1°ABph 160448

N° d'ordre : 3210

Université des Sciences et Technologies de Lille Villeneuve d'Ascq

THESE

présentée par

Yann GUERARDEL

pour l'obtention du grade de

Docteur de l'Université Sciences de la Vie et de la Santé Spécialité: Biochimie

Variabilité structurale et fonctionnelle des glycoconjugués

Etudes réalisées chez les mycobactéries, les nématodes,

les éponges et les amphibiens.



Soutenue le 28 Novembre 2002 devant la commission d'examen:

Président:

Rapporteurs:

Examinateurs:

Directeur de Thèse:

Dr Jean-Claude MICHALSKI Dr Germain PUZO Dr Rudolf GEYER Dr Rafael ORIOL Dr Gradimir MISEVIC Dr Gérard STRECKER

2	
Somm	01100
Some	анс

INTRODUCTION	4
PARTIE A: Les O-glycannes de type mucine	
I. GENERALITES	
1. Les O-glycannes de type mucine	
11. Définition	7
12. Structure des O-glycannes	7
121. Les O-glycannes de mammifères	7
122. Les O-glycannes d'autres modèles animaux	. 12
123. Spécificité d'espèce des O-glycannes	. 15
13. Biosynthèse des O-glycannes	. 17
2. Les mucines	
21. Définition	. 18
22. Caractéristiques générales	
221. Introduction	. 20
222. Les unités de répétition	. 21
223. Les domaines riches en cystéines de type "Domaine D"	
et "Cysteine-knot"	. 23
224. Autres domaines riches en cystéines	. 24
225. Domaines spécifiques aux mucines membranaires	. 25
226. Conclusion	. 25
23. Régulation de la glycosylation des mucines	
231. Introduction	. 26
232. Régulation de l'initiation de la glycosylation	. 26
233. Régulation de la structure des glycannes par l'apomucine	. 29
234. Notion de glycoforme	29
24. Rôles des mucines	
241. Introduction	31
242. Mucines sécrétées	31
243. Mucines membranaires	33
244. Conclusion	35
II. PROBLEMATIOUE	36
III. RESULTATS	
1. Le modèle amphibien	
11 Description	
111 Introduction	40
112 Structure de la matrice extracellulaire	40
112. Entretare de la matrice extracellulaire au cours de la fécondation	43
12 Etude des O-glycappes isolés de la gangue oviducale de <i>Venopus laevis</i>	. 15
12. Lidde des O-grycamies isoles de la gangue ovidueale de <i>Achopus luevis</i>	44
121. I resentation des resultais	
122. Resultats	- 5 61
13 Etude des Orgiveannes isolés de la gangue oviducale de <i>Yanonus tronicalis</i>	
13. Druce des O-grycallies isoles de la galigue ovidueale de <i>Achopus il opicalis</i>	66
137 Régultate	00 AA
132. Rosultais	00 RA
14 Snécificité zoologique de la fécondation et snécificité d'esnèce des O-alvoannes	
14. Specificite zoologique de la recondation et specificite d'espèce des O-grycalilles,	

vue personnelle	82
2. Le modèle Caenorhabditis elegans	
21 Introduction	87
22. Analyse des O-glycannes de C. elegans	
221. Résultats	88
222. Conclusions	105
23. Purification préliminaire des mucines de C. elegans	
231. Introduction	106
232. Résultats	108
PARTIE B: Lipoglycannes de mycobactéries	
IGENERALITES	
1. Les infections mycobactériennes	
11 Introduction	114
12 I a tuberculose	114
12. La tuberculose	114
13. Les mycobacterioses	115
2. L'enveloppe des mycobacteries	117
21. Introduction	11/
22. Le complexe mycolyl-arabinogalactane-peptidoglycanne (mAGP)	120
23. Les glycolipides extractibles	
231. Les glycolipides de type phosphatidyl-myoinositol	100
2311. Les phosphatidylinositol-mannosides (PIMs)	128
2312. Les lipomannannes (LM) et lipoarabinomannannes (LAM).	130
232. Les glycolipides de type tréhalose	
2321. Les tréhaloses mono- et dimycolates	141
2322. Les tréhaloses polyacylés	143
2323. Les sulfatides	144
2324. Les lipooligosaccharides (LOS)	145
233. Les glycosylphénolphtiocérol (PGL)	148
234. Les glycopeptidolipides (GPL)	152
24. Les polysaccharides libres	153
25. Ultrastructure de l'enveloppe	155
3. Biosynthèse du lipoarabinomannanne	158
4. Activités biologiques liées aux lipoarabinomannannes	
41. Modulation de l'expression des cytokines	161
42. Phagocytose des mycobactéries	166
43. Chimiotactisme	168
II PROBLEMATIQUE	169
III RESULTATS	
1. Etude structurale des lipomannannes et lipoarabinomannannes	5
isolés de M. chelonae et M. kansasii	
11. Présentation des résultats	172
12. Etude des lipoglycannes de <i>M. chelonae</i>	172
13. Etude des lipoglycannes de <i>M. kansasii</i>	187
14. Conclusions	215
2. Etude des propriétés immunomodulatrices des linouvcannes de	e
2. Liude des proprietes inimunomodulatilees des inpogrycannes de	-

M. chelonae et M. kansasii

21. Présentation des résultats	21	17	7
--------------------------------	----	----	---

22. Résultats	
· ·	

PARTIE C: Glyconectines d'éponge	
I. GENERALITES	
1. Introduction	249
2. Interactions sucre-sucre chez les eucaryotes supérieurs	
21. Interactions homophiliques Le ^x -Le ^x	251
22. Interactions hétérophiliques	
221. Adhésion dépendante de GM ₃	252
222. Interaction galactocérébroside-cérébroside sulfate (GalCer-CBS)	254
3. Agrégation des cellules d'éponge	
31. Historique	255
32. Nature des interactions cellulaires	257
4. Conclusion. Importance physiologique des interactions	
sucre-sucre	261
II. PROBLEMATIQUE et RESULTATS	
1. Présentation des travaux	263
2. Résultats	264
3. Conclusions	292
4. Annexe	295
CONCLUSION GENERALE	321
REFERENCES	323

INTRODUCTION

La glycosylation revêt les formes les plus diverses: glycoprotéines, glycolipides, polysaccharides ou oligosaccharides libres. Chacune de ces formes de glycosylation fait montre d'une extraordinaire diversité structurale, et ce à tous les niveaux d'organisation biologique, entre espèces, entre populations et individus d'une même espèce, entre cellules et molécules, voire entre molécules similaires dans le cas des glycoformes de glycoprotéines. En plus de cette diversité "typologique", les profils de glycosylations individuels ou cellulaires varient au cours du développement, de la différentiation, et d'une manière plus générale de l'activation cellulaire. La nature chimique même des glycannes les prédispose à acquérir les structures les plus diverses. En effet, les nombreux degrés de liberté au sein des glycannes nature des monosaccharides, séquence, anomérie, forme pyranne/furane, isoméries D/L, position des liaisons, substitutions (sulfates, phosphates, acétyles...)- ont la capacité de générer un nombre stupéfiant de combinaisons. Chaque forme de glycosylation se subdivise elle même en un grand nombre de familles différentes qui présentent des caractères communs. Pour ne citer qu'un seul cas, les glycannes associés aux glycoprotéine se divisent en familles caractérisées par la nature du monosaccharide lié sur la protéine (GlcNAc, GalNAc, Fuc, Glc, Xyl...), le type de liaison glycanne-protéine (via O, C ou N) et éventuellement les caractéristiques générales des glycannes. Au sein de chacune de ces familles, on peut encore observer une très importante diversité structurale.

Malgré le manque flagrant d'analyse systématique de la glycosylation chez les espèces animales et végétales, il est à noter que toutes les formes de glycosylation ne sont pas réparties de façon homogène dans le monde vivant, chaque ordre ou sous-ordre ayant privilégié certaines d'entre elles. Ainsi, les polysaccharides complexes et les glycolipides constituent les formes prédominantes de glycosylation chez les bactéries, alors que les glycoprotéines sont très peu représentées. Contrairement aux glycannes d'eukaryotes qui sont constitués d'un panel restreint de monosaccharides, plusieurs centaines de monosaccharides différents ont été mis en évidence chez les bactéries. Chez les végétaux, la forme majoritaire de glycosylation est constituée de polysaccharides homogènes -cellulose, agarose, amidon- et -arabinogalactanes, xyloglycanneshétérogènes mais de nombreuses N-glycosylglycoprotéines ont également été mises en évidence. Les eucaryotes supérieurs ont quant à eux privilégié les glycoprotéines et les glycolipides au cours de leur évolution. Ces formes de glycosylation sont présents en grandes quantité à la surface des cellules des eukaryotes supérieurs, soit intégrés dans la membrane, soit au sein de la matrice extracellulaire.

Les glycannes sont impliqués dans trois types de fonction: (i) des fonctions de réserve; (ii) des fonctions d'ordre structural, en participant à l'élaboration de la paroi des bactéries, de l'ultrastructure des végétaux ou de la matrice extracellulaire des eucaryotes; (iii) des fonctions liées aux interactions cellulaires en général. Les répercussions physiologiques souvent dramatiques dues aux déficiences de la synthèse des glycoconjugués sont une preuve directe de l'importance fondamentale des glycannes dans les interactions cellulaires. De plus, la découverte des interactions spécifiques sucre-protéine et sucre-sucre a fournie une base moléculaire à l'étude des rôles joués par les glycannes. Malgré le nombre croissant de phénomènes physiologiques identifiés dans lesquels les glycannes sont impliqués, une question fondamentale reste en suspens: pourquoi une telle diversité structurale? Plus précisément et sans faire de finalisme, il est légitime de se demander si l'existence d'une telle diversité structurale à tous les niveaux d'organisation peut, in fine, avoir une fonction ou tout du moins se justifier d'un point de vue évolutif. Ceci se justifie d'autant plus que les rôles exacts de nombreux glycannes identifiés sont encore largement inconnus. L'une des hypothèses les plus convaincantes intègre l'émergence de la diversité glycannique dans un processus permettant à la fois le maintien de l'intégrité individuelle (reconnaissance soi/non soi) et la défense contre les organismes pathogènes [Gagneux, P & Varki, A., 1999]. La pression de sélection exercée par les agents pathogènes, dont beaucoup reconnaissent des glycannes, sur un organisme donné serait un puissant facteur de diversification des glycannes, non seulement au sein de cet organisme, mais également au sein de la population dont il fait partie. Dans cette optique, la variation structurale des récepteurs glycanniques à tous les niveaux permettrait de garantir la préservation d'un partie de la population (de cellules ou d'individus) contre l'infection. La preuve expérimentale définitive d'une telle théorie est quasiinaccessible, car les causes, ainsi que les effets, de ces changements sont en constante évolution, mais certaines observations sont en accord avec elle. En particulier, dernièrement il a été observé que l'infection par Nippostrangylus brasiliensis, un parasite intestinal, modifiait le profil de glycosylation des mucines intestinales de rat, suite à l'induction d'une α -Nacétylgalactosaminyltransférase impliquée dans la synthèse du déterminant de groupe A [Olson, F.J et al., 2002]. Ces changements rapides ont été interprétés comme un mécanisme de protection contre l'infection bactérienne, et pourraient être le reflet sur une courte durée de

L'ensemble des travaux présentés dans ce mémoire concernent l'étude de la diversité structurale des glycannes. Nous nous sommes plus particulièrement intéressés à la notion de

processus évolutifs qui agissent sur une échelle de temps beaucoup plus longue.

spécificité d'espèce de la glycosylation au travers de l'étude de plusieurs modèles expérimentaux: les mycobactéries, les organismes pluricellulaires primitifs et les vertébrés. Ces travaux vont être successivement présentés dans trois parties distinctes. Ils relèvent tous de la même démarche expérimentale qui consiste à comparer les profils de glycosylation d'espèces animales ou bactériennes souvent proches, mais leurs finalités sont différentes. Dans la partie A, nous présenterons l'étude de la structure des O-glycannes de type mucine isolées d'espèces animales dont la génétique est bien définie (Xenopus laevis, Xenopus tropicalis, Caenorhabditis elegans). Ces travaux ont pour objectif de mettre en évidence de nouvelles activités glycosyltransférasiques et constituent un pré-requis à l'étude ultérieure de la spécificité et du rôle des glycosyltransférases chez ces modèles expérimentaux. Dans la partie B, nous présenterons l'étude structurale de lipoglycannes -les lipomannanes et les lipoarabinomannanes- isolés de la paroi de deux espèces de mycobactéries atypiques pathogènes (Mycobacterium chelonae, Mycobacterium kansasii). Ces composés sont considérés comme des facteurs de virulence majeurs et leur étude se justifie par la recherche de nouveaux variants structuraux permettant de mieux appréhender les relations entre leur structure et leur fonction. Dans la partie C, nous présenterons les études comparatives effectuées sur un nouveau type de glycoconjugué -les glyconectines- isolés de trois espèces d'éponge marine. Les glyconectines sont impliquées au travers d'interactions homophiliques sucre-sucre dans l'agrégation spécifique d'espèce des cellules dont elles sont isolées. Cette étude a pour objectif de mettre en évidence des variations structurales inter-espèce des glyconectines et ainsi donner des bases à la modélisation moléculaire de leur interaction.

I. GENERALITES

1. Les O-glycannes de type mucine 11. Définition

Trois types de glycannes, qui sont définies en fonction de la nature de la liaison sucreprotéine, substituent les glycoprotéines: les O-, N- et C-glycosides. Les O-glycosides peuvent être liés sur les protéines *via* une unité de N-acétylgalactosamine, N-acétylglucosamine, mannose ou fucose. Nous ne retiendrons ici que la famille des glycannes liés sur le groupement hydroxyle de la thréonine et de la sérine *via* une unité de α -Nacétylgalactosamine. Les membres de cette famille sont majoritairement associés aux glycoprotéines de type mucine mais substituent également de nombreuses glycoprotéines solubles et membranaires dans la presque totalité du règne animal. Ils ont ainsi été nommés par extension "O-glycannes de type mucines", mais sont le plus souvent simplement appelés "O-glycannes". Leur taille varie de un à une vingtaine de monosaccharides et ils présentent une hétérogénéité structurale extrême. De fait, plusieurs centaines de O-glycannes distincts ont été décrits dans l'ensemble du règne animal, le seul dénominateur commun à ces composé étant la présence d'un résidu de GalNAc à leur extrémité réductrice.

12. Structure des O-glycannes

121. Les O-glycannes de mammifères

La structure des O-glycannes a principalement été étudiée chez l'homme et chez quelques autres espèces de mammifères, surtout à partir des mucines excrétées. Au vu de l'hétérogénéité des O-glycannes et à des fins de classement et d'étude de leur biosynthèse, Hounsell et Feizi [Hounsell, E.F. & Feizi, T., 1982] ont identifié trois régions correspondant à des séries distinctes de réactions enzymatiques de biosynthèse: noyau glycannique formé par l'association du résidu de GalNAc terminal et d'un ou de deux monosaccharides, squelette résultant de l'élongation du noyau glycannique et périphérie (Fig. A1). Les sucres périphériques peuvent substituer soit le squelette, soit directement le noyau. Cette rationalisation de la structure des O-glycannes s'applique principalement aux O-glycannes de mammifères et a permis d'appréhender efficacement la biosynthèse des O-glycannes chez l'homme. Néanmoins, ce système de description montre rapidement ses limites dans le cas de composés plus "exotiques" trouvés dans d'autres modèles animaux.



Fig. A1: D'après Roussel, P. & Lamblin, G., 1996.

Selon ce système, huit noyaux 0glycanniques consensuels sont reconnus (Tableau A1). Certains sont ubiquitairement retrouvés tandis que d'autre le sont plus rarement. Ils sont numérotés dans l'ordre de leur découverte. Les noyaux 1, 2, 3 et 4 correspondent aux quatre types majoritairement retrouvés dans les mucines. Ils ont notamment tous été isolés des mucines intestinales et coliques de rat [Carlsson, H.E. et al., 1978; Brockhausen, I. et al., 1985], des mucines gastriques et respiratoires humaines [Hanisch, F.G. et al., 1993; Klein, A. et al., 1988], ou encore des kystes ovariens humains [Mutsaers, J.H. et al., 1986]. Les autres noyaux sont beaucoup moins fréquents. Les noyaux 5 et 6 ont été caractérisés dans le méconium humain [Hounsell, E.F. et al., 1985; Capon, C. et al., 1989] et dans la mucine sous-maxillaire bovine [Savage, A.V. et al., 1990], tandis que les noyaux 7 et 8 ont été mis en évidence au niveau des mucines sous-maxillaires bovines [Martensson, S. et al., 1998] et



Tableau A1: Structure des noyaux des Oglycannes reconnus.

bronchiques humaines [Van Halbeek, H. et al., 1994].

L'élongation du noyau glycannique forme le squelette des O-glycannes. Chez les mammifères en général et l'homme en particulier, elle résulte principalement de l'action successive d'une galactosyl- et d'une *N*-acétylglucosaminyl-transférase qui forment des unités disaccharidiques de séquence Gal-GlcNAc. Deux types d'unités sont principalement synthétisées: des unités de type 1 Gal(β 1-3)GlcNAc β et des unités de type 2 Gal(β 1-4)GlcNAc β . L'élongation débute sur n'importe lequel des noyaux glycanniques et conduit à la formation de polymères linéaires ou ramifiés, homogènes ou mixtes, en position C-3 ou C-6 d'une unité galactose.

La périphérie est constituée de monosaccharides liés le plus souvent selon une anomérie a, en positions terminales non-réductrices du squelette ou du noyau. Chez les mammifères ces monosaccharides consistent habituellement en des unités de Gal, Fuc et NeuAc. Chez l'homme, ces sucres périphériques constitueront le support de motifs antigéniques, autrefois connus sous le vocable d'antigènes de groupes sanguins, bien que le terme antigène tissulaire apparaisse aujourd'hui plus approprié. Ces motifs antigéniques, abondamment trouvés à la surface des mucines, caractérisent également les glycoprotéines et les glycolipides. Le tableau A2 résume la structure des principaux de ces motifs. De très nombreux autres motifs sont synthétisés par ajout, entre autre, d'acides sialiques et de groupements sulfates. L'acide sialique est plus communément lié en α -2,3 ou α -2,6 sur un Gal terminal et en α -2,6 sur le GalNAc du point d'attache. Bien que la majorité des mucines humaines renferment essentiellement de l'acide N-acétylneuraminique (Neu5Ac), des acides 9-O-acétyl-N-acétylneuraminique (Neu5,9Ac₂) et 7,9-di-O-acétyl-N-acétylneuraminique (Neu5,7,9Ac₃) ont été mis en évidence dans les mucines du colon humain [Muchmore, E.A. et al., 1987; Gold, D.V. & Shochat, D., 1989]. Quant aux groupements sulfates, ils sont présents en positions 3-, 4- ou 6- du galactose et en position 6- de la N-acétylglucosamine. Acide sialique, fucose et groupement sulfate peuvent coexister sur un même O-glycanne [Lo-Guidice, J.M. et al., 1994].

Antigàno	Structu	Ire
Antigene	Type 1	Type 2
i	Galβ – GlcNAcβ Galβ – GlcNAcβ	
I	Galβ – GlcNAcβ Galβ – GlcNAcβ Galβ – GlcNAcβ	
H1/H2	GlcNAcβ Galβ I Fucα	Galβ – GlcNAcβ Ι Fucα
Le ^a /Le ^x	Fuca – GlcNAc β Gal β	Galβ - GlcNAcβ Fucα
Le ^b /Le ^y	Fucα – GlcNAcβ Galβ I Fucα	Galβ – GlcNAcβ Ι Fucα Fucα
A1/A2	GlcNAcβ Galβ GalNAcα Fucα	Galβ – GlcNAcβ J GalNAcα Fucα
B1/B2	$GlcNAc\beta$ $Gal\beta$ $Gal\alpha$ Fuca	Galβ - GicNAcβ Galα Fucα
Т	, GalNAα Galβ	caSer/Thr
Tn	GalNAcaS	er/Thr
Sialyl-Tn	NeuAc2α GalN	AcaSer/Thr

Tableau A2: Structure des principaux motifs glycanniques d'antigènes tissulaires identifiés chez l'homme. | liaison 1-2; / liaison 1-3; — liaison 1-4; \ liaison 1-6.

La simple association des noyaux, squelettes et sucres périphériques permet d'envisager plusieurs milliers de structures O-glycanniques chez l'homme. Néanmoins, ce nombre est minoré par les schémas de biosynthèse décrits chez l'homme. Nonobstant, 88 Oglycannes différents ont été isolés et caractérisés à partir des seules mucines bronchiques humaines [Roussel, P. & Lamblin, G., 1996] et certains auteurs évaluent à plusieurs centaines le nombre de structures différentes présentes dans les mucines respiratoires humaines [Lamblin, G. *et al.*, 1991].



Fig. A2: Exemples de O-glycannes isolés de diverses espèces animales. a) Glycannes libérés de glycoprotéines membranaires de *Trypanosoma cruzii* (Todeschini, R.A. *et al.*, 2001). b) Glycannes extraits de mucines salivaires du martinet d'Asie (nid d'hirondelle) (Wieruszeski, J.M. *et al.*, 1987). c) Glycannes libérés des polysialoglycoprotéines (PSGP) de *Salmonidae* (d'après Inoue, S. & Inoue, Y., 1997).

122. Les O-glycannes d'autres modèles animaux

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 notre laboratoire, 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 les noyaux de types 2 ou 5 [Wieruszeski, J.M. et al., 1987; Strecker, G. 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. La présence d'un noyau de type complètement nouveau n'a été mis en évidence que chez le parasite Tripanosoma cruzi (Fig. A2) [Previato, J.O. et al., 1994]. Les O-glycannes sont liés à la protéine par une unité de GlcNAc, et ne répondent donc pas stricto sensu à la définition des O-glycannes de type mucine. Néanmoins, leur structure générale, ainsi que leur localisation sur des glycoprotéines hyper-O-glycosylées, suggèrent qu'ils sont les équivalents des O-glycannes de type mucine chez ce parasite. En fait, la variabilité structurale des O-glycannes tient essentiellement à l'élongation ultérieure des chaînes glycanniques, qui peut prendre les formes les plus diverses en fonction de l'organisme. Par exemple, les O-glycannes isolés de la mucine salivaire du martinet d'asie présentent un squelette constitué du disaccharide Gal(\beta1-4)Gal\beta, inconnu chez l'homme (Fig. A2). 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.

Ainsi, les O-glycannes de truite et de saumon sont constitués de corps glycanniques non répétitifs sur lesquels sont greffés des oligo- ou polymères d'acides sialiques [Inoue, S. & Inoue, Y, 1997]. De tels O-glycannes isolés des polysialoglycoprotéines d'œufs de *Salmonidae* sont présentés dans la figure A2.



 $\begin{array}{c} GalNAc \ \beta(1-4)GlcNAc \ \beta(1-6) \\ Fuc \ \alpha(1-3) \\ Gal \ \beta(1-3) \\ Fuc \ \alpha(1-4) \ Kdn \ \alpha(2-3) \\ Fuc \ \alpha(1-2) \\ Fuc \ \alpha(1-2) \\ \end{array}$

Ambystoma maculatum

Fuc $\alpha(1-5)$ Fuc $\alpha(1-4)$ Kdn $\alpha(2-6)$ GalNAc-ol Fuc $\alpha(1-5)$ Gal $\beta(1-3)$ Fuc $\alpha(1-4)$ Kdn $\alpha(2-3)$

Ambystoma tigrinum

Fuc $\alpha(1-4)$ Kdn $\alpha(2-6)$ Fuc $\alpha(1-3)$ GalNAc-ol Gal $\alpha(1-4)$ Gal $\beta(1-4)$ GlcNAc $\beta(1-3)$ Fuc $\alpha(1-2)$

Ambystoma mexicanum

GlcNAc $\beta(1-6)$ Gal $\alpha(1-4)$ Gal $\beta(1-3)$ Fuc $\alpha(1-2)$ Fuc $\alpha(1-2)$ Gal $\beta(1-3)$

Xenopus laevis



GlcNAc $\beta(1-6)$ Gal $\alpha(1-3)$ Gal $\beta(1-3)$ Gal $\beta(1-3)$ Gal $\beta(1-3)$ GalNAc-ol Fuc $\alpha(1-2)$ Fuc $\alpha(1-2)$

Bufo bufo

 $\begin{array}{c} Gal NAc-ol \\ Gal \beta(1-4)Gal \beta(1-3) \\ HSO_3 Gal \beta(1-3) \\ Fuc \alpha(1-2) \\ \end{array}$

Rana temporaria

Fig. A3: Exemples de O-glycannes isolés des gangues oviducales d'amphibiens.

Les acides sialiques, exclusivement localisés en position terminale non-réductrice dans les O-glycannes de mammifères, constituenttypiquement le domaine majeur des O-glycannes de poisson. De plus, on observe parmi ces acides sialiques de l'acide N-glycolylneuraminique (NeuGc) et des formes particulières telles que l'acide 2-céto-3-désoxynononique (Kdn) et l'acide 9-O-acétyl-2-céto-3-désoxynononique (Kdn9Ac) [Iwasaki, M. *et al.*, 1990].

L'étude des O-glycannes isolés des gangues oviducales de dix huit espèces d'amphibiens de l'ordre des anoures (grenouilles, crapauds, rainettes) et des urodèles (salamandres, tritons) a permis d'isoler et de caractériser quelque 350 O-glycannes originaux. L'ensemble de ces structures peuvent être visualisées sur le site internet à l'adresse suivante: ustl.univ-lille1.fr/glycobase/index.htm. Quelques exemples de ces structures sont décrits dans la figure A3. Ils renferment généralement entre deux et douze monosaccharides, et jusque vingt monosaccharides chez l'anoure Rana clamitans. De nombreux déterminants antigéniques trouvés chez l'homme ont été identifiés chez les amphibiens. En particulier, l'antigène tissulaire de groupe H est présent dans les trois espèces d'ambystomes étudiées, tandis que l'antigène tissulaire de groupe A est observé dans plusieurs O-glycannes de l'anoure Bufo arenarum [Morelle, W. et al., 1998]. De même, les motifs Le^x, Le^y et A Le^y substituent la majorité des O-glycannes isolés de Pleurodeles Waltl [Strecker, G. et al., 1992]. Ces motifs peuvent donc être terminaux comme chez l'homme, mais également substitués tel le dimère de l'antigène H substitué par une résidu de Gal(α 1-4) chez Xenopus laevis (Fig. A3). Malgré la présence de motifs communs, les O-glycannes d'amphibiens diffèrent substantiellement de ceux des mammifères par la présence de nombreux motifs complexes inédits. En effet, bien qu'ils soient constitués exclusivement de Gal, GlcNAc, GalNAc, Fuc, GlcA, Neu5Ac et Kdn, et comme le suggèrent les quelques structures présentées, aucune règle présidant leurs agencements n'a émergé de l'étude des 350 structures décrites. Tout au plus pouvons nous observer des motifs partagés par plusieurs espèces telles que les séquences -3[Fuc(α 1-2)₀₋₁Gal(β 1-]₁₋₃, Gal(β 1-4)Gal β , ou les disaccharides de types 1 et 2. Ces motifs peuvent être considérés comme l'équivalent du squelette des O-glycannes de mammifères, bien qu'ils exhibent une hétérogénéité plus importante. Au contraire des mammifères, la fucosylation chez les amphibiens n'intervient pas exclusivement en position terminale, mais peut former des motifs oligomérisés tels que Fuc(α 1-2)Fuc α et Fuc(α 1-3)[Fuc(α 1-2)]Fuc α comme chez Ambystoma maculatum [Fontaine, M.D. et al., 1995]. De même pour les acides sialiques qui peuvent être mono- ou disubstitués, entre autres chez Ambystoma tigrinum

[Maes, E. et al., 1995]. Comme chez les poissons, la diversité structurale est encore accrue par l'existence de nombreuses formes d'acides sialiques différentes. Ainsi, chez les ambystomes, le Kdn et le Kdn9Ac ont été détectés dans des proportions équivalentes, tandis que du Kdn7Ac et du Kdn7,9Ac₂ ont été mis en évidence chez *Pleurodeles waltl* [Klein, A. communication personnelle].

Au contraire de la débauche structurale des O-glycannes observée chez les poissons et les amphibiens, les insectes semblent posséder une O-glycosylation beaucoup plus simple. Ainsi, les mucines de guêpe sont substituées par des O-glycannes très courts. Le composé majoritairement exprimé possède la séquence du noyau de type 1, qui peut néanmoins être substitué par un groupement de phospho-éthanolamine [Maes, E., résultats non publiés]. De façon similaire, l'analyse structurale des O-glycannes synthétisés par des cellules de lépidoptère n'a révélé que l'existence de GalNAc et Gal(β1-3)GalNAc [Thomsen, D.R. *et al.*, 1990].

123. Spécificité d'espèce des O-glycannes

L'analyse systématique des O-glycannes chez différentes espèces animales a mis en évidence deux faits majeurs. Premièrement, comme nous l'avons démontré au chapitre précédent, la O-glycosylation est encore plus variable que ce que l'étude du modèle humain ne le suggérait déjà. Deuxièmement, les évidences principalement accumulées chez les poissons et les amphibiens suggèrent que les O-glycannes sont strictement spécifiques à chaque espèce. Ceci n'implique aucunement que, malgré l'énorme potentiel de variabilité des O-glycannes, chaque espèce présente des structures totalement différentes des autres espèces.

Chez les saumons, les O-glycannes isolés des polysialoglycoprotéines (PSGP) de nombreuses espèces présentent tous une architecture très similaire. La spécificité d'espèce apparaît principalement dans la structure caractéristique des acides sialiques intégrés aux chaînes polysialylées: nature (Neu5Ac, Neu5Gc ou Kdn), position des groupements O-acétyl et O-lactyl.

Chez les amphibiens, la variabilité inter-spécifique est plus marquée. Toutes les espèces étudiées synthétisent une, deux, voire trois séries de O-glycannes d'architectures distinctes. Les O-glycannes d'une espèce donnée sont caractérisés soit par la présence de

motifs glycanniques inédits, soit par un assemblage particulier de motifs communs à plusieurs espèces. Par exemple, chaque espèce d'ambystome étudiée se distingue des deux autres par la présence d'un motif apparenté mais différent: $Fuc(\alpha 1-5)[Fuc(\alpha 1-4)]Kdn$ pour A. tigrinum, Fuc(α 1-3)Fuc(α 1-4)Kdn pour A. mexicanum et Fuc(α 1-3)[Fuc(α 1-2)]Fuc(α 1-4)Kdn pour A. maculatum. Par contre Bufo bufo ne présente pas de motif inédit mais se caractérise par un arrangement particulier de motifs connus. Ainsi, c'est la substitution d'un dimère de galactose Gal(β 1-3)Gal, généralement fucosylé, par le motif Gal(α 1-3)GalNAc qui caractérise la glycosylation de Bufo bufo. De plus, la présence d'un résidu NeuGc augmente encore l'aspect spécifique de la glycosylation des mucines oviducales de Bufo bufo. Il est à noter que les Oglycannes d'espèces phylogénétiquement proches partagent souvent plus de points communs que des espèces éloignées. Ainsi, l'étude des O-glycannes de deux espèces de crapauds phylogénétiquement très proches, Bombina variegata et Bombina bombina, a démontré qu'ils ne se différenciaient que par la nature de leurs acides sialiques terminaux, Kdn ou Neu5Ac. Ceci n'est pas toujours vrai, comme l'a montré l'étude de Pleurodeles waltl et Pleurodeles poireti, qui présentent des structures glycaniques très différentes malgré leur degré de parenté proche.

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 d'amphibiens, le potentiel de diversité de la glycosylation semble être compatible avec une telle hypothèse.

13. Biosynthèse des O-glycannes

Au contraire de la N-glycosylation, qui résulte du transfert en bloc d'un glycanne présynthétisé sur l'axe peptidique, la O-glycosylation résulte de l'addition séquentielle de monosaccharides sur le peptide cible, par un ensemble de glycosyltransférases distinctes. Ces enzymes seraient réparties tout le long de l'appareil de Golgi. Ainsi, la maturation des Oglycannes s'effectuerait au cours du cheminement des glycoprotéines dans la voie de sécrétion [Brockhausen, I., 1995]. Chaque glycosyltransférase présente une spécificité très stricte envers le motif accepteur O-glycannique (ou peptidique dans le cas de l'initiation de la glycosylation), le nucléotide-sucre donneur et la liaison glycosydique formée (anomérie et position). L'identification de la plupart de ces enzymes a permis de décrypter les voies biosynthétiques conduisant à l'assemblage des structures majeures humaines. Nous ne détaillerons pas ces voies de biosynthèse qui sont décrites dans de nombreuses revues générales [e.g. Brockhausen, I., 1995; Brockhausen, I., 1999], mais exposerons les principales voies de contrôle de la O-glycosylation. Chez l'homme, ces études ont démontré que la synthèse de centaines de O-glycannes différents était régulée de manière très stricte, spécifiquement au tissu et au stade de développement,.

Le panel de glycosyltransférases exprimé par un tissu est l'élément clef du contrôle de la direction que prend la O-glycosylation. Ces enzymes sont très spécifiques pour leur substrat accepteur et toute addition d'un monosaccharide par une glycosyltransférase va influer sur l'activité des autres glycosyltransférases. Comme plusieurs enzymes pouvent entrer en compétition pour le même substrat, la quantité finale d'un O-glycanne donné va principalement dépendre des activités relatives de ces enzymes et de leurs compartimentations le long de l'appareil de Golgi. De plus, certains motifs glycanniques inhibent l'action de certaines ou de toutes les glycosyltranférases. Par exemple, l'addition d'un NeuAc en (α 2-6) sur le premier GalNAc d'un O-glycanne est un signal qui empêche toute élongation de ce composé. Au contraire, l'addition d'un monosaccharide peut être indispensable à l'activité d'autres glycosyltransférases. Ainsi, la structure finale de tout oligosaccharide va être dirigée dès les étapes précoces de sa synthèse.

Les activités des glycosyltransférases sont elles-mêmes régulées de façon spécifique au tissu par le micro-environnement de la membrane et du lumen de l'appareil de Golgi. Leur activité peut également être régulée par des phénomènes post-traductionnels. Enfin, l'activité glycosyltransférasique peut être indirectement régulée par la vitesse de transport des nucléotides-sucres dans le Golgi et donc par l'accessibilité de l'enzyme à son substrat donneur. Les voies de biosynthèse des O-glycannes n'ont été décrites en détail que chez l'homme. Au vu des homologies observées entre toutes les espèces animales dans les premières étapes de la synthèse (formation des noyaux), les processus de régulation décrits pour l'homme sont probablement communs à toutes les espèces. Comme nous l'avons introduit en début de ce chapitre, les O-glycannes sont majoritairement présent sur les glycoprotéines de type mucine. Ceci est particulièrement vrai en ce qui concerne les modèles d'études que nous avons utilisés, les gangues oviducales d'amphibiens et *Caenorhabditis elegans*. De fait nous allons décrire dans le chapitre suivant la structure des mucines et les fonctions qui leurs sont associées.

2. Les Mucines

21. Définition

Les mucines sont définies comme des glycoprotéines de haute masse moléculaire, abondamment O-glycosylées. Elles sont retrouvées à l'interface de nombreux épithéliums et de leur environnement extracellulaire. Originellement, le terme "mucine" était uniquement associé aux glycoprotéines constituant le mucus secrété par les épithéliums. Plus tard, il a été montré que ces mêmes épithéliums produisaient également des glycoprotéines transmembranaires qui ont été désignées, par extension, par le terme "mucine". De fait cette dénomination regroupe des molécules à la fois sécrétées et liées à la membrane qui ont été incluses, chez l'humain, dans la famille MUC sur la base d'informations de séquence [Strous, G.J. *et al.*, 1992]. Enfin, de nombreuses autres glycoprotéines ont été associées à la famille des mucines sur la simple base de leur haut degré de substitution par des O-glycannes sans toutefois présenter d'homologie avec la famille MUC et, pour cette raison, ne serons pas présentées ici.

Les mucines exhibent typiquement des masses moléculaires très importantes. Les chaînes peptidiques des MUCs, appelées apomucines, présentent des masses comprises entre 200 et 900 kDa, à l'exception de MUC7 dont l'apomucine a une faible masse de 39 kDa. La masse moléculaire d'une mucine glycosylée reste difficile à apprécier. A titre d'exemple, la masse moléculaire moyenne du polypeptide de MUC2 est d'environ 500 kDa [Gum, J.R., 1994]. Celle ci augmente à 2-4 MDa pour la forme complètement glycosylée et peut atteindre 40 MDa après polymérisation [Carlstedt, I. *et al.*, 1984]. A cause de la densité exceptionnelle de leur O-glycosylation (jusqu'a 80% du poids de la glycoprotéine), les mucines sont rétives à toute analyse de séquence peptidique sur la majorité de leur longueur. De fait, toutes les séquences polypeptidiques des mucines publiées ont été déduites du séquençage partiel ou complet de leur ADN complémentaire.



Fig. A4:Comparaison des ultrastructures d'un monomère de mucine sécrétée (MUC2) et de mucine membranaire (MUC1), d'après Dekker, J. et al., 20002.



Fig. A5: Séquences polypeptidiques partielles déduites des séquences nucléotidiques de quelques gènes MUCs humains parmi les mieux caractérisés.

La famille MUC est en rapide expansion: elle comptait onze membres dans la revue générale de Moniaux et al. d'octobre 2001, quatorze membres dans celle de Dekker et al. de mars 2002 et s'est agrandie depuis de deux membres supplémentaires, MUC17 [Gum, J.R. et al., 2002] et MUC15 [Pallesen et al., 2002]. En fonction de leur structure, les mucines humaines peuvent être classées en deux sous-familles distinctes: mucines sécrétées (MUC2, MUC5B, MUC5AC, MUC6, MUC7) et mucines liées à la membrane (MUC1, MUC3, MUC4, MUC12, MUC13, MUC15, MUC17) [Moniaux et al., 2001]. Les mucines MUC8, MUC11 et MUC16 ne sont pas encore classifiées en raison du manque d'information quant à leurs séquences. MUC7 se distingue des autre mucines sécrétées par le fait qu'elle est la seule à ne pas former des gels stables. Cette propriété est corrélée à des différences dans l'organisation de son apomucine que nous allons détailler. Chaque membre d'une même famille présente des caractéristiques structurales communes et probablement des fonctions spécifiques. En effet, pendant longtemps, l'unique fonction attribuée aux mucines était en relation avec leurs effets protecteurs et lubrifiants. Les études récentes sur les structures et les fonctions des mucines tendent à démontrer qu'elles seraient également impliquées dans d'autres processus tels que le recrutement des macrophages, la différentiation, le renouvellement de l'épithélium et la carcinogenèse. En particulier, les différences observées entre les conformations des mucines membranaires et sécrétées (figure A4) suggèrent qu'elles jouent des rôles différents.

Il est à noter que de nombreux homologues de cette famille ont été décrits chez les animaux et ont été utilisés de façon extensive pour étudier les propriétés physico-chimiques, l'organisation génomique et la biosynthèse des mucines. Parmi ceux ci, les plus étudiés sont les gènes RIM (mucine intestinale de rat), FIM (mucine tégumentaire de xénope), BSM (mucine sous-maxillaire de bœuf), CTM (mucine bronchique de chien) et PSM (mucine sousmaxillaire de porc).

22. Caractéristiques générales 221. Introduction

Jusqu'a présent, seules les séquences polypeptidiques de six mucines humaines (MUC1, MUC2, MUC4, MUC5B, MUC5AC et MUC7), ainsi que celle de PSM, ont été complètement déterminées. Malgré l'existence d'une grande diversité dans ces séquences, il apparaît que tous les membres de la famille MUC présentent une structure générale très

semblable (Fig. A5) et de fortes homologies dans certains domaines protéiques. Ainsi, ces données mettent en lumière plusieurs caractéristiques communes à l'ensemble des mucines humaines, ou propres à chacune des familles de mucines humaines ou animales.

222. Les unités de répétition

Quelque soit leur taille, toutes les mucines possèdent des régions fortement enrichies en proline, thréonine et sérine (régions PTS) abondamment substituées par des O-glycannes. Elles sont systématiquement localisées dans la partie centrale de la protéine et sont flanquées à chaque extrémité par d'autres types de domaines bien définis. La forte densité de glycosylation de ces régions confère à la partie centrale des mucines une structure filamenteuse étendue par les effets combinés de charge et d'encombrement stérique [Jentoft, N. et al., 1990]. La région PTS peut être ininterrompue (e.g. MUC1, MUC6, PSM) ou consister en plusieurs fragments (e.g. MUC2, MUC5B) séparés par des domaines riches en cystéines. Ces régions sont organisées en unités de répétitions dont le nombre, la taille et la séquence varient très fortement d'une mucine à l'autre [Gender, S.J. et al., 1995]. La taille de l'unité de répétition varie de 8 à 169 acides aminés chez les MUCs. Pour certaines mucines (MUC1-4, MUC6, PSM) les unités de répétition sont parfaitement identiques au sein d'une même mucine, tandis que pour d'autres elles sont similaires mais non identiques (MUC5AC, MUC5B) [Van Klinken, B.J.W et al. 1998a; Eckhardt, A.E. et al. 1997]. Ainsi, trois séquences répétitives distinctes ont été mises en évidences dans la régions PTS chez MUC5B [Desseyn, J.L. et al., 1997]. Les régions PTS sont également soumises à un polymorphisme inter-individuel, appelé VNTR pour "Variable Number of Tandem Repeats", portant sur le nombre d'unités de répétitions. Par exemple, la mucine épithéliale MUC1 présente une unité de répétition de 20 acides aminés répétée de 20 à 120 fois [Gendler, S. et al., 1988]. De même, PSM est codée par au moins trois allèles de 99, 110 et 135 unités de répétitions de 81 acides aminés chacune [Eckhardt, A.E. et al., 1997]. Il est a noter que l'effet de ce polymorphisme sur la masse moléculaire de la mucine est encore amplifié par la présence d'une abondante O-glycosylation sur les domaines de répétition.



Fig. A6: Schéma d'association des monomères de mucine sous-maxilaire porcine proposé par Perez--Vilar, J. et Hill, R.L., 1998b.

223. Les domaines riches en cystéines de type "Domaine-D" et "Cysteine-knot"

A l'exception de MUC7 [Bobek, L.A. *et al.*, 1993], tous les MUCs sécrétés, ainsi que de nombreuses mucines animales, possèdent des régions riches en cystéines, appelées domaines-D, souvent retrouvées aux deux extrémités de la protéine. Les domaines D des mucines présentent une homologie importante avec ceux du pré-pro facteur von Willebrand (vWB) humain [Gums, J.R. *et al.*, 1994; Joba, W. & Hoffmann, W., 1997] où ils ont été identifiés pour la première fois. La plupart des mucines sécrétées contiennent trois domaines-D homologues (D1, D2 et D3) dans leur partie N-terminale, et certaines un quatrième (D4) à leur extrémité C-terminale. Ces domaines d'environ 250 acides aminés qui contiennent jusqu'a 30 cystéines présentent de très fortes homologies d'une mucine à l'autre et avec le facteur vWB. Ces mêmes mucines contiennent également un autre domaine riche en cystéines (10% de cytéines) à leur extrémité C-terminale [Gum, J.R. *et al.*, 1994; Eckhardt, A.E., *et al.*, 1997; Joba, W. & Hoffmann, W., 1997; Verma, M. *et al.*, 1993]. Il présente une forte homologie avec un domaine C-terminal du facteur vWB de la famille des "cysteine-knots" (CK). Les domaines-D et CK renferment de nombreux sites de N-glycosylation. MUC2, par exemple, en possède 30 [Gum, J.R. *et al.*, 1994].

De nombreuses études, portant en particulier sur MUC2 et PSM, ont montré que les domaines-D et CK étaient directement impliqués dans le processus de polymérisation des mucines. Ainsi, comme illustré par la figure A6, toutes les mucines possédant les domaines-D et CK se dimériseraient d'abord par leurs extrémités C-terminales via des ponts disulfures entre les domaines CK, puis se multimériseraient par leurs extrémités N-terminales via des ponts disulfures entre les domaines-D. La dimérisation des domaines CK a été directement observée pour PSM [Perez-Vilar, J. et al., 1998b], MUC2 [Asker, N. et al., 1998b] et Muc2 de rat [Bell, S.L., et al., 1998]. De manière identique, MUC5AC, MUC5B et MUC6 se dimérisent juste après leur synthèse dans le réticulum endoplasmique [Sheehan, J.K. et al., 1996; Asker, N. et al., 1998a; Van Klinken, B.J.W. et al., 1998b], probablement selon un processus comparable. Pour MUC2 [Asker, N. et al., 1998b], Muc2 de rat [Dekker, J. & Strous, G.J., 1990] et MUC5AC [Asker, N. et al., 1998a], la N-glycosylation des domaines CK est un pré requis à leur dimérisation, tandis qu'elle semble accessoire à PSM [Perez-Vilar, J. et al., 1996]. Les études effectuées sur PSM ont mis en évidence que sur les onze résidus de cystéines que compte le domaine CK, deux, C¹³²⁴⁴ et C¹³²⁴⁶, étaient indispensables à la formation des ponts disulfures inter-chaines [Perez-Vilar, J. and Hill, R.L., 1998a]. Ces deux cytéines font partie d'une courte séquence $C^{13244}LC^{13246}C$ conservée dans toutes les mucines sécrétées formant des gels et le facteur vWB.

La multimérisation des mucines, quant à elle, s'effectuerait par l'intermédiaire des extrémités N-terminales seulement après que les domaines PTS aient été O-glycosylés (Fig. A6). Similairement au domaine CK, une séquence directement impliquée dans la formation de ponts disulfures inter-chaines a été identifiée dans le domaine D3 de PSM [Perez-Villar, J. & Hill, R.L., 1998c]. Elle est très conservée chez les mucines sécrétées et le facteur vWB, ce qui a permis de généraliser le processus de multimérisation observé chez PSM aux autres mucines possédant des domaines D et CK. Par contre, il apparaîtrait que les modalités précises de ce phénomène varient sensiblement d'une mucine à l'autre. Par exemple, la libération d'un polypeptide en position N-terminale contenant les domaines D1 et D2 précède toujours la multimérisation du facteur vWB [Sadler, J.E., 1998] et de MUC5B [Wickström, C. & Carlstedt, I., 2001]. Il a été postulé que les domaines D1 et D2 du pré-pro facteur vWB, et par extension de MUC5B, formeraient des interactions non covalentes entre monomères et permettraient leur bon alignement, indispensable à la formation de ponts disulfures entre les domaines-D3 [Wise, R. et al., 1988]. Un tel phénomène n'a par contre pas été observé pour PSM [Perez-Vilar, J. et al., 1998b]. Ce sont leurs capacités de polymérisation qui confèrent aux mucines leur viscosité. De fait, la présence des domaines-D et CK est un facteur déterminant qui différencie les mucines formant des gels des autres types de mucines. Il en est ainsi pour MUC7 chez l'homme qui se distingue des autres MUCs codant des mucines sécrétées par le fait que son produit ne forme pas de gel stable. La résistance de ces gels à la solubilisation peut être renforcée par des interactions covalentes non-réductibles entre les monomères. Ces liens supra-moléculaires ont étés observés dans MUC2, mais leur nature exacte n'a pas été encore déterminée [Herrmann, A. et al., 1999].

224. Autres domaines riches en cystéines

De nombreux autres domaines riches en cystéines mais ne présentant aucune homologie avec les domaines D et CK ont été identifiés sur les mucines. En particulier les domaines PTS hautement glycosylés de MUC2, MUC5AC et MUC5B sont interrompus par des régions d'une centaine d'acides aminés contenant 10 cystéines très conservées [Desseyn, J.L. *et al.*, 1997]. De même, la région centrale de la mucine tégumentaire de xénope FIM-B.1 contient de nombreuses régions riches en cystéines appelées SCRs, pour "short consensus repeats" [Joba, W. & Hoffmann, W., 1997], tandis que FIM-A.1 et FIM-C.1 contiennent des domains P, caractérisés par six cystéines conservées. Ces derniers domaines seraient impliqués dans la formation de ponts disulfures intra-moléculaires [De, A. *et al.*, 1994]. Par contre, FIM-A.1 et FIM-C.1 ne possèdent ni domaines D ni domaines CK et ne présentent donc pas de capacité de multimérisation.

225. Domaines spécifiques aux mucines membranaires

En plus du domaine trans-membranaire qui est leur caractéristique majeure, deux domaines sont souvent retrouvés dans les mucines membranaires: un où plusieurs domaines homologues au facteur de croissance épidermique (EGF) et un domaine de type "sea-urchin-sperm-protein-enterokinase-agrin" (SEA). Ces domaines sont disposés de manière très comparable dans toutes les mucines qui les possèdent (Fig. A5). Les domaines homologues à EGF ont été identifiés dans les séquences de MUC3, MUC4, MUC12, MUC13 et MUC17 [Dekker, J. et al., 2002; Williams, S.J. et al., 2001; Gum, J.R. et al., 2002]. Leur fonction exacte au sein des mucines reste obscure, bien que certaines données suggèrent leur implication dans la modulation de la croissance épithéliale [Carraway, K.L. et al., 1999]. Les domaines SEA sont présent à l'extrémité C-terminale, dans la partie extracellulaire de MUC1, MUC3, MUC12, MUC13 et MUC17 [Dekker, J. et al., 1997; Gum, J.R. et al., 2002]. Ils contiennent un site de coupure protéolytique qui permettrait la libération des mucines membranaires sous formes solubles [Wreschner, D.H. et al., 2002]. La séquence de coupure a été déterminée précisément dans MUC1 et des homologues de celle ci ont été retrouvés dans MUC3, MUC12 et MUC17 [Parry, S. et al., 2001; Gum, J.R. et al., 2002].

226. Conclusion

Ainsi, les mucines forment une famille de composés relativement homogène, dont chaque sous-famille présente des motifs structuraux conservés. Néanmoins, leur principale caractéristique consiste en l'existence d'une partie protéique, souvent de grande taille, organisée en unités de répétition et abondamment O-glycosylée. C'est l'importance de cette glycosylation qui confèrerait aux mucines la plupart de leurs propriétés mécaniques et biologiques. Aussi nous allons exposer brièvement les modalités de glycosylation des mucines avant de passer en revue les rôles qui ont été attribués à ces macromolécules.

23. Régulation de la glycosylation des mucines 231. Introduction

Un monomère d'une mucine sécrétée peut contenir plusieurs centaines de O-glycannes qui peuvent ainsi représenter jusqu'a 80% du poids des mucines. Considérant l'impact de la partie glycannique sur la fonctionnalité des mucines, un enjeu majeur depuis plusieurs années a été de révéler l'existence d'hypothétiques mécanismes de régulation de la O-glycosylation des mucines. Plus précisément, les diverses études effectuées ont pour finalité de répondre, entre autres, à ces quelques questions: peut on prédire la position des O-glycannes sur l'apomucine, comment est contrôlée la densité de glycosylation et quels facteurs interviennent dans la sélection des structures O-glycanniques sur chaque mucine?

232. Régulation de l'initiation de la O-glycosylation

Des études ont permis de mettre en évidence un ensemble de séquences consensus pour la majorité des types de glycosylation. L'exemple le plus connu est la N-glycosylation de l'asparagine pour laquelle deux séquences consensuelles, utilisée à différents degrés, ont été décrites: Asn-Xaa-Ser/Thr (où Xaa n'est pas Pro), Asn-Xaa-Cys [Marshall, R.D., 1972; Kasturi, L. et al., 1995]. De même, des séquences consensus ont été décrites pour les glycosaminoglycannes des protéoglycannes [Bourdon, M.A. et al., 1987], les glycannes de l'hydroxylysine Prockop, D.J. et al., 1979], la type collagène sur O-Nacétylglucosaminylation des sérines et thréonines [Haltiwanger, R.S. et al., 1992] ou la Ofucosylation des sérines et thréonines [Harris, R.J. & Spellman, M.W., 1993]. Des efforts considérables ont été développés pour définir des sites peptidiques consensuels pour la Oglycosylation de type mucine chez l'homme, soit en comparant les séquences connues pour intégrer des sérines et thréonines substituées par des O-glycannes de type mucine, soit en testant la spécificité de substrats d'activités GalNAc-transférases à partir de lysats cellulaires [e.g. Elhammer, A.P. et al., 1993; Wang, Y. et al., 1993; Nishimori, I. et al., 1994]. Par contraste avec les autres types de glycosylation cités, de tels motifs n'ont jamais été mis en évidence. De l'ensemble de ces travaux deux hypothèses ont été émises, postulant l'existence

soit d'une GalNAc-transférase présentant une spécificité de substrats très large, soit de nombreuses GalNAc-transférases avec des spécificités de substrats très différentes [Clausen, H & Bennett, E.P., 1996].

Les travaux récents ont démontré que la biosynthèse de O-glycannes de type mucine est une famille de glycosyltransférases appelées UDP-Ninitiée par acetylgalactosamine:polypeptide N-acetylgalactosaminyltransférases (GalNAc-transférases). Bien que toutes n'aient pas encore été caractérisées, douze GalNAc-transférases putatives ont été identifiées chez les mammifères, neuf chez Caenorhabditis elegans et treize chez Drosophila melanogaster [Schwientek, T. et al., 2002]. Chez les mammifères les activités de huit d'entre elles ont été finement caractérisées (ppGaNTase-T1 à -T7 et ppGaNTase-T9). Certaines sont exprimées de façon constitutive (ppGaNTase-T1, -T2 et -T3) tandis que d'autres présentent des profils d'expressions restreints à certains tissus (ppGaNTase-T4, -T5, -T6, -T7 et T-9) [Hagen, K.G. et al., 2001]. Il est à noter que tous les types de cellules animales semblent posséder une activité GalNAc-transférasique pouvant être attribuée à un panel différent de GalNAc-transférases [Clausen, H. et al., 1996].

L'isolement des GalNAc-transférases, couplée au séquençage d'un nombre croissant de gènes de mucines, a permis de tester directement l'activité de chaque membre de la famille des GalNAc-transférases sur des peptides de séquences homologues aux unités de répétition des mucines. Ces travaux confirment l'absence d'un site commun de O-glycosylation, mais tendent a démontrer que toutes les GalNAc-transférases posséderaient des substrats accepteurs spécifiques. Néanmoins, une certaine souplesse existe dans la sélection des substrats. Ainsi, un des premiers travaux de ce type a mis en évidence que les ppGaNTase-T1, -T2 et -T3 peuvent toutes utiliser l'unité de répétition de MUC1 comme accepteur, et transférer une GalNAc sur trois sites de glycosylation sur cinq, mais avec des cinétiques très différentes. Ceci suggère qu'in vivo ces trois enzymes ne glycosyleraient pas les mêmes acides aminés à cause de phénomènes de compétitions [Wandall, H.H. et al., 1997], et qu'elles auraient donc bien des fonctions différentes. A ce titre, l'utilisation d'un peptide dérivé de MUC2 possédant trois thréonines consécutives (PTTTPLK) fut encore plus démonstrative de l'existence de différences d'activité entre ces trois enzymes [Iida, S. et al., 1999]. En effet, les ppGaNTase-T1, -T2 et -T3 transféraient sur ce peptide un maximum de deux, un ou trois résidus de GalNAc respectivement, toujours avec des cinétiques différentes. Ces travaux suggèrent dont l'existence d'une régulation de la glycosylation par la séquence peptidique.

Ce schéma s'est complexifié par la découverte de GalNAc-transférases spécifiques des glycopeptides. En effet, pour ppGaNTase-T4 [Bennett, E.P. et al., 1998], -T6 [Hagen, K.G. et al., 1999], -T7 [Bennett E.P. et al., 1999] et -T9 [Hagen, K.G. et al., 2001], la présence préalable d'un résidu de GalNAc sur le substrat peptidique peut être un pré-requis à leur activité. Là encore, les activités spécifiques de ces enzymes sous-entendent des phénomènes de régulation très fins. Ainsi, l'action différentielle de ppGaNTase-T7 et -T9 sur un même glycopeptide dérivé de MUC1 conduit à des produits dont les sites de glycosylation sont occupés différemment. De même, il apparaît que seule l'action combinée de plusieurs GalNAc-transférases permet l'occupation de tous les sites de glycosylation sur une unité de répétition. Par exemple, les deux sites de glycosylation possibles de l'unité de répétition de MUC1 laissés libres après l'action de ppGaNTase-T1, -T2 et -T3 ne peuvent être glycosylés que par l'action ultérieure de ppGaNTase-T4 [Bennett, E.P. et al., 1998]. La sélectivité de ppGaNTase-T4 pour un substrat glycosylé lui serait conférée par un domaine lectinique présent à son extrémité C-terminale [Hassan, H. et al., 2000]. L'interaction entre ce domaine et une unité de GalNAc déjà présent sur l'axe peptidique est nécessaire à l'activité de l'enzyme. En résumé, le degré et les positions de glycosylation d'une mucine in vivo seraient déterminés par l'action séquentielle d'un grand nombre de GalNAc-transférases distinctes pouvant fonctionner en coopération et vont varier en fonction (i) de la séquence de son domaine PTS (ii) du répertoire de GalNAc-transférases exprimés dans ce même tissu.

Jusqu'à présent, l'apport de chaque activité GalNAc-transférasique à la glycosylation définitive d'une mucine naturelle *in vivo* n'a jamais été déterminée. Une telle étude présuppose de connaître le profil exact des activités GalNAc-transférasiques d'un tissu donné, de pouvoir les discriminer et d'étudier les sites de glycosylation d'une mucine purifiée. En particulier, les données disponibles sur ce dernier point sont encore rares. Deux études identifient les sites de glycosylation des unités de répétition de la forme soluble de MUC1 dans le lait [Müller, S. *et al.*, 1997], et de PSM [Gerken, T.A. *et al.*, 1997]. Il est très intéressant de constater que l'analyse des sites de glycosylation de MUC1 est en contradiction partielle avec les résultats de l'analyse *in vitro* sur peptides synthétiques [Nishimori, I. *et al.*, 1994], ce qui laisse présager l'existence d'autres modes de régulation que la séquence peptidique, et en particulier la taille du domaine PTS qui va influer sur l'accessibilité aux sites de glycosylation et l'influence des glycannes adjacents. Ce dernier point sera brièvement discuté au chapitre suivant. Ces travaux d'analyse *in vivo* n'ont pas encore été corrélés avec l'étude de l'expression

tissu-spécifique des GalNAc-transférases et ne permettent donc pas encore d'avoir une vision globale de la régulation du transfert des O-glycannes sur les mucines.

233. Régulation de la structure des glycannes par l'apomucine

Nous avons précédemment évoqué les différents facteurs contrôlant la structure des Oglycannes de type mucine. Ces facteurs ne sont pas spécifiques aux mucines et concernent la glycosylation de toute glycoprotéine substituée par ce type de glycannes. En fait, là encore, très peu d'études ont concerné l'influence possible de la séquence de l'unité de répétition des mucines sur l'élongation des O-glycannes. Les mesures d'activités in vitro de la β3-Galtransférase, responsable de la synthèse du corps de type 1, sur un ensemble d'unités de répétition de MUC2 plus ou moins glycosylées ont démontré que la séquence du peptide sousjacent, ainsi que la position du GalNAc, influençaient fortement l'allongement du O-glycanne [Granovsky, M. et al., 1994]. De plus, la présence préalable d'autres disaccharides sur le glycopeptide inhibait le transfert de galactose, probablement par encombrement stérique. Ces résultats ont été confirmés par une approche in vivo, en localisant sur l'unité de répétition de PSM les mono-(GalNAc-O-Ser/Thr), di-(Gal-B1,3-GalNAc-O-Ser/Thr) et trisaccharides (Fucα1,2-Gal-β1,3-GalNAc-O-Ser/Thr) [Gerken, T.A., et al., 2002]. Entre autres résultats, l'analyse démontre que l'activité de la ß3-Gal-transférase sur PSM est inversement proportionnelle à la densité locale de sérines et thréonines. Par contre, l'activité de l'a2-Fuctransférase ne montre pas de corrélation avec la séquence locale, mais semble fortement influencée par la nature de l'acide aminé substitué (sérine ou thréonine), ce qui suggère des différences de conformation ou de dynamisme des accepteurs disaccharidiques.

234. Notion de glycoforme

Les données que nous avons précédemment exposées démontrent que la régulation de la glycosylation des mucines est le résultat d'un grand nombre de facteurs et suggèrent que celle-ci est spécifique à chaque tissu, voire à chaque cellule. Cela implique qu'une même mucine devrait présenter divers états de glycosylation en fonction de son origine cellulaire et ainsi être à l'origine de plusieurs glycoformes. La notion de glycoforme est très commune à la majorité des glycoprotéines. Par exemple, la transferrine de rat est présente sous forme de quatre glycoformes facilement purifiées en fonction du profil d'occupation de ses deux sites de glycosylation par des N-glycannes bi- ou tri-antennés [Regoeczi, E., et al., 1987]. Par contre l'identification de glycoformes de mucines in vivo est un exercice très difficile car la densité importante de glycosylation gène l'accès du domaine PTS qui est la partie la plus distinctive des mucines, et de fait, limite l'utilisation d'anticorps spécifique à chaque mucine. De plus, la purification et la déglycosylation complète d'une mucine intacte restent très aléatoire. Malgré ces difficultés, il a été établi que deux isoformes distincts de MUC5B étaient produits dans les tissus bronchiques par les cellules des glandes mucipares [Thornton, D.J., et al., 1996; Thornton, D.J. et al., 1997]. Ces deux composés se différenciaient par leurs mobilités électrophorétiques, ce qui suggérait qu'il étaient chargés différemment et présentaient donc des profils de glycosylation distincts. L'analyse de la composition en sucres des deux mucines a confirmé cette hypothèse. D'autres études ont ensuite suggéré que ces deux glycoformes seraient synthétisés par des populations distinctes de cellules des glandes mucipares, qui exprimeraient des répertoires de glycosyltransférases différents [Wickström, C. et al., 1998]. Au vu du peu de données disponibles, les répercussions physiologiques de la présence de différents glycoformes d'une même mucine au sein d'un même tissu ou de tissus distincts sont inconnues. A titre d'exemple, une étude de la composition du bouchon muqueux isolé d'un individu asthmatique, décédé suite à l'obstruction de ses voies respiratoires, a révélé que seul l'un des deux glycoformes de MUC5B était présent. De plus ce glycoforme était surreprésenté dans le mucus comparé à celui d'un individu non asthmatique et il a été observé qu'il adoptait une morphologie particulière qui le rendait très résistant à la protéolyse [Sheehan, J.K. et al., 1999]. La relation directe de cause à effet entre la proportion anormalement haute de ce composé et les modifications rhéologiques du mucus respiratoire observées chez les patients atteints de mucoviscidose n'a pas été établie. Néanmoins, l'observation est suffisamment troublante pour que l'on s'intéresse de plus près aux répercussions possibles d'une variation de la glycosylation sur les propriétés physicochimiques d'une mucine et donc ses fonctions. Nous allons maintenant passer en revue quelques unes des fonctions attribuées aux mucines, en essayant de mettre en exergue les rôles que jouent les glycannes.

24. Rôles des mucines

241. Introduction.

Ainsi que nous l'avons mentionné précédemment, les fonctions des mucines diffèrent largement en fonction de leur nature sécrétée ou membranaire et sont très dépendantes de leur ultra-structure. D'un coté, la majorité des mucines sécrétées présentent une conformation étendue et se polymérisent pour former des chaînes linéaires (e.g. MUC5B, MUC5AC) ou ramifiées (e.g. PSM) dont la taille peut atteindre 5 µm [Sheehan, J.K, *et al.*, 1987]. Ces mucines sont les composants principaux des mucus qui forment *in vivo* des gels visco-élastiques dont les propriétés rhéologiques exactes varient apparemment en fonction de la nature des mucines qui les composent et de la présence d'autres composés associés à ces mucines tels certains lipides [Carlstedt, I. *et al.*, 1995; Houdret, N., *et al.*, 1986]. D'un autre coté, les mucines membranaires ne se polymérisent pas, mais interviennent dans la formation du glycocalix à la surface de certaines cellules. La présence d'une partie intra-cellulaire les rend susceptibles de transmettre des signaux.

242. Mucines sécrétées

La fonction primordiale des mucines sécrétées en milieu aqueux est liée à la forte viscosité qui leur permet d'assurer une protection mécanique, chimique et bactériologique des épithéliums qu'ils recouvrent. En effet, le mucus recouvre uniformément l'ensemble des tractus gastro-intestinal et uro-génital et forme ainsi la première barrière entre un environnement hostile et l'épithélium fragile. En plus de protéger directement les tissus contre les blessures, le mucus piège de manière non spécifique toute particule indésirable. Ainsi, le mucus et les mucines qui le composent sont essentiels au bon fonctionnement du système muco-ciliaire bronchique qui permet l'évacuation des particules inhalées en les faisant remonter vers les parties supérieures du tractus gastro-intestinal où elles sont éliminées par déglutition [Roussel, P. & Lamblin, G., 1996]. Dans l'estomac la protection mécanique est associée à une protection chimique de la paroi stomacale. La couche muqueuse, en emprisonnant des ions carbonate, permet la formation d'un gradient de pH entre le lumen stomacal à pH 1-2 et la surface de l'épithélium gastrique à pH 6-7, ce qui protège ce dernier de l'auto digestion par les sucs gastriques [Bhaskar, K.R. *et al.*, 1992]. Dans le même ordre

d'idées, le mucus stomacal apporte également une protection physique vis à vis des enzymes hydrolytiques.

Tous les mucus, localisés à l'interface des milieu extérieur et intérieur, sont constamment exposés à des interactions complexes avec les micro-organismes. Si certains tractus (génital, digestif) sont physiologiquement colonisés par un grand nombre d'espèces bactériennes, d'autres comme la partie basse des bronches doivent être rigoureusement stériles. Ainsi, il a été reconnu que les mucines intervenaient dans des phénomènes de reconnaissance spécifiques des micro-organismes, pathogènes ou non. Ces interactions s'établissent en majorité *via* une reconnaissance entre des lectines bactériennes ou virales et des motifs glycanniques précis portés par les mucines (tableau A3). La diversité des épitopes glycanniques portés par les mucines pourrait donc être envisagée comme une mosaïque de sites d'attachement pour les micro-organismes. A titre d'exemple, elle permet au mucus bronchique de piéger efficacement les bactéries inhalées, qui sont ensuite éliminées de la même manière que les particules. Certaines bactéries comme *Pseudomonas aeruginosa* seraient également capables d'interagir avec la partie peptidique des mucines [Reddy, M.S., 1992].

Microorganismes	Structures glycanniques
Entamoeba histolytica	Gal/GalNAc
Candida albicans	$Gal(\beta 1-3)[Fuc(\alpha 1-4)]GlcNAc$
Streptococus pneumoniae	GlcNAc(β1-3)Gal
Actinomyces naeslundii	Gal(β 1-3)GalNAc
Streptococcus sanguis	NeuAc(α2-3)Gal(β1-3)GalNAc
Pseudomonas aeruginosae	GlcNAc(β1-4)Gal(β1-3)GlcNAc(β1-
	$GlcNAc(\beta 1-3)Gal(\beta 1-3)GlcNAc(\beta 1-$
Mycoplasma pneumoniae	NeuAc(α 2-3)Gal(β 1-4)GlcNAc
	HSO3-O-Gal
Influenza virus	NeuAc(α 2-6)Gal
	NeuAc(α 2-3)Gal

Tableau A3: Exemples de structures glycannique présentes dans les mucines humaines reconnues par de microorganismes. D'après Roussel, P. & Lamblin, G., 1996

Une infection intervient quand une colonisation anormale par un micro-organisme passe les défenses normales de l'organisme, dont la barrière muqueuse, et induisent des altérations de la surface cellulaire sous jacente. Ceci peut intervenir quand la barrière muqueuse est inappropriée contre un type donné de micro-organisme, ou qu'elle est endommagée. A titre d'exemple, les virus de type influenza reconnaissent les acides sialiques terminaux de la surface cellulaire du tractus respiratoire grâce à une hémagglutinine virale dont la spécificité varie légèrement d'une espèce à l'autre. Ainsi le virus de l'influenza humain est spécifique du motif Sia(α 2-6)Gal, tandis que l'équivalent aviaire est spécifique du motif Sia(\alpha2-3)Gal [Rogers, G.N. & D'Souza, B.L., 1989]. Les motifs Sia(\alpha2-6)Gal et Sia(\alpha2-3)Gal sont respectivement exprimés chez l'homme sur les cellules épithéliales ciliées bronchiques et dans les mucines bronchiques. De fait les mucines bronchiques captent très efficacement le virus influenza aviaire, mais sont totalement inefficaces contre l'équivalent humain [Couceiro, J.N. et al., 1993]. Le cas le plus étudié d'altération des muqueuses entraînant une pathologie sérieuse est la mucoviscidose. Suite à une mutation (550 mutations différentes ont été identifiées) d'un canal au chlore de la membrane apicale des épithéliums sécréteurs, un important déséquilibre hydrique intervient dans le mucus bronchique. Celui ci devenant trop visqueux, le système muco-cilliare est incapable d'éliminer correctement le mucus qui s'accumule dans les bronches [Roussel, P. & Lamblin, G., 1996]. Du fait de la présence des motifs potentiellement reconnus par de nombreux micro-organismes, le mucus stagnant devient un foyer de propagation de bactéries pathogènes qui ne sont pas éliminés. La majorité de la morbidité et de la mortalité chez les patients est le fait d'infections récurrentes par Pseudomonas aeruginosa. De plus, des phénomènes secondaires encore inconnus modifient la glycosylation des mucines des patients atteints de la mucoviscidose, ce qui augmenterait leur affinité pour P. aeruginosa [Scharfman, A. et al., 2000]. D'autres pathologies ont été associées aux capacités d'adhésion des micro-organismes aux mucines. On peut citer les ulcères de l'estomac déclenchés suite à la colonisation du mucus gastrique par Helicobacter pylori, ou différentes pathologies liées à l'infection des intestins.

243. Mucines membranaires

Relativement peu d'informations sont disponibles quand aux fonctions associées aux mucines membranaires. MUC1 est le composé ayant reçu le plus d'attention car il a été reconnu très tôt comme un marqueur de tumorisation. MUC1 est exprimé sur la surface apicale des cellules sécrétrices de nombreux épithéliums. Ces molécules, de par leur conformation et leur taille, s'étendent jusqu'a 200 nm de la surface cellulaire [Taylor-Papadimitriou, J. *et al.*, 1999] et possèdent probablement des fonctions protectrices envers l'épithélium sous-jacent. De plus, comme leur taille excède de 5 à dix fois la taille de la plupart des autres molécules membranaires, et considérant leur densité importante, elles joueraient un rôle dans la régulation de l'adhésion cellulaire en masquant les molécules

d'adhésion cellulaire. Néanmoins, comme dans le cas des mucines sécrétées, MUC1 présente des épitopes glucidiques reconnus par nombre de molécules d'adhésions telle que ICAM1 sur les cellules endothéliales, ce qui suggère un rôle potentiel dans l'intravasation [Regimbald, H.L. *et al.*, 1996]. De même MUC1 est un ligand de la sialoadhésine des macrophages et pourrait ainsi être impliqué dans le recrutement des macrophages [Nath, A. *et al.*, 1999]. Enfin, la partie cytoplasmique de MUC1 interagit avec des protéines intracellulaires telles que la β -caténine [Li, Y. *et al.*, 1998] et le récepteur de l'EGF [Schroeder, J.A. *et al.*, 2001], et contient des sites potentiels de phosphorylation impliqués dans la transduction de signaux cellulaires [Pandey, P. *et al.*, 1995].

Dans plusieurs types de cancer, il est apparu que l'expression et la glycosylation de MUC1 étaient fortement altérées. En particulier, la densité de glycosylation de l'unité de répétition de MUC1 passe d'une moyenne de 2,6 sur 5 sites potentiels pour des cellules de la glande mammaire normale [Müller, S. et al., 1997] à 4,8 sites pour des cellules cancéreuses [Müller, S., et al., 1999]. De plus, la structure des O-glycannes de MUC1 est modifiée dans les lignées cancéreuses: ils sont plus courts et basés sur le corps de type-1 et non sur le corps de trisaccharide type-2, ce aui conduit à une sur-expression du NeuAc($\alpha 2,6$)Gal($\beta 1,3$)GalNAc (sialyl T) absent des cellules normales [Lloyd, K.O. et al., 1996] et à une perte d'expression des épitopes Le^a, Le^x ou sialyl Le^x. Dernièrement, il a été démontré que la modification de la structure des O-glycannes serait le résultat de la surexpression de la a3sialyl: galactosyle-transférase-I qui dévierait leur synthèse vers le motif sialyl T et empêcherait l'élongation normale des O-glycannes en entrant en compétition avec la ß6GlcNAc: N-acétylgalactosaminyle-transférase-I [Dalziel, M. et al., 2001]. Enfin, ces changements s'accompagnent d'une dérégulation totale de l'expression de MUC1 dont la localisation est étendue à l'ensemble de la surface des cellules. L'augmentation importante de la charge négative à la surface des cellules suite à la surexpression du sialyl T va inhiber leur capacités d'adhésion avec la matrice extracellulaire et induire des forces de répulsion entre les cellules [Wesseling, J. et al., 1995]. De plus, les changements dans le profil de glycosylation vont totalement modifier l'ensemble des interactions spécifiquement dépendantes des épitopes glycanniques exprimés à la surface de MUC1. Bien qu'il ait été reporté que la progression des tumeurs étaient ralentie chez des souris mutantes Muc1 [Spicer, A.P. et al., 1995], l'implication exacte des modifications de MUC1 dans la progression métastatique est encore incertaine.

244. Conclusion

Comme il a été décrit précédemment, la glycosylation des mucines exerce son influence de deux manières fondamentales qui présentent des degrés de spécificités probablement différents. Tout d'abord, c'est elle qui confère aux mucines leur conformation et leur dynamique moléculaire à la base de leurs propriétés physico-chimiques. Il est intéressant de noter que des études de modélisations moléculaires en cours (Sheehan, J.K., résultats non publiés) tendent à démontrer que les séquences polypeptidiques des régions PTS elles mêmes suffiraient à conférer à la mucine sa conformation étendue. La haute densité de glycosylation renforcerait ensuite la rigidité de la molécule et lui donnerait sa dynamique très particulière. Loin de relativiser l'importance de la glycosylation sur les mucines, cette étude met en évidence un aspect synergique encore inconnu entre les parties protéiques et glycanniques dans l'adoption de la conformation des mucines. Les glycannes exercent également leur activité au travers d'interactions très spécifiques avec des molécules d'adhésions endogènes et exogènes. Il parait donc que les changements dans le profil de glycosylation d'une mucine isolée dans un cadre physiologique (polymorphisme intra-spécifique, glycoformes, réponse au parasitisme) ou pathologique vont se répercuter différemment sur ces deux types d'activités. La moindre modification dans l'expression d'un épitope glucidique terminal va avoir des répercutions importantes dans la capacité d'interaction des mucines tandis qu'elle a toute les chances de ne pas changer leur conformation. D'une manière générale, le contrôle des activités biologiques liées à la conformation des mucines par leur glycosylation présente un aspect beaucoup plus permissif que le contrôle des activités liées aux interactions moléculaires, et a encore été peu étudié. D'un point de vue évolutif, on peut arguer que cette contrebalancer des permissivité permet de l'extrême variabilité activités glycosyltransférasiques observées dans le règne animal. Ainsi, des mucines substituées par des O-glycannes largement différents vont conserver leurs propriétés rhéologiques et leurs fonctions protectrices essentielles.
II. PROBLEMATIQUE

Notre équipe de recherche est impliquée depuis de nombreuses années dans l'étude de la diversité structurale des O-glycannes dans le règne animal. Ainsi que nous l'avons détaillé dans le chapitre consacré à la structure des O-glycannes, ces études systématiques, en conjonction avec les travaux d'autres équipes, ont permis de mettre en évidences deux faits majeurs. Premièrement, la diversité structurale des O-glycannes semble sans limite. Deuxièmement, la structure des O-glycannes est spécifique à chaque espèce. 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 pouvaient en découler.

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 sur les dix huit 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. 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(α 1-4)Kdn peut être à son tour fucosylé, tandis que les fucoses présents dans les motifs Fuc(α 1-2)Gal et Fuc(α 1-3)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: α 1,2-, α 1,3- et α 1,4-fucosyltransférases ayant pour substrat le di-N-acétylchitobiose GlcNAc(\beta1-4)GlcNAc ou le motif Gal(\beta1-3/4)GlcNAc [Oriol, R. et al., 1999].



Fig. A7: a) Modèle hypothétique d'évolution divergente des principaux gènes de fucosyltransférases. b) Substrats accepteurs communs des FUT1 à FUT8. D'après Oriol, R. *et al.*, 1999.

37

Ces fucosyltransférases résulteraient en grande partie d'une évolution divergente à partir d'un ou de deux ancêtres communs, comme représenté dans la figure A7. 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 inconnu. Bien que très spécifique, l'activité des glycosyltransférases semble très sensible aux mutations. Ainsi, il a été montré qu'une mutation unique (Trp¹¹¹ \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 qu'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 aue 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(\beta1-6)GalNAc dans la mucine sous-maxilaire 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.

Nos travaux ont consisté en l'étude de la structure des O-glycannes isolés d'espèces animales génétiquement définies. Ils ont pour but 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. Au sein de la thématique d'étude des amphibiens, nous nous sommes intéressés aux espèces *Xenopus laevis* et *Xenopus tropicalis*. Elles ont ainsi permis de mieux définir la nature biochimique des composants de la gangue oviducale des amphibiens et mis en évidence l'existence de polymorphismes intra- et inter-spécifiques portant à la fois sur la partie glycannique et protéique des glycoprotéines. Une nouvelle série de travaux a été initiée sur ce point précis. En marge du modèle amphibien, nous avons également procédé à l'étude du profil de glycosylation exprimé par le nématode *Caenorhabditis elegans*. Ces études ont également porté sur l'analyse des glycoprotéines porteuses des O-glycannes.

Nous présenterons successivement les modèles d'études et les résultats obtenus, en commençant par les amphibiens.

III. RESULTATS

1. Le modèle amphibien

11. Description du modèle

111. Introduction

Le matériel de base utilisé pour l'étude de la diversité des O-glycannes chez les amphibiens est la gangue entourant les œufs, appelée gangue oviducale. Cette gangue gélatineuse représente la couche la plus externe de la matrice extracellulaire des œufs d'amphibiens. Dans le cadre d'une étude structurale, ce matériel présente beaucoup d'avantages: il est disponible en quantité très importante et est facilement collecté, sa collecte ne nécessite pas de sacrifier l'animal, sa production est inductible en quelques jours par injection d'hormones pituitaires, il est très riche en O-glycannes et les glycoprotéines porteuses de O-glycannes sont facilement purifiées. Ainsi, après traitement, une ponte de xénope permet l'obtention d'environ un gramme d'un mélange de glycoprotéines, qui fournissent typiquement une centaine de milligrammes de O-glycannes de type mucine. Avant d'exposer les résultats de nos études, nous présenterons rapidement la structure de la matrice extracellulaire de l'œuf de xénope et ses fonctions au cours de la fécondation.

112. Structure de la matrice extracellulaire

La matrice extracellulaire qui entoure les œufs de xénopes est composée de plusieurs gangues gélatineuses, d'une enveloppe et d'un espace périvitellin séparant l'enveloppe de la membrane plasmique (Fig. A8). Ces gangues offrent aux oeufs une protection mécanique et un micro-environnement propice au développement de l'embryon. Elles servent aussi de réserve de cations divalents (Ca^{2+} , Mg^{2+}) indispensables à la pénétration des spermatozoïdes.



Fig. A8. La matrice extracellulaire de Xenopus laevis.

Les œufs de tous les vertébrés et invertébrés, sont entourés d'une enveloppe dont l'ultrastructure et la composition semblent similaires [Dumont, J.N. *et al.*, 1985]. Elle est composée d'un matériel protéique fibreux mal défini, imperméable aux spermatozoïdes . Chez *Xenopus laevis*, cette enveloppe peut prendre quatre formes différentes en fonction de l'état de développement de l'œuf et de sa localisation dans l'oviducte (tableau A4) [Hedricks, J.L. *et al.*, 1991]. Ces formes diffèrent par leur ultrastructure, leur composition et leur fonction. La conversion d'une enveloppe à l'autre est induite par des facteurs sécrétés par l'oviducte avant la fertilisation et d'autres exocytés par l'œuf après l'entrée du spermatozoïde.

Enveloppes	Etat de différentiation l'œuf
Ovarienne	Oocyte dans l'ovaire
Cœlomique	Œuf ovulé dans le coelome
Vitelline	Œuf pondu
De fertilisation	Œuf fertilisé

Tableau A4: Différentes formes de l'enveloppe

Les oocytes sont recouverts, au cours de leur transit dans l'oviducte, des produits de sécrétion des segments successifs de ce canal. Les gangues oviducales sont ajoutées dans la seconde partie de l'oviducte, appelée *pars convulata*, elle même divisée en quatre segments histologiquement distincts.

La pars convulata de Xenopus laevis sécrète trois couches de gangues gélatineuses (de J1, la plus interne, à J3) qui sont séquentiellement enroulées autour de chaque œuf au fur et à mesure de leur transport au travers de l'oviducte [Yoshizaki, N., 1985]. Chaque couche consiste en une matrice stable, composée en majorité de glycoprotéines fibreuses de très haut poids moléculaire, à laquelle sont associées de nombreuses protéines globulaires. [Bonnel, B.S. et al., 1996]. Au moins cinq glycoprotéines ont été observées, dont les masses apparentes en gel de polyacrylamide varient entre 450 et 900 kDa. La nature exacte de ces composés n'a jamais été précisément définie, néanmoins leurs propriétés physico-chimiques sont tout à fait comparables à celles des mucines, telles qu'elles ont été définies au chapitre précédent. Ces propriétés incluent: leur très haut poids moléculaire, leur conformation étendue et rigide [Bonnel, B.S. et al., 1996] et la haute teneur en O-glycannes de la gangue (60% du poids total des gangues) [Yurewicz, E.C. et al., 1975]. De plus, la très forte viscosité des gangues et leur insolubilité est compatible avec la présence de mucines sécrétées polymérisées.



Fig. A9: Représentation schématique des étapes de la fécondation chez le xénope: 1, pénétration du spermatozoïde dans la gangue gélatineuse de l'oeuf; 2, réactions de capacitation du spermatozoïde (incluant le déclenchement de la réaction acrosomienne) au contact des composants de la gangue; 3, fixation du spermatozoïde capacité aux récepteurs de l'enveloppe vitelline; 4, pénétration du spermatozoïde au travers de l'enveloppe vitelline; 5, formation de l'enveloppe de fertilisation empéchant la polyspermie, suite à l'exocytose de lectines des granules corticaux (CGL).

Le fait que les gangues ne se solubilisent qu'en présence d'un agent réducteur tel que le dithiotréitol ou le β -mercapto-éthanol suggère la présence de ponts disulfures stabilisant l'ultrastrusture du gel, ce qui est également en faveur de la présence de mucines sécrétées. L'étude du mucus tégumentaire de *X. laevis* a révélé la présence d'une mucine sécrétée pouvant former des gels, FIM-B.1 [Probst, J.C. *et al.*, 1990], ce qui rend plausible l'expression de tels composés par l'oviducte. L'analyse structurale des gangues d'amphibiens au laboratoire n'a de plus jamais révélée l'existence de protéoglycannes.

A cette matrice sont associées de nombreuses protéines globulaires de masse moléculaire inférieure à 180 kDa. Certaines d'entre elles sont fortement associées à la matrice et ne diffusent jamais hors de la gangue, tandis que d'autres diffusent librement dans le milieu une fois la gangue hydratée [Bonnel, B.S. *et al.*, 1996].

113. Fonctions de la matrice extracellulaire au cours de la fécondation

La fertilisation de l'œuf chez Xenopus laevis se déroule selon un processus remarquablement conservé au cours de l'évolution : le spermatozoïde est attiré vers l'œuf grâce à des substances chémoattractantes [Al-anzi, B. et al., 1998], pénètre la gangue, est activé au contact de celle ci, se lie à l'enveloppe vitelline qu'il finit par traverser et fusionne avec la membrane plasmique (Fig. A9) [Vacquier, V.D., 1998].

Au cours de la traversée de la matrice extracellulaire de l'œuf, le spermatozoïde va subir une série de transformations indispensables à son activité fertilisatrice (capacitation du spermatozoïde), dont la mieux connue est la réaction acrosomienne. Cette dernière se traduit par l'exocytose de nombreux composés suite à la fusion de la vésicule acrosomienne avec la membrane apicale du spermatozoïde. Parmi ces composés se trouvent des enzymes à activité protéolytique qui ont pour fonction de digérer l'enveloppe vitelline imperméable aux spermatozoïdes. La réaction acrosomienne est induite à proximité de l'enveloppe vitelline, dans la couche J3, sous l'action d'un produit de sécrétion de la partie supérieure de l'oviducte, appelée *pars recta* [Katagiri, C., 1987]. Le facteur déclenchant la réaction la réaction acrosomienne chez *Xenopus laevis* n'est pas connu. Chez les oursins, ce facteur a été reconnu comme un polymère de fucose sulfaté dont la structure et l'activité biologique sont strictement spécifiques à l'espèce [Alves, A.P. *et al.*, 1997], tandis que ce rôle est tenu par la glycoprotéine ZP3 chez la souris [Bleil, J.D. & Wasserman, P.M., 1983]. D'une manière générale la nature des composants de la gangue de *X. laevis* impliqués dans la capacitation des

spermatozoïdes est encore très controversée. Pour certains auteurs, ce rôle est uniquement dévolu aux protéines diffusibles associées à la matrice de la gangue [Olson, J.H. & Chandler, D.E., 1999], tandis que pour d'autres c'est la fraction macromoléculaire de la gangue qui joue un rôle prépondérant [Mozingo, N.M. & Hedrick, J.L., 1999].

Le spermatozoïde capacité se fixe à la membrane vitelline par l'intermédiaire d'un récepteur spécifique. Deux glycoprotéines, gp69 et gp64, jouent ce rôle chez *Xenopus laevis* [Tian, J. *et al.*, 1997]. Ce sont deux glycoformes d'une même protéine de 54 kDa homologue de la protéine ZP2 de la *zona pellucida* de l'œuf de souris [Tian, J. *et al.*, 1999]. Il a été montré que la déglycosylation de son récepteur inhibait la fixation du spermatozoïde, suggérant que les glycannes jouent un rôle majeur dans ces interactions. Bien que possédant des profils de O-glycosylation distincts, gp69 et gp64 semblent avoir une activité biologique identique [Tian, J. *et al.*, 1997a]. L'intérêt biologique d'un tel polymorphisme n'a pas encore été élucidé.

Suite à sa fixation sur son récepteur, le spermatozoïde fécondant traverse l'enveloppe vitelline localement digérée sous l'action des protéases libérées de l'acrosome et fusionne avec la membrane plasmique de l'œuf.

12. Etude des O-glycannes isolés de la gangue oviducale de Xenopus laevis 121. Présentation des résultats

Lorsque nous avons commencé cette étude, la structure des O-glycannes des gangues de *Xenopus laevis* avait déjà été partiellement déterminée au laboratoire [Strecker, et al., 1995]. *Xenopus laevis* étant un modèle animal communément utilisé pour de nombreuses applications biologiques, tout particulièrement en embryologie, cette espèce a été l'une des premières a être étudiées au laboratoire. Les structures de neufs O-glycannes neutres appartenant à deux séries distinctes, dont les archétypes sont présentés ci dessous, ont ainsi été déterminées.



Des résultats préliminaires avaient également suggéré l'existence d'un polymorphisme glycannique intra-spécifique chez X. laevis. Nos travaux ont donc consisté à mettre en évidence l'existence d'un tel polymorphisme intra-spécifique chez Xenopus laevis, et à en déterminer l'ampleur. En effet, s'il apparaissait que l'ampleur de la variabilité intra-spécifique était plus importante que celle de la variabilité inter-spécifique, ces résultats remettraient en cause la notion de spécificité d'espèce des O-glycannes. De plus, dans l'optique d'une étude des glycosyltransférases exprimées chez Xenopus laevis il était important d'obtenir des informations sur l'homogénéité de la répartition des activités glycosyltransférasiques au sein d'une population de Xenopus laevis.

Pour cette étude, nous avons donc comparé les profils de O-glycosylation individuels des gangues isolées de six spécimens de *Xenopus laevis* provenant du même élevage. Les O-glycannes ont été libérés de préparations de mucines de gangues par β -élimination réductive, isolés par HPLC et analysés par une combinaison de RMN, spectrométrie de masse et chromatographie en phase gazeuse. Le profil de glycosylation de chaque spécimen a ensuite été confirmé par l'utilisation de lectines spécifiques au motifs glycanniques présents sur les O-glycannes. Ces études ont également permis de préciser la nature des glycoprotéines porteuses des O-glycannes. Les résultats de cette étude sont présentés sous forme d'un article publié (article 1)

122. Résultats

Article 1: O-glycan variability of egg-jelly mucins from Xenopus laevis: characterization of four phenotypes that differ by the terminal glycosylation of their mucins.

Guérardel, Y., Kol, O., Maes, E. Lefebvre, T., Boilly, B., Davril, M. and Strecker, G.

O-glycan variability of egg-jelly mucins from *Xenopus laevis*: characterization of four phenotypes that differ by the terminal glycosylation of their mucins

Yann GUERARDEL*, Ossarath KOL*, Emmanuel MAES*, Tony LEFEBVRE*, Bénoni BOILLY†, Monique DAVRIL‡ and Gérard STRECKER*¹

*Laboratoire de Chimie Biologique et Unité Mixte de Recherche du CNRS 8576, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq, France, †Laboratoire de Biologie du Développement, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq, France, and ‡Unité INSERM no. 377, Place de Verdun, 59045 Lille Cedex, France

46

Eggs from Xenopus laevis are surrounded by several layers of jelly that are needed for proper fertilization. Jelly coat is composed of high-molecular-mass glycoconjugates to which are bound many globular proteins. O-glycans released from the jelly coat of X. laevis have been partially described in previous studies. In this study, we compared the glycosylation pattern of the egg jelly coat isolated from six specimens of X. laevis. The O-glycans were released from jelly coats by alkali/borohydride treatment. Structural characterization was performed through a combination of one- and two-dimensional ¹H-NMR and methylation analysis. This allowed the description of a new family of sulphated O-glycans present in jelly coats of all X. laevis. However, the jelly O-glycans showed a low extent of polymorphism be-

tween specimens. This intra-specific variability was restricted to the terminal substitution of O-linked oligosaccharides. The differential expression of two glycosyltransferase [an α -(1 \rightarrow 4) galactosyltransferase and an α -(1 \rightarrow 3) fucosyltransferase] activities resulted in the characterization of four phenotypes of X. laevis. Furthermore, electrophoretic analysis suggested that the high-molecular-mass fraction of jelly coat was mostly composed of mucin-type glycoproteins. Blot analysis with lectins confirmed that the glycan variability was borne by these mucin-type components. However, fertilization assays suggested that the glycan polymorphism had no repercussion on egg fertilizability.

Key words: amphibian, fertilization, glycosyltransferase, NMR.

INTRODUCTION

Xenopus laevis eggs are surrounded by a complex extracellular matrix that consists of a vitelline envelope and a thick jelly coat. The jelly coat contains three concentric layers (J1 to J3) which are successively deposited around the egg as it passes through the oviduct [1]. This matrix provides a protective environment for the developing embryo and is involved in many specific interactions such as, sperm binding to the vitelline envelope [2], block to polyspermy [3] and selection of sperm [4].

Carbohydrates themselves are involved in many of these interactions. Indeed, oligosaccharide chains of sperm receptor (gp69/64) seem to be directly involved in the binding function of the protein [2]. Also, the block to polyspermy is in part due to the agglutination of jelly-coat components and a cortical granule lectin (CGL) [3]. The CGL is released from the egg just after sperm penetration and binds to the carbohydrate moieties of two jelly-coat glycoproteins [5]. Thus it is suggested that the ability of sperm to traverse the jelly coat is dependent on interactions with the carbohydrate moieties of egg-jelly components [6].

If jelly coat has been recognized for a long time to be essential for fertilization, little is known about the macromolecules involved in the sperm-egg interaction. The jelly coat is composed of high-molecular-mass glycoproteins that act as a scaffold to which low-molecular-mass proteins, diffusible or not, are bound [7]. However, the respective part these two types of components play in fertilization is still controversial [8,9]. The exact nature of the high-molecular-mass fraction is still unclear. However, the structural and chemical characteristics of the jelly coat suggest that this fraction is mostly constituted by mucin-type glycoproteins. They include: the high content, up to 60%, of sugar in jelly coat [10], the very high-molecular-mass and the stiff, extended conformation of these constituents [7], the release of high quantities of O-linked mucin-type oligosaccharides after alkaline treatment [11,12] and the absence of uronic acids which are the main constituents of glycosaminoglycans, except keratan sulphates [11]. Furthermore, three mucin-type glycoproteins [frog integumentary mucin (FIM) A1, FIM B1 and FIM C1] have already been described in the X. *laevis* skin [13–15], one of which has the same structure as human mucin MUC 2 [16].

The analysis of X. laevis egg-jelly-coat O-glycans identifies 19 different neutral oligosaccharides, classified into two families [11,12]. These glycans appear characteristic of the X. laevis species. Indeed, studies of jelly-coat O-glycans from 15 amphibian species demonstrated two main points. First, these components showed a remarkable heterogeneity of structures, more than 300 different O-glycans were identified and approx. 20 new glycosyltransferase activities were demonstrated. Some of these glycans present ubiquitous epitopes such as the Lewis X determinant in jelly coat of the urodele Axololtl maculatum [17] or the A-Lewis Y epitope in the urodele Pleurodeles waltl [18], while others show non-predictable sequences such as Fuc(α 1-5)[Fuc(α 1-4)]Kdn(α 2-3/6) (where Kdn stands for 3deoxy-D-glycero-D-galacto-nonulosic acid) found in Ambystoma tigrinum [19]. Second, it appeared that these oligosaccharides were highly specific to each amphibian species. Each species analysed showed one to three families of structurally related O-glycans. In most cases, the species specificity is established through the

Abbreviations used: CGL, cortical granule lectin; GC, gas chromatography; GalNAc-ol, N-acetylgalactosaminitol; PAS, periodate-Schiff; BSI-B₄, Bandeiraea simplicifolia isoform B₄; Hm, Mm and Lm, high-, medium- and low-mobility glycoproteins; FIM, frog integumentary mucin.

¹ To whom correspondence should be addressed (e-mail Gerard.Strecker@univ-lille1.fr).



Figure 1 HPLC profiles of oligosaccharide alditols released from the egg jelly of six specimens of *X. laevis* on an amine-bonded silica column

Specimens: 1 (a), 2 (b), 3 (c), 5 (d), 4 (e) and 6 (f). All the designated peaks have been purified and the structure of the corresponding O-glycans analysed. Peaks containing multiple compounds were recycled on a reverse-phase column.

presence of one or a mixture of several original oligosaccharide sequences. X. laevis species was characterized by the presence of a blood group A epitope substituted by a GlcNAc residue in an α -(1 \rightarrow 3) linkage on one out of the two O-glycans families, while the other family presents the ubiquitous sequence in amphibians Fuc(α 1-2)Gal(β 1-3)[Fuc(α 1-2)]Gal(β 1-3).

The fact that these structural characteristics were found in X. laevis from different origins [11,12] is consistent with the hypothesis of a glycan species-specificity. However, such a species specificity is not incompatible with an intra-specific variability such as the 'blood group' polymorphism found in human. If it occurred, such a structural variability of the jelly-coat components might be of importance in the studies of their respective roles during fertilization. Also, if it appeared that this variability was to be of a higher extent than the one observed between species, it would challenge the hypothesis of a species specificity of O-glycosylation. Here, we report that the glycosylation variability of jelly-coat mucins in X. laevis is restricted to the terminal substitution of O-glycans and involves two distinct glycosyltransferase activities.

EXPERIMENTAL

Sampling of eggs

Eggs from X. laevis were obtained from spawnings induced by injection of 500 units of human chorionic gonadotrophin. Eggjelly coats from the same clutch were extracted into Dulbecco's phosphate buffered saline (Sigma) containing 10 mM EDTA, 1 mM PMSF and 0.5% 2-mercaptoethanol at 4 °C overnight. The mixture was centrifuged and the supernatant was then dialysed for 72 h against water and finally freeze dried.

Isolation of oligosaccharide alditols

The material was submitted to alkaline reductive degradation in 100 mM NaOH containing 1 M NaBH₄ at 37 °C for 72 h. The reaction was stopped by the addition of DOWEX 50 × 8 (25–50 mesh, H⁺ form) at 4 °C until pH 6.5, and after evaporation to dryness, boric acid was distilled as the methyl ester in the presence of methanol. Total material was submitted to a cationic exchange chromatography on DOWEX 50 × 2 (200–400 mesh, H⁺ form) to remove residual peptides. The oligosaccharide fraction was then purified on a Bio-Gel P2 column (Bio-Rad).

Fractionation of oligosaccharide alditols

Compounds were fractionated by HPLC on a primary aminebonded silica column (SupelcosylTM, LC-NH₂, 4.6 mm × 250 mm, Supelco Inc., Bellefonte, PA, U.S.A.) using a mixture of acetonitrile/30 mM KH₂PO₄/water (75:0:25 to 50:50:0, by vol., in 60 min) with a flow rate of 1 ml/min. Oligosaccharides were detected by UV spectroscopy at 206 nm using an LDC variablewavelength detector (Spectra Monitor D, Milton Roy, Riviera Beach, FL, U.S.A.) connected to a Spectra-Physics Model 4100 computing integrator. If necessary, fractions were recycled on a 5 μ m ODS Zorbax column (4.4 mm × 250 mm, DuPont Ins., Paris, TX, U.S.A.) using a mixture of water/acetonitrile (99:1, v/v).

PAGE

Dry egg jellies were solubilized in sample buffer (pH 6.8) containing 0.15 M Tris/HCl, 20 % (v/v) glycerol, 4% (w/v) SDS and 1% 2-mercaptoethanol [20]. Samples were loaded on to a gradient 2.5–10% polyacrylamide/SDS gel and run at 5 mA, overnight. For separation of proteins by two-dimensional gel electrophoresis, a mixture of 2.5–4 and 4–6 pH range ampholyte (Biochemika) was used in the first dimension. Gels were stained for carbohydrates using the periodate–Schiff (PAS) method [21]. For Western blotting analysis, proteins were electro-transferred to nitrocellulose sheets and probed with horseradish-peroxidase-labelled *Bandeiraea simplicifolia* isoform B_4 (BSI- B_4) and *Tetragonolobus purpureas* lectins (Sigma) at a dilution of 1:2000. Staining was performed using an ECL[®] detection kit (Amersham).

Agarose gel electrophoresis

Aliquots of the purified mucins (400 μ g) were subjected to agarose gel electrophoresis in veronal buffer, pH 8.2, as described previously [22]. Slides were stained for carbohydrates with PAS reagent and for acidic components with Toluidine Blue.

Enzymic digestions

Samples were dissolved in 200 mM Tris/acetate buffer, pH 7.5, containing 2 mM CaCl₂, and the following enzymes were added: hyaluronate lyase from *Streptomyces hyalurolyticus* (type IX, Sigma), chondroitinase ABC from *Proteus vulgaris* (Seikagaku, Tokyo), heparinase III from *Flavobacterium heparinum* (Sigma) in the amounts 100, 50 and 50 m-units/mg of material respectively. The mixture was stirred overnight at 37 °C and

Table 1	Structures of	major	O-glycans	isolated	from	six	specimens	of	Xenopus	laevis
---------	---------------	-------	------------------	----------	------	-----	-----------	----	---------	--------

Each O-glycan (A to W) is attributed to the specimens from which it has been isolated (+).

