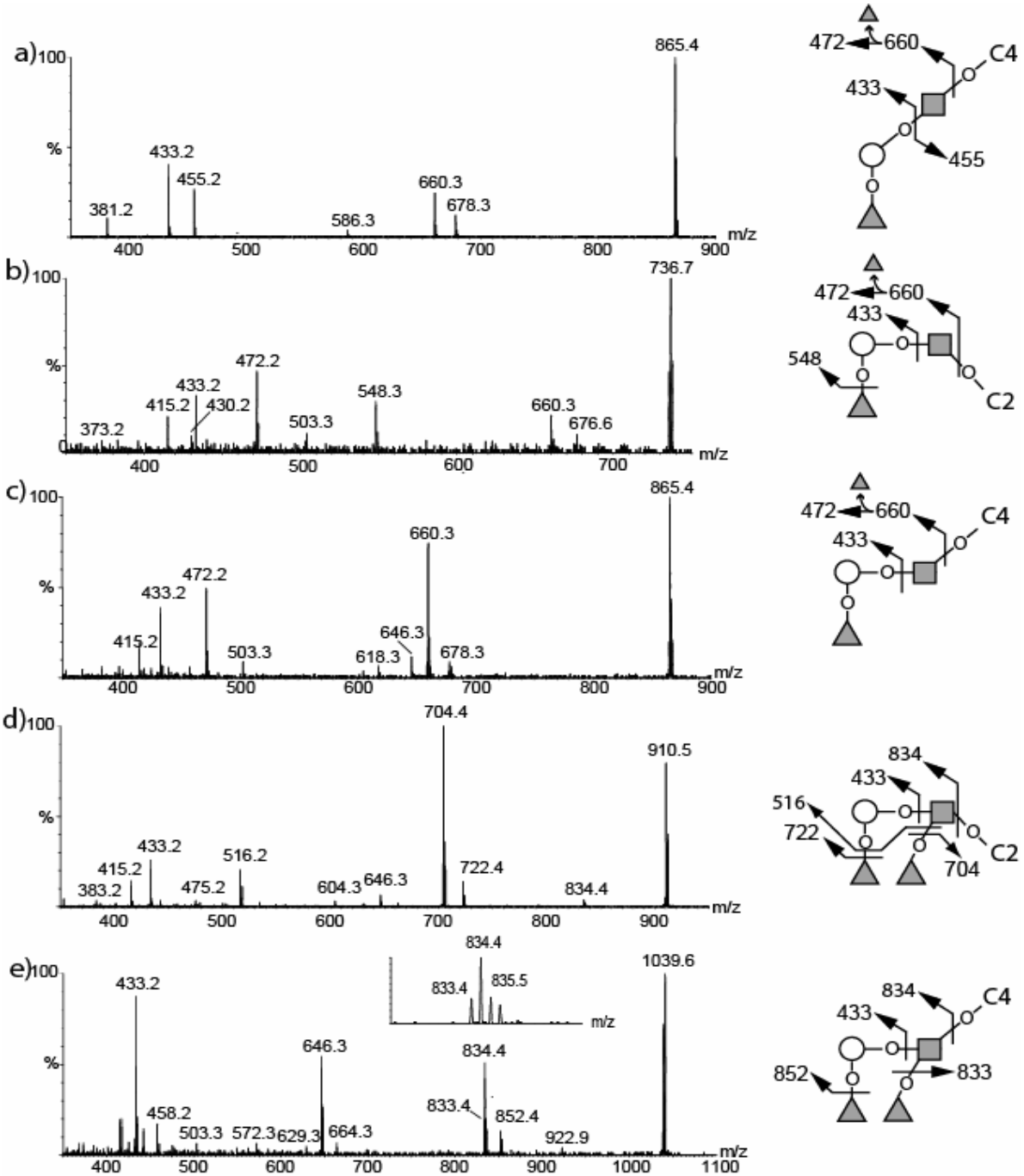
m/z 951. However, owing to secondary fragmentation, Le^x could be differently located on core 2 glycans (m/z 715) in all three sample but also on core 1 glycan in both CF and colo 205 (m/z 676), but not in BSM, in accordance with absence of core 1 glycan in this sample. Because the elimination of 1,3-linked Gal is not specific for Le^a, its presence was assessed in CF and colo 205 on the basis of secondary fragmentation at m/z 424. Its presence was further pinpointed in core 1 glycan by secondary ion at m/z 646 in CF and colo 205 as well as in core 2 glycan by secondary ion at m/z 685 exclusively in colo 205. Terminal H2 (^{3,5}A-ions at m/z 503) showed also a distribution restricted to CF sample, but absence of further fragmentation involving this motif does not permit to localize it definitively on either core 1, 2 or both. Then, presence of H1 on core 1 glycan was exclusively observed in CF, owing to the conjunction of Z ion at /z 921 and secondary ion at m/z 472. As shown in Fig. 5f, systematic search of individual structural motifs demonstrated a wide heterogeneity between samples; whereas CF includes all screened motifs, BSM and colo 205 only include subsets.

In a second step, we tried to assign to each mixture its final isobaric composition. For CF, fine analysis of secondary fragmentations demonstrated that some of the isobars were actually missing, despite that all motifs were individually identified in the isobaric mixture. In particular, none of the specific fragments for compounds **VIII**, **X** and **XII** (m/z 685 and 511) were identified in the corresponding spectrum. This result is consistent with the known biosynthetic pathway of O-glycans. Indeed, to our knowledge, type 1 LacNAc chain directly attached to C6 position of GalNAc has never been observed so far in Man, for it is considered as a signal of termination for upper chain extension. As a general rule, non fucosylated type 2 LacNAc and H-2 motifs are difficult to locate because they generate fewer distinctive secondary fragments. So, it is not possible to definitively discriminate between compound **V** and **XI**, although the presence of both compounds is highly probable. On the contrary, as deduced from the presence of fewer individual motifs, BSM appeared as a much simpler mixture of isobars. In particular, association of identified motifs only enables the generation of compounds **VI** and **IX**. Although mass spectrometry is not considered as a quantitative method, comparison of equivalent fragments for both components (486/660; 694/520; 747/921) clearly showed that compound **VII** was the major isobar of the mixture, and permitted to evaluate the **VII** to **IX** ratio to about 5 to 1. Then, colo 205 showed an isobaric distribution altogether very different from both previous samples. Association of types 1/2 cores with Le^a/Le^x gives rise to four potential compounds **III**, **IV**, **IX** and **X**, out of which compound **IV** was not actually observed, as deduced from the absence of the secondary fragment ion at m/z 646. Absence of the Le^a in core 1 containing glycan indicated the existence of compound **X** in colo 205. Indeed, a set of two ions at m/z 685 and 611 that did not occurred in any of the two other samples was

