

UNIVERSITE DE LILLE – FACULTE DES SCIENCES ET TECHNOLOGIES

ECOLE DOCTORALE BIOLOGIE SANTE – LILLE

## THESE DE DOCTORAT

Discipline : Aspects moléculaires et cellulaires de la biologie

Spécialité : Biochimie et biologie moléculaire

Présentée par

**Audrey AUGER**

En vue de l'obtention du titre de Docteur de l'Université de Lille

# ROLE DES OLIGOSACCHARIDES DANS LES INTERACTIONS VIRUS-HOTE

Application à la contamination des huîtres par les Norovirus

*Soutenance le 08/12/2023*

Commission d'examen :

<b>Présidente :</b>	<b>Pr. Sophie GROUX-DEGROOTE</b>	UGSF UMR CNRS 8576, Université de Lille
<b>Rapporteurs :</b>	<b>Pr. Muriel BARDOR</b>	GlycoMEV UR 4358, Université de Rouen Normandie
	<b>Pr. Abderrahman MAFTAH</b>	LABCiS UR 22722, Université de Limoges
<b>Examinatrice :</b>	<b>Dr. Marion DESDOUTS</b>	MASAE, Ifremer, Nantes
<b>Directeur :</b>	<b>Dr. Frédéric CHIRAT</b>	UGSF UMR CNRS 8576, Université de Lille





UNIVERSITE DE LILLE – FACULTE DES SCIENCES ET TECHNOLOGIES

ECOLE DOCTORALE BIOLOGIE SANTE – LILLE

## THESE DE DOCTORAT

Discipline : Aspects moléculaires et cellulaires de la biologie

Spécialité : Biochimie et biologie moléculaire

Présentée par

**Audrey AUGER**

En vue de l'obtention du titre de Docteur de l'Université de Lille

# ROLE DES OLIGOSACCHARIDES DANS LES INTERACTIONS VIRUS-HOTE

Application à la contamination des huîtres par les Norovirus

*Soutenance le 08/12/2023*

Commission d'examen :

<b>Présidente :</b>	<b>Pr. Sophie GROUX-DEGROOTE</b>	UGSF UMR CNRS 8576, Université de Lille
<b>Rapporteurs :</b>	<b>Pr. Muriel BARDOR</b>	GlycoMEV UR 4358, Université de Rouen Normandie
	<b>Pr. Abderrahman MAFTAH</b>	LABCiS UR 22722, Université de Limoges
<b>Examinatrice :</b>	<b>Dr. Marion DESDOUTS</b>	MASAE, Ifremer, Nantes
<b>Directeur :</b>	<b>Dr. Frédéric CHIRAT</b>	UGSF UMR CNRS 8576, Université de Lille







Ce travail a été encadré par le Dr. Frédéric Chirat au sein de l'Unité de Glycobiologie Structurale et Fonctionnelle (UGSF), Unité Mixte de Recherche (UMR) 8576 - Centre National de Recherche Scientifique (CNRS), Basée à l'Université de Lille en France et actuellement dirigée par le Dr. Yann Guérardel.

Cette thèse a bénéficié d'une allocation du Ministère de l'Enseignement Supérieur et de la Recherche (MESR) et d'un co-financement de l'UGSF.



# REMERCIEMENTS

---

En premier lieu, je tiens à remercier l'ensemble des membres de mon jury d'avoir accepté de participer à ma soutenance de thèse. Plus particulièrement, je remercie les Prs **Muriel Bardor** et **Abderrahman Maftah** de me faire l'honneur d'être mes rapporteurs. Je remercie également la **Dr Marion Desdouits** d'être examinatrice et ainsi, après avoir suivi une grande partie de ma thèse *via* le projet ANR, d'assister à la finalité de mes travaux. Enfin, je tiens à remercier grandement la Pr **Sophie Groux-Degroote** d'avoir accepté d'être la présidente de mon jury.

J'adresse également mes remerciements aux Drs **Jacques Le Pendu** et **Jean Dubuisson** pour leurs participations à mes Comités de Suivi Individuels. Je remercie plus particulièrement le Dr **Jacques Le Pendu** pour les discussions au travers du projet ANR ainsi que son accueil très chaleureux et très instructif lors d'un séjour au sein de son équipe à Nantes.