O-glycan structures		1	2	3	4	5	6	Ī	O-glycan structures		1	2	3	4	5	6
GicNAc(61-6) Gal(61-8) Fuc(a1-2)	A	+	+	+	+	. 4	÷ 4		NeuAc(a2-6) Gal(\$1-5) Gal(\$1-5) Fuc(a1-2) Fuc(a1-2)	м	+	+	+		+	+
Fuc(a 1-3) GlcNAc(β1-6) GalNAc-ol Gal(β1-3) Fuc(a 1-2)	B				+	•	4	-	GicNAc(B1-5), Gal(B1-3), Gal(B1-3), Gal(B1-3), Fuc(a1-2) Pus(a1-2)	N	+	+				
Gal(\$1-3) Gal(\$1-3) Fuc(a1-2) Gal(\$1-3)	с				+	•	4		Gal-ol GalNAc(a1-3) GleNAc(a1-3) Fuc(a1-2)	0	+	+	+		+	
GłcNAcβ1-6) GalNAc-ol Gal(β1-3) Duccal 9	D	+	+	-	4	-	-	۰	Qal-ol OnINAc(a 1-3) Fuc(a 1-3) Fuc(a 1-3)	P						+
Fuc(a1-2) Fuc(a1-3) Gal(β1-3) Gal(β1-3) Fuc(a1-2)	E				4	F	•	+	GalNAc-ol Cal(B1-8) GleNAc(a1-3) GleNAc(a1-3) Fuc(a1-2)	Q	+	+	+		+	+
GalNAc-ol Gal(α1-4) Gal(β1-3) / ↓ Fuc(α1-2)	F	+	•	• •	F		+	+	HSO.(6) GleNAc(61-6) GalNAc-ol GalNAc(01-3) (Ib Marco 1-3) Particular (Ib Marco 1-3)	R	+	+				
GlcNAc(\$1-6) GalNAc-ol Gal(c1-4)Gal(\$1-3) Bue(c1-2)	G	+		+ -	ł		+	+	HSQ.(6) (GLNAc(51-6) (GLNAc(51-6) (GLNAc(51-6) (GLNAc(51-6) (GLNAc(51-6) (GLNAc(51-6)) (GLNAC(51-6))	S						+
Fuc(α1-3) Fuc(α1-3) Gal(α1-4)Gal(β1-3) Fuc(α1-2) Fuc(α1-2)	н							+	Peoral-1-81 HSC, (0) SO, HSS Gal(01-4)—Old:VAc(01-6) Gal(01-5) Gal(01-5) Gal(01-5) Cal(01-5) Rev(01-5)	T	+	+	· -		+	
Galβ1-3) Galβ1-3) Galβ1-3) Galβ1-3) Galβ1-3) Galβ1-3)	I	٩	+ -	+	÷		+	+	(Herry Article 3) 1900(0) 0401-0-042(Aup) (0) 1900(0) 04010-0 04204-0 0400-0	U				+	-	+
GleNAc(\$1-6) Gal(\$1-3) Gal(\$1-3) Fuc(a1-2) Fuc(a1-2)	J	•	÷	+	+	+	+	+	HBQ(0) (abit-o-Gables).0 HBQ(0) Factors Gables, Gables, Gab	v				-	ŀ	÷
Gal(β1-3) Fuc(α1-3) Gal(β1-3) Fuc(α1-2) Fuc(α1-2)	K	•				+		+	HBC469 CallNAc(c1-5)-Cal(01-6)-CheNAc(01-6) CallNAc(c1-5) CallNA(c1-5) Call	w	' 1	- 4		F		
GicNAc(β1-6) Gal(β1-3) Gal(α1-4)Gal(β1-3) Fuc(α1-2) Fuc(α1-2)	I		+	+	+		+	+								

freeze dried before being subjected to isopycnic ultracentrifugation.

Alternatively, freeze-dried material was subjected to mild Pronase (Sigma) digestion at 37 °C in 0.01 M calcium acetate buffer, pH 7, at an enzyme ratio of 1:40 (w/w) for 24 h. The glycopeptides were then purified on a Bio-Gel P6 (Bio-Rad) column.

Disulphite reduction and alkylation

Dry material was solubilized in 6 M guanidinium chloride, 0.1 M Tris/HCl, pH 8.0, and reduced with 10 mM dithiothreitol for 5 h $\,$

at 37 °C. Iodoacetamide was added to a final concentration of 25 mM and left in the dark overnight at room temperature. Reduced and alkylated material was dialysed into SDS/PAGE sample buffer.

Isopycnic ultracentrifugation

Freeze-dried material was dissolved by stirring at 4 °C overnight in Dulbecco's phosphate buffered saline containing 0.02%sodium azide and 42% (w/v) CsBr. Solubilized material was centrifuged at 170000 g for 72 h at 12 °C, as described previously [23]. Fractions of 1 ml were recovered from each tube and tested

Table 2 ¹H NMR chemical shifts of the oligosaccharide alditols A to G

nd, not determined.

		Oligosaccharide alditols, & (p.p.m.)								
0-Glycan	Reporter group	A	В	C	D	E	F	G		
GalNAc-ol I	H-2	4.404	nd	4.337	4.342	4.336	4.404	4.409		
	H-3	4.084	4.403	4.108	4.099	4.098	4.109	4.102		
	H-4	3.501	4.084	3.574	nd	nd	4.646	3.532		
	H-5	4.252	4.252	4.134	nd	4.227	4.128	4.194		
	NAc	2.056	2.042	2.051	2.055	2.040	2.046	2.055		
Gal(<i>β</i> 1-3)	H-1	4.572	4.570	4.664	4.620	4.620	4.618	4.602		
., .	H-3	nd	nd	4.015	4.009	4.009	nd	nd		
	H-4	3.91	3.924	4.228	4.22	4.222	4.043	4.051		
	H-6	nd	-	nd	3.93	nd	nd	nd		
Gal(81-3)	H-1	_	-	4 623	4 640	4 651	-	_		
	H-4	-	-	3.925	3.92	4.925	-	-		
Gal(x1-4) III/IV	H-1		-	-	-	_	4.932	4.937		
	H-4	-	-	-	-	-	4.033	4.032		
	H-5	-	-	_	_		4.326	4.316		
GICNAC(131-6) II'	H-1	4.551	4.560	-	4.559	4.557	_	4.553		
•	H-3	nd	nd	-	nd	nd	-	3.546		
	H-4	nd	nd	-	3.45	nd	-	3.45		
	H-5	nd	nd	-	3.45	nd	-	3.45		
	H-6	nd	3.944		nd	nd	-	nd		
	NAc	2.056	2.054	-	2.055	2.045	-	2.060		
$Fuc(\alpha 1-2) F(II)$	H-1	5.222	5.221	5.383	5.361	5.357	5.261	5.231		
	H-3	nd	ndi	3.892	nd	nd	nd	nd		
	H-5	4.275	4.275	4.277	4.25	4.265	4.284	4.282		
	H-6	1.244	1.245	2.464	1.232	1.233	1.244	1.245		
Fuc(α 1-2) F(III)	H-1	-	-	-	-	_		-		
	H-5	-	-	-	-	-	-	-		
	H-6	-	-	-	-	-	-	-		
Fuc(a1-3) F(11')	H-1	-	4.991	-	_	4.989	· _	-		
	H-5	-	4.327	-	-	4.33	-	-		
	H-6	-	1.160	-	-	1.159		-		

for carbohydrates using the orcinol assay. The high-density fractions ($\rho > 1.4$ g/ml) were pooled and exhaustively dialysed against deionized water and freeze-dried.

400 MHz ¹H NMR spectroscopy

NMR experiments were performed on a Bruker ASX 400 WB spectrometer. Chemical shifts are expressed downfield from internal 4,4'-dimethyl-4-silapentane-1-sulphonate but were actually measured by reference to internal acetone ($\delta = 2.225$ in ²H₂O at 25 °C). The two-dimensional homonuclear COSY were performed using Bruker standard pulse sequences.

Methylation analysis

Permethylation was carried out as described by Ciacanu and Kerek [24]. After methanolysis (0.5 M HCl/methanol), the partially methylated methyl glycosides were peracetylated and analysed by gas chromatography (GC)–MS [25]. The sulphated oligosaccharide alditols were subjected to mild periodate oxidation prior to methylation analysis with sodium meta-periodate in imidazole buffer, at 0 °C for 30 min, as described previously [26]. This specifically cleaved the *N*-acetylhexosaminitol unit, which facilitated the chloroform extraction of unsulphated branchs of the sulphated O-glycans.

© 2000 Biochemical Society

Fertilization assay

Assays were done according to Olson and Chandler [8]. Briefly, male testes were removed and macerated in $1.5 \times O$ -R2 buffer (124 mM NaCl, 3.75 mM KCl, 1.5 mM CaCl₂, 1.5 mM MgCl₂, 1.5 mM Na₂HPO₄ and 10 mM Hepes). Sperm density was normalized to 5×10^7 sperm/ml. Just ovoposited eggs were covered with F-1 buffer (41.25 mM NaCl, 1.25 mM KCl, 0.25 mM CaCl₂, 0.06 mM MgCl₂, 0.5 mM Na₂HPO₄ and 2.5 mM Hepes, pH 7.8). Sperm was added to a final concentration of 1.0×10^6 sperm/ml in the fertilization assay medium. Egg development was followed for 6 h after fertilization. Fertilization was scored by counting cleavage stage embryos after 2 h.

RESULTS

Fractionation of oligosaccharide alditols

Egg jellies extracted from individual clutches of six X. laevis specimens were subjected to alkaline treatment. Purified Oglycans were sequentially injected on to an amino-bonded silica column, in equivalent conditions. Out of the six resulting chromatographic profiles (Figure 1), four were qualitatively identical (specimens 1, 2, 3 and 5; Figures 1a, 1b, 1c and 1d respectively) while two presented original patterns (specimens 4 and 6; Figures 1e and 1f respectively). This suggested the occurrence of three distinct patterns of glycosylation in the egg jelly of the six specimens. The use of two other chromatographic

Table 3 ¹H NMR chemical shifts of the oligosaccharide additois H to L

nd, not determined.

		Oligosaccharide alditols, δ (p.p.m.)							
0-Glycan	Reporter group	н	1	J	к	L			
GalNAc-ol I	H-1	nd	nd	3.81	3.75	3.80			
	H-1'	nd	nd	3.76	3.79	3 74			
	H-2	4 400	1 24	4.24	A 22A	A 24			
	11-2 LI D	4.000	4.065	4.05	4.044	4.64			
	п-з	4.099	4.000	4.00	4.044	4.04			
	H-4	3.525	4.019	3.62	3.612	3.62			
	H-5	4,191	4.089	4.19	4.190	4.19			
	H-6	nd	nd	3.92	3.924	3.92			
	H-6'	nd	nd	3.70	3.694	3.67			
	NAc	2.043	2.046	2.05	2.036	2.05			
Gal(<i>β</i> 1-3) II	H-1	4.601	4.696	4.68	4.684	4.69			
	H-2	nd	nd	3.82	3.813	3.84			
	H-3	nd	4.186	4.18	4.178	4.20			
	H-4	4.050	4.019	4.02	4.016	4 00			
	H-5	nd	nd	nd	nd	3.77			
Gal(81-3) III	H-1	_	4.899	4.89	4.898	5.03			
H-4 H-5 H-5 H-2 H-2 H-3 H-4 H-5 H-6 H-6' 3al(α1-4) III/IV H-1 H-2 H-3 H-4 H-5 H-6	H-2	_	nd	3.69	3.691	3.72			
	H-3	_	nd	3,89	3,895	3 95			
	H-4	_	3 906	3.92	3 015	4.05			
	115	-	0.000 nd	20.0 pd	0.010 od	2 70			
	11-3 11-6	-	HU	10	nu	3.13			
	л-о Н-б'	-	Dri ba	nd	nd	on ba			
0-1/ 1 0 10/01	11-0	-	110	110	110				
Gar(021-4) 111/1V	n-i u n	4.930	-	-		4.90			
	n-2	nu nd	-	-	-	3.03			
	n-3	10	-	-	-	3.90			
	H-4	4.032	-	-	-	4.04			
	H-5	4.274	-	-	-	4.38			
	H-6	nđ	-	-	-	3.71			
	H-6'	nd	-	-	-	3.71			
GIcNAc(<i>(</i> 31-6) 11'	H-1	4.563	-	4.56	4.569	4.56			
	H-2	nd	-	3.71	3.845	3.71			
	H-3	nd	-	3.53	3.641	3.53			
	H-4	nd		3.44	3.53	3.44			
	H-5	nd	-	3.45	3.53	3.45			
	H-6	nd	-	3.93	3,930	3.94			
	H-6'	nd	_	3 74	3 749	3 74			
	NAc	2.056	-	2.05	2.046	2.05			
$Fuc(\alpha 1-2) F(11)$	H-1	5,289	5.407	5.41	5.412	5.42			
_/ . (. ,	H-2	nđ	nd	3.77	3,767	3.78			
	H-3	nd	nd	3.86	3,866	3.86			
	11-5	od	nu	2.00	3.000 nd	3.00			
	11-4 11 E	4 200	4 200	4.07	4 976	3.01			
	n-5 H.6	4.309	4.299	4.27	4.270	4.30			
Evelut ON Edith	11-0	1.245	5.200	6.00	5.000	5.04			
$ruc(\alpha -2) r(m)$	n-1 u n	-	0.329 od	2.33	2.332	0.34			
	n-2	-	IN	3.00	3.799	3.60			
	H-3	-	па	3.74	3.739	3./1			
	H-4	-	กป	3.82	nd	3.81			
	H-5	-	4.336	4.33	4.333	4.31			
	H-6	-	1.197	1.23	1.197	1.24			
Fuc(a1-3) F(11')	H-1	4.992	-	-	4.984	-			
	H-2	nd	-	-	3.694	-			
	H-3	nd	-	-	3.830	-			
	H-4	nd	-	-	nd	-			
	H-5	4.326	-	-	4.334	-			

50

systems (high performance anion-exchange chromatography and TLC; results not shown) gave identical results.

In order to determine the extent of variability between these three groups, the major peaks from specimens 3, 4 and 6 were isolated for structural analysis. Furthermore, we could confirm the similarity of glycosylation between specimens 1, 2, 3 and 5, by also analysing major components from these specimens. After structural analysis, it appeared that specimens 1, 2, 3 and 5 formed a group presenting homogeneous glycosylation, while specimens 4 and 6 both presented an original glycosylation, as suggested by the chromatographic analyses. All the O-glycans structures analysed are described in Table 1. Depending on their

Table 4 ¹H NMR chemical shifts of the oligosaccharide aiditols M to P

		Oligosaccharide alditols, δ (p.p.m.)							
0-Glycan	Reporter group	M	N	0	Р				
GalNAc-ol 1	H-1.1′	3.78	3.78	_					
	H-2	4.24	4.228	-	-				
	H-3	4.023	4.045	_	-				
	H-4	3 576	3.625	_	_				
	H-5	415	4.188	-	-				
	H-6.6'	3 84/3 463	3.925/3.695	-	-				
	NAc	2.042	2.046	-	_				
Collect	11.4.47	2.012	2.010	0 70 /0 70	0.75				
031-01	H-1,1	-	-	3./3/3./b	3./5				
	H-2	-	-	4.18	4.187				
	H-3	-	-	3.98	3.989				
	H-4	-	-	3.89	3.892				
	н-э H-6.6′	-	-	3.67	3.760				
Coll (01.2) 11		4 600	4 600						
uai(ρ1-3) II	П-I И Л	4.099	4.002		-				
	п-2 11 р	3.017	3.000	-	-				
	H-3	4.182	4.204	-	-				
	H-4	4.009	4.004	-	-				
Gal(β1-3) III	H-1	4.893	4.877	-	-				
	H-2	3.686	3.799	-	-				
	H-3	3.891	4.189	-	-				
	H-4	3.919	4.007	-	-				
Gal(R1-3) IV	H-1	_	4 992	-	_				
	H-2	_	3 694	_	_				
	H.3	-	3 883	_	-				
	H-4	-	3.903		_				
Fuc(~1.2) F(0)	H.1	5 403	5 250	510	5 1 0 5				
	11-1 LL D	2.76	3.555	3.87	3 803				
Fuc(~1.2). Cal.of	11-2 11-2	3.76	3 607	3.81	2.003				
	11-3 11 A	3.70	3 825	3.87	2 915				
	0-9 U 5	3.02	3.023	3.02	3.013				
	n-5 H-6	4.30	9.327	4.21	4.203				
	11.0	5.000	5 200	1.20	1.100				
$ruc(\alpha 1-2)$ $r(11)$	∏*I U 1	3.320	3.390	-	_				
	11-2	3.790	3.// 1	-	-				
	n-3	3.73	3.000	-	-				
	n-4	3.01	3./93	-	-				
	п-э н с	4.33	4.270	-	_				
	n-0	1.201	1.195	-	-				
$FUC(\alpha 1-2) F(IV)$	H-1	-	5.348	-	-				
	H-2	-	3.782	-	-				
	H-3	-	3.697	-					
	H-4	-	3.82	-	-				
	H-5	-	4.382	-	-				
	H-6	-	1.219	-	-				
GicNAc(31-6) II'	H-1	-	4.552	-	-				
	H-2	-	3.717	-	-				
	H-3		3.533	-					
	H-4	-	3.434	-	-				
	H-5	-	3.450	-	-				
	H-6.6′	-	3.952/3.742	-	-				
	NAC	-	2.039	-	-				
GalNAc(a1-3) II	H-1	_	-	5.170	5,185				
	H-2	_	_	4.304	4.308				
	H-3	_	-	4 041	4 053				
	H-4	_	_	4 080	4 080				
	H-5	_	_	4 207	4.000				
	H-6 6'		_	3 60	3 75				
	NAr	_	_	2.05	2.12				
GicNAc(~1.2) III	H-1	_	_	105	¥ 000				
uning(21-3) III	H-2	_	_	3.00	7.500				
	11-2 H.9	-	_	3.53	9.113				
	H-4	_	_	3.00 3.55	3 630				
	H-5	-	-	412	J.U29 1 102				
	H-6.6'	_	-	382/272	9.100 2 85				
	NAc	_	_	0.00/0./0 0 AD	2.03				
	11/16	-	-	2.03	2.0/1				

© 2000 Biochemical Society

455

O-Glycan		Oligosaccharide alditols, δ (p.p.m.)						
	Reporter group	M	N	0	Р			
Fuc(a1-3) F(111)	H-1	_	_		5.046			
	H-2	-	-	-	3.722			
	H-3	-	-	-	3.846			
	H-4	-	-	-	3.804			
	H-5	-		-	4.351			
	H-6	-	-	-	1.169			
NeuAc(a2-6)	H-3ax	1.681	-	-				
	H-3ea	2.721	-	-	-			
	NAC	2.033	-	-	-			

Table 4 (contd.)

presence, these structures were differentially assigned to each group.

Structural analysis

¹H NMR parameters of compounds A to W are summarized in Tables 2, 3, 4 and 5. Compounds A, D, I, J, M, O and Q have been described in X. laevis previously [11]. Briefly, the core of the A, D and J compounds was Gal(β 1-3)[GlcNAc(β 1-6)]GalNAc-ol as shown by the H-2 and H-5 proton resonances of GalNAcol [27]. Moreover, the set of the GlcNAc H-2, H-3 and H-4 proton resonances was typical of a terminal non-reducing residue. For compound A, the ¹H NMR parameters of Fuc were typical of the α -(1 \rightarrow 2) linkage to Gal. Thus, the structure of A was deduced to be $Fuc(\alpha 1-2)Gal(\beta 1-3)[GlcNAc(\beta 1-6)]GalNAc-ol.$ For compound J, the pattern of Fuc III and Gal III H-2, H-3 and H-4 proton resonances indicated a terminal Fuc(α 1-2)Gal sequence, as observed for compound A. Furthermore, the unusually downfield-shifted value of Gal III anomeric proton at 4.89 p.p.m. was established as indicative of the sequence $Fuc(\alpha 1-2)Gal(\beta 1-$ 3)[Fuc(α 1-2)]Gal [11]. The compound D was established as an extension of A with a terminal Gal residue III linked in β -(1 \rightarrow 3) to a Gal II. From the ¹H NMR parameters relative to I, the compound exhibited the same unit $Fuc(\alpha 1-2)Gal(\beta 1-3)[Fuc(\alpha 1-2)Gal(\beta 1-3)]Fuc(\alpha 1-2)Gal(\beta 1-3)]Fuc(\alpha 1-2)Gal(\beta 1-3)[Fuc(\alpha 1-2)Gal(\beta 1-3)]Fuc(\alpha 1-2)Gal(\beta 1-3)]Fuc(\alpha 1-2)Gal(\beta 1-3)[Fuc(\alpha 1-2)Gal(\beta 1-3)]Fuc(\alpha 1-2)Gal(\beta 1-3)]Fuc(\alpha 1-2)Gal(\beta 1-3)[Fuc(\alpha 1-2)Gal(\beta 1-3)]Fuc(\alpha 1-3)[Fuc(\alpha 1-2)Gal(\beta 1-3)]Fuc(\alpha 1-3)[Fuc(\alpha 1-3)]Fuc(\alpha 1-3)$ 2)]Gal ascribed for J. However, the chemical shift of H-6 and H-6' of proton resonances, and the downfield shift of H-5 proton resonance from the GalNAc-ol residue showed the absence of C-6 substitution on this residue. It has to be noted that Fuc II and Fuc III H-1 assignments reported in [11] have been revised on the basis of rotating-frame Overhauser enhancement spectroscopy ('ROESY') experiments (results not shown). This established the structure of compound I as the pentasaccharide $Fuc(\alpha I)$ 2)Gal(β 1-3)[Fuc(α 1-2)]Gal(β 1-3)GalNAc-ol. The structure of C has been reported to be a carbohydrate chain of the anuran species Rana temporaria [28]. Oligosaccharide D is an extension of C with a β -GlcNAc residue at C-6 of GalNAc-ol, as shown by the additional GlcNAc II' H-1 proton resonance and the characteristic chemical shifts of the GalNAc-ol H-5 signal.

Compound M has already been characterized in the oviducal mucin of *Rana dalmatina* [29]. Compound N (Figure 2) was isolated from jelly mucin of specimens 1 and 2. From the twostep relayed COSY spectrum depicted in Figure 2, the presence of three α -Fuc, three β -Gal, one β -GlcNAc and one GalNAc-ol units was clearly established. The characteristic chemical shifts of GalNAc-ol H-6, H-6', as well as the H-2, H-3 and H-4 resonances of β -GlcNAc, are significant of the terminal GlcNAc in β -(1 \rightarrow 6) linkage to GalNAc-ol. Thus, Fuc and Gal units are necessarily part of the lower branch of the molecule. The three anomeric protons observed at 4.682, 4.877 and 4.992 p.p.m. were assigned to Gal II, Gal III and Gal IV respectively. The set of H-2, H-3 and H-4 resonances of Gal IV are characteristic of the terminal unit Fuc(α 1-2)Gal. The Gal II and Gal III H-3 and H-4 resonances are indicative of a 2,3-di-substituted β -Gal unit, while the Gal II H-1 signal is not significantly affected by the presence of the additional β -Gal IV unit. Moreover, the observation of the nuclear Overhauser effect ('NOE') contacts H-1 Fuc (IV)/H-2 Gal IV (results not shown) allowed us to determine the exact assignments of the anomeric protons of the three fucose units. These observations led to the proposal of the structure of compound N as an extension of compound J with a terminal Fuc(α 1-2)Gal unit.

Structures O and Q were also characterized through a combination of NMR and methylation analysis in X. laevis previously [11]. The newly isolated compound R presented the same ¹H NMR parameters as Q for the Gal II, GalNAc III and GlcNAc IV units, but also showed an additional β -GlcNAc residue. The H-6 and H-6' signals of the GalNAc-ol residue resonated at 3.917 and 3.711 p.p.m. respectively, which defines the presence of core type II. Thus, compound R appeared to be an extension of Q with a β -GlcNAc residue. The H-6 and H-6' proton resonances of GlcNAc II' are deshielded at 4.373 and 4.23 p.p.m. respectively, and the absence of an additional couplage with ³¹P confirmed the presence of a sulphate group attached at O-6 of this monosaccharide unit.

Compound T (Figure 3) presented identical characteristics to R but contained an additional β -Gal unit (Gal III') linked in $(1 \rightarrow 4)$ to the GlcNAc II' residue as shown by the deshielding of the GlcNAc II' H-5 resonance in compound T at 3.786 p.p.m. Furthermore, the downfield-shifted resonances of Gal III' H-3 and H-4 were attributable to the C-3 substitution by a sulphate group [30].

Oligosaccharide-alditol W contained two α -Fuc, two α -GalNAc, two α -GlcNAc, two β -Gal, one O-6 sulphated β -GlcNAc and one GalNAc-ol residues, as could be inferred from the integration of the anomeric signals and the set of vicinal coupling constants. The H-1 resonances of Gal II and Gal III', both observed at 4.709 p.p.m., were characteristic of a β -Gal unit involved in the 'blood group' A determinant [31]. These observations showed the structure of compound W to be an extension of R with the sequence GlcNAc(α 1-3)GalNAc(α 1-3)[Fuc(α 1-2)]Gal.

Structures containing the terminal α -(1 \rightarrow 4)-linked Gal unit

Compounds F, G and L have been described previously [11]. They were characterized by the presence of the Gal unit linked in

456 Y. Guerardel and others

Table 5 ¹H NMR chemical shifts of the oligosaccharide additols Q to W

nd, not determined.

		Oligosaccharide alditols, & (p.p.m.)									
0-Glycan	Reporter group	Q	R	S	T	U	v	W			
GalNAc-ol I	H-1	nd	nd	3.79	3.77	3.79	3.792	3.79			
	H-1'	nd	nd	3.79	3.77	3.79	3.792	3.79			
	H-2	4.310	4.287	4.287	4.30	4.325	4.293	nd			
	H-3	nd	4.070	4.073	nd	4.069	4.068	4.071			
	H-4	3.601	3.616	3.616	3.621	3.615	3.608	3.610			
	H-5	4.22	4.19	4.192	4.191	4.190	4.188	4.198			
	H-6	nd	3.917	3.924	nd	3.93	3.927	3.904			
	H-6'	nd	3.711	3.712	nd	3.70	3.698	3.690			
	NAc	2.038	2.043	2.043	2.044	2.044	2.044	2.047			
Gal(<i>β</i> 1-3) II	H-1	4.707	4.712	4.715	4.716	4.714	4.713	4.709			
	H-2	nd	3.911	3.913	3.918	3.916	3.917	3.911			
	H-3	nd	4.020	4.023	4.028	4.021	4.022	4.018			
	H-4	4.22	4.209	4.219	4.207	4.216	4.215	4.206			
GalNAc(a1-3) III	H-1	5.234	5.226	5.244	5.226	4.244	5.245	5.225			
	H-2	4.328	4.339	4.348	4.337	4.348	4.347	4,337			
	H-3	nd	4.067	4.058	4.043	4.058	4.057	4.042			
	H-4	4.061	4.062	4.074	4.062	4.072	4.095	4.063			
	H-5	4.158	4.24	4.211	4.231	4.27	4.232	4.278			
	H-6	nd	nd	3.63	3.73	3.74	3.633	3.71			
	H-6′	nd	nd	3.63	3.73	3.74	3.633	3.71			
	NAC	2.046	2.043	2.049	2.047	2.052	2.052	2.056			
GICNAC(a1-3) IV	H-1	4.941	4.941	4.886	4.941	4.889	4.888	4.940			
. ,	H-2	nd	3.921	4.114	3.926	4.113	4.113	3,927			
	H-3	nd	3.863	3.889	3.913	3.890	3.888	3.863			
	H-4	3.549	3.549	3.642	3.550	nd	3.632	3.549			
	H-5	4 112	nd	nd	4.173	nd	4 223	415			
	H-6	nd	4152	nd	3.85	3 845	3 848	3 828			
	H-6'	nd	4 154	nd	3 722	3 845	3 742	3 737			
	NAC	2.087	2 087	2.070	2.089	2.070	2,071	2 087			
	11.4	2.001	4.500	4.500	4.044	1.01.0	2.071	4.000			
GICNAC(p1-0) II	H-I	-	4.568	4.089	4.014	4.014	4.529	4.602			
	H-2	-	3./3/	3.738	3.788	3.793	3.951	3.//8			
	H-3	-	3.557	nd	na	na	4.044	na			
	H-4	-	4.230	no	na	nd	no	nd			
	H-5	-	3.677	3.676	3.786	nd	3.836	nd			
	H-6	-	4.373	4.3/1	4.450	4.452	4.416	4.379			
	H-6'	-	4.23	4.236	4.320	4.324	4.378	4.26			
	NAC	-	2.053	2.054	2.051	2.052	2.044	2.041			
Gal(β1-4) III'	H-1	-	-	-	4.648	4.647	4.647	4.709			
	H-2	-	-	-	3.671	3.673	3.622	3.871			
	H-3	-	-	-	4.345	4.346	4.330	3.970			
	H-4	-	-	-	4.294	4.296	4.276	4.221			
GalNAc(a1-3) IV'	H-1	-	-	-	-	-	-	5.225			
	H-2	-	-	-	-	-	-	4.310			
	H-3	-	-	-	-	-	-	4.058			
	H-4	-	-	-	-	-	-	4.042			
	H-5	-	-	-	-	-		4.233			
	H-6	-	-	-	-	-	-	3.72			
	H-6′		-	-	-	-	-	3.72			
	NAc	-	-	-	-	-	-	2.035			
GicNAc(a1-3) V'	H-1	-	-	-	-	-	-	4.940			
	H-2	-	-	-	-	-	-	3.927			
	H-3	-	-	-	-	-	-	3.852			
	H-4	-	-	-	-	-	-	3.549			
	H-5	-	-	-	-	-	-	4.15			
	H-6	-	-	-	-	-	-	3.828			
	H-6'	-	-	-	-	-	-	3.737			
	NAc	-	-	-		-		2.087			
Fuc(a1-2) F(11)	· H-1	5.387	5.385	5.387	5.385	5.385	5.384	5.380			
	H-2	nd	3.80	3.81	3.799	3.802	3.802	3.802			
	H-3	nd	3.80	3.81	nd	nd	3.8	nd			
	H-4	nd	nd	3.81	nd	nd	3.8	nd			
	H-5	4.328	nd	4.327	4.319	4.327	4.322	nd			
	H-6	1.232	1.228	1.231	1.232	1.230	1.235	1.226			

457

		Oligosaccharide alditols, δ (p.p.m.)									
0-Glycan	Reporter group	0	R	S	T	U	٧	W			
Fuc(a1-3) F(IV)	H-1	<u>.</u>	-	5.052	_	5.054	5.055	-			
. , . ,	H-2	_	-	3.723	~	nd	3.720	-			
	H-3	-	-	nd	-	nd	3.853	-			
	H-4	-	-	nd	-	nd	nd	-			
	H-5	-	-	4.356	-	4.357	4.356	-			
	H-6	-	-	1.171	-	1.170	1.171	-			
Fuc(α 1-3) F(II')	H-1	-	_	_	-	-	5.121	-			
	H-2	-	-	-	-	-	3.679	-			
	H-3	-	_	-	-	-	3.916	-			
	H-4	-	-	-	-	-	3.796	-			
	H-5	-	_	-	-		4.801	_			
	H-6	-	_	-	-	-	1.181	-			



Table 5 (contd.)

Figure 2 COSY spectrum of oligosaccharide alditol N

 α -(1 \rightarrow 4) to Gal II or Gal III. The compounds G (Figure 4) and L were extensions of A and J respectively, with an additional terminal Gal unit linked in α -(1 \rightarrow 4). The Gal unit in an α -(1 \rightarrow 4) linkage was clearly reflected by the anomeric proton observed at 4.932 or 4.937 p.p.m. for F and G respectively, and 4.98 p.p.m.

for L, whereas the chemical shift of the H-5 signal varied between 4.32 and 4.38 p.p.m.

Compound H (Figure 4), which concomitantly presents both terminal Gal in an α -(1 \rightarrow 4) linkage and terminal Fuc in an α -(1 \rightarrow 3) linkage, are discussed below.

Structures containing the sequence $Fuc(\alpha 1-3)GicNAc(\beta 1-6)$ or $Fuc(\alpha 1-3)GicNAc(\alpha 1-3)$

Compounds B, E and K were extensions of compounds A, D and J with a Fuc linked in α -(1 \rightarrow 3) to β -GlcNAc. Methylation analysis of compounds B, E and K confirmed this substitution through the presence of 4,6-di-O-Me-GlcNAc among the methyl ethers obtained from these compounds, while compounds A, D and J showed 3,4,6-tri-O-Me-GlcNAc. From the Tables 2 and 3, it can be noted that the chemical shifts of the Fuc H-1, H-5 and H-6 proton resonances are remarkably constant from one compound to another. Similarly, the structural-reporter-group signals of the constituting residues of the lower branch at GalNAc-ol match completely those of the same branch in A, D and J.

Four other compounds, namely P, S, U and V, also contained an additional α -Fuc unit, characterized by its H-1 (δ 5.05), H-5 (δ 4.36) and H-6(δ 1.17) proton resonances (Figure 3). A comparison of the COSY spectra of R and S on one hand, and S and T on the other hand, clearly indicated the sequence HSO₃(6)GlcNAc(β 1-6) for R and S upper branches and HSO₃(3)Gal(β 1-4)[HSO₃(6)]GlcNAc(β 1-6) for T upper branch. The lower branches of compounds S and U were fucosylated at O-3 of α -GlcNAc, as proved by the downfield shift of its H-2 ($\Delta\delta$ = +0.19) and H-2 ($\Delta\delta$ = +0.03) signals. Thus they appear as an extension of the R and T lower branch with a Fuc in α -(1 \rightarrow 3) linkage.

Methylation analysis was performed on compounds T and U to confirm the position of the Fuc in the α - $(1 \rightarrow 3)$ linkage. A mild periodate oxidation of GalNAc-ol residue of T and U, prior to the methylation, released the di-sulphated Gal(β 1-4)GlcNAc sequence from the lower branch. Thus, the lower branch could be effectively extracted in the chloroform phase after methylation, while the sulphated upper branch remained in the aqueous phase. GC-MS analysis of compound T lower branch showed one 4,6-di-O-Me-Gal, one 2,3,4-tri-O-Me-Fuc, one 4,6-di-O-Me-GalNAc and one 3,4,6-tri-O-Me-GlcNAc among the partially methylated and acetylated methyl glycosides. Compound U differed from T by the presence of two 2,3,4-tri-O-Me-Fuc and one 4,6-di-O-Me-GlcNAc instead of a 3,4,6-tri-O-Me-Fuc



Figure 3 COSY spectra of oligosaccharide aiditols T (left) and U (right)

GlcNAc. This confirmed the ¹H NMR parameters ascribed to Fuc in α -(1 \rightarrow 3) linkage for compounds P, S, U and V.

Compound V contained two sulphate groups, attached to O-3 of β -Gal III' (δ H-3 = 4.330; δ H-4 = 4.276) and O-6 of β -GlcNAc II' (δ H-6 = 4.416; δ H-6' = 4.378) respectively, as already described for compounds T and U. The presence of an α -Fuc unit O-3 linked to β -GlcNAc II' can be deduced from the Fuc H-1, H-5 and H-6 proton resonances, which are characteristic of the Lewis X determinant [27]. Since the NMR parameters relative to the lower branch perfectly match those observed for compounds S and U, the structure of compound V was fully established.

From these results, it appears that Gal in α -(1 \rightarrow 4) linkage and Fuc in α -(1 \rightarrow 3) linkage can be easily identified on the basis of one-dimensional ¹H NMR spectra. Particularly, the H-1 signals of α -Fuc O-linked to β -GlcNAc or α -GlcNAc are observed at 4.984–4.992 and 5.052–5.055 p.p.m. respectively. Similarly, the Gal in α -(1 \rightarrow 4) linkage is characterized by its H-1 resonance observed at 4.932–4.936 or 4.98 p.p.m., according to the nature of the core, Gal(β 1-3)GalNAc-ol (compounds F and G) or Gal(β 1-3)Gal(β 1-3)GalNAc-ol (compound L). Owing to these new structural reporter groups, the structure of compound H was established on the basis of the observation of signals at 4.936 (H-1 α Gal) and 4.992 p.p.m. (H-1 α Fuc). Compound H, which is a combination of B and G (Figure 4), is the one compound of the



series exhibiting both Gal in α -(1 \rightarrow 4) linkage and Fuc in α -(1 \rightarrow 3) linkage.

Thus, as shown in Table 1, the three groups of X. laevis specimens differed by the terminal substitution of their egg-jelly O-glycans. Specimens 1, 2, 3 and 5 only presented Gal in α -(1 \rightarrow 4) linkage, specimen 4 only presented Fuc in α -(1 \rightarrow 3) linkage, while specimen 6 showed both substitutions.

Electrophoretic analysis

Extracted egg-jelly coats from the six specimens of X. laevis were analysed on a gradient 2.5-10% polyacrylamide/SDS gel. As shown in Figure 5(a), PAS staining revealed three populations of very high-molecular-mass glycoproteins called Hm, Mm and Lm for high-, medium- and low-mobility glycoproteins respectively. It is noteworthy that all the O-glycans we have isolated from jelly, even the acidic ones, are susceptible to periodate oxidation and thus are stained by PAS. Lm populations showed typically two narrow bands, Mm populations a single broad band and Hm populations one to three thin bands. Mm and Hm populations present apparent molecular-mass variations between individuals. This polymorphic appearance is common to proteins with repetitive elements, such as mucin-type glycoproteins, where it is due to the expression of a variable number of tandem repeats [32] e.g. between 20 and > 100 in MUC 1 [33]. This phenomenon is



Figure 4 ¹H NMR spectra of oligosaccharide alditois G (top panel), B (middle panel) and H (bottom panel)



Figure 5 SDS/PAGE of egg jelly from six specimens of X. laevis (1 to 6)

A 2.5–10% polyacrylamide/SDS gradient gel was used. Gels were either (a) stained for carbohydrates with PAS or transferred on to nitrocellulose membranes. Membranes were then probed with either (b) BSI-B₄ lectin or (c) *T. purpureas* lectin, Lm, Mm and Hm: low, medium and high mobility respectively. In these experimental conditions, standard protein molecular-mass markers are excluded from the gel.

also typically responsible for the expression of polydisperse mRNAs in a single individual and synthesis of apomucins of different lengths. It has been observed for FIM B1 and FIM C1 in X. laevis [13,15] and for other mucins such as MUC 4 in humans [34]. The multiple bands observed in Lm and Hm populations may either result from the expression of such



Figure 6 Two-dimensional electrophoresis of egg jelly from specimen 6 of X. Jaevis

Gels were either (a) stained for carbohydrates with PAS or blotted with (b) BSI-B₄ lectin or (c) *T. purpureas* lectin. Lane 1 shows the one-dimensional electrophoresis of specimen 6. Only the compound labelled 'a' in panels (a) and (b) from the Hm population reacted to BSI-B₄ lectin, while the compound labelled 'b' in panel (a) did not react to either lectin. A third compound from the Hm population, labelled 'c' in panel (c), reacted intensively to *T. purpureas* lectin.

C 2000 Biochemical Society



Figure 7 SDS/PAGE of egg jelly stained for carbohydrates with PAS

A 2.5–10% polyacrylamide/SDS gradient gel was used. (a) Jelly coat from specimen 2 was analysed before (lane 1) and after (lane 2) mild pronase proteolysis. For increased clarity the arrows indicate the exact positions of the bands for both samples. (b) Lane 1, total jelly coat from specimen 1; lane 2, high-density fraction ($\rho > 1.4$ g/ml.) recovered after CSBr ultracentrifugation of this sample.

polydisperse mRNAs or come from the synthesis of different components. The occurrence in specimen 4 and 5 of three bands in the Hm population is compatible with the first hypothesis considering the tetraploid status of *X. laevis* [35].

Two-dimensional electrophoresis of specimen 6 (Figure 6a) showed that the Hm population contained two very acidic components of similar pHi and slightly different apparent molecular-mass. The Mm population consisted of a mixture of molecules with different isoelectric points, at least two of which were acidic, while the L population appeared as not acidic.

In order to know whether the bands observed in SDS/PAGE resulted from molecular aggregation, despite the reducing conditions, samples were reduced and subsequently alkylated. We could not observe any difference in mobility on SDS/PAGE between these samples and their untreated equivalents (results not shown). This suggested that the high-molecular-mass components we observed were in their monomeric form.

After pronase digestion, the Mm and Hm populations of glycoproteins showed a slight reduction in their molecular mass (Figure 7a). However, we did not observe the appearance of new bands. These two observations are consistent with the presence of a protease-resistant, highly O-glycosylated core in these components. However, we did not observe any molecular-mass reduction for the Lm population in the pronase digested preparation. Considering the fact that this population barely penetrates the separating gel, it is likely that a slight reduction in molecular mass would not be visible under these experimental conditions.

After agarose electrophoresis (results not shown), PAS intensively stained one cathodic and one anodic broad band. The anodic PAS-reactive band showed the same mobility as purified bronchial mucins. Two anodic bands were stained by Toluidine Blue, one of which corresponded to the PAS stained anodic band. All these bands were entirely recovered in the fraction of upper density (superior to 1.4) after isopycnic ultracentrifugation. Furthermore, SDS/PAGE of this high-density fraction showed the same high-molecular-mass glycoprotein populations as the total egg-jelly coat (Figure 7b). The susceptibility of the anodic PAS-reactive band to Toluidine Blue showed that it contained acidic groups. Toluidine Blue stained bands were insensitive to any of the glycosaminoglycan degrading enzymes cited in the experimental section.

Thus, from these data, it appears that the PAS-reactive bands observed in agarose and polyacrylamide gels are typical mucintype glycoproteins. Some of them exhibit acidic properties, which may be due to the presence of the sulphated O-glycans described previously, while others do not. This suggests that the mucin-type glycoproteins would support distinct O-glycans. Also, the most anodic Toluidine Blue-reactive band does not appear as either a mucin-type glycoprotein, due to its lack of reactivity toward PAS staining, or a proteoglycan, as confirmed by its insensitivity to any glycosaminoglycan degrading enzymes.

Blot analysis

After SDS/PAGE and two-dimensional electrophoresis, gels were blotted and the membranes were probed with either BSI-B₄ or *T. purpureas* lectins, in order to localize α -galactose residues or Fuc in α -(1 \rightarrow 3) linkage respectively, on egg-jelly glycoproteins.

As Figures 5(b) and 5(c) showed, we observed a perfect correlation between electrophoretic and structural, NMR based, analysis. Specimens 1, 2, 3, 5 and 6 were reactive to BSI-B₄ lectin, whilst specimens 4 and 6 were reactive to *T. purpureas* lectin. The *T. purpureas* lectin distinguished Fuc in α -(1 \rightarrow 3) linkage residues from the ubiquitous Fuc in α -(1 \rightarrow 2) linkage. Thus, the three phenotypes described previously are easily identified by this means. Systematic analysis of 13 other specimens of *X. laevis* permitted isolation of a specimen presenting the, not yet identified, $[(\alpha 1-4)Ga1(-); (\alpha 1-3)Fuc(-)]$ phenotype.

It appeared (Figures 6b and 6c) that the three populations of high-molecular-mass glycoproteins were not homogeneously substituted by Fuc in α - $(1 \rightarrow 3)$ linkage and Gal in α - $(1 \rightarrow 4)$ linkage. Indeed, except in specimen 1, where it also concerned the Mm population, α - $(1 \rightarrow 4)$ galactosylation was restricted to the Hm population of glycoproteins, while α - $(1 \rightarrow 3)$ fucosylation was more widely distributed in Mm and Hm populations. Figure 6(c) showed that a previously not described very acidic band from the Hm population-equivalent molecular-mass (label c) was most intensively substituted by fucose in α - $(1 \rightarrow 3)$ linkage. The Lm population did not support either substitution. Within the same population α galactosylation was not evenly distributed. Indeed, in specimen 6, only one band (label a), out of the two that the Hm population contained (Figures 5a and 5b), reacted with BSI-B₄ lectin.

Fertilization assays

Fertilization assays on eggs presenting [G(+), F(-)], [G(+), F(+)] and [G(-), F(-)] phenotypes repetitively showed scores of approx. 90 %. So, no difference of fertilizability was observed between the three phenotypes tested. Fertilized eggs were kept for 6 h, and all of them showed normal subsequent development upon visual examination, irrespective of their phenotype.

DISCUSSION

In this study, we have fully sequenced 76 O-linked oligosaccharides isolated from the egg-jelly coat of six X. laevis specimens. We have described two families of O-glycans composed of 23 different structures, out of which 11 presented novel sequences. These O-glycans showed a remarkable polymorphism restricted to their terminal substitution. The combination of NMR-based and lectin studies led us to describe four phenotypes characterized by the differential expression of two glycosyltransferase activities, one α -(1 \rightarrow 4) galactosyltransferase and one α -(1 \rightarrow 3) fucosyltransferase activities, on the egg-jelly mucins. The Gal in α -(1 \rightarrow 4) linkage substituted either the Gal II or the Gal III of one family of glycans, while the Fuc in α -(1 \rightarrow 3) linkage substituted any α or β GlcNAc residue from both families. Out of 19 X. *laevis* specimens we have studied, 13 presented the $[\alpha-1,4Gal(+);$ α -1,3Fuc(-)] phenotype, four presented the [G(+); F(+)] phenotype, one the [G(-); F(+)] phenotype and one the [G(-); F(-)] phenotype. Each of these two activities may be the result of one or several glycosyltransferases. It should be noted that each substitution always occurs on the same acceptor substrate: Fuc(α 1-2)Gal sequence for α -(1 \rightarrow 4) galactosylation and GlcNAc residue for α -(1 \rightarrow 3) fucosylation. Even so, transfer of Gal residue in α -(1 \rightarrow 4) linkage on Gal II or Gal III may result from two distinct galactosyltransferases, while transfer of Fuc residue in α -(1 \rightarrow 3) linkage on α -GlcNAc II or β -GlcNAc IV may also be due to distinct fucosyltransferases. Nevertheless, the fact that each substitution was always observed simultaneously in both positions in a specimen suggests that each one is the result of either a single glycosyltransferase, or more unlikely several glycosyltransferases which segregate 'en bloc' in each specimen.

To our knowledge, such a polymorphism has only be described for ABO, Hh and Lewis blood-group systems in humans. The resulting phenotypes observed in the ABO system and X. laevis jelly coat are very similar. Indeed, in both cases we observe four phenotypes due to different terminal glycosylations of complex oligosaccharide structures and which present the same pattern: [A(+); B(+)], [A(+); B(-)], [A(-); B(+)] and [A(-); B(-)]for ABO system, and [G(+); F(+)], [G(+); F(-)], [G(-);F(+)] and [G(-); F(-)] in X. laevis. However, it is most unlikely that the polymorphism observed in X. laevis results from the expression of allelic forms of a single gene, as observed in the ABO system [36,37]. Indeed, the polymorphism in X. laevis concerns both acceptor and donor substrates while it concerns only donor substrate in the ABO system. Furthermore, while all fucosyltransferases utilize GDP-associated fucose as the donor substrate [38], galactosyltransferases utilize UDP-associated substrate. Thus, considering the major differences observed between galactosyltransferase and fucosyltransferase activities, we postulate that the four phenotypes observed in X. laevis originate from the differential expression of at least two independent glycosyltransferases. Furthermore, it appeared that this polymorphism was tissue-specific. Western blot analysis (results not shown) showed that skin glycoproteins always presented α linked galactoses but no fucose in α -(1 \rightarrow 3) linkage, irrespective of the phenotype observed in egg jelly.

This work has brought new data about the exact nature of the high-molecular-mass fraction of X. laevis egg-jelly. It appears to be mainly composed of mucin-type glycoproteins. These components showed an important degree of inter- and perhaps intra-specific variability. The present study did not allow factors of variability such as synthesis of distinct apomucins, expression of polydisperse mRNAs and synthesis of glycoforms to be distinguished. However, considering the importance of sulphation in jelly O-glycans, it is most probable that the variations of pIs observed in Mm populations are due to variable degrees of sulphation in the constituting mucins. This was confirmed by the fact that we could separate two populations of mucins presenting different quantities of sulphate groups by CsCl centrifugation (results not shown). The presence of differently charged mucins was postulated by Yurewicz et al. [10] who observed that only the J, layer contained sulphate groups. Thus, components of Mm populations might consist of glycoforms differently substituted by sulphated O-glycans. Evidence of the synthesis of glycoforms of the same mucin within the same tissue was unambiguously demonstrated by the study of MUC 5B in bronchial secretions [39]. Blot analysis also showed that jelly mucins were differently glycosylated within the same specimen. In each jelly layer, synthesized by a distinct segment of oviduct [1], differential glycosylation may be the result of compartmentalization of glycosyltransferase activities in each segment or within a segment, or a strict regulation of glycosylation of each apomucin. The oviduct of X. laevis appears as an appealing model for the study of O-glycosylation regulation because (i) important amounts of mucins are available, (ii) mucins which differ in their glycosylation are sequentially synthesized by distinct segments of oviduct, (iii) O-glycan structures synthesized by X. laevis oviduct are now well known, (iv) jelly mucins from X. laevis show a low extent of intra-specific variability of glycosylation, which is easily identified.

Only the diffusible components of egg jelly from X. laevis have a fertilization-promoting activity [7]. However, it was suggested that the ability of sperm to traverse the jelly was dependent on specific interactions with the carbohydrate moeities of jelly highmolecular-mass glycoconjugates [6]. Considering the important variation of glycosylation that we have observed in egg jelly from amphibians, it is conceivable that these sperm-jelly interactions would be very species specific. Thus, carbohydrate speciesspecificity may, at least partially, account for the inability of heterologous sperm to penetrate egg-jelly coat [4] and for the maintenance of the barrier preventing inter-species crossing. If so, divergence of glycan structures in egg jelly through evolution might be an important factor in speciation, by exerting an early control on inter-species fertilizability. The low extent of variation of glycosylation between the specimens presenting the [G(+);F(-), [G(+); F(+)] and [G(-); F(-)] phenotypes did not seemed to influence fertilizability of the jellied egg. This suggests that Gal linked in α -(1 \rightarrow 4) and Fuc linked in α -(1 \rightarrow 3) in jelly are not involved in specific interactions with sperm. Thus, this intra-specific polymorphism showed by the jelly high-molecularmass mucins does not seem to have an essential role in the control of egg fertilizability. It is not known to what extent Oglycan structures have to be modified in order to disturb sperm-jelly interactions and thus play a role in the selection of sperm.

Previous studies have shown that α -Gal-substituted glycans were ligands of the CGL [40] and were necessary for the formation of the fertilization envelope. The formation of this structure, as well as the cleavage of sperm receptors [41,42], is involved in the maintenance of monospermy in X. *laevis*. Despite this fact, eggs from specimens lacking terminal α -Gal substitution presented monospermic fertilization, as proved by their normal devolpment. Thus, either the fertilization layer was formed in the absence of the CGL putative ligand, or proteolytic cleavage of the sperm receptor alone is sufficient to maintain monospermy.

We believe that an in-depth knowledge of the egg-jelly matrix will help greatly in understanding the molecular mechanisms underlying the different interactions in which it is involved (sperm-jelly interactions, selection of sperm, formation of the fertilization layer). This will require studies of the structure of individual mucins and their glycosylation, as well as their exact location within the jelly.

We thank C. Vandeperre and P. Humbert for valuable technical assistance, and Dr G. Lamblin and Dr C. Mariller for critically reading the manuscript. This research was supported by the Centre National de la Recherche scientifique (UMR 8576, Glycobiologie structurale et fonctionelle, Director Professor André Verbert) and by the Ministère de l'Enseignement Supérieur et de la Recherche.

REFERENCES

- Yoshizaki, N. (1985) Fine structure of oviducal epithelium of *Xenopus laevis* in relation to its role in secreting egg envelopes. J. Morphol. 184, 155–169
- 2 Tian, J., Gong, H., Thomsen, G. H. and Lennarz, W. J. (1997) Gametes interactions in *Xenopus laevis*: identification of sperm binding glycoproteins in the eggs vitelline envelope. J. Biol. Cell **136**, 1099–1108
- 3 Wyrick, R. E., Nishihara, T. and Hedrick, J. L. (1974) Agglutination of jelly coat and cortical granule components and the block to polyspermy in the amphibian *Xenopus laevis*. Proc. Natl. Acad. Sci. U.S.A. **71**, 2067–2071

- 4 Brun, R. and Kobel, H. R. (1977) Observations on the fertilization block between Xenopus borealis and Xenopus laevis. J. Exp. Zool. 201, 135–138
- 5 Tseng, K., Bonnel, B. S., Wang, H., Hedrick, J. L. and Lebrilia, C. B. (1998) The structural analysis of anionic oligosaccharide-alditols released from the 630 ligand that binds with cortical granule lectin by MALDI-MS. Abstracts from the 1998 meeting of the Society for Glycobiology in Glycobiology 11
- 6 Reinhart, D., Ridgway, J. and Chandler, D. E. (1998) *Xenopus laevis* fertilization: analysis of sperm motility in egg jelly using video light microscopy. Zygote 6, 173-182
- ⁷ Bonnel, B. S., Reinhart, D. and Chandler, D. E. (1996) *Xenopus laevis* egg jelly coat consist of small diffusible proteins bound to a complex system of structurally stable networks composed of high-molecular-mass glycoconjugates. Dev. Biol. **174**, 32–42
- 8 Olson, J. H. and Chandler, D. E. (1999) *Xenopus laevis* egg jelly contains small proteins that are essential to fertilization. Dev. Biol. 210, 401–410
- 9 Mozingo, N. M. and Hedrick, J. L. (1999) Distribution of lectin binding sites in Xenopus laevis egg jelly. Dev. Biol. 218, 428–439
- 10 Yurewicz, E. C., Oliphant, G. and Hedrick, J. L. (1975) The macromolecular composition of *Xenopus laevis* egg jelly coat. Biochemistry 14, 3101–3106
- 11 Strecker, G., Wieruszeski, J. M., Plancke, Y. and Boilly, B. (1995) Primary structure of 12 neutral oligosaccharide-alditols released from the jelly coats of the anuran *Xenopus laevis* by reductive β -elimination. Glycobiology **5**, 137–146
- 12 Tseng, K., Lindsay, L. L., Penn, S., Hedrick, J. L. and Lebrilla, C. B. (1997) Characterization of neutral oligosaccharide-alditols from *Xenopus laevis* egg jelly coats by matrix-assisted laser desorption Fourier transform mass spectrometry. Anal. Biochem. **250**, 18–28
- 13 Hoffmann, W. (1988) A new repetitive protein from Xenopus laevis skin highly homologous to pancreatic spamolytic polypeptide. J. Biol. Chem. 263, 7686–7690
- 4 Probst, J. C., Hauser, F., Joba, W. and Hoffmann, W. (1992) The polymorphic integumentary mucin B.1 from *Xenopus laevis* contains the short consensus repeat. J. Biol. Chem. **267**, 6310–6316
- 15 Hauser, F. and Hotkmann, W. (1992) P-domains as shuffled cysteine-rich modules in integumentary mucin C.1 (FIM-C.1) from *Xenopus laevis*. J. Biol. Chem. 267, 24620–24624
- 16 Joba, W. and Hoffmann, W. (1997) Similarities of integumentary mucin B.1 from Xenopus laevis and prepro-von Willebrand factor at their amino-terminal regions. J. Biol. Chem. 272, 1805–1810
- 17 Strecker, G., Wieruszeski, J. M., Fontaine, M. D. and Plancke, Y. (1994) Structure of the major neutral oligosaccharide-alditols released from the egg jelly coat of Axoloti maculatum. Characterization of the carbohydrate sequence GalNAc(β1-β4)[Fuc(α1α3)]GlcNAc(β1-β3/6). Glycobiology 4, 605–609
- 18 Strecker, G., Wieruszeski, J. M., Michalski, J. C. and Montreuil, J. (1992) ¹H- and ¹³C-n.m.r. spectroscopy of 2-oxo-3-deoxy-o-glycero-o-galactononulosonic acidcontaining oligosaccharide-alditols bearing Lewis X, Lewis Y and A-lewis Y determinant isolated from the jelly coat of *Pleurodeles waltl* eggs. Biochem. J. **287**, 905–909
- 19 Maes, E., Wieruszeski, J. M., Plancke, Y. and Strecker, G. (1995) Structure of the three Kdn-containing oligosaccharide -alditols released from oviducal secretions of *Ambystoma tigrinum*: characterization of the carbohydrate sequence Fuc(α1α5)[(Fuc(α1-α4)]Kdn(α2-α3/6). FEBS Lett. **358**, 205–210
- 20 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227, 680-685
- 21 Zacharius, R. M., Zell, T. E., Morrisson, J. H. and Woodlock, J. J. (1969) Glycoprotein staining following electrophoresis on acrylamide gels. Anal. Biochem. 30, 148–152
- 22 Rahmoune, H., Lamblin, G., Lafitte, J. J., Galabert, C., Filliat, M. and Roussel, P. (1991) Chondroitin sulfate in sputum from patients with cystic fibrosis and chronic bronchitis. Am. J. Respir. Cell Mol. Biol. 5, 315–320
- 23 Houdret, N., Perini, J. M., Galabert, C., Scharfman, A., Humbert, P., Lamblin, G. and Roussel, P. (1986) The high lipid content of respiratory mucins in cystic fibrosis is related to infection. Biochim. Biophys. Acta 880, 54–61
- 24 Ciacanu, I. and Kerek, F. (1984) A simple and rapid method for the permethylation of carbohydrates. Carbohydr. Res. 131, 209–217
- 25 Fournet, B., Strecker, G., Leroy, Y. and Montreuil, J. (1981) Gas-liquid chromatography and mass spectrometry of methylated and acetylated methyl glycosides. Application to the structural analysis of glycoprotein glycans. Anal. Biochem. **116**, 489–502
- 26 Morelle, W., Lemoine, J. and Strecker, G. (1998) Structural analysis of 0-linked oligosaccharide-alditols by electrospray-tandem mass spectrometry after mild periodate oxidation and derivatization with 2-aminopyridine. Anal. Biochem. 259, 16–27
- 27 Vliegenthart, J. F. G. (1992) High-resolution ¹H-nuclear magnetic resonance spectroscopy of oligosaccharide-alditols released from mucin-type O-glycoproteins. In Biological Magnetic Resonance (Berliner, L. J. and Reben, J., eds.), volume 10, pp. 1–287, Plenum Press, London

Glycosylation of mucins from Xenopus laevis

463

- 28 Coppin, A., Maes, E., Morelle, W. and Strecker, G. (1999) Structural analysis of 13 neutral oligosaccharide-alditols released by reductive beta-elimination from oviducal mucins of *Rana temporaria*. Eur. J. Biochem. **266**, 94–104
- 29 Morelle, W., Guyetant, R. and Strecker, G. (1998) Structural analysis of oligosaccharide-alditols released by reductive beta-elimination from oviducal mucins of *Rana dalmatina*. Carbohydr. Res. **306**, 435–443
- 30 Capon, C., Leroy, Y., Wieruszeski, J. M., Ricart, G., Strecker, G., Montreuit, J. and Fournet, B. (1989) Structures of 0-glycosidically linked oligosaccharides isolated from human meconium glycoproteins. Eur. J. Biochem. 182, 139–152
- 31 Dua, V. K., Rao, B. N., Wu, S. S., Dube, V. E. and Bush, C. A. (1986) Characterization of the oligosaccharide alditols from ovarian cyst mucin glycoproteins of blood group A using high pressure liquid chromatography (HPLC) and high field ¹H NMR spectroscopy. J. Biol. Chem. **261**, 1599–1608
- 32 Swallow, D. M., Gendler, S., Griffiths, B., Corney, G., Taylor-Papadimitriou, J. and Bramwell, M. E. (1987) The human tumour-associated epithelial mucins are coded by an expressed hypervariable gene locus PUM. Nature (London) 328, 82–84
- 33 Lancaster, C. A., Peat, N., Duhig, T., Wilson, D., Taylor-Papadimitriou, J. and Gendler, S. J. (1990) Structure and expression of the human polymorphic epithelial mucin gene: an expressed VNTR unit. Biochem. Biophys. Res. Commun. **173**, 1019–1029
- 34 Noliet, S., Moniaux, N., Maury, J., Petitprez, D., Degand, P., Laine, A., Porchet, N. and Aubert, J. P. (1998) Human mucin gene MUC4: organization of its 5'-region and polymorphism of its central tandem repeat array. Biochem. J. 332, 739–748

Received 19 April 2000/5 July 2000; accepted 19 September 2000

- 35 Kobel, H. R. and Du Pasquier, L. (1986) Genetics of polyploid *xenopus*. Trends Genet. 2, 310–315
- 36 Watkins, W. M. (1995) Molecular basis of antigenic specificity in the ABO, H and Lewis blood-group systems. In New Comprehensive Biochemistry volume 29a: Glycoproteins (Neuberger, A. and van Deenen, L. L. M., eds.), pp. 313–376, Elsevier, Utrecht
- 37 Yamamoto, F., Clausen, H., White, T., Marken, J. and Hakomori, S. (1990) Molecular genetic basis of the histo-blood group ABO system. Nature (London) 345, 229–233
- 38 Oriol, R., Mollicone, R., Cailleau, A., Balanzino, L. and Breton, C. (1999) Divergent evolution of fucosyltransferase genes from vertebrates, invertebrates, and bacteria. Glycobiology 9, 323–334
- 39 Wickstrom, C., Davies, J. R., Eriksen, G. V., Veerman, E. C. and Caristedt, I. (1998) MUC5B is a major gel-forming, oligomeric mucin from human salivary gland, respiratory tract and endocervix: identification of glycoforms and C-terminal cleavage. Biochem. J. 334, 685–693
- 40 Quill, T. A. and Hedrick, J. L. (1996) The fertilization layer mediated block to polyspermy in *Xenopus laevis*: isolation of the cortical granule lectin ligand. Arch. Biochem. Biophys. **333**, 326–332
- 41 Gerton, G. L. and Hedrick, J. L. (1986) The vitelline envelope to fertilization envelope conversion in eggs of *Xenopus laevis*. Dev. Biol. 11, 1–7
- 42 Lindsay, L. L. and Hedrick, J. L. (1989) Proteases released from *Xenopus laevis* eggs at activation and their role in envelope conversion. Dev. Biol. 131, 126–135

123. Conclusions

Ainsi, cette étude a pour la première fois mis en évidence un polymorphisme glycannique intra-spécifique dans un modèle animal. Il se limite à l'expression de deux activités glycosyltransférasiques différentes: α 1,4-galactosyl- et α 1,3-fucosyltransférasique. L'étude directe des O-glycannes a permis dans un premier temps de décrire trois phénotypes dépendants de l'expression différentielle des deux activités P1 [Gal(+); Fuc(-)], P2 [Gal(-); Fuc(+)] et P3 [Gal(+); Fuc(+)].

L'expression différentielle de ces activités devait donc théoriquement conduire au phénotype [Gal(-); Fuc(-)] dans certains spécimens. L'existence de ce quatrième phenotype P4 a été confirmé en testant une plus grande population à l'aide des lectines spécifiques des motifs Gala et Fuc(α 1-3) comme le montre la figure A10. Un seul spécimen présentant ce phénotype a été identifié, ici visible en piste 12. On note que le profil électrophorétique de ce spécimen est différent de celui des douze, autres. En particulier, la famille de glycoprotéines de haut poids moléculaire (Hm) disparaît au profit d'un composé glycosylé de plus haut poids, présent sous forme d'une traînée dans le gel de concentration. La famille Hm ne portant jamais ni Fuc(α 1,3) ni Gal α , cette



Figure A10: Analyse en SDS-PAGE des gangues de 13 spécimens de X. *laevis*. Le gel est révélé soit au PAS (a), soit transféré sur nitrocellulose et révélé par les lectines $BSI-B_4$ (b) ou T. purpuralis(c).

observation est difficilement interprétable. Nous ne savons pas si une corrélation entre les phénotypes protéique et glycannique peut être établie.

Les différences importantes qui existent entre les mécanismes réactionnels des galactosyl- et fucosyltransférases et la ségrégation des activités que nous avons observée dans



Fig. A12: Fractionnement des mucines de la gangue oviducale d'un spécimen de Xenopus laevis par ultracentrifugation isopicnique en bromure de césium (a) puis chlorure de césium (b). La fraction de haute densité en BrCs (1) a été dialysée, lyophilisée et sous fractionnée (fractions 1.1 et 1.2). Chaque fraction obtenue a été analysée par electrophorèse en gel d'agarose (b et e) et en gel de polyacrylamide (c et f), avec un standard de mucine bronchique ou de gangue non purifiée (T). Les électrophorèses ont été révélées avec une coloration spécifique aux sucres (PAS). L, M et H correspondent aux fractions de faible, moyenne et haute mobilité (voir article 1).▼ correspond à l'origine des électrophorèses en gel d'agarose.

62

chaque spécimen suggèrent que les activités glycosyltransférasiques à l'origine du polymorphisme sont portées par deux enzymes distinctes non apparentées: une α 1,4-galactosyltransférase utilisant le substrat Fuc(α 1-2)Gal β et une α 1,3-fucosyltransférase utilisant indifféremment les substrats GlcNAc α ou GlcNAc β . Ces enzymes ne rentrent donc pas en compétition lorsqu'elles sont simultanément exprimées dans le même spécimen mais peuvent agir sur le même O-glycanne. Ainsi, un même glycanne va être finalement décoré différemment en fonction du phénotype du spécimen duquel il est extrait, comme le représente la figure A11.



Fig. Al1 : Le glycanne A a été isolé des trois phénotypes P1, P2 et P3. Par contre, les glycannes B, G et H résultent de l'action de glycosyltransférases spécifiques à chaque phénotype.

La séparation des composants de la gangue par centrifugation isopycnique en bromure de césium (BrCs) démontre que la quasi totalité du matériel glucidique est associé à des composés de densité supérieure à 1,3 (Fig. A12). L'analyse de cette fraction en gel de polyacrylamide démontre que la totalité des glycoprotéines de haut poids moléculaire présentent une densité supérieure à 1,3. Ces composés ont pu être sous fractionnés par ultracentrifugation isopycnique en chlorure de césium (CsCl) qui permet d'obtenir des densités finales plus élevée que le BrCs (Fig. A12). Nous avons ainsi différencié deux populations (1.1 et 1.2) de composés présentant des densités et des taux de sulfatation très différents. Le caractère très anionique de la population 1.1 de plus haute densité est confirmé par électrophorèse sur agarose. La population 1.2 de moindre densité est composée d'une fraction majeure cationique et d'une fraction mineure anionique, qui est certainement à l'origine de la présence de sulfate dans la population totale. L'électrophorèse en gel de polyacrylamide montre que chaque population est constituée d'une mixture de glycoprotéines de haut poids moléculaire. La sulfatation de ces composés et par extension leur caractère anionique sont sans aucun doute apportés par la substitution par les nombreux O-glycannes sulfatés que nous avons décrit lors de cette étude. De fait, en accord avec la localisation des O-glycannes à l'aide de lectines spécifiques, ces résultats confirment que les O-glycannes sont répartis de manière hétérogène sur les différentes glycoprotéines de la gangue. Ceci démontre l'existence d'une régulation très stricte de la O-glycosylation sur les différents substrats peptidiques accepteurs, comme il a déjà été discuté précédemment.

Les comportements en ultracentrifugation et électrophorèse sur agarose des glycoprotéines de xénope ont été comparés à ceux exhibés par des mucines bronchiques isolées, et sont très similaires, ce qui est encore en faveur de leur appartenance à la famille des mucines. D'autres nouveaux arguments en faveur de cette attribution ont été fournis lors de cette études: ces composés présentent un domaine résistant à la protéolyse de grande taille, et ils ne pénètrent aucun gel de polyacrylamide sans agent réducteur.

Des sérums spécifiques dirigés contre les mucines tégumentaires de X. laevis nous ont été gracieusement donnés par le Dr Hoffmann de l'université de Magdeburg. Ces anticorps ont été obtenus par immunisation de lapin avec des peptides synthétiques dérivés des séquences C- terminales de FIM-A.1, FIM-B.1 et FIM-C.1, et se sont révélés très spécifiques à chacun de ces composés. Nous avons testé ces anticorps sur les glycoprotéines de gangues de plusieurs spécimens purifiées par ultracentrifugation avec l'espoir qu'ils exhibent une forme de spécificité envers les mucines oviducales de X. laevis. Malheureusement, ainsi que l'illustre la figure A13, après transfert



Fig. A13: Electrophorèse d'une gangue gélatineuse colorée au PAS (1), ou transféré sur nitrocellulose et révélée avec des sérums anti-FIM-A.1 (2), -FIM-B.1 (3) et FIM-C.1 (4).

sur nitrocellulose les sérums anti-FIM-A.1 et -C.1 reconnaissent systématiquement sans distinction les familles de glycoprotéines Mm et Lm, tandis que le sérum FIM-B.1 ne

reconnaissait aucune famille. Ces anticorps ne peuvent donc pas nous renseigner sur la structure des mucines oviducales.

Ainsi, les résultats de l'étude partielle des glycoprotéines de la gangue gélatineuse sont en accord avec l'hypothèse selon laquelle l'ultrastructure de la gangue est composée de mucines sécrétées.

13. Etude des O-glycannes isolés de la gangue oviducale de Xenopus tropicalis. 131. Présentation des résultats

Dans le cadre de nos travaux portants sur l'analyse des profils de glycosylation d'espèces animales en vue de l'étude de nouvelles glycosyltranférses, l'espèce *Xenopus tropicalis* s'est imposée comme l'un des modèles les plus prometteurs. 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, il souffre du sérieux handicap d'être pseudotétraploïde. Dernièrement, pour pallier cet inconvénient majeur, *Xenopus tropicalis*, une espèce diploïde voisine de *X. laevis*, a été développé dans de nombreux laboratoires pour remplacer ce dernier. Le mise en place d'outils génétiques spécifiques à cette espèce, dont fait partie un projet de séquençage de son génome, laisse espérer un développement rapide de *X. tropicalis* comme modèle d'étude. Ainsi, l'étude de la glycosyltransférasiques, dont une activité α -1,4-fucosyltransférasique impliquée dans la synthèse de l'épitope Le^a, dont les supports enzymatiques sont potentiellement beaucoup plus accessibles que pour toutes les autres espèces d'amphibiens précédemment étudiées.

132. Résultats

Les résultats de cette étude sont présentés sous forme d'un article en cours de rédaction.

Article 2: Identification of the blood group Le^a determinant in the oviducal mucins of Xenopus tropicalis

Guérardel, Y., Maes, E., Strecker, G. and Kol, O.

INTRODUCTION

Xenopus eggs are surrounded by a complex extracellular matrix which consists in a vitelline envelope and a carbohydrate-rich thick jelly coat. The jelly coat is composed of high-molecular weight glycoproteins that act as a scaffold to which low-molecular weight proteins -diffusible or not- are bound [Bonnel, B.S *et al.*, 1996]. From the studies completed on the related species *Xenopus laevis*, the physico-chemical characteristics of the jelly coat suggest that the high-molecular weight glycoproteins are mostly constituted by mucin-type glycoproteins. They include: the high content, up to 60%, of sugar in jelly coat [Yurewicz, E.C. *et al.*, 1977], the very high molecular-weight and the stiff, extended conformation of these constituents [Bonnel, B.S *et al.*, 1996], the release of high quantities of O-linked mucin type oligosaccharides after alkaline treatment [Strecker, G. *et al.*, 1995; Tseng, K. *et al.*, 1997], their electrophoretic mobility, the presence of a large glycosylated central domain and their apparent high density in CsCl ultracentriguation [Guérardel, Y. *et al.*, 2000]. Furthermore, the use of lectins enable us to localize mucin-type O-glycans of known structure on these glycoproteins.

Over the last decade, studies of jelly coat O-glycans from about 20 amphibian species demonstrated two main points. First, these components showed a remarkable heterogeneity of structures, more than 300 different O-glycans being identified and about 20 new glycosyltransferase activities demonstrated. Some of these glycans present ubiquitous epitopes such as the Lewis X determinant in jelly coat of the urodele Axololtl maculatum [Strecker, G. et al., 1994] or the A-Lewis Y epitope in the urodele Pleurodeles waltl [Strecker, G. et al., 1992a], while others show non predictable sequences such as Fuc(α 1-5)[Fuc(α 1-4)]Kdn(α 2-3/6) found in Ambystoma tigrinum [Maes, E. et al., 1995]. Second, it appeared that these oligosaccharides were highly specific to each amphibian species. Each species analysed showed 1 to 3 families of structurally related O-glycans. In most cases, the species-specificity is established through the presence of one or a mixture of several original oligosaccharidic sequences. Xenopus laevis species was characterised by the presence of a blood group A epitope substituted by a GlcNAc residue in α -(1 \rightarrow 3) linkage on one out of the two O-glycan families, while the other family presents the ubiquitous sequence in amphibians Fuc(α 1-2)Gal(β 1-3)[Fuc(α 1-2)]Gal(β 1-3).

Discovery of numerous new motives in amphibian O-glycans established the occurrence of complex panels of original glycosyltransferase activities in these organisms.

Most of these activities, such as the Kdn : fucosyltransferases or Fuc : fucosyltransferases do not fit in any family of yet described glycosyltransferase. Similarly, no data permits to evaluate if glycosyltransferases involved in synthesis of ubiquitous determinants such as Le^y or A blood group antigens in amphibians are somehow related to the mammalian enzymes, or originate from convergent evolution of their coding genes. However, studies of glycosyltransferases from most amphibian species are greatly hampered by the lack of both available molecular tools, and knowledge of their genetic. For several decades, the Xenopus laevis species has been extensively used as a model for all developmental sciences. Its two main advantages are that the X. laevis embryos develop externally, and their relatively large size allows micromanipulation in much easier way than with other vertebrates. However, neither X. laevis, or other amphibian species, has provided the complementary powerful insights permitted by extensive genetic analyses. The main reason is that X. laevis is a pseudotetraploid, as a results of an additional genome duplication (relative to other vertebrates), about 30 millions years ago [Bisbee, C.A. et al., 1977]. In fact, the Xenopus genus contains nearly twenty separate species, which are except one polyploid, ranging from tetraploid (4N) to dodecaploid (12N) [Amaya, E. et al., 1998]. Consequently, polyploidy brings into the system a level of redundancy that makes genetic studies virtually impossible. On the contrary, Xenopus tropicalis is the only diploid species of its genus. It shares all the advantages of X. laevis but its diploid status make genetic studies possible. Among other advantages it has the shortest generation time of all amphibians [Amaya, E. et al., 1998], which enables faster generation of stable transgenic lines. In addition, because of the close relationship of X. tropicalis and X. laevis, many sequences between the two species are highly conserved. All of these features make X. tropicalis an exceptional vertebrate model system where classical embryological approaches can be combined with modern molecular approach.

From our point of view, X. tropicalis appears as a very promising model for the study of amphibian glycosyltransferases. Indeed, the availability of genomic data provides the unique opportunity in amphibian to identify and clone new glycosyltransferases. Here, we report the first attempt to describe the glycosylation potentialities of this newly-developed model. This work aimed to define the structure of major O-glycans synthesised by X. tropicalis, in order to give an insight into the endogenous activity of its expressed glycosyltransferases. As shown in previous studies about amphibian glycosylation, oviducal mucus provides a reliable and abundant source of O-glycans. O-glycans were released from this material by reductive alkaline treatment, purified and subjected to analysis by NMR spectroscopy. This strategy led to the sequencing of twelve oligosaccharides, out of which three comprised the trisaccharidic Le^a epitope.

EXPERIMENTALS

Isolation of oligosaccharide-alditols. Egg jelly coats were extracted in Dulbecco's phosphate buffer saline (PBS) (Sigma)/ 10 mM EDTA/ 1 mM PMSF/ 0.5% β mercapto-ethanol at 4°C overnight. The mixture was centrifuged, the supernatant was then dialysed for 72h. against water and finally freeze dried. This material was submitted to alkaline reductive degradation in 100 mM NaOH containing 1.0 M NaBH₄ at 37 °C for 72 h. The reaction was stopped by the addition of DOWEX 50x8 (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 submitted to a cationic exchange chromatography on DOWEX 50x2 (200-400 mesh, H⁺ form) to remove residual peptides. Oligosaccharidic fraction was then purified on a TSK HW-40C (Toyo-Pearl) column. Neutral and acidic oligosaccharide-alditols were fractionated by HPLC on a primary amine-bonded silica column (Supelcosyl, LC-NH₂, 4.6x250 mm, Supelco Inc., Bellefonte, USA) using a mixture of acetonitrile/H₂PO₄K 30 mM/ H2O (75/0/25 to 50/50/0 in 60 min.) with a flow rate of 1 ml/min. Oligosaccharides were detected by UV spectroscopy at 206 nm using a LDC variable-wavelenght detector (Spectra Monitor D, Milton Roy, Riviera Beach, FL, USA) connected to a Spectra-Physics Model 4100 computing integrator.

NMR spectroscopy. The NMR experiments were performed on Bruker® ASX400 spectrometer equipped with a 5mm ${}^{1}\text{H}/{}^{13}\text{C}$ mixed probe-head, operating in the pulse Fourier transform mode controlled by an Aspect 3000 computer. Each oligosaccharide was dissolved in 400 µL ${}^{2}\text{H}_{2}\text{O}$ after three exchanges with ${}^{2}\text{H}_{2}\text{O}$ at pD 7 (99.97 % atom ${}^{2}\text{H}$, Euriso-top, CEA group, Gif-sur-Yvette France) and intermediate lyophilisations. The oligosaccharides were analyzed at 300 °K. The chemical shifts (δ) were referenced to internal acetone ($\delta^{1}\text{H}=2.225$ and $\delta^{13}\text{C}=31.55$ p.p.m. in the condition used). Two dimensional homonuclear (COSY90, one relayed COSY and double relayed COSY) experiments were performed by using standard Bruker® pulse programs (cosy, cosyr1, cosyr2). The main pulses and variable delays were optimized for each pulse program and sample.

RESULTS

Both neutral and acidic oligosaccharide-alditols released from the mucin by reductive β -elimination were fractionated by normal-phase HPLC (Fig.1). Thirteen fractions were finally collected, out of which four (fractions 1 to 4) contained non-carbohydrate material and two (fractions 11 and 13) contained too complex mixtures of oligosaccharides to be analysed. On the remaining fractions, four (fractions 5, 9, 10 and 12) were composed of mixtures of two or three oligosaccharide-alditols that were analysed as such, whereas the others contained single compounds. Oligosaccharides will be referred according to the labeling of the HPLC fractions out of which they were isolated, eventually completed by A, B or C for mixtures of compounds (5A, 5B, 5C, 9A, 9B, 10A, 10B, 12A and 12B). Their structures are summarised on Fig. 2.

Monosaccharides units were identified on the basis of the measurement of vicinal coupling constants, which discriminated the α , β -galacto and the α , β -gluco configurations, and GalNAc and GlcNAc were distinguished from Gal and Glc according to the downfield shift of their H-2 atom resonance. The substitution pattern of GalNAc-ol was easily inferred from the resonance position of its H-2, H-3, H-5 and H-6,6' signals according to Kamerling (1992). In particular GalNAc-ol residues of all compounds showed H-2 chemical shift values between 4.253 and 4.404 ppm and H-3 values between 3.987 and 4.100 ppm, establishing that they were all substituted in C-3 positions. The observation of a single H6/H6' signal at 3.64-3.66 ppm for compounds **5C**, **6** and **7** established that they were not substituted in C-6 positions. On the contrary, for compounds **5A**, **5B**, **8**, **9A**, **9B** and **10A**, the marked deshielding of their H-5 signals at 4.130-4.280 ppm and of their H6/H6' signals to average chemical shifts values of 3.93/3.68 ppm, established that they were substituted in C-6 positions by β -GlcNAc residues. Then, compounds **10A**, **12A** and **12B** showed H-5 signals between 4.118 and 4.126 ppm and H6/H6' signals with average chemical shifts values of 3.85/3.49 ppm, which were very distinctive of NeuAc or Kdn substitutions in C-6 positions.

Structure of carbohydrate appendages on C-3, and eventually C-6, positions of the GalNAc-ol units from all compounds were determined owing to known structural-reportergroups. Considering that the NMR values of one branch do not influence those of the other branch, their structures can be established independently. So characteristic structural-reportergroups of all described motives are given in table 1. Analysis of each compound revealed that most of them presented well known structure. In particular, compounds **5A**, **5B**, **5C** and **10B** are found in most mucins, and have already been extensively described [Kamerling, 1992]. Then, compounds **7**, **8**, **9B**, **10A** and **12B** were previously identified among the material released from the oviducal mucins of *Bufo bufo* and *Bufo arenarum* [Morelle, W. & Strecker, G., 1997; Morelle, W. *et al.*, 1998]. For this reason we will not describe these compounds again.

On the contrary, compounds 6, 9A and 12A showed original structure, and are characteristic of the X. tropicalis species. Indeed, these compounds exhibited NMR parameters typical of the presence of the immuno group Lewis A determinant. H-1 and H-5 atom resonances of the α -fucose unit present in these three compounds are observed at δ 5.00-5.01 and 8 4.86-4.87 ppm, which are typical of Lewis A determinant [Kamerling, JP., 1992], whereas the corresponding resonances for the immuno group Lewis X determinant are found at δ 5.13 - 5.14 and δ 4.81 - 4.85 [Strecker, G. et al., 1992b]. Moreover, the set of the β -GlcNAc H-2, H-3 and H-4 resonances, at δ 3.97, 4.09 and 3.73 respectively, is also typical of the Lewis A determinant, compared with the values at δ 3.98, 3.90 and 3.95 observed for the Lewis X determinant. In compound 6, a specific shift of the GalNAc-ol H-1,1' signals to δ 3.60-3.62 ppm confirmed the presence of the disacharidic sequence GlcNAc($\beta 1 \rightarrow 3$)GalNAc (Fig. 3) [Strecker, G. et al., 1992b]. Compound 12A presented identical NMR parameters except for the specific shifts for the GalNAc-ol H-6 and H-6' signals that was characteristic of a substitution of the GalNAc-ol unit in C-6 position by a residue of NeuAc. NeuAc was distinguished from Kdn which is commonly observed in amphibian O-glycans owing to its characteristic H-3 ax, eq chemical shift values and to the presence of a NAc signal at δ 2.033 ppm. Altogether, these data clearly established the sequence of compound 6 and 12A as:

6

12A



Although 9A and 9B were analysed in mixture, most of their ¹H-NMR signals were assigned. On the basis of NMR parameters from previously analysed compound 9B [Morelle,
W. & Strecker, G., 1997] the resonances belonging from **9A** and **9B** were easily distinguished (Fig. 3). The δ values of GalNAc-ol H-1,1' demonstrated the absence of GlcNAc(β 1 \rightarrow 3)GalNAc-ol sequence, which classified **9A** (and **9B**) as belonging to the Gal(β 1 \rightarrow 3)GalNAc-ol core structure. The attachment of α -Fuc residue *via* a (1 \rightarrow 4) linkage to GlcNAc II' is established from the chemical shifts of Fuc H-1 at δ 5.012 and H-5 at δ 4.861, and GlcNAc H-3 at δ 4.048 and H-4 at δ 3.729. The two β -Gal units present in compound **9A** occur in non-reducing terminal position, as inferred by the set of their H-2, H-3 and H-4 atom resonances. These data established the structure of compound **9A** as:

Fuc(α 1-4) GlcNAc(β 1-6) Gal(β 1-3) Gal(β 1-3) Gal(β 1-3)

REFERENCES

- Amaya, E., Offield, M.F. and Grainger, M. (1998) Frog genetics: *Xenopus tropicalis* jumps into the future. *TIG* 14, 253-255
- Bisbee, C.A., Baker, M.A., Wilson, A.C., Haji-Azimi, I. and Fischberg, M. (1977) Albumin phylogeny for clawed frogs (Xenopus). *Science* 195, 785-787
- Bonnel, B.S., Reinhart, D. and Chandler, D.E. (1996) *Xenopus laevis* egg jelly coat consist of small diffusible proteins bound to a complex system of structurally stable networks composed of high-molecular-weight glycoconjugates. *Dev. Biol.* 174, 32-42
- Guérardel, Y., Kol, O., Maes. E., Lefebvre, T., Bolilly, 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
- Kamerling, J.P. and Vliegenthart, J.F.G. (1992) High resolution ¹H-nuclear magnetic resonance spectroscopy of oligosaccharide-alditols released from mucin-type Oglycoproteins. *Biol. Magn. Reson.* 10, 1-194
- Maes, E., Wieruszeski, J.M., Plancke, Y. and Strecker, G. (1995) Structure of the three Kdn-containing oligosaccharide -alditols released from oviducal secretions of *Ambystoma tigrinum*: characterization of the carbohydrate sequence Fuc(α1-5)[(Fuc(α1-4)]Kdn(α2-3/6). FEBS. Lett. 358, 205-210
- Morelle, W. and Strecker, G. (1997) Structural analysis of oligosaccharide-alditols released by reductive β-elemination from oviducal mucins of *Bufo Bufo*: characterisation of the carbohydrate sequence Gal(α1-3)GalNAc(α1-3)[Fuc(α1-2)]Gal. *Glycobiology* 7, 1129-1151
- Morelle, W., Cabada, M.O. and Strecker, G. (1998) Structural analysis of oligosaccharide-alditols released by reductive β-elimination from the jelly-coats of the anuran *Bufo arenarum*. *Eur. J. Biochem.* 252, 253-260
- Strecker, G., Wieruszeski, J.M., Fontaine, M.D., Plancke, Y. (1994) Structure of the major neutral oligosaccharide-alditols released from the egg jelly coat of *Axolotl maculatum*. Characterization of the carbohydrate sequence GalNAc(β1-4)[Fuc(α1-3)]GlcNAc(β1-3/6). *Glycobiology* 4, 605-609
- Strecker, G., Wieruszeski, J.M., Michalski, J.C. and Montreuil, J.(1992a) ¹H- and ¹³Cn.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 determinant isolated from the jelly coat of *Pleurodeles waltl* eggs. *Biochem. J.* 287, 905-909

- Strecker, G., Wieruszeski, J.M., Michalski, J.C., Alonso, C., Boilly, B. and Montreuil, J. (1992b) Characterization of Le^x, Le^y and A Le^y antigen determinants in KDN containing O-linked glycan chains from *Pleurodeles Waltlii* jelly coat eggs. *FEBS lett.* 298, 39-43
- Strecker, G., Wieruszeski, J.M., Plancke, Y. and Boilly, B. (1995) Primary structure of 12 neutral oligosaccharide-alditols released from the jelly coats of the anuran *Xenopus laevis* by reductive β-elimination. *Glycobiology* 5, 137-146
- Tseng, K., Lindsay, L.L., Penn, S., Hedrick, J.L. and Lebrilla, C.B. (1997) Characterization of neutral oligosaccharide-alditols from *Xenopus laevis* egg jelly coats by matrix-assisted laser desorption Fourier transform mass spectrometry. *Anal Biochem* 250, 18-28
- Yurewicz, E.C., Oliphant, G. and Hedrick, J.L. (1975) The macromolecular composition of *Xenopus laevis* egg jelly coat. *Biochemistry* 14, 3101-3106

	Compounds	Characteristic structural- reporter-groups
Branches in C-3 position of GalNAc-ol	<u>n hay na ana kaong pang pang ang pang pang pang pang pan</u>	
Gal(β1-3)	5A, 9A, 10B	α-Gal H-2, H-3, H-4
Fuc(α1-2)Gal(β1-3)	5B	β-Gal H-1
		α-Fuc H-1, H-5
GalNAc(α1-3)[Fuc(α1-2)]Gal(β1-3)	5C, 8, 10A	α-Fuc H-1, H-5
		β-Gal H-1, H-4
		α -GalNAc H-1, H-2
$Gal(\alpha 1-3)GalNAc(\alpha 1-3)[Fuc(\alpha 1-2)]Gal(\beta 1-3)$	7, 9B, 12B	α-Fuc H-1, H-5
		α-Gal H-1, H-2, H-3, H-4
		α-GalNAc H-2, H-3, H-4
		β-Gal H-1, H-4
Gal(β1-3)[Fuc(α1-4)]GlcNAc(β1-3)	6, 12A	α-Fuc H-1, H-5
		β-GicNAc H-3, H-4
		GalNAc-ol H-1,1'
Branches in C-6 position of GalNAc-ol		
GicNAc(β1-6)	5A, 5B, 8, 9B	β-GIcNAc H-2, H-3, H-4
HSO3(6)GlcNAc(β1-6)	10A	β-GlcNAc H-6,6'
NeuAc(α2-6)	10B, 12A, 12B	α-NeuAc H-3ax, H-3e
		GalNAc-ol H-6,6'
Gal(β 1-3)[Fuc(α 1-4)]GlcNAC(β 1-6)	9A	α-Fuc H-1, H-5
		β-GicNAc H-3, H-4
		GalNAc-ol H-6,6'

 Table 1: Characteristic structural-reporter-groups used for structural determination of oligosaccharidic motives, separated according to their position on the GalNAc-ol residue.

Table 2: ¹H NMR chemical shifts of the oligosaccharide alditols isolated from oviducal secretions of *X*. tropicalis. An additional compound containing Le^x epitope (X=Gal(β 1-4)[Fuc(α 1-3)]GlcNAc(β 1-3)[Kdn(α 2-6)]GalNAc-ol) was inserted for reference.

	Chemi	cal shift of	H (ppm))										
		5A	5B	5C	6	7	8	9A	9B	10A	10B	12A	12B	Х
	H • 1	3.75	3.78	3.79	3.62	3.80	3.78	3.787	3.78	3.79	3.779	3.60	3.78	3.61
	H - 1'	3.75	3.78	3.79	3.62	3.80	3.78	3.730	3.78	3.79	3.725	3.60	3.78	3.59
	H - 2	4.391	4.404	4.301	4.278	4.300	4.29	4.391	4.303	4.297	4.379	4.253	4.296	4.244
GalNAc-ol	H - 3	4.060	4.084	4.095	3.999	4.100	4.083	4.060	4.088	4.076	4.058	3.987	4.093	3.978
1	H-4	3.468	3.501	3.599	3.573	3.598	3.58	3.461	3.586	3.611	3.536	3.636	3.613	3.634
	H - 5	4.280	4.252	4.118	4.126	4.118	4.223	4.280	4.216	4.198	4.243	4.131	4.181	4.161
	H - 6	3.93	3.93	3.66	3.64	3.65	3.928	3.926	3.930	3.932	3.858	3.849	3.849	3.814
	H - 6'	3.685	3.709	3.66	3.64	3.65	3.705	3.678	3.705	3.712	3.488	3.493	3.493	3.474
	NAC	2.066	2.054	2.047	2.033	2.048	2.048	2.036	2.048	2.046	2.048	2.027	2.042	2.023
	H-1	4 464	4 572	4 709	-	4 719	4 680	4 464	4 706	4711	4 473		4 724	-
Gai(81-3)	H-2	3 561	3 674	3 896		3 907	3 891	3 560	3 901	3 911	3 571	-	3 919	-
Щ, т	H - 3	3.672	3.873	4.017	-	4.051	4.013	3.675	4.045	4.015	3.668	-	4.046	-
"	H - 4	3.901	3,923	4.224	-	4.242	4.220	3.903	4.235	4.209	3.894	-	4.232	-
	H-1	•	•	•	4.637	-	-	-	•	-	-	4.638	-	4.645
	п·2 Ц 2	•	-	•	3.973	-	-	-	•	-	-	3.907	-	3,901
IcNAc(B1-3)	H-3	-	-	•	4.097	•	-	•	•	-	-	4.093	•	3.09/
	L. 5	-	•	•	3.720	•	-	-	•	-	-	3./43	-	3,900
11	H-5 L.6	-	•	•	3.392	•	-	-	-	-	-	3.56	-	3.003
	H - 6'	-	-	•	3 845	-	-	-	•	-	-	4.002	-	4,020
	NAc			2	2 082		-	•	-	-	-	2.075		2.066
					2.002							2.070		2.000
	H - 1	4.538	4.551	•	-	-	4.565	4.548	4.564	4.588	-	-	-	-
	H - 2	3.707	3.716	-	•	•	3.723	3.908	3.722	3.735	-	-	-	•
	H-3	3.54	3.54	-	-	-	3.541	4.048	3.54	3.556	-	-	-	-
ICNAC(p1-0)	H-4	3.46	3.46	-	-	-	3.457	3.729	3.46	n.d.	-	-	-	-
11'	H-5	3.46	3.46	-	•	-	3.46	3.557	3.46	3.66	-	-	-	-
	H-6	3.934	3.934	-	-	•	3.94	3.988	3.941	4.368	-	-	•	-
	H - 6'	3.743	3.743	•	•	-	3.75	3.849	3.745	4.233	-	-	-	-
	NAC	2.066	2.058	-	•	•	2.056	2.068	2.062	2.058	-	-	•	•
Gal(β1-3)	ы.1				4 520		_	4 497				4 512		
III. III'	H-2		-	-	3 480			3 480	-	-	-	3 483	-	3 494
or	H - 3	-	-	-	3,635		-	3 623	-		-	3 636		3 660
	H-4	-	-	-	3.889	-	-	3.882			-	3.89	-	3,902
Gal(p1-4)														
	H - 1	-	-	5.190	-	5.180	5.188	-	5.178	5.181	-	•	5.180	-
	H - 2	-	•	4.246	-	4.477	4.288	-	4.476	4.253	-	-	4.475	-
ialNAc(a1-3) H - 3	-	-	3.928	-	4.028	3.932	-	4.037	3.929	-	-	4.027	•
. เม้ (′H-4	-	-	4.021	•	4.264	4.021	-	4.263	4.023	-	-	4.264	-
	H-5	-	-	4.156	-	4.184	4.158	-	4.184	4.165	-	-	4.186	-
	H - 6.6'	-	-	3.78	-	3.79	3.77	-	3.79	3.77	-	•	3.79	-
	NAC	-	-	2.041	•	2.058	2.047	-	2.056	2.048	-	-	2.042	-
	H - 1	-	•	-	-	5.154	-	-	5.154	-	-		5.153	-
• • • • •	H - 2	-	-	-	-	3.831	-		3.830	-	-	• -	3.830	-
Gal(α1-3)	H - 3	-	-	-	-	3.755	•	-	3.755	-	-	-	3.755	-
IV	H - 4	-	-	-	•	3.987	-	-	3.990	-	-	-	3.992	-
	H - 5	-	•	-	-	4.184	-	-	4.184	-	-	-	4.186	-
	H - 6.6'	-	-	-	-	3.79	•	-	3.79	-	-	•	3.79	-
	H - 1	-	5 222	5 390	-	5 378		-	5 365	5 383		-	5 377	-
	H-2	-	3.804	3.803	-	3.81	-	-	3.81	3 80	-	-	3.81	-
	H - 3	-	3.911	3.80		3.81	-	-	3.81	3.80	-	-	3.31	-
$Fuc(\alpha^{1}-2)$	H - 4	-	3.827	3.817	-	3.81	-		3.803	3.814	-	-	3.80	-
	H - 5	-	4.275	4.337		4.34	•		4.327	4.328	-	-	4.334	•
	H - 6	-	1.243	1.235	•	1.232	•	•	1.230	1.233	-	•	1.235	-
					4 099			5 010				5 004		E 420
Euclard A	n•1 H_2	-	-	•	4.900	-	-	3.012	-	-	-	3.004	•	5.132
ruc(α1-4)	H-2	-	-	-	3 000	•	-	3,199	-	-	-	3./90	-	3,700
or	п•э ц л	-	-	-	3.009	•	•	3.88/	-	•	-	3.79	•	3.902
$Fuc(\alpha 1-3)$	п-4 Н.б	-	-	-	J./ 80	-	-	3.791	-	-	-	3./0/	-	J./82
	H-6	-	-		005	-	-	4.001	-	•	-	4.000 1 170	-	4.010
		-	-	-	1.170	-	-	1.170	-	-	-	1.173	-	1.170
NeuAc(a2-6)) H - 3 ax	-	-	-	-	-	•	-	-	-	1.692	1.692	1.692	1.652
or	H - 3 eq	-	-	-	•	-	-	-	-	-	2.728	2.731	2.731	2.676
	H-4	-	-	-	•	-	•	-	-	-	3.666	3.67	3.67	3.603
run(α2-0)	NAC	•	•	•	•	-	-	-	•	-	2.033	2.033	2.033	-



Fig. 1: HPLC profile of the oligosaccharides released by mild alkali treatment. Material was eluted from a primary amine-bonded silica column by H_2PO_4K gradient in a mixture of water and acetonitrile.



Fig. 2: Summary of structures of O-glycans isolated from oviducal secretions of *Xenopus tropicalis*, with their corresponding 'H NMR spectra.



Fig. 3: 'H-'H two dimension COSY 90 spectra of compounds 6 (left) and 9A-9B (right).

133. Conclusions

Cette étude a permit de définir la structure de douze O-glycannes majeurs synthétisés par l'espèce Xenopus tropicalis, dont neuf avaient déjà été identifié chez d'autres espèces d'amphibiens. Les composés les plus simples ont 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 dans aucune autre espèce animale. La présence, dans une espèce d'amphibien donnée, de structures glycanniques déjà observées chez d'autres amphibiens ne met aucunement en péril la notion de spécificité d'espèce, car il est reconnu que les motifs les plus simples, tels que les noyaux, sont ubiquitairement retrouvés dans l'ensemble du règne animal. Par contre, en accord avec cette notion, 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. 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, A. et al., 1989]. Par contre, le déterminant Le^a est très commun dans les glycoprotéines de plante [Palma, A.S. et al., 2001]. De plus, on observe sur le composé 9A la présence du déterminant Le^a branché en position C-6 de l'unité de GalNAc en position terminale réductrice. Ceci 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 substitué 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 α 3/4-fucosyltransférases permettant de synthétiser l'épitope Le^a [Oriol, R., 1995; Costache, M. *et al.*, 1997a]. Ces deux gènes n'ont jusqu'à présent été identifiés que chez les homoïdés (humains et chimpanzés), tandis que l'activité α 3/4-fucosyltransférases n'a été mise en évidence que chez les homoïdés et les singes du vieux monde (Cercopitécoides) [Costache,

M. et al., 1997b], mais jamais chez les singes du nouveau mode (Platyrrhines). Un orthologue de FUT3 dont le produit présente une activité α 4-fucosyltransférasique, FUT3 rh, a dernièrement été cloné d'un représentant des Cercopitécoides [Dupuy, F. et al., 2002]. Ces études ont suggéré 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 α 3-fucosyltransférase. Néanmoins, aucune donnée n'était jusqu'à présent disponible quant à l'existence de telles activités ailleurs que chez les primates. De fait, suite à la découverte de l'épitope L^a dans les mucines de X. tropicalis, nous avons mis en place une collaboration avec le groupe du Pr A. Maftah de l'Unité de Génétique Moléculaire Animale (UMR-INRA 1061) qui est spécialisé dans l'étude des α4-fucosyltransférases chez les primates. Le but des travaux en cours est de caractériser l'enzyme responsable de la synthèse du Le^a chez X. tropicalis et de déterminer les mécanismes moléculaires à l'origine de l'émergence de cette activité spécifiquement dans cette espèce. Ce projet est rendu plus accessible grâce au développement rapide de X. tropicalis en temps que modèle de la génétique des Xenope, ainsi qu'à la possibilité d'utiliser les substrats glycanniques endogènes isolés de cette espèce pour les études d'activité enzymatique.

14. Spécificité zoologique de la fécondation et spécificité d'espèce des O-glycannes, vue personnelle

Si des espèces étroitement apparentées et habitant la même région ne s'hybrident pas, c'est qu'un mécanisme d'isolement les en empêche. Nous entendons par là n'importe quelle propriété des deux espèces qui prévient l'hybridation. Cette barrière zoologique existante entre espèces est indispensable à la rétention de leurs caractéristiques propres et est la base même de la notion d'espèce. Ces mécanismes ont été classés selon leur caractère dits "prézygotique" ou "postzygotique", et font intervenir des processus à la fois d'ordre biochimique et écologique (tableau A5). Je me pencherai ici uniquement sur les phénomènes pouvant être associés, selon cette classification, à *l'isolement gamétique* qui conduit à la spécificité zoologique de la fécondation: le spermatozoïde d'une espèce donnée ne peut féconder que l'œuf d'une même espèce. Cette spécificité est néanmoins loin d'être stricte, en particulier chez les amphibiens ou les cas d'hybridismes, même féconds, sont légions.

- 1 Mécanismes *prézygotiques* qui préviennent la formation des zygotes hybrides:
 - (a) Isolement écologique ou d'habitat. Dans une même région géographique, les populations occupent des habitats différents.
 - (b) Isolement temporel. La copulation a lieu à des moments différents.
 - (c) *Isolement éthologique*. L'attraction sexuelle entre membres d'espèces différentes est faible ou nulle.
 - (d) Isolement mécanique. Une incompatibilité anatomique entre organes génitaux empêche la reproduction.
 - (e) Isolement gamétique. Chez les organismes à fécondation externe, les gamètes mâles et femelles d'espèces différentes ne s'attirent pas. Chez les espèces à fécondation interne, les gamètes ou les gamétophytes d'une espèce sont inaptes à survivre dans les conduits génitaux d'une autre espèce.
- 2 Mécanismes post-zygotiques qui réduisent la viabilité ou la fertilité des zygotes hybrides:
 - (a) Létalité des hybrides. La viabilité des zygotes hybrides est réduite ou nulle.
 - (b) *Stérilité des hybrides*. Les hybrides F1 sont incapables de produire des gamètes fonctionnels.
 - (c) *Dégénérescence des hybrides*. La viabilité ou la fertilité des hybrides de deuxième génération est réduite.

Tableau A5: Classification des mécanismes d'isolement reproductif reconnus par Dobzhansky, cité par Ridley,M., 1996.