observed from colo 205. These two were tentatively attributed as Z-type and its ^{0,4}X-type associated secondary fragmentations from Z-ion at m/z 921, respectively. These ions were compatible with the presence of a Le^a motif linked to the C-6 position of the GalNAc-ol residue of a core 2 O-glycan. As aforementioned, such a glycan has never been characterized in Man. In animal kingdom, Le^a motif has only been observed in the primate lineage on both glycolipids and glycoproteins, but never on O-glycans from the C-6 position of GalNAc-ol. Recently, by a combination of preparative HPLC separation and NMR analyses, we have isolated and characterized for the first time, such a compound from the oviducal secretions of the amphibian model *Xenopus tropicalis*, demonstrating the emergence of an independent (α1-4)fucosyltransferase activity in amphibians (Guérardel *et al.*, 2003). In order to confirm that the couple of unknown ions observed in Colo 205 sample originated from the fragmentation of a Le^a chain linked to C-6 position of GalNAc-ol, purified O-glycan from *X. tropicalis* was permethylated and fragmented by CID-MS/MS in identical conditions. As hypothesized, intense ions at m/z 685 and 611 are observed from fragmentation pattern of this glycan, confirming that their originated from the successive elimination of two (1-3)Gal residues (Fig. 5d). Accordingly, these fragments were not observed from equivalent glycans containing either H-2 or Le^x determinants (data not shown or sup data). In summary, as shown in Fig. 5g, CF mucus, BSM and colo 205 exhibited very different patterns of glycosylation.

In conclusion CID-MS/MS analysis enabled to decipher the major structural features of individual O-glycans from complex mixtures. It permitted, without prior extensive glycan separation, to easily distinguish between different samples according to their isobaric variabilities, from a simple mixture of two or three isomers to a very complex mixture of up to nine components, each sample containing distinctive sets of glycans. Furthermore, it demonstrated its capacity to identify a novel structure within a mixture of known structures. Indeed, for the first time, a Le^a motif directly linked to the reducing GalNAc residue was identified in a human carcinogenous cell culture O-glycan, suggesting a profound modification of glycan biosynthesis pathways in colonic cancer cells.