La barrière empêchant la fertilisation inter-spécifique des amphibiens a fait l'objet de nombreux travaux pour la plupart assez anciens. De manière intéressante, il est apparu que la gangue gélatineuse jouait un rôle important dans la sélection précoce des spermatozoïdes fécondants. Ainsi, le spermatozoïde de l'espèce Xenopus muelleri est dans l'incapacité de fertiliser l'œuf de X. laevis, sauf si ce dernier est entouré par la gangue de X. laevis [Blackler, A.W. & Gecking, C.A., 1972]. Le même phénomène a été observé entre X. laevis et Xenopus borealis, à la différence que chez ces deux espèces la spécificité de fécondation est unilatérale: le sperme de X. borealis féconde les deux espèces, tandis que celui de X. laevis ne peut fertiliser l'œuf de X. borealis que si ce dernier est entouré de la gangue de X. laevis [Brun, R. & Kobel, H.R., 1972]. Par contre, un œuf de X. laevis enfoui dans une gangue de X. borealis n'est plus fécondable par le sperme de X. laevis. Ceci démontre sans ambiguïté que chez des espèces apparentées la gangue gélatineuse, à elle seule, est capable de sélectionner les spermatozoïdes de manière spécifique à l'espèce. D'un point de vue phénoménologique, un parallèle intéressant peut être dressé entre la sélection spécifique d'espèce du spermatozoïde par les gangues et leur glycosylation spécifique. Le facteur responsable de cette sélection n'a pas été déterminé dans les deux exemples que j'ai cité. Deux hypothèses sont envisageables: soit la gangue d'une espèce est dépourvue en un élément spécifique indispensable à l'activation du spermatozoïde d'une autre espèce, soit la gangue inhibe la progression de tout spermatozoïde étranger. Dans le cas de X. laevis et X. borealis, il a été observé que le spermatozoïde de X. laevis était incapable de traverser la couche la plus interne de la gangue de X. borealis et s'accumulait à l'interface des couches J1 et J2, la cause d'un tel blocage n'ayant jamais été élucidée.

Plus récemment, il a été montré que la progression du spermatozoïde de X. laevis au travers de la gangue gélatineuse était inhibée de façon dose dépendante par incubation préalable des œufs dans la lectine de germe de blé (WGA) [Reinhart, D. et al., 1998] qui reconnaît principalement les résidus de GlcNAc en positions terminales. La WGA n'affecte pas la motilité propre du spermatozoïde mais inhiberait les interactions spécifiques entre le spermatozoïde et la partie glycannique de la matrice macromoléculaire (mucines) de la gangue. Cette inhibition n'est pas le résultat d'un encombrement stérique comme le prouve le fait que d'autres lectines interagissent fortement avec les macromolécules de la gangue sans pour autant inhiber la fertilisation [Mozingo, N.M. & Hedrick, J.L., 1999]. Ces études démontrent donc que le spermatozoïde doit interagir spécifiquement avec les glycannes de la gangue pour traverser cette dernière. Ce résultat apporte un éclairage *a posteriori* très révélateur aux études plus anciennes sur la sélection des spermatozoïdes par la gangue citées ci dessus. En effet, on peut émettre l'hypothèse selon laquelle la spécificité de glycosylation



Fig : Anneau d'espèce de la salamandre *Ensatina* dans l'ouest des Etats-Unis. Une espèce (*E. oregonensis*) se trouve dans le Nord, jusque dans l'état de Washington. A partir du Nord de la Californie, elle constitue vers le Sud un anneau à peu près continu autour de la vallée de San Joaquin. D'une localité à l'autre, les formes diffèrent et ont été décrites sous des noms différents. Dans le Sud de la Californie, dans les localités ou les formes occidentales et orientales se rencontrent (en noir sur la figure), les taxons occidentaux et orientaux de l'anneau se comportent comme des espèces à part entière. D'après Ridley, M., 1996.

interviendrait dans la sélection des spermatozoïdes par l'intermédiaire de mécanismes de reconnaissances spécifiques entre spermatozoïdes et gangues. D'un point de vue plus général, un tel phénomène s'apparenterait clairement au mécanisme d'isolement reproductif prézygotique dit "d'isolement gamétique" décrit par Dobzhansky. Néanmoins, ainsi que je l'ai décrit ci dessus pour *X. laevis* et *X. borealis*, cette sélection ne semble pas toujours suffisante pour conduire à un isolement reproductif total. Si l'existence d'interactions spécifiques entre spermatozoïde et gangue est une certitude, leurs bases moléculaires sont encore inconnues Soit elles sont de type 'récepteur-ligand' et font intervenir des composés membranaires à activité lectinique, soit elles interactions ont été mises en évidence dernièrement entre les spermatozoïdes et l'œuf de truite, mais leurs propriétés spécifiques d'espèce n'ont pas été étudiées [Yu, S. et al., 2002]. Ce dernier type d'interaction cellulaire spécifique d'espèce a également été mis en évidence chez les éponges. L'étude de ses bases moléculaires s'intègre dans les travaux que nous avons effectué et seront rapportés dans la suite de ce mémoire.

L'hypothèse selon laquelle la spécificité de glycosylation participerait à la formation de la barrière zoologique de fécondation chez les amphibiens a également des répercussions au niveau du phénomène de spéciation. En effet, l'évènement capital à l'origine d'une espèce nouvelle est son isolement reproductif. Dans l'abstrait, à partir d'une espèce unique composée d'individus interféconds, une ou plusieurs variations génétiques devraient se répandre dans une partie de l'espèce, conduisant les porteurs d'une même variation à procréer exclusivement entre eux. De nombreuses questions se posent quant aux circonstances et au mécanisme de la spéciation. Par exemple, qu'elles sont les modifications génétiques qui caractérisent ou qui entraînent l'isolement reproductif? Au vu des éléments que j'ai exposé précédemment, l'acquisition d'un profil de glycosylation spécifique par la gangue de l'œuf pourrait participer à l'isolement reproductif d'espèces nouvellement formées par sélection précoce des spermatozoïdes au travers d'un mécanisme de reconnaissance spécifique. L'acquisition de l'isolement reproductif est un phénomène graduel, comme l'est certainement l'acquisition de la spécificité d'espèce de la glycosylation. Elle peut être directement observée lors de la formation "d'anneaux d'espèces". C'est par exemple le cas dans le sud de la Californie où deux espèces distinctes de salamandres non interfécondes (Ensatina eschscholtzii et Ensatina klauberi) coexistent au Sud de la vallée de San Joaquin et dont l'origine commune a été déterminée. Schématiquement (Fig. A14), à partir d'une espèce unique vivant au Nord de la vallée, les salamandres se seraient étendues vers le Sud de part et d'autre de la vallée de San

Joaquin, les populations de la côte pacifique auraient acquis le phénotype de *E. eschscholtzii* alors que les populations orientales auraient acquis le phénotype de *E. klauberi*. En divers points de cette expansion, les formes occidentales et orientales se seraient rencontrées et auraient données naissance à des formes hybrides, démontrant qu'elles n'auraient pas encore acquis d'isolement reproductif. Par contre plus au Sud, les espèces auraient suffisamment divergé de part et d'autre pour devenir inter-stériles et constituer des espèces à part entière. La réalité de ces anneaux d'espèce démontre que des différences intraspécifiques peuvent s'accumuler au point de devenir interpécifiques. Il doit en être de même pour la glycosylation des gangues de *X. laevis*, sans que cela ne porte préjudice à la fécondation, en est une autre manifestation. Des différences intra-spécifiques discrètes dans les activités glycosyltransférasiques se transforment par accumulation en différences inter-spécifiques que nous observons.

La spéciation fait intervenir une multitude de facteurs qu'il est souvent impossible d'observer *a posteriori*. Je ne tente aucunement de réduire les causes de la spéciation des amphibiens aux changements intervenant dans le profil de glycosylation des gangues. En particulier, aucune donnée ne permet de déterminer si l'acquisition de la spécificité glycannique des gangues précède ou suit l'acquisition de l'isolement reproductif. Néanmoins, le fait que les gangues gélatineuses sont en "première ligne" de la fécondation par leur potentialité d'interaction et de sélection des spermatozoïdes me semblait suffisamment suggestif pour émettre l'hypothèse selon laquelle la glycosylation pourrait tenir un rôle dans la spéciation. Que ces deux phénomènes soient ou non reliés, on peut noter que les changements de glycosylation de la gangue entre espèces 'ne semblent pas influer sur les propriétés rhéologiques de la gangue. Cette dernière conserve donc sa fonction protectrice quelque soit son statut de glycosylation.

2. Le modèle Caenorhabditis elegans.

21. Introduction

Actuellement, *Caenorhabditis elegans* est l'animal le plus simple et le mieux caractérisé: son anatomie et son développement ont été totalement déterminés [Wood, W.B., 1998], et le séquençage de son génome est complet depuis plusieurs années [The *C. elegans* sequencing consortium, 1998]. 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. En effet, des gènes codant pour des polypeptides GalNAc-transférases [Hagen, F.K. & Nehrke, K., 1998], GlcNAc-transférases [Chen, S. *et al.*, 1999; Warren, C.E., *et al.*, 2001] et fucosyltransférases [DeBose-Boyd, R.A. *et al.*, 1998] actives ont été caractérisés. De fait certains auteurs estiment 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, J.W, *et al.*, 1999]. *C. elegans* apparaît donc comme un modèle idéal pour l'étude du rôle des glycoconjugués dans nombre de processus biologiques reliés au développement.

Si l'expression de glycosyltransférases a mis en évidence l'existence d'activités glycosyltransférasiques ubiquitaires, les activités exactes de certaines de ces enzymes sont encore incertaines. Ainsi, sur les trois homologues de l'UDP-N-acétylglucosamine: α-3-Dmannoside B-1,2-N-acétylglucosaminyltransférase I (GnT I) de mammifères qui ont été exprimées (gly-12, -13 et -14), seules gly-12 et gly-14 présentaient une activité GlcNActransférasique observable [Chen, S. et al., 1999]. L'inactivation de gly-12 ou/et gly-14, ne provoque aucun phénotype observable alors que celle de gly-13 provoque un arrêt du développement au stade larvaire [Chen, S. et al., 2001]. Bien que cette enzyme semble indispensable au développement de C. elegans son activité endogène n'a pas encore été définie. De même, bien que C. elegans ne synthétise pas le motif Le^x in vivo, un extrait de C.elegans adulte est capable de le synthétiser à partir du substrat accepteur exogène Gal (α) -4)GlcNAc β -R. Une α -1,3-fucosyltransférase putative (CEFT-1) de C. elegans exprimée en cellule COS-7 a permis d'isoler cette activité. De fait, bien qu'une activité précise ait été observée, l'activité endogène de CEFT-1 est toujours inconnue. Malgré l'existence de motifs conservés chez les différentes familles de glycosyltransférases, ces deux exemples illustrent la difficulté d'étude de glycosyltransférases isolées d'un modèle dont les substrats endogènes sont a priori inconnus.

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 établi chez *C. elegans*. Par contre la structure de six glycosphingolipides appartenant à la série arthro-, trois neutres et trois acides (Fig. A6), avait été établie à partir de lysats totaux de *C. elegans* [Gerdt, S. *et al.*, 1997; Gerdt, *et al.*, 1999]. Ces structures présentaient des homologies importantes avec le nématode parasite *Ascaris suum* [Lochnit, G. *et al.*, 1997], ce qui confortait la position de *C. elegans* en tant que modèle d'étude des nématodes parasites.

 $\label{eq:Glcb-Cer} Glc\beta-Cer \\ Man(\beta1-4)Glc\beta-Cer \\ GlcNAc(\beta1-3)Man(\beta1-4)Glc\beta-Cer \\ GalNAc(\beta1-4)[Pcho]GlcNAc(\beta1-3)Man(\beta1-4)Glc\beta-Cer \\ Gal(\alpha1-3)GalNAc(\beta1-4)[Pcho]GlcNAc(\beta1-3)Man(\beta1-4)Glc\beta-Cer \\ Gal(\beta1-3)Gal(\alpha1-3)GalNAc(\beta1-4)[Pcho]GlcNAc(\beta1-3)Man(\beta1-4)Glc\beta-Cer \\ Gal(\beta1-3)GalNAc(\beta1-4)[Pcho]GlcNAc(\beta1-3)Man(\beta1-4)Glc\beta-Cer \\ Gal(\beta1-3)GalNAc(\beta1-4)[Pcho]GalNAc(\beta1-3)Man(\beta1-4)GalAc(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1-4)GalAc(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1-4)[Pcho]GalAc(\beta1-3)Man(\beta1$

Tableau A6: Structure des domaines glycanniques de six glycolipides extraits de populations mixtes de C. elegans. D'après Gert, S. et al., 1997 et Gert, S. et al., 1999.

De même, deux études s'étaient penchées sur les glycosaminoglycannes et avaient montré que *C. elegans* synthétisait des quantités très importantes de chondroïtines non sulfatées et dans une moindre mesure des héparane-sulfates [Yamada, S. *et al.*, 1999; Toyoda, H. *et al.*, 2000].

22. Analyse des O-glycannes de *C. elegans* 221. Résultats

Les résultats sont présentés sous forme d'un article publié.

Article 3: The nematode Caenorhabditis elegans synthesises unusual O-linked glycans Identification of glucose-substituted mucin-type O-glycans and short chondroitin-like oligosaccharides

Guérardel, Y., Balanzino, L., Maes, E., Leroy, Y., Coddeville, B., Oriol, R. and Strecker, G.

The nematode *Caenorhabditis elegans* synthesizes unusual O-linked glycans: identification of glucose-substituted mucin-type O-glycans and short chondroitin-like oligosaccharides

Yann GUÉRARDEL*1, Luis BALANZINO[†], Emmanuel MAES*, Yves LEROY*, Bernadette CODDEVILLE*, Rafael ORIOL[†] and Gérard STRECKER*

*Laboratoire de Chimie Biologique et Unité Mixte de Recherche du CNRS 8576, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq, France, and †INSERM U 504/Université Paris Sud XI, Villejuif, France

The free-living nematode *Caenorhabditis elegans* is a relevant model for studies on the role of glycoconjugates during development of multicellular organisms. Several genes coding for glycosyltransferases involved in the synthesis of N- and O-linked glycans have already been isolated, but, apart from repetitive dimers of glycosaminoglycans, no detailed structure of either type of component has been published so far. This study aimed to establish the structures of the major O-glycans synthesized by *C. elegans* to give an insight into the endogenous glycosyltransferase activities expressed in this organism. By the use of NMR and MS, we have resolved the sequence of seven of these components that present very unusual features. Most of them were characterized by the type-1 core substituted on Gal and/or GalNAc by $(\beta 1-4)$ Glc and $(\beta 1-6)$ Glc residues. Another compound exhibited the GalNAc $(\beta 1-4)$ *N*-acetylglucosaminitol sequence in the terminal position, to which was attached a tetramer of β -Gal substituted by both Fuc and 2-*O*-methyl-fucose residues. Our experimental procedure led also to the isolation of glycosaminoglycan-like components and oligomannosyl-type N-glycans. In particular, the data confirmed that *C. elegans* synthesizes the ubiquitous linker sequence GlcA $(\beta 1-3)$ Gal $(\beta 1-3)$ Gal $(\beta 1-4)$ Xyl.

Key words: glycosaminoglycan, mass spectrometry, NMR.

INTRODUCTION

Until now, most studies of N- and O-glycosylation in nematodes concentrated on the identification of immunogenic glycoconjugates of parasitic nematodes. Indeed, it has been suggested that most parasite antigens are either excreted-secreted (ES antigens) or surface glycans with unique structures [1]. Systematic sequencing of glycans isolated from parasitic nematodes demonstrated that these organisms synthesize a wide variety of Nglycans presenting unusual substitutions. They include phosphocholine- and chito-oligomer-substituted glycans in filarial nematodes [2], D-tyvelose capped glycans in *Trichinella spiralis* [3] or highly fucosylated chitobiose core structures in *Haemonchus contortus* [4].

The only O-linked mucin-type glycans described in nematodes so far were isolated from the *Toxocara* excretory-secretory (TES) product released by the infective larval stage of *Toxocara canis* and *T. cati* [5]. The two major structures were trisaccharidic Oglycans characterized by the presence of 2-O-methyl-fucose (2-O-Me-Fuc) and 4-O-methyl-galactose.

The free-living nematode *Caenorhabditis elegans* has been recognized as a good model for the study of the role of glycans in eukaryotes as well as a prototypic model for parasitic nematodes because of the detailed information available on its genome, development and physiology. So far, only glycolipid[6,7] and glycosaminoglycan (GAG)-type structures have received proper attention. In particular, the presence of large amounts of unsulphated chondroitin and minute amounts of heparan sulphate-type GAGs have been identified in *C. elegans* [8,9].

No detailed structures of either O- or N-linked oligosaccharides isolated from C. elegans have been published. However, the identification of three genes (Gly-12, -13 and -14) homologous to mammalian N-acetylglucosamine: α -3-D-mannoside β -1,2-Nacetylglucosaminyltransferase (GnT1) suggests the presence of N-linked oligosaccharides in C. elegans [10]. Gly-12 and Gly-14 were shown to encode active glycosyltransferases, but only mutation of Gly-13 induced defects in embryogenesis, whereas Gly-12- and Gly-14-null mutants displayed wild-type phenotypes [11]. Similarly, C. elegans expresses a family of functional polypeptide N-acetylgalactosaminyltransferases (ppGaNTases), involved in the initiation of mucin-type O-glycosylation [12]. Inhibition of these ppGaNTases by antisense RNAs resulted in severe defects in embryonic development [13]. Thus, these data not only suggest the occurrence of O-linked glycans but also suggest that this type of glycosylation plays a critical role in the embryogenesis of C. elegans.

Also, in agreement with abundant N- and O-glycosylation in C. elegans, 30 different putative fucosyltransferase genes, including $24 \alpha - 2$, five $\alpha - 3$ and one $\alpha - 6$ fucosyltransferase gene, have

Abbreviations used: 2-O-Me-Fuc, 2-O-methyl-fucose; GalNAc-ol, *N*-acetylgalactosaminitol; GlcNAc-ol, *N*-acetylglucosaminitol; Xyl-ol, xylitol; Δ HexA, 4-deoxy- α -t-threohex-4-enepyranosyluronic acid; ESI-MS, electrospray ionization MS; TES, *Toxocara* excretory-secretory; GAG, glycosaminoglycan; MALDI-TOF, matrix-assisted laser-desorption ionization-time-of-flight; ROESY, rotating-frame Overhauser enhancement spectroscopy; NOE, nuclear Overhauser effect; HMQC, heteronuclear multiple quantum coherance; β -Glcp, β -glucopyranose; β -Galp, β -galactopyranose; Hex, hexose; HexA, hexuronic acid; HexNAc-ol, *N*-acetylexosaminitol.

¹ To whom correspondence should be addressed (e-mail Yann.Guerardel@univ-lille1.fr).

been identified in databanks ([14] and R. Oriol, unpublished work). One of these, CEFT-1, has been cloned; it can fucosylate unsialylated type 2 acceptors to generate the Le^x antigen *in vitro*, although *C. elegans* does not express such an antigen [15]. This example points out the difficulties of characterizing the exact activity of an isolated glycosyltransferase without any knowledge of its endogenous glycan substrate.

This work aimed to define the structure of major O-glycans synthesized by C. elegans, to give insight into the endogenous activity of glycosyltransferases cloned from this worm and to facilitate studies about their exact role during embryogenesis. Considering the extreme structural diversity of O-glycans, and the absence of real conserved motifs, the analysis of their detailed structure in a new model may be very complex. Thus, in this first study, we have worked on a total extract of a mixed-stage wildtype strain to obtain enough material for multiple NMR experiments. O-linked glycans were released from the C. elegans extract by a reductive alkaline treatment. Through the use of homoand heteronuclear NMR experiments, electrospray ionization MS (ESI-MS) and GC/MS analyses, we identified two families of unusual mucin-type O-glycans. Both families were characterized by the original presence of multiple β -Glc residues in nonreducing terminal positions. By analysing the GAG chains released by β -elimination, this study confirmed that C. elegans synthesizes the linker sequence $GlcA(\beta 1-3)Gal(\beta 1-3)G$ 4)Xyl-ol (where Xyl-ol is xylitol) that initiates the biosynthesis of most GAGs, except dermatan sulphate and hyaluronate. The identification of short chondroitin-type oligosaccharides with a GalNAc residue in the reductive terminal position also raised the possibility of GAG chains directly linked to the protein moieties through a GalNAc. The release of oligosaccharides in mild alkali also permitted us to isolate mannosyl-type N-glycans synthesized in large quantities by C. elegans.

EXPERIMENTAL

C. elegans cultures

C. elegans wild-type strain N2 Bristol was grown in either solid or liquid cultures. Solid cultures were achieved by growing the nematodes on NGM agar plates seeded with Escherichia coli OP50, a leaky uracil-requiring strain [16]. For yields of worms that are in the range of grams, C. elegans is conveniently grown in bulk by using E. coli strain NA22 as a food source [17]. Briefly, the bacteria were inoculated in 21 flasks containing 500 ml of 3XD medium (10.5 g/l of Na₂HPO₄, 4.5 g/l of KH₂PO₄, 0.6 g/l of NH₄Cl, 15 g/l of casein hydrolysate, 24 g/l of glycerol and 3 ml/l of 1 M MgSO₄) [18], and incubated in an orbital shaker overnight at 37 °C and 250 rev./min. The bacteria were harvested by spinning for 15 min at 4500 g in sterile 500 ml centrifuge bottles. The supernatant was discarded and the bacteria weighed.

The bacterial pellets were resuspended in 500 ml of S complete medium (5.9 g/l NaCl, 50 ml/l of 1 M potassium phosphate, pH 6.0, 5 mg/l cholesterol, 10 ml/l of 1 M potassium citrate, pH 6.0, 10 ml/l trace metals solution, 3 ml/l of 1 M CaCl₂ and 3 ml/l of 1 M MgSO₄) [19] at a final density of 30–50 g/l in 2 l flasks. Then, worms were washed off with sterile M9 buffer (6 g/l Na₂HPO₄, 3 g/l KH₂PO₄, 5 g/l NaCl and 0.25 g/l MgSO₄ · 7H₂O) [16] from fresh, uncontaminated solid cultures and inoculated to the bacteria.

The cultures were incubated in an orbital shaker at 20 °C and 200 rev./min for several days until they were clear of bacteria and had ceased to grow. The cultures were collected and allowed to settle overnight at 4 °C into 500 ml glass cylinders. Most of the medium was then aspirated, and the worms were transferred to 50 ml Falcon tubes, pelleted by spinning for 5 min at 1000 g, and

90

washed three times with 50 ml of M9 buffer. Then nematodes were harvested and decontaminated by sucrose floatation [17]. Briefly, the washed worms were resuspended in 50 ml Falcon tubes in 25 ml of ice-cold 0.1 M NaCl, and then 25 ml of ice-cold 60 % (w/w) sucrose was added and mixed by inversion. Quickly, the tubes were centrifuged at 1500 g for 10 min and the cap of floating nematodes was collected using a Pasteur pipette and distributed into two 15 ml pre-weighed Falcon tubes. The worms were washed three times with M9 buffer, weighed and stored at -80 °C. Freshly cleared cultures produced a yield of worms equal to about half the weight of the bacteria used [20].

Some nematode strains used in this work were provided by the *Caenorhabditis* Genetics Center, which is funded by the National Institutes of Health National Center for Research Resources (NCRR), Buenos Aires, Argentina.

Isolation of oligosaccharide-alditols

Freeze-dried C. elegans were ground and repeatedly delipidated by chloroform/methanol (2:1) extraction. The delipidated material was submitted to alkaline reductive degradation 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 cationexchange 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 submitted to cation-exchange chromatography on a Dowex 50 × 2 column (200–400 mesh, H⁺ form) to remove residual peptides and desalted on a Bio-Gel P2 column (Bio-Rad). The oligosaccharide fraction was then purified from macromolecular material on a Bio-Gel P6 column (Bio-Rad).

Fractionation of oligosaccharide-alditols

Compounds were fractionated by HPLC on a primary aminebonded silica column (Supelcosyl, LC-NH₂, 4.6×250 mm; Supelco, Bellafonte, PA, U.S.A.) using a mixture of acetonitrile/30 mM KH₂PO₄/water (80:0:20-50:50:0 in 65 min) at a flow rate of 1 ml/min. Oligosaccharides were detected by UV spectroscopy at 206 nm using a variablewavelength detector (Spectra Monitor D, Milton Roy, Riviera Beach, FL, U.S.A.) connected to a Spectra-Physics Model 4100 computing integrator.

Matrix-assisted laser-desorption ionization-time-of-flight (MALDI-TOF) MS

The molecular mass of the oligosaccharides was measured by MALDI on a Vision 2000 time-of-flight instrument (Finnigan Mat) equipped with a 337 nm UV laser. Samples were dissolved in water at a concentration of 100 pmol/ μ l; 1 μ l of the solution was mixed with an equal volume of the matrix solution on the target and then allowed to crystallize at room temperature. We used 2,5-dihydroxybenzoic acid (10 mg/ml; dissolved in water/ methanol, 4:1, v/v) and 3-aminoquinoline for the neutral and acidic oligosaccharides, respectively.

ESI-MS

All MS measurements were carried out in positive-ion mode on a triple quadrupole instrument (Micromass, Altrincham, Cheshire, U.K.) fitted with an atmospheric pressure ionization electrospray source. A mixture of polypropylene glycol was used to calibrate the quadrupole mass spectrometer. The sample (2μ) , dissolved in methanol/water (50:50) and 1% formic acid at a concentration of 10 pmol/ μ l, was infused using the nanoflow probe at 50 nl/min. Quadrupole was scanned from 600 to 1800 Da with a scan duration of 3 s and a scan delay of 0.1 s. The samples were sprayed using 1.4 kV needle voltage and the declustering (cone) was set at 70 V. For collision-induced dissociation (CID) experiments, the pressure of argon in the cell was set at 4×10^{-3} mbar and the collision energy was set to values ranging from 40 to 75 eV.

NMR spectroscopy

The NMR experiments were performed on Bruker® ASX400 and DMX600 spectrometers both equipped with a 5 mm ¹H/¹³C mixed-probe head, operating in the pulse Fourier-transform mode controlled by an Aspect 3000 computer. Each oligosaccharide was dissolved in 400 μ l of ²H₂O after three exchanges with ²H₂O at pH 7 (99.97% atom ²H, Euriso-top, CEA group, Gif-sur-Yvette, France) and intermediate lyophilizations. The oligosaccharides were analysed at 300 K. The chemical shifts (δ) were referenced to internal acetone ($\delta^{1}H = 2.225$ and $\delta^{13}C =$ 31.55 p.p.m. under the conditions used). Two-dimensional homonuclear [COSY90, single-relayed COSY and double-relayed COSY, TOCSY, rotating-frame Overhauser enhancement spectroscopy (ROESY)] and heteronuclear multiple quantum coherance (HMQC) experiments were performed by using standard Bruker® pulse programs (cosy, cosyr1, cosyr2, mlevtp, roesyprsh and invbtp respectively). The main pulses and variable delays were optimized for each pulse program and sample.

Composition of monosaccharides

Monosaccharides were systematically analysed by GC/MS as three different derivatives: per-trimethylsilyl [21], per-heptafluorobutyryl [22] and alditol-acetate derivatives. For alditolacetates analysis, oligosaccharide-alditols were hydrolysed in 4 M trifluoroacetic acid for 4 h at 100 °C. They were reduced with sodium borodeuterite in 0.05 M NH₄OH for 4 h. Reduction was stopped by dropwise addition of acetic acid until pH 6 was reached and borate salts were co-distilled by repetitive evaporation in dry methanol. Per-acetylation was performed in acetic anhydride at 100 °C for 2 h.

Carboxyl reduction

Before methylation, GlcA-containing oligosaccharide-alditols were alternatively treated as follows. Compounds were incubated in a mixture of dry DMSO (100 μ l) and iodomethane (100 μ l) for 2 h to convert carboxylic acid groups into methyl esters. After lyophilization, methyl ester function was reduced in a solution of sodium borodeuterite in dry methanol for 12 h. Acetic acid was added to pH 6 and boric acid was removed as methyl borate by repetitive evaporation in methanol. Dried material was further purified by gel filtration on Sephadex G-10 (Pharmacia) and subjected to methylation analysis.

Methylation

Native oligosaccharide-alditols or carboxyl-reduced oligosaccharides were per-methylated according to Ciucanu and Kerek [23]. Briefly, compounds were incubated overnight in a suspension of 200 mg/ml NaOH in dry DMSO (100 μ l) and iodomethane (100 μ l). The methylated products were extracted in chloroform and washed with water. Lyophilized product was subjected to methanolysis in 500 μ l of 0.5 M HCl in anhydrous methanol at 80 °C for 20 h. After drying, products were per-acetylated in 200 μ l of acetic anhydride and 50 μ l of pyridine overnight at room temperature. The reagents were evaporated, and the sample was dissolved in chloroform before analysis in GC/MS.

RESULTS

Purification

The purification scheme led to the isolation of 12 major fractions by normal-phase HPLC (Figure 1), containing 15 different oligosaccharides. Unlabelled peaks contained either mixtures that were too complex or compounds in quantities too small to be analysed exhaustively. Three fractions (fractions 6, 10 and 12) contained a mixture of two oligosaccharides that were then analysed as a blend. All these compounds were classified in three types, depending on their nature: compounds 2, 3, 4, 5, 6A, 6B and 7 appeared to be mucin-type O-glycans, compounds 1, 8, 12A and 12B were GAG-derived oligosaccharides and compounds 9, 10A, 10B and 11 were oligomannosyl-type N-glycans. Each one was analysed by NMR, and the chemical shifts of their constituent ¹H, and for some of them ¹³C, are given in Table 1 for mucin-type O-glycans, Table 2 for GAG-type oligosaccharides and Table 3 for N-glycans.

Mucin-type O-glycan analysis

The ¹H NMR spectrum of the oligosaccharide-alditols 2, 3, 4, 5, 6A, 6B and 7 were assigned using two-dimensional COSY, TOCSY and ROESY, and ¹H-detected ¹H-¹³C HMQC experiments also allowed the assignment of the ¹³C chemical shifts for compounds 4, 5, 6A and 6B. After structural analysis, it appeared that oligosaccharide-alditols 2, 3, 4, 6A, 6B and 7 were part of a homologous family of O-glycans, whereas compound 5 presented largely distinct features. Thus, analysis of this last compound will be discussed separately at the end of this section.

Based on the ${}^{3}J_{H,H}$ coupling constant values, the three sugarspin systems of β -glucopyranose (β -Glcp), β -glucopyranuronic acid and β -galactopyranose (β -Galp) were identified in all five samples 2, 3, 4, 6 and 7. The N-acetylgalactosaminitol (GalNAcol) unit was identified owing to the assignment of its characteristic ¹H and ¹³C atom resonances, and the confirmation was given by GC/MS of the corresponding hexosaminitol acetate. The two-dimensional COSY spectrum of compound 4 (Figure 2) clearly showed the presence of 3,6-di-substituted GalNAc-ol, as indicated by the significant downfield displacements of H-3 and H-6,6' at & 4.073, 3.986 and 3.788 p.p.m. respectively. The absence of downfield glycosylation shifts for the ring protons of β -Glc and β -GlcA indicated that these two units were terminal. On the contrary, downfield displacements were observed for the H-2, H-3 and H-4 atom resonances of β -Gal. The H-5 resonance of β -Gal was assigned from the H-1,5 cross-peak in the ROESY spectrum (Figure 3), and consequently found at & 3.97 p.p.m. in the ¹H-¹³C HMQC spectrum. Comparison of the ¹³C chemical shifts for β -Gal with literature values for β -methyl galactoside [24] identified the substitution position at C-3. In agreement with the significant deshielding of H-5 and the light shielding of C-5, a second substitution position was established at C-6. Finally, the H-5, H-6, C-5 and C-6 resonances of β -Gal were assigned by elimination after all the other ¹H and ¹³C resonances had been allocated.

The ROESY spectrum (Figure 3) showed the following contacts: (i) between GlcA III H-1 and signals at δ 3.52 p.p.m. (GlcA III H-3), δ 3.721 p.p.m. (GlcA III H-5) and δ 3.834 p.p.m. (Gal II H-3), (ii) between Glc IV H-1 and signals at δ 3.987 p.p.m. (Gal II H-6') and δ 4.073 p.p.m. (Gal II H-6), (iii) between Gal II H-1 and signals at δ 4.073 p.p.m. (Gal II H-6), (iii) between Gal II H-1 and signals at δ 4.073 p.p.m. (Gal II H-6), (iii) between Gal II H-1 and signals at δ 4.073 p.p.m. (Gal II H-6), and δ 3.97 p.p.m. (Gal II H-6) and \delta 3.97 p.p.m. (Gal II H-6) and δ 4.073 p.p.m. (Gal II H-6) and A 4.073



Figure 1 HPLC profile of the oligosaccharides released by mild alkali treatment

Material was eluted from a primary amine-bonded silica column by a KH2PO4 gradient in a mixture of water and acetonitrile.

(iv) between Glc II' H-1 and signal at δ 3.788 p.p.m. (GalNAcol H-6'). On the basis of the data obtained, it is concluded that compound 4 has the following structure:



Compound 2 showed a $[M-H]^-$ at m/z 721 in MALDI-TOF. In accordance with composition analysis, it suggested the presence of two hexose residues, one hexosaminitol residue and one hexuronic acid residue. Integration of signals from the twodimensional COSY spectrum of compound 2 confirmed that it was composed of one β -Glc, one β -Gal, one β -GlcA and one GalNAc-ol. Based on the set of its H-3, H-5 and H-6,6' atom resonances it was concluded that the GalNAc-ol unit is 3-linked. The H-2, H-3, H-4 and H-6,6' atom resonances of β -Gal were identical to those observed in compound 4, as well as the NMR parameters of terminal β -Glc and β -GlcA. Thus the NMR evidence permits us to write the structure of compound 2 as follows:



Compound 3 presented the same $[M-H]^-$ at m/z 721 as the latter compound. Based on the relayed COSY spectrum (Figure 2) as well as on the composition analysis, the oligosaccharide-alditol 3 was defined as a tetramer containing two β -Gal, one β -GlcA and one GalNAc-ol unit. The H-2, H-3 and H-4 atom resonances of the β -Gal II unit were observed at δ values that signify a 3substitution with β -GlcA, whereas the corresponding signals of β -Gal II' are characteristic of a terminal position. Resonance positions of the signals for GalNAc-ol H-3, H-6 and H-6' at δ 4.102, 3.977 and 3.793 p.p.m. respectively showed a 3,6-substitution. The ROESY spectrum showed contacts between (i) GlcA III H-1 and Gal II H-3 (strong), (ii) Gal II H-1 and GalNAc-ol H-3 and (iii) Gal II' H-1 and GalNAc-ol H6, H6'. Therefore, compound 3 has the following structure:



From the MALDI-TOF analysis, fraction 6 was found to be composed of two compounds, 6A and 6B, containing respectively four hexose (Hex), one hexuronic acid (HexA) and one Nacetylhexosaminitol (HexNAc-ol; $[M+Na]^+$ at m/z 1040) and four Hex and one HexNAc-ol ($[M + Na]^+$ at m/z 894). Based on the intensity of the H-1 signals and the ${}^{3}J_{H,H}$ coupling constants, the three sugar-spin systems of β -Glc, β -Gal and β -GlcA were identified in the ratio 6:2:1. Moreover, two types of H-3, H-4 and H-6,6' signal related to the GalNAc-ol unit were clearly observed in the COSY90 spectrum. A ROESY experiment (Figure 4) showed inter-residue cross-peaks between Gal II, H-1 and GalNAc-ol I_A H-3 and Gal II_B H-1 and GalNAc-ol I_B H-3. Therefore, it is concluded that oligosaccharide-alditols 6A and 6B have the following compositions: three β -Glc, one β -Gal, one β -GlcA and one GalNAc-ol for 6A and three β -Glc, one β -Gal and one GalNAc-ol for 6B. From the NMR spectra, 3,6and 3,4,6-substituted GalNAc-ol units were characterized. The TOCSY spectrum included signals at 8 4.077, 8 3.556, 8 3.89 and & 3.78 p.p.m., easily assignable to the H-3, H-4, H-6 and H-6' resonances of the 3,6-disubstituted GalNAc-ol I_A unit, whereas the signals at \$ 4.194, \$ 3.784, \$ 4.05 and \$ 3.86 p.p.m. could therefore be attributed to the H-3, H-4, H-6 and H-6' resonances of the 3,4,6-trisubstituted GalNAc-ol I_B unit. The nature of the substitution of the Gal II, and II, units was deduced from

Table 1 ¹H and ¹³C NMR chemical shifts of the mucin-type O-glycans (compounds 2, 3, 4, 5, 6A, 6B and 7)

			Chemical shift of ¹ H or ¹³ C (p.p.m.)									
Compound	Residue	H/C	1	1′	2	3	4	5	6	6′	NAc	0-CH3
2	GalNAc-ol 1	н	3.802	3.747	4.399	4.064	3.515	4.201	3.67	3.67	2.052	-
	Gal(β1-3) II	н	4.558	-	3.732	3.828	4.208	3.97	n.d.	n.d.	-	-
	GicA(<i>β</i> 1-3) III	н	4.674	-	3.421	3.52	3.52	3.723	-	-	-	-
	Gic(β1-6) IV	н	4.493	-	3.31	3.505	3.38	3.46	3.918	3.718	-	-
3	GaiNAc-ol I	н	≈ 3.8	3.75	4.413	4.102	3.574	4.38	3.977	3.793	2.046	-
	Gal(β1-3) II	н	4.537	-	3.726	3.828	4.186	n.d.	n.d.	n.d.	-	-
	GICA(β 1-3) III	Н	4.671	-	3.425	3.52	3.52	3.723	-	-	-	-
	Gal(<i>β</i> 1-6) II'	н	4.439	-	3.543	3.659	3.922	n.d.	n.d.	n.d.	-	-
4	GalNAc-ol I	н	≈ 3.8	3.742	4.401	4.073	3.55	4.37	3.986	3.788	2.045	-
		С	60.8	-	51.73	76.97	69.24	68.42	71.26	-	22.23	-
	Gal(β 1-3) II	н	4.564	-	3.734	3.834	4.208	3.97	4.073	3.987	-	-
		C	104.21	-	70.34	82.45	68.17	73.93	70.01	-	-	-
	GICA(β 1-3) III	н	4.669	-	3.42	3.52	3.52	3.721	-	-	-	-
		C	104.17	-	73.57	75.7	72.12	76.53	n.d.	-	-	-
	Gic(β1-6) IV	Н	4.49	-	3.312	3.505	3.383	3.462	3.92	3.718	-	-
		C	103.08	-	73.55	75.7	≈70	76.23	61.08	-	-	-
	Glc(β1-6) ΙΙ'	н	4.494	-	3.312	3.505	3.383	3.462	3.92	3.718	-	-
		C	103.08	-	73.55	75.7	≈ 70	76.23	61.08	-	-	-
5	GicNAc-ol 1	н	3.74	3.65	4.191	3.89	3.797	3.83	n.d.	n.d.	2.060	-
		С	62.27	-	53.66	69.82	81.58	74.3	≈ 62.3	-	23.7	-
	GalNAc(<i>β</i> 1-4) II	н	4.501	-	3.964	3.94	4.075	3.87	3.96	3.82	2.081	-
		C	103.61	-	52.63	77.45	70.08	74.79	69.37	-	23.7	-
	Gal(β1-3) III	н	4.676	-	3.761	3.966	4.282	3.813	4.058	3.93	-	-
		C	103.01	~	75.60	84.36	69.37	74.3	70.31	-	-	-
	Gal(β1-3) IV	н	4.606	-	3.63	3.63	3.918	3.690	n.d.	n.d,		-
		С	105.30	-	72.25	73.9	69.95	76.29	≈ 62.3	-	-	-
	Gal(<i>j</i> 31-6) V	н	4.542	-	3.612	3.820	.920	3.866	4.088	3.87	-	-
		С	103.26	-	77.17	72.8	69.95	74.79	70.87	-	-	-
	Gal(β1-6) VI	H	4.454	-	3.553	3.665	3.924	3.705	n.d.	n.d.	-	-
		С	104.56	-	72.01	73.9	69.95	76.29	≈ 62.3	-	-	-
	Gic(Н	4.538	-	3.310	3.508	3.389	3.471	3.951	3.724	-	-
		C	103.64	-	74.37	77.02	70.76	77.25	61.98	-	-	-
	Fuc(a1-2) VIII	н	5.325		3.815	3.617	3.695	4.185	1.207	-	-	-
		С	100.17	-	69.2	70.45	72.98	67.92	16.9	-		-
	2-0-Me-Fuc(a1-2) IX	н	5.443	-	3.502	3.879	3.837	4.284	1.223	-	-	3.511
		С	97.59	-	78.66	70.15	72.84	67.56	16.9	-	-	58.59
6A	GalNAc-ol IA	н	3.82	3.82	4.415	4.077	3.556	4.38	3.89	3.78	2.047	-
		С	61.61	-	52.49	78	69.95	69.3	71.86	_	23.08	
	Gal(<i>β</i> 1-3) IIA	н	4.578	-	3.843	3.93	4.415	3.989	≈ 4.1	4.01	-	-
	•	С	105.03	-	71.28	82.59	76.8	74.44	71.5	-	-	-
	GicA(<i>B</i> 1-3) 111A	н	4.673	-	3.419	3.51	3.51	≈ 3.7	-	-	-	-
	., ,	C	105.17	-	74.32	76.58	73.11	77.57	n.d.	-	-	-
	GIC(B1-4) IVA	н	4.832	-	3.306	3.535	≈ 3.4	≈ 3.4	3.93	3.72	-	-
	•	C	103.56	-	74.3	76.7	70.7	76.57	61.8	-	-	-
	GIc(61-6) VA	н	4.476	-	3.28	≈ 3.5	≈ 3.4	3.45	3.93	3.72	-	-
	v	С	103.95	-	74.3	76.7	70.7	76.9	61.8	-	-	-
	Gic(\$1-6) II'A	н	4.501	-	3.33	≈ 3.5	≈ 3.4	3.45	3.93	3.72	-	-
		С	103.73	-	74.3	76.7	70.7	76.9	61.8	-	-	-
6B	GalNAc-ol IB	н	3.82	3.82	4.45	4.194	3.784	4.38	4.05	3.86	2.067	-
		С	61.61	-	59.93	76.6	78.55	69.3	73.57	-	23.08	-
	Gal(B1-3) IIB	н	4.552	-	3.603	3.696	3.93	3.973	4.077	3.89	-	-
		C	105.5	-	71.85	73.16	69.55	74.85	70.51	-	-	-
	Glc(61-6) VB	Ĥ	4.501	-	3.33	≈ 3.5	≈ 3.4	3.45	3.93	3.72	-	-
		C	103.73	_	74.3	76.7	70.7	76.9	61.8	-	-	
	GIC(B1-4) VIB	Ĥ	4.442	-	3.33	n.d.	n.d.	n.d.	3.93	3.72	-	-
		C	103.44	-	74.3	n.d.	n.d.	n.d.	61.8	-	-	-
	Gic(81-6) II'B	Ĥ	4 479	_	3.28	≈ 3.5	≈ 3.4	3.45	3.93	3.72	-	-
		'n	103 95	_	3 33	76 7	nd	76 9	61 8	-	-	-
7	GalNAc-ol I	н	≈ 3 77	≈ 3.77	4.302	4 078	3 509	4 194	≈ 3.64	≈ 3.64	2.046	_
r	Gal/ R1.3) II	11	~ 0.11 A 520	~	3 700	2 812	A 149	nd	nd	n d	_	_
	CloA (/21 2) 11	л Ц	4.523	-	3.100	610.C 403 C	9.142	n.u. n.d	11. U .	n.u.	_	_
		п	4.000	-	J.440	3.004	3.120	n.u.				-
					3 700	0 10 4	n 4	n 4		r 4	2 1146	

the HMQC spectrum (Figure 4), which allowed us to distinguish the 3,4,6-trisubstituted β -Gal II_A unit (C-3, C-4, C-6 observed at δ 82.59, δ 76.8 and δ 71.5 p.p.m. respectively) from the 6-

mono-substituted β -Gal II_B unit (C-3, C-4, C-6 observed at δ 73.16, δ 69.55 and δ 70.51 p.p.m. respectively). Similarly, the HMQC spectrum also confirmed the C-3, C-4 or C-6 substitu-

Compound	Residue	H/C	Chemical shift of ¹ H or ¹³ C (p.p.m.)									
			1	1′	2	3	4	5	6	6′	NAc	
1	GaiNAc-ol I	н	3.77	3.696	4.378	4.039	3.475	4.095	3.618	3.588	2.03	
	GicA(<i>β</i> 1-3) II	н	4.509	-	3.415	3.627	3.75	3.72	-	-	-	
	GalNAc(B1-3) III	н	4.548	-	3.872	3.696	3.912	n.d.	n.d.	n.d.	2.038	
8	Xyl-ol I	н	n.d.	-	n.d.	n.đ.	3.986	n.d.	-	-	-	
	Gal(61-4) II	н	4.615	-	3.724	3.845	4.194	n.d.	n.d.	n.d.	-	
	Gal((31-3) III	н	4.669	-	3.749	3.805	4.154	n.d.	n.d.	n.d.	-	
	GICA(\$1-3) IV	н	4.669	-	3.456	3.623	3.75	3.72	-	-	-	
	GalNAc(B1-4) V	н	4.474	-	≈ 3.8	≈ 3.7	3.921	n.d.	n.d,	n.d.	2.060	
12A	GalNAc-ol IA	н	3.764	≈ 3.7	4.377	4.036	3.474	4.095	3.65	3.59	2.036	
GicA(,		С	62.52	-	52.86	77.06	70.31	70.55	65.31	-	23.17	
	GicA(B1-3) IIA	н	4.55	-	3.417	3.631	3.739	3.68	_	-	-	
		С	104.07	-	74.45	75.12	80.88	77.5	n.d.	-	-	
	GaINAc(B1-4) IIIA	н	4.548	-	4.014	3.803	4.169	3.69	3.76	3.74	2.019	
		С	102.12	-	52.71	81.79	69.08	76.4	62.52	-	23.83	
	GicA(\$1-3) IVA	н	4.486	-	3.309	3.47	3.47	3.679	-	-	-	
		C	105.58	-	73.99	76.71	73.29	77.5	n.d.	-	-	
12B	GalNAc-ol 1B	н	3.764	≈ 3.7	4.377	4.036	3.474	4.095	3.65	3.59	2.036	
		С	62.52	-	52.86	77.06	70.31	70.55	65.31	_	23.17	
	GIcA(B1-3) IIB	н	4.55	-	3.417	3.631	3.739	3.68	-	-	-	
	•	С	104.07	-	74.45	75.12	80.88	77.5	n.d.	-	-	
	GaINAc(81-4) IIIB	Н	4.543	-	3.987	3.793	4.115	3.69	3.76	3.74	2.012	
		C	102.12	-	52.76	81.79	69.12	76.4	65.52	-	23.83	
	GICA(B1-3) IVB	н	4.483	-	3.347	3.566	3.739	3.68	-	-	-	
	· · · ·	C	105.58	-	73.99	75.12	80.88	77.5	n.d.	-	-	
	GaINAc(B1-4) VB	Ĥ	4.461	-	3.876	3.688	3.907	3.7	3.76	3.74	2.041	
		C	102.25	-	53.7	76.4	69.12	76.4	62.52	-	23.82	

Table 3 ¹H chemical shifts of anomer protons of oligomannosyl N-glycans (compounds 9, 10A, 10B and 11)

	Chemical shift (p.p.m.)								
	11	10A	108	9					
	D3 B M~M_4' M-M_M D2 A M-GN-GN-o! M-M M 3 2 D1 C 4	M-M M M-GN-GN M-M-M	M-M IM-M ^{/M} /M-GN-GN-oł M-M	n - M M ⁻ M - GN - GN-oi M - M					
GICNAC 2	4.628	4.619	4.619	4.629					
Man 3	≈ 4.8	4.811	4.811	4.816					
Man 4	5.336	5.339	5.331	5.34					
Man 4'	4.686	4.866	4.866	4.872					
Man A	5.406	5.083	5.402	5.086					
Man B	5.145	5.143	5.143	5.054					
Man C	5.309	5.305	5.057	5.042					
Man D1	5.040-5.053	-	5.04	-					
Man D2	5.040-5.053	≈ 5.04	_	-					
Man D3	5.040-5.053	≈ 5.04	≈ 5.04	5.042					

tions of GalNAc-ol I_A or I_B . The assignment of the signals belonging to the six β -Glc units was performed on the ROESY spectrum by the observation of inter-residue cross-peaks between their respective, but not assigned, H-1 signals, and the perfectly assigned H-3, H-4, H-6 and H-6' resonances of Gal and GalNAc-ol units (Figure 4). Methylation analyses confirmed the presence of two different GalNAc-ol residues in this sample. As previously observed [25], the methanolysis of methylated 3,6-disubstituted GalNAc-ol leads to the formation of 1,4,5-Me₃-3,6-anhydro-GalNAc(Me)-ol (Figure 5). Similarly, the 3,4,6-trisubstituted GalNAc-ol formed an 1,5-Me₂-4-Ac-3,6anhydro-GalNAc(Me)-ol. Both residues showed the ions at m/z 130 and 88, but were distinguished by the ions at m/z 216, 174 and 142 for the tri-methylated residue or the ions at m/z 244, 202 and 170 for the di-methylated residue. However, the same analysis only showed the presence of a single 6-O-substituted Gal residue belonging to compound 6B. The absence of the trisubstituted Gal residue was attributed to the stabilization effect of the GlcA residue on the glycosidic bond in the dimer GlcA(β 1-3)Gal. A reduction of the carboxy group of the GlcA residue prior to the per-methylation and methanolysis permitted to liberate the trisubstituted Gal residue. Under these experimental conditions we could observe both mono- and trisubstituted Gal residues from both compounds (results not shown), as predicted by the

173



Figure 2 Extension of ¹H 400 MHz double-relayed COSY spectrum (left) of oligosaccharide 4 ($3.25-4.75 \times 3.25-4.75$ p.p.m.) and of ¹H 400 MHz single-relayed COSY spectrum (right) of oligosaccharide 3 ($3.3-4.8 \times 3.3-4.8$ p.p.m.)

NMR analysis. On the basis of the data obtained, it is concluded that the compounds 6A and 6B have the following structures:



Glc(β 1-6) Glc(β 1-6) Glc(β 1-4) GalNAc-ol Gal(β 1-3)

On the basis of composition analysis and integration of ¹H NMR analysis, compound 7 is composed of one β -GlcNAc, one β -GlcA, one β -Gal and one GalNAc-ol residue. Chemical shifts of GalNAc-ol H-6 and H-6', both observed at δ 3.64 p.p.m., and H-3 at 4.078 p.p.m. are indicative of a mono-substitution in C-3. The Gal II residue shows identical parameters to compound 2 and 3. The H-4 of GlcA III is deshielded to 3.728 p.p.m. compared with 3.51–3.52 p.p.m. in compounds 2, 3, 4 and 6A, suggesting that this residue is substituted by the β -GlcNAc residue in a (1-4) linkage. Thus, we propose the following structure for compound 7:

GalNAc-ol Gal(β1-3) GlcNAc(β1-4)GlcA(β1-3) Compound 5 presented features that were distinct from the six oligosaccharide-alditols analysed above. It showed pseudo-molecular ions $[M + Na]^+$ and $[M + K]^+$ at m/z 1566 and m/z 1582 on a MALDI-TOF analyser. In accordance with these data and composition analysis it was concluded that this compound constituted one GlcNAc-ol, one GalNAc, one Glc, three Gal, one Fuc and one O-Me-Fuc residue. The GalNAc residue was unambiguously distinguished from the GlcNAc-ol residue in GC/MS as an alditol-acetate derivative after a reduction with sodium borodeuterite. Indeed, in electronic-impact mode GlcNAc-ol residue showed two ions at m/z 144 and 145, attributed to CH2OCOCH3-CHNH+COCH3 and CH2OCOCH3-CHO+COCH₃ fragments, respectively. In contrast, deuterated GalNAc-ol residue showed both fragments C²HHOCOCH₃-CHNH⁺COCH₃ and CH₂OCOCH₃-CHO⁺COCH₃ at m/z 145. These data indicated that the non-deuterated GlcNAc-ol residue occurred in the reducing terminal position compared with the internal position of the GalNAc residue. Fuc and O-Me-Fuc residues were discriminated as alditol-acetate derivatives in chemical-ionization mode. Fuc was characterized by a $[M + NH_{4}]^{+}$ at m/z 395 and O-Me-Fuc by a $[M + NH_{4}]^{+}$ at m/z367. The position of the methyl group in the O-Me-Fuc residue was obtained in GC/MS by analysis of its fragmentation pattern in electronic-impact mode. It was characterized by an intense ion at m/z 118 corresponding to the mono-methylated, C²HHOAc- CH_2O^+Me fragment, as well as an ion at m/z 275 corresponding to the mono-methylated CH₂OAc-(CH₂OAc)₃-CH₂O⁺Me fragment. From these data, we concluded that it was substituted by





The ROESY spectrum was performed with a mixing time of 400 ms and with mild presaturation during relaxation time to suppress the residual water signal.

a methyl group in the C-2 position. These results were confirmed by ¹H NMR experiments.

The methylation results indicated that compound 5 contained terminal Gal and Glc residues, a 4-linked GlcNAc-ol, a 2,6-linked Gal and a 2,3,6-linked Gal. Fuc and 2-O-Me-Fuc residues were both observed as permethyl-Fuc derivatives.

Fragmentation of the ion $[M + Na]^+$ at m/z 1566 in ESI-MS/MS showed a series of informative ions (Figure 6). First, we observed an intense ion at m/z 1342 corresponding to the C ion $[M-HexNAc-ol + Na]^+$. This confirmed that the GlcNAc-ol residue in the terminal reducing position was mono-substituted. This is one of the only two fragments containing a Fuc residue that we identified, suggesting an important instability of the fucoseresidue linkage during fragmentation. Ions at m/z 1419 and 1196 correspond to the fragments [M-Fuc+Na]⁺ and [M-Fuc-HexNAc-ol+Na]⁺. Ions at m/z 1034, 872, 712, 550, 388 and 226 all correspond to decomposition fragments of the ion at m/z1196, lacking both Fuc and HexNAc-ol residues. X ions at m/z 611 and 449 are characteristic of the sequences Hex-HexNAc-HexNAc-ol and HexNAc-HexNAc-ol respectively. These data, concomitantly with methylation analysis, prove the presence of the dimer GalNAc(α/β 1-4)GlcNAc-ol in the terminal reducing position. The C and B ions at m/z 995 and 977 both indicate a fragment containing four Hex, one Fuc and one O-Me-Fuc residue. Intense C and B ions at m/z 849 and 831 correspond to the previous fragment lacking its Fuc residue. It is



Figure 4 Extension (50–110 × 65–3.1 p.p.m.) of HMQC (A) and extension (4.8–3.2 × 4.85–4.35 p.p.m.) of ROESY (B) of a mixture of oligosaccharide-alditols 6A and 6B

The ROESY spectrum was performed with a mixing time of 300 ms.

noteworthy that we did not observe the equivalent fragment devoid of O-Me-Fuc, suggesting a greater stability of this residue compared with the Fuc residue. Ions at m/z 687, 669, 525, 507, 365 and 347 correspond to further fragmentation of ions 849 and 831. Taken together, these data suggest that the GalNAc linked to the GlcNAc-ol residue is disubstituted in C-3 and C-6 by a sequence containing four Hex, one Fuc and one 2-O-Me-Fuc residues on one branch and by the remaining Hex residue on the other.

Assignment of the ¹H and ¹³C resonances of the sugar residues and the alditol was made from COSY, two-step relayed COSY, HMQC and ROESY spectroscopy. The H-1 to H-6 resonances for β -Glcp and the H-1 to H-4 resonances for β -Galp and β -GalNAc residues were assigned by referring to the twostep relayed COSY spectrum. The H-5 resonances for β -Gal III, IV, V and VI were detected by ROESY at 3.813, 3.69, 3.866 and 3.705 p.p.m., respectively, suggesting that β -Gal units III and V are 6-linked. The relative low-field position of some ¹³C resonances (Figure 7), compared with their corresponding nonsubstituted monosaccharides, indicated the presence of Glc (VII), Gal (IV), Gal (VI), Fuc (VIII), 2-O-Me-Fuc (IX) in the terminal position of 2,6-linked Gal (V) and 2,3,6-linked Gal (III). The GalNAc(β 1-4)GlcNAc-ol linkage gave rise to a strong nuclear Overhauser effect (NOE) signal between H-1 of GalNAc and H-4 of GlcNAc-ol (Table 4). Other strong NOE connectivities were observed between H-1 of Gal (V) and Gal (VI) and signals at 4.058 and 4.088 p.p.m., which were assigned as the H-6 resonances of Gal (III) and Gal (V). The exact location of Fuc VIII on Gal II was unambiguously established according to the successive NOE contacts observed between Fuc VIII H-1 and Gal III H-2, and between Gal III H-1 and GalNAc II H-3 (Table 4). Then 2-O-Me-Fuc IX was located on the remaining C-2 of the 2,6-disubstituted Gal V, which was confirmed by the NOE contact between 2-O-Me-Fuc IX H-1 and Gal V H-2. The resonances which remained to be assigned [namely C-5 of GalNAc (II) and GlcNAc-ol (I) as well as C-6 of GlcNAc-ol (I), Gal (IV) and Gal (VI)] were detected on the HMQC spectrum by comparison with ¹³C NMR chemical shifts of monosaccharides



Figure 5 GC/MS profile of methylation analysis of compounds 6A and 6B

(a) From top to bottom, is shown chromatogram reconstitution using the search for ions specific for 1,5-Me₂-4-Ac-3,6-anhydro-GalNAc(Me)-ol residue (ion at m/z 244) and 3,4,6-trisubstituted GalNAc-ol residue (ion at m/z 216) and a total ion count (TIC) chromatogram. (b) Fragmentation pattern of 1,5-Me₂-4-Ac-3,6-anhydro-GalNAc(Me)-ol. (c) Fragmentation pattern of 1,4,5-Me₃-3,6-anhydro-GalNAc(Me)-ol.



Figure 6 Fragmentation pattern of the proposed structure of oligosaccharide-aiditol 5 in ESI-MS

(a) Fragmentation of the pseudo-molecular ion $[M + Na]^+$ at m/z 1565.6. Only the most informative fragments are shown. (b) ESI-MS/MS spectrum.



Figure 7 Extension (110–50 \times 5.7–3.1 p.p.m.) of Inverse-detected ¹H-¹³C 600 MHz HMQC spectrum of oligosaccharide 5

Each experiment is an accumulation of 48 scans, and recorded in the phase-sensitive mode using states-TPPI (time-proportional phase incrementation) method.

Table 4 Summary of inter-residue contacts observed in homonuclear ROESY experiment for the oligosaccharide-alditol, compound 5

Inter-residue NOE connectivity GalNAc II H-1/GicNAc-ol H-4 Gal III H-1/GalNAc II H-3 Gal IV H-1/Gal III H-3 Gal VI H-1/Gal III H-6 Gal VI H-1/GalNAc II H-6 Fuc VIII H-1/GalNAc II H-6 Fuc VIII H-1/GalNAc II H-6 Fuc VIII H-1/Gal III H-2 2-0-Me-Fuc H-1/Gal V H-2



GAG-type oligosaccharide analysis

The second series of compounds (compounds 1, 8, 12A and 12B) is characterized by the presence of the dimer GalNAc(β 1-4)GlcA. From the MALDI-TOF analysis and the sugar ratio determination by routine GC/MS analysis, compound 8 was found to be composed of one GalNAc, one GlcA, two Gal and one Xylol residue. The ¹H NMR spectra led us to propose that this compound presented the following sequence:

[24]. Thus, on the basis of the above results, the structure of the compound 5 oligosaccharide may be represented as follows:



Figure 8 Extension (3.3-4.8 × 3.3-4.8 p.p.m.) of ¹H 400 MHz single-relayed COSY spectrum of oligosaccharide-alditol 8

, Gal(β1-4) Xy1-ol , Gal(β1-3) GalNAc(β1-4) GlcA(β1-3)

Indeed, the partial sequence GlcA(β 1-3)Gal(β 1-3)Gal(β 1-4)Xylol was easily identified on the basis of the ¹H NMR parameters, owing to a compilation of previous data [26]. The presence of the additional terminal GalNAc(β 1-4) unit was established on the basis of the ³J vicinal coupling constants of the GalNAc unit and the deshielding of GlcA H-3 and H-4 atom resonances to δ 3.623 and δ 3.75 p.p.m. respectively, as shown on the COSY spectrum (Figure 8).

On the basis of composition and MALDI-TOF analyses,

compounds 1, 12A and 12B contain one GalNAc-ol residue, together with GlcA and GalNAc residues. Compound 1 showed a $[M + Na]^+$ at m/z 623 and a $[M + K]^+$ at m/z 639. Compounds 12A and 12B were analysed as a mixture and showed $[M + Na]^+$ at m/z 799 and 1002, and $[M + K]^+$ at m/z at 815 and 1018. Thus, these compounds were respectively characterized as tri-, tetraand pentasaccharide-alditols. In each compound, the H-3 (δ 4.036–4.039 p.p.m.) and H-6,6' (δ 3.618–3.65 p.p.m. and δ 3.588– 3.59 p.p.m.) atom resonances of the GalNAc-ol unit indicated a mono 3-O-substitution of this residue (Figure 9). The attachment of a GlcA residue to the GalNAc-ol unit in compounds 12A and 12B was proved by a ROESY experiment (Figure 9). Starting from the H-4 atom resonance of GlcA IV_A and IV_B of compounds



Figure 9 Extension (3.15-4.8 × 3.15-4.8 p.p.m.) of ¹H 400 MHz double-relayed COSY spectrum (left) and extended ROESY (right; same scale as COSY) of mixture 12

The ROESY spectrum was performed with a mixing time of 400 ms and with mild presaturation during relaxation time to suppress the residual water signal.

12A and 12B, it was possible to assign the H-1 resonance of these two units. In a second step, a strong NOE effect was observed between these anomeric protons and both GalNAc III, and III, H-3, at & 3.803 p.p.m. and 3.793 p.p.m. respectively. In addition, the ROESY spectrum led us to distinguish and to assign the H-4 resonances of the GalNAc III_A and GalNAc III_B units, and consequently to determine the precise chemical shifts of their anomeric protons. Owing to a HMQC experiment, the 4-Osubstitution of GlcA II_A , II_B and IV_B was proved by the deshielding of their C-4 atom resonances to & 80.88 p.p.m., compared with terminal GlcA IV_{A} residue C-4 atom resonance at δ 73.29 p.p.m. Similarly, the C-3 atom resonances of GalNAc III_{A} and III_{B} were deshielded to δ 81.79 p.p.m. compared with the C-3 atom resonances of terminal GalNAc $V_{\rm B}$ at δ 76.4 p.p.m. From these observations, the sequences of compounds 12A and 12B were established as follows:



As shown by the COSY90 spectrum, the terminal non-reducing GalNAc III unit of compound 1 possesses identical H-2, H-3 and H-4 atom resonances to the GalNAc V_B unit of compound 12B. Similarly, the GlcA II protons show identical parameters to GlcA II_A and II_B of compounds 12A and 12B. Thus, we propose the following sequence for compound 1:

N-glycan analysis

MALDI-TOF and composition analyses indicated that fractions 9, 10 and 11 contained Man, GlcNAc and GlcNAc-ol in the proportions 7:1:1, 8:1:1 and 9:1:1. From one-dimensional ¹H NMR analysis, the ¹H NMR parameters matched perfectly those of oligomannoside-type N-glycans [27]. In particular, fraction 9 contained a single Man-7 N-glycan, fraction 10 consisted of a mixture of two Man-8 N-glycans (compounds 10A and 10B) and fraction 11 contained a Man-9 N-glycan. Thus the structures of these compounds were established as follows:

© 2001 Biochemical Society

 $Man(\alpha 1-2)Man(\alpha 1-6)$

 $Man(\alpha 1-3)$ $Man(\alpha 1-3)$ $Man(\alpha 1-2)Man(\alpha 1-3)$ $Man(\alpha 1-2)Man(\alpha 1-3)$

 $Man(\alpha 1-2)Man(\alpha 1-6)$

 $Man(\alpha 1-6)$ $Man(\alpha 1-3)'$ $Man(\beta 1-4)GlcNAc(\beta 1-4)GlcNAc-ol$ $Man(\alpha 1-2)Man(\alpha 1-2)'$

 $\begin{array}{l} Man(\alpha 1-2)Man(\alpha 1-6)\\ Man(\alpha 1-2)Man(\alpha 1-3)\\ Man(\alpha 1-2)Man(\alpha 1-3)\\ Man(\alpha 1-2)Man(\alpha 1-3)\\ \end{array}$

 $\begin{array}{l} \label{eq:man(al-2)Man(al-6)} \\ \mbox{Man(al-2)Man(al-3)} \\ \mbox{Man(al-2)Man(al-2)Man(al-2)} \\ \mbox{Man(al-2)Man(al-2)Man(al-3)} \\ \end{array}$

DISCUSSION

In this study, we have fully sequenced seven major highly unusual mucin-type O-glycans. Six of them were part of a homologous family, synthesized in abundance by C. elegans. They presented either type-1 core $Gal(\beta 1-3)GalNAc$ or the newly described core type $Gal(\beta 1-6)[Gal(\beta 1-3)]GalNAc$. Five out of six were substituted with a GlcA unit and showed the trisaccharidic sequence $GlcA(\beta 1-3)Gal(\beta 1-3)GalNAc$. Most of these structures had $(\beta 1-6)$ Glc and/or $(\beta 1-4)$ Glc residues on the GalNAc and/or Gal residues. To our knowledge, these are the only glucose-containing mucin-type O-glycans ever described. Similarly, compound 6B represents the only example of a trisubstituted GalNAc-ol residue ever described. All these unusual features contribute to the uniqueness of the glycosylation of C. elegans. Indeed, this few structures suggest the expression in C. elegans of five novel glycosyltransferase activities: four β -glucosyltransferase activities and a N-acetylgalactosamine β -1,6-galactosyltransferase activity.

The last mucin-type O-glycan we described (compound 5) presented a GlcNAc-ol residue in the terminal position, which suggests that it would be linked to the protein moieties through a GlcNAc and not a GalNAc residue. O-glycosidically linked GlcNAc-bound oligosaccharides have already been described in glycoproteins from Trypanosoma cruzi [28] and act as sialic acid acceptors for the α -2,3-trans-sialidase [29,30]. They present a simpler sequence than the one described in C. elegans, and are characterized by the oligomerization of β -Galf and β -Galp residues on the GlcNAc unit. Compound 5 was also characterized by the simultaneous presence of Fuc(β 1-2) and 2-O-Me-Fuc(β 1-2) residues, as well as a Glc(β 1-6) residue. 2-O-Me-Fuc residues were identified previously in the major mucin-type O-glycans substituting the TES protein of the parasitic nematodes T. canis and T. cati [5]. However, the overall, complex structure of compound 5 shows no equivalence with any other known Oglycan.

Considering the large amounts of these O-glycans that we have purified (up to 500 μ g each/10 g of dry worm), it is likely that

C. elegans overexpresses highly O-glycosylated glycoproteins. Previous studies on the parasitic nematode T. canis showed that it overexpresses and secretes a surface-associated highly O-glycosylated mucin-type glycoprotein, TES 120 [1,31]. A gene encoding a mucin-type glycoprotein, let-653 [32], as well as typical mucin-type cysteine-rich domains [33], have also been characterized in C. elegans, suggesting the occurrence of such glycoproteins. Moreover, preliminary results from the fractionation of the total components of C. elegans through isopycnic ultracentrifugation (results not shown) gave evidence for the presence of such components. They were characterized by a high density ($\rho \ge 1.4 \text{ g} \cdot \text{ml}^{-1}$), an electrophoretic mobility in agarose gel similar to purified mucins and a susceptibility to Toluidine Blue and periodic acid-Schiff reagents. Previous work showed that mutation of the let-653 gene in C. elegans resulted in larval death due to the dysfunction of the secretary/excretory apparatus [32], confirming the importance of such mucin-type glycoproteins in C. elegans.

The fact that we have released four major N-linked glycans among O-glycans by mild alkali treatment is not surprising. Indeed, if it is accepted that the N-glycosidic linkages are more stable under mild alkali conditions than are the O-glycosidic linkages, studies have pointed out that such treatments lead also to a slow degradation of GlcNAc-asparagine linkages [34] and to the release of a low percentage of total intact N-glycans.

We have also described fragments of an unsulphated chondroitin-like GAG by NMR. This component, as well as a minute amount of heparan sulphate-type GAG, have been studied previously by enzymic degradation [8,9]. We did not isolate any heparan sulphate-type GAG, probably due to the fact that the methodology we used only permits access to major components. However, the present study permitted us to isolate the tetrasaccharide GAG core substituted by a single GalNAc residue (compound 8), which was not achieved in previous work. This directly confirmed the presence of the ubiquitous GAG protein moiety linkage sequence $GlcA(\beta 1-3)Gal(\beta 1-3)G$ 4)Xyl in C. elegans. It is noteworthy that this compound does not seem to be the result of a chemical cleavage releasing the GAG chain from the core sequence. Indeed, to our knowledge, mild alkali degradation under reductive conditions cannot cleave the glycosidic bond in the GlcA(β 1-3)GalNAc dimer by β -elimination reaction. Similarly, stability studies of chondroitin sulphate demonstrated that no desulphation of this compound occurred in 0.1 M NaOH at 30 °C over a period of 1000 h [35]. This suggests that the unsulphated pentasaccharide we have isolated is definitely not a degradation product, and is abundantly synthesized by C. elegans, either as a precursor for longer GAG chains or as a final product.

The exact origin of compounds 1, 12A and 12B is still unclear. A previous study has demonstrated the possibility of a very slow decomposition process of chondroitin sulphate chains under non-reducing mild alkali conditions (0.1 M NaOH, at 30 and 60 °C) through a β -elimination reaction, forming a 4-deoxy- α -Lthreohex-4-enepyranosyluronic acid (Δ HexA) at the non-reducing end of each newly formed chain [35]. To assess this possibility, we subjected a chondroitin sulphate to the reductive alkaline conditions that we used on C. elegans extract but failed to detect any degradation product or Δ HexA-containing compound (results not shown). Furthermore, it is noteworthy that we did not detect any AHexA-substituted chondroitin-like oligosaccharide in C. elegans. Therefore, compounds 1, 12A and 12B were the result of either the selective release of oligosaccharide fragments exclusively from the non-reducing end of an unsulphated chondroitin-like GAG, or the release by β -elimination directly from the peptide backbone of a glycoprotein. This last hypothesis does not exclude the existence of classical nonsulphated chondroitin chains.

Previous studies isolated mutations that perturbed the worm vulval invagination without affecting vulval cell lineage [36]. Out of the eight genes involved in this defect sqv-3 and sqv-8 respectively encode a galactosyltransferase and a glucuronyltransferase directly involved in the biosynthesis of the tetrasaccharide core GlcA(β 1-3)Gal(β 1-3)Gal(β 1-4)Xyl [37-39]. Sqv-8 could also transfer GlcA to the Gal(β 1-3)GalNAc- α -Obenzyl sequence, although this activity was 10-fold lower than that observed on a Gal(β 1-3)Gal β -O-naphthalenemethanol. Considering the quantitative importance of the $GlcA(\beta)$ -3)Gal(β 1-3)GalNAc sequence in C. elegans, we believe that it is necessary to determine whether Sqv-8 is also involved in the biosynthesis of these mucin-type O-glycans. If it is involved, then inactivating mutations of sqv-8 would lead to the synthesis of truncated homologues of the mucin-type O-glycans we have observed in the present study. In this last case, modifications in the synthesis of these compounds might also be involved in the abnormal phenotypes observed in sqv mutants.

By defining major oligosaccharide structures expressed by *C. elegans*, our studies reported here should facilitate research concerning the specificity of glycosyltransferases in this organism and the role of glycosylation during development in multicellular organisms. Only knowledge of the *C. elegans* oligosaccharides will permit us to design proper substrates to assess the endogenous specificity of each glycosyltransferase cloned, especially in the case of unusual activities such as the ones observed in *C. elegans*. Similarly, comparison of stage-specific O-glycosylation patterns will be useful. Then, compilation of detailed data obtained from *C. elegans* development and specific-stage glycosyltransferase activities will offer the unique possibility in a multicellular organism to follow the synthesis of glycans throughout its entire embryogenesis.

We are deeply indebted to the late Professor André Verbert for his continuous support and encouragement. We thank Dr J. M. Wieruszeski, P. Timmerman and G. Ricard for valuable technical assistance. This research was supported by the Centre National de la Recherche scientifique (UMR 8576, Glycobiologie structurale et fonctionelle) and by the Ministère de l'Enseignement Supérieur et de la Recherche.

REFERENCES

- Maizels, R. M., Kennedy, M. W., Meghji, M., Robertson, B. D. and Smith, H. V. (1987) Shared carbohydrate epitopes on distinct surface and secreted antigens of the parasitic nematode *Toxocara canis*. J. Immunol. **139**, 207–214
- 2 Haslam, S. M., Houston, K. M., Harnett, W., Reason, A. J., Morris, H. R. and Dell, A. (1999) Structural studies of N-glycans of filarial parasites. Conservation of phosphorylcholine-substituted glycans among species and discovery of novel chito-oligomers. J. Biol. Chem. 274, 20953–20960
- 3 Reason, A. J., Ellis, L. A., Appleton, J. A., Wisnewski, N., Grieve, R. B., McNeil, M., Wassom, D. L., Morris, H. R. and Dell, A. (1994) Novel tyvelose-containing tri- and tetra-antennary N-glycans in the immunodominant antigens of the intracellular parasite Trichinella spiralis. Glycobiology 4, 593–603
- 4 Haslam, S. M., Coles, G. C., Munn, E. A., Smith, S., Smith, H. F., Morris, H. R. and Dell, A. (1996) *Haemonchus contortus* glycoproteins contain N-linked oligosaccharides with novel highly fucosylated core structures. J. Biol. Chem. 271, 30561–30570
- 5 Khoo, K. H., Maizels, R. M., Page, A. P., Taylor, G. W., Rendell, N. B. and Dell, A. (1991) Characterization of nematode glycoproteins: the major O-glycans of *Toxocara* excretory-secretory antigens are O-methylated trisaccharides. Glycobiology **12**, 163–171
- 6 Gerdt, S., Lochnit, G., Dennis, R. D. and Geyer, R. (1997) Isolation and structural analysis of three neutral glycosphingolipids from a mixed population of *Caenorhabditis elegans* (Nernatoda: Rhabditida). Glycobiology **7**, 265–275

- 7 Gerdt, S., Dennis, R. D., Borgonie, G., Schnabel, R. and Geyer, R. (1999) Isolation, characterization and immunolocalization of phosphoryicholine-substituted glycolipids in developmental stages of *Caenorhabditis elegans*. Eur. J. Biochem. **266**, 952–963
- 8 Yamada, S., Van Die, I., Van den Eijnden, D. H., Yokota, A., Kitagawa, H. and Sugahara, K. (1999) Demonstration of glycosaminoglycans in *Caenorhabditis elegans*. FEBS Lett. **459**, 327–331
- 9 Toyoda, H., Kinoshita-Toyoda, A. and Selleck, S. B. (2000) Structural analysis of glycosaminoglycans in *Drosophila* and *Caenorhabditis elegans* and demonstration that tout-velu, a *Drosophila* gene related to EXT turnor suppressors, affects heparan sulfate in vivo. J. Biol. Chem. 275, 2269–2275
- 10 Chen, S., Zhou, S., Sarkar, M., Spence, A. M. and Schachter, H. (1999) Expression of three *Caenorhabditis elegans* N-acetylglucosaminyltransferase I genes during development. J. Biol. Chem. 274, 288–297
- Chen, S., Spence, A. M. and Schachter, H. (2001) Phenotypes of *Caenorhabditis* elegans UDP-GicNAc:α-3-D-mannoside β1,2-N-acetylglucosaminyltransferase I (GnT I) null mutants. Glycoconj. J. 17, 71
- 12 Hagen, F. K. and Nehrke, K. (1998) cDNA cloning and expression of a family of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase sequence homologs from *Caenorhabditis elegans*. J. Biol. Chem. **273**, 8268–8277
- 13 Gentile, K., Layden, M., Berbach, K., Scwartz, A. and Hagen, F. (2001) O-linked glycoproteins and renal tubule development in *C. elegans.* Glycoconj. J. 17, 33
- 14 Oriol, R., Mollicone, R., Cailleau, A., Balanzino, L. and Breton, C. (1999) Divergent evolution of fucosyltransferase genes from vertebrates, invertebrates, and bacteria. Glycobiology 9, 323–334
- 15 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
- 16 Brenner, S. (1974) The genetics of Caenorhabditis elegans. Genetics 77, 71-94
- 17 Sulston, J. E. and Hodgkin, J. (1988) Methods. In The Nematode Caenorhabditis elegans (Wood, W. B., ed.), pp. 587–606, Cold Spring Harbor Press, Cold Spring Harbor
- 18 Epstein, H. F., Waterston, R. H. and Brenner, S. (1974) A mutant affecting the heavy chain of myosin in *Caenorhabditis elegans*. J. Mol. Biol. 90, 291–300
- 19 Sulston, J. E. and Brenner, S. (1988) The DNA of *Caenorhabditis elegans*. Genetics 77, 95–104
- 20 Lewis, J. A. and Fleming, J. T. (1995) Basic culture methods. In *Caenorhabdilis elegans*: Modern Biological Analysis of an Organism, Methods in Cell Biology, vol. 48 (Epstein, H. F. and Shakes, D. C., eds.), pp. 3–29, Academic Press, Orlando
- 21 Kamerling, J. P., Gerwig, G. J., Vliegenthart, J. F. and Clamp, J. R. (1975) Characterization by gas-liquid chromatography-mass spectrometry and protonmagnetic-resonance spectroscopy of pertrimethylsilyl methyl glycosides obtained in the methanolysis of glycoproteins and glycopeptides. Biochem. J. 151, 491–495
- 22 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
- 23 Ciucanu, I. and Kerek, F. (1984) A simple and rapid method for the permethylation of carbohydrates. Carbohydr. Res. 131, 209-217
- 24 Bock, K. and Pedersen, C. (1983) Carbon-13 nuclear magnetic resonance spectroscopy of monosaccharides. Adv. Carbohydr. Chem. Biochem. 41, 27–66
- 25 Wieruszeski, J. M., Michalski, J. C., Montreuil, J., Strecker, G., Peter-Katalinic, J., Egge, H., van Halbeek, H., Mutsaers, J. H. and Vliegenthart, J. F. (1987) Structure of the monosialyl oligosaccharides derived from salivary gland mucin glycoproteins of the Chinese swiftlet (genus *Collocalia*). Characterization of novel types of extended core structure, Gal beta(1–3)[GlcNAc beta(1–6)] GalNAc alpha(1–3)GalNAc(-ol), and of chain termination, [Gal alpha(1–4)]O–1[Gal beta(1–4)]2GlcNAc beta(1–.). J. Biol. Chem. **262**, 6650–6657
- 26 Sugahara, K., Ohi, Y., Harada, T., de Waard, P. and Vliegenthart, J. F. (1992) Structural studies on sulfated oligosaccharides derived from the carbohydrate-protein linkage region of chondroitin 6-sulfate proteoglycans of shark cartilage. 1. Six compounds containing 0 or 1 sulfate and/or phosphate residues. J. Biol. Chem. 267, 6027-6035
- 27 Vliegenthart, J. F. G., Dorland, L. and van Halbeek, H. (1983) High-resolution, 1Hnuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. Adv. Carbohydr. Chem. Biochem. 41, 209–374
- 28 Previato, J. O., Jones, C., Goncalves, L. P., Wait, R., Travassos, L. R. and Mendonca-Previato, L. (1994) O-glycosidically linked *N*-acetylglucosamine-bound oligosaccharides from glycoproteins of *Trypanosoma cruzi*, Biochem. J. **301**, 151–159

- 29 Vandekerckhove, F., Schenkman, S., Carvalho, L. P., Tomlinson, S., Kiso, M., Toshida, M., Hasegawa, A. and Nussenzweig, V. (1993) Substrate specificity of the *Trypanosoma cruzi* trans-sialidase. Glycobiology 2, 541–548
- 30 Ferrero-Garcia, M. A., Trombetta, S. E., Sanchez, D. O., Reglero, A., Frasch, A. C. and Parodi, A. J. (1993) The action of *Trypanosoma cruzi* trans-sialidase on glycolipids and glycoproteins. Eur. J. Biochem. **213**, 765–771
- 31 Maizels, R. M., de Savigny, D. and Ogilvie, B. M. (1984) Characterisation of surface and excretory-secretory antigens of *Toxocara canis* infective larvae. Parasite Immunol. 6, 23–37
- 32 Jones, S. J. and Baillie, D. L. (1995) Characterization of the *let-653* gene in *Caenorhabditis elegans*. Mol. Gen. Genet. **248**, 719–726
- 33 Gems, D. and Maizels, R. M. (1996) An abundantly expressed mucin-like protein from *Toxocara canis* infective larvae: the precursor of the larval surface coat glycoproteins. Proc. Natl. Acad. Sci. U.S.A. **93**, 1665–1670
- 34 Ogata, S. I. and Lloyd, K. O. (1982) Mild alkaline borohydride treatment of glycoproteins. A method for liberating both N- and O-linked carbohydrate chains. Anal. Biochem. **119**, 351–359

Received 12 February 2001/26 March 2001; accepted 19 April 2001

- 35 Volpi, N., Mucci, A. and Schenetti, L. (1999) Stability studies of chondroitin sulfate. Carbohydr. Res. 315, 345–349
- 36 Herman, T., Hartwieg, E. and Horvitz, H. R. (1999) Sqv mutants of *Caenorhabdilis elegans* are defective in vulval epithelial invagination. Proc. Natl. Acad. Sci. U.S.A. 96, 968–973
- 37 Herman, T. and Horvitz, H. R. (1999) Three proteins involved in *Caenorhabditis elegans* vulval invagination are similar to components of a glycosylation pathway. Proc. Natl. Acad. Sci. U.S.A. **96**, 974–979
- 38 Okajima, T., Yoshida, K., Kondo, T. and Furukawa, K. (1999) Human homolog of *Caenorhabditis elegans* sqv-3 gene is galactosyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. J. Biol. Chem. 274, 22915–22918
- 39 Bulik, D. A., Wei, G., Toyoda, H., Kinoshita-Toyoda, A., Waldrip, W. R., Esko, J. D., Robbins, P. W. and Selleck, S. B. (2000) Sqv-3, -7, and -8, a set of genes affecting morphogenesis in *Caenorhabditis elegans*, encode enzymes required for glycosaminoglycan biosynthesis. Proc. Natl. Acad. Sci. U.S.A. 97, 10838–10843

222. Conclusions

Ces 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. De plus, les quelques structures que non avons décrites démontrent l'existence de cinq nouvelles activités glycosyltransférasiques très particulières et imprévisibles: deux activités β -1,6-glucosyltransférasiques et une activité N-acétylgalactosamine β -1,6-galactosyltransférasique.

Ces prédictions ont déjà 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 \u00df6-GlcNAc-T(L) synthétise uniquement le noyau de type 2, la \u00df6-GlcNAc-T(M) les noyaux de types 2, 4 et l'antigène I, tandis que la I ß6-GlcNAc-T synthétise uniquement l'antigène I [Brockhausen, I. 1996]. Sur la base de comparaisons de séquences, ces gènes de C. elegans apparaissaient comme les meilleurs candidats pour coder des ß6-GlcNAc-T actifs sur les O-glycannes. Néanmoins, il n'était pas possible de déterminer si ils codaient pour des homologues de la ß6-GlcNAc-T(L), de la ß6-GlcNAc-T(M) ou de la I ß6-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(ß1-3)[Glc(\beta1-6)]GalNAc dans les O-glycannes de C. elegans, ils ont démontré que glv-1 transférait préférentiellement du glucose à partir de l'UDP-glucose sur un noyau de type 1 plutôt que de la N-acétylglucosamine [Warren, C.E. 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 très récemment caractérisé une UDPGalNAc:GlcNAcβ-R β1,4-N-acétylgalactosaminyltransférase de *C. elegans* [Kawar, Z.S *et al.*, 2002]. Les deux substrats accepteurs exogènes préférentiels testés de cette enzyme sont GlcNAc β -S-pNP et GlcNAc(β 1-6)Gal. De fait les auteurs suggèrent que cette enzyme puisse être impliqué dans la synthèse du noyau très inhabituel R-GalNAc(β 1-4)GlcNAc-Ser/Thr (LacdiNAc terminal) présent dans le O-glycanne 5 que nous avons décrit. 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*.

La structure de plusieurs N-glycannes a été publiée depuis nos travaux. Une étude, en particulier, démontre que *C. elegans* synthétise majoritairement des N-glycannes de type oligomannosidique et confirme la structure des N-glycannes que nous avons décris [Natsuka, S. *et al.*, 2002]. Plusieurs N-glycannes de type paucimannosisidique sont mono- ou difucosylés en position 6 ou/et 3 de l'unité GlcNAc proximale. De tels composés ont déjà été identifiés dans des glycoprotéines du venin d'abeille [Kubelka, V. *et al.*, 1993]. Enfin deux N-glycannes présentant un corps trimannosidique allongé par un unique résidu de GlcNAc ont été isolés en faible quantité, et sont les seuls N-glycannes de type complexe identifiés. Une autre étude rapporte la détection de N-glycannes contenant des résidus de 2-O-methyl fucose et 3- ou 4-O-methyl mannose, mais dont la structure n'est pas encore connue [Altmann, F. *et al.*, 2001]. Ces résultats confirment des observations personnelles qui n'étaient pas assez substantielles pour être présentées.

23. Purification préliminaire des mucines de *C. elegans*231. Introduction

Les quantités importantes de O-glycannes que nous avons isolées suggéraient la présence de composés hyper-glycosylés de type mucine chez *C.elegans*. Jusqu'a présent aucune étude ne s'est penchée sur la recherche systématique d'orthologues possibles des gènes MUCs chez *C. elegans*. Quelques études fonctionnelles isolées ont néanmoins mis en évidence des glycoprotéines putatives possédant un domaine riche en Ser/Thr pouvant correspondre aux domaines hautement O-glycosylés des mucines. En particulier, la protéine codée par *let-653* possède ce type de domaine et a été décrite par ses découvreurs comme une glycoprotéine de type mucine [Jones, S.J. & Baillie, D.L., 1995]. La mutation de ce gène engendre la mort à un stade larvaire suite au disfonctionnement du système de sécrétion. D'autre travaux ont mis en évidence le gène *lov-l* qui code une protéine membranaire de 3125



Fig. A15: Séparation des composés solubles totaux de *C. elegans* en douze fractions par ultracentrifugation isopycnique en chlorure de césium. La composition en monosaccharides de chaque fraction a été déterminée par analyse en chromatographie phase gazeuse. Les quantités de monosaccharides sont exprimées en une unité relative sur l'échelle 0 à 400 pour Glc et Gal (symboles pleins) et sur l'échelle 0 à 35 pour Man, GalNAc, GlcA et GlcNAc (symboles vides). La répartition des glycoconjugués est suggérée sur la base des densités et des compositions en monosaccharides.
acides aminés possédant un domaine extracellulaire similaire a une région PTS, et dont la délétion perturbe le comportement sexuel de l'animal [Barr, M.M. & Sternberg, P.W., 1999].

Par contre, les études effectuées sur le nématode parasite *Toxocara canis* ont permis d'identifier une famille de quatre gènes (Tc-MUC1 à 4) codant des glycoprotéines sécrétées de petite taille mais qui présentent une structure typique de mucines [Loukas, A. *et al.*, 2000]. Les quatre gènes codent des apomucines différentes qui sont toutes exclusivement synthétisées au stade larvaire infectieux. Le premier représentant de cette famille découvert, *nmuc1*, qui code la glycoprotéine TES-120 [Maizels, R.M *et al.*, 1984], possède une région riche en cystéines retrouvée dans de nombreux autres nématodes, dont *C. elegans* [Gems, D. & Maizels, R.M, 1996] ce qui confirme la présence de glycoprotéines de type mucine chez *C. elegans*.

Suite à notre première étude sur la structure des O-glycannes, nous avons cherché à isoler de tels composés de *C. elegans* dans le but d'étudier la glycosylation spécifique éventuelle à chaque glycoprotéine de type mucine. Comme nous l'avons rapporté dans un chapitre précédent, aucune étude ne s'est encore penchée sur l'identification du profil de glycosylation d'une mucine isolée et la traçabilité génétique de *C. elegans* rendait ce modèle très attractif pour débuter ce type d'étude. L'absence de travaux sur ce sujet tient essentiellement à la difficulté de purification de mucines naturelles à partir de milieux souvent très hétérogènes et à leurs propriétés physico-chimiques très particulières. La relative simplicité du modèle *C. elegans* laissait espérer une hétérogénéité glycoprotéique moindre, comparée aux autres modèles animaux. Cet avantage relatif est contrebalancé par le fait que l'on ne peut pas isoler de tissu individuel, ce qui oblige à travailler sur l'animal entier. Néanmoins, la facilité de culture du modèle et le fait que des quantités importantes de matériel soient disponibles rendaient cette étude réalisable.

232. Résultats

Les culots frais de *C. elegans* ont été mis en suspension dans une solution de Tris-HCl 50 mM pH 7.5, chlorure de guanidine 6M, triton 1% contenant un cocktail d'inhibiteurs de protéases, soniqués à plusieurs reprises et incubés sous agitation une nuit. Les composés solubles ont ensuite été séparés en fonction de leur densité par ultracentrifugation isopycnique dans une solution de chlorure de césium 2,4 M, 72h à 42000 t./mn.



Fig. A16: Séparation par ultracentrifugation isopycnique en CsCl des composés solubles de *C. elegans*.a) ler tour d'ultracentrifugation sur la fraction totale et c) 2ème tour sur les fractions de haute densité (3 à 9). Toutes les fractions après ultracentrifugation ont été testées par electrophorèses en gel d'agarose: électrophorèses après b) le ler tour puis le d) 2ème tour d'ultracentrifugation. Les electrophorèses sont colorés par le bleu de toluidine. **CS** correspond à la mobilité du chondroïtine sulfate.

Ce protocole a d'abord été testé sur une faible quantité d'échantillon (500 mg de culot). Douze fractions de densités différentes ont été collectées après ultracentrifugation et analysées. La composition en monosaccharides de chaque fraction suggérait la présence de plusieurs glycoconjugués différents (Fig. A15). En particulier, les trois fractions de plus haute densité (ρ >1,5) contenaient des quantité très importantes de glucose, ce qui démontrait la présence d'un polymère de glucose qui présentait toutes les caractéristiques structurales du glycogène après analyse (liaisons, densité, λ max). Les fractions 2 et 3 contenaient des quantités importantes d'acide glucuronique, suggérant la présence de protéoglycannes. Ensuite, par ordre décroissant de densité, les compositions suggéraient la présence de glycoprotéines de type mucine par la présence de galactose (1,442< ρ <1,5), de glycoprotéines substituées par des N-glycannes (1,348< ρ <1,374) et de glycolipides (ρ <1,318).

Cette expérience préliminaire avant suggéré la présence d'un matériel glucidique de haute densité, probablement de type mucine, nous avons utilisé les mêmes conditions expérimentales sur une quantité de nématodes importante (10 g.). Vingt et une fractions ont été collectées et testées pour les hexoses par une réaction à l'orcinol sulfurique et pour les acides uroniques par une réaction au carbazole, comme indiqué dans la figure A16a. Comme précédemment, nous observons un matériel de haute densité, très abondant, correspondant au glycogène. Les fractions 4-8 $(1,42 \le p \le 1,5)$ contiennent à la fois des quantités importantes d'héxose et d'acide uronique, ce qui suggère que les protéoglycannes et les mucines ont été copurifiées. Il faut néanmoins noter que si les mucines sont bien substituées par les O-glycannes contenant de l'acide glucuronique que nous avons précédemment analysés, elles devraient également réagir au carbazole L'électrophorèse sur gel d'agarose des fraction 3 à 9 révèle l'existence de deux composés majeur acides de mobilité différentes (Fig. A16b), le plus acide présentant une mobilité équivalente à celle d'un standard de chondroïtine sulfate. La digestion de la fraction 5 par une chondroïtinase dégrade en grande partie le composé de mobilité supérieur, ce qui confirme qu'il s'agit bien d'un protéoglycanne. Les acides nucléiques présentant également des densités équivalentes aux mucines, leur présence a été testée dans les échantillons. En effet, une électrophorèse sur gel d'agarose en présence de BET a confirmée la contamination par des acides nucléiques, qui sont totalement dégradés en présence de désoxyribonucléase.

Les fractions 4 à 10 ont été réunies et digérées par une DNase puis par un cocktail d'enzymes dégradant les protéoglycannes (voir section "experimental" de l'article 1) avant d'être soumis à une ultracentrifugation dans les mêmes conditions que précédemment. Le dosage des acides uroniques révèle la présence de composés de densités comprises entre 1,35 et 1,58 avec un maximum à une densité de 1,5 (Fig. A16c). Nous observons que ces fractions contiennent une mixture de glycoconjugués présentant des mobilités électrophorétiques comparables (Fig. A16d). Par contre, aucun composé de mobilité similaire aux chondroïtines sulfates n'est observé, confirmant que ces derniers ont été dégradés. L'analyse en présence de BET à également révélée l'absence totale d'ADN dans la préparation. Bien que comparables, toutes les fractions ne présentent pas des mobilités identiques, ce qui suggère la présence de

composés différents dans chaque fraction. En particulier, d'après leur mobilité et leur densité on peut distinguer deux populations qui semblent distinctes: la première est présente dans les fractions 3 à 7, la deuxième est présente dans les tubes de 6 à 12 et semble plus diffuse que la première. Néanmoins, ces composés semblent présenter des caractéristiques très proches et il est difficile de les distinguer.

Très peu d'anticorps caractéristiques des mucines sont actuellement disponibles pour aider à la caractérisation des mucines. Nous avons néanmoins testé les sérums spécifiques des mucines tégumentaires de *Xenopus laevis* sur les fractions de 1 à 11 dans l'éventualité d'un phénomène de reconnaissance croisé entre les mucines. De façon surprenante, non seulement les sérums dirigés contre FIM-A.1, -B.1 et -C.1 reconnaissaient les mucines présumées de *C. elegans*, mais elles démontrent une spécificité limitée envers celles ci (Fig. A17). Les sérums dirigés contre FIM-A.1 et –B.1 reconnaissent



Fig. A17: Electrophores en agarose après le 2^{ème} tour d'ultracentrifugation a) colorées au BT ou transférées sur membrane et révélées avec les sérums anti b) -FIM-A.1, c) -FIM-B.1 ou d) -FIM-C.1

majoritairement les fractions 6 à 11, avec une reconnaissance maximale pour les fractions 8 et 9 qui correspondent aux fractions majeures de la deuxième famille de mucines que nous avons définies plus haut, ce qui suggère que ces deux sérums seraient spécifiques à cette famille. Il est a noter que le sérum anti-FIM-A.1 reconnaît également une famille de composés de mobilité électrophorétique plus faible qui s'étale dans les fractions de 3 à 8. Le sérum anti-FIM-C.1 reconnaît majoritairement les fractions 8 à 11, avec un maximum pour la fraction 11. Ceci suggère la présence d'une troisième famille de composés de plus faible densité, distincte de la deuxième mais partiellement en mélange avec celle ci. Par contre, la famille de glycoprotéines de plus haute densité n'est reconnue par aucun des trois sérums. Ces trois sérums n'étant pas spécifiquement dirigés contre les mucines de *C. elegans*, on ne peut tirer aucune information précise d'ordre structural de ces expériences. Néanmoins, le fait qu'ils exhibent une forme de spécificité envers les glycoprotéines de C. elegans non seulement suggère l'existence de motifs communs entre les mucines de *C. elegans* non seulement suggère l'existence de motifs communs entre les mucines de xénope et deux des familles de glycoprotéines de haute densité de *C. elegans*, mais également démontre l'hétérogénéité structurale de ces composés.

Un deuxième anticorps, spécifique au régions riches en cystéines des mucines sécrétées, reconnaît spécifiquement les composés de la première famille de composés dans les fractions 3 à 8, ce qui démontre également l'hétérogénéité structurale des composés que nous avons isolés. Cette expérience confirme également l'existence de domaines typiquement de type mucine dans certains des composés de *C. elegans*.

En conclusion, nous avons purifié un mélange d'au moins trois familles de glycoprotéines a partir d'une fraction totale de *C. elegans*. Les différentes analyses auxquelles nous les avons soumises suggèrent fortement que ces composés partagent des propriétés communes avec les mucines. L'étape suivante consiste à séparer ces familles avant de pouvoir étudier la structure des glycannes et la séquence peptidique de chaque composé individuellement. La séparation des mucines étant très problématique, nous avons débuté une collaboration avec le Dr Sheehan de l'université de Caroline du Nord (UNC), qui a mis au point des techniques de purification adaptées à ce type de molécules et qui se chargera de cette partie de l'étude à partir du matériel que nous avons déjà préparé au laboratoire.

Alternativement, nous avons mis au point ensemble une variation de notre premier protocole de purification dans le but de purifier plus efficacement les différentes mucines de *C. elegans.* Nous avons utilisé le fait que *C. elegans* soient entouré d'une cuticule extrêmement résistante pour effectuer une première extraction sélective des mucines excrétées dans le milieu extérieur. Les nématodes intacts ont été incubés sous agitation dans une solution de chlorure de guanidine 6 M pendant une heure et centrifugés. Surnageant et culot ont été collectés. Les nématodes dans le culot ont été examinés par microscopie pour vérifier qu'ils étaient encore intacts puis pulvérisés après congélation dans l'azote liquide. Les composés solubles de ce matériel ont ensuite été extraits dans du chlorure de guanidine. Les deux extraits, avant et après pulvérisation, ont été déposés sur une colonne de gel filtration de très faible réticulation pour séparer les composés macromoléculaires. Les fractions exclues et incluses de la colonne ont ensuite été soumises à une première ultracentrifugation isopycnique en chlorure de césium. Les différentes fractions obtenues sont actuellement en cours d'analyse. Les premiers ont déjà révélé l'existence de glycoprotéines de haute densité différentes dans les fractions avant et après pulvérisation.

I GENERALITES

1 Les infections mycobactériennes 11. Introduction

Les mycobactéries peuvent être classées en trois groupes en fonction des pathologies qu'elles induisent: les mycobactéries responsables de la tuberculose (*Mycobacterium tuberculosis*, *M. africanum* et *M. bovis*), celles responsables de la lèpre (*M. leprae*) et les autres, dites mycobactéries atypiques ou non tuberculeuses. Les bactériologistes dénombrent des dizaines d'espèces de mycobactéries atypiques, dont la majorité sont présentes dans l'environnement. Ces mycobactéries peu ou pas pathogènes pour l'adulte sain, sont parfois responsables d'infections humaines sévères chez des individus immunodéprimés et par certains aspects similaires à la tuberculose, les mycobactérioses. Alors que des efforts importants sont consacrés à l'étude de *M. tuberculosis* et de la tuberculose, pleinement justifiés par l'ampleur considérable de l'épidémie de tuberculose et par sa gravité, relativement peu d'études d'ordre biochimique ou génétiques se penchent sur les pathogenèses liées aux mycobactéries atypiques. Nous allons brièvement décrire dans la suite de ce chapitre les principaux aspects liés aux infections dues à *M. tuberculosis* et les mycobactéries atypiques.

12. La tuberculose

La tuberculose est la première cause de mortalité liée à un agent infectieux unique. Sur les quinze dernières années, il a été estimé que 8 millions de personnes sont infectées et que 3 millions décèdent directement suite à la tuberculose chaque année. En termes généraux, la tuberculose serait actuellement responsable d'environ 7% de la mortalité humaine globale [Brennan, P.J., 1997].

L'infection tuberculeuse est le résultat de l'inhalation d'un faible nombre de bacilles de la tuberculose, majoritairement *M. tuberculosis,* également appelé bacille de Koch. Ces derniers sont alors immédiatement phagocytés par les macrophages alvéolaires des poumons. Les bacilles phagocytés survivent dans les macrophages où ils peuvent se répliquer de façon limitée au sein de vacuoles, les phagosomes. Rarement, chez les enfants ou chez les personnes immunodéprimées, les bacilles peuvent se disséminer précocement dans l'organisme et former de petite lésions ou des méningites. Néanmoins, dans la majorité des cas une réponse immunitaire cellulaire se développe 2 à 6 semaines suite à l'infection et entraîne le recrutement de lymphocytes et de macrophages activés qui vont en partie détruire les bacilles, et isoler le foyer infectieux par la formation d'une structure particulière, le granulome. L'infection est ainsi dans un premier temps stoppée mais une proportion inconnue de bacilles persistent au sein des macrophages, dans un état dormant. A ce stade, le seul indice de l'infection par *M. tuberculosis* est fournie par la réactivité à un test cutané contre un extrait de protéines purifiées du bacille (tuberculine ou PPD pour "purified protein derivative").

Dans un deuxième temps, les bacilles peuvent à tout moment déclencher une nouvelle infection qui provoquera la maladie. Les mécanismes régissant cette réactivation pulmonaire restent encore très mal connus. A partir du site initial d'infection, ils se répandront dans les poumons par les vaisseaux lymphatiques et sanguins pour générer une tuberculose pulmonaire, ou vers d'autres organes tels que la plèvre, le système lymphatique, les os, le système uro-génital, les méninges ou la peau pour donner naissance à des tuberculoses extrapulmonaires. L'ensemble des processus inflammatoires et pathologiques induisent des symptômes caractéristiques: faiblesse, fièvre, douleurs pulmonaires, toux et hémorragies. Le taux de mortalité suite à une tuberculose non traitée varie entre 40 et 60%. Un individu immunocompétent infecté par le bacille a environ 10% de chance de contracter la maladie suite à la réactivation des bacilles au cours de son existence. En revanche, ce pourcentage est d'environ 8% par an pour une personne infectée par le HIV qui présente un test positif à la tuberculine.

La résurgence de la tuberculose observée ces dernières années dans les pays en voie de développement et dans les pays industrialisés, est imputable à la dégradation générale des conditions sanitaires dans les grandes métropoles, à l'épidémie due au virus du SIDA ainsi qu'à l'émergence de souches multirésistantes aux antituberculeux.

13. Les mycobactérioses.

Chez l'homme, 95% des mycobactérioses sont liées à: *M. kansasii, M. avium, M. xenopi, M. scrofulaceum et M. marinum.* Ces germes sont tous présent dans l'environnement. Ils sont plus ou moins fréquents dans le sol et l'eau, selon les régions: *M. kansasii* est très souvent trouvé dans l'environnement aux Etats-Unis, tandis que *M. xenopi* n'est pratiquement jamais isolé aux Etats-Unis mais est fréquent en Europe de l'Est [Dautzenberg, B. & Mercat,

A., 1994]. La mycobactérie non tuberculeuse la plus pathogène, capable de provoquer de façon fréquente une maladie en absence de toute cause favorisante, si la charge bactérienne contaminante est élevée, est *M. kansasii*. Les autres germes ne deviennent pathogènes que dans trois circonstances: introduction accidentelle dans l'organisme (*M. marinum* en cas de plaie en milieu aquatique, *M. chelonae* en cas de petite chirurgie); lésions pulmonaires préexistantes (infection par *M. xenopi* ou *M. avium* sur séquelles de tuberculose ou de maladies respiratoires); immunodépression générale suite à un traitement immunosuppresseur ou chez les patients atteints du SIDA. Néanmoins, la résistance importante dont font preuve la majorité des mycobactéries atypiques aux antibiotiques habituellement efficaces sur *M. tuberculosis* et *M. leprae* pose de sérieux problèmes thérapeutiques.

Quatre symptômes cliniques sont observés suite à une mycobactériose: maladie pulmonaire, atteinte ganglionnaire, infection cutanée et infections disséminées. La pathologie d'une mycobactériose semble plus dépendre de l'hôte et du mode de contamination que de l'espèce mycobactérienne, bien qu'il existe quelques spécificités en fonction de la mycobactérie responsable (Tableau B1)

Espèce	Infection disséminée	Infection pulmonaire	Infection ganglionnaire	Infection cutanée
M. kansasii	***	****	0	****
M. marinum	*	0	0	* * * *
M. simiae	0	* * *	0	* *
M. szulgai	0	* * * *	0	* *
M. xenopi	**	* * * *	0	0
M. scrofulaceum	*	0	****	0
M. avium-intracellulare	****	* * * *	***	* *
M. malmoense	*	* * * *	0	*
M. ulcerans	0	0	0	****
M. fortuitum	**	0	0	**
M. chelonae	**	0	0	* * *
M. gordonae	0	0	0	*

Tableau B1: Pouvoir pathogène des mycobactéries atypiques, d'après Gbery, I.P. et al., 1996 ****: fort; ***: moyen; **faible; *: très faible; 0: non décrit

Avant la pandémie provoquée par le SIDA, les maladies pulmonaires représentaient la forme la plus commune des mycobactérioses. Les mycobactérioses pulmonaires et la tuberculose provoquent des symptômes très similaires et il peut être difficile de les distinguer. La majorité des patients atteints par cette forme de mycobactériose semblent présenter un terrain favorable aux infections suite à des lésions causées par d'autres maladies pulmonaires, dont la tuberculose. S'y ajoutent depuis plusieurs années les patients atteints par le VIH qui

ont contribués à la nette progression de la maladie. Non seulement, le nombre de personnes infectées par le SIDA est en constante augmentation, mais la proportion de cette population qui contracte des mycobactérioses pulmonaires atypiques semble augmenter également [Horsburgh, C.R., 1996].

Au contraire, les cas d'infections disséminées provoquées par les mycobactéries atypiques étaient extrêmement rares avant 1980. Ce type d'infection atteint presque exclusivement les personnes immunodéprimées, et l'extrême susceptibilité que développent les patients atteints du SIDA a conduit à une augmentation très importante du nombre recensé de cette affection. En effet, entre 15 et 24% des patients atteints du SIDA développeraient des infections disséminées. *M. avium* est de très loin le principal agent de ces infections aux Etats-Unis, suivi par *M. kansasii*, qui sévit principalement en Europe. L'infection s'effectue directement par l'environnement, par ingestion ou inhalation de bactéries et peut toucher tous les organes, en particulier le système gastro-intestinal, la moelle osseuse et les nodules lymphatiques.

2 L'enveloppe des mycobactéries

21. Introduction

La résistance exceptionnelle dont font preuve la majorité des mycobactéries envers les antibiotiques, les agents chimiothérapeutiques et de nombreux désinfectants chimiques trouve son origine en grande partie dans la présence d'une enveloppe de structure très inhabituelle et de la très faible perméabilité qu'elle impose. L'observation de cette enveloppe par microscopie électronique révèle l'existence de couches nettement différenciées présentant des densités électroniques variées (Fig. B1). L'un des enjeux de l'étude des enveloppes des mycobactéries est d'attribuer à chacune de ces couches une composition chimique définie grâce à l'étude individuelle de ses composants. Ces travaux, qui s'étalent sur plus de cinquante ans et qui sont toujours d'actualité, ont mis en évidence une prodigieuse hétérogénéité structurale au sein d'une même espèce et entre espèces différentes de mycobactéries. La préservation de l'enveloppe étant indispensable à leur survie, l'identification de structures lipidiques et glycolipidiques très originales laisse présager le développement de nouveaux agents antimycobactériens très spécifiques dirigés contre leurs voies de biosynthèse.



Fig. B1: a) Image de microscopie électronique de l'enveloppe de *M. kansasii*, d'aprés Paul, T.R & Beveridge, T.J., 1994. b) Interprétation de l'arangement des composés de l'enveloppe de mycobactéries, d'après Daffé, M. & Drapper, P,1998.

Actuellement, la plupart des anti-tuberculeux couramment utilisés pour traiter la tuberculose (éthambutol, éthionamide et isoniazide) inhibent la biosynthèse de composants majeurs de la paroi [Kremer, L. *et al.*, 2002]. De plus, il est apparu que nombre de composés isolés semblent directement impliqués dans les interactions hôte-pathogène et donc dans la virulence des mycobactéries. Néanmoins, les très nombreuses incertitudes concernant la localisation, et donc l'accessibilité, de ces composants rendent toute prédiction de leur activité *in vivo* difficile.

Sur la base des études structurales, les trois couches observées en microscopie ont été interprétées comme la membrane plasmique, la paroi et la capsule (Fig. B1) [Paul, T.R & Beveridge, T.J., 1992 et 1994]. Bien que vitale pour la mycobactérie, la membrane plasmique semble équivalente à celle de toutes les bactéries [Daffé, M. & Draper, M., 1998]. Elle présente une apparence asymétrique, la couche externe dense aux électrons étant plus épaisse que la couche interne. Cette différence à été associée à la présence spécifique de glycoconjugués dans la couche externe [Brennan, P. & Nikaido, H., 1995]. L'ultrastructure de la paroi est composée d'un complexe de trois macromolécules associées de manière covalente pour former une structure spécifique, le mycolyl-arabinogalactan-peptidoglycanne (mAGP). Cet assemblage qui est totalement insoluble est en grande partie responsable de la très faible perméabilité de l'enveloppe en générale et de la paroi en particulier. Cette caractéristique serait le résultat de la prévalence d'acides gras à longues chaînes, les acides mycoliques, qui sont retrouvés dans toutes les espèces de mycobactéries et d'actinomycètes [Minnikin, D.E. et al., 2002]. De très nombreux autres composés sont associés de façon non covalente à la paroi: quelques protéines mais essentiellement des lipides et des glycolipides. Enfin la capsule, dont la reconnaissance à été très tardive, est majoritairement composée de polysaccharides libres, de protéines et de quelques lipides et glycolipides [Daffé, M. & Etienne, G., 1999]. De fait, en dehors du complexe mAGP, l'ensemble des composants de l'enveloppe sont extractibles par des solvants appropriés, ce qui a grandement facilité leur étude. En effet, l'extraction exhaustive des composés libres de la paroi par une solution de SDS à 2% à chaud permet d'obtenir une fraction presque pure de mAGP [Daffé, M. et al., 1990]. Parmi les composés solubles on peut citer les glycolipides basés sur un phosphatityl-myo-inositolglycérol (LAM, LM et PIMs), sur un noyau de tréhalose (TMM, TDM, SL, LOS, DAT, TAT et PAT) ou sur un noyau phénolique (PGL), les glycopeptidolipides (GPL) et les polysaccharides neutres (AM et glucane). Certains de ces composés sont spécifiques d'espèce tandis d'autres sont trouvés ubiquitairement chez les mycobactéries, à quelques variantes près. Si de nombreuses informations structurales sont disponibles sur ces composés, leurs localisations exactes au sein de l'enveloppe est encore sujet à débat. Nous allons maintenant décrire succinctement les différentes familles de composés identifiés dans l'enveloppe des mycobactéries en faisant référence aux quelques activités biologiques qui leurs sont associées. Les lipoarabinomannanes (LAM) et lipomannanes (LM) ayant fait l'objet de nos études, nous nous attarderons plus longuement sur ces composés tandis que les activités biologiques qui leur sont associées seront examinées dans un chapitre séparé (Chapitre 4). Pour conclure, sur la base de la structure des composés isolés, nous présenterons un modèle d'organisation de l'enveloppe des mycobactéries.

22. Le complexe mycolyl-arabinogalactane-peptidoglycanne (mAGP)

Les mycobactéries produisent une paroi appelée paroi de chémotype IV. Le squelette insoluble de cette paroi est composé d'un complexe formé de trois macromolécules liées de façon covalente: le peptidoglycanne, l'arabinogalactane et les acides mycoliques. Cette macromolécule insoluble peut être facilement dissociée, ce qui a permis étude individuelle de ses trois constituants. Un long traitement en conditions acides douces permet de séparer le peptidoglycanne du mycolyl-arabinogalactane, tandis qu'un traitement alcalin à chaud permet d'isoler les acides mycoliques [Kenatsuna, F., 1968].

Le peptidoglycanne

Le squelette du peptidoglycanne est constitué d'un enchaînement de Nacétylglucosamines et d'acides N-glycolylmuramiques liés en β 1-4 [Adam, A. *et al.*, 1969]. Ces polymères glucidiques sont reliés entre eux par des tétrapeptides liés sur les acides Nglycolylmuramiques, formant ainsi une matrice tridimensionnelle stable. Dans toutes les mycobactéries examinées, à l'exeption de *M. leprae*, ce tétra-peptide présente une séquence Lalanyl-D-isoglucaminyl-méso-diaminopimélyl-D-alanine avec un acide diaminopimélique amidé [Weitzerbin-Falszpan, J. *et al.*, 1970]. Ce type de peptidoglycanne (Ala γ) est le type le plus commun chez les bactéries [Chatterjee, D., 1998]. Il se distingue néanmoins de deux façons chez les mycobactéries: d'une part l'acide muramique est N-glycolylé et non N-acétylé, d'autre part les polymères glucidiques sont non seulement reliés *via* des ponts entre l'alanine et l'acide diaminopimélique mais également par des ponts entre deux acides diaminopiméliques [Weitzerbin, J. et al., 1974]. L'acide muramique est également substitué à hauteur de 10 à 12% en position C-6 par l'arabinogalactane.

L'arabinogalactane

Plusieurs travaux anciens avaient suggéré la présence d'un groupement phospho-diester en position C-6 de l'acide muramique de plusieurs espèces de mycobactéries [Kanetsuna, F., 1968] pouvant être à l'origine de la liaison entre l'arabinogalactane et le peptidoglycanne. Pour résoudre la structure de la partie terminale réductrice de l'arabinogalactane liée au peptidoglycanne, le complexe mAGP a été per-méthylé, fragmenté par hydrolyse acide ménagée et analysé en GC/MS [McNeil, M. *et al.*, 1990]. Les fragments identifiés ont permis de démontrer que la liaison de l'arabinogalactane au peptidoglycanne s'effectuait via un trisaccharide -5)-D-Galf(1-4)-L-Rhap(1-3)-D-GlcNAc. L'analyse en RMN du phosphore du peptidoglycanne a de plus montré que l'unité de GlcNAc terminale était liée via une liaison phospho-di-ester au C-6 de l'acide muramique, ainsi que les premières analyses le suggéraient (Fig. B2).



Fig. B2: D'aprés McNeil, M. *et al.*, 1990. Structure proposée du motif de liaison entre l'arabinogalactane et le peptidoglycanne. La séquence du trisaccharide a été déterminée par l'analyse des fragments d'hydrolyse acide indiqués sur la figure par des barres horizontales. Les unités de Gal sont toutes du β-D-Galf, celles de Rha sont L et de GLcNAc sont D. Mur représente l'acide muramique N-glycolylé.

L'utilisation de la même technique de fragmentation et d'analyse a permis de mettre en évidence les principales caractéristiques structurales de l'arabinogalactane de *M. tuberculosis*. En dehors de la partie terminale, il est composé exclusivement de D-Gal*f* et de D-Ara*f*

[McNeil, M. et al., 1987] arrangés sous forme d'homopolymères distincts liés entre eux. Le domaine galactane est constitué d'une chaîne linéaire unique d'unités β-D-Galf liées entre elles alternativement en positions C-5 et C-6 [Daffé, M. et al., 1990]. Le domaine arabinane est quant à lui constitué de plusieurs chaînes ramifiées de α -D-Araf substituant le domaine galactane en position C-6 de certaines unités de -5)Galf. Chaque chaîne d'arabinose semble majoritairement être composé de 23 unités d'a-D-Araf. Les parties terminales non-réductrices de ces chaînes consistent en des motifs hexaarabinofuranosyles branchés [β -D-Araf-(1-2)- α -D-Araf]₂-3,5-α-D-Araf-(1-5)-α-D-Araf-(1-5)-α-D-Araf [Besra, G.S. et al., 1995]. Ces études ont permis d'estimer qu'une moyenne de trois chaînes arabinanes substituaient un domaine galactane, lui même composé d'une trentaine de résidus de galactose en moyenne. L'isolement d'une chaîne linéaire de 23 unités de Galf non substituées, identifiée comme la partie terminale non-réductrice du domaine galactane, a démontré que les chaînes arabinanes substituaient le domaine galactane à proximité du peptidoglycanne. Une interprétation divergente de cette expérience implique l'existence de polysaccharides liés aux peptidoglycannes composés uniquement d'un domaine galactane non arabinosylé. L'examen des domaines arabinogalactanes isolés de M. tuberculosis, M. bovis, M. avium, M. smegmatis et M. leprae n'a révélé aucune différence structurale, ce qui suggère que les caractéristiques décrites chez M. tuberculosis, y compris la nature du disaccharides Rha-GlcNAc terminal, sont communes à la majorité des espèces mycobactériennes [Daffé, M. et al., 1993]. Une étude assez ancienne avait également suggéré l'existence de galactosamine (GalN) dans la fraction non extractible de la paroi de plusieurs espèces de mycobactéries à croissance lente, dont M. lepraemurium [Draper, P., 1971]. Plus récemment, une étude menée sur les parois de quatre espèces pathogènes pour l'homme (M. tuberculosis, M. leprae, M. kansasii, and M. avium), a confirmé que la GalN était attachée de façon covalente sur le complexe mAGP, apparemment en position C-2 d'une unité d'arabinose substituée en 5 [Draper, P. et al., 1997].

Les acides mycoliques

Les acides mycoliques constituent le troisième type de macromolécules du complexe mAGP. Ce sont des acides gras de très haute masse moléculaire (C70-C90), α -alkylés et β -hydroxylés. Ils sont majoritairement associés au complexe insoluble mAGP au sein de la paroi des mycobactéries, mais sont également présent sous des formes extractibles associés

aux tréhalose mono- di- et trimycolates (TMM, TDM, TTM) (voir chapitre 2321). Les acides mycoliques sont élaborés par toutes les espèces de mycobactéries, et des composés équivalents, mais de plus faible masse moléculaire, ont été identifiés dans les genres *Corynebacterium, Nocardia* et *Rhodococcus* [Besra, G.S. & Chatterjee, D., 1994]. La présence d'un groupement hydroxyle en position C-3 rend les acides mycoliques susceptibles à la pyrolyse, ce qui a grandement facilité l'étude de leur structure [Brennan, P.J. & Nikaido, H., 1995]. Celle ci permet de générer deux chaînes alkylées, la branche α contenant la fonction hydroxyle et le méromycolate contenant le carbone C-3 (Fig.B3).



Fig. B3: Structure générale des acides mycoliques, d'aprés Brennan, P.J. & Nikaido, H., 1995. R, X, et Y représentent les parties variables de la molécule.

A l'exception de faibles variations de longueur la structure de la branche α est constante chez toutes les mycobactéries. Les variations structurales se concentrent donc au niveau du méromycolate et concernent surtout la longueur de la chaîne et la présence éventuelle d'insaturations, de groupements cyclopropanes, méthylènes, méthyles, cétones ou méthoxy. Six familles distinctes d'acides mycoliques sont communément décrites (tableau B2). Au sein de chaque famille, des modifications de structures interviennent, générant une grande diversité structurale. De fait, chaque espèce de mycobactérie synthétise un panel d'acides mycoliques unique [Watanabe, M. et al., 2002], dont l'analyse permet une détermination précise de l'espèce [Butler, W.R. & Guthertz, L.S., 2001]. A titre d'exemple, les structures des acides mycoliques majoritairement synthétisées par M. tuberculosis et M. smegmatis sont décrites dans la figure B4. M. tuberculosis modifie ses acides mycoliques par



Tableau B2: Exemples des principales familles d'acides mycoliques. D'après Besra, G.S. & Chatterjee, D., 1994. cyclopropanation alors que *M. smegmatis*, une espèce saprophyte à croissance rapide, n'en est pas capable.



Fig. B4: Structure des acides mycoliques majeurs de M. tuberculosis et M. smegmatis.

١.

Les acides mycoliques sont associés au domaine polysaccharidique du complexe mAGP sur la partie arabinane. Pour déterminer leur position exacte, le polysaccharide intact a été per-méthylé en milieu neutre, ce qui permet de conserver les groupements acyles alkalilabiles [Prehm., P., 1980], puis éthylé en milieu basique avant d'être hydrolysé, acétylé et analysé en GC/MS [McNeil, M. et al., 1991]. De cette manière, les auteurs ont obtenus des dérivés partiellement méthylés, éthylés et acétylés dans lesquels les groupements éthyle reflétaient les positions d'acylation. Cette méthode, couplée à l'utilisation de la dégradation de Smith par oxydation periodique et réduction, a démontré que les acides mycoliques substituaient spécifiquement les unités de β -Araf terminales et les unités d' α 2-Araf d'un même motif pentaarabinoside en position C-5 pour former des motifs tétra-mycolylpenta-arabinosides sont substituées par des acides mycoliques, tandis que les mAGP de *M. leprae*, *M. bovis BCG* et *M. smegmatis* présentent des taux de mycolylation plus faibles.



Fig. B6: D'après Besra, G.S. *et al.*, 1995. Modèle proposé de la structure du complexe mycolyle-arabinogalactane-peptidoglycanne des mycobactéries.



Fig. B5 : D'après McNeil, M. et al., 1991. Structure proposée des motifs penta-arabinosides mycolylées de la paroi de *M. tuberculosis*. La nature des chaînes d'acides mycoliques associées dans ce schéma sont représentatives des structures majoritairement synthétisées par cette espèce.

L'ensemble des données collectées sur chaque domaine du complexe mAGP a permis de définir la structure de la molécule native, proposé dans la figure B6. Les longues chaînes d'acides mycoliques seraient ordonnées de manière parallèle, tournées vers l'extérieure de l'enveloppe et formeraient ainsi une couche ininterrompue autours de la mycobactérie (Fig. B7) [Crick, D.C. *et al.*, 2001].



Fig. B7: Modèle d'arrangement du complexe mAGP à la surface de la membrane plasmique des mycobactéries, d'après le modèle de Minnikin, D.E., 1982 [Crick, D.C. et al., 2001].



Fig. B8: Structure des PIM₂ à PIM₆.

La quantification des acides mycoliques de BCG a démontré qu'ils étaient en nombre suffisant pour couvrir intégralemnt la surface de la mycobactérie, ce qui est en accord avec ce modèle [Nikaido, H. et al., 1993]. Cette couche très épaisse d'acides mycoliques formerait la couche transparente aux électrons à laquelle nous avons fait précédemment référence, tandis que la couche dense aux électrons sous-jacente contiendrait le peptidoglycanne et les arabinogalactanes. La transparence aux électrons serait liée à l'incapacité des sels métalliques solubles dans l'eau à pénétrer la couche très hydrophobe formée par les acides mycoliques. De nombreuses études, par l'utilisation de souches mutantes, ont démontré que les acides mycoliques ont directement impliqués dans la régulation de la fluidité et de la perméabilité de la paroi des mycobactéries [Jackson, M. et al., 1999]. En particulier, il a été montré que la réduction de 40% du taux d'acides mycoliques transférés sur le complexe mAGP par un mutant de M. tuberculosis, augmentait considérablement sa perméabilité envers des composés hydrophobiques et hydrophiliques. De même, un mutant de M. smegmatis synthétisant exclusivement des acides mycoliques réduits à la seule chaîne méromycolate non seulement synthétisait une paroi présentant une ultrastructure anormale, mais présentait une hypersensibilité aux antibiotiques hydrosolubles [Wang, L. et al., 2000]. La modification de la nature des acides mycoliques peut également entraîner des changements importants de virulence des mycobactéries. Ainsi, il a été observé que la substitution des kétomycolates par des methoxymycolates suite à la sur-expression d'une O-méthyltransférase dans une souche de M. tuberculosis diminuait grandement sa survie dans des cellules macrophagiques THP-1 [Yuan, Y. et al., 1998]. Par contre, malgré les différences marquées existantes entre les profils d'acides mycoliques des différentes espèces, aucune corrélation n'a jamais été établie entre la présence d'acides mycoliques particuliers et la pathogénie de la souche. De même, mis a part des différences mineures dans la structure du peptidoglycanne, aucune différence significative n'a été mise en évidence dans la structure du peptidoarabinogalactane pouvant être corrélée à des différences de virulence.

23. Les glycolipides extractibles

231. Les glycolipides de type phosphatidyl-*myo*inositol 2311. Les phophatidylinositol-mannosides (PIMs)

Dés 1939, les PIMs ont été isolés à partir de fractions phospholipidiques du bacille de Koch [Anderson, R.J. 1939]. L'analyse des hydrolysats acides et alcalins de ces composés avait déjà mis en évidence la présence de mannose, phosphate et inositol. Vingt cinq ans plus tard, l'équipe de Ballou en a établit la structure exacte en démontrant que les PIMs isolés de *M. bovis* BCG et *M. phlei* constituaient une famille de phosphatidyl-*myo*-inositol di-, tri-, tetra et penta-mannosylés (PIM₂ à PIM₅) [Ballou, C.E. *et al.*, 1963]. De plus, ces auteurs ont établis que l'inositol de PIM₂ était mannosylé en positions C-2 et C-6 [Lee, Y.C. & Ballou, C.E., 1964] et que l'élongation subséquente de la chaîne mannosylée dans les PIM₃ à PIM₅ s'effectuait 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 4 mannoses) qui détermine la nature du PIM (PIM₂ à PIM₅). Plus récemment, cette structure a été confirmée par l'analyse de PIMs hexa-mannosylés (PIM₆) déacylés, isolés de mycobactéries à croissance rapide [Chatterjee, D. *et al.*, 1992] et de deux souches de *M. bovis* [Severn, W.B *et al.*, 1998]. Ces études ont définitivement établi l'archétype de la structure de la partie oligosaccharidique des PIMs comme le montre la figure B8.

Il est néanmoins très vite apparu que l'acylation des PIMs ne se limitait pas à l'acylation des positions C-1 et C-2 du glycérol. En effet, plusieurs études ont rapporté la présence de formes di-, tri- et tétra-acylées des PIMs (Ac2-, Ac3- et Ac4PIMs) isolés de M. tuberculosis [Pangborn, M.C. & Mc Kinney, J.A., 1966; Leopold, K. & Fischer, W., 1993], de M. phlei [Brennan, P. & Ballou, C.E., 1967] et de Mycobacterium 607 [Khuller, G.K. & Subrahmanyam, D., 1968]. Ces études ne déterminaient néanmoins pas le positionnement des acides gras surnuméraires. L'analyse en FAB-MS des dérivés per-methylés en milieu neutre de PIM₂ et PIM₆ de *M. tuberculosis* et *M. leprae* a permis de mettre en évidence une troisième position d'acylation possible sur le C-6 de l'unité de mannose substituant l'inositol en C-2 (Fig B8) [Khoo, K.H. et al., 1995a]. Enfin, l'isolement et l'analyse par RMN de Ac₄PIM₂ isolé de M. bovis BCG a permi de préciser la quatrième position d'acylation en C-3 de l'unité de myoinositol [Gilleron, M. et al., 1999]. 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 différentes combinaisons possibles entre degré d'acylation, degré de glycosylation et nature des acides gras engendrent une famille de composés très hétérogène. Néanmoins, aucune variabilité dans la structure des PIMs n'a jamais été mise en évidence entre différentes espèces ou souches de mycobactéries.

2312. Les lipomannanes (LM) et lipoarabinomannanes (LAM)

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. Contrairement aux PIMs, ces deux composés n'ont été purifiés sous leur forme native que très tardivement. Les isolements du LAM de *M. tuberculosis* et *M. leprae* [Hunter, S.W. et al., 1986] et du LM de *M. tuberculosis* [Hunter, S.W. et al., 1990] ont permis d'envisager l'étude structurale complète de ces composés dont les parties polysaccharidiques n'avaient jusqu'alors été isolées qu'après un traitement alcalin drastique des mycobactéries [Azuma, I. et al., 1968; Misaki, A. et al., 1977]. Dans ces conditions expérimentales, les auteurs n'avaient pas pu distinguer l'origine des composés qui pouvaient provenir soit d'un polysaccharide neutre natif, en l'occurrence le mannane ou l'arabinomannane, ou d'un lipoglycanne.

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". La structure de ces deux composés est schématisée dans la figure B9.



Fig. B9: Représentation schématique de l'agencement des domaines du LM et du LAM.



Fig. B10: Profils d'acylation des ancres phosphatidyl-myoinositols des LM et LAM mis en évidence en RMN chez *M. tuberculosis* et *M. bovis* BCG, d'après la nomenclature de Nigou et collaborateurs, 1999a. Les phosphates P1, P3, P4 et P5 présentent des glissements chimiques différents en fonction du degré d'acylation du glycérol et du myo-inositol. Seules les positions d'acylation certifiées pour chaque type sont représentées. * correspond à la quatrième position mise en évidence par Khoo, K.H. *et al.*, 1995a.

Structure de l'ancre phospho-myoinositol

L'ancre phospho-myo-inositol des LM et LAM est basée sur celle des PIMs. La comparaison des structures de PIM₂ et du produit de digestion du LM de *M. tuberculosis* par une α -mannosidase a démontré sans ambiguïté que l'ancre lipidique du LM, et par extension celle du LAM, était constituée d'un glycérol-3-phospho-myo-inositol substitué en positions C-2 et C-6 de l'inositol par du mannose [Chatterjee, D. *et al.*, 1992]. La structure exacte de ce motif a été ensuite confirmée par une étude RMN exhaustive du LAM de *M. bovis* BCG [Venisse, A., *et al.*, 1995]. De façon concomitante, une étude en FAB-MS démontrait que les ancres lipidiques des PIMs, LM et LAM présentaient des états d'acylations identiques caractérisés par la présence de un à quatre acides gras [Khoo, K.H. *et al.*, 1995a]. Ces acides gras sont majoritairement les acides palmitique, tuberculostéarique et stéarique, et dans une moindre mesure les acides myristique (C₁₄), heptadécanoïque (C₁₇) et 10-méthylheptadécanoïque [Hunter, S.W. *et al.*, 1986; Leopold, K. & Fischer, W., 1993]. Ces

La définition exacte des différents états d'acylation des LAM et LM natifs n'a pu être abordée que grâce aux développements et à l'application des techniques hétéronucléaires ³¹P-¹H HMQC à l'étude de ces composés apportés par le groupe de G. Puzo. En particulier, ces études ont permis de caractériser quatre états d'acylation, dénommés P1, P3, P4 et P5, qui diffèrent par l'occupation des sites potentiels d'acylation du glycérol et de l'inositol (Fig. B10) [Nigou, J. et al., 1999a]. Seul le site d'acylation en C-3 du mannose terminal reste inaccessible à cette technique. L'application de cette technique aux fractions pariétales et cellulaires des LAMs de M. bovis BCG et M. tuberculosis H37Rv a mis en évidence des différences importantes dans les profils d'aculation des LAMs en fonction de leurs origines (tableau B3) [Nigou, J. et al., 1999; Gilleron, M. et al., 2000]. En particulier, les fractions pariétales, ne présentant qu'un seul état d'acylation, étaient systématiquement plus homogènes que les fractions cellulaires. De plus, l'analyse des acyl-glycérols partiellement acétylés libérés par acétolyse des LAMs de M. bovis BCG a démontré que la fraction cellulaire était substituée de manière homogène par un acide gras particulier uniquement retrouvé dans cette espèce et définie comme un acide 12-O-méthoxypropanoyl-12-hydroxystéarique [Nigou, J. et al., 1997]. La seule étude publiée sur l'état d'acylation du LM de BCG suggère que LM et LAM présentent des états d'acylation similaires au sein d'une même espèce [Gilleron, M. et al., 1999]:

	Diacylglycérol Monoacylglycéro			lglycérol
	P1	P3	P4	P5
ManLAM pariétal de H37Rv	0%	100%	0%	0%
ManLAM cellulaire de H37Rv	22%	67%	0%	11%
ManLAM pariétal de BCG	0%	0%	0%	100%
ManLAM cellulaire de BCG	38%	42%	8%	12%

Tableau B3: D'après Gilleron, M. et al., 2000. Abondance relative des différents états d'acylation des LAM isolés de *M. tuberculosis* et *M. bovis* BCG. Chaque échantillon de LAM a préalablement été séparé au cours du fractionnement en deux fractions distinctes.

De manière surprenante, il a été démontré que la forme majoritaire du LAM de *M. smegmatis* présentait un glycérol non acylé [Gilleron, M *et al.*, 1997]. Cette forme particulière de LAM est néanmoins différente de l'AM par la présence du motif phospho-inositol-glycérol, et a été dénommée GAM pour glycéro-arabinomannane.

Le corps polysaccharidique

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 D-Man $p(\alpha 1-6)$ partiellement substituées en α -1,2 par un résidu de D-Manp. Cet arrangement à été décrit à l'identique chez M. tuberculosis [Chatterjee, D. et al., 1991], M. bovis BCG [Venisse, A. et al., 1995] et M. smegmatis [Khoo, K.H. 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, comme le suggère la largeur de la bande correspondant au LM après électrophorèse sur gel de polyacrylamide. Ceci a été confirmé grâce à la mesure de la masse du domaine mannane du LM de M. smegmatis par MALDI qui a montré que sa taille varie de 17 à 35 mannoses, avec un maximum de 26 unités [Khoo, K.H., 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, D. & Khoo, K.H., 1998]. La taille du domaine mannane varie également en fonction de l'espèce: en moyenne 26 chez M. smegmatis, 20 chez M. tuberculosis Erdman [Chatterjee, D. et al., 1993] et 18 chez M. bovis BCG [Venisse, A. et al.,

1995]. De même, le degré de substitution des chaînes de mannose semble varier en fonction de l'espèce, comme le souligne le tableau B4.

	M. tub	erculosis	M. bovís	M. leprae	<u>М</u> . sp.	M. sm	egmatis
	H37Rv	Erdman	BCG			ATCC	mc ² 155
% substitution	64	69	71	66	60	57	47

Tableau B4: Comparaison des degrés de substitution des domaines mannanes des LAM calculées sur la base de la méthylation par la formule 2,6-Man/(2,6-Man + 6Man) [Khoo, K.H. et al., 1995b]

Néanmoins, ce résultat a été nuancé par l'utilisation d'une autre méthode de calcul basée sur la quantification des dimères Man(α 1-2)Man après acétolyse du domaine mannane [Nigou, J. *et al.*, 2000]. Les auteurs de cette étude n'ont pas observé de différence significative entre les LAMs isolés de *M. bovis* BCG, *M. tuberculosis* Erdman, H37Ra et H37Rv.

Au contraire du domaine mannane, le domaine arabinane est strictement spécifique au LAM. Celui ci est constitué d'un corps linéaire d'unités (α 1-5)D-Araf dont certaines sont substituées en α -1,3 par des chaînes latérales courtes d'unités de D-Araf. La structure exacte de ces chaînes latérales a été déterminée par analyse en GC/MS des fragments obtenus après dépolymérisation partielle de LAM per-méthylé [Chatterjee, D. *et al.*, 1991] puis confirmée par de nombreuses études RMN des LAM intacts. L'ensemble de ces études a montré que seuls deux types de chaînes latérales existaient chez les espèces mycobactériennes analysées: des chaînes linéaires tétrasaccharidiques (Ara₄) de séquence Ara(β 1-2)Ara(α 1-5)Ara(α 1-5)Ara(α 1-6)Ara(α 1-6)Ara(Ara(α



Fig. B11: Structure des chaînes latérales branchées en position C-3 d'une unité de 3,5-Ara du domaine arabinane. a) motif tétraarabinoside, b) motif hexaarabinoside.

On notera la ressemblance entre les motifs hexaarabinosides terminaux des LAMs et de l'arabinogalactane. Néanmoins, ces motifs sont finalement arrangés différemment au sein de leurs domaines arabinanes respectifs. L'analyse en MALDI révèle que le LAM est plus polydisperse que le LM [Vénisse, A. et al., 1993], ce qui suggère une hétérogénéité importante de la taille du domaine arabinane. Seules des différences minimes dans la taille de ce domaine ont été constatées en fonction de l'espèce: il est constitué d'une moyenne de 70 unités d'arabinoses chez M. bovis BCG [Venisse, A. et al., 1993] et M. smegmatis [Khoo, K.H. et al., 1996] et de 80 chez M. tuberculosis [Hunter, S.W. & Brennan, P., 1990]. Néanmoins, la structure de base du domaine arabinane semble être conservée chez toutes les espèces mycobactériennes étudiées à ce jour. Les premières études structurales du LAM natif ont mis en évidence la présence de divers acides organiques par dosage en GC/MS, en particulier des acides lactiques et succiniques [Hunter, S.W. et al., 1986]. Une étude ultérieure en RMN sur le LAM natif a permis de localiser l'acide succinique sur la position C-2 d'un arabinose disubstitué en positions C-3 et C-5, ce qui a définitivement démontré que ce substituant était associé de manière covalente au LAM [Delmas, C. et al., 1997]. Par ailleurs, le nombre d'acides succiniques par molécule de LAM varie entre 1 et 4 en fonction de la souche de M. bovis BCG utilisée. Par contre aucune étude n'est venu confirmer les premières observations quant à la présence éventuelle d'acide lactique.

Si la structure du corps arabinane est connu, ses modalités d'attachement sur le corps mannane restent encore obscurs. Les modèles de la littérature le présentent généralement sous la forme d'une chaîne unique substituant le corps mannane en position terminale, mais aucune donnée expérimentale ne permet de corroborer cette hypothèse. Ce point précis pourrait néanmoins être d'importance. En effet, si le domaine arabinane était composé de quelques chaînes et non d'une seule, leur longueur serait beaucoup plus faible, ce qui aurait des répercutions sur l'accessibilité de ses motifs terminaux (coiffe) en surface de la mycobactérie.

Les coiffes

Le corps arabinane du LAM peut être à son tour substitué par des motifs variés regroupés sous le terme général de "coiffe". Deux types de coiffes ont été mis en évidence chez les espèces étudiées. Les mycobactéries à croissance lente *M. tuberculosis* [Chatterjee, D. et al., 1992b], *M. bovis* [Prinzis, S. et al., 1993; Venisse, A. et al., 1993], *M. avium* [Khoo, K.H. et al., 2001] et *M. leprae* [Khoo, K.H. et al., 1995b] sont substituées par des coiffes de

$$Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Man\alpha \stackrel{2}{\rightarrow} Man\alpha \stackrel{5}{\rightarrow} Ara\beta \stackrel{2}{\rightarrow} Ara\alpha \stackrel{5}{\rightarrow} Ara\alpha$$

$$Ara\beta \xrightarrow{2} Ara\alpha \xrightarrow{5} Ara\alpha \xrightarrow{5} Ara\alpha$$

$$Man\alpha \xrightarrow{5} Ara\beta \xrightarrow{2} Ara\alpha \xrightarrow{5} Ara\alpha \xrightarrow{5} Ara\alpha$$

$$Man\alpha \xrightarrow{2} Man\alpha \xrightarrow{5} Ara\beta \xrightarrow{2} Ara\alpha \xrightarrow{5} Ara\alpha \xrightarrow{5} Ara\alpha$$

$$Man\alpha \xrightarrow{2} Man\alpha \xrightarrow{5} Ara\beta \xrightarrow{2} Ara\alpha \xrightarrow{5} Ara\alpha \xrightarrow{5} Ara\alpha$$

$$Man\alpha \xrightarrow{2} Man\alpha \xrightarrow{5} Ara\beta \xrightarrow{2} Ara\alpha \xrightarrow{5} Ara\alpha \xrightarrow{5} Ara\alpha$$



type oligomannosidiques, tandis que les espèces à croissance rapide dont *M. smegmatis* présentent une coiffe de type phophatidyl-*myo*inositol [Khoo, K.H. *et al.*, 1995b, Gilleron, M. *et al.*, 1997]. Les LAMs coiffés par des oligomannosyls ont été désignés ManLAM, par opposition aux LAMs non coiffés par des oligomannosyls nommés AraLAM. Cette dernière dénomination a été attribuée avant que la présence de coiffes de phosphatidyl-inositol n'ait été mise en évidence dans les LAM non coiffés par des oligomannosides. Depuis, certains auteurs ont proposé d'utiliser la dénomination PI-LAM pour les LAMs coiffés par de phosphatidyl-inositol, pour faire la différence avec un éventuel LAM ne présentant ni coiffe de phospho-inositol, ni coiffe de mannose [Gilleron, M. *et al.*, 1997].

Les coiffes oligomannosylées sont constituées de mono-, di- ou trimères de $(\alpha 1-2)D$ -Manp. Elles substituent en α -1,5 les arabinoses en positions terminales non-réductrices des motifs Ara4 et Ara6. L'analyse des motifs Ara4 et Ara6 mannosylés libérés du ManLAM de M. tuberculosis par l'action d'une endo-arabinosidase a permis de mettre en évidence un grand nombre d'arrangements possibles (Fig. B12) [Chatterjee, D. et al., 1993]. La quantification de chacune de ces formes a montré que les espèces Man₂Ara₄ et Man₄Ara₆ étaient les plus représentées. Ceci est confirmé par l'analyse des oligomannosides libérés de différentes espèces mycobactériennes qui a montré que le motif dimannosidique était systématiquement majoritaire par rapport aux deux autres formes [Nigou, J. et al., 2000]. Cette règle a récemment été battue en brèche par l'analyse du ManLAM d'un isolat clinique de M. avium résistant à l'éthambutol dont les chaînes Ara4 et Ara6 étaient majoritairement monomannosylées [Khoo, K.H. et al., 2001]. De même, il est apparu que le degré de substitution des chaîne arabinanes par les oligomannosides différait de manière significative en fonction de l'espèce (Tableau B5). On observe sur le tableau B5 que les deux méthodes de calcul utilisées dans la littérature donnent des résultats très différents. De plus, elles ne présument en rien du nombre d'unités oligomannosidiques par molécule de LAM puisqu'elles ne tiennent pas compte du pourcentage de branchement du corps arabinane. De fait, une méthode de quantification exacte par électrophorèse capillaire des coiffes mannosidiques, après hydrolyse sélective du corps arabinane, a été récemment développée [Monsarrat, B. et al., 1999]. Comme on l'observe dans le tableau B5, l'utilisation de cette technique confirme non seulement que le nombre d'unités oligomannosidiques est variable en fonction de l'espèce, mais également que ce dernier n'est pas forcement en relation avec le degré de substitution des chaînes Ara₄ et Ara₆.



Fig. B13: D'aprés Gilleron, M. *et al.*, 1997. Détermination par RMN de la position de la coiffe de phosphatidyl-*myo*inositol sur le LAM isolé de *M. smegmatis*. a) spectre ¹H-¹H HOHAHA, b) spectre ³¹P-¹H HMQC, c) ³¹P-¹H HMQC-HOHAHA, d)et e) interprétation des corrélations ${}^{3}J_{P-H}$ et ${}^{3}J_{H-H}$ observées sur les spectres c) et c) pour les phosphores P1 et P2 respectivement. Sur le spectre ³¹P-¹H HMQC P1corrèlait avec les protons H-1 d'un inositol glycosylé en C-2 et C-6 et H-3,3' d'un glycérol, ce qui a démontré qu'il fait partie de l'ancre phosphatidyl-*myo*inositol. P2 corrèlait directement avec les protons H-5,5' d'une unité d'arabinose et H-1 d'un inositol non substitué, démontrant que P1 ne faisait pas partie de l'ancre phosphatidyl-*myo*inositol. Sur le spectre ¹¹P-¹H HMQC-HOHAHA, P2 corrélait avec le proton anomérique d'une unité de β -arabinose terminale.

	BCG	Erdman	H37Rv	H37Ra	M. leprae ^a
(1)	73	71	44	43	0
% de substitution par la come (2)	79	78	64	51	20
Nombre de coiffes par molécule	6,5	10	7	9,4	n.d.

Tableau B5: Toutes les données proviennent de adapté de Nigou, J. et al., 2000, sauf ^a de Khoo, K.H. et al., 1995b. Les % de substitutions ont été calculés par (1) (2-Araf – t-Araf)/2-Araf [Khoo, K.H. et al., 1995b] ou (2) (3,5-Araf – t-Araf)/3,5-Araf [Nigou, J. et al., 1997].

Dernièrement, il a été suggéré qu'une faible proportion des coiffes oligomannosidiques du ManLAM de *M. tuberculosis* seraient substituées par une unité de 5-désoxy-5-méthyl-5thio-pentofuranose (TMP) [Treumann, A. *et al.*, 2002].

Le deuxième type de coiffe, qui est constitué par un motif phosphatidyl-myo-inositol, n'a été observé que chez des mycobactéries à croissance rapide, non pathogènes. Il a d'abord été mis en évidence dans une espèce indéfinie à croissance rapide contaminant une culture de M. tuberculosis H37Ra [Khoo, K. H. et al., 1995b]. La digestion enzymatique de la partie arabinane du LAM isolé de cette culture a généré une famille de d'oligoarabinosides (Ara4 à Ara₈) substitués par du phospho-inositol. L'analyse par méthylation de ces fractions suggéraient que le phospho-inositol était lié en C-5 sur l'arabinose en position terminale nonréductrice. La position exacte de ce substituant n'a pu être déterminée que par une étude RMN complète sur le LAM isolé de M. smegmatis [Gilleron, M. et al., 1997]. En particulier, sur la ³¹P-¹H HMQC-HOHAHA, deux groupements d'expériences hétéronucléaires base phosphatidyl-myo-inositol ont été distingués: l'un substituait le C-1 du glycérol et constituait l'ancre glycosylphosphatidyl-myo-inositol, tandis que l'autre substituait la position C-5 d'une unité ßAra terminale (Fig. B13). La quantification des deux espèces de phospho-diester à montré qu'une moyenne de quatre coiffes de phosphatidyl-inositol était présentes par molécule de LAM chez M. smegmatis.

L'ensemble de ces données ont permis d'envigeager un modèle de structure du LAM schématisé figure B14.



Fig. B14: Modèle de structure du LAM. L'ancre phosphatidyl-myoinositol est présenté avec le nombre maximum d'acides gras.

ē

*myo*inositol

232. Les glycolipides de type tréhalose

Le corps tréhalose est un motif communément retrouvé dans de nombreux types de glycolipides chez les mycobactéries. Quatre familles de composés associés au corps tréhalose ont été identifiés et sont représentées dans la figure B15: les α -D-tréhalose 6-monomycolate (TMM) et 6,6'-dimycolate (TDM), les α -D-tréhalose 2,3-diacyls (DAT), 2,3,6-triacyls (TAT) et 2, 2',3',4,6'-pentaacyls (PAT), les sulfolipides (SL) et les lipooligosaccharides (LOS).



Fig. B15:D'aprés Asselineau, J. & Lanéelle, G., 1998. Exemples des quatre types de glycolipides basés sur un corps tréhalose isolés des mycobactéries.

2321. Les tréhaloses mono- et dimycolates

Ces composés sont substitués par des acides mycoliques. Les TMM et TDM sont systématiquement extraits sous forme de mélanges de composés qui diffèrent par la nature de leurs acides mycoliques et peuvent être isolés par chromatographie en couche mince [Prome,



Fig. B16: Structure de tréhaloses polyacylés isolés de M. tuberculosis. DAT1, Besra, G.S. et al., 1992; TAT, Muñoz, M. et al., 1997; PAT, Daffé, M. et al., 1988.

÷

J, et al., 1976]. Ainsi, six TDMs ont été mis en évidence chez *M. tuberculosis*: les formes symétriques α - α , méthoxy-méthoxy et céto-céto, ainsi que trois formes assymétriques α -méthoxy, α -céto et méthoxy-céto [Strain, S. et al., 1977].

Le TDM a d'abord été identifié comme un composé indispensable à la croissance des cultures de M. tuberculosis sous une forme particulière ressemblant à des cordes et de fait a été appelé "cord factor" [Bloch, H., 1950]. Des études ultérieures ont contredit cette hypothèse, mais ce composé a conservé son appellation. Le TDM semble posséder de nombreuses fonctions biologiques associées à la pathogénicité des mycobactéries. En particulier, il induirait la production de granulomes dans les poumons de souris [Bekierkunst, A., 1968] et aurait des propriétés immunostimulatrices [Bekierkunst, A. et al., 1969] et antitumorales [Bekierkunst, A. et al., 1971]. Par ailleurs, il serait capable d'induire l'apoptose des thymocytes de souris [Ozeki, Y. et al., 1997]. Les mécanismes moléculaires à l'origine de ces activités sont encore largement inconnus et l'ensemble de ces propriétés sont très dépendantes du modèle animal choisi, de la voie d'injection du composé et des caractéristiques de la suspension de TDM [Asselineau, J. & Lanéelle, 1998]. Récemment une étude ont réévalué l'activité liée à la production des granulomes et ont démontré que les TDM induisaient à la fois une réponse non-spécifique par ses propriétés irritantes (immunité innée) et une réponse spécifique dépendante d'interactions avec les cellules T [Yamagami, H. et al., 2001]. De plus il a été démontré que le TDM pouvait promouvoir la néovascularisation du granulome formé en augmentant la production de VEGF (vascular endothelial growth factor) [Sakaguchi, I. et al., 2000]. Une autre étude a récemment montré que le TDM en faibles concentrations, ainsi que les acides mycoliques libres, augmentaient le survie des macrophages in vitro, alors que l'infection par M. tuberculosis ou M. avium induisait leur apoptose [Nuzzo, I. et al., 2002].

Au contraire du TDM, le TMM aurait une fonction dans le transfert d'acide mycolique au travers de la paroi [Belisle, J.T. *et al.*, 1997].

2322. Les tréhaloses polyacylés

Les di-, tri- et pentaacyl-tréhaloses (DAT, TAT et PAT) sont des glycolipides apparentés, substitués par des acides gras et non par des acides mycoliques. Trois exemples de tels glycolipides sont décrits dans la figure B16. Ces composés ont été décrits pour la première fois chez *M. tuberculosis* [Minnikin, M. *et al.*, 1985]. Les structures de six 2,3-di-
acyltréhaloses isolés de cette espèce ont finalement été complètement caractérisés [Besra, G.S. *et al.*, 1992]. Ils se différencient par la nature de leurs acides gras qui sont de trois types: acides saturés à chaînes linéaires C_{16} - C_{19} , acides mycosanoïque C_{21} - C_{25} saturés et acides mycolipanolique C_{24} - C_{28} saturés (tableau B6). Des tréhaloses tri- et pentaacylés, majoritairement substitués par des acides mycosanoïques insaturés en position C-2, ont également été décrits chez *M. tuberculosis* (Fig. B16) [Daffé, M. *et al.*, 1988; Muñoz, M. *et al.*, 1997]. Cette famille de glycolipides n'est pas spécifique de *M. tuberculosis*, mais a également été isolé de l'espèce à croissance rapide non pathogène *M. fortuitum* [Ariza, M.A. *et al.*, 1994].

2323. Les sulfatides

Les sulfatides sont des tréhalose tri- et tétraacylés sulfatés en position C-2', isolés initialement d'une souche virulente de *M. tuberculosis* [Goren, M. B., 1970]. Ils sont substitués par des acides gras polyméthylés: les acides phthiocéraniques et hydroxyphthiocéraniques (tableau B6), ainsi que par des acides stéariques et palmitiques [Goren, M. *et al.*, 1976a]. Les auteurs de ces travaux ont entièrement séquencé cinq sulfatides qui différaient par leur degré d'acylation (SL-I,-I',-II,-II' et -III dans tableau B7).

		Nature des groupements acyls		
Sulfatide	Positions d'acylation	Palmitate/ Stéarate	Phthiocéranate	Hydroxy- phthiocéranate
SL-II	2,4,6,6'	1	0	3
SL-II	2,3,6,6'	1	0	3
SL-I (majeur)	2,3,6,6'	1	1	2
SL-I'	2,3,6,6'	1	2	1
SL-III	2,3,6	1	0	2

Tableau B7: Différences dans les profils d'acylation des cinq familles de sulfatides identifiées chez Mycobacterium tuberculosis, d'aprés Goren, M.B. et al., 1976.

Le composé majeur, SL-I, a été identifié comme un 2,3,6,6'-tetra-O-acyltréhalose 2'sulfate. La recherche de ces composés dans plusieurs espèces de mycobactéries a suggéré que les sulfatides étaient hautement spécifiques à l'espèce *M. tuberculosis* [Dhariwal, K.R. *et al.*, 1984] et que la recherche spécifique de SL-I était une méthode fiable d'identification de cette espèce [Luquin, M. *et al.*, 1992]. De plus, une corrélation significative entre la présence de sulfatides et le degré de virulence chez des cobayes de la souche de *M. tuberculosis* a été établi [Goren, M.B., 1982], ce qui suggérait que ces composés puissent être des facteurs de virulence majeurs de *M. tuberculosis*. Une étude récente montre par marquage métabolique au ³⁵S que le SL-I n'est présent que dans les souches virulentes de *M. tuberculosis* H37Rv et Erdman, et absent de la souche atténuée H37Ra [Rivera-Marrero, C.A. *et al.*, 2002].

En 1976, l'équipe de Goren avait suggéré que les sulfatides inhibaient la fusion du phagosome et du lysosome, et amélioraient ainsi la survie intracellulaire du pathogène [Goren, M.B. *et al.*, 1976b]. Ces résultats ont été remis en cause par Goren dix ans plus tard [Goren, M.B. et al., 1987], mais confirmés par deux des auteurs de l'étude de 1976 en réponse aux dénégations de Goren [Hart, P.D. & Young, M.R., 1988], ce qui a laissé, à notre connaissance, le débat en suspens.

2324. Les lipooligosaccharides (LOS)

Les lipooligosaccharides sont les équivalents multiglycosylés des polyacyl-tréhaloses décrits plus haut. Ils ont été identifiés et étudiés en détail à partir de nombreuses espèces de mycobactéries. Cette famille de composés présente une extraordinaire diversité structurale, et n'a en commun que la présence du corps tréhalose. Ce dernier est substitué par un nombre variable en fonction de l'espèce d'acides gras linéaires et méthylés comparables à ceux des tréhaloses polyacylés et éventuellement de groupements méthyls: Glc(1-1)3,4,6-tri-O-acyl-2-O-Me-Glc chez *M. szulgai* [Hunter, S.H. *et al.*, 1988], 6-O-Me-Glc(1-1)-tri-O-acyl-Glc chez *M. tuberculosis canetti* [Daffé, M. *et al.*, 1991], 6-O-Me-Glc(1-1)-2,3,4,6-tétra-O-acyl-Glc

chez *M. gordonae* [Besra, G.S. *et al.*, 1993], 2-O-acyl-Glc(1-1)-4,6-di-O-acyl-Glc chez *M. gastri* [Gilleron, M. *et al.*, 1995] ou 2-O-acyl-Glc(1-1)-3,4,6-tri-Oacyl-Glc chez *M. fortuitum* [Besra, G.S. *et al.*, 1992b]. Néanmoins, c'est la partie glycannique des LOSs qui fait preuve de la plus grande hétérogénéité structurale. Chaque espèce synthétise une famille distincte de lipooligosaccharides comprenant jusqu'a six composés majeurs



Fig. B17: Structure des monosaccharides nouveaux identifiés dans les LOS de a) *M. gastri*, d'après Gilleron, M. *et al.*, 1994 et b) *M. kansasii*, d'après Hunter, S.W. *et al.*, 1986.

apparentés, présentant différents degrés de substitution d'une même structure de base. Les LOSs de chaque espèce diffèrent de ceux des autres par leur taille, leur composition et leur séquence. Plusieurs monosaccharides très inhabituels ont été identifiés uniquement dans ces composés. C'est par exemple le cas du 3,6-didésoxy-4-C-(1,3-diméthoxy-4,5,6,7-tétrahydroxy-heptyl)- α -xylohexopyranose chez *M. gastri* (Fig. B17a) [Gilleron, M. *et al.*, 1994] ou du 4,6-didésoxy-2-O-méthyl-3-C-méthyl-4-(2-méthoxy propanamido)-L-manno-*hexo*pyranose, appelé par ses découvreurs N-acylkansosamine, chez *M. kansasii* (Fig. B17b) [Hunter, S.W. *et al.*, 1984].

De fait, la structure de la partie oligosaccharidique des LOSs de toutes les espèces de mycobactéries examinées jusqu'à ce jour présente une stricte spécificité d'espèce. Sur la base de la structure des LOS il est ainsi possible de différentier des espèces mycobactériennes autrement très difficilement distinguables telles que *M. kansasii* et *M. gastri* [Gilleron, M. *et al.*, 1995]. L'étude des lipooligosaccharides de l'espèce *M. gordonae* a également démontré l'existence éventuelle de différences structurales importantes entre souches d'une même espèce [Besra, G.S. *et al.*, 1993]. Plusieurs exemples de structures de la partie saccharidique des LOSs sont donnés tableau B8. Comme l'ensemble des études que nous avons citées le suggèrent, les LOSs sont largement distribués dans de nombreuses espèces de mycobactéries. Ce n'est par contre pas le cas de *M. tuberculosis* dont la plupart des souches, exceptée la souche *canetti* [Daffé, M. *et al.*, 1991] en sont dépourvus.

Espèces et références	Séquences glycanniques
M. fortuitum Besra, G.S et al., 1992	Glc(β1-6)T
M. szulgai Hunter, S.W. et al., 1988	2-O-Me-Fuc(α 1-3)Rha(α 1-3)Rha(α 1-3)Glc(α 1-6)T
M. gordonae Besra, G.S. et al., 1993	Rha(a1-2)3-O-Me-Rha(a1-3)[Xyl(b1-2)]Rha(a1-3)Glc(β1-3)Rha(a1-3)T
<i>M. tuberculosis</i> Daffe, M. <i>et al.</i> , 1991	2-O-Me-Fuc(α1- -3)Glc(β1-3)2-O-Me-Rha(α1-3)2-O-Me-Rha(α1-3) Glc(β1-3)4-O- Me-Rha(α1-3)T
Tableau B8: E: mycobactériennes. T représ	xemples de séquences glycanniques de LOSs isolés de diverses espèces sente le corps de tréhalose polyacylé.

Les lipooligosaccharides des mycobactéries sont des composés extrêmement antigéniques dont les propriétés peuvent donc être utilisées à des fins diagnostiques. Par contre peu d'activités biologiques spécifiques leur ont été associées.



Fig B19: Structure du glycosylphénolphtiocérol dimicocérosate majeur de *M. leprae*, PGL-I. D'après Hunter, S.W. *et al.*, 1982.

De manière surprenante, il a été montré que le LOS de *M. phlei* constituait le recepteur de certains bactériophages [Besra, *et al.*,1996].

233. Les glycosylphénolphtiocérol dimicocérosates, ou glycolipides phénoliques (PGL)

Les PGLs sont une famille de glycolipides produits en quantités importantes par de nombreuses espèces de mycobactéries telles que *M. leprae*, *M. kansasii*, *M. bovis*, *M. gastri* ou *M. haemophilum*. Ils ont été identifiés pour la première fois chez *M. leprae* par la découverte de PGL-1 [Hunter, S.W. & Brennan, P., 1981] dont la structure a été élucidée peu de temps après [Hunter, S.W. *et al.*, 1982]. Ce composé représente à lui seul jusqu'à 2% de la masse bactérienne totale. De manière analogue aux LOS, à l'exception de la souche Canetti, *M. tuberculosis* semble dépourvue de PGL.

Les PGLs consistent tous en un noyau lipidique conservé de très grande taille substitué par une chaîne glycannique courte (entre un et quatre monosaccharides) de structure très variable et spécifique à l'espèce, sur le modèle présenté dans la figure B18.



Fig. B18: Structure générale des glycosylphénolphtiocérol dimycocérosates. R=acides mycocérosiques

Le noyau lipidique est typiquement constitué d'un phénolphthiocérol estérifié par deux chaînes d'acides gras polyméthylés, les acides mycocérosiques (tableau B6). A titre d'exemple, le corps lipidique de PGL-1 de *M. leprae* est représenté dans la figure B19. Son phénolphthiocérol est caractérisé par la présence de deux sections non-substituées de 18 et 4 groupes méthyléniques, tandis que les acides gras qui le substituent se composent d'un mélange d'acides 2,4,6,8-tétra-O-méthylmycocérosiques en C₃₀, C₃₂ et C₃₄. Cette partie n'est sujette qu'à des variations mineures d'une espèce à l'autre, principalement dans les longueurs des chaînes alkyles. Par exemple, le phénolphtiocérol majoritairement présent dans les PGL isolés de *M. haemophilum* est plus court de deux méthylènes [Besra, G.S. *et al.*, 1991], tandis que les PGL de *M. bovis* sont constitués d'un phénolphtiocérol identique à celui de *M. leprae*?



Fig. B20: Structure de la partie glycannique des sept PGL isolés de *M. kansasii*. KI, Fournié, J.J, *et al.*, 1987; KII, Rivière, M. *et al.*, 1987; KIV, Gilleron, M. *et al.*, 1990; K5-8, Watanabe, M. *et al.*, 1997. Les monosaccharides et les substituants indiqués en bleu sont à l'origine de la variabilité structurale observée.

ou d'un phénolphthiodolone, qui diffère par la présence d'un groupement cétone en position antépénultième, à la place du groupement méthoxy [Vercellone, A. & Puzo, G., 1989]. Chez cette dernière espèce, les acides gras sont constitués d'un mélange complexe d'acides 2,4,6-tri-O-méthylmycocérosiques en C_{26} , C_{27} et C_{29} , d'acides mycocérosiques non méthylés en C-2 et C-4 et de façon surprenante, d'acides palmitique et stéarique.

Par contre la partie glycannique possède au même titre que les LOSs, une hétérogénéité remarquable. Elle est constituée de un a quatre monosaccharides généralement partiellement méthylés ou acétylés. Le premier GPL dont la structure glycannique ait été élucidée provenait de *M. leprae* et présentait la séquence suivante: 3,6-di-O-Me-Glc $p(\beta 1-4)2,3$ -di-Me-Rha $p(\alpha 1-2)3$ -Me-Rha $p(\alpha 1-$ [Hunter, S.W. *et al.*, 1982]. De très nombreuses variations de ce tri-saccharide ont ensuite été mises en évidence chez les différentes espèces mycobactériennes. Chaque espèce semble en fait synthétiser un mélange de PGLs de structures sensiblement différentes. Ceci a particulièrement été montré chez *M. kansasii*, dont l'étude détaillée a permis de caractériser huit oligosaccharides distincts associés aux PGLs (Fig. B20). Comme dans le cas des LOS, l'étude de la structure des PGLs a permis de mettre en évidence l'existence de monosaccharides particuliers de structure inconnue jusqu'alors, tel que le 2,6-didésoxy-4-O-méthyl-arabino-hexopyranose dans plusieurs PGLs de *M. kansasii* [Fournié, J.J. *et al.*, 1987].

Depuis la découverte des PGLs, plusieurs rôles associés à la pathogénicité des mycobactéries leur ont été attribués. D'une manière générale, ces composés interviendraient dans la survie des mycobactéries pathogènes dans le phagosome en inhibant l'action délétère des radicaux oxygénés (O₂, H₂O₂ et OH) [Neill, M.A. & Klebanoff, S.J., 1988; Chan, J. *et al.*, 1989]. Cette action ne serait néanmoins pas spécifique aux PGLs. Par contre, plusieurs études ont mis en évidence des activités spécifiques à PGL-I qui seraient reliées à la pathogénie de *M. leprae*. En particulier, PGL-I serait directement impliqué dans la suppression sélective de l'immunité cellulaire dirigée contre les antigènes de *M. leprae* [Mehra, V. *et al.*, 1984]. Cette activité serait spécifiquement portée par le trisaccharide de PGL-I, des équivalents partiellement déglycosylés, déméthylés ou des PGLs d'autres espèces mycobactériennes ne présentant aucune activité immuno-suppressive.

a)

M. avium

$$\begin{array}{ccc} Acyl-NH-D-Phe-D-alloThr-D-Ala-Alaninol\\ & & & \\ O\\ & & & \\ O\\ I\\ O\\ I$$

b)

M. fortuitum

c)

M. xenopi

Acyl-NH-L-Ser-L-Ser-Phe-D-allo-Thr-O-Me O O O 1 0 0 1 0 0 1 1 0 1 13-O-Me-6-désoxy-L-Tal Me 2-O-Lau-L-Rha \leftarrow L-Rha

Fig. B21: Comparaison des séquences peptidiques des GPLs isolés de plusieurs espèces de mycobactéries. a) structure générale des GPLs du complexe *M. avium*; b) structure d'un GPL isolé de *M. fortuitum* serovar. perigrinum [Marin, L.M.L. *et al.*, 1991]; structure d'un GPL isolé de *M. xenopi* [Rivière, M. *et al.*, 1993].

Plus récemment il a été montré que PGL-I serait l'un des facteurs majeurs qui déterminent la prédilection de *M. leprae* pour le système nerveux périphérique, en interagissant spécifiquement avec la laminine-2 résente à la surface des cellules de Schwann [Ng, V. et al., 2000]. La dégradation récurrente de la partie glycannique de PGL-I abolit son interaction avec la laminine-2.

Des équivalents non glycosylés et non phénoliques des PGLs, les phthiocérols dimycocérosates (PDIM) ont été mis en évidence dans plusieurs espèces mycobactériennes à croissance lente [Brennan, P. & Nikaido, H., 1995]. Ils ont en particulier été décrits dans la plupart des souches de *M. tuberculosis* qui ne synthétisent pourtant pas de PGLs [Minnikin, D.E. *et al.*, 2002]. Une étude récente a montré que la présence de ces composés dans la paroi mycobactérienne était indispensable à la réplication de *M. tuberculosis* dans les poumons et qu'ils représentaient donc un facteur de virulence majeur dans cette espèce [Cox, J.S. *et al*, 1999].

234. Les glycopeptidolipides (GPL)

Les glycopeptidolipides sont une famille de glycolipides isolés de nombreuses espèces de mycobactéries, mais jamais présente en conjonction avec les PGLs. Ils ont principalement été étudiés dans le complexe M. avium, M. intracellulare, M. scrofulaceum (MAIS) dont ils forment les antigènes de surface majeurs. Ce complexe a été subdivisé en 31 sérotypes différents en fonction de la structure des GPLs présents à leur surface. De fait, les GPLs présentent une structure hautement variable spécifique à chaque sérotype. Dans le complexe MAIS, ils sont invariablement composés d'un corps peptidique réduit en position C-terminale de séquence D-Phe-D-alloThr-D-Ala-L-alaninol, acylé sur la phénylalanine et glycosylé sur les allo-Thréonine et alaninol par un oligo- et monosaccharide respectivement, ainsi que représenté dans la figure B21. Ces composés ont été divisés en deux familles, apolaire et polaire. Les GPLs apolaires sont synthétisés de manière extensive par tous les sérotypes et sont caractérisés par un oligosaccharide qui se réduit au simple 6-deoxy- α -L-Talose. Les GPLs polaires sont caractérisés par l'extension de la chaîne oligosaccharidique, d'abord par une unité invariable de α -L-Rha puis par un oligosaccharide de structure très variable qui va être responsable des différences observées entre sérotypes. La taille maximale observée de l'oligosaccharide total est de six monosaccharides. L'étude de ces composés dans des espèces à croissance rapide a mis en évidence l'existence de divergences dans la structure de base des GPLs en dehors du complexe MAIS. En particulier chez l'espèce *M. fortuitum*, les GPLs différent par la structure des deux parties glycanniques: le 6-désoxytalose est remplacé par une unité de 3-O-Me-Rha toujours terminale (Fig. B21), tandis qu'un disaccharide de rhamnose ou un rhamnose unique, sulfaté en position C-2, substitue le résidu d'alaninol [Lopez Marin, L.M. *et al.*, 1991, Lopez Marin, L.M. *et al.*, 1992]. Enfin, plusieurs études portant sur l'espèce *M. xenopi* ont montré que la structure du peptide et la position relative des parties glycanniques pouvaient également varier en fonction de l'espèce. Comme dans le cas du complexe MAIS, la structure des parties oligosaccharidiques des GPLs *M. xenopi* exhibe une variabilité importante en fonction de la souche (Fig. B22), ce qui suggère l'existence de sérotypes distincts au sein cette espèce.



Fig.B22: Comparaison des structures des oligosaccharides substituants le résidu d'alloThr des GPLs de trois souches de *M. xenopi*. a, souche NCTC 10042 [Rivière, M. *et al.*, 1993]; b, souche ATCC 19250 [Besra, G.S. *et al.*, 1993]; c, souche CIPT 140 [Rivière, M. & Puzo, G., 1991].

24. les polysaccharides libres

Plusieurs espèces de mycobactéries, dont *M. tuberculosis* et *M. bovis* BCG, synthétisent des quantités importantes de polysaccharides libres qui semblent associés à l'enveloppe. Ils consistent essentiellement en des glucanes, des arabinomannanes et des mannanes.

L'existence d'un glucane dans l'enveloppe de *M. tuberculosis* avait été mise en évidence par plusieurs études il y a de nombreuses années [Misaki, A. & Yukawa, S., 1966; Amar-Nacasch, C. & Vilkas, E., 1970].



Fig. B23: Modèle d'arrangement des composants majeurs de *M. tuberculosis*, montrant les interactions possibles des glycolipides (sulfatides, di- et tri-acyl tréhaloses, phthiocerol dimycocérosate) avec la monocouche d'acides mycoliques. La capsule n'est pas représentée à cause du peu d'information disponible sur son organisation. D'après Minnikin, D.E.*et al.*, 2002.

Il a été démontré récemment qu'il présentait une structure similaire au glycogène interne des mycobactéries, mais avait une masse moléculaire apparente beaucoup plus réduite (100 kDa) et exhibait des chaînes latérales de plus petite taille [Ortalo-Magné, A. *et al.*, 1995]. Il consistait en un homopolymère de 4- α -D-Glc, branché en C-6 par une unité mono- ou diglycoside, à hauteur de 16-20%.Les mannanes et arabinomannanes (AM) isolés de l'enveloppe des mycobactéries semblent présenter des structures en tous points identiques aux domaines polysacharidiques des lipomannanes et lipoarabinomannanes, respectivement. Ils ne diffèrent que par l'absence de l'ancre phosphatidyl-*myo*inositol. Les études des AMs isolés de *M. tuberculosis* et *M. bovis* BCG ont démontré que ces derniers sont coiffés par des oligomannosides de façon identique aux LAMs [Lemassu, A. & Daffé, M., 1994; Nigou, J. *et al.*, 1999b]. Il est à noter que l'extraction différentielle de deux fractions d'AM de l'enveloppe de *M. bovis* BCG a également suggéré qu'une partie de l'AM serait plus profondément enfouie dans la paroi et inaccessible au milieu extérieur [Nigou, J. *et al.*, 1999b].

25. Ultrastructure de l'enveloppe

Comme il a été exposé dans l'introduction, et bien que la structure des différents composés de la paroi des mycobactéries soit relativement bien connue, leur arrangement exact au sein de l'enveloppe n'a pas été établie avec certitude. En 1982, Minnikin a proposé un modèle d'arrangement cohérent des lipides associés à la paroi, selon lequel ceux ci s'intercaleraient dans la couche constituée par les acides mycoliques pour former une pseudobicouche lipidique externe [Minnikin, D.E., 1982]. La couche externe serait non seulement formée par les glycolipides spécifiques d'espèce qui ont été décrits au chapitre précédent, mais également par une grande quantité de phospholipides C₁₆₋₁₈ ubiquitaires. Les tréhaloses dimycolates et les différents tréhaloses polyacylés possèdent les caractéristiques idéales pour s'insérer dans une bicouche lipidique et seraient potentiellement intégrés dans cette structure selon le schéma proposé dans la figure B23 [Brennan, P.J. & Nikaido, H., 1995]. Ceci est d'autant plus probable que ces composés sont tous substitués par des acides gras multiméthylés dont la forte hydrophobicité leur permet d'interagir facilement avec la couche très hydrophobe d'acides mycoliques. L'existence d'une seconde couche de composés amphiphiles au dessus de la couche d'acides mycoliques expliquerait l'imperméabilité importante que la paroi des mycobactéries exhibe envers les molécules polaires. L'étude de la diffraction aux rayons X de la paroi purifiée de M. chelonae a montré l'existence de plusieurs zones de fluidité différentes au sein du domaine lipidique qu'explique le modèle de Minnikin: il passe d'un état presque cristallin dans la région la plus profonde à un état désordonné très fluide dans la partie externe [Nikaido, H. *et al.*, 1993]. La très faible fluidité dans la partie interne s'explique par l'arrangement ordonné parallèle des branches α et des méromycolates. En s'éloignant du groupe carboxyle, l'introduction de groupes cyclopropanes, de doubles liaisons et de groupements oxygénés le long des méromycolates est prompte à former des 'coudes' dans les chaînes hydrocarbonées et à augmenter la fluidité de la couche d'acides mycoliques. Enfin, la couche externe de la bicouche présente une fluidité similaire à celle de la membrane plasmique.

Plusieurs observations expérimentales sont en accord avec ce modèle. En particulier, l'existence d'un deuxième plan de cryo-fracture à l'extérieur de la membrane plasmique suggère l'existence d'une bicouche lipidique dans la paroi. Dans ce cas, la fracture s'effectuerait entre la couche d'acides mycoliques liés au complexe mAGP et la couche formée par les lipides extractibles [Brennan, P.J. & Nikaido, H, 1995]. D'autres auteurs interprètent ce résultat autrement et suggèrent plutôt que la fracture s'effectuerait à l'interface entre la paroi et la capsule [Daffé, M. & Draper, P., 1998]. Bien qu'aucune donnée ne permette de remettre en cause le concept général de ce modèle, plusieurs variantes ont été proposées. Dans une de celles-ci, les lipides extractibles ne s'intercaleraient pas entre les acides mycoliques mais formeraient une monocouche au dessus des acides mycoliques [Rastogi, N., 1991]. Considérant les différences de longueurs d'acides gras entre les phospholipides et les glycolipides, et les différences de tailles entre les deux branches des acides mycoliques, il est également possible que les composés les plus longs soient partiellement intercalés entre les méromycolates, contrairement aux phospholipides. Cette dernière hypothèse est d'ailleurs en accord avec l'épaisseur de la couche transparente aux électrons qui correspond à la pseudo bicouche: 70-80 atomes de carbones qui coïncident avec l'addition de la chaîne de méromycolate (C_{60}) et d'un phospholipide (C_{16-18}).

Au dessus de la paroi, une couche d'épaisseur et d'apparence très variable a été récemment reconnue par l'utilisation de plusieurs techniques de fixation et de coloration préservant l'intégrité de cette structure [Rastogi, N. *et al.*, 1986]. Elle est majoritairement composée de protéines et de polysaccharides, en particulier les glucanes et les arabinomannanes, mais contient également des quantités limités de glycolipides variés [Ortalo-Magné, A *et al.*, 1995]. Très peu d'informations sont actuellement disponibles quant à l'existence d'une ultrastucture conservée de la capsule, à l'image de celle observée dans la

paroi. Une étude s'est attachée à définir la répartition des différents glycolipides au sein de la capsule en analysant le matériel libéré par différents traitements mécaniques successifs de plus en plus drastiques [Ortalo-Magné, A. *et al.*, 1996]. Il est apparu que pour la majorité des espèces étudiées, la plupart des glycolipides synthétisés étaient retrouvés associés à la capsule. Par contre, des différences importantes sont apparues quant à la localisation présumée de chaque composé dans la capsule, ce qui suggère que la structure de cette couche externe varie fortement d'une espèce à l'autre. Bien que peu d'informations précises soient disponibles sur sa structure et que son existence semble encore contesté, la relevance biologique d'une telle structure est potentiellement très importante car les éléments qui la composent seraient les plus à même d'être directement impliqués dans les différentes interactions avec les cellules hôtes.

La localisation exacte de certains glycolipides au sein de l'enveloppe des mycobactéries est encore inconnue. C'est en particulier le cas des lipomannanes et des lipoarabinomannanes. L'utilisation d'anticorps monoclonaux dirigés contre le LAM semblaient suggérer qu'au moins la partie terminale de ce dernier était accessible à l'environnement extérieur [Gaylord, et al., 1987], mais la découverte dans la capsule d'arabinomannanes réagissant avec cet anticorps remet en cause ces résultats. Néanmoins, comme nous allons le voir dans le prochain chapitre, de nombreux rôles, en particulier dans les interactions hôtes-pathogènes, ont été dévolus aux LAMs, ce qui suggérait que les LAMs étaient bien en partie accessibles. La présence d'une partie lipidique suggère que ces molécules sont capables de s'intégrer dans une membrane lipidique, et de fait, dans la plupart des modèles publiés, les LMs et LAMs s'intègrent via leur domaine acylglycérols dans la membrane plasmique [Chatterjee, D., 1997]. Dans ce cas, pour être accessible, la partie polysaccharidique très hydrophile du LAM devrait traverser la couche d'acides mycolique très hydrophobe de part en part. Une deuxième hypothèse est que le domaine lipidique du LAM soit inséré dans la couche lipidique externe de la paroi [Chatterjee, D. & Khoo, K.H., 1998]. Dernièrement, l'utilisation de nouveaux protocoles d'extraction a permis de différentier le LAM en deux fractions, pariétale et cellulaire [Nigou, J. et al., 1997]. Il a été proposé que les différences d'extraction de ces deux fractions puissent refléter la localisation différentielle des LAMs, dans la paroi ou dans la membrane [[Vercellone, et al., 1998]. Par contre les études effectuées sur la capsule de différentes mycobactéries n'a jamais révélé l'existence de LAM ou de LM au sein de cette couche, alors que les PIMs étaient systématiquement présents [Ortalo-Magné, A. et al., 1996]. Enfin, le LAM peut être présent en dehors de l'enveloppe des mycobactéries. En particulier, il a été observé qu'après infection des macrophages par *M. avium*, le LAM et d'autres dérivés lipidiques s'intégraient dans des vésicules formées par bourgeonnement de la membrane phagosomique et pouvaient ainsi être dirigés vers les compartiments de la voie d'endocytose et vers la surface cellulaire du macrophage infecté [Xu, S. *et al.*, 1994]. Des vésicules uniformes contenant du LAM ont ainsi été retrouvées dans le milieu de culture de macrophages infectés et permettent l'intégration de LAM dans la membrane de macrophages voisins non-infectés [Strohmeier, G.R. & Fenton, M.J., 1999]. Il est donc probable que le LAM intégré dans ces vésicules puissent exercer certaines de ses fonctions biologique grâce à sa grande accessibilité sous cette forme.

3. Biosynthèse du lipoarabinomannane

Au regard des nombreuses études structurales effectuées sur les lipoarabinomannanes et de l'importance de ces composés dans les pathologies liées aux mycobactéries (voir chapitre 4), peu d'informations sont disponibles quant à leurs modes de biosynthèse. La reconnaissance d'une ancre phosphatidyl-*myo*inositol (PI) commune aux PIM, LM et LAM laisse néanmoins présager une relation biosynthétique directe entre ces trois composés selon le schéma PI \rightarrow PIM \rightarrow LM \rightarrow LAM. La complexité de ces structures présage l'intervention d'un grand nombre de glycosyltransférases spécifiques, dont une minorité seulement a été identifiée. Le schéma de biosynthèse des lipoarabinomannannes à partir des phosphatidylinositol (PI) est représenté dans la figure B24.

Les premiers travaux sur la biosynthèse du PIM effectués sur *M. phlei* ont permis de mettre en évidence l'existence de deux activités mannosyltransférasiques distinctes qui transfèrent à partir du GDP-Man une unité de mannose en position C-2 de l'inositol du Ac₂PI pour former le phosphatidylinositol monomannoside (Ac₂PIM₁), puis une unité de mannose en position C-6 pour former Ac₂PIM₂ [Hill, D.L. & Ballou, C.E., 1966; Brennan, P. & Ballou, C.E., 1967]. Suivant ce modèle, deux acides gras sont ensuite transférés successivement en C-6 de l'unité de mannose attaché en C-2 sur l'inositol, puis en C-3 de l'inositol [Ballou, C.E., 1972]. Dernièrement, les enzymes responsables de l'addition des mannoses en position C-2 et C-6 de l'inositol ont été identifiées et partiellement isolée de *M. tuberculosis*, et nommées PimB et PIMA, respectivement [Schaeffer, M.L. et al., 1999, Kordulakova, J. et al., 2002].



Fig. B24: Schéma de biosynthèse hypothétique des PIMs, LM et LAM de M. tuberculosis, d'après Kremer, L et al., 2002.

L'extension de la chaîne de mannose linéaire en α 1,6 pour former le LM nécessite quant à elle un autre donneur de mannose. En effet, Besra et collaborateurs ont montré que Ac₃PIM₂ était spécifiquement allongé par l'addition de Manp provenant d'un donneur polyprénol-monophosphorylmannose (C35/C50-P-Man) [Besra, G.S. et al., 1997]. Sur la base de leur structure, les PIM₃ à PIM₄ pourraient donc être des intermédiaires entre PIM₂ et le LM, alors que PIM₆ serait un produit terminal divergent car il contient une séquence linéaire de mannoses liés en α 1,2 qui n'est observée ni dans le LM ni dans le LAM. Dernièrement, une voie alterne permettant la synthèse de Ac₃PIM₃ à partir de Ac₃PIM₂ sous l'action d'une nouvelle mannosyltransférase codée par le gène pimC de M. tuberculosis CDC 1515 a été mise en évidence [Kremer, L. et al., 2002]. Au contraire de l'activité décrite plus haut, cette mannosytransférase utilise le GDP-Man comme donneur de monosaccharide. De manière surprenante, ce gène est absent chez M. tuberculosis H37Rv ainsi que dans de nombreux isolats cliniques de M. tuberculosis, suggérant qu'il existe des mécanismes de compensation capables de générer du PIM₃ à partir de PIM₂ et de GDP-Man. Suite à l'élongation de la chaîne de mannose en a1,6 la synthèse du LM mature nécessite l'action d'autres enzymes encore non identifiées pour former les substitutions en α 1,2. De même, les enzymes impliquées dans la synthèse du domaine arabinane, et des coiffes sont encore totalement inconnues.

4. Activités biologiques liées aux lipoarabinomannanes

Depuis leur découverte de très nombreuses activités biologiques ont été associées au lipoarabinomannane. Les plus étudiées sont sans aucun doute celles liées à leur capacités de moduler l'expression de nombreuses cytokines directement impliquées dans la réponse immunitaire de l'hôte aux infection mycobactériennes. Le LAM peut également influer de manière plus directe sur le système immunitaire de l'hôte en régulant les fonctions microbicides des macrophages et en modulant l'apoptose des macrophages qu'ils infectent. Nous allons détailler dans la suite de ce chapitre quelques activités biologiques associées au LAM parmi les mieux étudiées.

41. Modulation de l'expression des cytokines

Le LAM présente une vaste panoplie d'activités immunomodulatrices et semble impliqué dans l'immunopathologie de la lèpre et de la tuberculose. Il est capable d'induire l'expression d'une gamme variée de cytokines associées au macrophage, telles que le TNF-a [Moreno, C. et al., 1989], le GM-CSF, l'IL-1a, l'IL-1b, l'IL-6, l'IL-10 [Barnes, P.F. et al., 1992], l'IL-12 [Yashida, A. & Koide, Y., 1997] et le TGF-ß [Dahl, K.E. et al., 1996]. Certaines de ces cytokines interviennent dans la modulation de l'activation des macrophages et par conséquent dans la survie intracellulaire de la mycobactérie. Ainsi, le TNF-α et GM-CSF, par exemple, sont souvent associés à l'inhibition de la croissance, voire à la destruction de la mycobactérie, tandis que l'IL-6 semblerait promouvoir la croissance du bacille [Denis, M. & Gregg, E.O., 1991]. L'un des aspect les plus stimulants de l'étude des activités immunomodulatrices du LAM consiste en ce qu'il est vite apparu que ces activités étaient très dépendante de l'origine, et par extension de la structure des LAMs utilisés. En particulier, la nature des coiffes semble exercer une influence très importante sur la nature de l'activité biologique du LAM. De fait, un consensus semble émerger selon lequel les PILAMs seraient des molécules proinflammatoires stimulant la production de TNF- α , d'IL-6 et d'IL-12, tandis que les ManLAMs seraient d'une manière générale des molécules antiinflammatoires inhibant la production de TNF- α et IL-12 par les macrophages humains et les cellules dendritiques. Ceci est d'autant plus intéressant qu'un parallèle peut être dressé entre la nature et l'activité des LAMs et la capacité des mycobactéries correspondantes à survivre dans le macrophage. En effet, l'incapacité de M. smegmatis à survivre dans le macrophage activé corrèle avec l'activité proinflammatoire du PILAM, tandis que la capacité de M. tuberculosis et M. bovis de se reproduire dans le macrophage corrèle avec les propriétés antiinflammatoires du ManLAM [Nigou, J. et al., 2002]. Ces différences d'induction de cytokines par les LAMs purifiés se reflètent partiellement dans l'induction de cytokines par les mycobactéries entières. Des études sur un large panel d'espèces ont révélé que les niveaux d'induction du TNF-α et de GM-CSF par les espèces pathogènes (M. tuberculosis, M. avium, M. kansasii et M. xenopi) étaient de deux à huit fois inférieurs à ceux d'espèces non pathogènes (M. smegmatis, M. phlei) [Beltan, E. et al., 2000]. Par contre la différence d'induction de l'IL-6 par le PILAM et le ManLAM ne se retrouve pas au niveau des mycobactéries entières.

Nous décrirons succinctement dans la suite de ce chapitre les phénomènes liés à l'induction de trois des cytokines les plus étudiées, le TNF- α , l'IL-12 et le TGF- β . Nous

constaterons également que d'autres motifs structuraux en dehors des coiffes semblent également intervenir dans l'activation différentielle des cytokines en fonction de l'origine des LAMs.

Le TNF- α

Le TNF- α intervient dans l'activation des macrophages et stimule leurs propriétés antimicrobiennes. Il joue également un rôle actif dans la pathogénie tuberculeuse en participant à la formation de granulomes qui sont capable de contenir les mycobactéries, empêchant ainsi leur dissémination dans l'organisme infecté [Kindler, V. et al., 1989]. Présent en trop grande quantité, le TNF-a peut avoir un effet délétère sur l'organisme en induisant la nécrose des tissus pulmonaires sains, et ainsi participer à la pathologie associée aux mycobactéries. Les premières études effectuées sur le LAM isolé de M. tuberculosis ont démontré que celui ci était un inducteur de la sécrétion de TNF-a par les macrophages humains et murins [Moreno, C. et al., 1989]. Néanmoins, des études ultérieures ont montré que le PILAM isolé des espèces à croissance rapide, dont M. smegmatis, induisaient la production de l'ordre de cent fois plus de TNF- α dans ces mêmes cellules, suggérant que la présence de la coiffe de phospho-inositol était essentielle à cette activité [Dahl, K.E. et al., 1996; Barnes, P.F. et al., 1992.; Gilleron, M. et al., 1997]. D'autres éléments structuraux tel que le profil d'acylation semblent également avoir une influence sur cette activité. En effet, Gilleron et collaborateurs ont isolé deux fractions de PILAM à partir de M. smegmatis, dont l'une était dépourvue d'acides gras et qui était la seule à présenter une activité inductrice de TNF-α [Gilleron, et al., 1997; Vercellone, et al., 1998]. De même, la fraction pariétale du ManLAM de *M. bovis* BCG présente un forte activité inductrice de TNF- α par les cellules dendritiques humaines, comparé à la fraction cellulaire [Nigou, J. et al., 1997]. Comme dans le cas précédent, ces deux fractions ne diffèrent que par la nature de leur partie lipidique (voir chapitre 2312). En revanche, en présence de cellules microgliales humaines, le PILAM et le ManLAM induisent des niveaux comparables de TNF- α , suggérant que l'origine des macrophages constitue également un déterminant important dans la réponse vis-à-vis des différents types de LAMs [Peterson, P.K. et al., 1995].

Les bases moléculaires à l'origine de la modulation des activités inductrices de TNF- α en fonction de la nature de la coiffe ou du statut d'acylation sont à l'heure actuelle inconnues. Néanmoins, il a été montré que l'activité du PILAM était en partie dépendante d'interactions à la fois avec CD14 et des récepteurs de la famille Toll (Toll-like receptors, TLR). CD14 est une glycoprotéine membranaire de 55 kDa exprimée à la surface des monocytes, des macrophages et des cellules microgliales, et qui sert de corécepteur au LPS des bactéries Gram négatives. Le blocage de la sécrétion de TNF-α induite par le PILAM par des anticorps dirigés contre CD14 a fortement suggéré que ce dernier était impliqué dans l'induction [Zhang, Y. et al., 1993]. De plus, le fait que le PILAM puisse directement interagir avec la forme soluble de CD14 est en accord avec cette hypothèse [Pugin, J.D. et al., 1994]. La présence d'une activité résiduelle pour des quantités de LAM élevées, malgré la présence d'anticorps anti-CD14, a de plus suggéré l'existence d'un autre récepteur. Des travaux plus récents ont également démontré que le PILAM activait la voie NF-kB des macrophages via le récepteur TLR-2, mais non TLR-1 ou TLR-4, en synergie avec CD14 alors que le ManLAM en était incapable. Encore une fois, l'origine de ces différences est encore inconnue [Means, T.K. et al., 1999a; Means, T.K et al., 1999b].

Le TGF- β

De nombreux patients atteints de tuberculose pulmonaire montrent une suppression de la réponse immune de type cellulaire pour certains antigènes mycobactériens. Cette immunosuppression semble liée, d'une part, à une diminution de la lymphoprolifération cellulaire et de la production d'IL-2, et d'autre part, à un accroissement de la concentration en TGF- β [Hirsch, C.S. *et al.*, 1996]. Entre autres effets, le TGF- β réduit la production de cytokines inflammatoires, atténue la réponse des cellules T à l'IL-2 et inhibe la production d'IFN- γ impliqué dans l'activation macrophagique. Ainsi, le TGF- β apparaît comme une cytokine immunosupressive. De plus, la production excessive de TGF- β est associée à une importante fibrose et à une destruction tissulaire. Il a été montré que *M. tuberculosis* entier [Hirsch, C.S. *et al.*, 1994], ainsi que les LAMs isolés de diverses espèces de mycobactéries étaient des puissants inducteurs de TGF- β [Dahl, K.E. *et al.*, 1996], ce qui démontrait une fois de plus que le LAM était directement impliqué dans la modulation de la réponse immune de l'hôte.



Fig. B25: D'aprés Nigou, J. et al., 2001. Modèle d'inhibition de la sécrétion d'IL-12 par le ManLAM en fonction de son état d'acylation. MR, récepteur à mannose.

Par contre, aucune différence dans les niveaux d'induction entre les ManLAM isolés de différentes souches de *M. tuberculosis*, le ManLAM de *M. bovis* BCG et le PILAM de mycobactéries à croissance rapide n'a été observé.

L'IL-12

L'interleukine 12 est une cytokine indispensable au développement de la réponse immunitaire aux mycobactéries. Elle est impliquée dans la différenciation des cellules T naïves en Th1 en induisant la production d'IFL-y. Il a en effet été constaté que des souris mutantes IL-12 p40^{-/-} étaient dans l'incapacité de contrôler l'infection par M. tuberculosis, à cause de l'absence d'IFN-y [Cooper, A.M. et al., 1997]. De même, la déficience en récepteur de l'IL-12 augmente dramatiquement la susceptibilité des individus aux mycobactérioses [de Jong, R. et al., 1998]. Alors que le PILAM induit la production de quantités importantes d'IL-12 par les macrophages murins, le ManLAM non seulement n'induit pas sa production dans ces mêmes cellules, mais en inhibe la production par les cellules phagocytaires humaines [Yoshida, A. & Koide, Y., 1997; Knutson, K.L. et al., 1998]. Le ManLAM peut donc dans ce cas précis être considéré comme un facteur de virulence en inhibant partiellement la réponse immunitaire de l'hôte. Il a été récemment suggéré que le ManLAM interagirait avec le récepteur à mannose (RM) des cellules dendritiques humaines pour délivrer un signal inhibant la production d'IL-12 induite par des LPS bactériens [Nigou, J. et al., 2001]. Le récepteur à mannose (MR) est une protéine transmembranaire contenant huit domaines lectiniques (carbohydrate recognition domains, CRD). Seule une reconnaissance multivalente du ligand par le MR permet d'obtenir une reconnaissance de haute affinité. Ainsi, seul le ManLAM, par l'intermédiaire de la coiffe de mannose, et non le PILAM, serait capable d'interagir avec le MR. Comme dans le cas du TNF-a, cette étude a également démontré l'influence de la partie lipidique sur l'activité du LAM. En effet, pour présenter une activité inhibitrice, le ManLAM devait être substitué par au moins deux acides gras, quelque soit leur nature. Les auteurs suggèrent qu'une interaction efficace capable de déclencher un signal nécessite que le ManLAM soit présenté au RM sous forme d'agrégats, ce qui nécessiterait la présence d'au moins deux acides gras (Fig. B25). Ces observations ne rendent néanmoins pas forcement compte de l'activité du LAM au sein de la paroi mycobactérienne. Par contre, elles sont cohérente avec l'existence de vésicules de LAM relarguées par les macrophages infectés (voir chapitre 25) qui pourraient exercer d'une manière générale les fonctions liées à l'induction de la sécrétion des cytokines. On peut noter également que plusieurs études ont établi que l'arabinomannane en solution, et donc jamais agrégé, n'avait aucune influence sur l'induction de cytokine. Si les différentes activités associées au ManLAM requièrent une présentation multivalente des coiffes mannosylées, comme dans le cas du RM, et si la fonction de la partie lipidique soit de permettre cette présentation multivalente en solution, une réévaluation du rôle potentiel de l'arabinomannane *in vivo* est nécessaire. En effet, si ce dernier est présent en quantité importante dans la couche la plus externe de l'enveloppe des mycobactéries, il est probable qu'il se présente sous forme multimérique, ce qui en fait un autre candidat pour exercer certaines des fonctions jusqu'alors uniquement dévolues au LAM. Cette hypothèse n'est cependant consistante qu'en admettant que l'AM présent à la surface des mycobactéries puisse présenter une multivalence et une présentation comparables à celles des agrégats de LAM soluble.

Il est également possible que la partie lipidique puisse exercer son influence au travers d'interactions directes avec le récepteur. Il a par exemple été montré que la fixation du ManLAM sur la molécule présentatrice d'antigène CD1b nécessitait des interactions hydrophobes avec la partie lipidique du ManLAM [Ernst, W.A. *et al.*, 1998]. Bien que ce type d'interactions soit tout à fait envisageable sur des LAMs intégrées dans une vésicule soluble, les modèles actuels d'insertion du LAM au sein de la paroi des mycobactéries les rendent *a priori* improbables au vue du peu d'accessibilité présumée à l'ancre lipidique.

42. Phagocytose des mycobactéries

Le récepteur à mannose interviendrait également dans la phagocytose des mycobactéries, toujours par l'intermédiaire du ManLAM. En effet, il a été démontré que des microsphères enrobées de ManLAM, et non celles enrobées de PILAM, étaient phagocytés par des macrophages *via* le récepteur à mannose [Schlesinger, L. *et al.*, 1994]. L'observation que des microsphères enrobées de ManLAM dont l'extrémité mannosylée avait été digérée par une exo-mannosidase n'étaient pas internalisées a confirmé que le déterminant reconnu par le RM était bien la coiffe de mannose. L'intervention spécifique du RM a été démontrée en inhibant l'interaction par des anticorps anti-RM et des mannanes de *Sacharomyces*. Néanmoins, il existe des différences importantes dans les capacités d'adhésion de ManLAMs issus de souches de *M. tuberculosis* distinctes [Schlesinger, L. *et al.*, 1996]. Ainsi, les ManLAMs des souches Erdman, H37Rv et H37Ra, respectivement, présentent des capacités

d'adhésion décroissantes. Ces différences ne corrèlent pas parfaitement avec les différences des pourcentage de coiffage des ManLAM utilisés, et pas du tout avec le nombre de coiffes d'oligomannosides par molécule (voir tableau B5) ce qui suggère l'intervention d'autres déterminants structuraux dans l'interaction ou de différences de structure tertiaire pouvant influer sur la présentation des déterminants mannosylés au MR.

Le LAM pourrait également moduler la phagocytose des mycobactéries en interagissant avec des protéines du surfactant pulmonaire de type collectine, la SP-A et la SP-D. Les collectines sont des protéines mutimériques possédant chacune plusieurs domaines lectiniques, dix-huit pour la SP-A et douze pour la SP-D. Il a été démontré le SP-A et SP-D se fixaient de façon spécifique à M. tuberculosis mais auraient des effets antagonistes [Gaynor, C.D. et al., 1995; Ferguson, J.S. et al., 1999]. Alors que la fixation au SP-A augmente la phagocytose de M. tuberculosis par les macrophages alvéolaires, la fixation au SP-D la réduit. Des tests d'interaction entre la SP-A et des mycobactéries entières ont révélé que la SP-A ne discriminait pas les espèces pathogènes telle que M. bovis BCG des espèces non pathogènes telle que M. smegmatis [Sidobre, S. et al., 2000]. Par contre, le criblage de ses ligands potentiels parmi les différents glycoconjugués isolés des mycobactéries a permit d'identifier le ManLAM et le LM comme ligands préférentiels, mais non le PILAM, ce qui suggère l'implication d'autres ligands que le LAM, au moins chez M. smegmatis. Tout comme dans la cas du récepteur à mannose, ces études ont montré que la SP-A reconnaissait spécifiquement la coiffe mannosylée du ManLAM, ainsi que le domaine mannane du LM et ont mis en évidence une influence de la partie lipidique. Récemment, l'étude des bases moléculaires de l'interaction entre SP-A et LAM par résonance plasmonique de surface à révélé que la partie lipidique du LAM n'était pas directement reconnue par la SP-A, mais serait par contre nécessaire à la présentation multivalente de la partie mannane du ManLAM [Sidobre, S. et al., 2002].

La fixation du LAM sur la SP-D s'effectuerait selon des modalités similaires à celles décrites pour le SP-A, via la coiffe mannosylée [Ferguson, J.S. et al., 1999]. En effet, comme précédemment, les PILAMs isolés de *M. smegmatis* ou de mycobactéries à croissance rapide ne se fixent pas sur la SP-D. Par contre, au contraire de la SP-A, la SP-D ne se fixe pas sur *M. smegmatis*, et semble donc discriminer les espèces virulentes des avirulentes.

43. Chimiotactisme

Une activité chimiotactique du LAM, encore une fois variable en fonction de son origine, a été mis en évidence. En effet, il apparaît que le ManLAM de *M. tuberculosis* Erdman et le PILAM isolé de mycobactéries à croissance rapide induisaient la migration de lymphocytes T humains *in vitro* [Berman, J.S. *et al.*, 1996]. De manière surprenante, le ManLAM isolé de *M. bovis* BCG n'a pas démontré une telle capacité, suggérant que des différences structurales mineures, peut être l'état d'acylation, pouvaient engendrer des différences d'activité biologique et moduler la réponse immunitaire de l'hôte. Une seconde étude dans des conditions similaires a montré que le même PILAM induisait également le chimiotactisme des monocytes humains, contrairement au ManLAM de *M. tuberculosis* [Bernardo, J. *et al.*, 1998]. Malheureusement, cette activité n'a pas été définie pour le ManLAM de *M. bovis* BCG. Les auteurs de cette étude suggèrent que l'activation différentielle engendrée par le PILAM et le ManLAM résulterait d'une adaptation de *M. tuberculosis* lui permettant de minimiser le recrutement des cellules phagocytaires additionnelles sur le site d'infection et de favoriser sa persistance au sein des macrophages. Néanmoins la portée exacte de ces phénomènes *in vivo* est encore mal comprise.

II PROBLEMATIQUE

Les lipoarabinomannannes 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 a été 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. Ainsi que nous les avons présentées dans les chapitres précédents, 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. 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 est rendue délicate à cause de deux facteurs principaux: l'hétérogénéité des préparations de LAM et le nombre très 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 exhibaient des activités biologiques distinctes [Nigou, J. *et al*, 1997; Gilleron, M. *et al*, 1997; Gilleron, M. *et al*., 2000]. 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 [Nigou, J. *et al.*, 2001].

Une deuxième stratégie d'étude, complémentaire de la première, consiste à rechercher des LAMs présentant des déterminants structuraux différents de ceux utilisés jusqu'à présent. En effet, seules deux classes majeures de LAMs, qui différent par la nature de leurs coiffes, ont été décrits: 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 devrait fournir de nouveaux outils pour décrypter les fonctions de cette famille de molécules, et éventuellement mettre en évidence de nouvelles fonction biologiques. L'étude structurale et fonctionnelle de LAMs isolés de mycobactéries atypiques permettra également de fournir des informations quant aux mécanismes liés à la pathogenèse des espèces de mycobactéries atypiques.

Les travaux présentés ont d'abord consisté en l'étude structurale préliminaire des lipomannannes et liporabinomannannes extraits de deux espèces de mycobactéries: *M. chelonae* et *M. kansasii. 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 (voir chapitre 1) et nous est apparu comme un modèle intéressant par sa relevance clinique potentielle. Cet aspect a été renforcé par le fait que nous avons procédé à l'étude d'un isolat clinique de cette espèce provenant d'un patient atteint du SIDA de l'hôpital de New York. Il a en effet été montré récemment que le LAM purifié d'un isolat clinique de *M. tuberculosis* présentait des variations structurales comparées aux souches de laboratoire [Khoo, K.H. *et al.*, 2001]. 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.