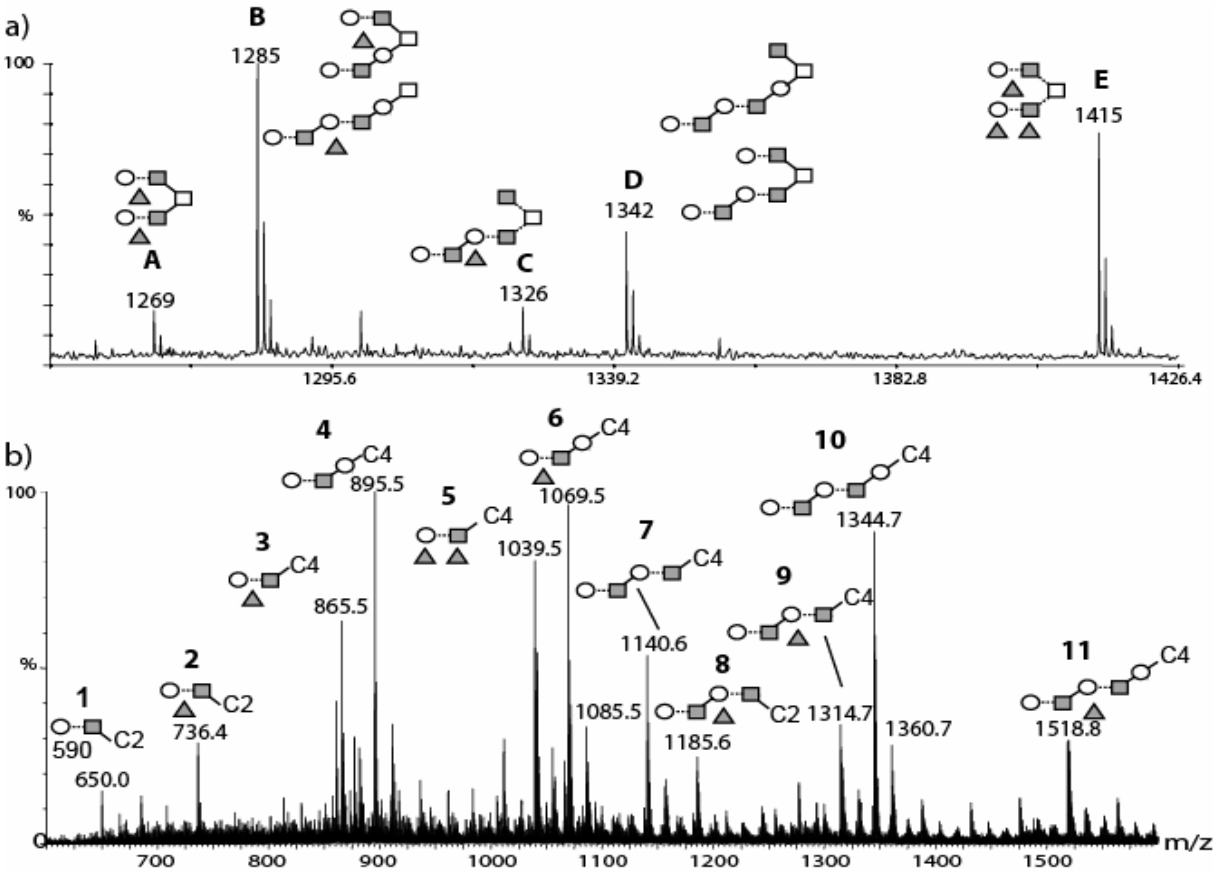
Fig. 6



Study of glycans after periodic oxidation

Mild periodic oxidation has already been used as an effective way to discriminate C-3 from C-6 linked branches of O-glycans by selectively cleaving C4-C5 bond of GalNAc-ol. Oxidation of a single reduced O-glycan generates after reduction two fragments in which GalNAc-ol residue is replaced either by a CH_2OH-CHNAc-CHOH-CH_2OH group (C4) for C-3 lower branch or by a $[CH2OH]_2$ group (C2) for C-6 upper branch. We will demonstrate that the aforementioned principles of analysis can be applied to oxidize glycans and give an additional mean to study mixture of complex glycans. In order to validate the method, a set of purified mucin O-glycans were oxidized, reduced, permethylated and subjected to CID-MS/MS in MALDI Q-TOF. As shown on Fig. 6, C2 and C4 containing product substituted by identical glycan moieties are easily distinguished in MS owing to the apparent molecular size of their parent ions, differing from 129 mu. These oxidation products show very similar fragmentation patterns, irrespective of the nature of the aglycon group, C2 or C4. Fragmentation patterns of glycan moieties of permethylated oxidized products are also similar to those from intact permethylated glycans as demonstrated by comparing Fig. 6a and 1a, Fig. 6b/c and 3a, Fig. 6d/e and 3c. However, care should be taken when analysing Le^x/Le^y containing C4 products. Indeed, release of the C4 fragment by a B-type cleavage results in the generation of a [M-205] ion very close to the [M-206] fragment ion generated by the specific elimination of 1,3 linked Fuc residue. As shown in Fig. 6e, both ions are easily distinguished on a high resolution spectrum, and closeness of signals never induced any ambiguity of interpretation in all the samples we analysed. Similarly, C-type ion at [M-189] and Z-type ions at [M-188] that result from the release of C4 fragment and the non specific release of Fuc residue, respectively, are very close from each other, but easily distinguished. Nevertheless, in order to ease interpretation, we tried to omit the reduction step prior to permethylation as an alternative process, which facilitated distinction of these fragments. In this case, release of C4 fragment generates a B-type ion at [M-189] which does not overlap anymore [M-206] ion, but shed in its isotopic cluster the [M-188] fragment ion. On the other hand, reduction of oxidized fragments with deuterium prevented any distinction between ions generated by release of Fuc residue and C4 fragments, as expected (data not shown).

Fig. 7



In order to evaluate the usefulness of periodate oxidation for the study of complex mixture of glycans, we used as starting material a partially purified fraction of O-glycans. Total neutral O-glycan fraction prepared from CF mucus was separated in HPLC on an amine bond column eluted by a 90 min gradient of H_2O in ACN. Ninety fractions were automatically collected and their composition was individually assessed by MALDI-TOF MS. This permitted to follow in the ninety fractions the distribution of sixty $[M+Na]^+$ molecular ions ranging from m/z at 611 (Hex1HexNac2-ol) to m/z at 2745 (Fuc6Hex5HexNAc4-ol) (Supplemental data). All ions were then fragmented in CID-MS/MS in order to partially assess the isobaric distribution of compounds in each fraction. Despite its low capacity to discriminate between isobars in mixture, MS/MS analysis of native O-glycans demonstrated that HPLC partially separated isobaric structures of low molecular weight compounds. On the other hand, this was difficult or impossible to assess for higher molecular weight compounds (above 1500 Da) due to decrease of the separation resolution and increase of the isobaric complexity along with increasing size.

We have arbitrarily chosen fraction **52** as a model of study. MALDI-MS analysis of this fraction showed five ions with m/z values ranging from 1269 to 1415 mass units (Fig. 7a). In a first step of analysis, CID-MS/MS analysis of native compounds based on already published data permitted to evaluate their general structure, including nature of cores, length of branches and distribution of eventual fucose residues. Results from these analyses are depicted in Fig. 7a. However, exact positioning of fucose, nature and distribution of LacNAc units could not be established by this technique as already discussed. Nevertheless, systematic fragmentation of these ions in surrounding fractions established that only a subset of isobaric structures was present in fraction **52** owing to partial isobaric separation after HPLC. In particular, for the ion at m/z 1269, extended core 3 containing isobars were identified in fractions **45** to **47** whereas $Le^{y/b}$ containing isobars in fraction **48** (data not shown). Fraction **52** exclusively contained core 4, mono-fucosylated LacNAc substituted O-glycans (Fig. 7a). Similarly for ion at m/z 1285, fractions **49**-**51** exclusively contained core 1 isobars, whereas fraction **52** contained both core 1 and core 2.

In a second step of analysis, fraction **52** was subjected to mild periodic oxidation, $NaBH_4$ reduction and permethylation. Eleven ions labeled from **1** to **11** and ranging from m/z 562 (Hex1HexNAc1-C2) to m/z 1518 (Fuc1Hex3HexNAc2-C4) were observed in MALDI-MS (Fig. 7b) and subsequently fragmented in CID-MS/MS. Use of aforementioned procedures of analysis permitted to identify the major isobaric sequences in each of the eleven compounds detected as summarized in Table 1. It is noteworthy that C2 and C4-containing fragments presenting identical monosaccharide compositions systematically exhibited distinct isobaric distributions. In accordance