Dans un deuxième temps, en collaboration avec le Dr E. Elass de l'Institut de Glycobiologie Structurale et Fonctionnelle de Villeneuve d'Ascq, les propriétés immunologiques des différents lipoglycannes que nous avions isolés et caractérisés ont été comparées avec celles des lipoglycannes de souches déjà étudiées. Pour ce faire, leur pouvoir inducteur de cytokines pro-inflammatoires a été testé sur des lignées monocytaires humaines THP-1. De manière surprenante, ces études ont permis de réévaluer l'importance respective des LAMs et des LMs dans l'induction d'une réponse inflammatoire.

III RESULTATS

1. Etude structurale des lipomannanne et lipoarabinomannanne isolés de *M. chelonae* de *M. kansasii*.

11. Présentation des résultats

Ce travail a consisté en l'étude préliminaire de la structure du LM et du LAM isolés de *M. chelonae* et *de M. kansasii*. Les lipoglycannes ont d'abord été purifiés des extraits mycobactériens par deux cycles d'extraction biphasiques au triton X-114 et au phénol aqueux. Ensuite, les LM et LAM ont été séparés des autres composés par gel filtration en milieu chaotropique. L'ancre lipidique et le domaine mannanne ont été décrits successivement à partir du LM qui présente une structure plus simple, puis du LAM dont la partie arabinane a ensuite été étudiée. Les analyses structurales ont nécessité l'utilisation de diverses techniques de résonance magnétique nucléaire homo- et hétéronucléaires (COSYs, TOCSY, ¹H-¹³C et ¹H-³¹P HMQC, HMBC, ROESY) et de spectrométrie de masse (EI-MS, CI-MS, MALDI). Certains des motifs structuraux mis en évidence par des méthodes physico-chimiques ont été confirmés par l'utilisation de lectines.

Les résultats issus de l'étude des lipoglycannes de *M. chelonae* sont présentés sous forme d'un article publié (Article 4). L'étude de l'espèce *M. kansasii* est encore en cours, mais les résultats partiels sont néanmoins présentés chapitre III 13. Les méthodes utilisées lors de cette dernière étude étant très similaires à celles figurant sur l'article 4, nous avons omis de les refaire figurer.

12. Etude des lipoglycannes de *M. chelonae* Résultats: Article 4

Structural Study of Lipomannan and Lipoarabinomannan from *Mycobacterium chelonae*

PRESENCE OF UNUSUAL COMPONENTS WITH α1,3-MANNOPYRANOSE SIDE CHAINS*

Received for publication, May 6, 2002, and in revised form, June 12, 2002 Published, JBC Papers in Press, June 12, 2002, DOI 10.1074/jbc.M204398200

Yann Guérardel‡, Emmanuel Maes‡, Elisabeth Elass‡, Yves Leroy‡, Philippe Timmerman‡, Gurdyal S. Besra§¶, Camille Locht, Gérard Strecker‡, and Laurent Kremer||**

From the ‡Laboratoire de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq Cedex, France, §School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT United Kingdom, and µLaboratoire des Mécanismes Moléculaires de la Pathogénie Microbienne, INSERM U447, Institut Pasteur de Lille/IBL, 1 Rue du Pr. Calmette, BP245-59019 Lille Cedex, France

Lipomannan (LM) and lipoarabinomannan (LAM) are major glycolipids present in the mycobacterial cell wall that are able to modulate the host immune response. In this study, we have undertaken the structural determination of these important modulins in Mycobacterium chelonae, a fast growing pathogenic mycobacterial species. One-dimensional and two-dimensional NMR spectra were used to demonstrate that LM and LAM from M. chelonae, designated CheLM and CheLAM, respectively, possess structures that differ from the ones reported earlier in other mycobacterial species. Analysis by gas chromatography/mass spectrometry of the phosphatidyl-myo-inositol anchor, which is thought to play a role in the biological functions of these lipoglycans, pointed to a high degree of heterogeneity based on numerous combinations of acyl groups on the C-1 and C-2 positions of the glycerol moiety. Characterization of the mannan core of CheLM and CheLAM revealed the presence of novel α 1,3-mannopyranosyl side chains. This motif, which reacted specifically with the lectin from Galanthus nivalis, was found to be unique among a panel of nine mycobacterial species. Then, CheLM and CheLAM were found to be devoid of both the mannooligosaccharide cap present in Mycobacterium tuberculosis and the inositol phosphate cap present in Mycobacterium smegmatis and other fast growing species. Tumor necrosis factor- α and interleukin-8 production were assessed from human macrophages with LAM preparations from different species. Our results suggest that the inositol phosphate capping may represent the major cytokineinducing component of LAMs. This work not only underlines the diversity of LAM structures among various mycobacterial species but also provides new structures that could be useful to dissect the structure-function relationships of these complex molecules.

Mycobacterium species are responsible for important human diseases including tuberculosis and leprosy. Infection and im-

This paper is available on line at http://www.jbc.org

munopathogenesis of these diseases widely implicate the mycobacterial cell wall (1) which is abundantly composed of mannoconjugates, notably polysaccharides and lipoglycans. The latter consist mainly of phosphatidyl-myo-inositol mannosides (PIMs),¹ lipomannan (LM), and the structurally related lipoarabinomannan (LAM). LAM is a major cell wall component and is considered as a modulin through its various immunoregulatory and anti-inflammatory effects, which favor the survival of the mycobacteria within the infected host. These effects include suppression of T lymphocyte proliferation through interference with antigen processing (2), inhibition of macrophage activation by interferon- γ (3, 4), and scavenging of oxygen-derived free radicals (5). LAM is not only a virulence factor responsible for macrophage deactivation, but it is also implicated in phagocytosis of mycobacteria into phagocytic cells (6). In addition, PIMs that are believed to be precursors of LM and LAM have recently been proposed to recruit NK T cells, which play a primary role in the granulomatous response (7, 8).

The biosynthetic relationship of phosphatidylinositol (PI), PIMs, LM, and LAM has recently been supported by biochemical (9, 10) and genetic studies (11, 12), but the details of this pathway remain highly speculative. However, the structures of LAM from several species including Mycobacterium tuberculosis, Mycobacterium leprae, Mycobacterium bovis BCG, and Mycobacterium smegmatis have been extensively described during the last decade. LAM is a complex glycolipid composed of Dmannan and D-arabinan attached to a PI moiety that anchors the glycolipid in the mycobacterial cell wall (13). The biosynthesis of LAM involves the addition of mannopyranosyl (Manp) residues to PI to produce both the short PIMs (2-5 Man residues) and LM, which is further glycosylated with arabinan to form LAM (9, 13-15). In all the species described so far, Dmannan consists of a highly branched structure with an α 1,6linked Manp backbone substituted at C-2 by single Manp units

^{*} This work was supported in part by the CNRS (UMR 8576, Glycobiologie Structurale et Fonctionnelle, Director Dr. Jean-Claude Michalski), the Ministère de l'Enseignement Supérieur et de la Recherche, and by INSERM. 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.

[¶] Lister Institute-Jenner Research Fellow and supported by the Wellcome Trust and the Medical Research Council.

^{**} To whom correspondence should be addressed. Tel.: 33-3-20-87-11-54; Fax: 33-3-20-87-11-58; E-mail: laurent.kremer@ibl.fr.

¹ The abbreviations used are: PIMs, phosphatidyl-myo-inositol mannosides; Araf, arabinofuranosyl; ConA, lectin from *Canavalia ensifor*mis; COSY, correlation spectroscopy; Cl, chemical ionization; El, electron impact; GC/MS, gas chromatography/mass spectrometry; Gro, glycerol; GNA, lectin from *Galanthus nivalis*; HMBC, Heteronuclear multiple bound correlation; HMQC, heteronuclear multiple quantum coherence; Ins, inositol; LM, lipomannan; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; LAM, lipoarabinomannan; MALDI MS, matrix-assisted laser-desorption ionization mass spectrometry; Manp, mannoyranosyl; NOESY, nuclear Overhauser enhancement spectroscopy; PI, phosphatidyl-myo-inositol; ROESY, rotating frame Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; TMS, trimethylsilyl; DIG, digoxigenin; IL, interleukin; TNF, tumor necrosis factor; GPI, glycosylphosphatidylinositol; PBS, phosphate-buffered saline.

30636

Lipoarabinomannan Structure from M. chelonae

(16). The mannan size and the degree of branching can vary depending on the species. The arabinan consists of a linear α 1,5-linked arabinofuranosyl (Araf) backbone punctuated by branching produced with 3,5-O-linked α -D-Araf residue. The lateral chains are organized either as linear tetra-arabinofuranosides β -D-Araf-(1,2)- α -D-Araf-(1,5)- α -D-Araf-(1, 5)- α -D-Araf or as a biantennary hexa-arabinofuranosides [β -D-Araf-(1,2)- α -D-Araf-(1-]₂-3 and 5)-α-D-Araf-(1,5)-α-D-Araf (17, 18). Comparative analyses of LAMs from different mycobacterial species have shown that the non-reducing termini of the arabinosyl side chains are differentially modified. M. tuberculosis and M. *leprae* modify the termini with Manp residues, thereby yielding "ManLAM," whereas the rapidly growing species M. smegmatis uses inositol phosphate, generating "AraLAM" (19). It is thought that these modifications are responsible for the marked differences in the biological activities of ManLAM and AraLAM (15, 19, 20). However, as mentioned above, all available LAM structures are derived from a very limited panel of mycobacterial species, and whether these structures are invariably present in most species remains to be investigated. Awareness that subtle differences among LAMs may affect their biological properties prompted us to establish the structure of LM/LAM from Mycobacterium chelonae, a rapidly growing pathogenic mycobacterium that is found in soil and fresh water throughout the world (21). M. chelonae infection typically causes localized skin lesions, often following penetrating trauma or injections. Disseminated disease usually occurs in patients with a significant immune compromised state that is most commonly attributable to exogenous steroid use (22, 23).

We report here the detailed structure of LM/LAM from *M. chelonae* and provide evidence for important differences such as the acylation composition of the PI, branching of the mannan core, and the absence of Manp and inositol phosphate caps. Because these structures were found to be unique among a panel of various mycobacterial species, we propose to designate these components as CheLM and CheLAM.

EXPERIMENTAL PROCEDURES

Strain and Culture Conditions—All mycobacterial species used were grown on plates containing Middlebrook 7H11 agar supplemented with 10% oleic acid-albumin-dextrose-catalase enrichment (Difco) or in liquid Sauton medium. Except for *M. chelonae* (ATCC 19536), which was grown under shaking in Sauton medium at 30 °C for several days, all other species were grown at 37 °C. Confirmation of the identity of the *M. chelonae* strain was done by analyzing its mycolic acid profile, which is rather unusual because it consists of 60% of *a*-mycolates and 40% of α' -mycolates (24, 25).

Purification of CheLAM and CheLM-Extraction of CheLM and CheLAM was adapted from Nigou et al. (26) based on the Triton X-114 phase partitioning. Briefly, cells were harvested, washed in PBS (20 mm K₂HPO₄ (pH 7.5), 0.15 M NaCl), and resuspended in lysis buffer (8% (v/v) Triton X-114 in PBS, 5 mM EDTA, 10 mM MgCl₂). Cells were then heat-inactivated, disrupted using a French pressure cell and stirred overnight at 4 °C. Cellular debris were removed by centrifugation $(27,000 \times g, 30 \text{ min}, 4 \text{ °C})$, and phase separation was induced at 37 °C. Lipoglycans present in the lower phase were precipitated by adding 5 volumes of cold ethanol and collected by centrifugation $(27,000 \times g, 30)$ min, 4 °C). The pellet was dissolved in water, and proteinase K was added to a final concentration of 10 µg/ml for 20 min at 55 °C. Proteins were extracted twice by adding saturated phenol. Combined aqueous phases containing lipoglycans were dialyzed for 72 h against water, lyophilized, and resuspended in Tris deoxycholate buffer (10 mm Tris HCl (pH 8.0), 10 mM EDTA, 0.2 M NaCl, 0.25% deoxycholate). CheLAM and CheLM were then separated by gel filtration on a Sephacryl S-200 (Amersham Biosciences) column (80×1.5 cm) in the same buffer. The eluted fractions were monitored by 13% SDS-PAGE stained for carbohydrates according to Tsai and Frasch (27). The appropriate LAM and LM fractions were pooled and dialyzed for 48 h against 10 mM Tris-HCl (pH 8.0) and then for 48 h against water prior to lyophilization. The endotoxin content of all reagents was measured in a chromogenic Limulus lysate assay (BioWhittaker). The LAM preparations contained insignificant amounts of endotoxin (<25 pg/ml)

NMR Analysis—Prior to NMR spectroscopic analysis, LM (15 mg) and LAM (5 mg) were repeatedly exchanged in ${}^{2}\text{H}_{2}\text{O}$ (99.97% purity, Euriso-top, CEA Saclay, France) with intermediate freeze-drying and then dissolved in 250 μ l of Me₂SO-d₆ (Euriso-top). Chemical shifts were expressed in ppm downfield from the signal of the methyl group of Me₂SO-d₆ ($\delta^{1}\text{H/TMS} = 2.52 \text{ ppm}, \delta^{13}\text{C/TMS} = 40.98 \text{ ppm at 343 K}$). The samples were analyzed in 200 × 5 mm BMS-005-B Shigemi@ tubes on a Bruker ASX-400 spectrometer (Centre d'Analyses RMN, Villeneuve d'Ascq) (¹H, 400.33; ¹³C, 100.66, ³¹P, 162.5 MHz) equipped with a double resonance (¹H/X) Broad Band Inverse z-gradient probe head. All NMR data were recorded without sample spinning.

The one-dimensional proton ¹H spectrum was measured using 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 16,384 complex data points. The one-dimensional ¹³C was recorded using a spectral width of 20,161 Hz, and 32,768 data points were collected to obtain a free induction decay resolution of 0.6 Hz per point. The ³¹P spectra of both compounds were acquired with a spectral width of 16,233 Hz collected in 16,384 data points. Both experiments were recorded using a composite pulse decoupling during acquisition using globally optimized alternating phase sequence at the carbon or phosphorus frequency (28). An exponential transformation (line broadening factor = 5 for ¹³C and 3 Hz for ³¹P) was applied prior to processing data points in the frequency domain.

Two-dimensional homonuclear (1H-1H) spectra (COSY-ROESY-TOCSY) were measured using standard Bruker pulse programs. ROESY spectra were acquired with various mixing times (50, 100, 200, and 400 ms) and acquired in States mode according to Bax and Davis (29), whereas both COSY and relayed COSY were acquired in the magnitude calculation mode. 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 35 μ s. The spectral width was 4000 Hz in both dimensions. 512 spectra of 4096 data points with 32 scans per t1 increment were recorded giving a spectral resolution of 0.9 Hz/point in F2 and \sim 8 Hz/point in F1. Heteronuclear experiments (¹H.¹³C and ¹H.³¹P) were obtained with standard Bruker pulse sequences such as HMQC (inv4tp), HMQC-HOHAHA (inv4mltp), and HMBC (inv4lrnd). HMQC and HMQC-HOHAHA were acquired in the phase-sensitive increment time proportional phase increment (TPPI) method, whereas HMBC was recorded in magnitude mode calculation. All parameters (pulse widths, pulse powers, and delays) were optimized for each experiment. Acquisitions and processing conditions are expressed in the figure legends.

Gas Chromatography Techniques—Monosaccharides were analyzed as alditol-acetate derivatives. Lipoglycans were hydrolyzed in 4 N trifluoroacetic acid for 4 h at 100 °C and reduced with NaBH₄ in 0.05 N NH₄OH for 4 h. Reduction was stopped by dropwise addition of acetic acid until the pH reached 6, and borate salts were co-distilled by repetitive evaporation in dry methanol. Per-acetylation was performed in acetic anhydride at 100 °C for 2 h, and derivatives were analyzed in GC on a BPX70 12 m × 0.22 mm ID column (Chrompak).

For exact quantification of glycerol and myo-inositol, 100 μ g of lipoglycans along with 1 μ g of scyllo-inositol, as an internal standard, were hydrolyzed with 6 N hydrochloric acid constant boiling (Pierce), at 110 °C for 24 h, and analyzed as TMS derivatives (30) using GC on a DB-1 60 m ×0.25 mm inner diameter by on-column injection.

Linkage analyses of monosaccharides was achieved by two steps of per-methylation according to Ciucanu and Kerek (31) and followed by derivatization with acetyl groups with acetic anhydride.

Acylglycerols were analyzed according to Nigou *et al.* (26) after cleavage of the phosphodiester bond by acetolysis. Briefly, 100 μ g of lipoglycans were treated with 400 μ l of anhydrous acetic acid/acetic anhydride, 3:2 (v/v), at 110 °C for 12 h. Acetylated acylglycerols were extracted by cyclohexane and analyzed by GC/MS on a WCOT fused silica 30 m × 0.25 mm inner diameter column (Chrompak).

Fatty acids were analyzed from intact lipoglycans as well as from extracted acetylated acylglycerol as pyrrolidine derivatives. They were released by methyl esterification with 0.5 M HCl in anhydrous methanol at 80 °C for 20 h, extracted with heptane, and derivatized with 200 μ l of pyrrolidine/acetic acid, 9:1 (v/v), at 80 °C for 2 h. Pyrrolidine-derivatized fatty acids were repetitively extracted by CHCl₉/H₃O, 1:1 (v/v), and analyzed by GC/MS on a WCOT fused silica 30 m \times 0.25 mm inner diameter column (Chrompak).

MALDI-Time-of-Flight Mass Spectrometry—The molecular mass of the lipoglycans was measured by matrix-assisted laser desorption ionization on a Vision 2000 time-of-flight instrument (Finnigan Mat) equipped with a 337-nm UV laser. Samples were dissolved in water at a concentration of 100 pmol/µl. One µl of the solution was mixed with

Lipoarabinomannan Structure from M. chelonae

30637

FIG. 1. Purification of CheLM and CheLAM. a, gel filtration elution profile of the Triton X-114 fraction containing lipoglycans. The fraction was loaded onto a Sephacryl S-200 column and eluted with Tris deoxycholate buffer. Elution of CheLAM, CheLM, and ChePIMs was monitored by SDS-PAGE, and pure fractions were pooled. b, 13% SDS-PAGE analysis of pooled mannoconjugate samples from M. chelonae: lane 1, total polysaccharide fraction before S-200 fractionation; lane 2, pooled CheLAM fractions; lane 3, pooled CheLM fractions. The gel was treated with periodic acid and stained with silver nitrate.



an equal volume of 2,5-dihydroxybenzoic acid (10 mg/ml; dissolved in water/methanol, 4:1 (v/v)) matrix solution on the target and then allowed to crystallize at room temperature.

Western Blot Analysis-Western blot analyses were conducted either on purified LMs and LAMs or on crude mycobacterial lysates. For purified LM and LAM samples, the sugar content was estimated by GC, and the equivalent of 0.5 μ g of mannose for each sample was analyzed by 13% SDS-PAGE. For crude lysates, mycobacterial cells were harvested, resuspended in 0.8 ml of PBS, and disrupted for 10 min with a Branson Sonifier 450. Protein concentrations were determined using the BCA Protein Assay Reagent kit (Pierce). Equal amounts of proteins (30 µg) were then separated by 13% SDS-PAGE and then transferred onto a Hybond-C Extra membrane (Amersham Biosciences). Membranes were then saturated with 5% bovine serum albumin in PBS, 0.1% Tween 20, and probed overnight with either ConA-DIG or GNA-DIG (Roche Molecular Biochemicals, dilution 1:1000). After washing, membranes were subsequently incubated with anti-DIG antibodies conjugated to alkaline phosphatase (Roche Molecular Biochemicals, 1:1000 dilution).

Cytokine Production-The human promonocytic THP-1 cell line was grown in RPMI 1640 (Invitrogen) supplemented with 10% fetal calf serum, 2 mM L-glutamine, and 2 10^{-5} M β -mercaptoethanol in an atmosphere of 5% CO2 at 37 °C. THP-1 cells were induced to express CD14 by treatment with 50 nM 1,25-dihydroxyvitamin D₈ (Calbiochem) for 48 h. Cells were then washed twice with RPMI 1640 and cultured in 96-well plastic culture plates at a density of 2×10^5 cells/well in RPMI 1640 supplemented with 10% fetal calf serum and glutamine. The purified molecules LAMs from M. Chelonae, M. tuberculosis Erdman, and M. smegmatis were added at a final concentration of 10 μ g/ml, and triplicates were performed in order to measure cytokine release. Culture supernatants were collected after 6 or 24 h for TNF- α or IL-8 production, respectively. Specific enzyme-linked immunosorbent assays commercial kits were used according the manufacturer's instructions. Human IL-8 and TNF- α kits were purchased from Bender Med systems Diagnostic and R & D Systems, respectively. Cytokine production was quantified with a microtiter plate reader in comparison with a standard curve generated with recombinant human cytokines.

RESULTS

Purification of CheLAM and CheLM

The experimental protocols used to extract LM and LAM from *M. chelonae* are based on successive detergent and phenol extractions, leading to the recovery of nucleic acid-, protein-, and lipid-free materials. Purity of the preparation was assessed by GC/MS and SDS-PAGE. CheLAM and CheLM were finally resolved on a gel permeation S200 column. Fig. 1*a* represents the elution profile of each compound and shows that CheLM was present in higher amounts than CheLAM. This was confirmed by SDS-PAGE analysis of the total Triton X-114 extract (Fig. 1*b*). The CheLM/CheLAM ratio (w/w) was also determined by routine monosaccharide analysis and revealed that CheLM was three times as abundant as CheLAM. This result was found for several independent extractions (data not shown). It is noteworthy that the relative abundance of these two compounds largely differs from the ones reported previously for other mycobacterial species, in which LAM represents the major component. For instance, an approximate LM/LAM ratio (w/w) of 1:2.5 was found in *M. bovis* BCG (26) compared with the 3:1 ratio observed in *M. chelonae*.

Structural Analysis

Because LM is thought to be a biosynthetic precursor of LAM, we first undertook the structural elucidation of LM, and we subsequently determined the structure of the more complex LAM molecule.

Structure of CheLM—The molecular weight of CheLM was investigated by MALDI analysis. It showed a broad unresolved peak centered at 6900 Da (data not shown). Quantitative analysis of the alditol-acetate derivatives of CheLM led to an average composition of 35 mannose, 1 arabinose, and 1 myo-inositol residues. The exact quantification of TMS derivatives by GC using on-column injection demonstrated the presence of 1.07 unit of glycerol per molecule of myo-inositol, which establishes the ratio of mannose/arabinose/inositol/glycerol as 35:1:1:1.

Analysis of partially methylated alditol-acetate derivatives revealed the presence of three major components (Table I) identified on the basis of their retention times and fragmentation patterns as t-Manp, 6-Manp, and 3,6-Manp. Another product, different from a mannitol-acetate, was characterized by $(M + NH_4)^+$ at m/z 338, indicative of a di-acetylated, tetramethylated inositol. EI/MS fragments at m/z 200, 191, and 75 identified this compound as a 2,6-Ac₂-1,3,4,5-Me₄-Ins, which is in agreement with an earlier published report (16). An (M + NH₄)⁺ ion at m/z 366 corresponding to a tri-acetylated, trimethylated inositol was detected as a minor component. However, because its fragmentation pattern was unclear, the detailed structure of this product was not analyzed further.

In order to establish the main features of both the polysaccharide and the putative GPI anchor of LM from *M. chelonae*, an exhaustive NMR-based study was conducted. NMR experiments were recorded successively in D_2O and in Me₂SO. As shown previously (26), an improved resolution was obtained in Me₂SO, which was used for most experiments.

Polysaccharide Moiety—¹H and ¹³C NMR parameters of the polysaccharide moiety from CheLM were assigned using onedimensional ¹H and ¹³C experiments as well as two-dimensional ¹H-¹H homonuclear and two-dimensional ¹H-¹³C and ¹H-³¹P heteronuclear experiments. Attributed ¹H and ¹³C

Lipoarabinomannan Structure from M. chelonae

Methylation analysis of CheLM and CheLAM
Values of molar ratio were obtained from peaks area integration of
GC analysis and corrected by the use of effective carbon-response fac-
tors as described previously (26)

TABLE I

Deserved adversations	m	ol (%)
rroposed structure	CheLM	CheLAM
t-Ara		13.5
2-Ara		6.2
5-Ara		45.8
3.5-Ara		14.3
t-Man	52.7	9.9
6-Man	16.6	4.7
3.6-Man	30.7	5.6

NMR parameters from CheLM are summarized in Table II.

The anomeric proton region is dominated by two signals at δ 4.94 and δ 4.71 ppm (Fig. 2a). The former signal correlates in the COSY 90 experiment with a single H-2 signal at δ 3.84 ppm, whereas the other correlates with two largely distinct H-2 signals at δ 3.94 ppm and δ 3.68 ppm (Fig. 2d), suggesting the presence of two anomeric protons of distinct origins at δ 4.71 ppm. The configuration of these three spin systems was unambiguously attributed to Manp in accordance with TOCSY and NOESY experiments. Moreover, the magnitude of the ${}^{1}J_{\rm H1,C1}$ coupling constant, 168–170 Hz, provided evidence for the α -anomeric configuration of these mannose units. This observation was confirmed by the presence of an intra-residual H-1/H-2 NOE contact (data not shown).

Based on the complete assignment of its ¹H and ¹³C NMR parameters and according to the published literature (14, 32), the mannose unit deriving from the anomeric proton at δ 4.94 was typified as a terminal α -Manp(IV). The signal at δ 4.71 ppm correlated with a single C-1 signal at δ 100.7 ppm and was tentatively assigned to the anomeric protons of the 6-O-substituted and 3,6-O-substituted mannose units previously observed through methylation analysis. Because of the ¹H-¹³C HMQC (Fig. 2b), HMQC-HOHAHA (data not shown), COSY and HMBC experiments (Fig. 2a), all the carbon parameters of these two units could be assigned. In particular, two distinct C-3 signals were identified, one at δ 72 ppm corresponding to an unsubstituted C-3, and one deshielded at δ 80.3 ppm indicative of a substituted C-3. On the other hand, a single deshielded C-6 resonance at δ 66.6 ppm, and a single slightly shielded C-5 resonance at δ 72.42 ppm could also be observed. These parameters confirm that the two mannosyl units are 6-O-substituted and 3,6-O-substituted. Altogether, these NMR data support the fact that the major ¹H signal at δ 4.71 corresponds to the anomeric proton of both 6- and 3,6-substituted α -Manp residues, labeled VI and VIII, respectively.

Through ${}^{s}J_{H-C}$ signal assignments, HMBC experiment (Fig. 2a) allowed us to define intra- and inter-residual correlations originating from the anomeric proton of t-Man (IV), 6-Man (VI), and 3,6-Man (VIII) residues. For t-Man, intra-residual correlations H-1/C-3 and H-1/C-5 were observed at δ 71.4 and δ 74.26 ppm, respectively, whereas an extra-residual 1-IV/3-VIII correlation was observed at δ 80.8 ppm. This unambiguously confirmed the attachment of t-Man residues to the C-3 of 3,6-Man residues. As observed above, the C-3 chemical shift is the only parameter significantly differing between 6-Man and 3,6-Man residues. Thus, the ${}^{3}J_{\mathrm{H-C}}$ signals starting from anomeric proton at δ 3.67 ppm were attributed as follows: the signal at δ 72 ppm was attributed to the intra-residue connectivities 1-VI/3-VI; signals at δ 80.31 and δ 72.42 ppm to intra-residue connectivities 1-VIII/3-VIII and 1-VIII/5-VIII, respectively; the signal at 66.56 ppm was attributed to inter-residue correlations between the anomeric protons of VI/VIII units and C-6s of VI/VIII units.

These signals demonstrate that VI and VIII residues are connected to each other through 6-O-substitutions. In addition, an intense signal at δ 70.5 ppm was also attributed to the 1-VI/2-VI correlation. Altogether, these data deriving from methylation and NMR analyses are consistent with the occurrence of a linear backbone of 6-substituted α -D-Manp residues partially substituted in C-3 by single α -D-Manp residues.

GPI Anchor—The structure of the GPI anchor of CheLM was first investigated through the use of ${}^{1}\text{H}{}^{-31}\text{P}$ HMQC and ${}^{1}\text{H}{}^{-31}\text{P}$ HMQC-HOHAHA experiments (Fig. 3). This method has been successfully applied to study the GPI anchors of LM and LAM isolated from *M. bovis* BCG by Nigou *et al.* (33). Coupled to HOHAHA experiments, they allow assigning glycerol, inositol, and mannose residues that constitute the GPI anchor.

The ³¹P spectrum of LM in Me₂SO showed only two sharp peaks at δ 1.65 and δ 1.85 ppm, indicative of the presence of two distinct GPI anchors. Integration of these signals gave us a P1/P3 ratio of 73:27. Assuming that the phosphorus atom of the GPI anchor is linked to both glycerol and myo-inositol, we used the ¹H-³¹P HMQC experiment to assign Gro H-3 and myo-Ins H-1 (Fig. 3b). As expected, each phosphorus atom correlated with two slightly different sets of signals that were identified by analogy to the literature as myo-Ins H-1 (at δ 4.03 and δ 3.93 ppm, respectively) and Gro H3/H3' (at δ 3.80 and δ 3.85 ppm, respectively). Then ¹H-³¹P HMQC-HOHAHA experiments (Fig. 3a) permitted to assign the proton spin systems of myo-Ins and Gro linked to each phosphorus atom. The assignment of these signals was confirmed by the use of sequential multirelayed ¹H-¹H COSY and HOHAHA (data not shown) sequences, using a mixing time of 120 ms. It appeared that the glycerol protons derived from both phosphorus atoms presented very similar parameters: H-1 at δ 4.32-4.35 ppm; H-1' at δ 4.09-4.11 ppm; H-2 at δ 5.07–5.09 ppm; and H3/H3' at δ 3.80–3.85 ppm (Fig. 3d). Their most relevant feature was the presence of a very deshielded H-2 proton at δ 5.07–5.09 ppm, characteristic of a 1,2-di-acylated glycerol (34). On the other hand, the two myoinositols exhibited distinct ¹H NMR parameters; in particular, the H-3 signals differed by about 1.4 ppm at δ 4.57 and δ 3.20 ppm. Such a deshielding is in agreement with an acylated position at C-3 occurring on one of the two myo-inositol types (Fig. 3c). Sets of parameters of Gro and myo-Ins matched perfectly with some of the previously analyzed GPI anchors (33, 35). By using the nomenclature proposed by these authors, we could typify the two phosphorus atoms at δ 1.65 ppm as P1 and at δ 1.85 ppm as P3, where P1 containing GPI is characterized by the presence of a diacylglycerol and a C-3-acylated myoinositol, whereas P3 containing GPI is characterized by the presence of a diacylglycerol and non-acylated myo-inositol.

Four minor deshielded anomeric protons were clearly observed in the ¹H spectrum at δ 5.15, δ 5.13, δ 5.19, and δ 5.11 ppm (Fig. 4). On the basis of their C-1 parameters at δ 98.7, δ 98.7, δ 100.96, and δ 101.5 ppm, respectively, as observed on the ¹H-¹³C HMQC spectrum (Fig. 5), and their H-2 parameters at δ 3.61, δ 3.58, δ 3.79, and δ 3.78 ppm as observed on the HOHAHA spectrum (Fig. 4c), these four residues were tentatively identified as Manp residues. On the ROESY and NOESY spectra (Fig. 4, a and b), all four anomers showed intra-residue correlations with their respective H-2, which substantiate their attribution as α -Manp residues. Moreover, starting from H-1 signals at δ 5.19 ppm, NOESY and ROESY spectra showed intense NOE effect with H-2 myo-Ins P1 at δ 4.22 ppm, whereas starting from signal at δ 5.11, NOESY and ROESY showed NOE effect with H-2 myo-Ins P3 at δ 4.19 ppm. A weaker correlation signal was also observed on the NOESY spectrum between anomeric proton at δ 5.19 ppm and H-3 Ins P1 at δ 4.57 ppm. These data clearly demonstrate the glycosylation of

Lipoarabinomannan Structure from M. chelonae TABLE II

CheLM 'H and ¹³C chemical shifts in Me₂SO-d₆ at 343 K Chemical shifts Residues 6′ 1 2 3 5 6 4 ¹H ¹³C 3.69 3.84 3.66 3.48 3.55 t-α-Manp IV 4.94 3.66 103.2 71.09 71.54 68.31 74.26 62.17 ιH 6-α-Manp VI 4.71 3.68 3.56 3.52 ND ND ND ¹³C 100.7 70.97 72 67 ND ND ND ¹H ¹³C 3,6-a-Manp VIII 4.71 3.94 3.65 3.64 3.62 3.76 3.60 100.7 70.33 80.31 66.56 66.28 72.42 α-Manp-1 ιH 5.15 3.61 ND ND ND ND ND ¹³C 98.7 ND ND ND **P1** ND ND ND 1H 5.13 3.58 ND ND ND ND ND ¹³C 98.7 ND ND ND **P**3 ND ND ND α-Manp-2 ιH 5.19 3.79 ND ND ND ND ND ¹³C **P1** 100.96 71.2 ND ND ND ND ND ¹H ¹³C ND ND ND ND 5.11 3.78 ND P3 101.5 71.5 ND ND ND ND ND Ins ιH 4.03 4.22 4.57 3.6 3.12 3.66 ¹³C **P1** 77.11 75.4 73.15 ND 80.3 ND ¹Н 13С 3.4 ND 3.93 4.19 3.05 3.62 3.2 ND ND ND ND 80.30 **P**3 Gro 1 1 2 3.3 ¹Н 13С 4.32 4.09 5.07 3.8 **P1** 63.32 71.64 63.6 63.32 ιH 4.35 4.11 5.09 3.85 P3 13C 63.32 63.32 71.64 63.6

" ND, not determined.

both myo-inositol types P1 and P3 on C-2 position by a mannose residue, labeled Man-2 (P1) and Man-2 (P3), respectively. Similarly, anomeric protons at δ 5.15 and δ 5.13 ppm correlated with their respective H-2 signals and with myo-Ins P1 H-6 at δ 3.66 ppm and myo-Ins P3 H-6 at δ 3.62 ppm, respectively. This confirmed the substitution of P1 and P3 myo-Ins at C-6 by mannose residues labeled Man-1(P1) and Man-1(P3). These data are in total agreement with previous studies (9, 16, 26) on GPI anchors of PIMs, LM, and LAM where inositol was invariably substituted at the C-2 by a single α -D-Manp residue and at C-6 by the mannan core. However, our experiments did not enable us to distinguish between a terminal mannose residue and substituted mannose residues. Considering the relative proximity of the acylation position on the myo-Ins residue with the two glycosylation positions, it is noteworthy that the parameters of Man-2 H-1 protons are highly affected by the acylation of the myo-Ins, whereas the Man-1 H-1 protons are not.

The acylation state of the glycerol moiety was subsequently studied by GC/MS by liberating the intact acylglycerol by acetolysis as described previously by Nigou et al. (26). To confirm the exact nature of substituting fatty acids, these were cleaved from the acetylated acylglycerol by methanolysis and analyzed as pyrrolidine derivatives by GC/MS. The occurrence of hydroxylated fatty acids was specifically investigated by derivation with heptafluorobutyric anhydride (36). Acetolysis products appeared as being predominantly constituted of diacylglycerols. Quantification of each form showed that monoacylglycerols represented less than 1% of the total acylglycerols analyzed. They included hexadecenoic acid $(C_{16:1})$, octadecenoic acid $(C_{18:1})$, and nonadecanoic acid (C_{19}) -substituted glycerol. This is in agreement with the ¹H-³¹P HMQC NMR experiments that only showed the presence of a diacylglycerol unit in CheLM through the observation of P1- and P3-type phosphorus atoms. The region of the EI total ion current chromatogram profile corresponding to the diacylated products exhibited 12 peaks that were all assigned by EI and CI/MS (Fig. 6a). The

three most abundant species were assigned as 1/2-tetradecanoyl-1/2-nonadecanoyl-3-acetylglycerol, 1/2-hexadecanoyl-1/2octodecenoyl-3-acetylglycerol, and 1/2-hexadecanoyl-1/2-nonadecanoyl-3-acetylglycerol, which represent 45, 34 and 13%, respectively, of the total acylglycerols isolated from CheLM. They were easily characterized because of the $(M + NH_4)^+$ ions at m/z 642, 654, and 670, respectively, and to the fragment ions resulting from the loss of each acyl group (m/z 327 and 397 for $\rm C_{14}\text{-}C_{19}$ Gro, m/z 355 and 381 for $\rm C_{16}\text{-}C_{18:1}$ Gro, and m/z 355 and 397 for C₁₆-C₁₉ Gro). Each of these compounds was attributed to two peaks showing a constant ratio of 1:5. These were tentatively attributed to the two possible positions of fatty acids, but the ions resulting from the fragmentation between C-1 and C-2 of the glycerol were not intense enough to unequivocally localize the acyl group on each carbon and thus to determine the predominant form. Similarly, the remaining peaks were attributed to a very heterogeneous family of diacylglycerols, representing less than 10% of the total products. They include C_{14} - $C_{18:1}$, C_{16} - $C_{16:1}$, C_{16} - C_{16} , C_{16} - $C_{17:1}$, C_{15} - C_{19} , C_{16} -C₁₈, C₁₆-C_{19:1}, and C₁₇-C_{18:1} di-substituted glycerols. Further fatty acid analysis based on acetolysis products indicated that the octadecenoic acid was present as a mixture of $C_{18:1}\Delta^8$ and $C_{18:1}\Delta^9$ and permitted the identification of nonadecanoic acid as tuberculostearic acid (10-methyloctadecanoic acid).

Structure of CheLAM—The structure of CheLAM was elucidated by the same experimental protocols used for CheLM. Analysis of CheLAM by MALDI showed a broad peak centered on m/z at 17,000 Da and ranging from 9000 to 23,000 Da. It is noteworthy that in identical experimental conditions CheLAM gives a much more intense response in MALDI analysis than CheLM. Composition analysis of this compound showed a mannose/arabinose/inositol/glycerol ratio of 31:80:1:1. Analysis of partially methylated and acetylated monosaccharides residues indicated the presence of t-Man, 6-Man, 3,6-Man, t-Ara, 5-Ara, 2-Ara, and 3,5-Ara (Table I). These experiments led to a first insight into the nature of the capping of LAM isolated from M. chelonae. Indeed, both CheLM and CheLAM showed a sim30640

Lipoarabinomannan Structure from M. chelonae



FIG. 2. NMR anchor analysis of CheLM. *a*, expanded region (δ¹H, 5.00-4.60, and δ¹³C, 85-60) of the ¹H-¹³C HMBC spectrum. *b*, expanded region (δ¹H, 4.05-3.40, and δ¹³C, 85-60) of the ¹H-¹³C HMQC spectrum. *c* and *d*, expanded regions δ¹H, 5.00-4.60 and 5.00-4.60, and δ¹⁴H, 4.05-3.40 and 5.00-4.60, respectively, of two relayed ¹H-¹H COSY spectra that were recorded with a spectral width of 3200 Hz in order to collect 4096 × 512 points data matrix with 32 scans per *t*1 increment. ¹H-¹³C HMBC spectrum (*a*) was recorded with a spectral width of 20,000 Hz in ¹³C and 4006 Hz in ¹⁴H dimensions in order to collect 4096 × 256 points data matrix with 128 scans per *t*1 value. ¹H-¹³C HMBC appetrum (*b*) was recorded with the same spectral width. HMBC and HMQC experiments were performed with 65- and 3.4-ms delays, respectively, for evolution of, in the first case, germinal and vicinal coupling (²J and ³J_{H-C}) and in the second case for direct coupling (¹J_{H-C}). Symbols (*) correspond to arabinosyls and mannosyls residues of traces of CheLAM, and *black arrows* (before number) on ¹H-¹³C HMBC spectrum indicate the long range coupling toward linkage. *IV*, t-α-Manp; *VI*, σ-α-Manp; *VIII*, 3,6-α-Manp.

ilar number of mannose units, suggesting the absence of additional mannose capping on arabinan branches, a well known feature present in LAM from M. bovis BCG and M. tuberculosis (14, 15). This observation was reinforced by the fact that proportions of t-Man, 6-Man, and 3,6-Man were similar in both CheLM and CheLAM and that no other type of Man unit was present in CheLAM (Table I). Calculation of the degree of capping based on the ratio of 3,5-Araf minus t-Araf to 3,5-Araf, providing that the capping motif is not alkali labile, gives a very low capping percentage of about 5%. This result is also in accordance with the absence of an oligomannosyl cap. The use of another calculation method, based on the ratio of 2-Araf minus t-Araf to 2-Araf (19), gives a capping percentage below zero due the ratio of 2-Araf to t-Araf which is inferior to 1. The significantly lower relative abundance of 2-Araf compared with t-Araf may be due either to the fact that some of the t-Araf residues may not be linked to 2-Araf residues or to an underestimation of the relative abundance of 2-Araf in our experimental conditions. However, such a discrepancy between the two methods has already been observed (37). Similarly, the

presence of a single myo-inositol residue per molecule of LAM suggests that no other myo-inositol residue than the one incorporated in the GPI anchor is present in the molecule, and consequently that arabinan chains are not terminated by phospho-myo-inositol groups as observed in fast growing Mycobacterium sp. (19) including M. smegmatis (38). Both observations were confirmed by multiple NMR experiments, as shown below.

Attributed ¹H and ¹³C NMR parameters from CheLAM are summarized in Table III. Comparison of the anomeric region of ¹H-¹³C HMQC spectra from CheLM and CheLAM (Fig. 5) allowed us to identify signals from the mannan core of CheLAM. On this basis, the intense ¹³C resonance at δ 103,1 ppm was tentatively identified as the t- α -Manp anomeric carbon (IV), whereas the ¹³C resonance at δ 100.55 was identified as the 6-α-Manp and 3,6-α-Manp anomeric carbons (VI and VIII). Two-dimensional ¹H-¹H homonuclear experiment confirmed that the ${}^{1}H/{}^{13}C$ signal at δ 4.92/103.1 ppm correlated with a single spin system characteristic of a t-Manp residue, whereas the ¹H/¹³C signal at δ 4.70/100.55 ppm correlated with two distinct spin systems characteristic of 6-Manp and 3,6-Manp units (data not shown). Assignment of their respective ¹³C resonances owing to ¹³C-¹H heteronuclear experiment confirmed this observation. Furthermore, as observed in CheLM, an HMBC experiment showed inter-residual correlations between IV H-1 and VI C-3, VI H-1 and VI/VIII C-6, and between VIII H-1 and VI/VIII C-6 (data not shown). These data confirm that the mannan cores of CheLM and CheLAM are similar.

Fig. 5 shows that four minor signals previously assigned to anomeric carbons of α -Man-1P1 and P3 and α -Man-2P1 and P3 in CheLM were also clearly observed in the CheLAM spectrum. The remaining anomeric signals of the spectrum were attributed by comparison with previous spectral NMR data of LAM from M. tuberculosis (32) and M. smegmatis (38). Spin systems deriving from each anomer was identified owing to ¹H-¹H homonuclear and ¹³C-¹H heteronuclear experiments. This way the remaining signals in the anomeric region of the spectra were all attributed to the anomeric signals of the arabinan core of CheLAM, in agreement with the methylation analysis: 3,5- α -Araf (I), 5- α -Araf (II₁ to II₄), 2- α -Araf (III₁, III₂, and III₃), and t- β -Araf (V₁ and V₂). No additional mannose residues could be identified. Most types of Ara units showed multiple anomeric signals presenting identical spin systems. This multiplicity of signals was previously attributed to the different positions that each unit may take within the arabinan core (14). HMBC spectrum showed a complex pattern of $^3J_{\rm H-C}$ correlations that was entirely resolved, except for the region between δ 4,95 and 5 ppm where all intra-residue connectivities from III_1 , III_2 , II_3 , and II₄ residues could not be unambiguously attributed because of the very close chemical shift of their anomeric protons (not shown). Nevertheless, this shed some light onto the general sequence of the arabinan core. In particular all the anomeric protons of 3,5- α -Araf residues (I) and 5- α -Araf residues (II₁ to II₄) showed intense ${}^{3}J_{\rm H-C}$ correlation signals with a substituted C-5 at δ 67.9 ppm, unambiguously attributed as inter-residue connectivities with C-5 I and C-5 II. Similarly, both H-1 of t- β -Araf residues (V₁ and V₂) showed inter-residue connectivities with a very deshielded carbon at δ 88.29 ppm attributed to the C-2 of $2-\alpha$ -Araf, confirming the attachment of all t- β -Araf at C-2 of 2- α -Araf residues. On the other hand, it is noteworthy that no other connectivity with the C-2 of $2-\alpha$ -Araf residues could be observed suggesting that no arabinose residues other than t- β -Araf residues are linked to 2- α -Araf units. HMBC experiment also enabled us to distinguish III₃ residues from III₁ and III₂ residues on the basis of their respective ${}^{3}J_{H-C}$ connectivities. Indeed, whereas H-1 III₁/III₂ showed an intense inter-residue correlation with C-5 I/IIs at 67.9 ppm, H-1 III₃ did



FIG. 3. NMR analysis of LM phosphate substituents in Me₂SO-d₆ at 343 K. a and b, expanded regions (δ^{1} H, 5.20 to 3.00, and δ^{31} P, 2.40 to 1.20) of the heteronuclear ³¹P-decoupled, ¹H-detected HMQC-HOHAHA and HMQC spectra, respectively. HMQC-HOHAHA was recorded using MLEV-17 mixing sequence of 120 ms. The data matrix for (a and b) experiments were 2048 × 64 time proportional phase increment (TPPI) points with 256 scans for (a) experiment or 64 scans for (b) experiment per t1 value. In both cases, the spectral window was 2592 Hz in F1 (³¹P) dimension and 4003 Hz in F2 (¹H). The original data matrix were expanded to 4096 × 512 real matrix moreover for processing a sine-bell window shifted by IV2 was applied in both dimensions. c, expanded region (δ^{1} H, 5.20–3.00, and 4.60–4.54); d, expanded region (δ^{1} H, 5.20–3.00, and 4.40–4.28) of homonuclear two-dimensional clean TOCSY spectrum. This experiment was recorded with a spin-lock time of 120 ms. The data matrix was 2048 × 512 (states) points with 32 scans per t1 increment and was expanded to 4096 × 1024.

not show any correlation at δ 67.9 ppm but an intense and broad signal around δ 83.5 ppm attributed to C-4-III₃ and to C-3-I. These data establish that III₁ and III₂ residues are linked to C-5 of 3,5- or 5- α -Araf residues, whereas III₃ is linked to C-3 of 3,5- α -Araf. Therefore, NMR data arising from the arabinan chain of CheLAM show very little discrepancy with published NMR data of LAM from *M. tuberculosis* (32) and *M. smegmatis* (38), strongly suggesting that a similar arabinan core is shared by the three mycobacterial species. On the other hand, methylation and composition analyses have previously suggested that M. chelonae differs in its capping status from M. tuberculosis and M smegmatis. Indeed, as already mentioned, homonuclear ¹H-¹H and heteronuclear ¹H-¹³C NMR experiments showed no evidence of the presence of other mannose residues than the one found in the mannan core, suggesting the absence of oligomannosyl caps on the arabinan side chains as observed in M. tuberculosis and M. bovis. Furthermore, it was observed on the HMBC spectra
Lipoarabinomannan Structure from M. chelonae



FIG. 4. Part of NMR spectra from CheLM analysis in Me_2SO-d_s at 343 K. a and b, expanded regions (δ^1 H, 4.70-3.00 and 5.23-4.8 ppm) of 400-ms ROESY spectrum and 300-ms NOESY spectrum, respectively. c and d, expanded regions (δ^1 H, 4.70-3.00 and 5.22-5.07, and δ^1 H, 4.70-3.00 and 3.24-3.02) of 120-ms TOCSY spectrum. *Man-1* corresponds to O-6 linked mannosyl and *Man-2* to O-2 linked mannosyl on inositol unit.

that no additional ${}^{3}J_{\rm H-C}$ correlation occurred between any anomeric protons of mannose residues (IV, V, and VIII) and any carbon of arabinose residues. The possibility of the occurrence of phospho-*myo*-inositol type capping motifs is discussed below. *M. chelonae*, we used an experimental approach used above for CheLM. The ³¹P spectrum of CheLAM showed two sharp peaks at δ 1.65 and δ 1.85 ppm. Similarly, ¹H-³¹P HMQC-HOHAHA experiments established that glycerol and *myo*-inositol residues substituting both phosphorus types showed ¹H-NMR parameters identical to those of CheLM. Thus, it was concluded

In order to study the GPI anchor of CheLAM isolated from

Lipoarabinomannan Structure from M. chelonae



'H Chemical shifts (ppm)

FIG. 5. Parts of heteronuclear ¹H-¹³C HMQC spectra. a and b, expanded regions (δ^{1} H: 5.30–4.60, and δ^{13} C: 110.0–97.0) of LM-HMQC and LAM-HMQC spectra, respectively. *I*, 3,5- α -Araf; *II*, 5- α -Araf; *III*, 2- α -Araf; *IV*, t- α -Manp; V, β -Araf; VI, 6- α -Manp; VIII, 3,6- α -Manp. *, unidentified signals.

that GPI of CheLAM and CheLM shared identical acylation states, characterized by a mixture of P1 and P3 phosphates, in similar ratios. Detailed analysis of the acetolysis products of CheLAM by GC/MS showed again that diacylglycerols of CheLM and CheLAM present the same heterogeneity, reinforcing the hypothesis that CheLM and CheLAM share strictly identical GPI anchors.

It is noteworthy that the ${}^{1}H{}^{31}P$ HMQC spectrum did not reveal any other type of phosphorus than those attributed to the GPI anchors of the lipopolysaccharide. In particular, no evidence for the presence of an additional phospho-myo-inositol group substituting the arabinan side chains was found, as observed in LAM from *M. smegmatis* (38). Altogether, with the results of the composition analysis showing a myo-inositol/ glycerol ratio of 1:1, the NMR data confirmed that CheLAM does not possess a phospho-myo-inositol capping motif.

Mapping with Lectins

Occurrence of a new type of mannan core in lipoglycans isolated from *M. chelonae* was investigated by screening purified LM and LAM from various mycobacterial species with several mannose-recognizing lectins. The most relevant results were obtained with lectins isolated from *Canavalia ensiformis* (ConA) and from *Galanthus nivalis* (GNA). Although ConA is relatively unspecific toward mannose-containing glycoconjugates, GNA is known to be highly specific for non-reducing terminal α 1,3-linked mannose residues (39). As expected, ConA interacted with purified LM from *M. chelonae*, *M. bovis* BCG, and *M. smegmatis* (Fig. 7*a*). This suggests that ConA does not differentiate mannan cores branched at C-2 (*M. bovis* BCG and *M. smegmatis*) from mannan cores branched at C-3 (*M. chelo*- nae). Surprisingly, this lectin only recognized purified LAM from M. tuberculosis (Fig. 7a), but not LAM from M. chelonae or LAM from M. smegmatis, although LM and LAM share common mannan cores. The absence of reactivity between ConA and LAM from M. chelonae or M. smegmatis may be explained by steric hindrance of the mannan core by the arabinan polymer. Specific interaction observed between ConA and LAM from M. tuberculosis would originate from the presence of an oligomannosyl-type capping only in this species. These results are in agreement with findings from Prinzis et al. (40) who reported that ConA recognizes purified LAM from M. tuberculosis Erdman but not from a rapidly growing strain.

30643

GNA exhibited a much more restricted specificity than ConA. As observed in Fig. 7b, GNA recognizes exclusively CheLM. Considering the described specificity of GNA toward a Man(α 1-3) determinant, this result is in total agreement with direct structural studies in this report that distinguish the mannan core of CheLM from those of other mycobacterial species on the basis of their respective branching positions. As observed for ConA, GNA did not bind to CheLAM, again suggesting that arabinan masks the α 1,3-Manp side chains of the mannan core.

Given the specificity of GNA for the $Man(\alpha 1,3)$ -substituted mannan core of CheLM, we addressed the question whether this lectin may be particularly useful to rapidly detect the presence of similar structures in various other mycobacterial species, including rapid, slow growing, pathogenic, and non-pathogenic species. Crude lysates of *M. bovis* BCG, *M. smegmatis*, *M. chelonae*, *Mycobacterium scrofulaceum*, *Mycobacterium thermoresistibile*, *Mycobacterium fortuitum*,



FIG. 6. GC/MS analysis of actetolysis products of CheLM. a, total ion count of acetylated diacylglycerol obtained after acetolysis of native CheLM. According to their EI and CI spectra the 12 peaks were attributed as follows: 1, C-14—C-18:1; 2, C-16—C-16:1; 3, C-16—C-16; 4 and 5, C-19—C-14; 6, C-16—C-17:1; 7, C-15—C-19 and C-16—C-18; 8 and 9, C-16—C-18:1; 10 and 11, C-16—C-19; 12, C-16—C-19:1 and C-17—C-18:1. b, fragmentation pattern of peak 4 in EI mode. c, $[M + NH_4]^+$ ion of peak 4 in CI mode. d, deduced fragmentation pattern of peak 4 in EI mode.

Mycobacterium xenopi, Mycobacterium gastri, and Mycobacterium malmoense were transferred to a membrane, incubated with GNA-DIG, and revealed with anti-DIG antibodies. Western blot analysis shows that Man(α 1,3)-substituted mannan core was only present in *M. chelonae* (Fig. 7c). Interestingly, GNA did not interact with LM from *M. fortuitum*, known to be a *M. chelonae* related species. This indicates that GNA constitutes a useful tool for a rapid screening of LM and LAM structures with α 1,3-linked mannose residues without the need of a purification step. It also suggests that among the nine mycobacterial strains analyzed, the structure of *M. chelonae* appears to be unique.

Fig. 8 summarizes all the data collected from the structural analysis of CheLAM and highlights the main structural features and differences between CheLAM and LAMs from the already studied species M. tuberculosis/M. bovis and M. smegmatis.

TNF-a and IL-8 Secretion by Various LAM Preparations

To investigate the consequences of the structural differences established between CheLAM, ManLAM, and AraLAM, we compared TNF- α and IL-8 secretion by differentiated human macrophages stimulated with LAMs isolated from *M. tuberculosis, M. smegmatis,* and *M. chelonae.* As shown in Fig. 9, incubation THP-1 cells with LAM isolated from *M. smegmatis* resulted in a sharp increase of both TNF- α and IL-8 secretion. In contrast, LAMs from *M. chelonae* and *M. tuberculosis* were not able to induce cytokine production.

Lipoarabinomannan Structure from M. chelonae

CheLAM ¹H and ¹³C chemical shifts in Me₂SO-d₆ at 343 K Chemical shifts Residues 1 2 3 5 6 4 3,5-α-Araf I ιH 4.87 4.02 3.823.85° 3.57/3.61 13C 109.35 81.2 83.64 84.9 67.9 ιH 4.82 3.85° 3.69 3.934 3.57/3.70 5- α -Araf II₁ 13C 82.4 108.6 78.2 82.7 67.9 ιH 5-α-Araf II, 4.86 3.87° 3.73° 3.934 3.57/3.70 ¹³C 108.7 82.4 78.2 62.7 67.9 ιH 3.76 3.57/3.70 5-α-Araf II_a 4.97 3.86° 3.93 13C 106.65 82.4 78.2 82.7 67.9 5-α-Araf II ιH 4.95 3.86 3.76* 3.93 3.57/3.70 13C 108.8 82.4 78.2 82.7 67.9 Η^ι 2-α-Araf III, 4.97 4.00^a 3.88 3.8 3.5/3.6 ¹³C 106.65 88.29 76.2 83 62.2 ιH 3.889 3.5/3.6 $2-\alpha$ -Araf III₂ 4.95 4.03 3.8 13C 106.8 76.2 88.09 83 62.2 ιH 5.09 2-a-Araf III3 4 02 3 879 3.789 3.5-3.6 18C 106.2 76.2 88.46 83.3 62.2 ΊH t-α-Manp IV 4.92 3.83 3.45 3.65 3.66 3.6213C 103.1 71.2 71.9 74.3 67.9 61.8 ^{1}H 4.94 3.85 3.84 3.65 3.6 β -Araf V₁ 13C 101.8 78.2 75.7 84.25 64.01 β-Araf V₂ ιH 4.90 3.85 3.84 3.65 3.6 13C 101.1 75.7 78.2 84.25 64.01 6-α-Manp VI ιH 4.70 3.70 ND^b ND ND ND ¹³C 100.55 71.2 ND ND ND ND ¹Н 13С 3.69/3.56 3,6-a-Manp VIII 4.702 3.93 3.64-3.66 3.66 3.65 100.55 70.57 80.6 66.18 72.4 67.03 1H ND 5.20 ND ND ND a-Manp 2 **P1** 3.79 ¹³C 100.5 ND ND ND ND ND ¹H ¹³C 3 789 ND ND P3 5.12ND ND α -Manp 2 ND ND 101.7 ND ND ND P1/P3 1H ND ND ND ND ND 5.1 a-Manp 1 ¹³C 98.7 ND ND ND ND ND ιH 3.97 4.23 4.58 3.63 3.68 Ins **P1** 3.14 ¹³C ND 75.18 73.05 ND ND 80.48 Ins **P**3 ιH 3.91 4.20 3.22 3.45 3.08 3.64 ¹³C ND ND ND ND ND ND 1 1' 2 3.3 Gro **P1** Ϋ́ 4.31 4.1 5.09 3.85 13C 66.69 66.69 ND 71.46 Gro **P**3 ιH 4.36 4.125.123.8 13C ND ND ND ND

" Value derived from HMBC analysis.

^b ND, not determined.

" Value derived from ROESY analysis.

Previous studies (38, 41, 42) have been reported that mannose-capped LAM from *M. tuberculosis* was inactive toward TNF- α and IL-8 secretion, whereas phosphoinositol-capped LAMs from a fast growing *Mycobacterium* species and from *M. smegmatis* were strong cytokine inducers. Considering the fact that CheLAM is neither capped by oligomannosyl nor by phosphoinositol groups, our results strongly suggest that TNF- α and IL-8 secretion by LAM was very likely modulated by the presence of the phosphoinositol-type capping. This also suggests that the mannose type capping does not influence cytokine induction.

DISCUSSION

LM and LAM are two polysaccharides of the mycobacterial cell wall that display strong antigenicity and that are widely distributed in Mycobacterium sp. LAM has been shown to participate in the virulence and immunopathogenesis of tuberculosis (43). Extensive studies performed on LAM from M. bovis BCG or M. tuberculosis have demonstrated that important biological effects are linked to the degree and chemical nature of capping functions and other substituents on the arabinan portion of LAM. Arabinans of LAMs from M. bovis BCG, M. tuberculosis, M. leprae, and M. avium (44) are mannosecapped (ManLAM) to varying degrees. Studies based on a fast growing Mycobacterium sp. and on M. smegmatis demonstrated that the arabinan termini are uncapped (AraLAM), whereas a minor portion terminates with inositol phosphate caps (19, 38, 40).

So far, LAM structures from only a restricted panel of mycobacterial species have been determined. The relatively low structural diversity encountered in these molecules may limit the possibilities of studying relationships between structure and biological activities. The aim of the present work was to decipher new sources of structural variability in LM and LAM in order to generate new models for such studies. We therefore used *M. chelonae* as a model. *M. chelonae* is a nontuberculous mycobacterium that is usually encountered in water sources 30646

Lipoarabinomannan Structure from M. chelonae



FIG. 7. Reactivity of lectins with LMs and LAMs. Western blot analysis of purified LMs or LAMs were blotted with either ConA-DIG (a) or GNA-DIG (b) and incubated with anti-DIG antibodies conjugated with alkaline phosphatase. Lanes 1 and 4, M. chelonae; lane 2, M. bovis BCG; lanes 3 and 6, M. smegmatis; lane 5, M. tuberculosis H37Rv. c, Western blot analysis with GNA-DIG of crude lysates from various mycobacterial species as indicated.



FIG. 8. Structural models of LAMs from *M. chelonae* (a), *M. tuberculosis/M. bovis* BCG (b), or *M. smegmatis* (c). The emphasis was put on the parts of the molecules showing inter-specific heterogeneity, in particular the nature of capping, the substitution of the mannan core, and the acylation state of the GPI anchor. Despite subtle structural differences between LAM from *M. tuberculosis* and *M. bovis* BCG, both species do not show important discrepancies concerning the sus-mentioned points and were therefore associated. When known, nature of the fatty acids that substitute glycerol was indicated. *C-X* represents 12-O-(methoxypropanoyl)-12-hydroxystearic acid. Only certified acylation positions are indicated.

and soil. Although it has originally been proposed to be Friedmann's turtle tubercle bacillus (45), it sometimes causes disease in humans. *M. chelonae* presents also the advantage of being a rapidly growing species. Furthermore, we have recently shown that it can be transformed, thus making it attractive for future genetic studies (24). Considering the intrinsic polydispersity of mycobacterial lipoglycans, structural studies were undertaken, consisting of an independent comparison of each structural feature of these molecules with their homologs that have been studied in other species.

The first observation was that LAM from M. chelonae shares a similar overall structure with other LAMs, composed of an arabinan chain and a mannan chain linked to a GPI anchor. However, it shows unique features that unequivocally distinguishes it from other LAMs. Significant differences observed in M. chelonae are illustrated in Fig. 8. Considering the growing inter-specificity heterogeneity observed in LAMs and, as proposed by Khoo *et al.* (19), we referred to lipoglycans isolated from M. chelonae CheLM and CheLAM, according to their origin, rather than to their structural features.

In this study, we demonstrate for the first time that a virulent mycobacterium species exhibits a LAM structure that is devoid of mannose and inositol phosphate capping motifs (Fig. 8). Therefore, both CheLM and CheLAM represent suitable molecules to use along with ManLAMs or AraLAMs to evaluate the role of mannose and inositol phosphate capping in relation to various biological functions attributed to LAM. As an example, it has been observed for a long time that LAMs extracted from avirulent mycobacterial species were more potent inducers of TNF- α than those extracted from virulent species (46). A correlation was rapidly made between the absence of an oligomannosyl capping and the potency to trigger TNF- α release. It was also reported that the activity of LAM from the avirulent strain *M. smegmatis* was diminished after alkali treatment (38), suggesting the importance of alkali-labile phosphoinositol groups in eliciting cytokine production. The present study confirms this observation by showing that a LAM naturally lacking mannose and phosphoinositol caps does not induce cytokine secretion.

In addition, our results demonstrate that both CheLM and CheLAM possess an unusual mannan core characterized by the presence of $\alpha 1,3$ -Manp side chains, rather than the $\alpha 1,2$ -Manp substitution as found in all mycobacterial LAM structures reported so far (Fig. 8). Analysis by Western blotting using GNA, an $\alpha 1,3$ -Manp-specific lectin, revealed that $\alpha 1,3$ -Manp side chains containing LM were only present in *M. chelonae* but not in other mycobacterial species. Whether these substitutions are important for biological activities of CheLM or CheLAM Lipoarabinomannan Structure from M. chelonae



FIG. 9. TNF- α and IL-8 release by human macrophages in response to various mycobacterial LAM preparations. Differentiated human promonocytic THP-1 cells were stimulated with CheLAM, LAM from *M. tuberculosis* (ManLAM), or LAM from *M. smegmatis* (AraLAM) at 10 µg/ml. Control cells were incubated without stimulation (*Medium*). Culture supernatants were collected after 6 and 24 h of incubation and assessed for TNF- α (a) and IL-8 (b) secretion, respectively. Data presented are means ± S.E. for one representative experiment of three with comparable results (p < 0.05).

remains to be investigated. Another substantial difference arising from the present study concerns the LM/LAM ratio. It is generally assumed that LM is a precursor of LAM and is usually present in smaller amounts than LAM (26). We observed that M. chelonae contains a LM/LAM ratio of 3:1 (w/w), corresponding to ~10 times more molecules of CheLM than CheLAM. The presence and the possible physiological impact of larger amounts of CheLM than CheLAM in M. chelonae is presently unknown. This difference in the CheLM/CheLAM balance may arise from a deficiency in arabinosyltransferase(s) expression in M. chelonae compared with M. tuberculosis and M. bovis BCG. On the other hand, arabinosyltransferase(s) involved in arabinan polymerization may be less active in M. chelonae than in M. tuberculosis. At present, this remains very difficult to investigate because no arabinostransferases involved in LAM biosynthesis have been clearly identified.

Another unexpected difference between CheLAM and Man-LAMs concerns the lipid part of the PI anchors. Nigou *et al.* (26) have demonstrated that cellular ManLAM anchors from *M. bovis* BCG mainly contain 1- or 2-palmitoylglycerol, 1- or 2-tuberculostearoylglycerol, 1,2-dipalmitoylglycerol, and 1-tuberculostearoyl-2-palmitoylglycerol. Surprisingly, the same authors found that the PI anchor of parietal ManLAM from *M. bovis* BCG is composed of a single type of acylglycerol consisting of 1-[12-O-(methoxypropanoyl)-12-hydroxystearoyl]glycerol. In addition, parietal ManLAM was found to be a better inducer of TNF- α and IL-8 by human dendritic cells than cellular LAM, thus suggesting that the unusual 1-[12-O-(methoxypropanoyl)-12-hydroxystearoyl]-glycerol part is the major cytokine-regulating component of the ManLAMs (26). Moreover, the fact that biological activities of ManLAMs are abrogated after deacylation and that ManAMs devoid of the PI anchor are unable to stimulate cytokine production are consistent with the assumption that the lipid part of the PI anchor is essential for the immunological properties of the ManLAMs. The present study shows that the lipid part of the PI anchor of CheLAM is fundamentally different from that of previously described PI anchors because it presents a very large heterogeneity, due to a combination of fatty acids ranging from myristic acid to tuberculostearic acid with eventual unsaturations (Fig. 8). This heterogeneity generates a large pool of molecules with various PI anchors, which may influence the biological activities of the whole molecule. Whether this heterogeneity is due to numerous acylglycerol acyltransferases with different specificities in M. chelonae remains an open question. It is also noteworthy that CheLM and CheLAM show perfectly identical mixtures of diacylglycerols. This observation substantiates the hypothesis that LM is a biosynthetic precursor of LAM.

Mycobacteria represent a major source of microbial antigens that are presented by CD1 molecules. Most of the lipids that can be presented by CD1 molecules have two hydrophobic tails and a polar head group. This led to the suggestion that each hydrophobic tail fits into a pocket of the CD1 groove (47). Binding studies also revealed that the lipid portion of the

30648

Lipoarabinomannan Structure from M. chelonae

antigen is required for CD1 binding (48, 49). Sieling et al. (50) have shown that ManLAMs from M. tuberculosis Erdman and M. leprae can be presented in the context of CD1 molecules and are able to stimulate CD4/CD8 double negative $\alpha\beta$ T cells. How different CD1 molecules bind, in particular LM/LAM, is not known. Available data suggest that human CD1b and CD1d molecules require a minimal hydrocarbon chain length but that they are not as selective with regard to the type of lipid they will bind (48, 49). Therefore, it would be interesting to study the PI anchors of CheLM or CheLAM in relation to CD1 binding for presentation of these glycolipids to T cells. Given their unusual structures, CheLM and CheLAM constitute new tools to compare the biological activities of LMs and LAMs such as for antigen presentation via CD1 molecules, cell adhesion, or T cell activation and should help to clarify their roles in mycobacterial virulence.

REFERENCES

- 1. Daffé, M. & Draper, P. (1998) Adv. Microb. Physiol. 39, 131-203
- Dialey and C. Mehlert, A. & Lamb, J. (1988) Clin. Exp. Immunol. 74, 206-210
 Sibley, L. D., Hunter, S. W., Brennan, P. J. & Krahenbuhl, J. L. (1988) Infect. Immun. 56, 1232-1236
- 4. Sibley, L. D., Adams, L. B. & Krahenbuhl, J. L. (1990) Clin. Exp. Immunol. 80, 141-148
- 5. Chan, J., Fan, X. D., Hunter, S. W., Brennan, P. J. & Bloom, B. R. (1991) Infect. Immun. 59, 1755-1761 6. Schlesinger, L. S., Hull, S. R. & Kaufman, T. M. (1994) J. Immunol. 152,
- 4070-4079
- 4070-4079
 Apostolou, I., Takahama, Y., Belmant, C., Kawano, T., Huerre, M., Marchal, G., Cui, J., Taniguchi, M., Nakauchi, H., Fournie, J. J., Kourilsky, P. & Gachelin G. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 5141-5146
 Gilleron, M., Ronet, C., Mempel, M., Monsarrat, B., Gachelin, G. & Puzo, G.
- (2001) J. Biol. Chem. 276, 34896-34904
- 9. Khoo, K. H., Dell, A., Morris, H. R., Brennan, P. J. & Chatteriee, D. (1995)
- Glycobiology 5, 117-127
 Besra, G. S., Morehouse, C. B., Rittner, C. M., Waechter, C. J. & Brennan, P. J. (1997) J. Biol. Chem. 272, 18460-18466
- Schaeffer, M. L., Khoo, K. H., Besra, G. S., Chatterjee, D., Brennan, P. J., Belisle, J. T. & Inamine, J. M. (1999) J. Biol. Chem. 274, 31625–31631
- Kremer, L., Gurcha, S. S., Bifani, P., Hitchen, P. G., Baulard, A., Morris, H. R., Dell, A., Brennan, P. J. & Besra G. S. (2002) *Biochem. J.* 363, 437–447
 Hunter, S. W. & Brennan, P. J. (1990) *J. Biol. Chem.* 265, 9272–9279
 Venisse, A., Berjeaud, J. M., Chaurand, P., Gilleron, M. & Puzo, G. (1993) *J. Biol. Chem.* 268, 12401–12411
 Chetterian, D. Hawell, K. Brenin, P. M. Neil, M. P. & Bargar, P. J. (1990)
- Chatterjee, D., Lowell, K., Rivoire, B., McNeil, M. R. & Brennan, P. J. (1992) J. Biol. Chem. 267, 6234-6239
- Chatterjee, D., Hunter S. W., McNeil, M. & Brennan, P. J. (1992) J. Biol. Chem. 267, 6228-6233
- 17. Chatterjee, D., Bozic, C. M., McNeil, M. & Brennan, P. J. (1991) J. Biol. Chem. 266, 9652-9660
- 18. Chatterjee, D., Khoo, K. H., McNeil, M., Dell, A., Morris, H. R. & Brennan, P. J.

- (1993) Glycobiology 3, 497-506
 19. Khoo, K. H., Dell, A., Morris, H. R., Brennan, P. J. & Chatterjee, D. (1995) J. Biol. Chem. 270, 12380-12389
 20. Roach, T. I., Barton, C. H., Chatterjee, D. & Blackwell, J. M. (1993) J. Immu-1000 (1996) 1000 (1997)
- nol. 150, 1886-1896
- Wolinsky, E. (1979) Am. Rev. Respir. Dis. 119, 107-159
 Wallace, R. J., Jr., Brown, B. A. & Onyi, G. O. (1992) J. Infect. Dis. 166,
- 405-412 23. Ingram, C. W., Tanner, D. C., Durack, D. T., Kernodle, G. W., Jr. & Corey,
- G. R. (1993) Clin. Infect. Dis. 16, 463-471
 24. Kremer, L., Dover, L., Carrère, S., Nampoothiri, M., Lesjean, S., Brown, A. Brennan, P. J., Minnikin, D. E., Locht, C. & Besra, G. S. (2002) Biochem. J., 364. 423-430
- Minnikin, D. E., Minnikin, S. M., Goodfellow, M. & Stanford, J. L. (1982) J. Gen. Microbiol. 128, 817-822
- Gen. Microbiol. 128, 817-822
 Nigou, J., Gilleron, M., Cahuzac, B., Bounéri, J. D., Herold, M., Thurnher, M. & Puzo, G. (1997) J. Biol. Chem. 272, 23094-23103
 Tsai, C. M. & Frasch, C. E. (1982) Anal. Biochem. 119, 115-119
 Shaka A. J., Barker P. B. & Freeman R. (1985) J. Magn. Reson. 64, 547-552
 Bax, A. D. & Davis, J. (1985) J. Magn. Reson. 63, 207-213

- Kamerling, J. P., Gerwig, G. J., Vliegenthart, J. F. & Clamp, J. R. (1975) Biochem. J. 151, 491-495
- Ciucanu, I. & Kerek, F. (1984) Carbohydr. Res. 131, 209-217
 Ciuleron, M., Bala, L., Brando, T., Vercellone, A. & Puzo, G. (2000) J. Biol. Chem. 275, 677-684
- Nigou, J., Gilleron, M. & Puzo, G. (1999) Biochem. J. 337, 453-460
 Wang, Y. & Hollingworth, R. I. (1995) Anal. Biochem. 225, 242-251
 Gilleron, M., Nigou, J., Cahuzac, B. & Puzo, G. (1999) J. Mol. Biol. 285,
- 2147-2160
- 36. Zanetta, J. P., Timmerman, P. & Leroy, Y. (1999) Glycobiology 9, 255-266
- Vercellone, A., Nigou, J. & Puzo, G. (1988) Front. Biosci. 3, 149–163
 Gilleron, M., Himoudi, N., Adam, O., Constant, P., Venisse, A., Rivière, M. &
- Puzo, G. (1997) J. Biol. Chem. 272, 117-124
 Shibuya, N., Goldstein, I. J., Van Damme, E. J. & Peumans, W. J. (1988) J. Biol. Chem. 263, 728-734
- 40. Prinzis, S., Chatterjee, D. & Brennan, P. J. (1993) J. Gen. Microbiol. 139, 2649-2658
- 41. Dahl, K. E., Shiratsuchi, H., Hamilton, B. D., Ellner, J. J. & Toossi, Z. (1996)
- Infect. Immun. 64, 399-405 42. Zhang, Y., Broser, M., Cohen, H., Bodkin, M., Law, K., Reibman, J. & Rom, W. N. (1995) J. Clin. Invest. 95, 586-592
- Chatterjee, D. & Khoo, K. H. (1998) Glycobiology 8, 113–120
 Khoo, K. H., Tang, J. B. & Chatterjee, D. (2001) J. Biol. Chem. 276, 3863–3871
 Friedmann, F. F. (1903) Zentr. Bakt. Parasitenkr. Infektionskr. Hyg. Orig. I 34, 647-658
- 46. Chatterjee, D., Roberts, A. D., Lowell, K., Brennan, P. J. & Orme, I. M. (1992)
- Chatterjee, D., Roberts, A. D., Loweit, K., Brennan, F. J. & Orme, I. M. (1992) Infect. Immun. 60, 1249-1253
 Moody, D. B., Reinhold, B. B., Guy, M. R., Beckman, E. M., Frederique, D. E., Furlong, S. T., Ye, S., Reinhold, V. N., Sieling, P. A., Modlin, R. L., Besra, G. S. & Porcelli, S. A. (1997) Science 278, 283-286
 Forth W. A. When, J. G. O. Nicital K. B. Guttaria, D. Marka, D. R.
- 48. Ernst, W. A., Maher, J., Cho, S., Niazi, K. R., Chatterjee, D., Moody, D. B.,
- Bernst, W. A., Maner, J., Uno, S., Mazi, K. K., Chatterjee, D., Moody, D. B., Besra, G. S., Watanabe, Y., Jensen, P. E., Porcelli, S. A., Kronenberg, M. & Modlin, R. L. (1998) *Immunity* 8, 331-340
 Naidenko, O. V., Maher, J. K., Ernst, W. A., Sakai, T., Modlin, R. L. & Kronenberg, M. (1999) *J. Exp. Med.* 190, 1069-1080
 Sieling, P. A., Chatterjee, D., Porcelli, S. A., Prigozy, T. I., Mazzaccaro, R. J., Soriano, T., Bloom, B. R., Brenner, M. B., Kronenberg, M., Brennan, P. J. & Modlin, R. I. (1995) *Science* 260 272-230 Modlin, R. L. (1995) Science 269, 227-230

13. Etude de la structure des lipoglycannes de M. kansasii

LAM general features

Mannan domain

In order to ease attribution of the NMR signals, and considering the biosynthetic filiations hypothesized between both compounds, LM and LAM have been studied simultaneously. Due to its greater simplicity, LM enabled an easy deciphering of the structure of mannan and glycero-phosphatidyl-*myo*inositol domains. Then, structure of both domains were compared in KanLM and KanLAM, and structure of arabinan domain was studied on KanLAM.

Quantitative analysis of the alditol-acetates derivatives from KanLM and KanLAM led to Ara/Man/myoinositol molar ratios of 4/38/1 and 65/56/1, respectively. Detection of arabinose residues in LM traditionally devoid of arabinose suggested a slight contamination of KanLM by KanLAM. Indeed, these arabinose units were identified as t-Ara, 2-Ara, 5-Ara and 3,5-Ara by methylation analysis. Subsequent NMR analyses showed that arabinose units presented identical parameters in KanLM and KanLAM samples, which demonstrated that detected arabinose in KanLM sample arose from a contamination by KanLAM, estimated to about 5% (mol/mol). Analysis of partially methylated alditol-acetate derivatives (Fig. 1) revealed the presence of three major types of mannose units in KanLM identified owing to their retention times and fragmentation patterns as t-Manp, 6-Manp and 2,6-Manp (Table I). Small amounts of 5-Araf and 2-Manp were also identified in similar quantities. Considering the facts that contamination of KanLM by KanLAM represents only 5% (mol/mol) of the total, and that 5-Araf constitutes about 50% of KanLAM (Table I), the fact that 2-Manp and 5-Araf were present in equal amounts in KanLM sample strongly suggested that 2-Manp was associated to KanLM molecule and did not originated from contaminating KanLAM. Exact arrangement of these units was further studied by multiple NMR experiments. ¹H and ¹³C NMR parameters of the mannan domain from KanLM were assigned using one-dimensional ¹H and ¹³C experiments as well as two-dimensional ¹H-¹H homonuclear, and two-dimensional ¹H-¹³C heteronuclear experiments. Attributed ¹H and ¹³C NMR parameters from KanLM are summarized in Table II. Based on their NMR parameters and on literature [Gilleron, M. et al., 1999], spin systems of major t- α -Manp, 6- α -Manp and 2,6- α -Manp units were identified and labelled as IV, VIII and VI units respectively. Also, spin system of a minor 2-\alpha-Manp unit (VII) was clearly identified on two-dimensional ¹H-¹H homonuclear spectra out of which its ¹H parameters were defined, and matched with those of 2- α -Manp from oligomannosyl capping of ManLAM [Gilleron, M. et al., 2000]. Due to overlap with 2,6-α-Manp. its carbon chemical shifts were not attributed. As seen on the anomer region of ¹H homonuclear and ¹³C-¹H heteronuclear NMR spectra (Fig. 2b), VI unit showed a marked heterogeneity, that led to the identification of two distinct 6- α -Man residues labelled VI₁ and VI₂ owing to their H-1 chemical shifts at δ 4.70 and δ 4.78 ppm, respectively. This suggested that both 6- α -D-Manp units presented distinct arrangements within the mannan core of KanLM. HMBC experiment unambiguously demonstrated that mannan domain of KanLM was composed of a linear backbone of α 1,6-linked α -D-Manp residues, partially substituted in C-2 position mostly by α -D-Manp residues (Fig. 3). In particular, the occurrence of an extra-residual ${}^{3}J_{H-C}$ correlation between IV H-1 and VIII C-2 at 8 79.4 ppm established that all t-Man units (IV) substituted 2,6-Man units (VIII) or 2-Man units (VII) in C-2 positions, while VIII H-1/VIII C-6,6' and VIII H-1/VI C-6,6' correlations at δ 67.3 and 66.5 ppm, respectively, established that VIII units substituted both VIII and VI units in C-6 positions (data not shown). Interestingly, it was observed that H-1 of VI₁ and VI₂ units established distinct ${}^{3}J_{H-C}$ correlations with either VIII C-6 at δ 67.0 ppm or VI C-6 at δ 66.5 ppm, respectively. This distinct positions within the mannan backbone may in part account for the differences observed in VI1 and VI2 H-1 chemical shifts. Furthermore, the slight differences of chemical shifts observed between C-6s of VIII residues (δ 67.0 and 67.3) revealed the occurrence of two distinct populations of 2,6-Man residues only differing in their C-6 parameters, VIII₁ and VIII₂, depending on the nature of the substituting unit in C-6 position, 6-Man or 2,6-Man residues. A distinctive set of ${}^{3}J_{H-C}$ correlations were clearly observed from H-1 VII at 8 4.86 ppm with C-3 and C-5 VII at δ 71.5 and 74.7 ppm respectively, and with a deshielded C-2 signal at δ 79.5 ppm attributed to either 2- or 2,6-Man residues (not shown). These data suggested the occurrence of few oligomannosyl motifs substituting 6-linked mannan core, presumably in C-2 position.

In order to investigate this issue, KanLM was submitted to complete acetolysis that specifically cleaves the α 1,6-Man linkages. Resulting oligosaccharides were then analysed in MALDI-MS. As shown Fig 4, we observed ions at m/z 989, 1005, 1277 and 1293 that corresponded to [M+Na]⁺ and [M+Na+K]⁺ ions of per-acetylated Hex₃ and Hex₄ oligosaccharides. This result suggested the presence of di- and trimannosidic branches in accordance with NMR data. However, these compounds might also result from an incomplete acetolysis of α (1-6) linkages.

The IV, VI_1 , VI_2 , VII and VIII spins systems were all retrieved in LAM as shown on the anomeric region of ¹³C-¹H heteronuclear NMR spectra of KanLAM (Fig. 2c), establishing that LM and LAM shared a common mannan domain. This was confirmed in HMBC by the identification of extra-residual correlations identical to those observed for KanLM (Fig. 5).

Arabinan domain

From methylation analysis, it was shown that arabinan domain was composed by 3,5-Ara, 5-Ara, 2-Ara and t-Ara units (Fig. 1). By comparing ¹³C-¹H heteronuclear NMR spectra of KanLM and KanLAM (Fig. 3 and 5), and in agreement with previous studies [Gilleron, M. et al., 2000; Guérardel, Y. et al., 2002], anomeric protons of all arabinose types identified by gas chromatography were attributed, and their spin systems determined. Attributed ¹H and ¹³C NMR parameters from KanLAM are summarized in Table III. As already observed from previous studies on mycobacterial LAMs, some Ara units with identical spin systems presented multiple H-1 signals that led to the attribution of one $3,5-\alpha$ -D-Araf unit (I), three 5- α -D-Araf units (II₁ to II₃), two 2- α -D-Araf units (III₁ and III₂) and a single t- β -D-Araf unit (V). Such a multiplicity of signals was attributed to the different positions that each unit may take within the arabinan core [Venisse, A. et al., 1993]. Altogether, NMR parameters derived from arabinan domain of KanLAM, including ${}^{3}J_{H-C}$ correlations determined by HMBC experiment, showed no marked discrepancy with those derived from other mycobacterial species such as M. bovis BCG [Nigou, J. et al., 1997], M. tuberculosis H37Rv [Gilleron, M. et al., 2000], M. smegmatis [Gilleron, M. 1997] and M. chelonae [Guérardel, Y. et al., 2002], suggesting that all these species shared very similar arabinan domains. Briefly, extra residual correlation V H-1/III C-2 at δ 88.4-88.9 ppm established that β -Ara units substituted 2- α -Ara units, while III₁ H-1/I C-5 and III₂ H-1/I C-3 correlations at δ 67.2 and δ 83.1 ppm respectively, established that 2- α -Ara units substituted 3,5- α -Ara units in both C-3 and C-5 positions.

Capping structure

Composition analysis established that KanLAM contained more mannose residues that KanLM (see above). Methylation analysis also established that more 2-Man residues occurred in KanLAM, suggesting the presence of an additional mannan domain in KanLAM compared to KanLM. In accordance with literature, anomeric proton and carbon of 2- α -D-Manp residue (VII) were attributed on anomeric region of ¹H-¹³C heteronuclear spectrum at δ 4.85/99.6 ppm (Fig. 2). Complete NMR parameters of this residue were then determined by multi-relayed COSY and TOCSY experiments and matched perfectly with those previously described for 2- α -D-Manp residues of mannose-capped LAM from *M. tuberculosis* [Gilleron, M. *et al.*, 2000]. However, based on their NMR parameters, we could not distinguish 2- α -D-Manp residues of mannosyl capping from those described in KanLM.

Comparison of HMBC spectra from KanLM and KanLAM showed a correlation between a H-1 signal at δ 4.7 ppm (attributed to VI₁ in KanLM) and an intense signal at δ 69.61 ppm that occurred exclusively in KanLAM but not in KanLM (Fig. 3 and 5). In KanLAM, this correlation occurred in addition to ${}^{3}J_{H-C}$ extra- and intra-residual correlations associated with VI1 residue, demonstrating that two distinct units contained an anomeric proton with identical chemical shifts at δ 4.7 ppm in KanLAM. This hypothesis was reinforced by the fact that relative intensity of signal at δ 4.7 ppm was higher in KanLAM than in KanLM. Two-dimensional ¹H-¹H experiments did not permit to distinguish the two units according to their other protons parameters, which were identical in KanLM and KanLAM. Similarly, two-dimensional ¹³C-¹H experiments showed that their C-1 had identical chemical shifts. Examination of the two dimensional ¹H-¹H NMR spectrum did not reveal any trace of deshielded H-2 (data not shown) indicative of a hexopyranoside substituted in C-2 position, which established that the additional unit was not 2- nor $2,6-\alpha$ -Man residues. This was also confirmed by the absence of internal nOe contact with such H-2 in ROESY experiment. Also, previous studies on ManLAM structure in similar experimental conditions established that all t- α -Manp units from both manno-oligomannosyl caps and from mannan domain presented very similar H-1/C-1 chemical shifts [Gilleron, M. et al., 2000] around 103 ppm. These parameters are largely different from those of the anomeric carbon of the additional unit observed at δ 4.7/100.7 ppm, which strongly suggested that it was not a t- α -Manp unit. Then, the fact that we observed on the HMBC spectrum a single intra-residual correlation with a shielded C-5 at δ 72.6 ppm established that this additional unit was also substituted in C-6 position. Altogether, NMR data suggested that in addition to VI₁ and VI₂ units, a 6- α -Manp unit labelled VI₃ with identical NMR parameters than VI₁ was specifically present in KanLAM.

Signal at δ 69.64 ppm observed in HMBC spectrum was unambiguously attributed from ¹H-¹³C heteronuclear spectrum to an extra-residual correlation with V C-5. Linkage of VI₃ unit with V unit in C-5 position was then confirmed by ROESY experiment (data not shown). Indeed, starting from signal at δ 4,7 ppm we observed four major correlations at δ 3.55, 3.60, 3.70 and 3.84 ppm in ROESY spectrum of KanLAM. Signal at δ 3.70 ppm corresponded to VI₁-VI₃ H-2 internal correlations that systematically occurs in α -mannoses, whereas signals at δ 3.60 and 3.84 ppm were interpreted as VI₁ external correlation with VIII H-6/6', in agreement with the VIII C-6 extra-residual ³J_{H-C} correlation observed in HMBC. Then, remaining signal at δ 3.55 ppm perfectly matched with V H-5, confirming the occurrence of a linkage between some 6- α -D-Manp residues (VI₃) and β -D-Araf residues in C-5 position. Correlation of VI₃ H-1 with V H-5' could not be directly observed due to an overlap with VI₁-VI₃ H-2 signal. As observed for HMBC spectra, ROESY contact with V H-5 was only present in KanLAM and not KanLM. Thus, these data suggested that another type of 6- α -D-Manp labelled VI₃ substituted some β -D-Araf residue in C-5 position and may be part of a new capping motif.

On HMBC spectrum, another correlation was established between VII H-1 and V C-5 suggesting the presence of oligomannosyl caps constituted by 2-linked mannose residues along with eventual 6-linked mannosyl capping as described above. No correlation was observed between IV H-1 and V C-5, suggesting either the absence of mono-mannosyl capping or that they occurred in too minute amount to be observed.

In order to precise the actual status of capping in KanLAM, we liberated eventual capping motives from arabinan side chains by mild hydrolysis, using the highest susceptibility of arabinose linkage to hydrolysis as previously described [Venisse, A. *et al.*, 1993]. Cleaved oligosaccharides were reduced, per-methylated and analysed in GC/EI-MS. As shown Fig. 6, four permethylated oligosaccharides were identified from the hydrolysate. Major species were identified according to their fragmentation patterns as Hex-Pent-ol, Hex-Hex-Pent-ol and Hex-Hex-Hex-Ara-ol, which indicated that mono-, di- and tri-mannosyl units substituted

arabinan domain. In agreement with previous works on mannose capping from *M. tuberculosis* and *M. bovis* BCG, di-mannosylated species was the most abundant, while trimannoside species was barely observable. The presence of mono-mannosyl unit confirmed that $t -\alpha$ -Manp (IV) unit substituted β -Araf unit, so the absence of observable HMBC correlation between these two units was attributed to the too small amount of this motif. EI/MS fragmentation pattern of the mono- and dimannosyl motif permitted to confirm that the arabinose residue was substituted in C-5 position owing to observation of ions resulting from RCHOMe-RCHOMe cleavages within the arabinitol residue. However, this did not allow us to unambiguously discriminate between α -1,6 or α -1,2 linked mannose residues. At most, we observed from Man-Man-Ara-ol motif a fragment ion at m/z 456 corresponding to MeO⁺=CH-O-Man-O-Ara-ol which indicated that the Man residue attached to Ara-ol residue was not substituted in C-3 position.

Altogether, the data unambiguously established that KanLAM belongs to the ManLAM class. Considering the NMR evidences we propose the existence of a -6)Man(α 1-5)Ara(β 1-motives at the non reducing end of some arabinan chains. These may be part of a new oligomannosyl capping motif in KanLAM. However further analyses are needed to confirm this hypothesis.

Phosphatidyl-myoinositol anchor

Structure of the anchor was first analysed on KanLM by one-dimensional ³¹P and twodimensional ³¹P-¹H HMQC NMR experiments. On one-dimensional ³¹P spectrum, we observed two intense signals at δ 1.6 and 1.9 ppm, tentatively attributed to phosphorus atoms of distinct phosphatidyl-*myo*inositol anchors, as confirmed by two-dimensional NMR experiments (data not shown). ³¹P-¹H HMQC experiment established that both phosphorus were linked to inositol in position C-1 and glycerol in position C-3, in agreement with data from literature [Nigou, J. *et al.*, 1999]. ¹H parameters of glycerols were attributed owing to ³¹P-¹H HMQC-HOHAHA experiment and showed that glycerols deriving from both phosphorus atoms had very similar parameters (Table II). Their very deshielded H-2 proton typified them as 1,2-di-acylated glycerols [Wang, Y. et al., 1995]. On the contrary, the ¹H NMR parameters of the two *myo*inositols showed slightly different parameters. In particular their H-3 chemical shifts differed of 1.6 ppm at δ 71.3 and 72.9 ppm which is in agreement with an acylation position in C-3 of one of the inositols, as previously observed in LAMs from other mycobacterium species [Nigou, J. *et al.*, 1999; Gilleron, M. *et al.*, 2000; Guérardel, Y. *et al.*, 2002]. Altogether, these data established that phosphorus atom at δ 1.6 ppm was part of a phosphatidyl-*myo*inositol anchor containing a di-acylated glycerol and a mono-acylated *myo*inositol, whereas phosphorus atom at δ 1.9 ppm was part of an phosphatidyl *myo*inositol anchor containing a di-acylated glycerol and a mono-acylated *myo*inositol anchor containing a di-acylated glycerol and a non-acylated *myo*inositol. Using nomenclature proposed by Nigou J. and collaborators, we typified the phosphorus atom at δ 1.6 ppm as P1 and the one at δ 1.9 ppm as P3. Integration of both signals showed that P3 was the major form (P3/P1 ratio of 9:1).

Nature of the acyl groups associated with acylglycerols units was then established by liberation of intact acetylated acylglycerols by acetolysis and study by GC/MS in electronic impact mode, as described previously [Nigou, J. *et al.*, 1997]. In agreement with NMR data, this methodology confirmed that KanLM anchor exclusively contained di-acylglycerol, and consequently no mono-acylglycerol. As shown Fig. 7, di-acylglycerol region of TIC gas chromatogram was dominated by only two peaks presenting identical EI/MS spectra. Nature of the substituting fatty acids was easily established by the observation of m/z ions at 239 and 281 typifying hexadecanoyl and nonadecanoyl acids, respectively. This attribution was confirmed by the occurrence of fragment ions resulting from the loss of each acyl appendage: [M-C₁₆] ion at m/z 397, [M-C₁₉] ion at m/z 355 and [M-CH₃COO] ion at m/z 592. On these bases, the two peaks were attributed to the two possible arrangements of fatty acids on the C-1 and C-2 positions of glycerol. These experiments established that phosphatidyl-*myo*inositol of KanLM was substituted exclusively by 1/2-hexadecanoyl-1/2-nonadecanoyl-di-acylglycerol.

Glycosylation patterns of inositols associated to P1 and P3 were defined in KanLM owing to NOESY experiment (Fig. 8). Starting from ¹H signals of both inositols on TOCSY spectrum, we observed in NOESY intense nOe contacts between H-6 Ins(P3), H-6 Ins(P1), H-2 Ins(P3) and H-2 Ins(P1) with four presumably mannose anomeric signals at δ 4.98, 4.99, 5.12 and 5.19 ppm, respectively, suggesting that Ins(P1) and Ins(P3) were equally glycosylated in C-2 and C-6 positions. From signal at δ 4.98 ppm nOe contacts of decreasing intensities with Ins(P3) H-5, H-4 and H-3 were also observed, whereas only contact between signal at δ 4.99 ppm and Ins(P1) H-5 could be observed due to smallest quantity of Ins(P1). Similarly, low intensity nOe contacts with Ins(P3) H-1, H-3, H-4 and H-5 were observed from signal at δ 5.12 and with Ins(P1) H-1 from signal at δ 5.19 ppm, confirming their attachment

to either Ins C-2 or C-6. NOESY spectra also showed other contacts attributed to H-2 of each four units. Then, partial attribution of their spin systems from TOCSY spectrum (H-2,-3 and - 4 for units associated to Ins(P3) and H-2,-3 for those associated to Ins(P1)) confirmed that inositol-substituting units were α -Manp, which were labelled Man1(P1) and Man1(P3) for units linked to C-2 position and Man2(P1) and Man2(P3) for units linked in C-6 position. These data are in total agreement with previous studies on GPI anchors of PIMs, LM and LAM where inositol was invariably substituted at the C2 by a single α -D-Manp residue and at the C6 position by the mannan core [Khoo, K.H. *et al.*, 1995; Chatterjee, D. *et al.*, 1992; Nigou, J. *et al.*, 1997].

Similar methodology was used in order to study the structure of phosphatidylmyoinositol of KanLAM. Results from these experiments were undistinguishable from those obtained with KanLM (Table III), demonstrating that KanLM and KanLAM shared identical lipid anchors.

Substitution by thio-methylpentose residues

Observation of the anomeric region of ¹H-¹³C heteronuclear spectrum of KanLM revealed a last unattributed signal at δ 5.23-5.24/104.2 ppm (Fig. 2b). Close examination of the ¹H signals in one and two-dimensional experiments revealed the presence of two distinct signals with very closely related NMR parameters. The likeness of their parameters established that both signals originated from identical components that occurred in slightly different environments. The entire spin system of this unit was established by twodimensional multi-relayed COSYs and TOCSY ¹H-¹H experiments, as shown Fig. 9 and Fig. 10. Then, chemical shifts of the corresponding carbons were attributed owing to ¹H-¹³C heteronuclear and HMBC spectra (Fig. 10) and summarised Table II. Identification of only five ring carbons and corresponding protons established that this residue was a pentose. Furthermore the important downfield chemical shift of the C-4 claimed for a pentofuranoside unit. However, the very highfield chemical shift of C-5 from about δ 65-75 ppm for a C-5 -CH₂OH of a pentofuranose to δ 34.1 ppm suggested that this carbon would be substituted by another element than oxygen. H-5/5' signals clearly appears as a quadruplet, that established the absence of coupling. Altogether, these data ruled out the possibility of the presence of substututing phosphorus atom. Starting from H-5/5' on HMBC spectrum (Fig. 10), in addition to typical ${}^{3}J_{\text{H-C}}$ and ${}^{2}J_{\text{H-C}}$ correlations with C-3 and C-4 respectively, we observed a correlation with an carbon at δ 16.6 ppm. Comparison with standard chemical shifts from literature revealed that CH₂- γ and -S-CH₃ from methionine present average ${}^{1}\text{H}/{}^{13}\text{C}$ chemical shifts of δ 2.47/32.13 and 1.84/17 ppm in D₂O. This values are in agreement with H-5/C-5 values measured in DMSO of the pentofuranoside unit and with additional carbon observed in HMBC, respectively. On this basis, signal observed in HMBC at δ 16.6 ppm was interpreted as a ${}^{3}J_{\text{H-C}}$ correlation between H-5/5' and carbon atom of substituting -S-CH3 group. Altogether, these data suggested that the additional pentafuranose unit observed in LM would be a 5-methylthiopentose (5-MTP). Furthermore, its ring carbon and proton parameters matched with those of 5-methylthiopentose previously observed in minute amount in LAM isolated from *M. tuberculosis*, owing to a ${}^{13}\text{C}$ enrichment [Treumann, A. *et al.*, 2002], which confirmed its attribution as 5-MTP.

As observed Fig. 2, signals from anomeric proton of 5-MTP occurred both in KanLM and KanLAM, but not in KanPIM, suggesting that 5-MTP would be localized on the mannan domain of both KanLM and KanLAM. However, due to a slight contamination of KanLM by KanLAM, we could not rule out the possibility that the 5-MTP observed in KanLM would originate from KanLAM contamination. Upon visual examination of the KanLM spectra, it clearly appeared that 5-MTP signal had a similar intensity than the most abundant contaminating arabinose type (I) and a higher intensity than the other arabinose units, rendering highly unlikely that it would be associated to KanLAM. This was confirmed by the integration of ¹H signals on one-dimensional ¹H spectra of 5-MTP H-1 and of the well individualized Gro H-1s. Each KanLM and KanLAM molecule containing a single Gro H-1, calculation of the ratio 5-MTP H-1 to Gro H-1 established that KanLM contained a average of 0.9 5-MTP residue per molecule, whereas KanLAM contained only an average of 0.5 5-MTP residue. This lower value for KanLAM was confirmed by the calculation of 5-MTP H-5 to H-1 Gro ratio that gave an identical result. Altogether, these data clearly established that 5-MTP substituted KanLM, presumably on mannan domain, as well as KanLAM.

As shown on HMBC spectra (Fig. 10), 5-MTP H-1 established a set of four correlations, three of which were attributed to intra-residual ${}^{2}J_{\text{H-C}}$ and ${}^{3}J_{\text{H-C}}$ correlations with its own C-2, C-3 and C-4 at δ 78, 76.4 and 80 ppm, respectively. The fourth signal was assigned to an extra-residual correlation with a carbon easily identified on the ${}^{1}\text{H-}{}^{13}\text{C}$ heteronuclear spectrum owing to its distinctive ${}^{13}\text{C}/{}^{1}\text{H}$ chemical shift at δ 75.5/3.67 ppm (Fig.

3). The linkage of 5-TMP to this carbon was confirmed in a ROESY experiment owing to the observation of a nOe contact between 5-MTP H-1 and a proton at δ 3.67 ppm. A second nOe contact at δ 4.917/4.898 ppm was attributed to its own H-2, whereas a third contact at δ 3.723/3.693 ppm still remains unattributed.

The chemical shift of the linkage carbon at δ 75.5 ppm clearly discarded the possibility that the 5-MTP would be linked to an Ara residue. This was in agreement with previous observations that stated that 5-MTP was linked to mannan domain of KanLM. Similarly, in accordance with data from lipoglycans studies and litterature, this carbon was far too downfield shifted to be a C-6 mannose substituted, and its associated proton too highfield shifted to be associated with a C-2 mannose substituted [Gilleron, M. et al. 2000; Bock, K. & Pedersen, C., 1983; , Vliegenthart, J.F.G. et al., 1983]. Therefore, we concluded that 5-MTP unit was necessarily linked to a mannose unit either in C-3 or in C-4 position. ¹³C/¹H chemical shifts of unsubstituted mannose C-3 and C-4 of KanLM and KanLAM in DMSO were observed at δ 71.8/3.58 and 68.2/3.47 ppm respectively. Considering that substitution of C-3 carbon typically has very little influence on H-3 chemical shift value, we could not unambiguously assign the carbon observed at δ 75.5/3.67 ppm to substituted mannose C-3 or to substituted mannose C-4. As shown on HMBC and HMQC spectra of KanLAM, an identical correlation with a carbon at δ 75.5/3.67 ppm was observed from 5-MTP H-1, establishing that 5-MTP residues in KanLM and in KanLAM shared identical locations. Altogether, these data demonstrated that some KanLM and KanLAM molecules are substituted within their respective mannan domain by a 5-methylthiopentose residue on a mannose unit in C-3 or C-4 position.

Substitution by succinic acid

Presence of succinic acid in KanLM and KanLAM was assessed by standard gas chromatography technique. Ester linked substituents were released from both compounds by methanolysis and analysed in gas chromatography. This revealed that only KanLAM comprised significant amounts of succinic acid, evaluated owing to the use of adipic acid as internal standard to an average of two succinic acid per molecule of KanLAM (Fig. 11).

 13 C-¹H heteronuclear HMQC spectrum of KanLAM revealed the presence of a broad signal at δ 2.5-2.6/30.5 ppm (Fig. 12), that was tentatively assigned to the two distinct

methylene groups of succinic acid, according to data from literature [Delmas, C. *et al.*, 1997]. However, our data differed from previous by the fact that carbons from both groups exhibited identical chemical shifts. This was attributed to the effect of different solvents (DMSO *versus* D₂O), which was confirmed by repeating the experience in D₂O. Experiments in DMSO did not permit to differentiate the two methylenic protons while in D₂O protons of the two methylene groups were unambiguously assigned at δ 2.51 and 2.66 ppm (data not shown) [Delmas, C. *et al.*, 1997]. From signal at δ 2.51 ppm, ¹³C-¹H heteronuclear HMBC experiment in D₂O revealed a ³J_{H-C} correlation at δ 176.06 ppm and a ²J_{H-C} correlation at δ 181.7 ppm attributed according to literature to carbons of ester and carboxyl groups, respectively. Inversely, proton at δ 2.66 ppm established ³J_{H-C} correlation with carboxyl group carbon at δ 181.7 ppm and ²J_{H-C} correlation with ester group at δ 176.06 ppm. These data definitively established the attribution of protons at δ 2.51 and 2.66 ppm as methylenic protons of succinic acid.

From ester group carbon at δ 176.06 ppm, no additional correlation was visible in D₂O, which prevented us to localize the ester position on the KanLAM molecule in these experimental conditions. Despite the fact that methylenic protons were not resolved in DMSO, HMBC experiment showed that they correlated with two intense signals at δ 172.4 and 173.6 ppm, assigned by comparison with data obtained in D₂O to ester and carboxyl groups, respectively (Fig. 12). From signal at δ 172.4 ppm, a single additional correlation was observed with a proton at δ 4.83 ppm. On ¹³C-¹H heteronuclear HMQC spectrum, we observed a corresponding signal at δ 80.5/4.84 ppm that was tentatively attributed to a non anomeric carbon bearing the succinyl group. This remained speculative since 4.8-4.9 ppm region is fairly crowded. However, this hypothesis was in accordance with previous study which demonstrated that succinyl-substituted carbon was up-shifted to δ 82.2 ppm [Delmas, C. *et al.*, 1997].

From proton at δ 4.84 ppm, the complete proton spin system of the monosaccharide was assigned owing to ¹H-¹H homonuclear COSY 90, single and double relayed COSY experiments. On this basis, the proton at δ 4.84 ppm was unambiguously attributed as a H-3 of a C-5 substituted Araf residue. On ¹H-¹H homonuclear COSY 90 spectrum, proton at δ 4.84 ppm correlated with two protons at δ 4.02 and 4.15 ppm attributed to H-2 and H-4, respectively (Fig. 9). The absence of correlation with an anomeric proton in COSY 90 experiment clearly established that proton at δ 4.84 ppm was not a H-2. This was confirmed by the apparition of an additional correlation with an anomeric proton at δ 4.96 ppm in single relayed COSY experiment, which suggested that proton at δ 4.84 ppm was a H-3. Nature of the H-4 at δ 4.13 ppm was confirmed by the observation in COSY 90 of cross-peaks with H-5 and H-5' at δ 3.75 and 3.59 ppm, respectively. ¹H parameters of H-5/H-5' established that Ara*f* residue was substituted in C-5 position. Furthermore, H-5' correlated with proton at δ 4.84 ppm in single relayed COSY experiment, but not in COSY 90, confirming that this proton was a H-3. So, ¹H NMR parameters of this residue were established as: H-1, δ 4.96 ppm; H-2, 4.01; H-3, 4.84; H-4, 4.13; H-5/H-5', 3.75/3.59. These data demonstrated that succinyl groups substituted some 5-Ara*f* residues of the arabinan core of KanLAM in C-3 position.

References

- Bock, K. and Pedersen, C. (1983) Carbon-13 nuclear magnetic resonance spectroscopy of monosaccharides. In Adv. Carbohydr. Chem. Biochem., 41, 27-66
- Chatterjee, D., Hunter, S.W., Mc Neil, M. and Brennan, P. (1992) Lipoarabinomannan. Multiglycosylated form of the mycobacterial mannosylphosphatidylinositols. J. Biol. Chem. 267, 6228-6239
- Gilleron, M., Bala, L., Brando, T., Vercellone, A. & Puzo, G. (2000). *Mycobacterium tuberculosis* parietal and cellular lipoarabinomannans. J. Biol. Chem. 275, 677-684.
- Gilleron, M., Himoudi, N., Adam, O., Constant, P., Venisse, A., Rivière, M. & Puzo, G. (1997). *Mycobacterium smegmatis* phosphoinositols-glyceroarabinomannans. Structure and localization of alkali-labile and alkali-stable phosphoinositides. J. Biol. Chem. 272, 117-124.
- Gilleron, M., Nigou, J., Cahuzac, B. & Puzo, G. (1999). Structural study of the lipomannans from *Mycobacterium bovis* BCG: characterisation of the multiacylated forms of the phosphatidyl-*myo*-inositol anchor. J. Mol. Biol. 285, 2147-2160.
- Khoo, K.H., Dell, A., Morris, H.R., Brennan, P.J. and Chatterjee, D. (1995) Structural definition of acylated phosphatidyl inositol mannosides from *Mycobacterium tuberculosis*: definition of a common anchor for lipomannan and lipoarabinomannan. *Glycobiology* 5, 117-127
- Khoo, K.H., Dell, A., Morris, H.R., Brennan, P.J. and Chatterjee, D. (1995) Structural definition of acylated phosphatidyl inositol mannosides from *Mycobacterium*

tuberculosis: definition of a common anchor for lipomannan and lipoarabinomannan. Glycobiology 5, 117-127

- Nigou, J., Gilleron, M. and Puzo, G. (1999) Lipoarabinommannan: characterization of the multiacylated forms of the phosphatidyl-myo-inositol anchor by NMR spectroscopy. Biochem. J. 337, 453-460
- Nigou, J., Gilleron, M., Cahuzac, B., Bounéri, J. D., Herold, M., Thurnher, M. & Puzo, G. (1997). The phospho-myo-inositol anchor of the lipomannans from Mycobacterium bovis bacillus Calmette Guérin. J. Biol. Chem. 272, 23094-23103.
- Venisse, A. Berjeaud, J.M., Chaurand, P., Gilleron, M. and Puzo, G. (1993) Structural features of lipoarabinomannan from *Mycobacterium bovis* BCG. J. Biol. Chem. 268, 12401-12411
- Vliegenthart, J.F.G., Dorland, L and van Halbeek, H. (1983) High-resolution, ¹H-nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. In *Adv. Carbohydr. Chem. Biochem.* 41, 209-374
- Wang, Y & Hollingworth, R. I. (1995). A solvent system for the high-resolution proton nuclear magnetic resonance spectroscopy of membrane lipids. *Anal. Biochem.* 225, 242-251





Fig. 1: Composition analysis of partially methylated and acetylated reduced monosaccharides liberated from KanLM and KanLAM by hydrolysis. FID gas-chromatography chromatograms from a) KanLM and b) KanLAM



Fig. 2: Anomer regions of ¹³C-¹H heteronuclear spectra of a) KanPIM, b) KanLM and c) KanLAM. * stands for Ara residues.



Fig 3.: Details of ¹³C-¹H heteronuclear HMQC and HMBC, and of ¹H-¹H ROESY spectra from KanLM. VIII H-1/VI C-6,6' correlation as well as all signals from VII are not apparent on the HMBC spectrum because of the used processing and/or too small quantities but are indicated nevertheless. • arabinosyl signal, *non sugar signal



Fig. 4: MALDI analysis of acetolysis products from KanLM.



Fig.5: Details of ¹³C-¹H heteronuclear HMQC (left) and HMBC (right) spectra from KanLAM.



Fig. 6: GC/EI-MS analysis of KanLAM hydrolysed, reduced and permethylated. a) TIC chromatogram is dominated by four peaks attributed to I, Ara-Ara-ol, II, Man-Ara-ol, III, Man2-Ara-ol, IV, Man3-Ara-ol. EI-MS fragmentation patterns and attribution of the major ions for b) peak II and c) peak III. Compound III is given with an 1,6-Man linkage according to NMR data, but the exact position of this linkage was not unambiguously confirmed.



Fig. 7: GC/MS analysis of acetolyseproducts of KanLM. a) Total Ion Count of acetylated diacylglycerol after acetolysis of KanLM. b) Fragmentation pattern of peak 1 in EI mode, characteristic of a C16-C19 acetylated diacylglycerol. Peak 2 presented an identical fragmentation pattern. c) Deduced fragmentation pattern of peaks 1 and 2 in EI mode.







Fig. 9: ¹H-¹H homonuclear COSY 90 (in black) and relayed R1 COSY (in red) spectra of KanLAM. Only spin systems from succinvlated arabinose (Ara) and thiomethyl-pentose (TMP) are assigned.



Fig. 10: Details of a) TOCSY spectrum, b) and c) HMBC, d) HMQC from KanLM showing entire 1 H/ 13 C spin system of thiomethylpentose residue. Red arrows, ${}^{3}J_{H-C}$ HMBC contacts; blue arrows, ${}^{2}J_{H-C}$ HMBC contacts.

209



Fig. 11: Succinic acid content of KanLAM. 50 μ g of Kan LAM were treated with 2% sulfuric methanol, 80°C, 30 mn. Reaction was stopped by water, and methylated organic acids were extracted by chloroform and injected through splitless injector on a capillary CARBOWAX column. 2 μ g of adipc acid were used as internal standards. Response factor of succinic to adipic acid was determined: RF=0.64.



Fig. 12: Expanded regions of ¹³C-¹H heteronuclear a) HMQC, b) c) d) HMBC spectra of KanLAM.

211

	mol %		
	KanLM	KanLAM	
t-Ara		3	
2-Ara		6	
5-Ara	4*	38	
3,5-Ara		9	
t-Man	44	21	
2-Man	4	6	
6-Man	19	7	
2,6-Man	29	10	

Table I: Methylation analysis of CheLM and CheLAM. Values of molar ratio were obtained from peaks area integration of GC analysis and corrected by the use of effective carbon-response factors. * originates from KanLAM contamination.

			Chemical s	hifts (ppm))				
			1	2	3	4	5	6	6'
t-α-Manp IV		¹ H	4,91	3,80	3,58	3,47	3,52	3,68	3,54
		¹³ C	103,2	71,2	71,8	68,2	74,2	62	,3
6-α-Manp VI₁		1H	4,70	3,75	3,57	3,49	3,52	3,76	3,65
		¹³ C	100,6	71,4	72,3	68,1	72,4	66	,5
VI ₂		1H	4,78	3,74	3,56	n.d.	n.d.	n.	d.
		¹³ C	100,9	71,2	72,3	n.d.	72,4	66	,5
2,6-α-Manp VIII		¹ H	4,89	3,80	3,69	3,49	3,62	3,81	3,53
		¹³ C	99,5	79,4	71,9	68,1	72,4	67/67,3	
TMP		¹ H	5,24/5,23	3,92/3,91	4,02/3,99	4,21/4,18	2,69/2,51	2,	11
		¹³ C	104,2	78,0	76,4	80,0	34,1	16	,6
α- Manp -1	¹ H P1 ¹³ (¹ Η	4,99	3,96	n.d.	n.d.	n.d.	n.d.	n.d.
		¹³ C	102,6	n.d.	n.d.	n.d.	n.d.	n.d.	
	P3 ¹ H	¹ H	4,98	3,92	3,60	3,54	3,51	n.d.	n.d.
		¹³ C	102,5	71,0	n.d.	n.d.	n.d.	n.	d.
α- Manp-2	1 P1 13	¹ H	5,19	3,80	n.d.	n.d.	n.d.	n.d.	n.d.
		¹³ C	101,01	n.d.	n.d.	n.đ.	n.d.	n.	d.
	P3 13	Ή	5,12	3,78	3,59	3,50	3,50	n.d.	n.d.
		¹³ C	101,5	71,8	70,8	n.d.	71,4	n.	d.
Ins	¹ P1 ₁₃	¹ H -	4,09	4,19	4,59	n.d.	3,16	3,	64
		¹³ C	n.d.	n.d.	72,9	n.d.	n.d.	n.	d.
	P3 ¹ †	¹ H	4,01	4,15	3,24	3,43	3,08	3,	61
		¹³ C	76,3	78,4	71,3	74,5	74,8	80	,3
			1 1'	2	3,3'				
-	P1 ¹ H	¹ H	4,35/4,12	5,11	3,85	-			
		¹³ C	63,5	71,6	63,7				

5,11

71,6

3,85

63,7

4,35/4,12

63,5

¹H

¹³C

Gro

P3

Table II: KanLM ¹H and ¹³C chemical shcifts in DMSO-d6 at 343°K

ģ

	Chemical shifts		······································			
	1	2	3	4	5	6
3.5-x-Arafi ¹ H	4.87	4.02	3.91	3.87	3.58/3.69	-
5,5-α-Arar 13C	109.3	81.3	83.1	82.7	67.2	-
¹ Η 5-α-Araf II ₁ ¹³ C ¹ Η II ₂ ¹ 3C	4.81	3.85	3.69	3.9	3.6/3.7	-
	109.1	82.7	78.5	83	68	-
	4.83	3.87	3.73	3.9	3.6	-
	109.1	82.7	78.5	83	68	-
" ¹ H	4.97	3.86	3.7	3.9	3.6	-
^{II3} 1 ³ C ¹ H 2-α-Araf III ₁ 1 ³ C ¹ H III ₂ 1 ³ C	108.1	83	78.3	83.5	68	-
	4.97	3.99	3.88	3.84	3.63	-
	106.8	88.4	76.3	84.6	~62	-
	5.10	3.98	3.8	3.9	~3.6	-
	106.2	88.9	76.5	84	~62	-
	4.91	3.78	3.51	3.42	3.5	3.70/3.51
	103.1	71.1	71.8	68.4	74.7	62.4
	4.85	3.73	3.7	3.45	3.45	~3.7/3.5
$2-\alpha$ -iviality vir 13C	99.6	78.4	71.5	68.4	74.6	~62
¹ H	4.93	3.85	3.81	3.78	3.70/3.56	-
β-Arat V 13C	101.9	77.9	76.5	81.8	~70	-
¹ H	4.70	3.78	3.54	3.45	3.60	3.76/3.63
6-α-Manp VI ₁ /VI _{3 13} C	101.0	71.6	72.5	68.2	72.6	66.7
¹ H VI ₂ 13C	4.79	3.74	n.d.	n.d.	n.d.	n.d.
	100.7	n.d.	72	n.d.	n.d.	n.d.
1 H 2,6- $lpha$ -Manp VIII 13 C	4.88	3.75	3.67	3.5	3.5	3.82/3.53
	99.5	78.4	71.6	68.2	72	67.03
¹ Η α-Manp 2 P1 ₁₃ C	5.19	3.79	n.d.	n.d.	n.d.	n.d.
	100.6	n.d.	n.d.	n.d.	n.d.	n.d.
¹ H	5.12	3.78	n.d.	n.d.	n.d.	n.d.
α-Manp 2 P3 13C	101.15	n.d.	n.d.	n.d.	n.d.	n.d.
α-Manp 1 P1/ P3 ¹ H ₁₃ C Ins P1 ¹ H ¹³ C	5.0/5.0	3.94/3.91	n.d.	n.d.	n.d.	n.d.
	102.1	70.7	n.d.	n.d.	n.d.	n.d.
	n.d.	4.2	4.6	n.d.	3.16	n.d.
	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
¹ H Ins P3 ₁₃ C	4,00	4.16	3.24	3.44	3.09	n.d.
	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
1 H TMP 13 C Ara succinylé en 1 H 3 1 C 1 A 1 3 13 C 1 C 3 13	5.23	3.90	3.99	4.18	2.70/2.53	2.1
	103.9	78	76.3	79.7	34.4	16.5
	4.96	4.01	4.84	4.13	3.75/3.59	-
	; 108.1	79.7	80.5	82.09	67.3	-
	1 1'	2	3,3'			
Gro P1 ¹ H 13	4.34/4.01	5.13	3.9			
	63.6	71.8	63.9			
¹ H Gro P3 ¹³ (4.36/4.14	5.13	3.86			
	63.6	71.8	63.9			

Table III: KanLAM ¹H and ¹³C chemical shcifts in DMSO-d6 at 343°K

14. Conclusions

Ces travaux représentent la première tentative d'analyse des lipoglycannes de deux espèces mycobactériennes pathogènes, M. chelonae et M. kansasii. Ils ont tout d'abord 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 myoinositol, 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 thiométhyle-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. Enfin, des résultats encore préliminaires suggèrent 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'a-1,2-Man, motifs normalement cantonnés à la coiffe oligomannosidique.

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 dont la structure exacte reste à définir, 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.

L'ensemble de ces résultats démontre qu'entre espèces de mycobactéries existent des variations discrètes dans l'activité des enzymes responsables de la synthèse des lipoglycannes. Ainsi, il est probable que l' α -1,3-mannosyltransférase impliquée de la synthèse du mannanne chez *M. chelonae* résulte de la mutation de l' α -1,2-mannosyltransférase présente chez les autres espèces de mycobactéries. Cette hypothèse n'est pour l'instant pas vérifiable, en cause l'absence de données accessibles sur ces enzymes chez les mycobactéries. Il est à noter que chez les eucaryotes, des études phylogénétiques ont démontré que les α -1,2- et α -1,3-mannosyltransférases utilisant le Dol-P-Man comme substrat donneur faisaient parties de deux super-familles distinctes, dont les gènes ont divergé très tôt [Oriol, R. *et al.*, 2002]. Ceci rend très improbable le passage d'une activité à l'autre par simple mutation d'une enzyme, chez les eucaryotes.

Comme nous l'avons déjà exposé dans la partie consacrée aux O-glycannes, les glycannes sont des molécules intrinsèquement polymorphes, et il est important de distinguer les formes de glycosylation dites 'futiles', dont les variations structurelles ne vont pas modifier la fonction, des formes de glycosylation spécifiques, dont la moindre variation va se répercuter sur la fonctionnalité. Néanmoins, ces deux formes de glycosylation cohabitent souvent au sein de la même structure, un même composé ayant souvent de nombreuses fonctions différentes, plus ou moins indépendantes. Ceci est vrai pour les O-glycannes, dont la simple présence, en grande quantité et quelle que soit leur structure, est indispensable aux propriétés rhéologiques des mucines, mais dont les séquences terminales non réductrices sont également impliquées dans des interactions spécifiques. Comme nous le verrons dans la partie suivante, ceci est vrai pour les polysaccharides d'éponges responsables de l'auto-agrégation des cellules d'éponges et dont seuls quelques épitopes précis semblent impliqués dans leur fonctionnalité. Ceci est également vrai pour les lipoglycannes qui jouent probablement un double rôle au sein des parois mycobactériennes, à la fois d'ordre structurel et comme facteur

de virulence. Plusieurs études ont démontré que la structure des coiffes des LAMs régulait directement leurs fonctions immunomodulatrices, tandis qu'un nombre minimum d'acides gras était essentiel à certaines de ces fonctions, établissant que les variations structurales de certains domaines avait des répercutions importantes sur leur fonctionnalité. Dans une deuxième étape de notre travail, nous avons donc tenté d'établir si les différences structurales que nous avons mises en évidence dans les lipoglycannes de *M. chelonae* et *M. kansasii* étaient de nature à modifier leurs fonctions.

2. Etude des propriétés immunomodulatrices des lipoglycanes de *M. chelonae* et *M. kansasii*

21. Présentation des résultats

Les propriétés immunomodulatrices, en particulier les propriétés d'induction des cytokines, sont les activités biologiques associées aux lipoglycannes de mycobactéries les mieux étudiées (voir chapitre III 4.). En accord avec une étude précédante qui suggérait que le potentiel d'induction de cytokines pro-inflammatoires était liée à la présence de coiffes de type inositol-phosphate [Gilleron, M. et al. 1997], nos travaux ont montré que les LAMs de M. chelonae et M. kansasii n'induisaient pas la production d'IL-8 et de TNF- α à partir d'une lignée leucémique promonocytaire humaine THP-1 différenciée vers un stade macrophagique [Guérardel, Y. et al., 2002; Guérardel, Y. et al., résultats non publiés]. Par contre, des résultats préliminaires indiquaient que les LMs isolés de ces deux espèces induisaient une importante production de ces deux cytokines, ce qui contredisait le modèle actuel selon lequel les lipomannannes seraient des molécules biologiquement inactives. L'examen de la littérature a révélé qu'aucune étude ne s'était véritablement penchée sur ce point précis, ce qui nous empéchait de valider ces résultats. En collaboration avec le Dr Elass nous avons donc exploré le nouveau concept selon lequel les lipomannannes présentent des propriétés immunologiques distinctes de celles des lipoarabinomannannes et mis en évidence plusieurs acteurs moléculaires indispensables à leurs activités.

22. Résultats

Les résultats de cette étude sont présentés sous forme d'un article en préparation: Article 5

Lipomannans, but not lipoarabinomannans, purified from *Mycobacterium* chelonae and *Mycobacterium kansasii* induce TNF-α and IL-8 secretion by a CD14-TLR2 dependent mechanism.

¹Vignal C., ¹Guerardel Y., ²Kremer L., ¹Masson M., ¹Legrand D., ¹Mazurier J. and ¹Elass E[§].

1-Laboratoire "Glycobiologie structurale et fonctionnelle", Unité Mixte de Recherche n°8576 du Centre National de la Recherche Scientifique, Université des Sciences et Technologies de Lille, 59 655 Villeneuve d'Ascq cedex, France. INSERM U447, Institut Pasteur de Lille/IBL, Lille, France.

[§]Corresponding author. E-mail: Elisabeth._Elass_@univ-lille1.fr

Key words: Lipomannan, lipoarabinomannan, CD14, TLR2, inflammation, cytokines, mycobacteria.

ABSTRACT

Lipoarabinomannans (LAMs) are complex glycolipids from the mycobacterial cell wall that exhibit a vast panel of biological activities, including pro- and anti-inflammatory properties. However, little is known about the immunomodulation activities of lipomannans (LMs), known to be precursor of LAMs.

In this study, we provide evidence that LMs purified from several pathogenic and non pathogenic mycobaterial species, including *Mycobacterium chelonae* and a clinical strain of *Mycobacterium kansasii*, stimulate mRNA expression and secretion of TNF- α and IL-8 from human macrophage-like differentiated THP-1 cells. In contrast to LMs, LAMs were not able to induce significant cytokine-inducing effect. After differentiation by 1,25-dihydroxy-vitamin D3, the THP-1 cells express at the cell surface mCD14, a receptor for multiple bacterial products. The mechanism of activation of differentiated cells by LMs was investigated using different antibodies raised against surface receptors. Incubation of THP-1 cells with specific anti-CD14 (My–4) or anti-Toll-like receptor 2 (TLR2) antibodies profoundly affected production TNF- α and IL-8, suggesting that CD14 and TLR2 participate in the LM-mediated activation process. Furthermore our experiments revealed that the stimulation of differentiated cells is highly dependent on the presence of the LPS-binding protein (LBP), a plasma protein which was shown to transfer glycolipids to mCD14. Furthermore, we have detected, by native gel electrophoresis assays, the formation of ternary complexes between soluble recombinant CD14, LBP and LMs.

We have also shown that chemical degradation of the arabinan domain from the ManLAM from *M. kansasii*, that presented no cytokine-eliciting effect, restored the cytokine-inducing activity level of LMs. These results support the hypothesis that the presence of an arabinan, and eventually terminal mannan domain, abrogates the capacity of LAMs to stimulate cytokine secretion. We have shown that phosphatidylinositol dimannosides (PIM₂) isolated from *M. kansasii* did not induce cytokine secretion. Altogether, this study suggests that LMs isolated from various mycobacterial species may be potentially involved in the immunomodulation of the infected host and suggests that the D-mannan core of this glycolipid is essential for this function.

INTRODUCTION

Lipoarabinomannans (LAMs) and lipomannans (LMs) are complex lipoglycans considered as major constituants of mycobacterial cell wall (Daffe *et al.*, 1998; Chatterjee *et al.*, 1997). These glycoconjugates are virulence factors which play a key role in human immune system through their interactions with various cells. (Chatterjee and Khoo 1998; Nigou *et al.* 2002). Indeed, they exhibit a wide variety of biological activities which enhance antimycobacterial immune defenses or facilitate mycobacterial survival through inhibition of the immune response. These effects include regulation of the pro- and anti-inflammatory cytokines production (Chatterjee *et al.*, 1992; Barnes *et al.* 1992; Zhang *et al.*, 1993; Zhang *et al.*, 1995; Yoshida *et al.*, 1997; Nigou *et al.*, 2001), inhibition of microbicidal activity of macrophages (Sibley *et al.* 1988) and suppression of T-lymphocytes proliferation (Moreno *et al.*, 1988).

So far, studies have focused on LAMs isolated from pathogenic species such as M. *tuberculosis, M. leprae, M. bovis* BCG or from fast-growing and non-virulent mycobacteria such as M. *smegmatis*. However, the biological properties of these glycolipids depend on their structural features (Chatterjee *et al.*, 1992; Khoo *et al.*, 1995a). LAM is composed by a phosphatidyl *myo*inositol anchor to which is attached a D-mannan core branched by single mannopyrannosyls residues followed by a D-arabinan domain (Hunter and Brennan, 1990; Venisse *et al.*, 1993; Khoo *et al.*, 1995b). LM, which is considered as being a direct precursor of LAM lacks the arabinan domain. According to the structure of caps located at the terminal extremity of arabinan domain, LAMs were classified into two types: _mannosylated LAMs (Man-LAMs) which are characterized by the presence of $\alpha(1-2)$ manno-oligosaccharides as found in *M. tuberculosis*, *M. leprae*, *M. bovis* BCG and *M. avium* and phosphoinositol-capped LAMs (PILAM) identified in *M. smegmatis* (Chatterjee *et al.*, 1992; Khoo *et al.*, 1995a; Gilleron *et al.*, 1997; Venisse *et al.*, 1993; Khoo *et al.*, 1995; manno-oligosaccharides as found in *M. tuberculosis*, *M. leprae*, *M. bovis* BCG and *M. avium* and phosphoinositol-capped LAMs (PILAM) identified in *M. smegmatis* (Chatterjee *et al.*, 1992; Khoo *et al.*, 1995a; Gilleron *et al.*, 1997; Venisse *et al.*, 1993; Khoo *et al.*, 2001). These glycolipids are also heterogenous in their acylation state and branching patterns of both arabinan and mannan domains.

It has been shown that purified PILAM activates macrophages in a Toll-like receptor (TLR2) and CD14 dependent manner leading to the production of TNF- α , IL-8 and IL-12 (Zhang *et al.*, 1993; Savedra *et al.*, 1996; Yu *et al.*, 1998; Means *et al.*, 1999; Orr and Tobias 2000; Heldwein and Fenton, 2002). In contrast, ManLAM failed to induce these cytokines. Moreover, Ma-LAM from *M. tuberculosis* inhibited TNF- α and IL-12 productions in human

mononuclear phagocytes through its binding to the mannose receptor (Venisse *et al.* 1995; Knutson *et al.*, 1998; Nigou *et al.*, 2001). However, despite the relative abundance of LM in the mycobacterial cell wall, the biological activities of LMs have been poorly reported.

In this context, a further insight in the relationship between structure and possible properties of LMs was needed. As experimental models, we used LMs isolated from two pathogen mycobacteria, M. chelonae and M. kansasii, commonly found in soil and water. M. kansasii is involved in serious pulmonary diseases in patients with immuno deficiencies and systemic illness, particularly those infected with HIV (Rafael et al., 1997; Pintado et al., 1999; Jacobson, 2000). M. chelonae infections are more frequently associated with cutaneous lesions in patients who are immunosuppressed by malignancy, corticosteroid therapy or drugs (Wallace et al., 1992; Ingram et al., 1993). LMs are six-fold (mol./mol.) more abundant in the cell wall of *M. chelonae* than LAM whereas the levels of LAM and LM in *M. kansasii* are similar. We have recently purified and determined the structures of both LM and LAM from M. chelonae (Guérardel et al., 2002) and from a clinical strain of M. kansasii isolated from a HIV⁺-postive patient; (Guérardel et al., in submission). These structures displayed significant discrepancies regarding to those of LMs and LAMs from other mycobacterial species. In particular, we found that the arabinan moiety of LAM from M. chelonae (CheLAM) was neither capped with oligomannosyl motives nor with phosphatidyl-myoinositol, thus defining a novel class of LAM. In addition to the absence of capping, the mannan domain was characterized by the presence of unusual α -1,3-mannosyl side chains instead of α -1,2 commonly found, in all other mycobacterial species analyzed to date. Lastly, in contrast to lipoglycans from other species that only contain two major types of fatty acids (C_{16} and C_{19}) [Nigou, et al., 1997; Gilleron et al., 1997; Gilleron et al., 2000], the phosphatidyl-myoinositol anchor of both CheLAM and CheLM showed a wide lipidic heterogeneity (Guérardel et al., 2002). From a biological point of view, CheLAM did not induce secretion of TNF-α and IL-8 from THP-1 cells. The recent structural determination of LAM from M. kansasii (KanLAM) has shown that it belongs to the general Man-LAM family of molecules, although it contains subtle differences compared to ManLAMs from M. tuberculosis or M. bovis BCG (Guérardel et al., in submission). For instance, the mannan domains of KanLAM and KanLM were characterized by the presence of an unusual thiomethyl-pentose residue.

In the present study, we have investigated the ability of LMs purified from M. chelonae and M. kansasii to stimulate the TNF- α and IL-8 secretion and mRNA expression in the human monocytic cell line THP-1. Cells were first differentiated with 1,25-dihydroxyvitamin D3, for expressing at the cell surface, the glycosylphosphatidyl-inositol anchored protein CD14 (mCD14), a pattern recognition receptor for multiple microbial products. Availability of LMs with defined structural variations prompted us to study their structure/activity relationships. Assays were performed with LMs from four species (*M. chelonae, M. kansasii, M. bovis* BCG, *M. smegmatis*) in comparison with LAM standards. The possible roles of mCD14, TLR-2 and the LPS-binding protein (LBP), a serum acute-phase protein which facilitates the transfer of glycolipids to mCD14, in the stimulation of cells by LMs were also established. Moreover, in order to define the domain of LM that appears to be crucial for the cell activation process, a set of molecules presenting variable arabinan chain lengths were prepared by mild hydrolysis of KanLAM. The activities of the LAM truncated variants but also the phosphatidyl *myo*inositol dimannosides (PIM₂), a precursor molecule of LM, were compared to intact LM with regard to cytokine production.

EXPERIMENTAL PROCEDURES

Materials- RPMI 1640 medium, L-glutamine were purchased from Gibco-BRL (Eragny, France), and foetal calf serum (FCS) from Dominique Dutscher S.A. (Brumath, France). Dulbecco's phosphate saline (PBS), bovine serum albumin (BSA), isotype control mouse IgG_{2a} and IgG_{2b}, humain IgG, FITC-conjugated goat anti-mouse IgG and peroxydase-conjugated goat anti-rabbit IgG were obtained from Sigma Chemical Co. (St Louis, Mo.). Anti-CD14 monoclonal mouse IgG_{2b} (clone My4) and anti-TLR2 monoclonal mouse IgG2a (clone TL2-1) were from Coulter Immunology (Miami, Floride) and Cascade Biosciences (Winchester, Massachusetts), respectively. Rabbit anti-human sCD14 antiserum was purchased from Biometec. The 1,25-dihydroxy-vitamin D3 provided from Calbiochem (Darmstadt, Germany) and the apyrogen water was from Cooper (Melun, France). Recombinant human soluble CD14 (sCD14) was purchased from Biometec (Greifswald, Germany) and human IL-8/Nap-1 module set (BSM204MST) from Bender Medsystems Diagnostic (Vienna, Austria). Human recombinant LBP, neutralizing goat anti-human TNF- α and human TNF- α protein detection Kit and Quantikine mRNA Kits were from R&D systems.

All chemicals used were of the highest analytical grade and LPS contamination was evaluated with the *Limulus* amoebocyte lysate assay kit (QCL1000; BioWhittaker, Walkersville, Md.).

Extraction and purification of LMs and LAMs from M. chelonae and M. kansasii-Lipoglycans from these two pathogenic mycobacterial species were purified by successive detergent and phenol extractions leading to the obtention of a protein, lipid and nucleic acid free material, as described previously (Guérardel *et al.* 2002). Then, lipoglycans were resuspended in Tris deoxycholate buffer (Tris-HCl 10 mM [pH 8.0], EDTA 10 mM, NaCl 0.2 M, deoxycholate 0.25%) and LAMs, LMs, PIMs were separated by gel filtration on a Sephacryl 200 column. LAMs and LMs were extensively dialysed. Purity of preparations was assessed by GC/MS routine experiments and SDS-Page. The endotoxin content of all reagents was measured in a chromogenic Limulus lysate assay (BioWhittaker). The LMs and LAMs preparations contained insignificant amounts of endotoxin (<20 pg LPS / 10µg of LMs and LAMs)-. Purification of PIM₂ from M. kansasii- PIMs were obtained from Tris deoxycholate insoluble This material repetitively lipids fraction. was extracted by chloroform/methanol/water (50/50/5) and applied to a silica gel (KG60 0,063-0,200 mm, MERCK) column irrigated with chloroform, then with chloroform/methanol/water eluant by increasing proportions of methanol/water. Glycolipids, including important quantities of lipooligosaccharides (LOSs) and PIMs were eluted from 50/50/5 eluant. PIMs were finally irrigated separated on а DEAE cellulose (acetate form) column with chloroform/methanol/ammonium acetate eluant increasing proportions of by methanol/ammonium acetate. Fractions containing purified PIMs were eluted with chloroform/methanol 1/2 0.05M ammonium acetate. Purity of the samples was assessed by GC/MS analysis and ¹H homonuclear and ¹³C-¹H heteronuclear HMQC experiments were used to typify this fraction as PIM₂.

Preparation of truncated LAM fractions- 5 mg of purified KanLAM was submitted to 2M acetic acid hydrolysis at 80°C for 2h 30. Sample was lyophilized, resuspended in Tris deoxycholate buffer and hydrolysed compounds were fractionated on a Sephacryl 200 column according to their size. The extend of hydrolysis was assessed by SDS PAGE. Collected fractions were pooled in five fractions (F1-F5), which compositions were determined by GC/MS analysis.

Cell culture- Human promonocytic leukemia THP-1 cells (ECACC n°88081201) were grown in RPMI 1640 supplemented with 10% foetal calf serum (FCS), 2 mM L-glutamine and 2 10^{-5} M β -mercaptoethanol in 5% CO₂-air humidified atmosphere at 37 °C. To express high amount of mCD14 at the cell surface, THP-1 cells were differentiated by treatment for 72 h with 50 nM 1,25-dihydroxy-vitamin D3 (Vey *et al*, 1992). Viability of cells was over 96 % as determined by trypan blue dye exclusion.

mCD14 and TLR2 expressions at the surface of THP-1 cells- mCD14 and TLR2 expressions; on differentiated and undifferentiated THP-1 cells; was analyzed by flow cytometry. Cells (300,000) were incubated 20 min at 4°C with 20 μ g/ml of human IgG, washed three times and incubated for 45 min with 12 μ g/ml of anti-CD14 (MY4) or anti-TLR2 (clone TL2-1) monoclonal antibodies in PBS containing 0.04% NaN₃ and 0.05%_BSA. Mouse isotype immunoglobulins were used as negative controls. After three washings, the cells were stained with second FITC-conjugated goat anti-mouse IgG (1/64) for 30 min at

4°C. Cells were analyzed with a Becton Dickinson FACScan flow cytometer. Cells were gated for forward- and side-angle light scatters, and 10,000 particles of the gated population were measured. The fluorescence channels were set on a logarithmic scale, and the mean fluorescence intensity was determined.

Induction of the secretion of TNF- α and IL-8 by LAM and LM preparations- Promonocytic and differentiated THP-1 cells were seeded into 96-well plastic culture plates at a density of $2x10^5$ cells/well, in RPMI with 10% FCS and glutamine. Various increasing concentrations of purified LAMs and LMs were added for 6h and 24h, in triplicate to cells to induce TNF- α and IL-8 protein releases, respectively. Culture supernatants were then collected, centrifuged prior to quantitation of cytokines by enzyme-linked immunosorbent assay (ELISA).

Inhibition of IL-8 secretion by neutralizing anti-TNF- α antibodies- In order to determine to which extent IL-8 secretion is mediated by TNF- α released by LMs, differentiated THP-1 cells were stimulated, for 24 h at 37°C with 1 or 10 µg/ml of LMs, in the absence or presence of neutralizing goat polyclonal antibodies directed against TNF- α . The supernatants were collected and IL-8 concentration was measured using ELISA.

Inhibition of TNF- α and IL-8 releases by anti-CD14 and anti-TLR2 monoclonal antibodies- Monoclonal anti-CD14 (My4) and anti-TLR2 (clone TL2.1) antibodies have been used to investigate whether the LM-inducing cytokine activity was CD14/TLR-2 dependent. Cells, incubated in RPMI with 10% FCS, were pretreated for 30 min at 37°C, with either 10 μ g/ml of anti-CD14, anti-TLR2, or with isotype antibodies as controls. Then, 10 μ g/ml of CheLM or KanLM were added for 6 h and 24 h, to allow production of TNF- α and IL-8, respectively.- Supernatants were analyzed for cytokines release by ELISA.

Effect of recombinant human LBP on the TNF- α and IL-8 releases induced by LMs- In some experiments, 0.5 ng/ml purified recombinant human LBP was preincubated 30 min at room temperature with 1 and 10 µg/ml of LMs and then added to cells (200, 000 cells/well) at 37°C in RPMI without FCS. After 6 h and 24 h incubations, respectively, TNF- α and IL-8 concentrations, present in supernatants were measured in conditions similar to those described above. Controls were performed without LBP.

Quantification of TNF-\alpha and IL-8 by ELISA- Cell culture supernatants were processed for TNF- α and IL-8 quantification by sandwich ELISA using commercially available kits and according to manufacturer's instructions. Briefly, microtiter plates were coated overnight at 4°C with anti-IL-8 or anti-TNF- α monoclonal antibodies, washed with PBS / 0.05 % (v/v) Tween-20 and blocked 2 h with PBS / 0.05 % Tween containing 0.5 % (w/v) BSA for IL-8 and PBS/ 1% (w/v) BSA/ 5% (w/v) sucrose for TNF. Supernatants diluted in PBS buffer / 0.05 % Tween / 0.1 % Triton X-100 were then transferred to the wells. After washings, peroxidase-conjugated anti-IL-8 (1/6500) or biotin-conjugated anti-TNF polyclonal antibodies (300 ng/ml) were added to wells for 2h. After addition of streptavidin-peroxydase (1/200) for TNF- α , cytokine detection was performed using o-phenylenediamine-dihydrochloride. The reaction was stopped with 50 µl of 2 M H₂SO₄ per well, and the absorbance at 490 nm was measured on a microplate reader. Cytokines in the cell culture supernatant were quantitated in comparison with a standard curve generated with recombinant human TNF- α and IL-8.

Cytokines mRNA expression by LMs and LAMs- The effect of LAMs and LMs isolated from M. chelonae and M. kansasii, on TNF-a and IL-8 mRNA expression by differentiated THP-1 cells, was investigated. 1 10⁶ cells/ml were incubated into 6-well plates in RPMI/ 10% FCS /glutamine in the presence of 10 µg/ml of these glycoconjugates. After 2h and 5h incubations, respectively, at 37°C, the cells were harvested by centrifugation and the level of human TNF and IL-8 mRNA was measured by colorimetric mRNA quantitation (R&D systems). This novel method uses a conventional microplate reader with colorimetric detection to quantitate gene-specific mRNA at low levels. Briefly, stimulated cells were resuspended in "Quantikine mRNA Cell lysis Diluent" (2 x 10⁶ cells/ml). Cell lysates were diluted and RNA contained in samples were hybridized with gene-specific biotin-labeled capture probes and digoxigenin-labeled detection probes in a microplate, during 1h at 65°C. The hybridization solution was then transferred to a streptavidin-coated microplate and the RNA/probe hybrid was captured at room temperature for 1h. After washings to remove unbound material, anti-digoxigenin alkaline phosphatase was added. In the last step, a substrate solution and then an amplifier solution were added. Color was developed in proportion to the amount of gene-specific mRNA in the original sample. Results were measured spectrophotometrically. A serial dilution of specific TNF-a and IL-8 mRNA

calibrators were performed to determine the concentration of mRNA in each sample and results were expressed in 10^{-18} mol/ml (amol/ml).

Analysis of LM/LBP/CD14 complexes- Purified human recombinant sCD14 (0.5 μ g) were mixed with KanLM or CheLM (5 μ g) in presence or absence of human recombinant LBP (0.3 μ g) in a final volume-of 30 μ l PBS and incubated at 37°C for 1h. Mixtures were submitted to electrophoresis on 0.9% agarose gels (veronal buffer, pH 8.2) in non-denaturing conditions. Complexes were then transferred for 24 h to nitrocellulose membranes and sCD14 was detected by western-blot, after blocking with PBS containing 3% SAB/0.05% Tween-20, and incubation with rabbit anti-human CD14 antiserum (1/1500).

RESULTS

LMs, but not LAMs, induce secretion of pro-inflammatory cytokines in differentiated THP-I- The human myelomonocytic leukemia cell line THP-1 is known to over-express mCD14 during differentiation of cells with 1,25-dihydroxy-vitamin D3 and to release cytokines after stimulation with pathogen-derived molecules. As illustrated in Fig 1, flow cytometry analysis confirmed the presence of higher levels of mCD14 to the surface of differentiated THP-1 cells as compared to undifferentiated THP-1 cells. In contrast, cell differentiation did not affect the expression level of TLR-2.

The capacity of LAMs and LMs isolated from *M. kansasii* and *M. chelonae*, to trigger the secretion of TNF- α and IL-8, was next investigated on both undifferentiated and differentiated THP-1 cells. As shown in Fig. 2, various concentrations of LMs were found to induce the release of TNF- α and IL-8 from differentiated cells, whereas no eliciting effect could be detected with undifferentiated THP-1. In differentiated cells, a peak of induction was obtained in the presence of 10 µg/ml LMs. The level of cytokines induced by LMs was about 17-fold higher for TNF- α and 8 fold-higher for IL-8, as compared to unstimulated cells. Moreover, CheLM and KanLM displayed similar cytokine-inducing responses. When polymyxin B, a cationic peptide, known to block the biological function of LPS, was added to the cultures, it did not inhibit cytokine secretion (data not shown). This suggests that the cytokine-inducing responses observed with our LM preparations cannot be attributed to an eventual LPS contamination.

Since TNF- α stimulate the secretion of IL-8 in monocytes/macrophages (Friedland *et al.*, 1992), we investigated whether the IL-8 production detected in THP-1 cells stimulated with LMs may be a consequence of the induction of TNF- α . To adress this question, experiments were undertaken by incubating cells with LMs as well as a neutralizing anti-TNF antibody. As reported in Table 1, the addition of this antibody reduced only partially (21 % and 6 %) the IL-8 production induced by 10 µg/ml of CheLM or KanLM, respectively. This result suggests that the vast majority of IL-8 released upon stimulation with LMs does not result from the secretion of TNF- α . In contrast to LMs, the mannose-capped LAM from *M. kansasii* or the non-capped LAM from *M. chelonae* did not induce any cytokine release even in the presence of 20 µg/ml.

We next compared the cytokine-inducing activities of various LMs purified from four mycobacterial species. PILAM from *M. smegmatis* was also added to this experiment as a positive control since it has been reported to be a strong TNF- α and IL-8 inducing molecule

on THP-1 cells (Gilleron, M. et al., 1997). As shown in Fig. 3, all LMs were potent inducers of TNF- α and IL-8 secretion but no significant differences were observed with respect to their induction capacity. At the same concentrations (10 µg/ml), LMs were slightly more potent in inducing cytokines than PILAM. However, taking into account the different molecular weights of LMs and PILAM -an average of 17 kDa for PILAM and 6 kDa for LMs- they exhibit similar activities toward induction of these cytokines.

To demonstrate that secretion of TNF- α and IL-8-by differentiated THP-1 cells stimulated with LMs results from an induction of specific ARNm synthesis, colorimetric mRNA quantitation method was applied, as described on Materials and Methods. As illustrated in Fig. 4, both CheLM and KanLM strongly up-regulated the TNF- α and IL-8 mRNA production, after 2 h and 5 h induction, respectively, whereas their respective LAMs failed to induce RNAm production. These data indicate that the enhanced levels of cytokines detected in the supernatant of LM-stimulated cells are the consequence of specific mRNA upregulation activities.

LMs induce cytokine secretion through a CD14/TLR-2 and LBP dependent pathway- Toll like receptors, essentially TLR-2 and TLR-4 have been shown to be essential for the recognition of distinct bacterial cell wall components. These bacterial components elicit the activation of an intracellular signaling cascade via TLR, which ultimatly leads to the activation of transcription factors that initiate the transcription of pro-inflammatory cytokine genes (Takeuchi and Akira, 2001). Therefore, TLRs play a key role in innate immune recognition and in subsequent activation of adaptive immunity. Underhill et al. (1999) have demonstrated that upon exposure to M. tuberculosis, macrophages are stimulated to produce TNF- α in a TLR-2-dependent manner. These authors also showed that TLR-2 dependent signaling mediates responses to mycobacterial cell wall fractions enriched for LAM, mycolylarabinogalactan-peptidoglycan, or *M. tuberculosis* total lipids. In addition, CD14 associated to TLR-2 has previously been described as a receptor for LAM capped with phosphoinositol (Yu et al., 1998; Means et al., 1999). CD14 is the major receptor responsible for the effects of LPS on macrophages, monocytes and neutrophils (Fenton and Golenbock, 1998). LBP, an acutephase protein present in serum, was shown to mediate the transfer of glycolipids to CD14, leading to activation of cells (Savedra et al., 1996; Yu et al., 1998; Orr and Tobias, 2000). Given the importance of CD14 and TLR-2 in inducing intracellular signalling to promote a

inflammatory immune response, we investigated whether CD14 and TLR-2 may be involved

in the pro-inflammatory immune response of LM-stimulated THP-1 cells. Studies were conducted by measuring the inhibitory effect of cytokine production using specific anti-CD14 (My4) and anti-TLR2 (clone TL2-1) monoclonal antibodies. As demonstrated in Fig. 5, whereas differentiated THP-1 cells were not activated by anti-CD14 and anti-TLR2 antibodies in the absence of LM, -anti-CD14 antibodies prevented the LM-induced expression of TNF- α and IL-8 by cells. TNF- α secretion was inhibited down to 90.2 and 78.7% for CheLM and KanLM, respectively. IL-8 expression was inhibited at 50.8 and 52.5 % for CheLM and KanLM, respectively. Regarding to the possible participation of TLR-2 signaling, an inhibition of 86% was obtained when cells were pre-treated with anti-TLR-2 antibodies. Isotype monoclonal antibodies did not affect the cytokines production (data not shown). These results clearly suggest the possible participation of both CD14 and TLR-2 in the activation process leading to TNF- α and IL-8 production by LMs.

We next examined whether LBP, an acute-phase protein present in serum that mediates the transfer of glycolipids to CD14, ultimately leading to activation of cells (Savedra *et al.*, 1996; Yu *et al.*, 1998; Orr and Tobias, 2000) may modulate the cytokine-inducing activity of LMs. Experiments were performed in serum-free medium with 1 and 10 μ g/ml LMs supplemented or not with 0.5 ng/ml human recombinant LBP. As shown in Fig. 6, preincubation of CheLM or KanLM with LBP resulted in a significant release of cytokines from differentiated cells, whereas no stimulation was detected in the absence of recombinant LBP, underlining an important of LBP in inducing cytokine produciton by LMs. Altogether, these results suggest that LMs-induced signaling events in differentiated THP-1 cells is mediated through a CD14/TLR2/LBP-dependent mechanism.

Given the importance of CD14 and LBP in mediating the pro-inflammatory response, we investigated whether sCD14, LBP and LMs may interact physically. This was achieved by incubating CD14 with either CheLM alone or in combination with CheLM and LPB, and by analyzing the electrophoretic mobility of the eventual complex formed by agarose-gel electrophoresis in native conditions and western-blot analysis with anti-CD14 polyclonal antibodies. When CD14 was incubated with CheLM, a shift the mobility was observed, suggesting the presence of a bound complex consisting of CD14/CheLM. Interestingly, with the addition of LM and LBP, the mobility was even lower, suggesting the presence of a ternary complex consisting of CD14/CheLM/LBP. However, no alteration of the mobility was found when CD14 was incubated with LBP alone. Similar results were obtained with KanLM.

In agreement with the known functions of CD14 and LBP in relation to LPS, we propose that LBP may transport LMs and transfers it to CD14.

Removal of the arabinan moiety of KanLAM restores the ability to induce cytokines- Since LMs, but not LAMs, stimulate the TNF- α and IL-8 porduciton, we have hypothesized that the presence of arabinan domain and/or the mannan cap in LAMs may prevent this mechanism of action. To address this question, we took advantage of the higher susceptibility to hydrolysis of arabinose linkages to specifically release the arabinan chains from LAM by mild hydrolysis. We therefore generated five fractions containing LAMs with partially truncated arabinan domains by mild hydrolysis, purified them by gel exclusion chromatography on a S200 column. and analyzed them by electrophoresis on a polyacrylamide gel (Fig. 8A). Purified KanLAM and KanLM were also loaded on the gel as standard reference molecules. Sizes of the five purified fractions ranged from nearly intact LAM (F1) to highly truncated molecules (F5) that ran faster than purified KanLM. As expected, assessment of their mannose and arabinose contents by routine gas chromatography (Fig.8 B) established that the decrease of molecular mass of fractions F1 to F5 correlated with a drastic shortening in their arabinan domain. Consequently, it was assumed that simultaneous reduction of their mannose content originated from the release of oligomannosyl domains capping arabinosyl side chains following degradation of arabinan domain. It is noteworthy that the most truncated molecules (F4 and F5) had smaller molecular weights than native LM. This suggests that mannan domain was slightly reduced during hydrolysis, presumably through the loss of Man(α 1-2) side chains due to the lower stability in acidic conditions of $\alpha 1,2$ glycosidic linkage compared to $\alpha 1.6$ linkage.

We next analyzed the cytokine-inducing activity of these LAM truncated molecules. As shown in Fig. 9, the less degraded variants corresponding to fractions F1 and F2 (arabinose/mannose ratio of 1.16 and 0.81, respectively), were unable to induce the TNF- α and IL-8 secretion, as reported for native KanLAM. However, fractions F3 to F5, characterized by shorter arabinan domains (arabinose/mannose ratio ranging from 0.65 to 0.6), induced a strong cytokine-production response. In addition, the capacity of these three fractions to stimulate IL-8 secretion was found to be similar to native KanLM. Interestingly, despite partial degradation of its mannan domain, F5 showed similar cytokine-inducing activity than F3, F4 or native KanLM. These results clearly establish that arabinan moiety of

native KanLAM exhibits a negative effect on the induction of a pro-inflammatory response. Thus, we suggest that the presence of arabinan domain, and eventually terminal mannosyl caps, in LAMs modulates biological properties by masking mannan domain, presumably through sterical hindrance.

PIM₂ fails to stimulate TNF- α and IL-8 secretions- Since the mannan core of LM seems to be required in order to induce an efficient pro-inflammatory response, we thought to define more precisely the domain of LM which is important to explain this biological activity. Phosphatidyl myoinositol mannosides (PIMs) are considered as being the precursors of LM (Chatterjee *et al.*, 1992). The biosynthesis of LM involves the addition of mannopyranosyl (Manp) residues to phosphatidyl myoinositol to produce the short PIMs (2-5 Manp residues) which are further elongated to form an α 1,6-linked Manp backbone substituted by Manp units. Thus, PIM₂ was purified from *M. kansasii* and increasing concentrations of this glycolipid were added to THP-1 cells in order to evaluate capacity of this minimal glycolipidic structure to induce cytokine production (Fig. 10). In contrast to LM, no production of TNF- α or IL-8 secretion could be detected in the supernatant from differentiated THP-1 stimulated with PIM₂, even at the highest concentration of 10µg/ml. Levels of cytokines secreted in the presence of PIM₂ were comparable to those found in supernatants of unstimulated cells (medium alone). These data suggest that the presence of only two Manp residues bound to the phosphatiyl inositol ancher are not sufficient to trigger a pro-inflammatory response. Structures with a higher number of Manp residues are therefore necessary for eliciting this pro-inflammatory response.

DISCUSSION

Up to now, most studies have focused on the functional properties of ManLAMs and PILAM (Chatterjee and Khoo 1998; Nigou *et al.* 2002). LM is a very abundant in mycobacterial cell wall component and represents one of the major lipoglycan fractions. However, probably because it is considered as a direct precursor of LAMs and therefore not such as interesting to study its biological relevance in the host defense has not thoroughly been investigated.

In the present work, we have studied the activity of LMs purified from two pathogen mycobacteria, *M. kansasii* and *M. chelonae*. We have demonstrated that LMs induced the

TNF- α and IL-8 secretions from the human macrophage-like THP-1 cell line. This response was highly dependent on the cell differentiation state since undifferentiated promonocytic THP-1 cells were not activated by LMs. During the differentiation of cells, mCD14 was over-expressed at the cell surface, as previously reported by Vey *et al.* (1992). The involvement of mCD14 in the LM-mediated cell activation process was confirmed by using specific anti-CD14 antibodies (My4). This monoclonal antibody recognizes an epitope involving amino-acid residues 39-44 which are involved in the binding to LPS (Stelter *et al.*, 1997). Since My4 also blocked the cytokine-inducing activity of LMs on THP1 cells, one may assume that residues 39-44 of CD14 also contribute to the interaction with LMs.

The TLR transmembrane signal transducing proteins, are involved in the activation of cells by LPS and other microbial products (Landmann *et al.*, 2000). In particular, PILAM activates macrophages through both TLR2 and CD14 receptors (Means *et al.*, 1999). In the present study, differentiated cells were pretreated with monoclonal anti-TLR-2 antibodies (clone TL2.1) known to interfere in the activation of cells with *Mycobacterium avium* and purified lipoproteins/lipopeptides from spirochetes (Lien *et al.*, 1999). These antibodies also inhibited the pro-inflammatory response elicited by LMs, suggesting that the signaling response mediated by LMs are CD14/TLR-2 dependent. Recently, it has been proposed that TLR2 may dimerize with others TLRs to induce a signal (Underhill *et al.*, 1999; Ozinski et al., 2000). Hence, the involvement of other TLRs in host responses to LMs cannot be excluded. The availability of knock-out mice defective in the different TLRs will help us to define the real participation of these receptors in the inflammatory response induced by LMs.

LBP, a serum protein with a lipotransferase activity, may be required for endotoxininduced cell activation (Tobias *et al.*, 1988; Schumann *et al.*, 1990; Juffermans *et al.*, 1998). More particularly, the importance of LBP in promoting the activity of PILAM on HL60 cells was reported (Yu *et al.*, 1998). We demonstrate here that LBP is essential for the action of LMs on differentiated THP-1 cells, thus suggesting that LBP may mediate the transfer of LM to mCD14. Our experiments showing molecular interactions between mCD14 and LMs in the presence of LBP are in accordance with these results. Hence, *in vivo*, LBP might modulate the bio-availability of LM and therefore regulate the inflammatory response induced by LMs.

Surprisingly, our data demonstrate that despite significant structural variations both in polysaccharidic core and in lipid anchor, all LMs tested showed very similar eliciting properties. This suggests that their immunological modulating properties are solely linked to the presence of either a mannan core or a phosphatidyl-*myo*inositol anchor, independently of

their structural variability. The fact that PIM₂ failed to induce any response favours the involvement of the sole mannan core in the activity mediated by LMs and excludes the possibility of the PI anchor to be responsible for this activity. However, numerous studies have reported that biological activities of lipoglycans were highly dependent on the integrity of their lipid moieties (Nigou et al., 1997). Deacylation of CheLM or KanLM completely abrogated their TNF-a secretion inducing properties toward differentiated THP-1 cells for concentrations below 10 µg. ml⁻¹ (data not shown). Unexpectedly, their activity was partially restored from concentrations of 10 µg ml⁻¹. Recently, two studies have established that lipid anchor of ManLAMs exerted its influence on IL-12 induction and SP-A fixation through formation of micelles and that it was not directly implicated in a receptor binding event [Sidobre et al., 2002; Nigou et al., 2001]. Presentation of lipoglycans integrated in micelles would then be a prerequisite for the acquisition of multivalency and the efficient presentation of ManLAM to mannose receptor and SP-A. Indeed, multivalency appears as a common feature in interactions involving carbohydrates, either lectin-carbohydrate or carbohydratecarbohydrate based interactions. In the light of these results, if multivalence is also required for LMs inducing properties, the partial activity restoration for high concentrations of deacylated LMs may simply be the result of acquisition of a pseudo-multivalency effect due to high local concentration.

The lack of activity of PIM₂ fraction is difficult to interpret in the light of previous studies. Indeed, some authors have reported a stimulation of TNF- α secretion by PIMs in macrophages (Jones *et al.*, 2001; Zhang J. *et al.*, 1995; Barnes *et al.*, 1992) while others have demonstrated a 100-fold lower induction by PIM₆ than PILAM (Chatterjee *et al.*, 1992). It is noteworthy, that most studies have used mixtures of PIMs, which render direct comparisons of data difficult. Indeed, if the mannan domain is involved in the eliciting activity, as our results tend to demonstrate, the heterogeneity of PIMs mannan chain length is expected to lead to various effects. In this context, the fact that PIM₂ did not show any activity established that presence of only two mannose residues -one substituting *myo*inositol group in C-2 position and the other in C-6 position- was not sufficient to trigger TNF- α and IL-8 release from differentiated THP-1 cells. This suggests that the degree of mannosylation modulates the eliciting effect of lipoglycans, supporting the hypothesis that the biological property of LM is associated with mannan core. This was further supported by the demonstration that arabinan domain of LAM inhibited the cytokine production of lipoglycans. Indeed, starting from a KanLAM that presented no cytokine eliciting effect, we were able to restore the optimal

activity level of LMs by progressively degrading its arabinan domain. Complete degradation of the arabinan domain was not necessary for restoring the activity, an effect being observed from 70% of arabinan degradation and release of 20 mannose residues (fraction F3). Twenty mannose residues correspond incidentally to the calculated average number of mannose residues of the oligomannoside capping domain (unpublished results Guérardel *et al.*, in submission). We observed a dramatic increase of activity between F2 and F3, despite a minimal difference of arabinan chain length between these two components. This effect was observed for both IL-8 and TNF- α , suggesting the occurrence of a functional 'turning point' between F2 and F3. From F3 to F5, the TNF- α release increased with the extent of arabinan degradation, whereas IL-8 release did not. On this basis, we hypothesize that arabinan domain exerts its inhibiting effect not through a specific interaction with a macrophage receptor but through sterical hindrance that precisely prevent interaction of the underlying mannan domain.

As presented in this study, LMs and PILAM exhibited similar cytokines inductions levels through pathways involving both CD14 and TLR-2, whereas ManLAMs and noncapped LAMs do not. Presently, no available data allows to understand the molecular bases of the sorting out of the lipoglycans with fairly similar structures but different functions. On the basis of their respective structures, it is unlikely that LMs and PILAM exhibit a single common structural feature that would explain they specific activity. Previous studies strongly suggested that the inositol-phosphate capping may represent the major cytokine-inducing component of PILAM [Gilleron et al., 1997; Guérardel et al., 2002]. One tentative hypothesis would be that the *myo*inositol-phosphate motif incorporated in the lipid anchor of lipoglycans, if accessible, may mimic the role of inositol-phosphate capping of PILAM, presuming that it is accessible in LMs but not in LAMs due to steric hindrance by the arabinan domain. However, due to decoration by mannose residues in C-2 and C-6, and eventually by an additional acyl group in C-3 position, it is highly improbable that the myoinositol-phosphate of the lipid anchor may mimic the terminal, unsubstituted myoinositol-phosphate. Furthermore, the lack of activity of PIM₂ definitively rules out this hypothesis. Presently, it is very difficult to define the exact recognition patterns of TLRs. This family of receptors seems to interact with a broad range of structurally unrelated microbial products including peptidoglycans, lipopolysaccharides, lipoglycans and lipoproteins (Lien et al., 2002). On this basis, it is likely that LMs and PILAM are recognized by a different panel of TLRs, including TLR-2, presenting distinct structural requirements, in association with CD14 and LBP.

In conclusion, our study strongly suggest that LMs are not only precursors of LAM biosynthesis but do participate in innate inflammatory response through a CD14-TLR2 dependent mechanism. The presence of large amounts of LMs in the cell wall of various pathogenic mycobacteria may influence the host immune defenses. Our studies is also paving the way to further investigation about the functions of LMs in virulence and mycobacterial pathogenesis and for the development of new therapeutic approaches.

Acknowledgements: This work was supported in part by the Université des Sciences et Technologies de Lille, the Centre National de la Recherche Scientifique (UMR n°8576, director: Dr J.C. Michalski) and by INSERM.

REFERENCES

- Baggiolini, M., Walz, A. and Kunkel, S.L. 1989. Neutrophil-activating peptide-1/inteleukine 8, a novel cytokine that activates neutrophils. J. Clin. Invest. 84 : 1045-1049
- Barnes, P.F., Chatterjee, D., Abrams, J.S., Lu, S., Wang, E., Yamamura, M., Brennan, P.J., and Modlin, R.L. 1992. Cytokine production induced by *Mycobacterium tuberculosis* lipoarabinomannan. Relationship to chemical structure. J. Immunol. 149: 541-547
- Bekker, L.G., Moreira, A.L., Bergtold, A., Freeman, S., Ryffel, B., and Kaplan, G.
 2000. Immunopathologic effects of tumor necrosis factor alpha in murine mycobacterial infection are dose dependent *Infection and immunity* 68 : 6954-6961
- Chatterjee, D. 1997. The mycobacterial cell wall : structure, biosynthesis and sites of drug action. *Curr. Opin. Chem. Biol.* 1 : 579-588
- Chatterjee, D., and Khoo, K.H. 1998. Mycobacterial lipoarabinomannan : an extraordinary lipoheteroglycan with profound physiological effects. *Glycobiology* 8 : 113-120
- Chatterjee, D., Lowell, K., Rivoire, B., McNeil, M.R., and Brennan P.J. 1992. Lipoarabinomannan of *Mycobacterium tuberculosis*. Capping with mannosyl residues in some strains. J. Biol. Chem. 267 : 6234-6239

- Chatterjee, D., Roberts, A.D., Lowell, K., Brennan, P.J., and Orme, I.M. 1992. Structural basis of capacity of lipoarabinomannan to induce secretion of tumor necrosis factor. *Infect. Immun.* 60 : 1249-1253
- Daffe, M., and Draper, P. 1998. The envelope layers of mycobacteria with reference to their pathogenicity. *Adv. Microb. Physiol.* 39 : 131-203
- Fenton, M. J., and Golenbock, D. T. 2001. LPS-binding proteins and receptors. J. Leukocyte Biol. 64: 25-
- Friedland, J.S., Remick, D.G., Shattock, R., and Griffin, G.E. 1992. Secretion of interleukine-8 following phagosytosis of *Mycobacterium tuberculosis* by human monocyte cell lines. *Eur. J. Immunol.* 22: 1373-1378
- Gilleron, M., Bala, L., Brando, T., Vercellone, A. & Puzo, G. 2000. *Mycobacterium tuberculosis* parietal and cellular lipoarabinomannans. J. Biol. Chem. 275: 677-684.
- Gilleron, M., Himoudi, N., Adam, O., Constant, P., Venisse, A., Riviere, M., and Puzo, G. 1997. *Mycobacterium smegmatis* phosphoinositols-glyceroarabinomannans. Structure and localization of alkali-labile and alkali-stable phosphoinositides. J. Biol. Chem. 272: 117-124
- 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 α 1,3 mannopyranose side chain J. Biol. Chem. 277, 30636-30648
- Heldwein, K.A., and Fenton, M.J. 2002. The role of Toll-like receptors in immunity against mycobacterial infection. *Microbes Infect.* 4 : 937-944
- Hunter, S.W., and Brennan, P.J. 1990. Evidence for the presence of a phosphatidylinositol anchor on the lipoarabinomannan and lipomannan of *Mycobacterium tuberculosis. J. Biol. Chem.* 265 : 9272-9279
- Ingram, C.W., Tanner, D.C., Durack, D.T., Kernodle, G.W., Jr, and Corey, G.R. 1993.
 Disseminated infection with rapidly growing mycobacteria. *Clin. Infect. Dis.* 16: 463-471
- Jacobson, K.L., Teira, R., Libshitz, H.I., Raad, I., Rolston, K.V.I., Tarrand, J., and Whimbey, E. 2000. *Mycobacterium kansasii* infections in patients with cancer. *Clin. Infect. Dis.* 30: 965-969

- Jones, B.W., Means, T.K., Heldwein, K.A., Keen, M.A., Hill, P.J., Belisle, J.T., and Fenton, M.J. 2001. Different Toll-like receptor agonists induce distinct macrophage responses. *Journal of leukocyte biology* 69: 1036-1044
- Juffermans, N.P., Verbon, A., van Deventer, S.J.H., Buurman, W.A., van Deutekom, H., Speelman, P., and Van Der Poll, T. 1998. Serum concentration of lipopolysaccharide activity-modulating proteins during tuberculosis. *The journal of infectious diseases* 178: 1839-1842
- Khoo, K.H., Dell, A., Morris, H.R., Brennan, P.J., and Chatterjee, D. 1995a. Inositol phosphate capping of the nonreducing termini of lipoarabinomannan from rapidly growing strains of *Mycobacterium*. J. Biol. Chem. 270: 12380-12389
- Khoo, K.H., Dell, A., Morris, H.R., Brennan, P.J., and Chatterjee, D. 1995b. Structural definition of acylated phosphatidylinositol mannosides from *Mycobacterium tuberculosis*: definition of a common anchor for lipomannan and lipoarabinomannan. *Glycobiology* 5: 117-127
- Khoo, K.H., Tang, J.B. and Chatterjee, D. 2001 Variation in mannose-capped terminal arabinan motifs of lipoarabinomannans from clinical isolates of *Mycobacterium tuberculosis* and *Mycobacterium avium* complex. J. Biol. Chem. 276, 3863-3871
- Knutson, K.L., Hmama, Z., Herrera-Velit, P., Rochford, R., and Reiner N.E. 1998. Lipoarabinomannan of *Mycobacterium tuberculosis* promotes protein tyrosine dephosphorylation and inhibition of mitogen-activated protein kinase in human mononuclear phagocytes. Role of the Src homology 2 containing tyrosine phosphatase. J. Biol. Chem. 273: 645-652
- Landmann, R., Muller, B., and Zimmerli, W. 2000. CD14, new aspects of ligand and signal diversity. *Microbes Infect.* 2 : 295-304
- Lien, E., Sellati, T.J., Yoshimura, A., Flo, T.H., Rawadi, G., finberg, R.W., Carroll, J.D., Espevik, T., Ingalls, R.R., Radolf, J.D., and Golenbock, D.T. 1999. Toll-like receptor 2 functions as a pattern recognition receptor for diverse bacterial products. J. Biol. Chem. 274: 33419-33425
- Means, T.K., Lien, E., Yoshimura, A., Wang, S., Golenbock, D.T., and Fenton, M.J. 1999. The CD14 ligands lipoarabinomannan and lipopolysaccharide differ in their requirement for Toll-like receptors. J. Immunol. 163: 6748-6755

- Moreno, C., Mehlert, A., and Lamb, J. 1988. The inhibitory effects of mycobacterial lipoarabinomannan and polysaccharides upon polyclonal and monoclonal human T cell proliferation. *Clin. Exp. Immunol.* 74 : 206-210
- Murdoch, C. and Finn, A. 2000. Chemokine receptors and their role in inflammation and infectious diseases. *Blood* 95 : 3032-3043
- Nigou, J., Gilleron, M., Cahuzac, B., Bounéri, J. D., Herold, M., Thurnher, M. & Puzo, G. (1997). The phospho-myo-inositol anchor of the lipomannans from Mycobacterium bovis bacillus Calmette Guérin. J. Biol. Chem. 272: 23094-23103.
- Nigou, J., Gilleron, M., Rojas, M., Garcia, L.F., Thurnher, M., and Puzo, G. 2002. Mycobacterial lipoarabinomannans : modulators of dendritic cell function and the apoptotic response. *Microbes Infect.* 4 : 945-953
- Nigou, J., Zelle-Rieser, C., Gilleron, M., Thurnher, M., and Puzo, G. 2001. Mannosylated lipoarabinomannans inhibit IL-12 production by human dendritic cells : evidence for a negative signal delivered through the mannose receptor. J. Immunol. 166 : 7477-7485
- Orr, S.L., and Tobias, P. 2000. LPS and LAM activation of the U373 astrocytoma cell line : differential requirement for CD14. *Journal of endotoxin research* 6 : 215-222
- Ozinsky, A., Underhill, D.M., Fontenot, J.D., Hajjar, A.M., Smith, K.D., Wilson, C.B., Schroeder, L., and Aderem, A. 2000. The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between Toll-like receptors. *PNAS* 97 : 13766-13771
- Pintado, V., Gomez-Mampaso, E., Martin-Davila, P., Cobo, J., Navas, E., Quereda, C., Fortun, J., and Guerrrero, A. 1999. *Mycobacterium kansasii* infection in patients infected with human immunodeficiency virus. *Eur. J. Clin. Microbiol. Infect. Dis.* 18: 582-586
- Rafael, E., and Carlos, E. 1997. *Mycobacterium kansasii* disease in patients infected with human immunodeficiency virus. *Clin. Infect. Dis.* 24 : 1233-1238
- Savedra, R., Delude, R.L., Ingalls, R.R., Fenton, M.J., and Golenbock, D.T. 1996. Mycobacterial lipoarabinomannan recognition requires a receptor that shares components of the endotoxin signaling system. J. Immunol. 157: 2549-2554
- Schumann, R.R., Leong, S.R., Flaggs, G.W., Gray, P.W., Wright, S.D., Mathison, J.C., Tobias, P.S., and Ulevitch, R.J. 1990. Structure and function of lipopolysaccharide binding protein. *Science* 249 : 1429-1431

- Sibley, L.D., Hunter, S.W., Brennan, P.J., and Krahenbuhl, J.L. 1988. Mycobacterial lipoarabinomannan inhibits gamma interferon-mediated activation of macophages. *Infect. Immun.* 56: 1232-1236
- Sidobre, S., Puzo, G. and Rivière, M. (2002) Lipid-restricted rcognition of mycobacterial lipoglycans by human pulmonary surfactant protein A: a surfaceplasmon-resonance study. *Biochem. J.* 365, 89-97
- Stelter, F., Bernheiden, M., Menzel, R., Jack, R.S., Witt, S., Fan, X., Pfister, M., and Schütt, C. 1997. Mutation of amino acids 39-44 of human CD14 abrogates binding of lipopolysaccharide and *Escherichia coli. Eur. J.Biochem.* 243 : 100-109
- Takeuchi, O., and Akira, S. 2001. Toll-like receptors; their physiological role and signal transduction system. *Int. Immunopharmacol.* 1: 625-635
- Tobias, P.S., Mathison, J.C., and Ulevitch, R.J. 1988. A family of lipopolysaccharide binding proteins involved in responses to gram-negative sepsis. J. Biol. Chem. 263: 13479-13481
- Tsenova, L., Bergtold, A., Freedman, V.H., Young, R.A., and Kaplan, G. 1999. Tumor necrosis factor α is a determinant of pathogenesis and disease progression in mycobacterial infection in the central nervous system. *Proceedings of the National Academy of Sciences of the USA* 96 : 5657-5662
- Underhill, D. M., Ozinsky, A., Smith, K.D., and Aderem, A. 1999. Toll-like receptor-2 mediates mycobacteria-induced proinflammatory signaling in macrophages. *Proc. Natl. Acad. Sci. USA* 96: 14459-14463
- Underhill, D.M., and Ozinsky, A. 2002. Toll-like receptors : key mediators of microbe detection. *Curr. Opin. Immunol.* 14 : 103-110
- Venisse, A., Berjeaud, J.M., Chaurand, P., Gilleron, M., and Puzo, G. 1993. Structural features of lipoarabinomannan from *Mycobacterium bovis* BCG. Determination of a molecular mass by laser desorption mass spectrometry. *J. Biol. Chem.* 268 : 12401-12411
- Venisse, A., Fournie, J.J., and Puzo, G. 1995. Mannosylated lipoarabinomannan interacts with phagocytes. *Eur. J. Biochem.* 231: 440-447
- Vey, E., Zhang, J.H., and Dayer, J.M. 1992. INF-γ and 1,25(OH)₂D₃ induce on THP-1 cells distinct patterns of cell surface antigen expression, cytokine production, and responsiveness to contact with activated T cells. J. Immunol. 149 : 2040-2046

- Wallace, R.J., Jr, Brown, B.A., and Onyi, G.O. 1992. Skin, soft tissue, and bone infections due to *Mycobacterium chelonae chelonae* : importance of prior corticosteroid therapy, frequency of disseminated infections, and resistance to oral antimicrobials other than clarithromycin. *J. Infect. Dis.* 166 : 405-412
- Wright, S.D., Ramos, R.A, Tobias, P.S., Ulevitch, R.J., and Mathison, J.C. 1990.
 CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS-binding protein. Science 249 : 1431-1433
- Yoshida, A., and Koide, Y. 1997. Arabinofuranosyl-terminated and mannosylated lipoarabinomannans from *Mycobacterium tuberculosis* induce different levels of interleukine-12 expression in murine macrophages. *Infect. Immun.* 65 : 1953-1955
- Yu, W., Soprana, E., Cosentino, G., Volta, M., Lichenstein, H.S., Viale, G., and Vercelli, D. 1998. Soluble CD14 1-152 confers responsiveness to both lipoarabinomannan and lipopolysaccharide in a novel HL-60 cell bioassay. J. Immunol. 161: 4244-4251
- Zhang, Y., Broser, M., Cohen, H., Bodkin, M., Law, K., Reibman, J., and Rom, W.N. 1995. Enhanced interleukine-8 release and gene expression in macrophages after exposure to *Mycobacterium tuberculosis* and its components. *J. Clin. Invest.* 95, 586-592
- Zhang, Y., Doerfler, M., Lee, T.C., Guillemin, B., and Rom, W.N. 1993. Mechanisms of stimulation of interleukine-1β and tumor necrosis factor-α by *Mycobacterium tuberculosis* components. J. Clin. Invest. 91: 2076-2083

Table 1: Effect of anti-TNF α antibodies on the release of IL-8 (ng/ml) from differentiated THP-1 cells. Differentiated THP-1 cells were stimulated for 24h with KanLM and CheLM in the absence or presence of neutralizing anti-TNF α antibodies. The supernatants were collected and IL-8 release measured using IL-8 ELISA. The results were expressed as the mean +/- SD of one experiments performed in triplicate and representative of three independent experiments.

	Without antidody IL-8 ng/ml	With anti-TNFa IL-8 ng/ml	Percent inhibition
Unstimulated cells	1.63 ± 0.22	_	_
CheLM M. chelonae			
1µg ml ⁻¹	3.70 ± 0.12	2.63 ± 0.18	29
$10 \ \mu g \ ml^{-1}$	7.03 ± 0.11	5.53 ± 0.15	21
KanLM M. kansasii			
1μg ml ⁻¹	3.01 ± 0.24	2.31 ± 0.2	23
$10 \ \mu g \ ml^{-1}$	6.06 ± 0.11	5.66 ± 0.31	7



FIG.1: Cell Surface expression of TLR-2 and mCD14. Promonocytic THP-1 cells were exposed to 1,25dihydroxy-vitamin D_3 for 72 h, as described in materials and methods. Analysis of mCD14 and TL-2 expression at the surface of undifferentiated (A) and differentiated cells (B) was performed by flow cytometry -using anti-CD14 (MY4) and anti-TLR-2 (clone TL2-1) mouse antibodies or an isotype control antibody, followed by incubation with FITC-labelled anti-mouse IgG. Expression of mCD14 is represented in shadded, whereas expression of TLR-2 is designated as a solid line. Cells incubated with the control isotype are shown as a dotted line. Results are shown as linear-log scale fluorescence histogramme. Results shown are representative from two independent experiments.



FIG. 2: Pro-inflammatory cytokine secretion from THP-1 cells stimulated either with LMs-LAMs. Undifferentiated (white bars) or differentiated (black bars) human promonocytic THP-1 cells were incubated in 10% FCS-RPMI medium with increasing concentrations of CheLM, KanLM, CheLAM or KanLAM. Culture supernatants were collected after 6 h or 24 h to assay TNF- α (A) or IL-8 (B) production, respectively. The results presented are from one representative experiment out of three independent experiments with similar results. Data are expressed as mean +/- SD of triplicate wells.



FIG. 3. Release of pro-inflammatory cytokines from differentiated THP-1 cells in response to different LMs and PILAM. LMs isolated from *M. smegmatis*, *M. bovis* BCG, *M. chelonae*, or *M. kansasii* as well as PILAM purified from *M. smegmatis* were incubated at a concentration of 10 μ g/ml with differentiated THP-1 cells in presence of 10% FCS. Production of TNF- α and IL-8 present in supernatants were measured by ELISA. Results are presented as means of triplicate wells +/- SD from one representative experiment out of two with comparable results.



FIG. 4: Expression of specific-TNF- α and IL-8 mRNA induced by LMs in differentiated THP-1 cells. Differentiated THP-1 cells were incubated in 10% FCS-medium, alone or in combination with 10 µg/ml of CheLM, CheLAM, KanLM or KanLAM. After cells lysis, total-mRNA were hybridized with gene specific probes and quantified, as described in Materials and Methods. Expression levels of TNF- α mRNA are presented with white bars, whereas expression levels of IL-8 mRNA are presented with black bars. Results correspond to the mean of triplicate wells +/- SD from one representative experiment out of three.



FIG. 5: Inhibition of the TNF- α and IL-8 production by anti-CD14 and anti-TLR-2 antibodies. Differentiated THP-1 cells were pretreated with 10 µg/ml of either an anti-CD14 (white bars), an anti-TLR2 (grey bars), or a control isotype mAb (not shown) for 30 min at 37°C in 10% FCS medium, prior to the addition of CheLM or KanLM (10 µg/ml). Results with untreated cells are shown in black bars. TNF- α (A) and IL-8 (B) were measured in the cell culture supernatants, as previously described. The results from a representative experiment are expressed as a mean +/- SD of triplicate wells. Similar data were obtained in three independent experiments.



FIG. 6. Effect of human recombinant LBP on the LMs activity. LMs (1 and 10 μ g/ml) were preincubated for 30 min at room temperature alone (white bars) or with 0.5 ng/ml of purified human recombinant LBP (black bars), before addition to differentiated THP-1 cells. After 6 h and 24 h incubation in RPMI without FCS, TNF- α (A) and IL-8 (B) were measured in the cell culture supernatants. Results correspond to the mean of triplicate wells +/- SD from one representative experiment out of three.



FIG. 7. Analysis of complexes between sCD14 and LMs in the presence of LBP. Human recombinant sCD14 (0.5 μ g) was incubated with either 5 μ g CheLM (A) or KanLM (B), with or without human recombinant LBP (0.3 μ g), in PBS at 37°C for 1h. Complexes were runned on a 0.9% Agarose-gel (veronal buffer pH 8.2) in nondenaturing conditions and transferred to nitrocellulose membranes. The migration of sCD14 was visualizated by using rabbit anti-human CD14 IgG and peroxydase-conjugated anti- IgG. The experiment was performed three time and the result of one experiment is shown.



FIG 8. Acid hydrolysis of KanLAM. KanLAM was hydrolyzed 2h30 at 80 °C with 2 M acetic acid, and fractionated on Sephacryl 200 column, as described under "Experimental Procedures". The five fractions (F1 to F5) collected were analyzed by SDS-13% PAGE gel, followed by silver staining after periodic oxydation (A). Average mannose (white bars) and arabinose (black bars) content per molecule were defined by routine gas chromatography (B).



FIG. 9. Activity of hydrolyzed fractions of KanLAM. Differentiated THP-1 cells were stimulated with 10 μ g/ml of each purified fraction (F1 to F5), or with native KanLM and KanLAM, in RPMI with 10% FCS. Culture supernatants were assayed for TNF- α (A) and IL-8 (B) production, as previously described. Data presented are means +/- SD for one representative experiment of three with comparable results.



FIG. 10. Capacity of PIM₂ purified from *M. kansasii to* elicit TNF- α and IL-8 secretion. Differentiated THP-1 cells were activated by various concentrations of PIM₂, or 10µg/ml LM of *M. kansassi*, in RPMI with 10% FCS. The release of TNF- α (A) and IL-8 (B) was quantified, as previously described. Data presented are means +/- SD for one representative experiment of three with comparable results.

248

I GENERALITES

1. Introduction

Les processus moléculaires qui régulent la sélectivité des interactions cellule-cellule et cellule-substratum sont à la base de l'embryogenèse et de la morphogenèse, indispensables à l'avènement des organismes pluricellulaires, animaux ou végétaux. Ils ont également des implications multiples dans la distinction du "soi" et du "non soi". La perte de sélectivité des interactions cellule-cellule, qui s'observe dans de nombreuses pathologies, est, par exemple, impliquée dans les propriétés métastasiques des cellules cancéreuses. De très nombreuses familles de molécules d'adhésion ont été décrites à ce jour qui rendent compte de la plupart des phénomènes d'interactions connus entre cellules: (i) de nombreuses protéines d'adhésion (e.g. fibronectine, laminine, thrombospondine, N-CAM, E-cadherine) et leurs récepteurs protéiques, en particuliers les membres de la famille des intégrines, (ii) les membres de la famille des intégrines, (ii) les membres de la famille des intégrines de la familles et galectines-, (iv) les sucres qui interagissent entre eux.

Le rôle des sucres dans les interactions cellulaires a été très bien documenté ces vingt dernières années, mais presque exclusivement dans l'optique des interactions sucre-protéine et ce n'est que très récemment que le concept d'interaction spécifique sucre-sucre s'est imposé comme un phénomène incontournable dans de nombreux processus biologiques. Les prémices de ce concept ont principalement émergé de l'étude des polysaccharides et des glycosphingolipides (GSL). Dés 1980, il a été observé que des micro-billes substituées par des chaînes de hyaluronate s'agrégeaient à des billes recouvertes de chondroïtine sulfate. Par contre, chaque type de bille ne s'auto-agrégeaient pas [Turley, E.A. & Roth, S., 1980]. Le fait que cette interaction spécifique était sensible à l'oxydation periodique et à la digestion par une galactose oxydase, mais insensible à la protéolyse suggérait fortement l'existence d'interaction entre les chaînes glycanniques. Parallèlement, il avait été démontré que les fragments glycopeptidiques isolés du facteur d'agrégation des cellules d'éponges présentaient des capacités d'auto-agrégation spécifiques similaires à celles de la molécule native [Misevic, G.N. et al., 1982]. Ceci suggérait que les capacités d'adhésion de ce facteur était portées par sa partie glycannique. Concernant les glycolipides, l'hypothèse selon laquelle il existerait des interactions sucre-sucre avait clairement été édictée par Endo et collaborateurs suite à leurs travaux sur les liposomes [Endo, T. et al., 1982]. Ils avaient observé que la reconnaissance

spécifique de la glycophorine intégrée dans des liposomes par des lectines était inhibée par l'intégration de glycolipides de Forssman dans ces mêmes liposomes. Ces résultats suggéraient fortement que des interactions sucre-sucre entre les parties glycanniques des glycolipides et de la glycophorine empêchaient la fixation des lectines sur la glycophorine. Enfin, la démonstration définitive de l'existence de ces interactions par l'équipe du Dr Hakomori trouve son origine dans l'observation des phénomènes de compaction des cellules embryonnaires de mammifère. Le fait que l'expression transitoire du déterminant Le^x à la surface cellulaire au stade 8-32 cellules (morula) corrèle dans le temps avec la compaction de ces cellules [Gooi, H.C. et al., 1981; Hakomori, S. et al., 1981] et que la compaction est spécifiquement inhibée par des molécules porteuses de Le^x [Fenderson, B.A. et al., 1984; Bird, J.M. & Kimber, S.J., 1984], avait déjà suggéré l'implication de ce motif glycannique dans les interactions cellulaires. Bien que certains auteurs aient suggéré l'existence de lectines complémentaires au Le^x [Bird, J.M. & Kimber, S.J., 1984], de tels composés n'ont put être isolés. Sur cette base, les études ultérieures ont démontrées sans ambiguïté que ces interactions étaient en partie le résultat d'interactions moléculaires multivalentes homophiliques et homotypiques entre les déterminants Le^x exprimés par chaque cellule [Eggens, I. et al., 1989], ce qui a définitivement entériné le concept d'interaction sucre-sucre. Rétrospectivement, il apparaît que cette découverte a révolutionné la conceptualisation des interactions moléculaires impliquant les glycannes. Alors que ces derniers étaient jusqu'alors cantonnés à interagir avec des lectines, de nombreux exemples d'interactions sucre-sucre ont été mis en évidence dès après 1990. La majorité de ces interactions impliquent des glycosphingolipides, mais l'existence d'interaction entres glycoprotéines et polysaccharides a été également fermement établie. Dans la suite de ce chapitre, nous allons passer en détail les exemples les mieux étudiés d'interactions sucre-sucre, homophiliques ou hétérophiliques, impliquées dans les interactions cellulaires homotypes et hétérotypiques, en revenant tout d'abord sur les interactions Le^x-Le^x. Nous décrirons également les interactions homophiliques spécifiques d'espèce chez les cellules d'éponges marines qui impliquent un nouveau type de glycoconjugués, les glyconectines. Nos travaux s'intègrent dans l'étude de la spécificité d'espèce des structures des glyconectines, qui serait à la base de leurs capacités d'interaction spécifiques d'espèce.

2. Interactions sucre-sucre chez les eucaryotes supérieurs

21. Interactions homophiliques Le^x-Le^x

Les interactions homophiliques sucre-sucre sont définies ici comme des interactions spécifiques entre glycannes de structure identique, quelque soit la molécule porteuse de ces déterminants glycanniques. Il est fort probable que la molécule porteuse du déterminant glycannique impliqué dans l'interaction puisse elle même y jouer un rôle, ne serait ce que dans la présentation du déterminant, ainsi que cela a été démontré pour les interactions impliquant les galactocéramides (voir chapitre III 222.). Néanmoins, le fait que dans plusieurs cas la nature de la molécule porteuse n'ait pas été déterminée nous a incité à ne tenir compte que du glycanne impliqué dans l'interaction pour la classification des interactions homo- ou hétérophiles. L'existence d'interactions spécifiques Le^x-Le^x a été démontrée par l'étude des propriétés d'agrégation de liposomes formés de glycosphingolipides (GSL) purifiés d'adénocarcinomes humains portant le déterminant Le^x (III³FucnLc₄, V³FucnLc₆ et III³FucV³FucnLc₆) [Eggens, I. et al., 1989]. Les composés renfermant le déterminant Le^x seront dénommés dans la suite du chapitre sous les termes "liposome Le^x" et "glycosphingolipides Le^x" ou "GSL Le^x". La démonstration était basée sur les observations suivantes: (i) les liposomes Le^x interagissaient avec les glycosphingolipides Le^x adsorbés sur support solide, cette interaction étant inhibée par l'ajout de Le^x; (ii) les liposomes Le^x s'agrégeaient en solution, au contraire de tous les autres liposomes testés; (iii) les liposomes Le^x interagissaient avec du lacto-N-fucopentaitol III (LFP III). L'ensemble de ces résultats démontrait l'existence d'une interaction très spécifique entre déterminants Le^x. Ces interactions sont très dépendantes de la présence de cations divalents et nécessitent une présentation multivalente. La modélisation des épitopes Le^x a démontré que ces derniers présentaient une surface hydrophobe pouvant servir de base à une interaction hydrophobe spécifique entre deux molécules. Elle ferait intervenir un ion Ca²⁺ pouvant coordonner quatre oxygènes [Kojima, N. et al, 1994]. Une étude par RMN a récemment confirmé qu'en présence de Ca²⁺, un seul mode d'assemblage bien défini était observable entre deux molécules de Le^x [Geyer, A. et al., 2000]. De plus, la force d'interaction entre deux molécules individuelles de Le^x en solution a été estimée par microscope de force atomique à 20 pN (pico Newton), ce qui correspond à la force inter-moléculaire la plus faible jamais mesurée entre deux molécules biologiques [Tromas, C. et al., 2001]. Ce résultat démontre la nécessité d'une interaction multivalente pour obtenir une interaction significative au niveau biologique.
Ainsi que nous l'avons exposé précédemment, les interactions Le^x-Le^x seraient indispensables à l'initiation de la compaction des cellules dès les premiers stades du développement embryonnaire. Ces cellules embryonnaires précoces sont encore identiques, leur compaction est donc le résultat d'interaction homophiliques et homotypiques. La lignée F9 de cellules embryonnaires cancéreuses de souris présente toutes les caractéristiques des cellules de blastula murines, en particulier l'expression de Le^x membranaire uniquement au stade non-différencié. Leur utilisation pour l'étude des propriétés d'agrégation des cellules embryonnaires a démontré que l'auto-agrégation des cellules F9 via les interactions Le^x-Le^x n'était pas sous le contrôle de glycolipides Le^x, mais de polylactosaminoglycannes abondamment substitués par l'épitope Le^x [Kojima, N. et al., 1994]. Toutefois, il a également été montré que des cellules non adhérentes pré-incubées avec des glycosphingolipides Le^x s'agrégeaient de façon spécifique en présence de calcium, de manière analogue aux cellules F9. Ces résultats suggèrent que les interactions Le^x-Le^x peuvent indifféremment s'établir à partir de glycolipides ou de glycoprotéines, tant que l'interaction est de type multivalente [Boubelik, M. et al., 1998]. De nombreuses études ont démontré que les glycosphingolipides s'organisaient en microdomaines (clusters) à la surface des cellules [Tillack, T.W. et al., 1983; Rock, P. et al., 1990]. Ces microdomaines seraient donc à même de supporter les interactions multivalentes GSL Le^x-GSL Le^x entre cellules.

La mise en place de ces interactions sucre-sucre dès les premières phases de l'embryogenèse serait un pré-requis à l'induction de la différentiation cellulaire. Elles permettraient l'adhésion rapide de cellules identiques avant l'implication d'autre interactions plus fortes basées sur les protéines d'adhésion suite à l'induction de signaux intracellulaires. Cet aspect particulier des interactions sucre-sucre a été étudié plus en détail dans le cas des interactions GM₃-Gg₃ ainsi que nous allons l'exposer dans le chapitre consacré aux interactions hétérophiliques (Chap. 22.).

22. Interactions hétérophiliques

221. Adhésion dépendante de GM₃

Les glycosphingolipides sont impliqués dans de nombreux phénomènes d'adhésion, à la fois *via* des interactions sucre-protéine comme dans le cas des sélectines et galectines, mais également *via* des interactions sucre-sucre. Au contraire des interactions homophiliques Le^x-

Le^x, les interactions sucre-sucre dépendantes du sialosyllactocéramide GM₃ (NeuAca2-3Gal
\beta1-4Glc
\beta1-Cer) sont de type hétérophilique. En particulier, il a été démontré que GM3 interagissait spécifiquement sous forme multimérique avec d'autres sphingolipides: Gg3 (GalNAc\beta1-4al\beta1-4Glc\beta1-Cer) et à un moindre degré LacCer (Gal\beta1-4Glc\beta1-Cer) et Gb4 (GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer)- [Kojima, N. & Hakomori, S., 1989; Kojima, N. & Hakomori, S., 1991a]. Par contre une interaction négative a été observée entre molécules de GM₃ [Kojima, N. & Hakomori, S., 1991b]. Comme dans le cas des interactions Le^x-Le^x, les interactions dépendantes de GM3 nécessitent la présence d'ions divalents, et sont basées sur la reconnaissance de zones hydrophobes complémentaires [Kojima, N. & Hakomori, S., 1989]. Ces interactions spécifiques induisent non seulement l'adhésion mais amplifient également l'étalement et la mobilité des cellules qui expriment GM₃ [Kojima, N. & Hakomori, S., 1991a]. Ainsi, l'interaction initiale à la base de l'adhésion des mélanocytes -cellules qui expriment des quantités importantes de GM₃- sur les cellules endothéliales non activées serait dépendante d'interactions sucre-sucre entre microdomaines riches en GM₃ et microdomaines riches en LacCer [Kojima, N. et al., 1992]. Dans des conditions expérimentales mimant l'environnement micro-vasculaire, les interactions GSL-GSL sont beaucoup plus fortes que les interactions protéine-protéine et protéine-sucre engageant les intégrines et les lectines. Ainsi, il a été proposé qu'au cours de l'adhésion des cellules cancéreuses sur l'endothélium micro-vasculaire, les interactions GSL-GSL précèderaient celles dépendantes des protéines d'adhésion (sélectines, ICAM). Ces dernières renforceraient les interactions GSL-GSL après activation des endothéliums. De plus, l'adhésion des cellules riches en GM₃ sur les cellules endothéliales via LacCer stimulerait la migration trans-endothéliale de ces dernières,

Les fonctions d'adhésion des microdomaines riches en GSL sont directement couplées à des fonctions transductrices de différents signaux intracellulaires. Ainsi, l'adhésion dépendante de GM₃ induit l'augmentation de la phosphorylation des tyrosines de différents récepteurs intracellulaires (cSrc, FAK) et de la fixation de GTP à Rho A et Ras dans les mélanocytes [Iwabuchi, K. *et al.*, 1998]. Tous ces récepteurs interagissent directement avec les microdomaines des GSL et sont co-purifiés avec ces derniers. Ces signaux intracellulaires pourraient être à l'origine des modifications physiologiques des cellules métastatiques suite à leur adhésion aux cellules endothéliales [Hakomori, S., 2001].

phénomène à la base de la dissémination des cellules métastatiques par voie sanguine.

Dernièrement, il a été montré que l'adhésion du spermatozoïde de truite sur l'œuf était dépendante d'interactions spécifiques entre un GM₃ désaminé ((KDN)GM₃) exprimé en quantité importante sur la surface apicale du spermatozoïde et la partie glucidique du Gg₃ dans l'enveloppe de l'œuf [Yu, S. *et al.*, 2002]. La nature exacte du composé portant l'épitope glucidique complémentaire de (KDN)GM₃ n'a pas été définie; il pourrait s'agir soit du Gg₃ lui même soit d'une glycoprotéine de type mucine présente dans l'enveloppe vitelline de l'œuf riche en glycoprotéines (se référer au chapitre d'introduction des O-glycannes I. 12.). Contrairement aux autres interactions sucre-sucre, l'adhésion du spermatozoïde *via* (KDN)GM₃ ne nécessite pas la présence de Ca²⁺ ou de Mg²⁺, mais est renforcée par la présence d'ions Mn²⁺.

222. Interaction galactocérébroside-cérébroside sulfate (GalCer-CBS)

De manière analogue aux interactions dépendantes de GM3, des interactions spécifiques GalCer-CBS ont été mises en évidence par l'observation des propriétés d'auto agrégation de liposomes constitués par ces glycolipides [Hakomori, S. et al., 1991]. Bien que ces interactions soient le résultat d'interactions entre leur partie glycannique, il a été noté que la nature des chaînes lipidiques en influençait la force. En particulier la force d'interaction augmentait proportionnellement à la longueur des acides gras qui constituaient la partie céramide, probablement grâce à une meilleure accessibilité du glycanne à la surface des liposomes [Stewart, R.J. & Boggs, J.M., 1993]. Il a été proposé par les auteurs de ces travaux que les interactions GalCer-CBS seraient essentielles à la compaction de la myéline dans les systèmes nerveux central et périphérique. A ce titre, il mettaient en évidence une corrélation significative entre le taux de compaction de la myéline chez différentes espèces animales et la présence de GalCer et de CBS dans leurs membranes, ce qui suggère que la présence simultanée de ces deux éléments serait indispensable à l'obtention d'une myéline compactée. Cette hypothèse est en accord avec le fait que la myéline de souris déficiente en synthèse de galactocérébrosides est partiellement décompactée dans le système nerveux central, alors qu'elle présente une ultrastucture normale dans le système nerveux périphérique [Dupree, J.L. et al., 1998]. De plus l'analyse électro-physiologique démontrait des déficiences de conduction des axones, corrélées avec une réduction du pouvoir isolant de la myéline [Coetzee, T. et al., 1996].

3. Agrégation des cellules d'éponge

31. Historique

Les éponges sont les représentants les plus primitifs des animaux pluricellulaires. Ils ne possèdent aucun organe au sens classique du terme, mais leurs tissus sont constitués de quelques classes de cellules très mobiles au sein d'une matrice extracellulaire abondante. Les éponges ont été le premier modèle d'expérimentation de la notion de reconnaissance cellulaire. Dès 1907, Wilson a décrit l'existence d'un phénomène de réagrégation spécifique d'espèce des cellules d'éponges qui avaient été dissociées mécaniquement par passage au travers d'un tissus fin [Wilson, H.V., 1907]. Après désagrégation, les cellules prenaient contact avec le substrat, étendaient des pseudopodes et se mouvaient les unes vers les autres. Elles entraient ainsi en contact, adhéraient les unes aux autres et formaient des agrégats pour régénérer finalement des répliques fonctionnelles en miniature des éponges d'origines [Galtsoff, PS., 1925]. L'adhésion des cellules entre elles nécessitait la présence de Ca²⁺. L'aspect spécifique à l'espèce de la réassociation avait déjà été observé par Wilson (1907): en désagrégeant des espèces d'éponge de couleurs différentes, il constatait que les cellules se réagrégeaient uniquement avec des cellules de même couleur. En 1954, Spiegel préparait des sérums contre deux espèces, M. prolifera et Cliona celata, et observait que le sérum anti-M. prolifera n'inhibait que la réagrégation de M. prolifera mais non celle de C. celata, et vice versa [Spiegel, M. 1954]. Ceci démontrait clairement l'implication de composants présents à la surface cellulaire dans l'adhésion spécifique à l'espèce des cellules d'éponge. Malgré ces exemples frappants, la ségrégation inter-spécifique parfaite n'est pas une règle générale chez les éponges. Dans la plupart des cas, les cellules dissociées d'espèces différentes s'associent d'abord de manière non spécifique, puis ségréguent partiellement ou complètement dans un deuxième temps [Fernandez-Busquets, X. and Burger, M.M., 1999]. Par contre, la formation d'une éponge chimère fonctionnelle n'a jamais été observée.

Suite aux travaux de Spiegel (1954) qui avaient suggéré l'existence de composés impliqués dans l'agrégation à la surface des cellules, Humphreys observait que des cellules d'éponges lavées à l'eau de mer sans Ca^{2+} ni Mg^{2+} perdaient leurs capacités d'adhésion [Humphreys, T., 1963].



256





ПM

L'addition du surnageant de lavage leur restituait leurs capacités d'adhésion spécifique, ce qui démontrait l'existence de facteurs d'agrégation associés aux cellules indispensable aux interactions cellulaires. Ces facteurs d'agrégation ont ensuite été partiellement purifiés de plusieurs espèces d'éponge. Tous contenaient des macromolécules qui présentaient des masses moléculaires de plusieurs millions de Daltons et une ultrastructure proche de celle des protéoglycannes [Müller, W.E.G & Zahn, R.K., 1973].

32. Nature des interactions cellulaires

Les molécules responsables de l'agrégation ont été purifiées de plusieurs espèces et ont été appelés glyconectines. Les glyconectines de l'espèce M. prolifera (GN1) sont les mieux caractérisées de ces composés. Cette espèce d'éponge présente un pouvoir de réagrégation très spécifique comparé à celui d'autres espèces, et ses glyconectines sont très facilement isolées, ce qui en fait un modèle d'étude idéal. GN1 est constitué d'un complexe multi-protéique de Mr $= 2 \times 10^7$, composé à 60-70% (pds/pds) par des sucres. L'examen au microscope électronique a révélée qu'elle présentait une structure très particulière en couronne d'un diamètre d'environ 200 nm, de laquelle radiaient une vingtaine de bras de 180 nm de long (Fig. C1) [Humphreys, S. et al., 1977; Dammer, U. et al., 1995]. Elle contient plusieurs centaines de sites de liaison pour Ca²⁺ qui sont indispensables à la cohésion du complexe multiprotéique et à ses fonction d'adhésion [Jumblatt, J.E. et al., 1980]. Il est à noter que les facteurs d'agrégation isolés d'autres espèces telles que Halichondria panicea, Ficulina ficus ou C. chelata présentent des structures linéaires et non circulaires comme celui de M. prolifera (Fig. C1) [Humphreys, S. et al., 1977; Jarchow, J. et al., 2000]. GN1 est constitué par au moins deux protéines distinctes, MAFp3 et MAFp4 associées de façon non covalente [Fernandez-Busquets, X. & Burger, M., 1997; Fernandez-Busquets, X. et al. 2000]. Sa protéolyse complète libère deux polysaccharides de 6 et de 200 kDa (g6 et g200) qui sont chacun impliqués dans des interaction différentes [Misevic, G. & Burger, M., 1990; Misevic, G. & Burger, M., 1993]. Ces polysaccharides sont tous deux composés de Fuc, Man, Gal, GlcNAc et GlcA qui formeraient des séquences de répétitions distinctes de celles qui caractérisent les protéoglycannes. Néanmoins, leur structure exacte reste largement inconnue. Seuls quatre oligosaccharides issus de l'hydrolyse de GN1 ont été séquencés, mais leur représentativité de la structure polysaccharidique et leur arrangement en son sein n'ont pas été définis (tableau C1) [Spillmann, D. et al., 1993; Spillmann, D. et al., 1995].