Table 1: Summary of oxidation generated fragment identified by MS/MS sequencing

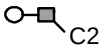
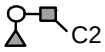
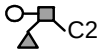
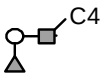
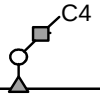
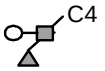
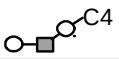
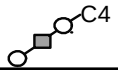
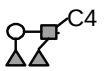
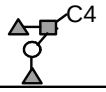
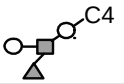
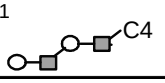
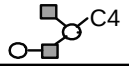
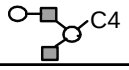
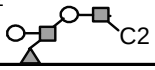
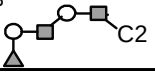
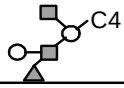
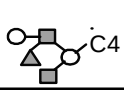
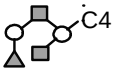
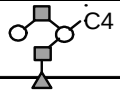
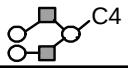
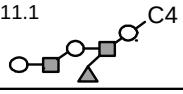
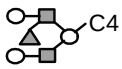
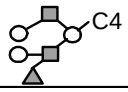
Cpds	m/z	Isobars	Fragments
1	**562**	1.1 	486
2	**736**	2.1 	660;548;503; 472;433
		2.2 	660; 530;259
3	**865**	3.1 	677;660;503; 472;433
		3.2 	660;586;455;433
		3.3 	660;659;259
4	**895**	4.1 	708;486;432;329
		4.2 	708;659;486;432
5	**1039**	5.1 	834;833;646; 628;503;458
		5.2 	834;760;629;572
6	**1069**	6.1 	660;454;432; 882;863;
7	**1140**	7.1 	935;677;486;472
		7.2 	953;881;490; 486;416;342
		7.3 	953;881;694;677; 620;546;486
8	**1185**	8.1 	979;903;722; 516;486
		8.2 	979;903;660;454; 660;548;472
		8.3 	433;503;660; 548;472
9	**1314**	9.1 	868;490;416;342
		9.2 	868;794;588; 720;514
		9.3 	904;694;620;546; 490;458;384
		9.4 	1078;694;458;620; 384;546;664
10	**1344**	10.1 	1157;881;694; 620;458;384
11	**1518**	11.1 	1332;1110;1055; 849;646;432
		11.2 	1332;1313;868;794; 588;720;514;418
		11.3 	1332;694;620; 546;418;384

Table 2: Possible combinations of oxidation generated fragments for each intact molecule

Cpds	Combinations
A	2+3
B	1+6; 2+4; 11
C	8; 9
D	1+7; 10
E	2+5

with previous observations, no type 1 LacNAc chain was identified directly attached to C2 fragments (compounds **2** and **8**), whereas it was on C4 fragment (compounds **3** and **5**). As a result only Le^x and H2 motifs were identified in product **2** (ion at m/z 736), whereas Le^x, H1 and H2 were identified in product **3** (ion at m/z 865) (data not shown). Indeed, fragmentation pattern of coumpound **2** showed typical fragment ions of terminal H2 (m/z at 433, 472, 548), as observed in standard molecule (Fig. 6a), as well as additional Z and B ions (m/z at 530 and 259) that clearly indicate the presence of Le^x containing isobar. However, no fragment indicative of H1 (Z ion at m/z 410) or Le^a (Z ion at m/z 500) epitopes could be observed, establishing the complete absence of such motifs. Contrarily, in addition to H2 and Le^x, H1 motif is clearly observed in compound **3** owing to the presence of Z ion at m/z 455.

As shown Fig. 8, a new set of characteristic ions deriving from secondary Y/C and internal fragmentations was used to identify branching patterns of oligo-lactosaminyl sequences. Fragmentation of compound **10** generated an intense secondary ion at m/z 418, which in conjunction with the absence of B/Y ions pair at m/z 935/432, established the presence of a branched sequence to the exclusion of its linear counterpart. This observation was confirmed owing to the set of ions at m/z 694 (Y/C), 620 (^{0,2}A) and 546 (^{4,0}A) that demonstrated the presence of a Gal residue di-substituted in 3 and 6 positions by a LacNAc motif. Elimination of a Gal residue from pseudomolecular ion, ^{0,2}A ion, ^{4,0}A ions and Y ion at 881 generates secondary ions at m/z 1108, 458, 384 and 645, respectively, which then established the presence of terminal type 1 LacNAc chain. In parallel, presence of terminal type 2 LacNAc was attested by the presence of ion at m/z 329. Although we established the presence of the two LAcNAc types, it was not possible to determine which of the four possible isobars were actually present in the mixture. On this basis, we attributed to product **10** a single isobaric structure presenting both types of LAcNAc, although presence of others is most probable. A similar strategy of analysis permitted to identify four different isobaric branched structures for product **9** (Fig. 8b). Distinction of different isobars is easier compared to the previous ion owing to the dissymmetry of the molecule. However, exact positioning of fucose is still difficult for isobaric structure 9.1 considering its similarity with structure 9.2. It is noteworthy that compounds **9** and **10** exclusively contained branched structures and no ‘core 1 type’ linear ones as revealed by the complete absence of Y ion at 432. In contrast, as already observed for products **2** and **3**, product **8** exhibited an altogether very different structural variability than its C2 counterpart products **9**. In particular, none of the specific ions for branched sequences were observed on fragmentation spectrum of product **8** (supplemental data). Instead, an intense B-type ion at m/z 979 was indicative of the cleavage of GlcNAc-C2 bond, altogether establishing the presence of ‘core 2 or 4 type’ linear fucosylated di-LacNAc sequences in place of branched ones.