Fig. C2: Modèle des interactions sucres-sucres impliquées dans l'adhésion des cellules d'éponges. RM, récepteur membranaire putatif; 68/210, protéines de haute affinité pour la glyconectine; CP,corps protéique. D'aprés Misevic, M. and Burger, M., 1993.

Oligosaccharides	Block 1	Block 2
py(4,6)Galβ1-4GlcNAcβ1-3Fuc	+	_
Gal ^{β1-4} GlcNAc ^{β1-3} Fuc	-	_
SO ₃ -3GlcNAcβ1-3Fuc		+
$Gal\alpha 1-2[SO_3-3]Gal\beta 1-4GlcNAc\beta 1-3Fuc$	-	-

Tableau C1: Structure des oligosaccharides générés par hydrolyse ménagée de GN1 et leur réactivité envers les anticorps bloquant l'interaction GN1-GN1 Block 1 et Block 2, d'après Spillmann, D. et al., 1995.

La répartition des parties protéiques et polysaccharidiques dans les différents domaines de GN1 (anneau central et bras) et donc la composition de ces derniers restent également sujets à débat. Néanmoins, la structure globale de la molécule laisse présager que les bras étendus seraient impliqués dans les interactions inter-moléculaires.

GN1 induit l'agrégation des cellules d'éponge via deux phénomènes d'interaction indépendants [Jumblatt, J.E. et al., 1980] qui impliquent deux domaines distincts de la molécule: (1) une interaction indépendante du Ca²⁺ entre GN1 et la surface cellulaire, (2) une interaction dépendante du Ca²⁺ entre molécules de GN1. L'observation qu'un anticorps (Block 1) dirigé contre un épitope glycannique de GN1 bloquait spécifiquement l'auto-association de MAF a dans un premier temps démontré que le domaine glycannique était impliqué dans cette interaction [Misevic, G. et al., 1987]. L'épitope reconnu par Block 1 a été ultérieurement identifié comme un oligosaccharide présentant la structure originale suivante: py(4,6)Galß1-4GlcNAcB1-3Fuc [Spillmann, D. et al., 1993]. g200 a été identifié comme le glycanne responsable de l'auto-association de GN1 en présence de Ca²⁺ [Misevic, G. & Burger, M., 1993]. Ce glycanne purifié, une fois couplé à des billes de verre ou encore multimérisé était capable de s'auto-agréger en l'absence de toute protéine, ce qui démontrait définitivement que l'adhésion homotypique des cellules d'éponge était dépendante d'interactions homophiliques sucre-sucre. L'interaction entre GN1 et la surface cellulaire s'effectuerait quant à elle par l'intermédiaire du glycanne g6 sur un ou plusieurs récepteurs membranaires [Misevic, G. & Burger, M., 1990]. Un modèle d'action de GN1 a ainsi pu être proposé dans lequel deux molécules bifonctionnelles attachées à deux cellules différentes relieraient ces deux dernières par l'intermédiaire d'interactions homophiliques g200-g200 (Fig. C2). La nature des récepteurs membranaires de GN1 est encore incertaine. Deux protéines non membranaires de 68 et 210 kDa présentant une forte affinité pour GN1 ont été isolées de la matrice extracellulaire des cellules de M. prolifera et caractérisées [Varner, J., 1995; Varner, J., 1996]. Toutes deux inhibaient l'auto-agrégation des cellules d'éponge en perturbant l'interaction GN1-cellule et pourraient servir de ponts moléculaires entre GN1 et d'éventuels récepteurs membranaires encore non identifiés.

Les interactions homophiliques GN1-GN1 et hétérophiliques GN1-cellule sont de nature polyvalente. Ceci a été démontré par le fait que des chaînes glycanniques individuelles de g200 ou g6 ne présentaient aucune affinité mesurable entre elles ou avec les cellules, et n'inhibaient pas l'agrégation des molécules entières [Misevic, M. & Burger, M., 1986; Misevic, M. & Burger, M., 1990, Misevic, M. & Burger, M., 1993]. Les propriétés d'adhésion de ces molécules sont par contre restaurées après polymérisation de leurs composants individuels. Leur affinité est strictement proportionnelle à leur valence: la polymérisation de glycopeptides issus d'une digestion trypsique de la molécule entière a permis d'obtenir un polymère de taille équivalente à celle de la molécule d'origine et qui présentait une affinité similaire. La polyvalence des interactions sucre-sucre a été directement observée en microscopie de force atomique [Dammer, U. et al., 1997]. La force d'interaction entre deux GN1s en présence d'ions Ca²⁺ a ainsi été estimée à une moyenne de 125 pN, pouvant atteindre 400 pN. La courbe d'interaction en microscopie de force atomique lors de l'éloignement contrôlée de deux molécules permet d'observer des sauts successifs d'une ampleur de 40 ± 15 pN qui correspondraient à la séparation de deux bras préalablement impliqués dans l'interaction inter-moléculaire. L'adhésion de deux bras serait elle même la résultante de plusieurs centaines d'interactions individuelles de très faibles amplitudes. Pas moins de 2500 épitopes glycanniques par molécule sont potentiellement impliqués dans ces interactions [Misevic, G. & Burger, M., 1993].

L'utilisation des anticorps bloquants Block 1 et Block 2 a démontré que l'adhésion homophilique GN1-GN1 impliquait deux épitopes glycanniques acides (voir C1). Dernièrement, il a été démontré par résonance plasmonique de surface que le disaccharide sulfaté synthétique (SO3-3GlcNAcB1-3Fuc) multimérisé sur une protéine porteuse s'autoagrégeait de manière spécifique mais n'interagissait pas avec d'autres déterminants glycanniques sulfatés [Haseley, S.R. et al., 2001]. Ceci confirmait définitivement que GN1 interactions homophiliques s'auto-associait via des sucre-sucre impliquant des oligosaccharides acides de structure définie. Néanmoins, les bases moléculaires de ces interactions sont encore inconnues. La spécificité d'interaction observée démontrait toutefois qu'elles n'étaient pas le fait d'une simple attraction de type ionique.

Comme nous l'avons exposé, de nombreux travaux ont permis de démontrer que le facteur d'agrégation isolé de *M. prolifera* était directement impliqué dans l'association

spontanée des cellules de cette espèce d'éponge. Toutefois, une question restait en suspens: ce facteur était il également responsable de la spécificité d'association des cellules en fonction de l'espèce? Pour y répondre définitivement Misevic et collaborateurs ont purifié les glyconectines de trois espèces d'éponges qui présentent une parfaite ségrégation (M. prolifera, Halichondria panicea et Cliona chelata) et les ont couplées à des billes de latex colorées [Misevic, M. & Popescu, O., 1995]. En présence de concentrations physiologiques en Ca²⁺, sous agitation, les billes s'associaient de façon strictement spécifique à l'espèce d'origine des molécules adsorbées. Ceci démontrait que les glyconectines étaient responsables non seulement de l'agrégation des cellules d'éponge, mais également de la sélectivité d'adhésion des cellules en fonction de l'espèce, selon toutes probabilités au travers d'interactions sucresucre. Là encore, les bases moléculaires responsables de cette sélectivité sont totalement inconnues. Plusieurs paramètres pouvant être à l'origine de cette sélectivité d'interaction ont été proposés: (1) les différences de séquence des structures glycanniques, (2) l'espacement entre les charges négatives au sein des glycannes, (3) le type de charge (acides carboxyliques, sulfates...), (4) l'hydrophobicité des glycannes, (5) le degré de répétitivité des épitopes impliqués dans l'interaction, (6) la distance entre les chaînes polysaccharidiques sur la protéine porteuse. Pour appréhender ce phénomène et en comprendre les bases moléculaires, il paraît donc essentiel d'obtenir des informations structurales sur les chaînes polysaccharidiques des différentes glyconectines choisies comme modèle d'étude.

4. Conclusion. Importance physiologique des interactions sucre-sucre

Le modèle d'interaction cellulaire sucre-sucre est basé sur la présence de multiples sites d'interactions de très faible affinité pour obtenir une adhésion de haute affinité, ou plus exactement de haute avidité. Au contraire des sites d'interactions protéine-protéine qui présentent une affinité individuelle importante, ce modèle permet un contrôle subtil de l'affinité totale du système au travers d'une régulation de la polyvalence. Cette dernière peut s'effectuer par diverses voies: contrôle de la densité des épitopes, de la force ionique, de la structure des glycannes exprimés. Ce système apparaît donc a priori comme très flexible et permet des modification rapides des forces d'interactions mises en œuvre en fonction des besoins de la cellule: adhésion, déplacement, examen de surface. Les interactions sucre-sucre homo- ou hétérophiliques peuvent être schématisées simplement sur la base du fonctionnement des fermetures "velcro". Ainsi, l'existence d'un système d'adhésion flexible permet par exemple d'expliquer la migration des cellules d'éponges au sein de l'organisme malgré l'existence d'interactions constantes avec les cellules environnantes. Dans le même ordre d'idée, il permet de fournir une adhésion cellulaire dynamique compatible avec les différents phénomènes de développement de la cellule tel que l'embryogenèse ou la formation de la myéline.

D'un point de vue général, l'avènement des interactions sucre-sucre, par la simplicité du phénomène mis en œuvre -deux sucres, identiques ou différents, interagissent sans intermédiaire- a sans aucun doute été un événement déterminant dans l'acquisition de la pluricellularité à partir des cellules primordiales. Il permet d'expliquer en partie que deux cellules strictement identiques aient pu interagir via des interactions homophiliques pour former les premiers agrégats cellulaires à la base de l'essor des êtres multicellulaires. Ce point de vue est renforcé par le fait que les plus primitifs des organismes pluricellulaires, les éponges, ont conservé ce mode d'association au travers de macromolécules qui ont sûrement elles mêmes beaucoup évolué depuis l'avènement de ce type d'association. De manière significative, ce processus a également survécu dans les premiers stades du développement embryonnaire, chez des cellules encore non différenciées, et dans les cellules cancéreuses dédifférenciées. Néanmoins, la portée exacte de ces interactions est encore totalement méconnue, en partie due aux difficultés inhérentes à leur observation. La découverte d'interactions sucre-sucre indispensables à la compaction de la myéline, présentes exclusivement chez les vertébrés supérieurs, laisse présager leur implication dans de nombreux autres phénomènes physiologiques.

II PROBLEMATIQUE et RESULTATS

1. Présentation des travaux

Nos travaux au sein de cette thématique ont consisté en l'étude préliminaire de la structure des glyconectines isolées des espèces d'éponge M. prolifera (GN1) Halichondria panicea (GN2) et Cliona chelata (GN3). Comme nous l'avons exposé précédemment, ces composés sont directement responsables de l'auto-association des cellules dissociées de ces trois espèces. Cette association ne dépend que de la présence de Ca²⁺ et peut s'effectuer indifféremment à partir de cellules vivantes, de cellules fixées par le glutaraldéhyde, de glyconectines purifiées, voire de fragments glycanniques isolés de glyconectines, après multimérisation. Les études effectuées sur M. Prolifera ont démontré que l'agrégation des molécules de GN1 était le résultat d'interactions sucre-sucre homophiliques, entre leurs domaines glycanniques respectifs. Deux courts déterminants glucidiques acides, impliqués dans ces interactions, ont ainsi été identifiés. Les glyconectines GN2 et GN3 isolées des espèces Halichondria panicea et Cliona chelata présentent des propriétés physicochimiques très comparables à celles de GN1 et se comportent d'une manière identique en présence de Ca²⁺. Par extension, il a donc été proposé que leurs capacités d'auto-agrégation étaient le résultat d'interactions sucre-sucre, mais ce point précis n'a jamais été clairement démontré comme dans le cas de GN1.

Les glyconectines sont responsables non seulement de l'adhésion des cellules d'éponge entre elles, mais également de leur ségrégation en fonction de l'espèce. Les trois glyconectines GN1, GN2 et GN3 présentent la propriété de ségréguer complètement quand elles sont en mélanges. Elles apparaissent ainsi comme un modèle d'étude parfait des phénomènes responsables de l'auto-agrégation spécifique d'espèce des cellules d'éponge. A ce jour, très peu d'informations sont disponibles quant à la structure et l'organisation de ces molécules. Les études structurales effectuées concernent uniquement l'espèce *M. prolifera* (GN1), mais reste néanmoins très fragmentaires. En particulier, la structure des domaines glycanniques de GN2 et GN3, directement impliqués dans la reconnaissance et l'agrégation des cellules, est encore totalement inconnue. De fait, aucune donnée ne permet actuellement de comprendre les bases moléculaires responsables de la spécificité d'interaction que manifestent ces composés. Est elle le résultat de l'existence de séquences glycanniques ou de charges différentes dans les trois glyconectines, ou alors de l'organisation différente des mêmes éléments de base au sein de chaque complexe supra-moléculaire?

Les travaux que nous avons effectués représentent la première phase d'une série d'études sur la structures des domaines glycanniques des glyconectines. Leur objectif premier était d'établir si GN1, GN2 et GN3 présentaient des domaines glycanniques identiques ou non, et le cas échéant, d'évaluer l'ampleur des différences structurales. Ce travail devait servir de base à la définition d'une stratégie cohérente d'étude de la structure et de l'organisation de ces molécules. Les glyconectines sont des complexes supra-moléculaires qui présentent des masses de plusieurs millions de Daltons. Les études antérieures ont révélé que le domaine polysaccharidique de GN1 était constitué d'au moins deux polysaccharides de 6 et 200 kDa. La structure de ce domaine a été très partiellement déterminée (Tableau C1) mais l'arrangement des oligosaccharides séquencés au sein des polysaccharides est encore inconnu. Au laboratoire, des études préliminaires sur la structure de g6 et g200 par RMN ont été effectuées. Néanmoins, l'étude des deux polysaccharides natifs par RMN a fourni peu d'informations en raison de leur complexité et de leur hétérogénéité intrinsèque. En conséquence, dans l'optique de comparer la structure des polysaccharides des trois glyconectines, ces derniers ont été fragmentés dans des conditions identiques avant analyse. L'ensemble des résultats obtenus sont présentés dans le chapitre suivant. Ces résultas ont été partiellement incorporés dans un article en préparation, qui est présenté pour référence en annexe.

2. Résultats

MATERIAL AND METHODS

Isolation of Glyconectin Proteoglycans. Proteoglycans were extracted from fresh cuts of sponges with artificial Ca²⁺ and Mg²⁺free seawater (462 mM NaCl, 10.7 mM KCl, 7 mM Na2SO4 and 2.1 mM NaHCO3) at +4°C for 12h. Proteoglycans were purified by sequential centrifugation, 20 mM CaCl₂ precipitation and ultracentrifugation [1,2].

Chemical degradation of glyconectins. Materials were repetitively hydrolysed in mild conditions for short times (TFA 0.1 M, 80°C, 1 h.). After each hydrolysis step, four volumes of ethanol were added and liberated oligosaccharides were collected in the ethanol soluble fraction. Release of oligosaccharides from glyconectins was assessed after each step by TLC analysis of ethanol soluble fraction on silica gel using /acetic acid/H₂O (40/30/20) as solvent. Typically, four hydrolysis steps were sufficient to liberate all hydrolysable material from glyconectins. Prepared oligosaccharides were further purified by gel filtration on a Bio-Gel P2 column (Bio-rad) before analysis. Oligosaccharides were reduced by either NaBH₄ for GN1 or NaBD₄ for GN2 and GN3 in aqueous NH₄OH 1% solution, and eventually permethylated according to Ciucanu [5].

Matrix-assisted laser-desorption ionization-time of-flight (MALDI-TOF) MS. Molecular mass of the oligosaccharides was measured by MALDI-MS on a Vision 2000 time-of-flight instrument (Finnigan Mat) equipped with a 337 nm UV laser: 1 μ l of sample at a concentration of 100 pmol/ μ l was mixed with an equal volume of matrix and allowed to crystallize. We used 2,5-dihydroxybenzoic acid and 3-aminoquinoline for neutral and acidic oligosaccharides, respectively.

Electrospray mass spectrometry. All MS measurements were carried out in positive-ion mode on a triple quadrupole instrument (Micromass Ltd, Altrincham, UK) fitted with an atmospheric pressure ionization electrospray source. A mixture of polypropylene glycol was used to calibrate the quadrupole mass spectrometer. Sample were dissolved in methanol/water (50/50) for underivated oligosaccharide or acetonitrile for permethylated oligosaccharides at a concentration of 10 pMole. μl^{-1} and infused using the nanoflow probe at 50 nl.min⁻¹. Quadrupole was scanned from 200 to 2000 Da with a scan duration of 3 s and a scan delay of 0.1 s. The samples were sprayed using 1.4 KV needle voltage and the declustering (cone) was typically set at 70 V. For collision-induced dissociation (CID) experiments, the pressure of argon in the cell was set at 4.10⁻³ mbar and the collision energy was set to values ranging from 25 to 75_eV.

NMR spectroscopy. The NMR experiments were performed on Bruker® ASX400 and DMX600 spectrometers both equipped with a 5mm ¹H/¹³C mixed probe-head, operating in the pulse Fourier transform mode controlled by an Aspect 3000 computer. Each oligosaccharide

sample was dissolved in 400 μ L ²H₂O after three exchanges with ²H₂O (99.97 % atom ²H, Euriso-top, CEA group, Gif-sur-Yvette France) and intermediate lyophilisations. The oligosaccharides were analyzed at 300 °K. The chemical shifts (δ) were referenced to internal acetone (δ^{1} H=2.225 and δ^{13} C=31.55 p.p.m. in the condition used). Two dimensional homonuclear (COSY90, TOCSY) and heteronuclear (HMQC) experiments were performed by using standard Bruker® pulse programs. The main pulses and variable delays were optimized for each pulse program and sample.

Gas-chromatography. Monosaccharides were analyzed by GC/MS as per-trimethylsilyl [3] and per-heptafluorobutyryl derivatives [4]. Methylation of oligosaccharides was conducted according to Ciucanu and Kerek [5] before analysis on GC/MS.

Other analytical methods. Size of the oligosaccharides was assessed by gel filtration on a Bio-Gel P2 or P4 column (Bio-Rad) equilibrated in acetic acid 0,5 % and calibrated with hydrolyzed dextran, in combination with thin layer chromatography as already described.

RESULTS

Carbohydrate compositional analysis of native glyconectins. The three purified native glyconectins were analysed for their sugar compositions. They were all found to contain fucose, galactose, mannose, glucuronic acid, N-acetylglucosamine and N-acetylglalactosamine in variable amounts (table 1). Arabinose was detected in one sample out of three, while a single type of pyruvated hexose was detected in the two others. The last residue was visualised in gas chromatography as methyl-glycoside heptafluoryl derivative, as two peaks tentatively attributed to its α and β anomers. It was first characterised as a pyruvated hexose in EI-MS, owing to a very intense [M-COOMe] ion at m/z 611, and in CI-MS as a [M+NH4] m/z at 688. The exact nature of this residue was established as 4,6-pyruvylated galactose by TMS derivatization of methyl glycoside and comparison with standard EI-MS spectra from literature (Fig. 1) [6]. Indeed, spectra was dominated by an ion at m/z 243, characteristic of a 4,6-O-(1-carboxyethylidene)-D-galactose structure [6], and by an intense [M-COOMe] at m/z 363. The intense ion at m/z 204, resulting from the cleavage of two adjacent trimethylsilyl groups, also characterized the 4,6-linkage position of the pyruvate group. Subsequently, this

attribution was confirmed by NMR experiments of glyconectin hydrolysis fragments from GN2 (see below). This is in agreement with previous work that identified a 4,6pyr-Gal containing oligosaccharide, obtained by hydrolysis of glyconectin from *M. prolifera* [7].

The stability of the 4,6-O-(1-carboxyethylidene)-galactose pyranose ring to methanolysis [6], enabled us to estimate, among the other monosaccharides, the amount of this particular component in each sample. As shown in table 1, glyconectins from the three analysed species are constituted by the same basic components, but showed substantial differences regarding to their sugar compositions. GN1 species was characterised by a high quantity of fucose (about one third of the total monosaccharide content) and minute amounts of Py(4,6)Gal, while GN2 was characterised by a low proportion of fucose and a percentage of Py(4,6)Gal three times as important as in GN1. On the other hand, GN3 was the only one to contain arabinose (10% mol/mol) as well as important quantities of fucose but no Py(4,6)Gal at all. This last species was also the only one to exhibit significant amounts (> 5%) of GalNAc.

Generation of oligosaccharides. In order to compare their carbohydrate moieties, the three glyconectins were fragmented by partial acid hydrolysis. Materials were repetitively hydrolysed in mild conditions for short times (TFA 0.1 M, 80°C, 1 h.). After each hydrolysis step, four volumes of ethanol were add and liberated oligosaccharides were collected in the ethanol soluble fraction. These hydrolysis conditions were devised to preserve most hexose and N-acetylhexosamine linkages as well as pyruvate and sulfate groups, but to cleave more fragile fucose and pentose linkages. Besides, repeated removing of hydrolysis products was adopted in an attempt to preserve some of the labile sequences involving Fuc or Ara residues.

Release of oligosaccharides from glyconectins was assessed after each step by TLC analysis of ethanol soluble fraction. This permitted to observe that liberated oligosaccharides showed a constant size from the first to the last hydrolysis step, suggesting that identical material was sequentially cleaved off glyconectins by partial hydrolyses. Consequently, all ethanol soluble fractions of each species were pooled for analysis. Typically, four hydrolysis steps were sufficient to liberate all hydrolysable material. After completion of hydrolyses, ethanol insoluble fractions were extracted by water in order to separate the insoluble remaining glycoproteins from eventual ethanol-insoluble oligosaccharides. In GN1, a minor water soluble fraction (6% of the total initial sugar content) was extracted this way, while in GN2 and GN3 no water soluble component were observed, as revealed by TLC and confirmed by sugar analyses (data not shown). This suggested that the vast majority of hydrolysed sugars were ethanol soluble and that the ethanol insoluble fraction was essentially composed of un-hydrolysed, denatured glycoprotein.

In similar conditions, the three glyconectins exhibited distinct susceptibilities to acid hydrolysis, as reflected by the proportions of carbohydrate in the ethanol soluble fraction after hydrolysis: 68 % of the total carbohydrate content for GN1, 56 % for GN2 and 41 % for GN3. Systematic analysis of each fraction (table 2) revealed that hydrolysed materials had substantially distinct sugar compositions compared to un-hydrolysed materials. In particular, cleaved oligosaccharides were systematically enriched in fucose and arabinose, if any (table 2), compared to un-hydrolysed fractions. This reflect the preferential cleavage by partial hydrolysis of domains of the molecule rich in acid-labile linkages. On the contrary, glucuronic acid was predominantly retrieved in hydrolysis resistant fractions. All these data strongly suggested that the carbohydrate moieties of glyconectins are not constituted by homogeneous polymers, but rather by structurally unrelated domains: an acid labile, presumably external, domain and an acid resistant polysaccharidic core. This work concentrated on the comparison of the acid labile domains of glyconectins from the three species by sequencing the hydrolysed oligosaccharides.

Sequencing of GN1. From TLC analysis, hydrolysate from GN1 was composed of oligosaccharides which apparent sizes range from DP 2 to 6 (Fig. 2.). It almost exclusively contained Fuc, Gal and GlcNAc. Traces of Man, GlcA and Py(4,6)Gal were also detected. Considering that Gal represents about 95 % of the total detected hexoses, all hexose residues within the described sequences were assimilated to Gal. This fraction was analysed in ES-MS/MS in positive and negative modes, before and after reduction with NaBH₄ in order to ease sequencing of its components (Fig. 3). Structural information collected from these experiments are summarised in table 3. Seven different oligosaccharides (labelled A to G) were identified in negative modes. MS/MS spectra showed that they were all characterised by an intense ion at m/z 97, attributable to the formation HSO₄ groups during fragmentation which demonstrated that all oligosaccharides were sulphated.

Compounds B, C and G exhibited related sequences. The presence of an intense ion at m/z 282 in their fragmentation spectra, corresponding to [(SO₃)HexNAc-H₂O] localised the sulphate group on GlcNAc residue in these three compounds. The smallest oligosaccharide, B, showed a very simple fragmentation pattern characterised by an parent ion at m/z 446 (native) or 448 (reduced) and an B type ion [M-deHex] or [M-deHex-ol] at m/z 282. This unambiguously demonstrated that the deHex residue was in terminal reducing position, which characterised compound B as the dimer (SO₃)GlcNAc-Fuc. Compound C was typified by the occurrence of a Fuc residue at both the reducing and the non reducing ends, as established by the presence of Y [M-deHex] ions at m/z 445.5 and 447.8 for unreduced and reduced oligosaccharides, respectively and secondary ions at [M-deHex-deHex] and [M-deHex-deHex-ol] (Fig. 4). Thus, sequence of compound G was established as Fuc- [SO₃]GlcNAc-Fuc. Then, fragmentation pattern of compound G, was characterised by the simultaneous presence of an Y type ion [M-HexNAc] at m/z 592 and an B type ion [M-deHex] at m/z 631. On this basis, compound G was described as an extension of compound C with an additional GlcNAc residue in terminal non-reducing position.

On the other hand, compounds D, E and F differed from previous compounds by sulphation position on a hexose residue, according to the ubiquitous presence of the $[(SO_3)Hex-H_2O]$ ion at m/z 241. Sequence of compound D was easily established as GlcNAc- $(SO_3)Gal$ -Fuc owing to the presence of [M-deHexol] B ion and [M-HexNAc] Y ion at m/z 444 and 407, respectively. Similarly, for compound E, [M-HexNAcol] B and C ions at m/z 403 and 421, respectively, demonstrated the presence of an GlcNAc residue in terminal-reducing position. Occurrence of the Y ion [M-Hex] at m/z 464 establishes the sequence as Gal(SO₃)Gal-GlcNAc. Finally, compound F was described as the extension of compound E with a Fuc residue in terminal-reducing position, as demonstrated by the [M-deHexol] ion at m/z 606. Simultaneous presence of [M-Hex], [M-deHexol-HexNAc] and [M-deHexol-Hex] at m/z 610, 403 and 444, respectively, confirmed the proposed sequence.

These results are in total accordance with previously published results on the structure of GN1 glyconectin [8]. In particular, the authors described two oligosaccharides which sequences, $SO_3(3)GlcNAc(\beta 1-3)Fuc$ and $Gal(\alpha 1-2)[SO_3(3)]Gal(\beta 1-4)GlcNAc(\beta 1-3)Fuc$, perfectly match with compounds B and F. Furthermore, two new motives were described, illustrated by the sequences GlcNAc-(SO₃)Gal-Fuc and Fuc-(SO₃)Gal-Fuc. The last motif originated from the incomplete hydrolysis of Fuc labile linkages. Electrospray experiment in positive mode on the same fraction only showed few components that could not be sequenced in ES/MS due to low quantities. In particular, a minor ion at m/z 622 was observed. Its calculated composition (1 Py-Gal, 1 HexNAc and 1 deHex) matches with the acid-labile tri-saccharide $Py(4,6)Gal(\beta 1-4)GlcNAc(\beta 1-3)Fuc$ previously described in GN1 [7]. The low amounts of Py(4,6)Gal detected in total oligosaccharidic fraction confirmed that this oligosaccharide was a very minor component.

Altogether, these data suggest that GN1 acid labile domain is not composed by single repetitive short units as observed in glycosaminoglycans. Indeed, we have shown that Fuc residues may be substituted by either GlcNAc, (SO₃)GlcNAc or (SO₃)Gal units. Similarly, (SO₃)Gal units may be substituted by either Gal or GlcNAc residues. Therefore, GN1 polysaccharide present a novel and highly heterogeneous structure.

GN2 sequencing. Unlike GN1, partial hydrolysis by TFA 0.1 M did not release small oligosaccharides but a material which important size prevented it to migrate on a TLC (Fig. 2). It contained large quantities of Gal (>40 % of the total monosaccharides), and Py(4,6)Gal (>20 %), as well as Man, GlcNAc, Fuc and traces of GalNAc. It is noteworthy that the fucose content (7 %) was much lower than the one observed in GN1 hydrolysate (50 %). Gel filtration analysis of this material revealed that its sized ranged from 15 to 20 Glc DP. Mass spectrometry analyses in MALDI or ES/MS did not enable us to precise the exact mass of these compounds, presumably due to their important size. However several relevant structural data were inferred from other analytical techniques.

First, nature of the reducing monosaccharides, was determined by NaB³H₄ labelling and hydrolysis of material. TLC and HPAEC analyses revealed that only galactose and fucose were converted to polyols (Fig. 5). Surprisingly, tritiated Gal-ol was three times as abundant as Fuc-ol, suggesting that more oligosaccharides originated from the cleavage of galactoselinkages than fucose-linkages.

Then, considering the size of the oligosaccharides, the high proportion of Py(4,6)Gal residues suggested that each molecule comprised, in average, more than one of these residues. This means that either oligosaccharides are linear with internal Py(4,6)Gal residues or oligosaccharides have branches with Py(4,6)Gal residues in terminal positions. These hypotheses were first assessed owing to multiple NMR experiments. On the HOHAHA

spectrum, among other signals, we observed a β -anomeric signal at δ 4,817 ppm, which correlated to a broad multiplet at δ 3,656-3,696 ppm and to a pseudo singlet at δ 4,155 ppm, respectively attributed to H-2/H-3 and to H-4 of a β-galactose residue (Fig. 6). H-6 and H-6' were assigned at 8 4,051-4,060 and 3,923-3,942 ppm respectively, owing to an intense correlation between a very shielded H-5 at δ 3,543 ppm and the H-4 and H-6/6'. These ¹H NMR parameters, matched with those previously described for terminal β -D-Py(4,6)Galp [9]. Furthermore, the observation of an intense singlet at $\delta 1.465/26.32$ ppm (¹H/¹³C) was in accordance with the occurrence of a pyruvate group in the oligosaccharide mixture (data not shown). As shown by a ¹H-¹³C HMQC experiment, C-4 and C-6 are noticeably shifted compared to terminal gal residue [10] to δ 72,21 ppm and of to δ 66,37 ppm respectively, suggesting that C-4 and C-6 positions of this residue are substituted by acetal groups, in agreement with published results for the terminal Py(4,6)-Galp-containing acidic exopolysaccharides from Burkholderia cepacia [11] and Lactobacillus rhamnosus [12]. Then, the important negative shift of C-5 to δ 67,62 ppm confirmed that this residue was substituted on C-4 and C-6 positions. Thus, compiled NMR data from GN2 hydrolysate clearly established the presence of terminal-Py(4,6)Gal residues. Furthermore, NMR experiments showed no evidence to substantiate the presence of substituted Py(4,6)Gal residues in the oligosaccharide mixture. Methylation analysis permitted to identify permethylated galactose and permethylated pyruvated hexose in the ration 1:1. This last residue was distinguished owing to a very intense [M-COOMe] ion at m/z 247 in EI-MS (Fig. 1). Fragmentation spectrum of this residue matched perfectly with published data from Bacillus circulans exopolysaccharide [13], confirming its attribution as terminal Py(4,6)Gal. In accordance with NMR data, no substituted Py(4,6)Gal residue was observed in methylation analysis. Altogether, results from composition, linkage and NMR analyses demonstrated that large oligosaccharides released from glyconectins of GN2 by mild hydrolysis are branched structures, capped in equal amount by terminal Py(4,6)Gal residues and Gal residues. Also, the fact that no trace of internal Py(4,6)Gal could be detected in the material released in mild hydrolysis conditions strongly suggest that Py(4,6)Gal residues are always found in terminal non-reducing position on the native glyconectin isolated from GN2.

In order to obtain sequence information from GN2 glyconectin, it was subjected to a recurrent partial hydrolysis in TFA 1 M. Such hydrolysis conditions should preserve pyruvate groups but not eventual sulphate groups. Hydrolysed fragments were collected as previously

described. TLC analysis revealed that hydrolysis generated small quantities of oligosaccharides that were separated from large saccharidic material by gel filtration on Biogel P-4 (Biorad). Half of the oligosaccharidic material was reduced with BD₄Na and permethylated. Native and per-methylated reduced materials were analysed by ES/MS-MS, permethylated material was also analysed in GC/MS. ES/MS analysis of native material in negative mode did not show any ion, suggesting that no negatively charged oligosaccharides were present. In contrast, ES/MS-MS in positive mode permitted to sequence six oligosaccharides, H to M (table 4), while analysis of per-methylated reduced oligosaccharides in positive mode gave the sequence of fifteen compounds (table 5). Only the three major compounds (H, I and J) could be sequenced in GC/MS (table 6), due to the low sensibility of the method. Compiled data from ES/MS-MS and GC/MS permitted to identify sixteen oligosaccharides from GN2 hydrolysate. In ES/MS-MS all native oligosaccharides showed intense [M-COOH+Na] ions at m/z M-44 and [M-PyHex+Na] Y type ions at m/z M-232, whereas all per-methylated oligosaccharides showed [M-PyHex+Na] Y type ions at m/z M-274, demonstrating that all compounds were substituted in terminal position by at least one Py(4,6)Gal residue. Taking into account data obtained from analysis of material released in milder conditions, this tends to indicate that oligosaccharides were always released from the extremities of the molecule and never originated from an internal domain.

Two minor compounds (N and O) were observed in ES/MS only as per-methylated derivatives. Both presented deHex residue in reducing-terminal position, as shown by the Y ion at m/z 246. Furthermore, for compound O, Y ion at m/z 449 demonstrated the occurrence of the dimer Hex-deHex in reducing position, while B and C ions at m/z 501 and 519 demonstrated the occurrence of the dimer PyHex-Hex in terminal non-reducing position. These data established the sequence of N as PyHex-deHex and O as PyHex-Hex-deHex.

All the other compounds were shown to possess either Hex or HexNAc residues in terminal reducing-position. Compounds I, J, K, L and M possessed un-substituted terminal-reducing HexNAc residues as shown by ES/MS-MS experiment of native oligosaccharides, owing to [M-HexNAc+Na] C ion at m/z M-203, [M-HexNAc-COOH+Na] C ion at m/z M-247 and [HexNAc+Na] Y ion at m/z 244. In per-methylated reduced oligosaccharides, HexNAc residue in reducing position was characterised by [M-HexNAcol+Na] C ion and [HexNAcol+Na] Y ion at m/z M-276 and m/z 317 respectively (Fig. 7). Similarly, unsubstituted terminal Hex residues were characterised in compounds H, P, T, U and W owing to [M-Hex+Na] C ion at m/z M-162, [M-Hex-COOH+Na] C ion at m/z M-206 and [Hex+Na] Y ion at m/z 203 for native oligosaccharide and/or to [Hexol+Na] Y ion at m/z 276 for per-

methylated oligosaccharides. In the case of compounds T and U, Hex-ol residues were characterised owing to the secondary fragment at m/z 465 distinctive of the substituted dimer Hex-Hex-ol. On these bases, sequences of the two simplest compounds, H and I, were easily established in ES/MS-MS as PyHex-Hex and PyHex-HexNAc respectively. In GC/EI-MS spectrum of compound H, ions deriving from the fragmentation of Hex-ol residue (m/z at 89, 90, 134, 437 and 480) demonstrated that PyHex residue substituted the terminal hexose residue in C-4 position. Similarly for compound I, fragmentation of HexNAc-ol residue (ions at m/z 89, 131, 435, 479, 523 and 524) clearly established that it was substituted in C-3 position. For compound J, fragmentation pattern established the HexNAc-ol residue was substituted in C-3 position (Fig. 8). Also, we identified a ion at m/z 323 that originates from the fragmentation of internal Hex residue, indicative of a C-3 substitution.

All other compounds appeared as linear or branched variations of oligosaccharides H and I with additional hexose residues. They differed by the number of hexose residues, and the position of substitution in case of branched oligosaccharides. Branched oligosaccharides were differentiated owing to the simultaneous presence in fragmentation spectra of [M-PyHex+Na] and [M-Hex+Na] ions. Compound U was the only one to exhibit two PyHex residues, according to its calculated composition. Successive m/z losses of 274 demonstrated that both residues were in terminal positions. Branching position was deduced from recurring Y type and secondary fragmentations of the oligosaccharides.

Compiled data from structural analyses conducted on oligosaccharides released in two different hydrolysis conditions strongly suggest that polysaccharide acid-labile domain of GN2 glyconectin is constituted by an highly heterogeneous, branched polymer. The saccharidic core is mostly constituted by a polymer of Gal, Man, GlcNAc and Fuc residues in internal position, and ended by Gal and Py(4,6)Gal residues. No real repetitive unit could be deduced from the fine sequencing of released oligosaccharides.

GN3 sequencing. As shown in table 2, GN3 hydrolysate notably differed from the two previous samples in regard to the unique presence of arabinose residues along fucose residues. Also, it is the only species showing significant amount of GalNAc, but no Py(4,6)Gal at all. TLC analysis revealed that this material was constituted by a mixture of small oligosaccharides which sizes were comparable to those hydrolysed from GN1. Thus, as already described for

GN1, this material was analysed in ES/MS-MS, first as native oligosaccharides, then after reduction by BD₄Na.

Ten oligosaccharides (compounds X to AG, see table 7), out of which nine were sequenced, were observed in negative mode, whereas no oligosaccharides were observed in positive mode. As observed in GN1, fragmentation pattern of each compound in negative mode was characterised by the presence of a very intense ion at m/z 97, indicating that they were all sulphated. Sulphate groups were easily positioned either on a deHex residue owing to the intense ion at m/z 225, or on a pent residue owing to the intense ion at m/z 211. Considering total composition analysis, compounds X, Z, AB, AD, AF and AG contained sulphated Ara residues while compounds Y, AA, AC and AE contained sulphated Fuc residues.

Compounds X and Y were the smallest oligosaccharides sequenced from GN3. From their calculated compositions and fragmentation patterns, it was deduced that X was composed of one HexNac and one sulphated Pent residues and that Y was composed of one HexNAc and one sulphated deHex residues. Surprisingly, on the contrary of all other sequenced compounds, reduction by BD₄Na did not affect these two oligosaccharides that presented identical fragmentation patterns before and after reduction. This prevented us to directly attribute the monosaccharide in terminal reducing position. However, we hypothesised that sulphation of the monosaccharide in terminal reductive position would prevent reduction of the oligosaccharide by BD₄Na, presumably through electrostatic interactions, and proposed that (SO₃)deHex and (SO₃)pent residues were in reducing positions. Indeed, if HexNAc residue was in reducing position it would have been reduced, as confirmed latter by the sequencing of of compounds AB and AC that presented HexNAc residues in reducing position. The two motives proposed for X and Y were subsequently found in the seven other sequenced compounds.

Then, compounds Z, AA, AB and AC appeared as extensions of X and Y with an additional deHex residue in reducing position for Z and AA, or an additional HexNAc residue in reducing position for AB and AC. Fragmentation spectra of Z and AA were unclear, but presence of HexNAc in terminal non-reducing position was demonstrated by the presence of [M-HexNAc+Na] Z ions, while direct attribution of deHex in reducing position was only done for compound AA owing to an small [M-deHexol+Na] B ion at m/z 428. Sequences of compounds AB and AC were easily established owing to intense [M-HexNAc+Na] Y ions at m/z 435 and 449 and [M-HexNAc-HexNAcol+Na] ions at 211 and 225, respectively.

In a similar way, compounds AD and AE were described as extensions of AB and AC with an additional deHex residue in terminal position. For AD, deHex and HexNAc residues were easily positioned owing to intense [M-deHexol+Na] B ion and [M-HexNAc] Y ion at m/z 617 and 581 respectively. Fragmentation pattern did not enable us to directly position the remaining HexNAc residue in position 2 or 3 from the reducing monosaccharide. However, in the second case (sequence HexNAc-HexNAc-(SO₃)Pent-deHex), we would have observed the [M-HexNAc-HexNAc+Na] Y ion, indicative of the presence of two consecutive HexNAc residues, which was not the case for both native and reduced oligosaccharide. On the contrary, we only observed the ions [M-deHexol-HexNAc+Na] ions at m/z 414 and 432, which was in favour with the sequence HexNAc-(SO₃)Pent-HexNAc-deHex. Identical fragmentation pattern was observed for compound AE, that only differed by the presence of a sulphated deHex residue instead of a sulphated pent residue.

Then, compound AF appeared as an extension of compound AD with an additional Hex residue. The [M-Hex+Na] and [M-Hex-HexNAc+Na] Y ions at m/z 784 and 581 positioned the disaccharide Hex-HexNAc dimer at the terminal non-reducing end of the oligosaccharide. Similarly, [M-deHexol+NA] and [M-deHexol-HexNAc+Na] at m/z 779 and 576, confirmed the presence of the motive HexNAc-deHex in reducing position, establishing the sequence as Hex-HexNAc-(SO3)pent-HexNAc-deHex. A minor peak at m/z 960 was observed in ES/MS, but could not be sequenced due to too small intensity. Considering the eight above sequenced compounds, its calculated composition (1 Hex, 1 HexNAc, 2 deHex 1 SO₃ group) strongly suggested that this oligosaccharide was the equivalent of compound AF with a sulphated deHex instead of a sulphated Pent residue. Indeed, it appeared that all oligosaccharides were systematically present as pairs of similar sequences: one containing a sulphated arabinose and one a sulphated fucose. Considering all the sequences only two different motives were distinguished that way: the pentasaccharide Hex-HexNAc-(SO3)[Ara/Fuc]-HexNAc-Fuc and the minor variant HexNAc-(SO₃)[Ara/Fuc]-Fuc. No extension of this last motif was identified. These data suggested that acid labile domain of GN2 glyconectin presented a relatively simple polymeric structure with homogeneous repetitive motives.

REFERENCES

- 1. Misevic, G. N., Finne, J. & Burger, M. M. (1987) J. Biol. Chem. 262, 5870-5877.
- 2. Misevic, G. N. & Burger, M. M. (1993) J. Biol. Chem. 268, 4922-4929.
- Kamerling, J.P., Gerwig, G.J., Vliegenthart, J.F. & Clamp, J.R. (1975) Biochem. J. 151, 491-495
- 4. Zanetta, J.P., Timmerman, P. & Leroy, Y. (1999) Glycobiology 9, 255-266
- 5. Ciucanu, I. & Kerek, F. (1984) Carbohydr. Res. 131, 209-217
- 6. Dudman, W.F. and Lacey, M.L. (1986) Carbohydr. Res. 145, 175-191
- Spillmann, D., Hard, K., Thomas-Oates.J., Vliegenthart, J.F.G., Misevic, G., Burger, M.M. and Finne, J. (1993) J. Biol. Chem. 268, 13378-13387.
- Spillmann, D., Thomas-Oates.J., van Kuik, J.A., Vliegenthart, J.F.G., Misevic, G., Burger, M.M. and Finne, J. (1995) J. Biol. Chem. 270, 5089-5097
- 9. Yang, B.Y, Brand, J. and Montgomery, R. (2001) Carbohydr. Res., 331:59-67
- 10. Bock, K. and Pedersen, C. (1983). In Adv. Carbohydr. Chem. Biochem., 41, 27-66
- 11. Cérantola, S., Marty, N. and Montrozier, H. (1996) Carbohydr. Res. 285: 59-67
- Van Calsteren, M.R., Pau-Roblot, C. Bégin, A. and Roy, D. (2002) *Biochem. J.* 363, 7-17
- 13. Fontaine, T., Wieruszeski, J.M., Talmont, F., Saniez, M.H., Duflot, P., Leleu, J.B. and Fournet, B. (1991) Eur. J. Biochem. 196, 107-113

Tables

Monosaccharide	GN1 mol%	GN2 mol%	GN3 mol%
Ara	0	0	11
Fue	28	7	31
Man	9	16	20
Gal	29	43	12
GalNAc	1	1	7
GlcNAc	17	14	12
Py(4,6)Gal	4	16	0
GlcA	11	3	6

Table 1: Carbohydrate composition (mol%) of glyconectins established by standard gas chromatography

Table 2. Fragmentation of Glyconectin Polysaccharides. Glyconectins were digested by mild acid treatment (0,1 N TFA, 80°C, 1 h.). Carbohydrate analysis was established by standard gas chromatography on liberated oligosaccharides (supernatant) and on acid-resistant core (pellet).

	GN1 m	ol%	GN2 m	ol %	GN3	%
	Supernatant	Pellet	Supernatant	Pellet	Supernatant	Pellet
Ara	0	0	0	0	10	5
Fuc	36	11	4	5	22	14
Man	4	12	18	33	19	18
Gal	27	37	28	33	8	15
GalNAc	0	0	1	0	15	6
GlcNAc	28	29	19	27	26	38
Py(4,6)Gal	2	4	29	2	0	0
GlcA	3	7	0	0	0	5

Table 3: Compositions and sequences deduced from the ES/MS sequencing in negative mode of TFA 0.1N hydrolysed native and BH_4Na reduced fragments isolated from GN1. Sequence information comes usually from ES/MS-MS fragmentation pattern of both native and reduced oligosaccharides. Origin of reported fragment ions (native or reduced) is indicated by *. Cmpd stands for compound.

Cmpd	m/z native	m/z reduced	Calculated composition	Proposed sequence	ES/MS-MS fragments [M-H]	Infered structure
	300*	-	HexNAc SO3	(SO ₃)HexNAc	97	(SO ₃)GlcNAc
B	446	448*	HexNAc deHex SO3	(SO₃)HexNAc-deHex	97; B, 282	(SO₃)GlcNAc-Fuc
C	592	594*	HexNAc 2xdeHex SO ₃	deHex-(SO₃)HexNAc-deHex	97; B, 282, 448	Fuc-(SO3)GlcNAc-Fuc
D	608	610*	HexNAc Hex SO3	HexNAc-(SO₃)Hex-deHex	97; B, 444; Y, 407; secondary fragment, 241	GlcNAc-(SO3)Gal-Fuc
E	624	626*	HexNAc 2xHex SO3	Hex-(SO3)Hex-HexNAc	97; B, 403; C, 421; Y, 464; secondary fragment, 241	Gal-(SO3)Gal-GlcNAc
	770	772*	deHex HexNAc 2xHex SO ₃	Hex-(SO3)Hex-HexNAc-deHex	97; B, 403, 606; Y, 444; secondary fragment, 241	Gal-(SO3)Gal-GlcNAc-Fuc
G	793*	795	2xdeHex 2xHexNAc	HexNAc-deHex-(SO ₃)HexNAc-deHex	97; B, 631; C, 649; Y, 592; secondary fragment, 282, 428, 446	GlcNAc-Fuc-(SO ₃)Gal-Fuc

Cmpd	m/z native	Calculated composition	Proposed sequence	ES/MS-MS fragments [M+Na] ⁺
	435	PyrHex Hex	НехРуг-Нех	[M-44], 391; [C-44], 229; Y, 203
Ĩ	476	PyrHex HexNAc	HexPyr-HexNAc	[M-44], 432; [C-44], 229; Z, 226; Y, 244
J	638	PyrHex Hex HexNAc	PyrHex-Hex-HexNAc	[M-44], 594; B, 417, [B-44], 373; C , 435; [C-44] 391; Z, 226, 388; Y, 244, 406
K	800	PyrHex 2xHex HexNAc	PyrHex-Hex-Hex-HexNAc	[M-44], 756; [B-44], 373, 535; C, 597; [C-44], 391, 553; Z, 226, 388, 550; Y, 244, 406, 568
L	800	PyrHex 2xHex HexNAc	PyrHex-Hex-(Hex)HexNAc	[M-44], 756; [B-44], 373, 576; C, 203; [C-44], 391, 594; Z, 388, 550, 620; Y, 406, 568, 638
M	962	PyrHex 3xHex HexNAc	PyrHex-Hex-Hex-Hex-HexNAc	[M-44], 918; [B-44], 535, 697; C, 759; [C-44], 391, 553, 715; Z, 388, 550, 712; Y, 244, 406, 568, 730; secondary fragments, 347, 365, 509, 527

.

Table 4: Compositions and sequences deduced from the ES/MS sequencing in positive mode of TFA 1N hydrolysed native fragments isolated from GN2.

Cmpd	m/z reduced	Calculated composition	Proposed sequence	ES/MS-MS fragments [M+Na] ⁺
Ň	520	PyrHex deHex	PyrHex-deHex	Y, 246
H	550	PyrHex Hex	PyrHex-Hex	Y, 276
i i i i i i i i i i i i i i i i i i i	591	PyrHex HexNAc	PyrHex-HexNAc	[M-59], 532; B, 299; C, 315; Y, 317
0	724	PyrHex Hex deHex	PyrHex-Hex-deHex	B, 501; C, 519; Y, 246, 449
	754	PyrHex 2xHex	PyrHex-Hex-Hex	[M-59], 695; B, 501; C, 519; Y, 276, 480
Q	754	PyrHex 2xHex	PyrHex-(Hex)Hex	[M-59], 695; Y, 480, 536
J	795	PyrHex Hex HexNAc	PyrHex-Hex-HexNAc	B, 501; C, 519; Z, 299, 503; Y,317, 521
R	958	PyrHex 3xHex	PyrHex-Hex-(Hex)Hex	Y, 480, 684, 740
K	999	PyrHex 2xHex HexNAc	PyrHex-Hex-Hex-HexNAc	B, 501, 705; C, 519, 723
	999	PyrHex 2xHex HexNAc	PyrHex-Hex-(Hex)HexNAc	B, 501; C, 519; Y 781
S	1162	PyrHex 4xHex	PyrHex-Hex-Hex-(Hex)Hex	Y, 480, 684, 888
T	1162	PyrHex 4xHex	Hex-Hex-(PyrHex)Hex-Hex	Y, 944, 740; secondary fragment, 465
J. U. S.	1218	2xPyrHex 3xHex	PyrHex-Hex-(PyrHex)Hex-Hex	Y, 944, 740; secondary fragment, 465
	1366	PyrHex 5xHex	Hex-Hex-Hex-Hex-(PyrHex)Hex	Y, 1148, 944, 740, 536
W	1366	PyrHex	HexPyr-Hex-Hex-Hex-Hex-Hex	Y, 1092, 888, 684, 276

Table 5: Compositions and sequences deduced from the ES/MS sequencing in positive mode of TFA 1N hydrolysed BD_4Na reduced, permethylated fragments isolated from GN2.

Cmpd	[M+H] ⁺	Calculated composition	Proposed sequence
<u>ÌÌ</u> ,	528	PyrHex Hex	Py(4,6)Hex1-4Hex
I	569	PyrHex HexNAc	Py(4,6)Hex1-3HexNAc
	773	PyrHex Hex HexNAc	Py(4,6)Hex1-3Hex1-3HexNAc

Table 6: Compositions, sequences and glycosidic linkages deduced from GC/MS analysis of TFA 1M hydrolysed oligosaccharides from GN2. [M+H]+ was established in CI mode and sequence and linkage in EI mode.

Table 7a: Compositions and sequences deduced from the ES/MS sequencing in negatve mode of TFA 0.1N hydrolysed native and BD₄Na reduced oligosaccharides isolated from GN3.

Cmpd	m/z native	m/z reduced	Calculated composition	Proposed sequence	ES/MS-MS fragments [M-H]
X	432	432	HexNAc	HexNAc-(SO ₃)Pent	97; Z, 211; Y, 229
			Pent		
			SO ₃		
Y	446	446	HexNAc	HexNAc-(SO3)deHex	97; Z, 225; Y, 243
			deHex		
			SO ₃		
Z	578		HexNAc	HexNAc-(SO3)Pent-deHex	97; Z, 357; secondary fragment, 211
		-	deHex		
			Pent		
			SO ₃		
ÅÅ	592	595	HexNAc	HexNAc-(SO3)deHex-deHex	97; B, 428; Z, 375; secondary fragment, 225
			2xFuc		
			SO ₃		
AB	635	638	2xHexNAc	HexNAc-(SO3)Pent-HexNAc	97; Y, 435; secondary fragment, 211
			Pent		
			SO ₃		
AC	649	652	2xHexNAc	HexNAc-(SO₃)deHex-HexNAc	97; Y, 449; secondary fragment, 225
			deHex		
			SO ₃		

Table 7b: Compositions and sequences deduced from the ES/MS sequencing in negatve mode of TFA 0.1N hydrolysed native and BD₄Na reduced oligosaccharides isolated from GN3.

Cmpd	m/z native	m/z reduced	Calculated composition	Proposed sequence	ES/MS-MS fragments [M-H]
AD	781	784	2xHexNAc Pent deHex SO ₃	HexNAc-(SO3)Pent-HexNAc-deHex	97; B, 414, 617; C, 432; Y, 581; secondary fragment, 211
AE	795	798	2xdeHex 2xHexNAc SO3	HexNAc-(SO3)deHex-HexNAc-deHex	97; B, 428, 631; C, 446; Y, 595; secondary fragment, 225
AF	943	946	2xHexNAc Hex deHex SO3	Hex-HexNAc-(SO3)Pent-HexNAc-deHex	97; B, 779, 576; Y, 784 secondary fragment, 211, 414, 581
AG	984	987	3xHexNAc Pent deHex	[(SO3)Pent+2xHexNAc+deHex]-HexNAc	97; B, 763; secondary fragment, 211, 617, 560

.

.



Fig. 1: EI-MS identification of 4,6-O-(1-carboxyethylidene)-D-galactose as a) TMS derivative and b) permethylated derivative.



Fig. 2: TLC analysis of total hydrolysis fractions from GN1 and GN2 stained by sulfuric orcinol 1) standard glc DP, 2) TFA 0.1N hydrolysis of GN1, 3) TFA 0.1 and 4) 1M hydrolysis of GN2. Due to too small quantities of sample available, such analysis could not be done on GN3.



Fig. 3: ES/MS profiling of ethanol soluble oligosaccharides from purified glyconectins.a) reduced oligosaccharides from TFA 0.1M hydrolysis of GN1, b) reduced, permethylated oligosaccharides from TFA 1M hydrolysis of GN2, c) reduced oligosaccharides from TFA 0.1 M hydrolysis of GN3.



Fig. 4: ES/MS-MS fragmentation patterns of a) native and b) reduced oligosaccharide C from GN1 hydrolysis at m/z 591.8 and 594, respectively.


Fig. 5: Analysis of reducing monosaccharides from TFA 0.1 N hydrolysis of GN2. Sample was reduced by $B^{3}H_{4}K$ and hydrolysed by TFA 4N, 100 °C, 4 h. Rsulting monosaccharides were analysed by a) HPAEC and b) TLC. Each eluted from HPAEC was collected and tadioactivity counted. TLC was autoradiographed.

ź



Fig. 6: a) TOCSY spectrum and b) HMQC spectrum of TFA 0.1 M hydrolysis of GN2. Only the spin system of terminal β -D-Py(4,6)Galp was assigned.

289



Fig. 7: Fragmentation pattern in ES/MS-MS oligosaccharide J from GN2 TFA 1M hydrolysis. a) native oligosaccharide and b) reduced and permethylated oligosaccharide.





3. Conclusions

Ce travail représente la première étude comparative de la structure des domaines glycanniques des glyconectines isolées des espèces M. prolifera (GN1) Halichondria panicea (GN2) et Cliona chelata (GN3). Dans cette première étape, seules les parties sensibles à l'hydrolyse ménagée ont été étudiées. Celles ci représentent entre 40 et 70% des domaines glycanniques totaux des glyconectines, en fonction de leur espèce d'origine. Cette étude a clairement établi que la structure glycannique de ces molécules d'adhésion était très spécifique à chacune de ces trois espèces. Elles différaient par (i) leur composition en monosaccharides, en particulier par la présence de Fuc, d'Ara et de py(4,6)Gal; (ii) leur susceptibilité à l'hydrolyse acide, reflétée par les quantités de matériel libéré lors de l'hydrolyse; (iii) la taille des oligosaccharides libérés lors de l'hydrolyse ménagée ATFA 0.1M (1-4 monosaccharides pour GN1, >10 pour GN2 et 1-5 pour GN3); (iv) la nature de la charge négative (essentiellement SO₃⁻- pour GN1 et GN3, ⁻OOC- pour GN2) et leur position (Gal et GlcNAc pour GN1, Ara et Fuc pour GN3); (v) la séquence des oligosaccharides libérés. Sur la base de la composition et du séquençage des oligosaccharides libérés par hydrolyse ménagée (ATFA 0.1 M pour GN1 et GN3 ou ATFA 1M pour GN2), et en ne prenant en compte que les oligosaccharides identifiés les plus longs, nous pouvons résumer les motifs structuraux présents sur les domaines glycanniques de chaque glyconectines de la façon suivante:

GN1	GN2	GN3
GlcNAc-Fuc-[SO3]GlcNAc-Fuc	py(4,6)Gal-(Hex) ₀₋₁ -Fuc	HexNAc-[SO ₃]Ara/Fuc-Fuc
GlcNAc-[SO3]Gal-Fuc	py(4,6)Gal-(Hex) ₀₋₃ -GlcNAc	HexNAc-[SO3]Ara/Fuc-HexNAc-Fuc
Gal-[SO3]Gal-GlcNAc-Fuc	py(4,6)Gal-(Hex)1-5	
py(4,6)Gal-GlcNAc-Fuc	py(4,6)Gal-(Hex) ₀₋₂ [Hex]Hex	
	py(4,6)Gal-[Hex]Hex-Hex	
	py(4,6)Gal-Hex-[py(4,6)Gal]Hex	
	(Hex) ₄ -[py(4,6)Gal]Hex	

Tableau C2: Résumé des motifs glycanniques présents sur GN1, GN2 et GN3. Hex: Gal ou Man, HexNAc: GlcNAc ou GalNAc

Ces résultats démontrent que les domaines polysaccharidiques des glyconectines ne sont pas constitués d'un enchaînement identique de courtes unités de répétition, et sont donc beaucoup plus hétérogènes que les glycosaminoglycannes. Cette étude a également établi que ces polysaccharides sont des molécules mixtes, composées d'un domaine sensible à l'hydrolyse, et d'un domaine résistant à l'hydrolyse. Dans les trois espèces, chacun des deux domaines présente des compositions très différentes, ce qui démontre que la fraction résistante à l'hydrolyse ne résulte pas d'une hydrolyse incomplète du polysaccharide mais qu'elle présente une structure glycannique distincte de la fraction sensible à l'hydrolyse. L'agencement des domaines sensibles et résistants à l'hydrolyse au sein de la molécule est encore inconnu. Pour GN1 et GN3 deux modèles sont envisageables:

(i) un polysaccharide linéaire composé soit d'un enchaînement non répétitif de courts oligosaccharides, soit d'unités de répétition de grande taille comportant des liaisons sensibles à l'hydrolyse (Fuc- et/ou Ara-). La stratégie que nous avons adoptée ne nous permet pas de différentier ces deux possibilités car les liaisons Fuc- et Ara- sont presque intégralement coupées. Cette chaîne polysaccharique sensible à l'hydrolyse serait distincte du domaine résistant ou liée sur ce dernier (Fig. C3a).

(ii) un polysaccharide mixte ramifié, composé d'un axe glycannique résistant à l'hydrolyse sur lequel sont greffés *via* un Fuc ou un Ara des oligosaccharides libérés par l'hydrolyse (Fig. C3b)

a)



Fig. C3: Modèles d'organisation possibles des chaînes polysaccharidiques des glyconectines GN1 et GN3 déduits de l'analyse des oligosaccharides libérés par hydrolyse ménagée.

Aucune donnée expérimentale ne nous permet, pour l'instant, de trancher entre ces deux modèles. Il faut noter que les différents domaines supposés peuvent également être indépendants et faire partie de glycannes distincts.

Au contraire de GN1 et GN3, l'hydrolyse de GN2 a libéré des oligosaccharides de grande taille (supérieurs à 10 monosaccharides). Leur analyse a révélé qu'ils étaient constitués de chaînes glycanniques très ramifiées, terminées en positions non réductrices par des unités py(4,6)Gal et Gal. L'hydrolyse plus poussée de ces polysaccharides a permis d'obtenir un grand nombre d'oligosaccharides plus courts qui ont été séquencés. Ils consistent en des oligomères d'hexoses (Gal et Man), linéaires ou branchés, terminés par des unités de py(4,6)Gal et interrompus par des unités GlcNAc et Fuc. L'hétérogénéité de GN2 semble donc encore plus importante que celle de GN1 et GN3, mais le fait que les polysaccharides de GN2 étaient moins sensibles à l'hydrolyse nous a donné accès à plus d'informations quant à leur arrangement. Ces donnés permettent d'envisager un modèle plausible de la structure de la partie hydrolysable de GN2 (Fig. C4):



Fig. C4: Modèle de structure des chaînes glycanniques sensibles à l'hydrolyse de GN2. Les segments encadrés correspondent à des exemples d'oligosaccharides libérés par hydrolyse ATFA 1M.

De nombreux travaux sont encore nécessaires pour déterminer de manière plus rigoureuse la structure des chaînes polysaccharidiques associées aux glyconectines d'éponges. Nous envisageons au laboratoire d'étudier la structure des domaines glycanniques qui ont résisté aux hydrolyses ménagées que nous avons effectuées sur les molécules natives. La fragmentation des chaînes polysaccharidiques par une autre méthode que l'hydrolyse acide nous permettra d'obtenir des informations complémentaires sur leur arrangement par analyse

de fragments glycanniques chevauchants. En particulier une hydrazinolyse des polysaccharides, suivie par une désamination nitreuse coupera spécifiquement les liaisons HexNAc-X en transformant les résidus de GlcNAc en 2,5-anhydromannose et les résidus de GalNAc en 2,5-anhydrotalose. Cette méthode sera appliquée aux glyconectines natives, ainsi qu'aux domaines résistants à l'hydrolyse. Au vu de la proportion importante de N-acétylhexosamine dans les glyconectines, la désamination nitreuse génèrera un mélange de courts oligosaccharides qui pourront être analysés d'une manière similaire aux fragments provenant de l'hydrolyse ménagée.

Sur la base des travaux de structure, des études complémentaires sur les propriétés d'adhésion des glyconectines sont actuellement en cours, en collaboration avec le Dr G. Misevic. Ces travaux consistent dans un premier temps à multimériser les oligosaccharides issus de l'hydrolyse des trois glyconectines et de tester leurs propriétés d'adhésion par microscopie de force atomique. S'il apparaît que les oligosaccharides en mélange interagissent entre eux, nous purifierons chaque oligosaccharide par chromatographie liquide haute performance et testerons leurs capacités individuelles d'interaction homophiliques. Ces études fourniront les premières informations systématiques détaillées sur les mécanismes moléculaires de régulation des interactions cellulaires chez les éponges, en fonction de la structure des glycannes de surface.

5. Annexe

Nous faisons figurer ici un article soumis, dans lequel ont été intégrés les résultats préliminaires de l'étude structurale que nous venons de présenter.

Classification:

Biological Sciences: Biochemistry

Title:

Molecular Recognition Between Glyconectins as a Adhesion Self-Assembly Pathway to Multicellularity

Gradimir N. Misevic*, Yann Guerardel, Xavier Czeszak, Lazar Sumanovski, Maurice Demarty, Camille Rippol, Yannis Karamanos†, Octavian Popescu‡, Gerard Strecker Authors Affiliations:

‡Department of Genetics and Ecology, University Babes-Bolyai and Institute of Biological Research, 3400 Cluj-Napoca, Romania.

Laboratoire de Chimie Biologique, Universite de Sciences et Technologies de Lille, UMR 8576 CNRS, 59655 Villeneuve D'Ascq, France

Corresponding author:

*To whom reprint requests should be addressed at: Gradimir N. Misevic. <u>gradimir@gradimir.com</u> Tel + 332 35 14 69 08; fax + 332 35 14 70 20

Manuscript information:

Number of text pages is . Number of figures is . Number of tables is .

Word and character count:

Number of words in abstract is 205. Total number of characters in the paper is .

Abbreviation footnote:

§ Sea Water Tris (SWT) is 0.5 M NaCl, 2 mM CaCl₂, 20 mM tris-HCl pH 7.4.

GC/MS Gas Chromatography Coupled Mass Spectrometry

EI-MS Electronic Impact Mass Spectrometry

CI-MS Chemical Ionization Mass Spectrometry

TMS Trimethylsilyl

TFA Trifluoro-Acetic Acid

HOHAHA Homonuclear Hartmann-Hahn Spectroscopy

ABSTRACT

The appearance of multicellular forms of life has been tightly coupled to the ability of the organism to retain its own anatomical integrity and to distinguish self from non-self. Large glycoconjugates as the outermost cell surface layer of all Metazoans (1-3), are the primary candidates for the primordial adhesion and recognition functions in biological self-assembly systems. Atomic force microscopy experiments demonstrated that the binding strength between a single molecular pair of Porifera cell surface glyconectin glycoconjugates can hold the weight of 1600 cells (3), proving their adhesion functions. Here, measurement of molecular self-recognition of glyconectins purified from three different Porifera species was used as an experimental model for primordial xenogeneic self-non-self discrimination. Physical characterization and chemical fingerprinting structural analyses of the three glyconectins showed that they define a new family of proteoglycan-like molecules exhibiting species-specific structures with complex and repetitive acidic carbohydrate motives different to the classical proteoglycans and mucins. In functional self-assembly color coded bead, cell and blotting assays glyconectins displayed species-specific recognition and adhesion. The specificity of homophilic glyconectin interactions in Porifera approaches binding selectivity of evolutionarily advanced immunoglobulin superfamily system. Xeno-selectivity of primordial glyconectin to glyconectin recognition may be a new paradigm in the self-assembly and non-self discrimination pathway of cellular adhesion leading to multicellularity.

INTRODUCTION

The emergence of multicellularity required the simultaneous development of cell adhesion and recognition properties. In addressing the unresolved question of what may have been the molecular basis for primordial self-recognition and non-self discrimination we focused our attention on the role of proteoglycan-like glyconectins in Porifera xenogeneic cellular interactions, as evolutionary most compatible model system for ancestors of Metazoans. Electron microscopic, X-ray diffraction and biochemical analyses showed that large glycoconjugates such as glyconectins, classical type proteoglycans and mucins, are the largest multi-million molecular weight macromolecules extending above the cell surface at least ten times higher then any other cell adhesion glycoprotein (1-3). Based on this evidence, here we forward the hypothesis that Porifera glyconectins as the most peripheral cell surface environment sensors may have provided the initial key recognition and adhesion functions during the emergence of Metazoan ancestors. This would imply that Porifera as the simplest Metazoans alive today should have preserved, at least in part, glyconectin adhesion and recognition mechanism guiding the initial phase of xenogeneic selectivity of cellular interactions. In fact, in recent atomic force microscopy experiments we have demonstrated with purified molecules that glyconectin 1 to glyconectin 1 binding strength supplies fundamental cell cohesion forces in the sponge Microciona prolifera, thus assuring the adhesion role ascribed to them following previous functional investigations (3-8).

To examine whether homophilic glyconectin to glyconectin binding has been also involved in primordial cellular adhesion and recognition self-assembly processes associated with the emergence of multicellularity, we have studied the molecular basis of specific xenogeneic cell interactions in phylum Porifera. Early work on dissociated marine sponge cells provided important phenomenological evidence for cell sorting (4-7). However, these and following experiments used semi purified and chemically undefined extracts named aggregation factors, and thus lacked quantitative and biochemical evidence about the underlying molecular mechanisms (4-14). Here we have isolated glyconectins from three marine sponge species, *Microciona prolifera* (GN1), *Halichondria panacea* (GN2) and *Cliona celata* (GN3). Physical, chemical and structural fingerprinting characterization of the three purified glyconectin macromolecules revealed species-specific proteoglycan-like structures defining a new class of glycocojugates. Using quantitative functional approach with color coded beads, cells and blotting, we have shown that highly specific cell surface glyconectin to glyconectin interactions indeed mediate cell recognition and adhesion in the three Porifera species. This structure to function related study establishes a new paradigm about the pivotal role of primordial glyconectins in the self-assembly and non-self discrimination pathway of cellular adhesion foremost to evolution of multicellularity.

MATERIALS AND METHODS

Isolation of Glyconectin Proteoglycans. Proteoglycans were extracted from fresh cuts of sponges with artificial Ca²⁺ and Mg2+free seawater (462 mM NaCl, 10.7 mM KCl, 7 mM Na2SO4 and 2.1 mM NaHCO3) at $+4^{\circ}$ C for 12h. Proteoglycans were purified by sequential centrifugation, 20 mM CaCl₂ precipitation and ultracentrifugation (13,14).

Ultracentrifugation Analysis. Molecular weight and sedimentation coefficient of sponge proteoglycans were determined by sedimentation equilibrium and sedimentation velocity analyses in seawater Tris (SWT) (0.5 M NaCl, 2 mM CaCl₂ 20 mM Tris pH 7.40 at +20°C using Beckman Model E analytical ultracentrifuge according to procedures described previously (15).

Agarose Electrophoresis. Electrophoretic separation of sponge proteoglycan was performed on a 0.75% agarose gel in 50 mM tris-acetate pH 7.3 for 2 hours at 100 V and 4°C (16). Gels were stained with 0.02% Toluidine blue in 3% acetic acid, followed by 0.1% Amido black 10B in 3% acetic acid.

Chemical degradation of glyconectins. Materials were repetitively hydrolysed in mild conditions for short times (TFA 0.1 M, 80°C, 1 h.). After each hydrolysis step, four volumes of ethanol were added and liberated oligosaccharides were collected in the ethanol soluble fraction. Release of oligosaccharides from glyconectins was assessed after each step by TLC analysis of ethanol soluble fraction on silica gel using /acetic acid/H₂O (40/30/20) as solvent. Typically, four hydrolysis steps were sufficient to liberate all hydrolysable material from glyconectins. Prepared oligosaccharides were further purified by gel filtration on a Bio-Gel P2 column (Bio-rad) before analysis.

Matrix-assisted laser-desorption ionization-time of-flight (MALDI-TOF) MS. Molecular mass of the oligosaccharides was measured by MALDI-MS on a Vision 2000 timeof-flight instrument (Finnigan Mat) equipped with a 337 nm UV laser: 1 µl of sample at a



concentration of 100 pmol/ μ l was mixed with an equal volume of matrix and allowed to crystallize. We used 2,5-dihydroxybenzoic acid and 3-aminoquinoline for neutral and acidic oligosaccharides, respectively.

NMR spectroscopy. The NMR experiments were performed on Bruker® ASX400 and DMX600 spectrometers both equipped with a 5mm ${}^{1}\text{H}/{}^{13}\text{C}$ mixed probe-head, operating in the pulse Fourier transform mode controlled by an Aspect 3000 computer. Each oligosaccharide sample was dissolved in 400 μ L ${}^{2}\text{H}_{2}\text{O}$ after three exchanges with ${}^{2}\text{H}_{2}\text{O}$ (99.97 % atom ${}^{2}\text{H}$, Euriso-top, CEA group, Gif-sur-Yvette France) and intermediate lyophilisations. The oligosaccharides were analyzed at 300 °K. The chemical shifts (δ) were referenced to internal acetone ($\delta^{1}\text{H}$ =2.225 and δ^{13} C=31.55 p.p.m. in the condition used). Two dimensional homonuclear (COSY90, TOCSY) and heteronuclear (HMQC) experiments were performed by using standard Bruker® pulse programs. The main pulses and variable delays were optimized for each pulse program and sample.

Gas-chromatography. Monosaccharides were analyzed by GC/MS as pertrimethylsilyl (17) and per-heptafluorobutyryl derivatives (18). Methylation of oligosaccharides was conducted according to Ciucanu and Kerek (19) before analysis on GC/MS.

Other analytical methods. Size of the oligosaccharides was assessed by gel filtration on a Bio-Gel P2 or P4 column (Bio-Rad) equilibrated in acetic acid 0,5 % and calibrated with hydrolyzed dextran, in combination with thin layer chromatography as already described. Amino acid composition was assessed by Pico-Tag HPLC and SO₄- content by Dionex ionexchange HPAEC both after 6M HCl hydrolysis at 100 °C for 12h.

Preparation of Cells and Glyconectin Coated Beads. Sponge cells were dissociated and washed in Ca^{2+} and Mg^{2+} free seawater at 0°C; in some experiments cells were fixed with 1% glutaraldehyde (13,14).

A volume of 200 ml 2% latex-amidine bead suspension, 0.6 mm in diameter (Molecular Probes, Inc., Eugene, OR, USA) was washed three times with 1 ml SWT by centrifugation at 3000g for 10 min. Beads were resuspended in 200 ml SWT and bath sonicated for 5 min. 75 μ g glyconectin/75 μ l bead suspension was incubated at 22°C for 15

min, followed by 3 washes as above. The glyconectin coated beads were resuspended in 75 ml SWT.

Recognition, Adhesion and Overlay Assays with Cells and Beads. Cell-cell and cell-bead assays were performed with 10^6 cells bearing surface proteoglycans in 100 ml SWT without or with 10 mM Ca²⁺. 5 to 10 ml of 2% bead suspension in SWT was bath sonicated for 1 min before use in a cell-bead, bead-bead and bead overlay aggregation assay (100 ml final volume). Since the diameter of beads is 600 nm and the size of glyconectins are in the same range, at least 1 molecule of glyconectin was bound to 1 bead. Bead-bead and bead overlay

RESULTS

Purification and Physical Characterization of Glyconectins of Three Sponge Species. Glyconectins (GNs) were purified by CaCl₂ precipitation and sequential ultracentrifugation of isotonic Ca²⁺ free sea water extracts from three different marine sponge species, *Microciona prolifera* (GN1), *Halichondria panacea* (GN2) and *Cliona celata* (GN3). This procedure is similar to that one we have originally used for *Microciona prolifera* and is thus different to guanidinium hydrochloride or urea extractions techniques applied for classical mammalian proteoglycans (20). Analytical ultracentrifugation and agarose electrophoresis of isolated glyconectins showed that each species expressed a unique type of multi-million macromolecule with sedimentation larger then 20 s composed of one or few highly negatively charged glycosylated subunits (Table 1 and Fig. 1).

Carbohydrate Compositional Analysis of Native Glyconectins.

The three purified native glyconectins were analysed for their sugar compositions in gas chromatography as shown in Table 1. A single type of pyruvated hexose was also detected as methyl-glycoside heptafluoryl derivative owing to a $[M+NH_4]$ ion at m/z 698 in CI-MS and a very intense [M-COOMe] ion at m/z 611 in EI-MS. Nature of this residue was established as 4,6-pyruvylated galactose (Py(4,6)Gal) by TMS derivatization of methyl glycosides and comparison with standard EI-MS spectra from literature (21).

As shown in Table 1, glyconectins from the three analyzed species showed substantial differences regarding their sugar compositions. GN1 species was characterized by a high quantity of fucose (about one third of the total monosaccharide content) and minute amounts

of Py(4,6)Gal, while GN2 was characterized by a lower proportion of fucose and a molar ratio of Py(4,6)Gal three times higher than GN1. On the other hand, GN3 was the only one to contain arabinose (10% mol/mol) as well as high quantities of fucose but no Py(4,6)Gal at all. This last species was also the only one to exhibit significant amounts (> 5%) of GalNAc.

Fragmentation of Glyconectin Polysaccharides

In order to compare the carbohydrate moieties of the three glyconectins, they were fragmented by partial acid hydrolysis. Each glyconectin exhibited different susceptibilities to acid treatment, as reflected by the distinct proportion of released oligosaccharides soluble in the ethanol. These fractions represented 68 % of the total carbohydrate content for GN1, 56 % for GN2 and 41 % for GN3. Further analysis revealed that hydrolyzed materials had substantially distinct sugar compositions compared to unhydrolyzed materials (Table 2). In particular, cleaved oligosaccharides were systematically enriched in fucose and/or arabinose, compared to un-hydrolyzed fractions. This reflects the preferential cleavage by partial hydrolysis of domains of the molecule rich in acid-labile linkages such as fucose and arabinose linkages. On the contrary, glucuronic acid was predominantly retrieved in hydrolysis resistant fractions. These data suggested that the carbohydrate moieties of glyconectins are not constituted by simple homogeneous polymers, but are rather more complex repeats of acidic glycan moieties containing fucose.

Mass-Spectrometry Fingerprinting of Glyconectin Oligosaccharides.

Since the most peripheral parts of glyconectins must be the first encountering structures during cell recognition, we performed the structural fingerprinting comparison between acid labile domains of three species. Estimation of apparent size of this material (expressed in DP Glc), by gel filtration and TLC analyses (data not shown) confirmed differences in the nature of the glyconectins peripheral domains. Oligosaccharides size ranged from DP 2 to DP 6 for GN1, from DP 15 to DP 20 for GN2 and from DP 2 to DP 5 for GN3. Exact measurement of oligosaccharides molecular weight by MALDI mass spectrometry in positive and negative modes confirmed that the three components showed very different structural features (Table 3). GN1 and GN3 consisted in relatively simple mixtures of short oligosaccharides. The vast majority of these oligosaccharides were observed in negative mode, suggesting that they exhibited a substantial negative charge. A single oligosaccharide showing a [M+Na] mass at 622 was identified in positive mode in GN1, whereas no one was observed in GN3. Apart from this component, calculated composition of all oligosaccharides

established that they contain a sulfate group, in accordance with the high sulfate content of native glyconectins isolated from GN1 and GN3. However, mass spectrometry fingerprinting clearly demonstrated that despite these common features (size, nature of the charge) peripheral oligosaccharide chains from these two species showed distinct species-specific structures.

As previously suggested by size fractionation and TLC analysis, acid labile material from GN2 appears as very different from GN1 and GN3. In mass spectrometry analysis, this material appeared as a very complex mixture of about 25 acidic oligosaccharides which molecular mass ranged from 1698 to 3076. The large size of these components did not allow a precise calculation of their individual composition. Thus we used NMR and methylation analyses to obtain structural information. Considering the size of the oligosaccharides, the high proportion of Py(4,6)Gal residues (29 %) suggested that each molecule comprised from 4 to 6 of these residues. A combination of ¹H-¹H HOHAHA and ¹H-¹³C NMR experiments clearly established the presence of terminal β -D-Py(4,6)Galp in this material (H-1, 4.817) ppm; H-2, 3.656; H-3, 3,696; H-4, 4.155; H-5, 3.543; H-6, 4.051-4.060; H-6', 3.923-3.942; C-4, 72.21 ppm; C-5, 67.62; C-6, 66.37) in accordance with published ¹³C and ¹H-NMR parameters (22-24). Furthermore, these experiments showed no evidence to substantiate the presence of substituted Py(4,6)Gal residues in the oligosaccharide mixture, proving that all Py(4,6)Gal residues were present in terminal non-reducing position. Methylation analysis on the oligosaccharides allowed identification of permethylated galactose and permethylated pyruvated hexose in the ration 1:1. Altogether, results from composition, linkage and NMR analyses demonstrated that material released by mild hydrolysis from glyconectins isolated from GN2 are large, highly branched oligosaccharides, capped in equal amount by terminal Py(4,6)Gal residues and Gal residues.

In conclusion biochemical analysis of Glyconectin 1, 2 and 3 revealed a novel glycoconjugate nature residing in 20-60 % of carbohydrate mass with over 5000 mol/mol of negative charge contributed by carboxyl and sulfate groups (Table 1). Presence of over 10% of fucose beside glucuronic acid, N-acetyl glucosamine and galactose indicated a more complex blockopolymeric glycan sequences to that of classical glycosaminoglycans of mammalian proteoglycans which are characterized by repeats of disaccharide units. Amino acid analysis of three sponge proteoglycans showed over 7% of serine and threonin and over 10% of acidic amino acids Table 1. Determined physico-chemical properties of highly glycosylated and blockoplymeric acidic sponge macromolecules define them as a new class of primordial proteoglycan-like molecules.

The cell adhesive function of two new sponge glyconectin proteoglycans purified from *Halichondria panicea* (GN2) and *Cliona celata* (GN3) was tested and compared to that of *Microciona prolifera* (GN1) in a rotary reaggregation assay with live metabolically attenuated and/or fixed cells depleted of proteoglycans. All three glyconectin proteoglycans mediated cell adhesion in the presence of physiological sea water with 10 mM CaCl₂, and not at concentrations below 1mM CaCl₂ (Fig. 2A). Magnesium ions could not replace Ca²⁺ (Fig. 2D), indicating essential and specific role of glyconectins and Ca²⁺ in sponge cell adhesion.

Ca²⁺ Dependent Glyconectin to Glyconectin Interactions Mediate Cell Adhesion.

We have recently shown by atomic force microscopy and glyconectin coated bead adhesion that Ca²⁺ dependent carbohydrate to carbohydrate self-interactions between glyconectin 1 from Microciona prolifera sponge provide the major driving force for cell adhesion (12,14, 3). To quantitatively test whether the glyconectins 2 and 3 from sponges Halichondria panicea and Cliona celata also utilize the same adhesive molecular mechanism via Ca²⁺ dependent glyconectin to glyconectin interactions we have used two assay systems, Ca²⁺ induced gelation of glyconectins (Fig. 2B) and aggregation of glyconectin coated bead aggregation (Fig. 2C). All three glyconectins revealed 100% self-interaction at 10 mM CaCl₂ and no interactions in the absence of calcium ions in both assay systems (Fig. 2B and C), see also bead recognition assay in Fig. 4 and 5). CaCl₂ could not be substituted with MgCl₂ indicating specific role of Ca ions (Fig. 2E and F) Titration experiments of Ca²⁺ concentration dependence of sponge glyconectin self-interactions revealed a transition at 5mM and 100% interactions at physiological 10 mM CaCl₂ (Fig. 2B and C), identical to that of Ca²⁺ dependent glyconectin promoted cell adhesion (Fig. 2A) These experiments indicated that a Ca²⁺ dependent glyconectin to glyconectin interactions play a pivotal role in cell adhesion of the three selected marine sponge species.

Specific Glyconectin to Glyconectin Interactions Mediate Biological Self-Assembly in Porifera Cellular Recognition and Adhesion.

In order to test whether cell adhesive glyconectin-glyconectin interactions can also mediate xenogeneic recognition in sponges, we have performed three new types of controlled *ex vivo* color coded cell-cell, bead-bead and cell-bead recognition experiments. This is essential step towards understanding molecular mechanism of previously reported phenomenological and

biochemical studies about the role of proteoglycans in binary xenogeneic reaggregation assays of dissociated sponge cells (Wilson, Humphreys and others). In our first assay we used adhesion of proteoglycan bearing cells in a trinary species combination. Live metabolically attenuated cells (at 0°C) or fixed cells of three different species *Microciona prolifera* (Fig. 3A), *Halichondria panacea* (Fig. 3B) and *Cliona celata* (Fig. 3C), were mixed in the presence (Fig. 3E), and absence (Fig. 3D) of 10 mM Ca²⁺ (physiological concentration in seawater). In a rotary assay within 5-15 min species-specific recognition and adhesion occurred only with 10 mM Ca²⁺, as monitored through different natural pigments of sponge cells (Fig. 3E). Upon removal of proteoglycans from cell surface by repetitive washing none of the three species displayed aggregation. Adding the purified proteoglycans back to the same live cells at 0°C or to the proteoglycan depleted fixed cells completely restored species-selective cohesion (results same as in Fig. 3D). These experiments indicated that specific cell surface glyconectin proteoglycans and Ca²⁺ are beside cell adhesion also instrumental in the initial steps of xenogeneic cell recognition of the three species, leading us to name them glyconectin 1 for *M. prolifera*, glyconectin 2 for *H. panicea*, and glyconectin 3 for *C. celata*.

The search for the molecular mechanism underlying such self-recognition, even in the mostly studied marine sponge Microciona prolifera, remained only on the hypothesis that not characterized and not cloned cell surface receptor may be involved in recognition of the proteoglycan (25,26). Taking in account that proteoglycans are most peripheral cell surface macromolecules we rather proposed a novel mechanism of species-specific proteoglycan to proteoglycan interactions as a primary event mediating both self-recognition and adhesion. If indeed glyconectin-glyconectin interactions are solely instrumental also for cell recognition, glyconectin functionalized and color coded beads should show species-specific binding and sorting identical to glyconectin bearing cells. We have, thus, performed three types of assays, first and second combined glyconectin color coated beads aggregation and overlay to blotted glyconectins, and second one utilized color coded glyconectin bead interactions with glyconectin separated on agarose gel. For all assays we used glyconectin 1 attached to fluorescent pink, glyconectin 2 to fluorescent green-yellow, and glyconectin 3 to fluorescent blue latex-amidine beads. In combined first and second assay not labeled glyconectins were immobilized on a nitrocellulose membrane in such a manner that the three molecules were used to draw the subsequent letters of the words GLYCONECTIN RECOGNITION. The three bead types were mixed and added to the coated membrane in the presence or absence of 10 mM CaCl₂ (Fig. 4). Within 5-15 min of constant rotation species-specific bead aggregation and homophilic recognition between membrane-bound and bead glyconectins were observed through three

separate colored aggregates and selective staining of each letter (Fig. 3B). Both processes occurred at apparently similar rates for each of the three glyconectins. In control experiments with glyconectin 1 separately attached to pink, yellow and white beads, as expected, mixed orange color aggregates were formed upon addition of 10 mM CaCl₂ (not shown). In the absence of 10 mM CaCl₂, bead aggregation did not occur either in the mixture of three glyconectins or of one glyconectin coated to three color beads (same as Fig. 4A). The third assay used same type of color coated glyconectin beads but this time overlaid on agarose gel containing electrophoretically separated three glyconectins. After over night incubation at room temperature in the presence of 10 mM CaCl₂ under gentle agitation species specific staining of gel glyconectin bands identical to ones stained with Toluidine blue and Amido black showed that glyconectin to glyconectin interactions are highly species specific and resemble the degree of selectivity similar to those observed for antibody antigen interactions in similar assay system. The combinations of the above described experiments demonstrate species-specific molecular self-recognition of glyconectins in an elementary bead adhesion system which fully resembles glyconectin mediated cell-cell and cell-bead recognition and adhesion. This simple bead, membrane and gel overlay assay systems mimic interactions between sessile and free cellular forms in a primeval environment. They, thus, show that glyconectins could enable self and non-self discrimination in both solution and on surface, and support our theory that selective homophilic glyconectinglyconectin associations play solely a crucial role in the initial step of sponge cell adhesion and xenogeneic recognition.

Self-Assembly of Biological and Artifical Cell-Bead System via Species-Specific Glyconectins Recognition.

To further evaluate the self-recognition role of glyconectins in *ex vivo* experiments and to bridge the gap between complex and multicomponent cell recognition systems to that of glyconectin bead experiments we have determined specificity of homophilic glyconectin-glyconectin interactions in a color coded cell-bead adhesion assay. Each of the three glyconectin attached to color coded beads, as in previous experiment, was mixed in nine combinations with glyconectin bearing cells of the three sponge species, in the presence of 10 mM Ca²⁺. In order to simultaneously follow cells and beads visible light was used for microscopy. Thus, fluorescent pink for glyconectin 1 corresponds to pink, fluorescent green-yellow for glyconectin 2 to yellow, and fluorescent blue for glyconectin 3 to white latex-amidine beads. Bead-cell coaggregation was observed for homotypic combinations as could be microscopically observed by color and shape of uniformly mixed beads and cells (Fig. 5A, E, I). The heterotypic combinations always resulted in bead and cell segregation (Fig. 5B, C, D, F, G and H). Cell-bead recognition and adhesion occurred at apparently equal rates and ended in similar aggregate sizes for all three species. Such recognition occurred only when cells had glyconectin on their surface and in the presence of 10 mM CaCl₂. This process did not require live cells since same results were obtained with glutaraldehyde fixed and/or metabolically attenuated cells at 0° C. These experiments showed that independently to what is glyconectins attached, as in examples of either living, or metabolically attenuated, or fixed cell, or even to nonnatural latex bead, they are able to recognize self and distinguish from non-self via cohesive proteoglycan to proteoglycan interactions.

DISCUSSION

The distinct functional and physicochemical properties of isolated cell surface sponge macromolecules suggest that, although they share general features of proteoglycans such as large size and long sulfated negatively charged glycans, they define a new class of primordial cell recognition and adhesion proteoglycans, named glyconectins. Recently reported partial sequence analysis of carbohydrate moiety (27, 28), cloning of the gene coding for the protein core of the *M. prolifera* proteoglycan-like molecule (29), carbohydrate analysis, enzymological and immunological studies of all three sponge proteoglycans (Table 1). Glycosaminoglycans of mammalian proteoglycans are characterized by a single type of repetitive disaccharide unit (20) while sponge glycans revealed higher complexity of acidic and fucosilated blockopolymeric sequences. These results suggest that highly glycosylated and acidic sponge macromolecules define a new class of primordial proteoglycans named glyconectins.

The degree of selectivity observed in the evolutionarily advanced heterophilic self and non-self molecular recognition within the immunoglobulin superfamily is closely approached by the specificity of homophilic glyconectin to glyconectin interactions in Porifera. However, structural differences between these two systems imply conceptually distinct molecular mechanisms. First, glyconectins are 100 times larger and 10 times more extended from cell surface than immunoglobulin molecules. Second, the glyconectin 1 specificity and tight binding of >10⁹ M⁻¹ reside in polyvalent carbohydrate-carbohydrate interactions of >1000 sites, with a low affinity for the single site (<10³ M⁻¹) (13, 14, 30), whereas immunoglobulins recognize antigens via higher affinity ranging from 104 to 10⁹ M⁻¹, with low valency binding (30-34, 38).

Glyconectin-glyconectin interactions may thus provide a new paradigm for initial molecular selfrecognition.

The close evolutionary relationship between Porifera today and ancestors of multicellular organisms (35-37) indicates that xenogeneic discrimination of self-cohesive glyconectins may be essential for the appearance of multicellularity. Clearly the evolution of primordial Metazoans to complex organisms required development of additional cell recognition and adhesion mechanisms mediated via immunoglobulin, lectin, integrin, and cadherin families of molecules. Such an increase of molecular diversity enables functional expansion to more demanding self and non-self discrimination, embryonal morphogenesis, renewal of adult tissues and defense from foreign pathogens. In spite of the evolving complexity of components participating in cellular interactions, the glyconectin-like structures have been preserved in mammalian systems suggesting versatile recognition and adhesion functions.

On the basis of using the simplest Metazoans of the phylum Porifera as the model system for primordial xenogeneic self-recognition and adhesion we suggest that the establishment of the first multicellular organisms, as well as the further divergence of species and appearance of more complex multicellular forms of life may have been achieved by the evolution of glyconectin-like proteoglycan molecules with the capacity for self-recognition and adhesion.

ACKNOWLEDGEMENTS. We thank to V. Norris, and Z. Housley, for helpful discussions. This work was mainly supported by private funds of G. N. Misevic, and in part by Conseil Régional Nord-Pas de Calais, CNRS, University of Rouen.

REFERENCES

- 1. Varki, A. (1993) Glycobiology 3, 97-130.
- 2. Dwek, R. A. (1996) Chem. Rev. 96, 683-720.
- Dammer, U., Popescu, O., Wagner, P., Anselmetti, D., Güntherodt, H.-J. & Misevic, G. N. (1995) Science 267, 1173-1175.
- 4. Wilson, H. V. (1907) J. Exp. Zool. 5, 245-258.
- 5. Wilson, H. V. (1910) Bull. Bur. Fisheries 30, 1-30.

- 6. Galstoff, P. S. (1925) J. Exp. Zool. 42, 183-221.
- 7. Curtis, A. S. G. Nature (1962) 196, 245-248.
- 8. Moscona, A. A. (1968) Dev. Biol. 18, 250-277.
- 9. Müller, W. E. G. & Zahn, R. K. (1973) Exp. Cell Res. 80, 95-104.
- 10. Humphreys, T. (1963) Dev. Biol. 8, 27-47.
- 11. Caudwell, C. B., Henkert, P. & Humphreys, T. (1973) Biochemistry 12, 3051-3055.
- 12. Jumblatt, J. E., Schlup, V. and Burger, M. M. (1980) Biochemistry 19, 1038-104212.
- 13. Misevic, G. N., Finne, J. & Burger, M. M. (1987) J. Biol. Chem. 262, 5870-5877.
- 14. Misevic, G. N. & Burger, M. M. (1993) J. Biol. Chem. 268, 4922-4929.
- 15. Papakonstantinou, E. A. and Misevic, G. N. (1993) J. Cell. Biochem. 53, 98-113.
- 16. Popescu, O. and Misevic, G. N. (1997) Nature 386, 321-322.
- 17. Kamerling, J.P., Gerwig, G.J., Vliegenthart, J.F. & Clamp, J.R. (1975) Biochem. J. 151, 491-495
- 18. Zanetta, J.P., Timmerman, P. & Leroy, Y. (1999) Glycobiology 9, 255-266
- 19. Ciucanu, I. & Kerek, F. (1984) Carbohydr. Res. 131, 209-217
- 20. Hassell, J. R., Kimura, J. H. & Hascall, V. C. (1986) Annu. Rev. Biochem. 55, 539-567.
- 21. Dudman, W.F. and Lacey, M.L. (1986) Carbohydr. R., 145: 175-191
- 22. Yang, B.Y, Brand, J. and Montgomery, R. (2001) Carbohydr. Res., 331:59-67
- 23. Cérantola, S., Marty, N. and Montrozier, H. (1996) Carbohydr. Res. 285: 59-67
- 24. Van Calsteren, M.R., Pau-Roblot, C. Bégin, A. and Roy, D. (2002) Biochem. J., 363: 7-17.
- 25. Burger
- 26. Burger
- 27. Spillmann, D. et al. (1993) J. Biol. Chem. 268, 13378-13387.
- 28. Spillmann, D. et al. (1995) J. Biol. Chem. 270, 5089-5097.
- Fernandez-Busquets, X., Kammerer, R. A. & Burger, M. M. (1996) J. Biol. Chem. 271,
 23558-23565&22. Hardingham, T. E. & Fosang, A. J. (1992) FASEB J. 6, 861-870.
- 30. Misevic, G. N. (1991) TIGG 3, 400-405.
- 31. Gallagher, J. T. in *Glycoimmunology* (eds Alavi, A. & Axford, J. S.) 125-134 (Plenum Press, New York, 1995).
- 32. Iozzo, R. V. & Murdoch, A. D. (1996) FASEB J. 10, 598-614.
- 33. van der Merwe, P. A. & Barclay, A. N. (1994) TIBS 19, 354-358.
- 34. Corr, M. et al. (1994) Science 265, 946-949.
- 35. Garrone, R. Phylogenesis of Connective Tissue (S. Karger-Verlag, Basel, 1978).

- 36. Bergquist, P. R. Sponges (Hutchinson and Co., London, 1978).
- 37. Simpson, T. L. The Cell Biology of Sponges (Springer-Verlag, New York, 1984).
- 38. Misevic, G. N. & Popescu, O. (1995) J. Mol. Recogn. 8, 100-105.

FIGURE LEGENDS

Fig. 1. Electrophoretic separation of sponge glyconectins. A) 0.75% agarose gel stained with 0.02% Toluidine blue followed by 0.1% Amido black 10B. *a-c*, GNs from *M. prolifera* GN1, *H. panicea* GN2, *C. celata* GN3, respectively (10 μ g each). B) 0.75% agarose gel stained with color coded fluorescent beads coated with *a*) GN1 pink, *b*) GN2 green; and *c*) GN3 blue in the presence of SWT with 10 mM CaCl₂.

Fig. 2. Calcium ions dependent GN promoted cell aggregation and self assembly gelation glyconectins and glyconectin coated beads. A-D) Cell aggregation promoted by GNs was performed as previously described (13,14). B-E) Gelation assay was performed with 10 μ g GNs in 100 μ l CSW. B) CaCL₂ range from 0 to 20 mM, (B) 10 mM MgCl₂. After 20 min incubation at room temperature samples were centrifuged at 5000 x g and pellet and supernatant were analyzed in a colorimetric and immunodot assay for the presence of GNs. C-F) Bead aggregation assay was done as described in Material and Methods C) CaCl₂ concentration ranged from 0 to 20 mM, (F) 10 mM MgCl₂. All three assays used similar conditions for self interactions of glyconectins.

Fig. 3. Glyconectin glycoconjugates are cell adhesion and recognition molecules. Ca^{2+} dependent glyconectin to glyconectin interactions mediate species-specific cell-cell recognition and adhesion. A) *M. prolifera* sponge, B) *H. panacea*, and C) *C. celata* living sponges. Self and <u>non-self</u> discrimination and adhesion in the suspension of mixed *M. prolifera* (orange), *H. panicea* (white) and *C. celata* (brown) live cells bearing glyconectins. D) SWT without 10 mM Ca^{2+} , and E), SWT with 10 mM Ca^{2+} at 0°C after 20 min rotation. Microscopically observed color of cells is somewhat different than that of the whole sponge. Early cell sorting experiments were usually done with binary sponge combinations at room temperature without rotation. The sorting is thus dependent on the presence of recognition molecules at the cell surface, cell motility, and speed of new synthesis and/or secretion of additional recognition molecules. Our rotary assays using either metabolically attenuated or fixed cells reduce the number of variable parameters.

Fig. 4. Simultaneous species-specific glyconectin to glyconectin recognition in a suspension and blotting assay. Letters were drawn using 4 ml of 1.5 mg/ml glyconectins on a *Hybond-C extra* nitrocellulose membrane (Amersham) and probed in SWT with pink, blue and green fluorescent

beads coated with glyconectin 1, 2 and 3, respectively. A) SWT without 10 mM Ca^{2+} and B) SWT with 10 mM Ca^{2+} . All photographs were taken after 30 min of mixing.

Fig. 5. Species-specific glyconectin to glyconectin interactions mediate bead-cell recognition and adhesion. Xenogeneic glyconectin self-recognition in a mixture of glutaraldehyde-fixed cells and glyconectin coated beads in the presence of 10 mM Ca²⁺. *M. prolifera* cells bearing glyconectin 1 were incubated with: A, glyconectin 1 pink beads, D, glyconectin 2 yellow beads, and G, glyconectin 3 white beads. *H. panicea* cells bearing glyconectin 2 were incubated with: B, glyconectin 2, and H, glyconectin 3 color coded beads. *C. celata* cells bearing glyconectin 3 were incubated with: C, glyconectin 1, F, glyconectin 2, and I, glyconectin 3 color coded beads (Glutaraldehyde fixation changes cell colors i.e. *M. prolifera* orange to yellow-white, *H. panicea* white to yellow-brown, and *C. celata* brown to brown-orange. We did not observe differences in adhesion properties between fixed and live metabolically attenuated cells in a rotary assay). Magnification 40x.

 Table 1. Physicochemical properties of sponge glyconectin

	Glyconectins purified from			
	M. prolifera	H. panicea	C. celata	
M _r x 10 ⁶	19	10	8	
s _{20,w} (S)	58	42	46	
Carbohydrate/protein	ratio (w/w) 63/37	21/79	36/64	
Aminoacid composition	n (mol%)			
Asx	11.1	9.1	7.8	
Glx	10.4	9.2	9.5	
Ser	6.7	7.0	11.3	
Gly	9.0	9.9	10.9	
Arg	5.6	7.6	6.0	
Thr	9.9	10.2	14.1	
Ala	7.1	7.0	7.7	
Pro	7.3	8.2	10.7	
Tyr	2.8	4.0	0.7	
Val	9.2	8.6	6.1	
Met	2.4	2.5	2.4	
Cys	0.1	0.1	0.2	
Ile	5.4	6.0	3.9	
Leu	7.8	5.5	5.1	
Phe	5.1	4.8	3.6	
Lys	0.2	0.2	0.1	
Carbohydrate compos	ition (mol%)			
Ara	0	0	11	
Fuc	28	7	31	
Man	9	16	20	
Gal	29	43	12	
GalNAc	1	1	7	
GlcNAc	17	14	12	
Py(4,6)Gal	4	16	0	

Molecular weight (M_r) and sedimentation coefficient $(s_{20,w})$ of sponge proteoglycans were determined by analytical centrifugation (6, 7). Carbohydrate/protein ratio, content of uronic acids and SO_4^{2-} were calculated from amino acid, monosaccharide and ion composition. Standard deviation from three analysis was maximaly 1% of each value (see Methods).

GlcUA

 SO_4^{2-} (mol/mol)

	GN1 mol%		GN2 mol%		GN3 mol%	
	Supernata	n Pellet	Supernatant	Pellet	Supernatant	Pellet
				ļ		
Ara	0	0	0	0	10	5
Fuc	36	11	4	5	22	14
Man	4	12	18	33	19	18
Gal	27	37	28	33	8	15
GlcA	3	7	0	0	0	5
GlcNAc	28	29	19	27	26	38
GalNAc	0	0	1	0	15	6
4,6pyr-Gal	2	4	29	2	0	0
					1	

Table 2. Fragmentation of Glyconectin Polysaccharides

Glyconectins were digested by mild acid treatment (0.1 N TFA, 80°C, 1 h.). Carbohydrate analysis was conducted as described in Methods on liberated oligosaccharides (supernatant) and on acid-resistant core (pellet).

	GN1	GN2		GN3
Average mass	Calculated composition	Average mass	Average mass	Calculated composition
300	HexNAc, SO3	1697 1738 1850	432	HexNAc, Pent SO3
446	HexNAc, deHex SO ₃	1900 2021 2062	446	HexNAc, deHex SO ₃
592	HexNAc, 2xdeHex	2132 2183 2224	578	HexNAc, deHex, Pent
608	HexNAc, Hex	2265 2294 2335	592	HexNAc, 2xdeHex SO3
624	HexNAc, 2xHex SO ₃	2427 2456 2497	635	2xHexNAc, Pent SO3
770	deHex, HexNAc, 2xHex	2589 2751 2913	649	2xHexNAc, deHex
	SO3	3075		SO3
793	2xdeHex, 2xHexNAc SO3		781	2xHexNAc, Pent, deHex SO3
622	Py-Hex, HexNAc, Fuc		795	2xdeHex, 2xHexNAc SO3
			943	2xHexNAc, Hex, deHex SO3
			984	3xHexNAc, Pent, deHex SO3

Average molecular mass of the liberated oligosaccharides from GN1 and GN3 was measured by MALDI-MS. Compositions were calculated from mass values. They correspond to [M-H] ions, except for compound of mass 622 [M+Na] ion from GN1. The high mass of oligosaccharides from GN2, did not permit exact evaluation of composition (see results).

.

Table 3. Mass-Spectrometry Fingerprinting of Glyconectin Oligosaccharides.



Fig. 1. Electrophoretic separation of sponge glyconectins. A) 0.75% agarose gel stained with 0.02% Toluidine blue followed by 0.1% Amido black 10B. *a-c*, GNs from *M. prolifera* GN1, *H. panicea* GN2, *C. celata* GN3, respectively (10 μ g each). B) 0.75% agarose gel stained with color coded fluorescent beads coated with *a*) GN1 pink, *b*) GN2 green, and *c*) GN3 blue in the presence of SWT with 10 mM CaCl₂.



Fig. 2. Calcium ions dependent GN promoted cell aggregation and self assembly gelation glyconectins and glyconectin coated beads. A-D) Cell aggregation promoted by GNs was performed as previously described (13,14). B-E) Gelation assay was performed with 10 μ g GNs in 100 μ l CSW. B) CaCL₂ range from 0 to 20 mM, (B) 10 mM MgCl₂. After 20 min incubation at room temperature samples were centrifuged at 5000 x g and pellet and supernatant were analyzed in a colorimetric and immunodot assay for the presence of GNs. C-F) Bead aggregation assay was done as described in Material and Methods C) CaCl₂ concentration ranged from 0 to 20 mM, (F) 10 mM MgCl₂. All three assays used similar conditions for self interactions of glyconectins.



Fig. 3. Glyconectin glycoconjugates are cell adhesion and recognition molecules. Ca^{2+} dependent glyconectin to glyconectin interactions mediate species-specific cell-cell recognition and adhesion. A) *M. prolifera* sponge, B) *H. panacea*, and C) *C. celata* living sponges. Self and <u>non-self</u> discrimination and adhesion in the suspension of mixed *M. prolifera* (orange), *H. panicea* (white) and *C. celata* (brown) live cells bearing glyconectins. D) SWT without 10 mM Ca^{2+} , and E), SWT with 10 mM Ca^{2+} at 0°C after 20 min rotation. Microscopically observed color of cells is somewhat different than that of the whole sponge. Early cell sorting experiments were usually done with binary sponge combinations at room temperature without rotation. The sorting is thus dependent on the presence of recognition molecules at the cell surface, cell motility, and speed of new synthesis and/or secretion of additional recognition molecules. Our rotary assays using either metabolically attenuated or fixed cells reduce the number of variable parameters.



Fig. 4. Simultaneous species-specific glyconectin to glyconectin recognition in a suspension and blotting assay. Letters were drawn using 4 ml of 1.5 mg/ml glyconectins on a *Hybond-C extra* nitrocellulose membrane (Amersham) and probed in SWT with pink, blue and green fluorescent beads coated with glyconectin 1, 2 and 3, respectively. A) SWT without 10 mM Ca²⁺ and B) SWT with 10 mM Ca²⁺. All photographs were taken after 30 min of mixing.



Fig. 5. Species-specific glyconectin to glyconectin interactions mediate bead-cell recognition and adhesion. Xenogeneic glyconectin self-recognition in a mixture of glutaraldehyde-fixed cells and glyconectin coated beads in the presence of 10 mM Ca²⁺. *M. prolifera* cells bearing glyconectin 1 were incubated with: A, glyconectin 1 pink beads, D, glyconectin 2 yellow beads, and G, glyconectin 3 white beads. *H. panicea* cells bearing glyconectin 2 were incubated with: B, glyconectin 1, E, glyconectin 2, and H, glyconectin 3 color coded beads. *C. celata* cells bearing glyconectin 3 were incubated with: C, glyconectin 1, F, glyconectin 2, and I, glyconectin 3 color coded beads (Glutaraldehyde fixation changes cell colors i.e. *M. prolifera* orange to yellow-white, *H. panicea* white to yellow-brown, and *C. celata* brown to brown-orange. We did not observe differences in adhesion properties between fixed and live metabolically attenuated cells in a rotary assay). Magnification 40x.

CONCLUSION GENERALE

L'ensemble des travaux présentés dans ce mémoire concernent l'étude de la diversité structurale des glycannes. Ainsi, les variabilités inter- et intra-spécifiques des O-glycannes de type mucine, des lipoglycannes et d'un nouveau type de glycoconjugués, les glyconectines, ont successivement été observées dans divers modèles expérimentaux. L'analyse de la structure des O-glycannes de type mucines chez deux espèces d'amphibiens et chez le nématode C. elegans rend compte de l'absolue nécessité de définir la structure des glycannes endogènes avant de procéder à l'étude des activités glycosyltransférasiques de tout modèle inconnu. Ce point de vue est illustré par les O-glycannes de C. elegans dont la biosynthèse requièrt des activités glycosyltransférasiques qui étaient jusqu'alors inconnues et dont la caractérisation partielle n'a été possible que sur la base d'informations structurales. De même, la mise en évidence de structures O-glycanniques supportant le déterminant Le^a chez l'amphibien X. tropicalis a démontré l'existence d'une nouvelle glycosyltransférase dans le règne animal capable de transférer une unité d'a-1,4-Fuc. Considérant les distances génétiques considérables séparant les amphibiens et les primates, qui sont les seuls représentants du règne animal chez qui cette activité avait jusqu'alors été détectée, il est fort probables que leurs α-1,4-fucosyltransférases respectives aient des origines très différentes. Il est à prévoir que l'étude de ces enzymes fournissent de nombreuses informations quant à l'émergence et à l'évolution des activités fucosyltransférasiques en général. Enfin, l'étude comparée des profils de glycosylation de X. tropicalis et de X. laevis a révélé que malgré leur proximité phylogénétique, ces deux espèces synthétisaient des structures glycanniques très différentes, ce qui confortait la notion de spécificité des structures glycanniques en fonction de l'espèce, chez les amphibiens.

L'analyse de la structure des lipoarabinomannanes et lipomannanes chez deux espèces de mycobactéries pathogènes a permis de mettre en évidence une variabilité structurale significative de ces composés tant au niveau de l'ancre lipidique que du corps polysaccharidique. L'impact potentiel de ces variations structurales sur la fonctionnalité de ces composés reste à définir. Néanmoins, les premières études que nous ayons effectuées sur les relations entre structure et fonctions des lipoglycannes ont permis de réévaluer l'importance des lipomannanes dans l'établissement de la réponse inflammatoire. En effet, dans les conditions expérimentales utilisées, il a été démontré que le domaine arabinane atténuait le potentiel des lipoglycannes à induire la sécrétion de cytokines pro-inflammatoires par des cellules macrophagiques. De fait, ces résultats suggèrent qu'en ce qui concerne les activités pro-inflammatoires, à l'exception des LAMs substitués par des groupements phosphatidylinositol (PILAM), les LAMs seraient des formes atténuées des LMs. S'il en est ainsi, nous émettons l'hypothèse selon laquelle l'activité pro-inflammatoire globale d'une mycobactérie puisse être régulée par la proportion relative des LMs et des LAMs présents au sein de sa paroi. Cette proportion pourrait elle-même être contrôlée par l'activité des arabinosyltransférases impliquées dans la synthèse du domaine arabinane. La validation d'une telle théorie devra attendre que les arabinosyl-transférases aient été identifiées, ce qui devrait permettre d'en contrôler l'activité, *in vivo*.

Enfin, l'étude des glyconectines isolées de trois espèces d'éponge a démontré que la structure de leur partie saccharidique était également très spécifique à chaque espèce. Ces travaux ont, pour la première fois, donné une base moléculaire à la compréhension de la spécificité de reconnaissance des glyconectines entre elles. En effet, la définition exacte de l'arrangement du corps polysaccharidique de ces composés permettra de modéliser avec exactitude les interactions de type sucre-sucre présidant à l'auto-agrégation spécifique d'espèce des cellules d'éponge dissociées.

REFERENCES

- Adam, A., Petit, J.F., Weitzerbin-Falszpan, J., Sinay, P., Thomas D.W. and Lederer, E. (1969) L'acide N-glycolyl-muramique, constituant des parois de *Mycobacterium smegmatis*; identification par spectrométrie de masse. *FEBS Lett.* 4, 87-92
- Al-Anzi, B. and Chandler, D. (1998) A sperm chemoattractant is released from *Xenopus laevis* egg jelly during spawning. *Dev. Biol.* 198, 366-375.
- Altmann, F., Fabini, G., Ahorn, H. and Wilson, I.B.H. (2001) Genetic model organisms in the study of N-glycans. *Biochimie* 83, 703-712
- Alves, AP., Mulloy, B., Diniz, J.A. and Mourao, P.A. (1997) Sulfated polysaccharides from the egg jelly layer are species specific inducers of acrosomal reaction in sperm of sea urchins. J. Biol. Chem. 272, 6965-6971
- Amar-Nacasch, C. and Vilkas, E. (1970) Etude des parois de Mycobacterium tuberculosis. II. Mise en évidence d'un mycolate d'arabinobiose et d'un glucane dans les parois de M. tuberculosis H₃₇Ra. Bull. Soc. Chim. Biol. 52, 145-151
- Ariza, M.A., Martin-Luengo, F. and Valero-Guillen, P.L. (1994) A family of diacyltrehaloses isolated from *Mycobacterium fortuitum*. Microbiology 140, 1989-1994
- Asker, N., Axelsson, M.A.B, Olofsson, S.O and Hansson, G.C. (1998b) Dimerization of the human MUC2 mucin in the endoplasmic reticulum is followed by a *N*-glycosylation-dependent transfer of the mono- and dimers to the Golgi apparatus. J. Biol. Chem. 273, 18857-18863
- Asker, N., Axelsson, M.A.B, Olofsson, S.O and Hansson, G.C. (1998a) Human MUC5AC mucin dimerizes in the rough endoplasmic reticulum, similarly to MUC2 mucin. *Biochem. J.* 335, 381-387
- Asselineau, J. and Lanéelle, G. (1998) Mycobacterial lipids: a historical perpective. Front. Biosci. 3, 164-174
- Azuma, I., Kimura, H., Ninaka, T., Aoki, I. and Yamamura, Y. (1968) chemical and immunological studies on mycobacterial polysaccharides. Purification and properties of polysaccharides from human tubercle bacilli. J. Bacteriol. 95, 263-271
- Ballou, C.E. (1972) Biosynthesis of mannophosphoinositides in Mycobacterium phlei. In Methods in Enzymology XXVIII, Ginsburg V., Academic Press, 493-500
- Ballou, C.E., Vilkas, E. and Lederer, E. (1963) Structural studies on the myo-inositol phospholipids of Mycobacterium tuberculosis (var. bovis, strain BCG). J. Biol. Chem. 238, 69-76
- Barnes, P.F., Chatterjee, D., Abrams, J.S., Lu, S., Wang, E., Yamamura, M., Brennan, P.J. and Modlin, R.L. (1992) Cytokine production induced by *Mycobacterium tuberculosis* lipoarabinomannan. Relationship to chemical structure. J. Immunol. 149, 541-547
- Barr, M.M. and Sternberg, P.W. (1999) A polycystic kidney disease gene homologue required for male mating behaviour in C. elegans. *Nature* 401, 386-389
- Bekierkunst, A. (1968) Acute granulomatous response produced in mice by trehalose-6,6'-dimicolate. J. Bacteriol 96, 958-951
- Bekierkunst, A., Levij, I.S., Yarkoni, E., Vilkas, E. and Lederer (1971) Suppression of urethaneinduced lung adenome in mice treated with trehalose-6,6'-dimycolate (cord factor) and living bacillus Calmette-Guérin. *Science* 174, 1240
- Bekierkunst, A., Levij, I.S., Yarkoni, E., Vilkas, E., Adam, A. and Lederer (1969) Granuloma formation induced in mice by chemically defined mycobacterial fractions. *J. Bacteriol.* 100, 95-102
- Belisle, J.T., Vissa, V.D., Sievert, T., Takayama, K., Brennan, P. and Besra, G.S. (1997) Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. *Science* 276, 1420-1422
- Bell, S.L., Khatri, I.A, Xu, G. and Forstner, J.F. (1998) Evidence that a peptide corresponding to the rat Muc2 C-terminus undergoes disulphide-mediated dimerization. *Eur. J. Biochem.* 253, 123-131
- Beltan, E., Horgen, L. and Rastogi, N. (2000) Secretion of cytokines by human macrophages upon infection by pathogenic and non-pathogenic mycobacteria. *Microbial Pathogenesis* 28, 313-318
- Bennett, E.P., Hassan, H., Hollingsworth, M.A. and Clausen, H. (1999) A novel human UDP-Nacetylgalactosamine:Polypeptide N-Acetylgalactosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates. FEBS lett. 460, 226-230
- Bennett, E.P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M.A., Merkx, G., van Kessel, A.G., Eiberg, H., Steffensen, R. and Clausen, H. (1998) Cloning of a human UDP-*N*-acetyl-α-D-galactosamine:Polypeptide N-Acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. J. Biol. Chem. 273, 30472-30481
- Berman, J.S., Blumenthal, R.L., Corfeld, H., Cook, J.A., Cruikshank, W.W., Vermeulen, M.W., Chatterjee, D., Belisle, J.T. and Fenton, M.J. (1996) Chemotactic activity of mycobacterial lipoarabinomannans for human blood T lymphocytes in vitro. J. Immunol. 156, 3828-3835
- Bernardo, J., Billingslea, A.M., Blumenthal, R.L., Seetoo, K.F., Simons, E.R. and Fenton, M.J. (1998) Differential responses of human monoclear phagocytes to mycobacterial lipoarabinomannans: role of CD14 and the mannose receptor. *Infect. Immun.* 66, 28-35
- Besra, G.S. and Chatterjee, D. (1994) Lipids and carbohydrates of Mycobacterium tuberculosis. Tuberculosis: Pathogenesis, Protection and Control. Bloom B.R., 285-305
- Besra, G.S., Bolton, R.C., McNeil, M.R., Ridell, M., Simpson, K.E., Glushka, J., van Halbeek, H., Brennan, P. and Minnikin, D.E. (1992) Structural elucidation of a novel family of acyltrehalose from *Mycobacterium tuberculosis. Biochemistry* 31, 9832-9837
- Besra, G.S., McNeil, M. and Brennan, P.J. (1992b) Characterization of the specific antigenicity of Mycobacterium fortuitum. Biochemistry 31, 6504-6509
- Besra, G.S., McNeil, M., Khoo, K.H., Dell, A., Morris, H.R. and Brennan, P.J. (1993) Trehalosecontaining lipooligosaccharides of *Mycobacterium gordonae*: presence of a mono-O-methyltetra-Oacyltrehalose "core" and branching in the oligosaccharide backbone. *Biochemistry* 32, 12705-12714
- Besra, G.S., McNeil, M., Minnikin, D.E., Portaels, F., Ridell, M. and Brennan, P.J. (1991) Structural elucidation and antigenicity of a novel phenolic glycolipid antigen from *Mycobacterium haemophilum*. Biochemistry, 7772-7777
- Besra, G.S., Morehouse, C.B., Rittner, C.M., Waechter, C.J. and Brennan, P.J. (1997) Biosynthesis of mycobacterial lipoarabinomannan. J. Biol. Chem. 272, 18460-18466
- Bhaskar, K.R., Garik, P., Turner, B.S., Bradley, J.D., Bansil, R., Stanley, H.E and LaMont, J.T. (1992) Viscous fingering of HCl through gastric mucin. *Nature* 360, 458-461
- Bird, J.M. and Kimber, S.J. (1984) Oligosaccharide containing fucose linked alpha(1-3) and alpha(1-4) to N-acetylglucosamine cause decompaction of mouse morulae. *Dev. Biol.* 104, 449-460
- Blackler, A.W. and Gercking, C.A. (1972) Transmission of sex cells of one species through the body of a second species in the genus Xenopus. I & II Interspecific matings. *Dev. Biol.* 27, 376-394
- Bleil, J.D. and Wassarman, P.M. (1983) Sperm-egg interactions in the mouse: sequence of events and induction of the acrosome reaction by a zona pellucida glycoprotein. *Dev. Biol.* 95, 317-324
- Bloch, H. (1950) Studies on the virulence of tubercle bacilli. Isolation and biological properties of a constituent of virulent organisms. J. Exp. Med. 91, 197-218
- Bobek, L.A., Tsa, H., Biesbrock, A.R. and Levine, M.J. (1993) Molecular cloning, sequence, and specificity of expression of the gene encoding the low molecular weight human salivary mucin (MUC7). J. Biol. Chem. 268, 20563-20569
- Bonnel, B.S., Reinhart, D. and Chandler, D.E. (1996) *Xenopus laevis* egg jelly coat consists of small proteins bound of structurally stable network composed of high molecular weight glycoconjugates. *Dev. Biol.* 174, 32-42.
- Boubelik, M., Floryk, D., Bohata, J., Draberova, L., Macak, J., Smid, F. and Draber, P. (1998) Lex glycosphyngolipids-mediated cell aggregation. *Glycobiology* 8, 139-146
- Bourdon, M.A., Krusius, T., Campbell, S., Schwartz, N.B and Ruoslathi, E. (1987) Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. *Proc. Natl. Acad. Sci. USA* 84, 3194-3198
- Brennan, P. and Ballou, C.E. (1967) Biosynthesis of mannophosphoinositides by *Mycobacterium phlei*. The family of dimannophosphoinositides. J. Biol. Chem. 242, 3046-3056
- Brennan, P. and Ballou, C.E. (1967) Biosynthesis of mannophospholipids by *Mycobacterium phle*. The family of dimannophosphoinositides. J. Biol. Chem. 242, 3046-3056
- Brennan, P. and Nikaido, H. (1995) The envelope of mycobacteria. Annu. Rev. Biochem. 64, 29-63
- Brennan, P.J. (1997) Tuberculosis in the context of emerging and reemerging diseases. FEMS 18, 262-269
- Brockhausen, I. (1995) Biosynthesis of O-glycans of the N-acetylgalactosamine-a-Ser/Thr linkage type. In Glycoprotein. J. Montreuil, H. Schachter and J.F.G. Vliegenthart (Eds), Elsevier Science B.V., 201-259
- Brockhausen, I. (1999) Pathways of O-glycan biosynthesis in cancer cells. *Biochim. Biophys. Acta.* 1473, 67-95
- Brockhausen, I., Matta, K.L., Orr, J. and Schachter, H. (1985) Mucin synthesis. UDP-GlcNAc:GalNAc-R beta 3-N-acetylglucosaminyltransferase and UDP-GlcNAc:GlcNAc beta 1-3GalNAc-R (GlcNAc to GalNAc) beta 6-N-acetylglucosaminyltransferase from pig and rat colon mucosa. *Biochemistry* 24, 1866-1874

- Brun, R. and Kobel, H.R. (1972) Observations on the fertilization block between Xenopus borealis and Xenopus laevis laevis. J. Exp. Biol. 201, 135-138
- Butler, W.R. and Guthertz, L.S. (2001) Mycolic acid analysis by high performance liquid chromatography for identification of Mycobacterium species. *Clin. Microbiol. Rev.* 14, 704-726
- Capon, C., Leroy, Y., Wieruszeski, J.M., Ricart, G., Strecker, G., Montreuil, J. and Fournet, B. (1989) Structures of O-glycosidically linked oligosaccharides isolated from human meconium glycoproteins. *Eur. J. Biochem.* 182, 139-152
- Carlsson, H.E., Sundblad, G., Hammarstrom, S. and Lonngren. J. (1978) Structure of some oligosaccharides derived from rat-intestinal glycoproteins. *Carbohydr. Res.* 64, 181-188
- Carlstedt, I., Herrmann, A., Hovenberg, H., Lindell, G., Nordman, H., Wickström, C. and Davies, J.R. (1995) Soluble and insoluble mucins-Identification of distinct populations. In "Mucins: their structure and biology"
- Carlstedt, I. and Sheehan, J.K. (1984) Is the molecular architecture of cervical, respiratory and gastric mucins the same? *Biochem. Soc. Trans.* 12, 615-617
- Carraway, K.L., Rossi, E.A., Price-Sciavi, S.A., Huang, D., Guy, P.M., Carvajal, M.E., Fregien, N. and Carraway, C.A. (1999) An intramembrane modulator of the erbB2 receptor tyrosine kinase that potentiates neuregulin signaling. *J. Biol. Chem.* 274, 5263-5266
- Chan, J., Fujiwara, T., Brennan, P., McNeil, M., Turco, S.J., Sibille, J.C., Snapper, M., Aisen, P. and Bloom, B. (1989) Microbial glycolipids: possible virulence factors that scavenge oxygen radicals. *Proc. Natl. Acad. Sci. U.S.A.* 86, 2453-2457
- Chatterjee, D. (1998) The mycobacterial cell wall: structure, biosynthesis and sites of drug action. Curr. Opin. Chem. Biol., 579-588
- Chatterjee, D. and Khoo, K.H. (1998) Mycobacterial lipoarabinomannan: an extraordinary lipoheteroglycan with profound physiological effects. *Glycobiology* 8, 113-120
- Chatterjee, D., Bozic, C., McNeil, M. and Brennan, P. (1991) Structural features of the arabinan component of the lipoarabinomannan of *Mycobacterium tuberculosis. J. Biol. Chem.* 266, 9652-9660
- Chatterjee, D., Hunter, S.W., Mc Neil, M. and Brennan, P. (1992a) Lipoarabinomannan. Multiglycosylated form of the mycobacterial mannosylphosphatidylinositols. J. Biol. Chem. 267, 6228-6239
- Chatterjee, D., Khoo, K.H., McNeil, M., Dell, A., Morris, H.R. and Brennan, P.J. (1993) Structural definition of the nonreducing termini of mannose-capped LAM from *Mycobacterium tuberculosis* through selective enzymatic degradation and fast atom bombardment-mass spectrometry. *Glycobiology* 3, 497-506
- Chatterjee, D., Lowell, K., Rivoire, B., Mc Neil, M. and Brennan, P. (1992b) Lipoarabinomannan of *Mycobacterium tuberculosis*. Capping with mannosyl residues in some strains. J. Biol. Chem. 267, 6234-6239
- Chen, S., Spence, A.M. and Schachter, H. (2001) Phenotypes of Caenorhabditis elegans UDP-GlcNAc:α-3-D-mannoside β1,2-N-acetylglucosaminyltransferase I (GnT I) null mutants. Glycoconj. J. 17, 71
- Chen, S., Zhou, S., Sarkar, M., Spence, A.M. and Schachter, H. (1999) Expression of three *Caenorhabditis elegans* N-acetylglucosaminyltransferase I genes during development. J. Biol. Chem. 274, 288-297
- Clausen, H and Bennett, E.P. (1996) A family of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases control the initiation of mucin-type O-linked glycosylation. *Glycobiology* 6, 635-646
- Coetzee, T., Fujita, N., Dupree, J., SHI, R., Blight, A., Suzuki, K., Suzuki, K. and Popko, B. (1996) Myelination in the absence of galactocerebroside and sulfatide: normal structure with abnormal function and regional instability. *Cell* 86, 209-219
- Cooper, A.M., Magram, J., Ferrante, J. and Orme, I.M. (1997) Interleukin-12 (IL-12) is crucial to the development of protective immunity in mice intravenously infected with *Mycobacterium tuberculosis*. J. Exp. Med. 186, 39-45
- Costache, Apoil, P., Cailleau, A., Elmgren, A., Larson, G., Henry, S., Blancher, A., Iordachescu, D., Oriol, R. and Mollicone, R. (1997b) Evolution of fucosyltransferase genes in vertebrates. J. Biol. Chem. 272, 29721-29728
- Costache, M., Cailleau, A., Fernandez-Mateos, P., Oriol, R. and Mollicone, R. (1997a) Advances in molecular genetics of α-2 and α-3/4-fucosyltransferases. *Transfus. Clin. Biol.* 4, 367-382
- Couceiro, J.N., Paulson, J.C. and Baum, L.G. (1993) Influenza virus strain selectively recognize sialooligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. *Virus Res.* 29, 155-165

- Cox, J.S., Chen, B., McNeil, M. and Jacobs, W.R. (1999) Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* 402, 79-83
- Crick, D.C., Mahapatra, S. and Brennan, P.J. (2001) Biosynthesis of the arabinogalactan-peptidoglycan complex of *Mycobacterium tuberculosis*. *Glycobiology* 11, 107-118
- Daffé, M. and Draper, M. (1998) The envelope layer of mycobacteria with reference to their pathogenicity. Adv. Microb. Physiol. 39, 131-203
- Daffé, M. and Etienne, G. (1999) The capsule of mycobacterium tuberculosis and its implications for pathogenicity. *Tuber. Lung. Dis.* 79, 153-159
- Daffé, M., Brennan, P.J. and McNeil, M. (1990) Predominant structural features of the cell wall arabinogalactan of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by ¹H and ¹³C NMR analyses. J. Biol. Chem. 265, 6734-6743
- Daffé, M., Lacave, C., Lannéelle, M.A., Gillois, M. and Lanéelle, G. (1988) Polyphthienoyl tréhalose, glycolipids specific for the virulent strains of the tubercle bacillus. *Eur. J. Biochem.* 172, 579-584.
- Daffé, M., McNeil, M. and Brennan, P.J. (1991) Novel type-specific lipooligosaccharide from Mycobacterium tuberculosis. Biochemistry 30, 378-388
- Daffé, M., McNeil, M. and Brennan, P.J. (1993) Major structural features of the cell wall arabinogalactans of *Mycobacterium*, *Rhodococcus*, and *Nocardia* spp. *Carbohydr. Res.* 249, 383-398
- Dahl, K.E., Shiratsuchi, H., Hamilton, B.D., Ellner, J.J. and Toossi, Z. (1996) Selective induction of transforming growth factor b in human monocytes by lipoarabinimannan of *Mycobacterium* tuberculosis. Infect. Immun. 64, 399-405
- Dalziel, M., Whitehouse, C., McFarlane, I, Brockhausen, I., Gschmeissner, S., Schwientek, T., Clausen, H., Burchell, J.M. and Taylor-Papadimitriou, J. (2001) The relative activities of the C2GnT1 and ST3Gal-I glycosyltransferases determine O-glycan structure and expression of a tumor-associated epitope on MUC1. J. Biol. Chem. 276, 11007-11015
- Dammer, U., Popescu, O., Wagner, P., Anselmetti, D., Güntherodt, H.J. and Misevic, G.N. (1995) Binding strength between cell adhesion proteoglycans measured by atomic force microscopy. *Science* 267, 1173-1175
- D'arcy Hart, P.D. and Young, M.R. (1988) Polyanionic agents inhibit phagosome-lysosome fusion in cultured macrophages. J. Leukoc. Biol. 43, 179-182
- Dautzenberg, B. and Mercat, A. (1994) Mycobactérioses atypiques. Presse Med. 23, 1483-1488
- 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
- de Jong, R., Altare, F., Haagen, I.A., Elferink, D.G., Boer, T., van Breda Vriesman, P.J., Kabel, P.J., Draaisma, J.M., van Dissel, J.T., Kroon, F.P., Casanova, J.L. and Ottenhoff, T.H. (1998) Severe mycobacterial and *Salmonella* infections in interleukin-12 receptor deficient patients. *Science* 280, 1435-1438
- De, A., Brown, D.G, Gorman, M.A, Carr, M., Sanderson, M.R. and Freemont, P.S. (1994) Crystal structure of a disulfide-linked "trefoil" motif found in a large family of putative growth factors. *Proc. Natl. Acad. Sci. U.S.A.* 91, 1084-1088
- Dekker, J. and Strous, G.J. (1990) Covalent oligomerization of rat gastric mucin occurs in the rough endoplasmic reticulum, is N-glycosylation-dependent and precedes initial O-glycosylation. J. Biol. Chem. 265, 18116-18122
- Dekker, J., Rossen, J.W.A and Einerhand, A.W.C (2002) The MUC family: an obituary. TRENDS Bioch. Sci. 27,126-131
- Delmas, C., Gilleron, M., Brando, T., Vercellone, A., Rivière, M. and Puzo, G. (1997) Comparative structural study of mannosylated-lipoarabinomannans from *Mycobacterium bovis* BCG vaccine strains: characterization and localization of succinates. *Glycobiology* 7, 811-817
- Denis, M. and Gregg, E.O. (1991) Recombinant interleukin-6 increases the intracellular and extracellular growth of *Mycobacterium avium*. Can. J. Microbiol. 37, 479-483
- Dennis, J.W., Granovsky, M. and Warren, C.E. (1999) Protein glycosylation in development and disease. *BioEssays* 21, 412-421
- Desseyn, J.L., Guyonnet-Duperat, V., Porchet, N., Aubert, J.P. and Laine, A. (1997) Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternate subdomains resulting in a super-repeat. structural evidence for a 11p15.5 gene family. J. Biol. Chem. 272, 3168-3178
- Dhariwal, K.R., Dhariwal, G. and Goren, M.B. (1984) observation on the ubiquity of the *Mycobacterium tuberculosis* sulfatides in mycobacteria. *Am. Rev. Respir. Dis.* 130, 641-646
- Draper, P. (1971) The wall of Mycobacterium lepraemurium: chemistry and ultrastructure. J. Gen. Microbiol. 69, 313-324

- Draper, P., Khoo, K.H., Chatterjee, D., Dell, A. and Morris, H.R. (1997) Galactosamine in walls of slow-growing mycobacteria. *Biochem. J.* 327, 519-525
- Dumont, J.N., Brummet A.R. (1985) Egg envelop in vertebrates. Dev. Biol. 1, 235-288
- Dupree, J.L., Coetzee, T., Suzuki, K. and Popko, B. (1998) Myelin abnormalities in mice deficient in galactocerebroside and sulfatide. J. Neurocytol. 27, 649-659
- Dupuy, F., Germot. A., Marenda, M., Orilo, R., Blancher, A., Julien, R. and Maftah, A. (2002) Alpha1,4-fucosyltransferase activity: a significant function in the primate lineage has appeared twice independently. *Mol. Biol. Evol.* 19, 815-824
- Dupuy, F., Petit, J.M., Mollicone, R., Oriol, R., Julien, R. and Mafta, A. (1999) 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. 274, 12257-12262
- Eckhardt, A.E., Timpte, C.S., Deluca, A.W. and Hill, R.L. (1997) Complete cDNA sequence and structural polymorphism of the polypeptide chain of porcine submaxillary mucin. J. Biol. Chem. 272, 33204-33210
- Eggens, I., Fenderson, B., Toyokuni, T., Dean, B., Stroud, M. and Hakomori, S. (1989) Specific interarction between Le^x and Le^x determinants. J. Biol. Chem. 264, 9476-9484
- Elhammer, A.P., Poorman, R.A., Brown, E. Maggiora, M.M., Hoogerheide, J.G. and Kezdi, F.J. (1993) The specificity of UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase as inferred from the database of *in vivo* substrates and from *in vitro* glycosylation of proteins and peptides. *J. Biol. Chem.* 268, 10029-10038
- Endo, T., Nojima, S. and Inoue, K. (1982) Intermolecular interaction between glycolipids and glycophorin on liposomal membranes. J. Biochem. 92, 1883-1890
- Ernst, W.A., Maher, J., Cho, S., Niazi, K.R., Chatterjee, D., Moody, D.B., Besra, G.S., Watanabe, Y., Jensen, P.E., Porcelli, S.A., Kronenberg, M. and Modlin, R.L. (1998) Molecular interaction of CD1b with lipoglycan antigens. *Immunity* 8, 331-340
- Fenderson, B.A., Zehavi, U. and Hakomori, S. (1984) A multivalent lacto-N-fucopentaose IIIlysyllysine conjugate decompacts preimplantation mouse embryos, while the free oligosaccharide is ineffective. J. Exp. Med. 160, 1591-1596
- Ferguson, J.S., Voelker, D.R., McCormack, F.X. and Schlesinger, L.S. (1999) Surfactant protein D binds to *Mycobacterium tuberculosis* bacilli and lipoarabinomannan via carbohydrate-lectin interactions resulting in reduced phagocytosis of the bacteria by macrophages. J. Immunol. 163, 312-321
- Fernandez-Busquets, X. and Burger, M.M. (1997) The main protein of the aggregation factor responsible for species-specific cell adhesion in the marine sponge *Microconia prolifera* is highly polymorphic. J. Biol. Chem. 272, 27839-27847
- Fernandez-Busquets, X. and Burger, M.M. (1999) Cell adhesion and histocompatibility in sponges. Microscopy research and technique 44, 204-218
- Fontaine, M.D., Wieruszeski, J.M., Plancke, Y., Delplace, F. and Strecker, G. (1995) Structure of six 2deoxy-D-Glycero-D-galacto-nonulosonic acid (Kdn)-containing oligosccharide-alditols released from oviducal secretion of Ambystoma maculatum. Eur. J. Biochem. 231, 424-433
- Fournié, J.J., Rivière, M. and Puzo, G. (1987) Structural elucidation of the major phenolic glycolipid from *Mycobacterium kansasii*.I. Evidence for tetrasaccharide structure of the oligosaccharidic moiety. J. Biol. Chem. 262, 3174-3179
- Fournié, J.J., Rivière, M., Papa, F. and Puzo, G. (1987) Structural elucidation of the major phenolic glycolipid from *Mycobacterium kansasii*. II. Presence of a novel dideoxyhexose. J. Biol. Chem. 262, 3180-3184
- Galtsoff, P.S. (1925) Regeneration after dissociation (an experimental study on sponges). I Behaviour of dissociated cells of *Microciona prolifera* under normal and altered conditions. J. Exp. Zool. 42, 183-221
- Gaynor, C.D., McCormack, F.X., Voelker, D.R., McGowan, S.E. and Schlesinger, L.S. (1995) Pulmonary surfactant prtein A mediates enhanced phagocytosis of *Mycobacterium tuberculosis* by a direct interaction with human macrophages. J. Immunol. 155, 5343-5351
- Gbery, I.P., Djeha, D., Yobouet, P., Aka, B. and Kanga, J.M. (1996) Atypical mycobacterial skin infections. Santé 6, 317-322
- Gems, D. and Maizels, R.M.(1996) An abundantly expressed mucin-like protein from *Toxocara canis* infective larvae: the precursor of the larval surface coat glycoproteins. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1665-1670
- Gendler, S., Taylor-Papadimitriou, J., Duhig, T., Rothbard, J. and Burchell, J. (1988) A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up tandem repeats. J. Biol. Chem. 263, 12820-12823
- Gendler, S.J. and Spicer, A.P. (1995) Epithelial mucin genes. Annu. Rev. Physiol. 57, 607-634

- Gerdt, S., Dennis, R.D., Borgonie, G., Schnabel, R. and Geyer, R. (1999) Isolation, characterization and immunolocalization of phosphorylcholine-substituted glycolipids in developmental stages of *Caenorhabditis elegans. Eur. J. Biochem.* 266, 952-963
- Gerdt, S., Lochnit, G., Dennis, R.D. and Geyer (1997) Isolation and structural analysis of three neutral glycosphingolipids from a mixed population of *Caenorhabditis elegans* (Nematoda:Rhabditida). *Glycobiology* 7, 265-275
- Gerken, T.A., Gilmore, M. and Zhang, J. (2002) Determination of the site-specific oligosaccharide distribution of the *O*-glycans attached to the procine submaxillary mucin tandem repeat. J. Biol. Chem. 277, 7736-7751
- Gerken, T.A., Owens, C.L. and Pasumarthy, M. (1997) Determination of the site-specific Oglycosylation pattern of porcine submaxillary mucin tandem repeat glycopeptide. Model proposed for the polypeptide:GalNAc transferase peptide binding site. J. Biol. Chem. 172, 9709-9719
- Gerton, G.L. and Hedrick, J.L. (1986) The vitelline envelope to fertilization envelope conversion in egg of *Xenopus laevis*. *Dev. Biol.* 11, 1-7.
- Geyer, A., Gege, C. and Schmitd (2000) Calcium-dependent carbohydrate-carbohydrate recognition between Lewis^x blood group antigens. *Angew. Chem. Int. Ed.* 39, 3246-3249
- Gilleron, M. and Puzo, G. (1995) Lipooligosaccharidic antigens from Mycobacterium kansasii and Mycobacterium gastri. Glycoconj. J. 12, 298-308
- Gilleron, M., Bala, L., Brando, T., Vercellone, A. & Puzo, G. (2000). Mycobacterium tuberculosis parietal and cellular lipoarabinomannans. J. Biol. Chem. 275, 677-684.
- Gilleron, M., Himoudi, N., Adam, O., Constant, P., Venisse, A., Rivière, M. & Puzo, G. (1997). *Mycobacterium smegmatis* phosphoinositols-glyceroarabinomannans. Structure and localization of alkali-labile and alkali-stable phosphoinositides. J. Biol. Chem. 272, 117-124.
- Gilleron, M., Nigou, J., Cahuzac, B. & Puzo, G. (1999). Structural study of the lipomannans from *Mycobacterium bovis* BCG: characterisation of the multiacylated forms of the phosphatidyl-*myo*-inositol anchor. J. Mol. Biol. 285, 2147-2160.
- Gilleron, M., Venisse, A., Rivière, M., Servin, P. and Puzo, G. (1990) Carbohydrate epitope structural elucidation by 1H-NMR spectroscopy of a new *Mycobacterium kansasii* phenolic glycolipid antigen. *Eur. J. Biochem.* 193, 449-457
- Gilleron, M., Vercauteren, J. and Puzo, G (1994) Lipo-oligosaccharidic antigen from *Mycobacterium* gastri. Complete structure of a novel C4-branched 3,6-dideoxy-alpha-xylo-hexopyranose. *Biochemistry* 33, 1930-1937
- Gold, D.V. and Shochat, D. (1989) Studies on the structure of the organ-specific determinant of human colonic mucin. *Mol. Immunol.* 26, 769-77
- Gooi, H.C., Feizi, T., Kapadia, A., Knowles, B.B., Solter, D. and Evans, M.J. (1981) Stage-specific embryonic antigen involves alpha 1 goes to 3 fucosylated type 2 blood group chains. *Nature* 292, 156-158
- Goren, M.B. (1970) Sulfolipid I of Mycobacterium tuberculosis strain H37Rv. I. Purification and properties. Biochim. Biophys. Acta. 210, 116-126
- Goren, M.B. (1982) Immunoreactive substances of mycobacteria. Am. Rev. Respir. Dis. 125, 50-69
- Goren, M.B., Brokl, O., Roller, P., Fales, H.M. and Das, B.C. (1976a) Sulfatides of Mycobacterium tuberculosis. The structure of the principle sulfatide (SL-I). Biochemistry 15, 2728-2735
- Goren, M.B., D'arcy Hart, P.D., Young, M.R. and Armstrong J.A. (1976b) Prevention of phagosomelysosome fusion in cultured macrophages by sulfatides of *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U.S.A. 73, 2510-2514
- Goren, M.B., Vatter, A.E. and Fiscus, J. (1987) Polyanionic agents do not inhibit phagosome-lysosome fusion in cultured macrophages. J. Leukoc. Biol. 41, 122-129
- Granovsky, M., Bielfeldt, T., Peters, S., Paulsen, H., Meldal, M., Brockhausen, I. and Brockhausen, J. (1994) UDPgalactose:glycoprotein-N-acetyl-D-galactosamine 3-b-D-galactosyltransferase activity synthesizing O-glycan core 1 is controlled by the amino acid sequence and glycosylation of glycopeptide substrates. *Eur. J. Biochem.* 221, 1039-1046
- Grey, R.D., Bastiani, M.J., Webb, D.J. and Shertel, E.R. (1982) An electric block is required to prevent polyspermy in eggs fertilized by natural mating of *Xenopus laevis*. Dev. Biol., 475-484.
- Grey, R.D., Working, P.K. and Hedrick, J.L. (1976) Evidence that the fertilization envelope blocks sperm entry in eggs of *Xenopus laevis*. Dev. Biol. 54: 52-60.
- Gum, J.R., Crawley, S.C., Hicks, J.W., Szymkowski, D.E. and Kim, Y.S. (2002) MUC17, a novel membrane-tethered mucin. *Biochem. Biophys. Res. Commun.* 291, 466-475

- Gum, J.R., Hicks, J.W., Toribara, N.W., Siddiki, B. and Kim, Y.S. (1994) Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of the amino terminus and overall sequence similarity to prepro-von Willebrand factor. J. Biol. Chem. 269, 2440-2446
- Hagen, F.K. and Nehrke, K. (1998) cDNA cloning and expression of a family of UDP-N-acetyl-Dgalactosamine:polypeptide N-acetylgalactosaminyltransferase sequence homologs from *Caenorhabditis* elegans. J. Biol. Chem. 273, 8268-8277
- Hagen, K.G., Bedi, G.S., Tetaert, D., Kingsley, P.D., Hagen, F.K., Balys, M.M., Beres, T.M., Degand, P. and Tabak, L.A. (2001) Cloning and characterization of a ninth member of the UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferase family, ppGaNTase-T9. J. Biol. Chem. 276, 17395-17404
- 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. (1999) Characterization of a UDP-GalNAc:Polypeptide *N*-Acetylgalactosaminyltransferase that displays glycopeptide *N*-Acetylgalactosaminyltransferase activity. J. Biol. Chem. 274, 27867-27874
- Hakamori, S. (2001) Cell adhesion/recognition and signal transduction through glycosphingolipid microdomain. *Glycoconj. J.* 17, 143-151
- Hakamori, S.I., Clausen, H. and Levery, S.B. (1987) A new series of blood group A and H antigens expressed in human erythrocytes and the incompatible A antigens expressed in tumours of blood group O and B individuals. *Biochem. Soc. Trans.* 15, 593-596
- Hakomori, S., Igarashi, Y., Kojima, N., Okoshi, H., Handa, K. and Fenderson, B. (1991) Functional role of cell surface carbohydrates into ontogenesis and oncogenesis. *Glycoconj. J.* 8, 178
- Hakomori, S., Nudelman, E., Levery, S., Solter, D. and Knowles, B.B. (1981) The hapten structure of a developmentally regulated glycolipid antigen (SSEA-1) isolated from human erythrocytes and adenocarcinoma: a preliminary note. *Biochem. Biophys. Res. Commun.* 100, 1578-1586
- Haltiwanger, R.S., Blomberg, M.A. and Hart, G.W. (1992) Glycosylation of nuclear and cytoplasmic proteins. Purification and characterisation of an uridine diphospho-N-acetylglucosamine: polypeptide beta-N-acetylglucosaminyltransferase. J. Biol. Chem. 267, 9005-9013
- Hanisch, F.G., Chai, W., Rosankiewicz, J.R., Lawson, A.M., Stoll, M.S. and Feizi, T. (1993) Coretyping of O-linked glycans from human gastric mucins. Lack of evidence for the occurrence of the core sequence Gal1-6GalNAc. *Eur. J. Biochem.* 217, 645-655
- Harris, R.J. & Spellman, M.W. (1993) O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology* 3, 219-224
- Haseley, R.S., Vermeer, H.J., Kameling, J.P. and Vliegenthart, J.F.G. (2001) Carbohydrate selfrecognition mediates marine sponge cellular adhesion. *Proc. Natl. Acad. Sci. USA*. 98, 9419-9424
- Hassan, H., Reis, C.A., Bennett, E.P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M.A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (2000) The lectin domain of UDP-N-acetyl-Dgalactosamine:polypeptide N-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities. J. Biol. Chem. 275, 38197-38205
- Hedrick, J.L. and Nishihara, T. (1991) Structure and fonction of the extracelleular matrix of anuran eggs. J. Electron. Microsc. Tech. 17, 319-335
- Herrmann, A., Davies, J.R., Lindéll, G., Martensson, S., Packer, N.H., Swallow, D.M. and Carlstedt, I. (1999) Studies on the insoluble glycoprotein complex from human colon. Identification of reductioninsensitive MUC2 oligomers and C-terminal cleavage. J. Biol. Chem. 274, 15828-15836
- Hill, D.L. and Ballou, C.E. (1966) Biosynthesis of mannophospholipids by Mycobacterium phlei. J. Biol. Chem. 241, 895-902
- Hirsch, C.S., Hussain, R., Toossi, Z. and Ellner, J.J. (1996) Cross-modulatory role for transforming growth factor beta in tuberculosis; suppression of antigen driven interferon gamma production. *Proc. Natl. Acad. Sci. USA* 93, 3193-3198
- Hirsch, C.S., Yoneda, T., Averill, L.E., Ellner, J.J. and Toossi, Z. (1994) Enhancement of intracellular growth of *Mycobacterium tuberculosis* in human monocytes by transforming growth factor β1. J. Infect. Dis. 170, 1229-1237
- Horsburgh, C.R (1996) Epidemiology of disease caused by nontuberculous mycobacteria. Seminars in respiratory infections 11, 244-251
- Houdret, N., Perini, J.M., Galabert, C., Scharfman, A., Humbert, C., Lamblin, G. and Roussel, P. (1986) The high lipid content of respiratory mucins in cystic fibrosis is related to infection. *Biochem. Biophys. Acta.* 880, 54-61
- Hounsell, E.F. and Feizi, T. (1982) Gaxtrointestinal mucins. Structures and antigenicities of their carbohydrate chains in health and disease. *Med. Biol.* 60, 227-236

- Hounsell, E.F., Lawson, A.M., Feeney, J., Gooi, H.C., Pickering, N.J., Stoll, M.S., Lui, S.C. and Feizi, T. (1985) Structural analysis of the O-glycosidically linked core-region oligosaccharides of human meconium glycoproteins which express oncofoetal antigens. *Eur. J. Biochem.* 148, 367-377
- Humphreys, S., Humphreys, T. and Sano, J. (1977) Organisation and polysaccharides of sponge aggregation factor. J. Supramol. Struct. 7, 339-351
- Humphreys, T. (1963) Chemical dissolution and in vitro reconstruction of sponge cells adhesion. I Isolation and functional demonstration of the components involved. *Dev. Biol.* 8, 27-47
- Hunter, S. W. and Brennan, P. J. (1990). Evidence for the presence of a phosphatidylinositol anchor on the lipoarabinomannan and lipomannan of *Mycobacterium tuberculosis. J. Biol. Chem.* 265, 9272-9279.
- Hunter, S.W. and Brennan, P.J. (1981) A novel phenolic glycolipid from *Mycobacterium leprae* possibly involved in immunogenicity and pathogenicity. J. Bacteriol. 147, 728-735
- Hunter, S.W., Barr, L.V., McNeil, M., Jardine, I. and Brennan, P.J. (1988) Trehalose-containing lipooligosaccharide antigens of *Mycobacterium sp.*: presence of a mono-O-methyltri-O-acyltrehalose "core". *Biochemistry* 27, 1549-1556
- Hunter, S.W., Fujiwara, T. and Brennan, P.J. (1982) Structure and antigenicity of the major specific glycolipid antigen of *Mycobacterium leprae. J. Biol. Chem.* 257, 15072-15078
- Hunter, S.W., Fujiwara, T., Murphy, R.C. and Brennan, P.J. (1984) N-acylkansosamine. A novel acylamino sugar from the trehalose-containing lipooligosaccharide antigens of *Mycobacterium kansasii*. J. Biol. Chem. 259, 9729-9734
- Hunter, S.W., Gaylord, H. and Brennan, P.J. (1986) Structure and antigenicity of the phosphorylated lipopolysaccharide antigens from the leprosy and tubercle bacilli. J. Biol. Chem. 261, 12345-12351
- Iida, S., Takeuchi, H., Helle, H., Clausen, H. and Irimura, T. (1999) Incorporation of *N*-acetylgalactosamine into consecutive threonine residues in MUC2 tandem repeat by recombinant human N-acetyl-D-galactosamine transferase-T1, T2 and T3. *FEBS lett.* 449, 230-234
- Inoue, S. and Inoue, Y. (1997) Fish glycoproteins. *In Glycoprotein II*. J. Montreuil, J.F.G. Vliegenthart and H. Schachter (Eds), Elsevier Science B.V., 143-161
- Iwabuchi, K., Yamamura, S., Prinetti, A., Handa, K. and Hakomori, S. (1998) GM3-enriched microdomain involved in cell adhesion and signal transducing through carbohydrate-carbohydrate interaction in mouse melanoma B16 cells. J. Biol. Chem. 273, 9130-9238
- Iwasaki, M., Inoue, S. and Troy, F.A. (1990) A new sialic acid analogue, 9-O-acetyl-deaminated neuraminic acid, and alpha-2,8-linked O-acetylated poly (N-glycolylneuraminyl) chains in a novel polysialoglycoprotein from salmon eggs. J. Biol. Chem. 265, 2596-2602
- Jackson, M., Raynaud, C., Lanéelle, M.A., Guilhot, C. Laurent-Winter, C., Ensergueix, D., Gicquel, B. and Daffé, M. (1999) Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the *Mycobacterium tuberculosis* cell envelope. *Mol. Microbiol.* 31, 1573-1587
- Jarchow, J., Fritz, J., Anselmetti, D., Calabro, A., Hascall, V.C., Gerosa, D., Burger, M.M. and Fernandez-Busquets, X. (2001) Suprmolecular structure of a new family of circular proteoglycans mediating cell adhesion in sponges. J. Struct. Biol. 132, 95-105
- Jarchow, J., Fritz, J., Anselmetti, D., Calabro, A., Hascall, V.C., Gerosa, D., Burger, M.M. and fernadez-Busquets, X. (2000) Supramolecular structure of a new family of circular proteoglycans mediating cell adhesion in sponges. J. Struct. Biol. 132, 95-105
- Jentoft, N. (1990) Why are proteins O-glycosylated. TRENDS Bioch. Sci. 15, 291-294
- Joba, W. and Hoffmann, W. (1997) Similarities of integumentary mucin B.1 from *Xenopus laevis* and prepro-von Willebrand factor at their amino-terminal regions. J. Biol. Chem. 272, 1805-1810
- Jones, S.J. and Baillie, D.L. (1995) Characterization of the *let-653* gene in *Caenorhabditis elegans*. *Mol. Gen. Genet.* 248, 719-726.
- Jumblatt, J.E., Schlup, V. and Burger, M. (1980) Cell-cell recognition: specific binding of *Microconia* sponge aggregation factor to homotypique cells and the role of calcium ions. *Biochemistry* 19, 1038-1042
- Kasturi, L., Eshleman, J.R., Wunner, W.H. and Shakin-Eshleman, S.H. (1995) The hydroxy amino acid in an Asn-X-Ser/Thr sequon can influence N-linked core glycosylation efficiency and the level of expression of a cell surface glycoprotein. J. Biol. Chem. 270, 14756-14761
- Katagiri, C. (1987) role of oviducal secretions in mediating gamete fusion in anuran amphibians. Zool. Scien. 4, 1-14.
- Kawar, Z.S., Van Die, I. and Cummings, R.D. (2002) Molecular cloning and enzymatic characterization of a UDPGalNAc:GlcNAcβ-R β1,4-N-acetylgalactosaminyltransferase from C. elegans. J. Biol. Chem. sous presse.
- Kenatsuna, F. (1968) Chemical analysis of mycobacterial cell-wall. *Biochim. Biophys. Acta.* 158, 130-143

- Khoo, K.H., Dell, A., Morris, H.R., Brennan, P.J. and Chatterjee, D. (1995a) Structural definition of acylated phosphatidyl inositol mannosides from *Mycobacterium tuberculosis*: definition of a common anchor for lipomannan and lipoarabinomannan. *Glycobiology* 5, 117-127
- Khoo, K.H., Dell, A., Morris, H.R., Brennan, P.J. and Chatterjee, D. (1995b) Inositol phosphate capping of the nonreducing termini of lipoarabinomannan from rapidly growing strain of Mycobacterium. J. Biol. Chem. 270, 12380-12389
- Khoo, K.H., Douglas, E., Azadi, P., Inamine, J.M., Besra, G.S., Mikusova, K., Brennan, P. and Chatterjee, D. (1996) Truncated structural variants of lipoarabinomannan in ethambutol drug-resistant strains of *Mycobacterium smegmatis. J. Biol. Chem.* 271, 28682-28690
- Khoo, K.H., Suzuki, R., Dell, A., Morris, H.R., McNeil, M.R., Brennan, P.J. and Besra, G.S. (1996) Chemistry of the lyxose-containing mycobacteriophage receptors of *Mycobacterium phlei/ Mycobacterium smegmatis. Biochemistry* 35, 11812-11819
- Khoo, K.H., Tang, J.B. and Chatterjee, D. (2001) Variation in mannose-capped terminal arabinan motifs of lipoarabinomannans from clinical isolates of *Mycobacterium tuberculosis* and *Mycobacterium avium* complex. J. Biol. Chem. 276, 3863-3871
- Khuller, G.K. and Subrahmanyam, D. (1968) On the mannophosphoinositides of *Mycobacterium* 607. *Experentia* 24, 851-852
- Kindler, V., Sappino, A.P., Grau, G.E., Piguet, P.F. and Vassali, P. (1989) The inducing role of tumor necrosis factor in the development of bactericidal granulomas during BCG infection. *Cell* 56, 731-740
- Klein, A., Lamblin, G., Lhermitte, M., Roussel, P., Breg, J., van Halbeek, H. and Vliegenthart, J.F.G. (1988) Primary structure of neutral oligosaccharides drived from respiratory-mucus glycoproteins of a patient suffering from bronchiectasis, determinad by comparison of 500-MHz ¹H-NMR spectroscopy and quantitative sugar analysis. *Eur. J. Biochem.* 171, 631-642
- Knutson, K.L., Hmama, Z., Herrera-Velit, P., Rochford, R. and Reiner, N.E. (1998) Lipoarabinomannan of *Mycobacterium tuberculosis* promotes protein tyrosine dephosphorylation and inhibition of mitogen-activated protein kinase in human mononuclear phagocytes. J. Biol. Chem. 273, 645-652
- Kojima, N. and Hakomori, S. (1989) Specific interaction between gangliotriasylceramide (Gg₃) and sialosyllactoceramide (GM₃) as a basis for specific cellular recognition between lymphoma and melanoma cells. J. Biol. Chem. 264, 20159-20162
- Kojima, N. and Hakomori, S. (1991a) Cell adhesion, spreading and motility of GM₃-expressing cells based on glycolipid-glycolipid interaction. J. Biol. Chem. 266, 17552-17558
- Kojima, N. and Hakomori, S. (1991b) Synergetic effect of two cell recognition systems: glycosphingolipid-glycosphingolipid interaction and integrin receptor interaction with pericellular matrix protein. *Glycobiology* 1, 623-630
- Kojima, N., Fenderson, B.A., Stroud, M.R., Goldberg, R.I., Habermann, R., Toyokuni, T. and Hakomori, S. (1994) 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.* 11, 238-248
- Kojima, N., Shiota, M., Sadahira, Y., Handa, K. and Hakomori, S. (1992) Cell adhesion in a dynamic flow system as compared to static system. J. Biol. Chem. 267, 17264-17270
- Kordulakova, J., Gilleron, M, Mikusova, K., Puzo, G., Brennan, P.J., Gicquel, B. and Jackson, M. (2002) Definition of the first mannosylation step in phosphatidylinositol mannoside synthesis. PIMA is essential for growth of mycobacteria. J. Biol. Chem. 277, 31335-31344
- Kremer, L. and Besra, G. (2002) Current status and future development of antitubercular chemotherapy. *Expert. Opin. Investig. Drugs.* 11, 1-17
- Kremer, L. Gurcha, S.S., Bifani, P., Hitchen, P.G., Baulard, A., Morris, H.R., Dell, A., Brennan, P.J. and Besra, G.S. (2002) Characterization of a putative a-mannosyltransferase involved in phosphatudylinositol trimannoside biosynthesis in *Mycobacterium tuberculosis*. *Biochem. J.* 363, 437-447
- Kremer, L., Besra, G.S., Brennan, P.J. and Baulard, A.R. (1999) Le lipoarabinomannane: structures et fonctions d'un glycolipide impliqué dans la pathogénie tuberculeuse. *Médecine/Science* 15, 842-850
- Kubelka, V., Altmann, F., Staudacher, E., Tretter, V., Marz, L., Hard, K., Kamerling, J.P. and Vliegenthart, J.F. (1993) Primary structure of the N-linked carbohydrate chains from honeybee venom phospholipase A2. *Eur. J. Biochem.* 213, 1193-1204
- Lamblin, G., Lhermitte, M., Klein, A., Houdret, N., Scharfman, A., Ramphal, R. and Roussel, P. (1991) The carbohydrate diversity of human respiratory mucins: a protection of the underlying mucosa? *Am. Rev. Respir. Dis.* 144, S19-S24

- Lee, Y.C. and Ballou, C.E. (1964) Structural studies on the myo-inositol mannosides from the glycolipidsof *Mycobacterium tuberculosis* and *Mycobacterium phlei. J. Biol. Chem.* 239; 1316-1327
- Lee, Y.C. and Ballou, C.E. (1965) Complete structures of the glycophospholipids of mycobacteria. *Biochemistry* 4, 1395-1404
- Lemassu, A. and Daffé, M. (1994) Structural features of the exocellular polysaccharides of Mycobacterium tuberculosis. Biochem. J. 297, 351-357
- Leopold, K. and Fischer, W. (1993) Molecular analysis of the lipoglycans of Mycobacterium tuberculosis. Anal. Biochem. 208, 57-64
- Li, Y., Bharti, A., Chen, D., Gong, J. and Kufe, D. (1998) Interaction of glycogen synthase kinase 3β with the DF3/MUC1 carcinoma-associated antigen and β-catenin. *Mol. Cell. Biol.* 18, 7216-7224
- Lindsay, L.L. and Hedrick, J.L. (1989) Proteases released from *Xenopus laevis* eggs at activation and their role in envelope conversion. *Dev. Biol.* 131, 126-135.
- Lindsay, L.L., Wieduwilt, M.J. and Hedrick J.L. (1999) Oviductin, the *Xenopus laevis* oviducal protease that processes egg envelope glycoproteine gp43, increases sperm binding to envelopes, and is translated as part of an unusual mosaic protein composed of two protease and several CUB domains. *Biol. Reprod.* 60, 989-995.
- Lloyd, K.O., Burchell, J., Kudryashov, V., Yin, B.W.T. and Taylor-Papadimitriou, J. (1996) Comparison of O-linked carbohydrate chains in MUC-1 mucin from normal breast epithelial cell lines and breast carcinoma cell lines. Demonstration of simpler and fewer glycan chains in tumor cells. J. Biol. Chem. 271, 33325-33334
- Lochnit, G., Dennis, R.D., Zähringer, U. and Geyer, R. (1997) Structural analysis of neutral sphingolipids from *Ascaris suum* adults (Nematoda: Ascaridida) *Glycoconjugate J.* 14, 389-399
- Lo-Guidice, J.M., Wieruszeski, J.M., Lemoine, J., Verbert, A., Roussel, P. and Lamblin, G. (1994) Sialylation and sulfation of the carbohydrate chains in respiratory mucins from a patient with cystic fibrosis. J. Biol. Chem. 269, 18794-18813
- Lopez, L.M., Lanéelle, M.A., Promé, D., Daffé, M., Lanéelle, G. and Promé, J.C. (1991) Glycopeptidolipids from *Mycobacterium fortuitum*: a variant in the structure of C-mycoside. *Biochemistry* 30, 10536-10542
- Lopez, L.M., Lanéelle, M.A., Promé, D., Lanéelle, G., Promé, J.C. and Daffé, M (1991) Structure of a novel sulfate-containing mycobacterial glycolipid. *Biochemistry* 31, 11106-11111
- Loukas, A., Hintz, M., Linder, D., Mullin, N.P., Parkinson, J., Tetteh, K.K. and Maizels, R.M. (2000) A family of secreted mucins from the parasitic nematode *Toxocara canis* bears divers mucin domains but shares similar flanking six-cysteine repeat motifs. *J. Biol. Chem.* 275, 50
- Luquin, M., Papa, F. and David, H.L. (1992) Identification of sulfolipid I by thin-layer chromatography in the rapid identification of *Mycobacterium tuberculosis*. *Res. Microbiol.* 143, 225-227
- Maes, E., Wieruszeski, J.M., Plancke, Y. and Strecker, G. (1995) Structure of three Kdn-containing oligosaccharide-alditols released from oviducal secretions of *Ambystoma tigrinum*. FEBS lett. 358, 205-210
- Maizels, R.M., de Savigny, D. and Ogilvie, B.M. (1984) Characterisation of surface and excretorysecretory antigens of *Toxocara canis* infective larvae. *Parasite Immunol.* 6, 23-37
- Marshall, R.D. (1972) Glycoproteins. Ann. Rev. Biochem. 41, 673-702
- Martensson, .S, Levery, S.B., Fang, T.T. and Bendiak, B. (1998) Neutral core oligosaccharides of bovine submaxillary mucin. Use of lead tetraacetate in the cold for establishing branch positions. *Eur. J. Biochem.* 258, 603-622
- Martensson, S., Levery, S.B., Fang, T.T. and Bendiak, B. (1998) Neutral core oligosaccharides of bovine submaxillary mucin. *Eur. J. Biochem.* 258, 603-622
- McNeil, M., Daffé, M. and Brennan, P.J. (1990) Evidence for the nature of the link between arabinogalactan and peptidoglycan of mycobacterial cell walls. J. Biol. Chem. 265, 18200-18206
- McNeil, M., Daffé, M. and Brennan, P.J. (1991) Location of the mycolyl ester substituents in the cell walls of mycobacteria. J. Biol. Chem. 266, 13217-13223
- McNeil, M., Wallner, S.J., Hunter, S.W. and Brennan, P. (1987) Demonstration that the galactosyl and arabinosyl residues in the cell-wall arabinogalactan of *Mycobacterium leprae* and *Mycobacterium tuberculosis* are furanoid. *Carbohydr. Res.* 166, 299-308
- Means, T.K., Lien, E., Yoshimura, A., Wang, S., Golenbock, D.T. and Fenton, M. (1999b) The CD14 ligands lipoarabinomannan and lipopolysaccharide differ in their requirement for Toll-like receptor. J. Immunol. 163, 6748-6755
- Means, T.K., Wang, S., Lien, E., Yoshimura, A., Golenbock, D.T. and Fenton, M. (1999a) Human Tolllike receptor mediate cellular activation by Mycobacterium tuberculosis. *J. Immunol.* 163, 3920-3927

- Mehra, V., Brennan, P.J., Rada, E., Convit, J. and Bloom, B. (1984) Lymphocyte suppression in leprosy induced by unique *M. leprae* glycolipid. Nature 308, 194-196
- Minnikin, D.E., Dobson, G., Sesardie, D. and Ridell, M. (1985) Mycolipenates and mycolipanolates of threhalose from *M. tuberculosis. J. Gen. Microbiol.* 131, 1369-1374
- Minnikin, D.E., Kremer, L., Dover, L.G. and Besra, G.S. (2002) The methyl-branched fortification of Mycobacterium tuberculosis. *Chemistry & Biology* 9, 545-553
- Misaki, A. and Yukawa, S. (1966) Studies on cell walls of mycobacteria. II. Constitution of polysaccharides from BCG cell walls. J. Biochem. 59, 511-520
- Misaki, A., Azuma, I. and Yamamura, Y. (1977) Structural and immunochemical studies on D-arabino-D-mannan of mycobacterium tuberculosis and other mycobacterium species. J. Biochem (Tokyo) 82, 1759-1770
- Misevic, G. and Burger, M. (1990) the species-specifique cell-binding site of the aggregation factor from the sponge Microciona prolifera is a highly repetitive novel glycan containing glucuronic acid, fucose and mannose. J. Biol. Chem. 265, 20577-20584
- Misevic, G. and Burger, M. (1993) Carbohydrate-carbohydrate interactions of a novel acidic glycan can mediate sponge cell adhesion. J. Biol. Chem. 268, 4922-4929
- Misevic, G., Finne, J. and Burger, M. (1987) Involvement of carbohydrate as multiple low affinity interaction sites in the self-association of the aggregation factor from the marine sponge *Microciona* prolifera. J. Biol. Chem. 262, 5870-5877
- Misevic, G.N., Jumblatt, J.E. and Burger, M.M. (1982) Cell binding fragments from a sponge proteoglycan-like aggregation factor. J. Biol. Chem. 257, 6931-6936
- Misevic, M. and Burger, M. (1986) reconstitution of high cell binding affinity of a marine sponge aggregation factor by cross-linking of small low affinity fragments into a large polyvalent polymer. J. Biol. Chem. 261,2853-2859
- Moniaux, N., Escande, F., Porchet, N., Aubert, J.P. and Batra, S.K. (2001) Structural organization and classification of the human mucin genes. *Front. Biosci.* 6, 1192-1206
- Monsarrat, B., Brando, T., Condouret, P., Nigou, J. and Puzo, G. (1999) Characterization of mannooligosaccharide caps in in mycobacterial lipoarabinomannan by capillary electrophoresis/electrospray mass spectrometry. *Glycobiology* 9, 335-342
- Morelle, W. and Strecker, G. (1997) Structural analysis of oligosaccharide-alditols released by reductive β-elemination from oviducal mucins of *Bufo Bufo*: characterisation of the carbohydrate sequence Gal(α1-3)GalNAc(α1-3)[Fuc(α1-2)]Gal. *Glycobiology* 7, 1129-1151
- Morelle, W., Cabada, M.O. and Strecker, G. (1998) Structural analysis of oligosaccharide-alditols released by reductive β-elimination from the jelly-coats of the anuran *Bufo arenarum*. Eur. J. Biochem. 252, 253-260
- Moreno, C., Taverne, J., Mehlert, A., Bate, C.A., Brealey, R.J., Meager, A., Rook, G.A. and Playfair, J.H. (1989) Lipoarabinomannan from *Mycobacterium tuberculosis* induces the production of tumour necrosis factor from human and murine macrophage. *Clin. Exp. Immunol.* 76, 240-245
- Mozingo, N.M. and Hedrick, J.L. (1999) Distribution of lectin binding sites in *Xenopus laevis* egg jelly. *Dev. Biol.* 210, 428-439
- Muchmore, E.A., Varki, N.M., Fukuda, M. and Varki, A. (1987) Developmental regulation of sialic acid modifications in rat and human colon. *FASEB J.* 1, 229-235
- Müller, S., Alving, K., Peter-Katalinic, J., Zachara, N., Gooley, A.A. and Hanisch, F.G. (1999) High density O-glycosylation on tandem repeat peptide from secretory MUC1 of T47D breast cancer cells. J. Biol. Chem. 274, 18165-18172
- Müller, S., Goletz, S., Packer, N., Gooley, A.A., Lawson, A.M. and Hanisch, F.G. (1997) Localisation of O-glycosylation sites on glycopeptides fragments from the lactation-associated MUC1. J. Biol. Chem. 272, 24780-24793
- Müller, W.E.G and Zahn, R.K. (1973) Purification and characterisation of a species-specific aggregation factor in sponges. *Exp. Cell. Res.* 80, 95-104
- Muñoz, M., Lanéelle, M.A., Luquin, M., Torrelles, J., Julian, E., Ausina, V. and Daffé, M. (1997) Occurence of an antigenic triacyl trehalose in clinical isolates and reference strains of Mycobacterium tuberculosis. *FEMS Microbiol. lett.* 157, 251-259
- Mutsaers, J.H., van Halbeek, H., Vliegenthart, J.F., Wu, A.M. and Kabat, E.A. (1986) Typing of core and backbone domains of mucin-type oligosaccharides from human ovarian-cyst glycoproteins by 500-MHz 1H-NMR spectroscopy. *Eur. J. Biochem.* 157, 139-146
- Nath, P., Hartnell, A., Happerfield, L., Miles, D.W., Burchell, J., Taylor-Papadimitriou, J. and Crocker, P.R. (1999) Macrophage-tumour cell interaction: identification of MUC1 on breast cancer cells as

potential counter-receptor for the macrophage-restricted receptor, sialoadhesin. *Immunology* 98, 213-219

- Natsuka, S., Adachi, J., Kawaguchi, M., Nakakita, S.I., Hase, S., Ichikawa, A. and Ikura, K. (2002) Structural analysis of N-linked glycans from C. elegans. J. Biochem. 131, 807-813
- Neill, M.A. and Klebanoff, S.J. (1988) The effect of phenolic glycolipid-1 FROM Mycobacterium leprae on the antimicrobial activity of human macrophages. J. Exp. Med. 167, 30-42
- Ng, V., Zanazzi, G., Timpl. R., Talts, J.F., Salzer, J.L. and Rambukkana, A. (2000) Role of the cell wall phenolic glycolipid-1 in the peripheral nerve predilection of *Mycobacterium leprae*. Cell 103, 511-524
- Nigou, J., Gilleron, M. and Puzo, G. (1999a) Lipoarabinommannan: characterization of the multiacylated forms of the phosphatidyl-myo-inositol anchor by NMR spectroscopy. *Biochem. J.* 337, 453-460
- Nigou, J., Gilleron, M., Brando, T., Vercellone, A. and Puzo, G. (1999b) Structural definition of arabinomannans from *Mycobacterium bovis* BCG. Glycoconj. J. 16, 257-264
- Nigou, J., Gilleron, M., Cahuzac, B., Bounéri, J. D., Herold, M., Thurnher, M. & Puzo, G. (1997). The phospho-myo-inositol anchor of the lipomannans from *Mycobacterium bovis* bacillus Calmette Guérin. J. Biol. Chem. 272, 23094-23103.
- Nigou, J., Gilleron, M., Rojas, M., Garcia, L.F., Thurnher, M. and Puzo, G. (2002) Mycobacterial lipoarabinomannans: modulators of dendritic cell function and the apoptotic response. *Microbes Infect.* 4, 945-953
- Nigou, J., Vercellone, A. and Puzo, G. (2000) 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. 299, 1353-1362
- Nigou, J., Zelle-Rieser, C., Gilleron, M., Thurnher, M. and Puzo, G. (2001) Mannosylated lipoarabinomannans inhibit IL-12 production by human dendritic cells: evidence for a negative signal delivered through the mannose receptor. J. Immunol. 166, 7477-7485
- Nikaido, H., Kim, S.H. and Rosenberg, E.Y. (1993) Physical organisation of lipids in the cell wall of *Mycobacterium chelonae. Mol. Microbiol.* 8, 1025-1030
- Nishihara, T., Wyrick, R.E., Working, P.K., Chen, Y.H. and Hedrick, J.L. (1986) Isolation and characterization of a lectin from the cortical granules of *Xenopus laevis* eggs. *Biochemistry* 7, 6013-6020.
- Nishimori, I., Johnson, N.R., Sanderson, S.D., Perini, F., Montjoy, K., Cerny, R.L., Gross, M.L. and Hollingsworth, M.A. (1994) Influence of acceptor substrate primary amino acid sequence on the activity of human UDP-N-acetylgalactosamine:polypeptide N-acetylgalactosaminyltransferase. Studies with the MUC1 tandem repeat. J. Biol. Chem. 269, 16123-16130
- Nuzzo, I., Galdiero, M., Bentivoglio, C., Galdiero, R. and Carrateli, C.R. (2002) Apoptosis modulation by mycolic acid, tuberculostearic acid and trehalose 6,6'-dimycolate. J. Infect. 44, 229-235
- Olson, J.H. and Chandler, D.E. (1999) Xenopus laevis egg jelly contains small proteins that are essential to fertilization. Dev. Biol. 210, 401-410
- Oriol, R. (1995) ABO, Hh, Lewis and secretion. Serology, genetics and tissue distribution. In *Molecular* basis of major blood group antigens. Cartron, J.P. and Rouger P. (eds). Plenum Press, London
- 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, 1451-1463
- Oriol, R., Mollicone, R., Cailleau, A., Balanzino, L. and Breton, C. (1999) Divergent evolution of fucosyltransferase genes from invertebrates, invertebrates, and bacteria. *Glycobiology* 9, 323-334
- Ortalo-Magné, A., Dupont, M.A., Lemassu, A., Andersen, A.B., Gounon, P. and Daffé, M. (1995) Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiology* 141, 1609-1620
- Ozeki, Y., Kadena, K., Fujiwara, N., Morimoto, M., Oka, S. and Yano, I (1997) In vivo induction of apoptosis in the thymus by administration of mycobacterial cord factor (trehalose 6,6'-dimycolate). Infect. Immun. 65, 1793-1799
- Pallesen, L.T., Berglund, L., Rasmussen, L.K., Petersen, T.E and Rasmussen, J.T. (2002) Isolation and characterization of MUC15, a novel membrane-associated mucin. *Eur. J. Biochem.* 269, 2755-2763
- Palma, A.S., Vila-Verde, C., Pires, A.S., Fevereiri, P.S. and Costa, J. (2001) A novel plant α4fucosyntransferase synthesises the Lewis^a adhesion determinant. *FEBS Lett.* 499, 235-238
- Pandey, P., Kharbanda, S. and Kufe, D. (1995) Association of the DF3/MUC1 breast cancer antigen with Grb2 and the Sos/Ras exchange protein. *Cancer Res.* 55, 4000-4003
- Pangborn, M.C. and Mc Kinney, J.A. (1966) Purification of serologically active phosphoinositides of Mycobacterium tuberculosis. J. Lipid. Res. 7, 627-633

- Parry, S., Silverman, H.S, McDermott, K., Willis, A., Hollingsworth, M.A. and Harris, A. (2001) Identification of MUC1 proteolytic cleavage site in vivo. Biochem. Biophys. Res. Commun. 283, 715-720
- Paul, T.R and Beveridge, T.J. (1992) Reevaluation of envelope profiles and cytoplasmic ultrastructure of mycobacteria processed by conventional anbedding and freeze-substitution protocols. J. Bacteriol. 174, 6508-6517
- Paul, T.R and Beveridge, T.J. (1994) Preservation of surface lipids and determination of ultrastructure of *Mycobacterium kansasii* by freeze substitution. *Infect. Immun.* 62, 1542-1550
- Perez-Vilar, J. and Hill, R.L. (1998a) The carboxyl-terminal 90 residues of porcine submaxillary mucine are sufficient for forming disulfide-bonded dimers. J. Biol. Chem. 273, 6982-6988
- Perez-Vilar, J. and Hill, R.L. (1998c) Identification of the half-cystine residues in porcine submaxillary mucin critical to multimerization through the D-domains. Roles of the CGLCG motif in the D1- and D3-domains. J. Biol. Chem. 273, 34527-34534
- Perez-Vilar, J., Eckhardt, A.E. and Hill, R.L. (1996) Porcine submaxillary mucin forms disulfidebonded dimers between its carboxyl-terminal domains. J. Biol. Chem. 271, 9845-9850
- Perez-Vilar, J., Eckhardt, A.E., DeLuca, A. and Hill, R.L. (1998b) Porcine submaxillary mucin forms disulfide-linked multimers through its amino-terminal D-domains. J. Biol. Chem. 273, 14442-14449
- Peterson, P.K., Hu, S., Weng, W.S., Kravitz, F.H., Molitor, T.W., Chatterjee, D. and Chao, C.C. (1995) Thalidomide inhibits tumor necrosis factor alpha production by lipopolysaccharide- and lipoarabinomannan-stimulated human microglial cells. J. Infect. Dis. 172, 1137-1140
- Pierce-Crétel, A., Decottignies, J.P., Wieruszeski, Strecker, G., Montreuil, J. and Spik, G. (1989) 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 ¹H-NMR spectroscopy. *Eur. J. Biochem.* 183, 457-476
- Previato, J.O., Jones, C, Goncalves, L.P., Wait, R., Travassos, L.R. and Mendonca-Previato, L. (1994)
 O-glycosidically linked N-acetylglucosamine-bound oligosaccharides from glycoproteins of *Trypanosoma cruzi. Biochem. J.* 301, 151-159
- Prinzis, S., Chatterjee, D. and Brennan, P. (1993) Structure and antigenicity of lipoarabinomannan from *Mycobacterium bovis* BCG. J. Gen. Microbiol. 139, 2649-2658
- Probst, J.C., Gertzen, E.M. and Hoffmann, W. (1990) An integumentary mucin (FIM-B.1) from *Xenopus laevis* homologous with von Willebrand factor. *Biochemistry* 29, 6240-6244
- Prockop, D.J., Kivirikko, K.I., Tuderman, L. and Guzman, N.A. (1979) Medical progress. The biosynthesis of collagen and its disorders. N. Engl. J. Med. 301, 13-23
- Prome, J.C., Lacave, C., Ahibo-Coffy, A. and Savagnac, A. (1976) Séparation et étude structurale des espèces monomoléculaires de monmycolates et dimicolates de α-D-tréhalose présents chez Mycobacterium phlei. Eur. J. Biochem. 63, 543-552
- Pujin, J., Heumann, D., Tomasz, A., Kravchenko, V., Akamatsu, Y., Nishijima, M., Glauser, M., Tobias, P. and Ulevitch, R. (1994) CD14 is a pattern recognition receptor. *Immunity* 1, 509-516
- Rastogi, N. (1991) Recent observations concerning structure and functions relationships in the mycobacterial cell enveloppe: elaboration of a model in term of mycobacterial pathogenicity, virulence and drug resistance. *Res. Microbiol.* 142, 464-476
- Rastogi, N., Fréhel, C. and David H.L. (1986) Triple-layered structure of mycobacterial cell wall: evidence for the existence of a polysaccharide-rich outer layer in 18 mycobacterial species. *Curr. Microbiol.* 13, 237-242
- Reddy, M.S. (1992) Human tracheobronchial mucin: purification and binding to *Pseudomonas* aeruginosa. Infect. Immun. 60, 1530-1535
- Regimbald, L.H., Pilarski, L.M., Longenecker, B.M., Reddish, M.A., Zimmermenn, G. and Hugh, J.C. (1996) The breast mucin MUC1 as a novel adhesion ligand for endothelial inter-cellular adhesion molecule in breast cancer. *Cancer Res.* 56, 4244-4249
- Regoeczi, E., Chindemi, P.A., Rudolph, J.R., Spik, G. and Montreuil, J. (1987) The chromatographic heterogeneity of rat transferrin on immobilized concanavalin A and lentil lectin. *Biochem. Cell. Biol.* 65, 948-954
- Reinhart, D., Ridgway, J. and Chandler, D.E. (1998) *Xenopus laevis* fertilization: analysis of sperm motility in egg jelly using video light microscopy. *Zygote* 6, 173-182
- Ridley, M. (1996) In Evolution, second edition, Blackwell Scientific Publications Ltd Oxford.
- Rivera-Marrero, C.A., Ritzenlhaler, J.D., Newburn, S.A., Roman, J. and Cummings, R.D. (2002) Molecular cloning and expression of a novel glycolipid sulfotransferase in *Mycobacterium tuberculosis*. *Microbiology* 148, 783-792

- Rivière, M., Fournié, J.J. and Puzo, G. (1987) A novel mannose containing phenolic glycolipid from Mycobacterium kansasii. J. Biol. Chem. 262, 14879-14884
- Rock, P., Allietta, M., Young, W.W.J., Thompson, T.E. and Tillack, T.W. (1990) Organisation of glycosphingolipids in phosphatidyl-choline bilayers: use of antibody molecules and Fab fragments as morphologic markers. *Biochemistry* 29, 8484-8490
- Rogers, G.N. and D'Souza, B.L. (1989) Receptor binding properties of human and animal H1 influenza virus isolates. *Virology* 173, 317-322
- Roussel, P. and Lamblin, G. (1996) Human mucosal mucins in diseases. In *Glycoproteins and Diseases*. Neuberger, A. & Van Deenen, L.L.M. Elsevier Science B.V., 351-394
- Sadler, J.E. (1998) Biochemistry and genetics of von Willebrand factor. Annu. Rev. Biochem. 67, 395-424
- Sakaguchi, I., Ikeda, N., Nakayama, M., Kato, Y., Yano, I. and Kaneda, K. (2000) Trehalose 6,6'dimycolate (cord factor) enhances neovascularization through vascular endothelial growth factor production by neutrophils and macrophages. *Infect. Immun.* 68, 2043-2052
- Savage, A.V., Donoghue, C.M., D'Arcy, S.M., Koeleman, C.A. and van den Eijnden, D.H. (1990) Structure determination of five sialylated trisaccharides with core types 1, 3 or 5 isolated from bovine submaxillary mucin. *Eur. J. Biochem.* 192, 427-432
- Schaeffer, M.L., Khoo, K.H., Besra, G.S., Chatterjee, D., Brennan, P.J., Belisle, J.T. and Inamine, M. (1999) The *pimB* gene of *Mycobacterium tuberculosis* encodes a mannosyltransferase involved in lipoarabinomannan biosynthesis. J. Biol. Chem. 274, 31625-31631
- Scharfman, A. Delmotte, P., Beau, J., Lamblin, G., Roussel, P. and Mazurier, J. (2000) Sialyl-Le(x) and sulfo-sialyl-Le(x) determinants are receptors for *P. aeruginosa. Glycoconj. J.* 17, 735-740
- Schlesinger, L.S., Hull, S.R. and Kaufman, T.M. (1994) Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of *Mycobacterium tuberculosis* to human macrophages. J. *Immunol.* 152, 4070-4079
- Schlesinger, L.S., Kaufman, T.M., Iyer, S., Hull, S.R. and Marchiando, L.K. Differences in mannose receptor mediated uptake of lipoarabinomannan from virulent and attenuated strains of *Mycobacterium tuberculosis* by human macrophages. J. Immunol. 157, 4568-4575
- Schroeder, J.A., Thompson, M.C., Gardner, M.M. and gendler, S.J. (2001) Transgenic MUC1 interacts with epidermal growth factor receptor and correlates with mitogen-activated protein kinase activation in the mouse mammary gland. J. Biol. Chem. 276, 13057-13064
- Schwientek, T., Bennett, E.P., Flores, C., Thacker, J., Hollmann, M., Reis, C.A., Behrens, J., Mandel, U., Keck, B., Schafer, M.A., Haselmann, K., Zubarev, R., Roepstorff, P., Burchell, J.M., Taylor-Papadimitriou, J., Hollingsworth, M.A. and Clausen, H. (2002) Functional conservation of subfamilies of putative UDP-N-acetylgalactosamine:Polypeptide N-Acetylgalactosaminyltransferases in Drosophila, Caenorhabditis elegans, and mammals. J. Biol. Chem. 277, 22623-22638
- Severn, W.B., Furneaux, R.H., Falshaw, R. and Atkinson, P.H. (1998) Chemical and spectroscopic characterization of the phosphatidylinositol manno-oligosaccharides from *Mycobacterium bovis* AN5 and Wag201 and *Mycobacterium smegmatis* mc² 155. *Carbohydr. Res.* 308, 397-408
- Sheehan, J.K, Thornton, D.J., Howard, M., Carlstedt, I., Corfield, A.P. and Pareskeva, C. (1996) Biosynthesis of MUC2 mucin: evidence for a slow assembly of fully glycosylated units. *Biochem. J.* 315, 1055-1060
- Sheehan, J.K., Howard, M., Richardson, P.S., Longwill, T and Thornton, D.J. (1999) Physical characterization of a low-charge glycoform of the MUC5B mucin comprising the gel-phase of an asthmatic respiratory mucous plug. *Biochem. J.* 338, 507-513
- Sheehan, J.K., Oates, K. and Carlstedt (1987) Electron microscopy of cervical, gastric and bronchial mucus glycoproteins. *Biochem. J.* 239, 147-153
- Shimamura, M., Endo, T., Inoue, Y. and Inoue, S. (1983) A novel neutral oligosaccharide chain found in polysialoglycoproteins isolated from Pacific salmon eggs. Structural studies by secondary ion mass spectrometry, proton nuclear magnetic resonance spectroscopy, and chemical methods. *Biochemistry* 22, 959-963
- Sidobre, S., Nigou, J., Puzo, G. and Rivière, M. (2000) 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. 275, 2415-2422
- Sidobre, S., Puzo, G. and Rivière, M. (2002) Lipid-restricted rcognition of mycobacterial lipoglycans by human pulmonary surfactant protein A: a surface-plasmon-resonance study. *Biochem. J.* 365, 89-97
- Song, Y., Withers, D.A. and Hakomori, S. (1998) Globoside dependent adhesion of human embryonal carcinoma cells, based on carbohydrate-carbohydrate interaction, initiates signal transduction and induces enhanced activity of transcription factors AP1 and CREB. J. Biol. Chem. 273, 2517-2525

- Spicer, A.P., Rowse, G.J., Lidner, T.K. and Gendler, S.J. (1995) Delayed mammary progression in Muc-1 null mice. J. Biol. Chem. 270, 30093-30101
- Spiegel, M. (1954) The role of specific surface antigens in cell adhesion. I The reaggregation of sponge cells. *Biol. Bull.* 107, 130-148
- Spillmann, D., Hard, K., Thomas-Oates.J., Vliegenthart, J.F.G., Misevic, G., Burger, M.M. and Finne, J. (1993) Characterization of a novel pyruvilated carbohydrate unit implicated in the cell aggregation of the marine sponge *Microciona prolifera*. J. Biol. Chem. 268, 13378-13387
- Spillmann, D., Thomas-Oates.J., van Kuik, J.A., Vliegenthart, J.F.G., Misevic, G., Burger, M.M. and Finne, J. (1995) Characterization of a novel sulfated carbohydrate unit implicated in the carbohydrate-carbohydrate-mediated cell aggregation of the marine sponge *Microciona prolifera*. J. Biol. Chem. 270, 5089-5097
- Stewart, R.J. and Boogs, J.M. (1993) A carbohydrate-carbohydrate interaction between galactosylceramide-containing liposomes and cerebroside sulfate-containing liposomes: dependence on the glycolipid ceramide composition. *Biochemistry* 32, 10666-10674
- Strain, S.M., Toubiana, R., Ribi, E. and Parker, R. (1977) Separation of the mixture of trehalose 6,6dimicolates comprising the mycobacterial glycolipid fraction. *Biochem. Biophys. Res. Commun.* 77, 449-456
- Strecker, G., Wieruszeski, J.M., Cuvellier, O., Michalski, J.C. and Montreuil, J. (1992) ¹H and ¹³C-NMR assignments for sialylated oligosaccharide-alditols related to mucins. Study of thirteen components from hen ovomucin and swallow nest mucin. *Biochimie* 74, 39-52
- Strecker, G., Wieruszeski, J.M., Michalski, J.C., Alonso, C., Boilly, B. and Montreuil, J. (1992) Characterization of Le^x, Le^y and A Le^y antigen determinants in Kdn-containing O-linked glycan chains from *Pleurodeles waltlii* jelly coat eggs. *FEBS lett.* 298, 39-43
- Strecker, G., Wieruszeski, J.M., Plancke, Y. and Boilly, B. (1995) Primary structure of 12 neutral oligosaccharide-alditols released from the jelly coats of the anuran *Xenopus laevis* by reductive β-elimination. *Glycobiology* 5, 137-146
- Strous, G.J. and Dekker, J. (1992) Mucin-type glycoproteins. Crit. Rev. Biochem. Mol. Biol. 27, 57-92
- Taylor-Papadimitriou, J., Burchell, J., Miles, D.W. and Dalziel, M. (1999) MUC1 and cancer. *Biochim. Biophys. Acta.* 1455, 301-313
- Thomsen, D.R., Post, L.E. and Elhammer, A.P. (1990) Structure of O-glycosidically linked oligosaccharides synthesized by the insect cell line Sf-9. J. Cell. Biochem. 43, 67-79
- Thornton, D.J., Carlstedt, I., Howard, M., Devine, P.L., Price, M.R. and Sheehan, J.K. (1996) Respiratory mucins: identification of core proteins and glycoforms. *Biochem. J.* 316, 967-975
- Thornton, D.J., Howard, M., Khan, N. and Sheehan, J.K. (1997) Identification of two glycoforms of the MUC5B mucin in human respiratory mucus. J. Biol. Chem. 272, 9561-9566
- Tian, J., Gong, H., G.H. and Lennarz, W.J. (1999) *Xenopus laevis* sperm receptor gp69/64 glycoprotein is a homolog of the mammalian sperm receptor ZP2. *Pro. Natl. Acad. Sci. USA* 96, 829-834.
- Tian, J., Gong, H., Thomsen, G.H. and Lennarz, W.J. (1997) Gamet interaction in *Xenopus laevis*: identification of sperm binding glycoproteins in the egg vitelline enveloppe. J. Biol. Cell. 136, 1099-1108.
- Tillack, T.W., Allietta, M., Moran, R.E. and Young, W.W.J. (1983) Localization of globoside and forsman glycolipids on erythrocyte membranes. *Biochim. Biophys. Acta*. 733, 15-24
- Todeschini, R.A., da-Silveira, E.X., Jones, C., Wait, R., Previato, J.O. and Mendonça-Previato (2001) Structure of O-glycosidically linked oligosaccharides from glycoproteins from *Trypanosoma cruzi* CL-Brenner strain: evidence for the presence of O-linked sialyl oligosaccharides. *Glycobiology* 11, 47-55
- Toyoda, H., Kinoshita-Toyoda, A. and Selleck, S.B. (2000) Structural analysis of glycosaminoglycans in Drosophila and *Caenorhabditis elegans* and demonstration that tout-velu, a Drosophila gene related to EXT tumor suppressors, affects heparan sulfate *in vivo. J. Biol. Chem.* 275, 2269-2275
- Treumann, A., Xidong. F., McDonnel, L., Derrick, P.J., Aschcroft, A.E., Chatterjee, D. and Homans, S.W. (2002) 5-methylthiopentose: a new substituent on lipoarabinomannan in *Mycobacterium* tuberculosis. J. Mol. Biol. 316, 89-100
- Tromas, C., Rojos, J., de la Fuente, J.M., Barrientos, A.G., Garcia, R. and Penades, S. (2001) Adhesion forces between Lewis^x determinant antigens as measured by atomic force microscopy. Angew. Chem. Int. Ed. 40, 3052-3055
- Turley, E.A. and Roth, S. (1980) Interactions between the carbohydrate chains of hyaluronate and chondroitin sulphate. *Nature* 283, 268-271
- Vacquier, V.D. (1998) Evolution of gamete recognition proteins. *Science* 281, 1995-1998.
- van de Vyver, G. (1975) Phenomena of cellular recognition in sponges. Curr. Top. Dev. Biol. 10, 123-140

- van Halbeek, H., Strang, A.M., Lhermitte, M., Rahmoune, H., Lamblin, G and Roussel, P. (1994) Structures of monosially oligosaccharides isolated from the respiratory mucins of a non-secretor (O, Lea+b-) patient suffering from chronic bronchitis. Characterization of a novel type of mucin carbohydrate core structure. *Glycobiology* 4, 203-219
- Van Klinken, B.J.W, Einerhand, A.W.C, Büller, H.A. and Dekker (1998a) Strategic biochemical analysis of mucins. *Anal. biochem.* 265, 103-116
- Van Klinken, B.J.W, Einerhand, A.W.C, Büller, H.A. and Dekker, J. (1998b) The oligomerization of four genetically clustered human grastrointestinal mucins. *Glycobiology* 8, 67-75
- Varner, J. (1995) Cell adhesion in sponges: potentiation by a cell surface 68 kDa proteoglycan-binding protein. J. Cell. Science 108, 3119-3126
- Varner, J. (1996) Isolation of a sponge-derived extracellular matrix adhesion protein. J. Biol. Chem. 271, 16119-16125
- Venisse, A. Berjeaud, J.M., Chaurand, P., Gilleron, M. and Puzo, G. (1993) Structural features of lipoarabinomannan from *Mycobacterium bovis* BCG. J. Biol. Chem. 268, 12401-12411
- Venisse, A., Rivière, M., Vercauteren, J. and Puzo, G. (1995) Structural analysis of the mannan region of lipoarabinommannan from *Mycobacterium bovis* BCG. J. Biol. Chem. 270, 15012-1521
- Vercellone, A. and Puzo, G. (1989) New-found phenolic glycolipids in *Mycobacterium bovis* BCG. J. Biol. Chem. 264, 7447-7454
- Verma, M. and Davidson, E.A. (1993) Molecular cloning and sequencing of a canine tracheobronchial mucin cDNA containing a cysteine-rich domain. *Proc. Natl. Acad. Sci. U.S.A.* 90, 7144-7148
- Vinall, L.E., Hill, A.S., Pigny, P., Pratt, W.S., Toribara, N., Gum, J.R., Kim, Y.S, Porchet, N., Aubert, J.P. and Swallow, D.M. (1998) Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5. *Hum. Genet.* 102, 357-366
- Wandall, H.H., Hassan, H., Mirgorodskaya, E., Kristensen, A.K., Roepstorff, P., Bennett, E.P., Nielsen, P.A., Hollingsworth, A.K., Burchell, J., Taylor-Papodimitriou and Clausen, H. (1997) Substrate specificities of three members of the human UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. J. Biol. Chem. 272, 23503-23514
- Wang, L., Slayden, R.A., Barry III, C.E. and Liu, J. (2000) Cell wall structure of a mutant of *Mycobacterium smegmatis* defective in the biosynthesis of mycolic acids. J. Biol. Chem. 275, 7224-7229
- Wang, Y., Agrwal, N., Eckhardt, A.E., Stevens, R.D. and Hill, R.D (1993) The acceptor substrate specificity of porcine submaxillary UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase is dependent on the amino acid sequences adjacent to serine and threonine residues. J. Biol. Chem. 268, 22979-22983
- Warren, C.E., Krizus, A. and Dennis, J.W. (2001) Complementary expression patterns of six nonessential *Caenorhabditis elegans* core 2/I N-acetylglucosaminyltransfearse homologues. *Glycobiology* 11, 979-988
- Warren, C.E., Krizus, A., Partridge, E.A. and Dennis, J.W. (2002) *Caenorhabditis elegans gly-1*, a core 2/I N-acetylglucosaminyl-transferase homologue, is a glucosyltransferase. *Glycobiology* 12, 8G-9G
- Watanabe, M., Aoyagi, Y., Mitome, H., Fujita, T., Naoki, H., Ridell, M. and Minnikin, D.E (2002) Location and functional groups in mycobacterial meromycolate chains; the recognition of new structural principles in mycolic acids. *Microbiology* 148, 1881-1902
- Watanabe, M., Aoyagi, Y., Ohta, A. and Minnikin, D.E. (1997) Structures of phenolic glycolipids from Mycobacterium kansasii. Eur. J. Biochem. 248, 93-98
- Weitzerbin, J., Das, B.C., Petit, J.F., Ledere, E., Leyh-Buille, M. and Ghuysen, J.M. (1974) Occurence of D-alanyl-(D)-meso-diaminopimelic acid and meso-diaminopimelic-meso-diaminopimelic acid interpeptide linkages in the peptidoglycan of mycobacteria. *Biochemistry* 13, 246-252
- Weitzerbin-Falszpan, J., Das, B.C., Azuma, I., Adam, A., Petit, J.F. and Lederer, E. (1970) Isolation and mass spectrometric identification of the peptide subunits of mycobacterial cell walls. *Biochim. Biophys. Res. Comm.* 40, 57-63
- Wesseling, J., Van der Valk, S.W., Vos, H.L., Sonnenberg, A. and Hilkens, J. (1995) Episialin (MUC1) overexpression inhibits integrin-mediated cell adhesion to extracellular matrix components. J. Cell. Biol. 129, 255-265
- Wickstrom, C and Carlstedt, I. (2001) N-terminal cleavage of the salivary MUC5B mucin. Analogy with the Von Willebrand propolypeptide? J. Biol. Chem. 276, 47116-47121
- Wickström, C., Davies, J.R., Eriksen, G.V., Veerman, E.C.I. and Carlstedt, I. (1998) MUC5B is a major gel-forming, oligomeric mucin from human salivary gland, respiratory tract and endocervix: identification of glycoforms and C-terminal cleavage. *Biochem. J.* 334, 685-693

- 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. (1987) Structure of the monosialyl oligosaccharides derived from salivary gland mucin glycoproteins of the chines swiftlet (Genus *Collocalia*). J. Biol. Chem. 262, 6650-6657
- Williams, S.J., Wreschner, D.H., Tran, M., Eyre, H.J., Sutherland, G.R. and McGuckin, M.A. (2001) MUC13, a novel human cell surface mucin expressed by epithelial and hemopoieitic cells. *J. Biol. Chem.* 276, 18327-18336
- Wilson H.V. (1907) On some phenomena of coalescence and regeneration in sponges. J. Exp. Zool. 5, 245-258
- Wise, R.J., Pittman, D.D., Handin, R.I., Kaufman, R.J. and Orkin, S.H. (1988) The propeptide of van Willebrand multimers. *Cell* 52, 229-236
- Wood, W.B. (ed.) (1998) The nematode *Caenorhabditis elegans*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Wreschner, D.H., McGuckin, M.A., Williams, S.J., Baruch, A., Yoeli, M., Ziv, R., Okun, L., Zaretsky, J., Smorodinsky, N., Keydar, I., Neophytou, P., Stacey, M., Lin, H.H. and Gordon, S. (2002) Generation of ligand-receptor alliances by "SEA" module-mediated cleavage of membrane-associated mucin proteins. *Protein Sci.* 11, 698-706
- Wyrick, R.E., Nishihara, T. and Hedrick, J.L. (1974) Agglutination of jelly coat and cortical granule components and the block to polyspermy in the amphibian *Xenopus laevis*. *Proc. Natl. Acad. Sci. USA* 71, 2067-2071.
- Yamada, S., Van Die, I., Van den Eijnden, D.H., Yokota, A., Kitagawa, H. and Sugahara, K. (1999) Demonstration of glycosaminoglycans in *Caenorhabditis elegans*. *FEBS Lett.* 459, 327-331
- Yamagami, H., Matsumoto, T., Fujiwara, N., Arakawa, T., Kaneda, K., Yano, I. and Kobayashi, K. (2001) Trehalose 6,6'-dimycolate (cord factor) of *Mycobacterium tuberculosis* induces foreign-bodyand hypresensitivity-type granulomas in mices. *Infect. Immun.* 69, 810-815
- Yoshida, A. and Koide, Y. (1997) Arabinofuranosyl- terminated and mannosylated lipoarabinomannans from *Mycobacterium tuberculosis* induce different levels of interleukin-12 expression in murine macrophage. *Infect. Immun.* 65, 1953-1955
- Yoshizaki, N. (1985) Fine structure of oviducal epithelium of *Xenopus laevis* in relation to its role in secreting egg envelopes. J. Morphol. 184, 155-169
- Yu, S., Kojima, N., Hakomori, S., Kudo, S., Inoue, S. and Inoue, Y. (2002) Binding of rainbow trout sperm to egg is mediated by strong carbohydrate-carbohydrate interaction between (KDN)GM3 (deaminated neuraminyl ganglioside) and Gg3-like epitope. *Proc. Natl. Acad. Sci. USA* 99, 2854-2859
- Yuan, Y., Crane, D.D. and Barry III, C.E. (1998) The effect of oxygenated mycolic acid composition on cell wall function and macrophage growth in *Mycobacterium tuberculosis*. *Mol. Microbiol.* 29, 1449-1458
- Yurewicz, E.C., Oliphant, G. and Hedrick, J.L. (1975) The macromolecular composition of *Xenopus laevis* egg jelly coat. *Biochemistry* 14, 3101-3106
- Zhang, Y., Doefler, M. Lee, T.C., Guillemin, B. and Rom, W.N. (1993) Mechanisms of stimulation of interleukin-1β and tumor necrosis factor-α by *Mycobacterium tuberculosis* components. J. Clin. Invest. 91, 2076-2083