Fig. 8

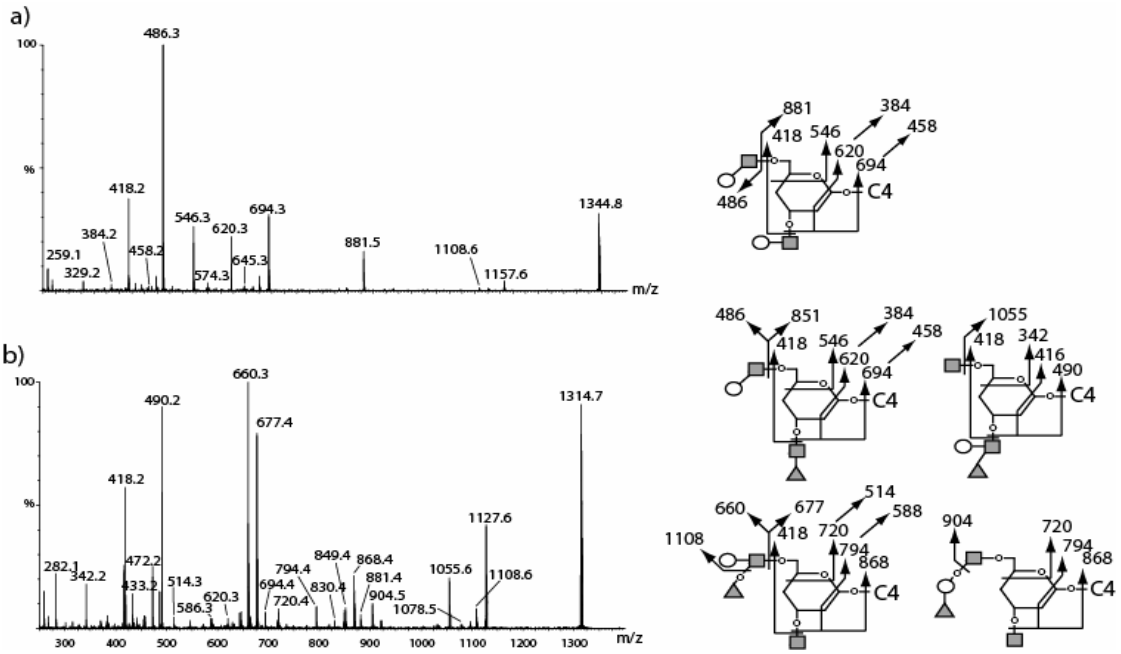


Table 3: Summary of identified O-glycan structures from fraction **52**

A	2+3
B	1+6 2+4 11
C	8 9
D	10 1+7
E	2+5

Three distinct linear isobars differing according to the fucose position could be distinguished within the mixture; two with Le^x motifs and one with terminal H2 motif (Table 1). Absence of Z-type [M-236] and [M-410] fragment ions established that it did not contain terminal Le^a or H1 motifs. Then, compound **11** was differing from three previous ones by the fact that it contained both linear (11.1) as demonstrated by the presence of B/Y ion pair at m/z 432/1109, and branched (11.2 and 11.3) isobars as demonstrated by B ion at m/z 418 and above described Y/C, $^{0,2}A$ and $^{4,0}A$ sets of ions. Similarly, compound **7** appeared as a mixture of linear 'core 3 or 4 type' isobar (7.1), as demonstrated by the presence of an intense Y ion at m/z 935 and branched isobars (7.2 and 7.3). In conclusion, exhaustive MS/MS analysis of oxydized HPLC fraction **52** permitted to definitely identify 25 different compounds. As already pointed out, isobaric variability is most probably underestimated due to the difficulty to distinguish very close isobaric structures, particularly for those which contain non fucosylated type 2 LacNAc motifs.

In a third step of analysis, independent sequences of C-3 and C-6 branches were combined according to the results of fragmentation analysis of intact glycans. As summarized in table 2, each compound (**A** to **E**) may be the result of one to three different combinations of oxidation products (1.1 to 11.3). It is noteworthy that most of the oxidation products are specific to a single final compound. Indeed, only oxidized products **1** and **2** can possibly occur in more than a single O-glycan (Table 2). On these bases, combination of the 25 identified oxidation products generated up to 29 distinct oligosaccharides from the five MS signals observed in the studied mixture (Table 3).

Références

- Ballou CE, Ballou L, Ball G. *Schizosaccharomyces pombe* glycosylation mutant with altered cell surface properties. Proc Natl Acad Sci U S A. 1994 91:9327-31

- Ballou CE. 1990 Isolation, characterization, and properties of *Saccharomyces cerevisiae* mnn mutants with nonconditional protein glycosylation defects. Methods Enzymol. 185:440-70.

- Banerjee S, Vishwanath P, Cui J, Kelleher DJ, Gilmore R, Robbins PW, Samuelson J. The evolution of N-glycan-dependent endoplasmic reticulum quality control factors for glycoprotein folding and degradation. Proc Natl Acad Sci U S A. 2007 104:11676-81

- Bardor M, Nguyen DH, Diaz S, Varki A. Mechanism of uptake and incorporation of the non-human sialic acid N-glycolylneuraminic acid into human cells. J Biol Chem. 2005 280:4228-37

- Chaban B, Voisin S, Kelly J, Logan SM, Jarrell KF. Identification of genes involved in the biosynthesis and attachment of Methanococcus voltae N-linked glycans: insight into N-linked glycosylation pathways in Archaea. Molecular Microbiology 2006 61**:** 259-268

- Chatterjee, D. and Khoo, K.H. Mycobacterial lipoarabinomannan: an extraordinary lipoheteroglycan with profound physiological effects. Glycobiology 1998 8: 113-120

- Chatterjee, D., Bozic, C., McNeil, M. and Brennan, P. Structural features of the arabinan component of the lipoarabinomannan of *Mycobacterium tuberculosis*. J. Biol. Chem. 1991 266: 9652-9660

- Chen, S., Zhou, S., Sarkar, M., Spence, A.M. and Schachter, H. Expression of three *Caenorhabditis elegans* N-acetylglucosaminyltransferase I genes during development. J. Biol. Chem. 1999 274:288-297

- Chiba A, Matsumura K, Yamada H, Inazu T, Shimizu T, Kusunoki S, Kanazawa I, Kobata A, Endo T. Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. J Biol Chem. 1997 272:2156-62.

- Cipollo JF, Awad AM, Costello CE, Hirschberg CB. *srf-3*, a mutant of *Caenorhabditis elegans*, resistant to bacterial infection and to biofilm binding, is deficient in glycoconjugates. J Biol Chem. 2004 279:52893-903.

- Costache, M., Cailleau, A., Fernandez-Mateos, P., Oriol, R. and Mollicone, R. Advances in molecular genetics of α-2 and α-3/4-fucosyltransferases. Transfus. Clin. Biol. 1997 4:367-382

- De Bose-Boyd, R.A., Nyame, A.K. and Cummings, R.D. (1998) Molecular cloning and characterization of an alpha 1,3 fucosyltransferase, CEFT-1, from *Caenorhabditis elegans. Glycobiology* 8:905-917

- Dennis, J.W., Granovsky, M. and Warren, C.E. Protein glycosylation in development and disease. BioEssays 1999 21:412-421
- Dobos, K. M., K. H. Khoo, K. M. Swiderek, P. J. Brennan, and J. T. Belisle. 1996. Definition of the full extent of glycosylation of the 45-kilodalton glycoprotein of Mycobacterium tuberculosis. J. Bacteriol. 178:2498–2506.

- Dobos, K. M., K. Swiderek, K. H. Khoo, P. J. Brennan, and J. T. Belisle. 1995. Evidence for glycosylation sites on the 45-kilodalton glycoprotein of Mycobacterium tuberculosis. Infect. Immun. 63:2846–2853.

- Dupuy, F., Germot. A., Marenda, M., Orilo, R., Blancher, A., Julien, R. and Maftah, A. Alpha1,4-fucosyltransferase activity: a significant function in the primate lineage has appeared twice independantly. Mol. Biol. Evol. 2002 19:815-824

- Dupuy, F., Petit, J.M., Mollicone, R., Oriol, R., Julien, R. and Mafta, A. A single amino acid in the hypervariable stem domain of vertebrate a1,3/1,4-fucosyltransferases determines the type 1/type 2 transfer. J. Biol. Chem. 1999 274:12257-12262

- Ecker M, Mrsa V, Hagen I, Deutzmann R, Strahl S, Tanner W. O-mannosylation precedes and potentially controls the N-glycosylation of a yeast cell wall glycoprotein. EMBO Rep. 2003 4:628-32.

- Eggens, I., Fenderson, B., Toyokuni, T., Dean, B., Stroud, M. and Hakomori, S. Specific interarction between Le^x and Le^x determinants. J. Biol. Chem. 1989 264:9476-9484

- Engel JC, Parodi AJ. *Trypanosoma cruzi* cells undergo an alteration in protein N-glycosylation upon differentiation. J Biol Chem. 1985 260:10105-10.

- Freeze HH, Aebi M. Altered glycan structures: the molecular basis of congenital disorders of glycosylation, Curr Opin Struct Biol. 2005 15:490-8.

- Fukazawa Y, Kagaya K, Shinoda T. Cell wall polysaccharides of pathogenic yeasts. Curr Top Med Mycol. 1995 6:189-219

- Funk VA, Thomas-Oates JE, Kielland SL, Bates PA, Olafson RW. A unique, terminally glucosylated oligosaccharide is a common feature on *Leishmania* cell surfaces. Mol Biochem Parasitol. 1997 84:33-48.

- Garbe,T., Harris,D., Vordermeir,M., Lathigra,R., Ivanyi,J. and Young,D. Expression of the *Mycobacterium tuberculosis* 19-kilodalton antigen in *Mycobacterium smegmatis*: immunological analysis and evidence of glycosylation. Infect. Immun. 199361: 260-267.

- Gemmill TR, Trimble RB. Schizosaccharomyces pombe produces novel pyruvate-containing N-linked oligosaccharides. J Biol Chem. 1996 271:25945-9.

- Gerdt, S., Dennis, R.D., Borgonie, G., Schnabel, R. and Geyer, R. Isolation, characterization and immunolocalization of phosphorylcholine-substituted glycolipids in developmental stages of *Caenorhabditis elegans.* Eur. J. Biochem. 1999 266:952-963

- Gerdt, S., Lochnit, G., Dennis, R.D. and Geyer :Isolation and structural analysis of three neutral glycosphingolipids from a mixed population of *Caenorhabditis elegans* (Nematoda:Rhabditida). Glycobiology 1997 7:265-275

- Gilleron M, Jackson M, Nigou J, Puzo, G Structure, Biosynthesis, and activities of the phosphatidyl-myo-inositol-based lipoglycans. In The Mycobacterial Cell Envelope, Ed. M. Daffé and J.M. Reyrat, ASM Press, Washington.

- Gilleron M, Nigou J, Nicolle D, Quesniaux V, Puzo G. The acylation state of mycobacterial lipomannans modulates innate immunity response through toll-like receptor 2. Chem Biol. 2006 13:39-47.

- Gilleron, M., Bala, L., Brando, T., Vercellone, A. & Puzo, G. *Mycobacterium tuberculosis* parietal and cellular lipoarabinomannans. J. Biol. Chem. 2000 275, 677-684.

- Girrbach V, Strahl S. Members of the evolutionarily conserved PMT family of protein O-mannosyltransferases form distinct protein complexes among themselves. J Biol Chem. 2003 278:12554-62.

- Hagen, K.G., Tetaert, D., Hagen, F.K., Richet, C., Beres, T.M., Gagnon, J., Balys, M.M., Van Wuyckhuyse, B., Bedi, G.S., Degand, P. and Tabak, L. Characterization of a UDP-GalNAc:Polypeptide *N*-Acetylgalactosaminyltransferase that displays glycopeptide *N*-Acetylgalactosaminyltransferase activity. J. Biol. Chem. 1999 274:27867-27874

- Helenius A, Aebi M Roles of N-linked glycans in the endoplasmic reticulum, Annu Rev Biochem. 2004 73:1019-49.

- Herrmann JL, O'Gaora P, Gallagher A, Thole JE, Young DB. Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kDa antigen from *Mycobacterium tuberculosis.* EMBO J. 1996 15:3547-54.

- Horn C, Namane A, Pescher P, Rivière M, Romain F, Puzo G, Bârzu O, Marchal G. Decreased capacity of recombinant 45/47-kDa molecules (Apa) of Mycobacterium tuberculosis to stimulate T lymphocyte responses related to changes in their mannosylation pattern. J Biol Chem. 1999 274:32023-30.

- Inoue, S. and Inoue, Y. Fish glycoproteins. *In Glycoprotein II*. 1997 J. Montreuil, J.F.G. Vliegenthart and H. Schachter (Eds), Elsevier Science B.V.:143-161

- Joe M, Sun D, Taha H, Completo GC, Croudace JE, Lammas DA, Besra GS, Lowary TL. The 5-deoxy-5-methylthio-xylofuranose residue in mycobacterial lipoarabinomannan. absolute stereochemistry, linkage position, conformation, and immunomodulatory activity. J Am Chem Soc. 2006 128:5059-72.

- Ju T, Brewer K, D'Souza A, Cummings RD, Canfield WM. Cloning and expression of human core 1 beta1,3-galactosyltransferase. J Biol Chem. 2002 277:178-86

- Ju T, Zheng Q, Cummings RD. Identification of core 1 O-glycan T-synthase from *Caenorhabditis elegans*. Glycobiology. 2006 16:947-58.

- Jurado LA, Coloma A, Cruces J. Identification of a human homolog of the Drosophila rotated abdomen gene (POMT1) encoding a putative protein O-mannosyl-transferase, and assignment to human chromosome 9q34.1. Genomics. 1999 58:171-80.

- Kawar ZS, Van Die I, Cummings RD. Molecular cloning and enzymatic characterization of a UDP-GalNAc:GlcNAc(beta)-R beta1,4-N-acetylgalactosaminyltransferase from Caenorhabditis elegans. J Biol Chem. 2002 277:34924-32

- Kelly J, Jarrell H, Millar L, Tessier L, Fiori LM, Lau PC, Allan B, Szymanski CM. Biosynthesis of the N-linked glycan in Campylobacter jejuni and addition onto protein through block transfer. J Bacteriol. 2006 188: 2427-34.

- Khoo, K.H., Douglas, E., Azadi, P., Inamine, J.M., Besra, G.S., Mikusova, K., Brennan, P. and Chatterjee, D. Truncated structural variants of lipoarabinomannan in ethambutol drug-resistant strains of *Mycobacterium smegmatis*. J. Biol. Chem. 1996 271: 28682-28690

- Kojima, N., Fenderson, B.A., Stroud, M.R., Goldberg, R.I., Habermann, R., Toyokuni, T. and Hakomori, S. Further studies on cell adhesion based on Le^x-Le^x interaction, with new approaches: embryoglycan aggregation of F9 teracarcinoma cells, and adhesion of various tumour cells based on Le^x expression. Glyconj. J. 1994 11, 238-248

- Koles K, Irvine KD, Panin VM. Functional characterization of Drosophila sialyltransferase. J Biol Chem. 2004 279:4346-57.

- Kornfeld R, Kornfeld S Assembly of asparagine-linked oligosaccharides, Annu Rev Biochem 1985 54: 631-64.

- Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC, Hernandez M, Kelly JF, Wacker M, Aebi M. Definition of the bacterial N-glycosylation site consensus sequence. EMBO J. 2006 25: 1957-66

- Kuntz C, Sonnenbichler J, Sonnenbichler I, Sumper M, Zeitler R. Isolation and characterization of dolichol-linked oligosaccharides from *Haloferax volcanii*. Glycobiology. 1997 7: 897-904.

- Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR. In vivo imaging of membrane-associated glycans in developing zebrafish. Science. 2008 320(5876):664-7

- Lechner J, Wieland F, Sumper M. Biosynthesis of sulfated saccharides N-glycosidically linked to the protein via glucose. Purification and identification of sulfated dolichyl monophosphoryl tetrasaccharides from halobacteria. J Biol Chem. 1985 260: 860-6.

- Lee, Y.C. and Ballou, C.E. Complete structures of the glycophospholipids of mycobacteria. Biochemistry 1965 4: 1395-1404

- Lengeler KB, Tielker D, Ernst JF. Protein-O-mannosyltransferases in virulence and development. Cell Mol Life Sci. 2008 65:528-44.

- Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi M, Wren BW. Functional analysis of the Campylobacter jejuni N-linked protein glycosylation pathway. Mol Microbiol. 2005 55:1695-703.

- Lommel M, Willer T, Strahl S. POMT2, a key enzyme in Walker-Warburg-Syndrome: somatic sPOMT2 but not testis specific tPOMT2 is crucial for mannosyltransferase activity in vivo. Glycobiology. 2008, sous presse

- Lussier M, Sdicu AM, Bussey H. The KTR and MNN1 mannosyltransferase families of *Saccharomyces cerevisiae*. Biochim Biophys Acta. 1999 1426:323-34.

- Martensson, .S, Levery, S.B., Fang, T.T. and Bendiak, B. Neutral core oligosaccharides of bovine submaxillary mucin. Use of lead tetraacetate in the cold for establishing branch positions. Eur. J. Biochem. 1998 258:603-622

- Matsumoto M, Hirata J, Hirohashi N, Hoshi M. Sperm-egg binding mediated by sperm alpha-L-fucosidase in the ascidian, *Halocynthia roretzi*. Zoolog Sci. 2002 19:43-8

- Michell SL, Whelan AO, Wheeler PR, Panico M, Easton RL, Etienne AT, Haslam SM, Dell A, Morris HR, Reason AJ, Herrmann JL, Young DB, Hewinson RG. The MPB83 antigen from *Mycobacterium bovis* contains O-linked mannose and (1-->3)-mannobiose moieties. J Biol Chem. 2003 278:16423-32.

- Monk CR, Sutton-Smith M, Dell A, Garden OA. Preparation of CD25(+) and CD25(-) CD4(+) T cells for glycomic analysis-a cautionary tale of serum glycoprotein sequestration. Glycobiology. 2006 16:11G-13G.

- Nigou, J., Gilleron, M., Cahuzac, B., Bounéri, J. D., Herold, M., Thurnher, M. & Puzo, G. The phospho-*myo*-inositol anchor of the lipomannans from *Mycobacterium bovis* bacillus Calmette Guérin. J. Biol. Chem. 1997 272, 23094-23103

- Nigou, J., Vercellone, A. and Puzo, G. New structural insights into the molecular deciphering of mycobacterial lipoglycan binding to C-type lectins: lipoarabinnomannan glycoform characterization and quantification by capillary electrophoresis at subnanomole level. J. Mol. Biol. 2000 299, 1353-1362

- North SJ, Koles K, Hembd C, Morris HR, Dell A, Panin VM, Haslam SM. Glycomic studies of Drosophila melanogaster embryos. Glycoconj J. 2006 23:345-54.

- Ong E, Kilburn DG, Miller RC Jr, Warren RA. *Streptomyces lividans* glycosylates the linker region of a beta-1,4-glycanase from *Cellulomonas fimi*. J Bacteriol. 1994 176:999-1008.

- Oriol, R., Mollicone, R., Cailleau, A., Balanzino, L. and Breton, C. Divergent evolution of fucosyltransferase genes from invertebrates, invertebrates, and bacteria. Glycobiology 1999 9:323-334

- Palma, A.S., Vila-Verde, C., Pires, A.S., Fevereiri, P.S. and Costa, J. A novel plant α4-fucosyntransferase synthesises the Lewis^a adhesion determinant. FEBS Lett. 2001 499: 235-238

- Paschinger K, Gutternigg M, Rendić D, Wilson IB. The N-glycosylation pattern of *Caenorhabditis elegans*. Carbohydr Res. 2008 343:2041-9

- Paschinger K, Rendic D, Lochnit G, Jantsch V, Wilson IB. Molecular basis of anti-horseradish peroxidase staining in *Caenorhabditis elegans*. J Biol Chem. 2004 279:49588-98.

- Paschinger K, Staudacher E, Stemmer U, Fabini G, Wilson IB. Fucosyltransferase substrate specificity and the order of fucosylation in invertebrates. Glycobiology. 2005 15:463-74.

- Pierce-Crétel, A., Decottignies, J.P., Wieruszeski, Strecker, G., Montreuil, J. and Spik, G. Primary structure of twenty three neutral and monosialylated oligosaccharides O-glycosidically linked to the human secretory immunoglobulin A hinge region determined by a combination of permethylation analysis and 400-MHz ^{1}H-NMR spectroscopy. Eur. J. Biochem. 1989 183:457-476

- Ragas A, Roussel L, Puzo G, Rivière M. The Mycobacterium tuberculosis cell-surface glycoprotein apa as a potential adhesin to colonize target cells via the innate immune system pulmonary C-type lectin surfactant protein A. J Biol Chem. 2007 282:5133-42.

- Reinhold BB, Hauer CR, Plummer TH, Reinhold VN. Detailed structural analysis of a novel, specific O-linked glycan from the prokaryote *Flavobacterium meningosepticum*. J Biol Chem. 1995 270:13197-203.

- Rivière M, Moisand A, Lopez A, Puzo G. Highly ordered supra-molecular organization of the mycobacterial lipoarabinomannans in solution. Evidence of a relationship between supra-molecular organization and biological activity. J Mol Biol. 2004 344:907-18.

- Robbins PW, Samuelson J. Asparagine linked glycosylation in Giardia. Glycobiology. 2005 15:15G-16G.

- Romain F, Horn C, Pescher P, Namane A, Riviere M, Puzo G, Barzu O, Marchal G. Deglycosylation of the 45/47-kilodalton antigen complex of Mycobacterium tuberculosis decreases its capacity to elicit in vivo or in vitro cellular immune responses. Infect Immun. 1999 67:5567-72.

- Samuelson J, Banerjee S, Magnelli P, Cui J, Kelleher DJ, Gilmore R, Robbins PW. The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. Proc Natl Acad Sci U S A. 2005 102:1548-53.

- Schachter H. Protein glycosylation lessons from *Caenorhabditis elegans*. Curr Opin Struct Biol. 2004 14:607-16
- Séveno M, Bardor M, Paccalet T, Gomord V, Lerouge P, Faye L. Glycoprotein sialylation in plants? Nat Biotechnol. 2004 22:1351-2

- Sidobre S, Nigou J, Puzo G, Rivière M. Lipoglycans are putative ligands for the human pulmonary surfactant protein A attachment to mycobacteria. Critical role of the lipids for lectin-carbohydrate recognition. J Biol Chem. 2000 275:2415-22.

- Sidobre S, Puzo G, Rivière M. Lipid-restricted recognition of mycobacterial lipoglycans by human pulmonary surfactant protein A: a surface-plasmon-resonance study. Biochem J. 2002 365:89-97.

- Strecker, G., Wieruszeski, J.M., Cuvellier, O., Michalski, J.C. and Montreuil, J. ^{1}H and ^{13}C-NMR assignments for sialylated oligosaccharide-alditols related to mucins. Study of thirteen components from hen ovomucin and swallow nest mucin. Biochimie 1992 74:39-52

- Szymanski CM, Burr DH, Guerry P. Campylobacter protein glycosylation affects host cell interactions. Infect Immun. 2002 70:2242-4

- Taniguchi T, Mizuochi T, Banno Y, Nozawa Y, Kobata A. Carbohydrates of lysosomal enzymes secreted by *Tetrahymena pyriformis*. J Biol Chem. 1985 260:13941-6.

- Teintenier-Lelièvre M, Julien S, Juliant S, Guerardel Y, Duonor-Cérutti M, Delannoy P, Harduin-Lepers A. Molecular cloning and expression of a human hST8Sia VI (alpha2,8-sialyltransferase) responsible for the synthesis of the diSia motif on O-glycosylproteins. Biochem J. 2005 392:665-74

- Treumann A, Xidong F, McDonnell L, Derrick PJ, Ashcroft AE, Chatterjee D, Homans SW. 5-Methylthiopentose: a new substituent on lipoarabinomannan in Mycobacterium tuberculosis. J Mol Biol. 2002 316:89-100.

- Venisse, A., Rivière, M., Vercauteren, J. and Puzo, G. Structural analysis of the mannan region of lipoarabinommannan from *Mycobacterium bovis* BCG. J. Biol. Chem. 1995 270: 15012-1521

- Voisin S, Houliston RS, Kelly J, Brisson JR, Watson D, Bardy SL, Jarrell KF, Logan SM. Identification and characterization of the unique N-linked glycan common to the flagellins and S-layer glycoprotein of *Methanococcus voltae*. J Biol Chem. 2005 280: 16586-93

- Warren, C.E., Krizus, A. and Dennis, J.W. Complementary expression patterns of six nonessential *Caenorhabditis elegans* core 2/I N-acetylglucosaminyltransfearse homologues. Glycobiology 2001 11:979-988

- Warren, C.E., Krizus, A., Partridge, E.A. and Dennis, J.W. *Caenorhabditis elegans gly-1*, a core 2/I N-acetylglucosaminyl-transferase homologue, is a glucosyltransferase. Glycobiology 2002 12:8G-9G

- Wieruszeski, J.M, Michalski, J.C., Montreuil, J., Strecker, G., Peter-Katalinic, J., Egge, H., van Halbeek, H., Mutsaers, J.H.G.M. and Vliegenthart, J.F.G. Structure of the monosialyl oligosaccharides derived from salivary gland mucin glycoproteins of the chines swiftlet (Genus *Collocalia*). J. Biol. Chem. 1987 262:6650-6657

- Willer T, Amselgruber W, Deutzmann R, Strahl S. Characterization of POMT2, a novel member of the PMT protein O-mannosyltransferase family specifically localized to the acrosome of mammalian spermatids. Glycobiology. 2002 12:771-83.

- Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St Michael F, Aberg E, Szymanski CM. Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, Campylobacter jejuni. J Biol Chem. 2002 277:42530-9

- Zheng Q, Van Die I, Cummings RD. A novel alpha1,2-fucosyltransferase (CE2FT-2) in Caenorhabditis elegans generates H-type 3 glycan structures. Glycobiology. 2008 18:290-302.

- Zheng Q, Van Die I, Cummings RD. Molecular cloning and characterization of a novel alpha 1,2-fucosyltransferase (CE2FT-1) from *Caenorhabditis elegans*. J Biol Chem. 2002 277:39823-32.