Un énorme merci à mon directeur de thèse, le Dr **Frédéric Chirat**, pour m'avoir accepté en stage de master après une longue réflexion et s'être lancé dans cette aventure qu'est la thèse avec moi. Merci de m'avoir fait confiance, de m'avoir guidé, d'avoir été si pédagogue et si conciliant. Ça a été un réel travail d'équipe et un énorme plaisir au quotidien ! Je pense que peu de doctorant peuvent se vanter d'avoir autant manipuler côte à côte avec son directeur de thèse ! Du grand sérieux accompagné de beaucoup de taquineries au quotidien qui a constitué un environnement très agréable et propice à mon épanouissement scientifique. Je n'oublie pas qu'il t'aura fallu énormément de patience avec moi et que tu as toujours fait en sorte de me mettre en avant et de m'encourager pour que je donne le meilleur de moi-même.

Je tiens à remercier vivement le Dr **Yann Guérardel**, tout d'abord pour m'avoir accepté dans son équipe en stage de master et puis pour s'être démené pour que je puisse réaliser cette thèse. Je n'oublie pas ton aide pour l'élaboration de mon sujet de thèse initial mais surtout celle apportée lorsqu'il a fallu en changer. Je t'en suis très reconnaissante et j'ai pu prendre un grand plaisir à faire avancer ce projet. Merci également pour ton implication durant ces années, tes conseils et toutes ces discussions scientifiques si enrichissantes.

Il y a toutes les personnes de mon équipe qui ont, d'une manière ou d'une autre, contribué à mon épanouissement au laboratoire et à ma réussite. Merci à tous ces enseignants-chercheurs qui sont toujours de bons conseils : merci à **Isabelle, Elisabeth, Christophe B, Christophe M, Florence, Fred K** et **Cédric**. Merci également à **Yannick R, Zina, Clément, Lucie** et **Olivier** pour avoir amené une ambiance encore meilleure au 2<sup>ème</sup> étage dès leur arrivée. Je remercie tout particulièrement le Dr **Yannick R** pour ses conseils avisés et pour toute la passion qu'il dégage envers la science et qui est une source d'inspiration.

Je remercie également **Charlotte, Bernadette** et **Shin Yi** qui ont pu me suivre du stage à la thèse et qui m'ont aidé à me former. Je remercie vivement **Shin Yi** qui a été d'une très grande aide pour les analyses de spectrométrie de masse : les compétences que j'ai développé pendant ma thèse sont en grande partie grâce à toi ! Je te remercie pour ta gentillesse et ta grande disponibilité.

Et puis il y a ces personnes qui ont aussi le courage de vivre l'aventure de la thèse. Merci à **Réginald, Laura, Awatef** et **Jodie**. Merci à **Louis-David**, tu as eu le courage de te lancer dans une thèse et tu l'as réussi avec brio ! Merci également pour toutes ces discussions, tes conseils mais aussi ces taquineries qui ont rendu le quotidien agréable. Il y a également ces personnes avec qui on passe tellement de temps que certaines en deviennent des amis avec qui on partage nos misères, on se soutient mais aussi avec qui on passe des soirées mémorables ! Merci à **Vincent** et **Yannick M** d'avoir été là en toute circonstance (on n'oubliera pas ces souvenirs aux Barnett's et au Stairway of course !). On se souviendra de l'invention du tableau des scores de blagues carambar qui, j'ai bien l'impression, sera notre héritage ! Et puis également merci à **Ryma** du fin fond du couloir du 2<sup>ème</sup> étage d'être aussi simple, vraie et d'avoir cette capacité à nous faire mourir de rire en une parole ! On n'oubliera pas toutes ces fameuses histoires de rat ! Cette thèse n'aurait pas été aussi bien sans vous ! Merci également aux anciens doctorants dont j'ai pu croiser la route et qui ont été une source d'inspiration : merci **Lin, P-A, Clémence, Marine, Maxence, Justine** et **Sumeyye**.

J'aimerais remercier les personnes de la plateforme PAGès, **Nao, Manu** et **Louise**, pour leur disponibilité, leurs conseils et également leur aide sur les différentes machines parfois bien capricieuses !

Je remercie tous les stagiaires qui ont pu croiser ma route au laboratoire et qui y ont amenée encore plus de vie. Je remercie particulièrement **Horgan**, avec qui ça a été un réel plaisir de travailler. Je remercie **Amandine** et **Alexandra** qui m'ont faites confiance pour les encadrer pour leur stage de BTS et de licence. Merci à **Geoffroy** et **Patrick** de m'avoir donné ma chance en enseignement et d'avoir été, avec **Rozenn**, si bienveillants et de si bons conseils autant envers moi qu'envers les étudiants. Ces expériences d'encadrements ont été pleines d'apprentissages.

Je tiens à remercier toutes les personnes qui ont été là d'une manière ou d'une autre au laboratoire notamment à l'occasion d'un échange au détour d'un couloir. Je tiens particulièrement à remercier **Tony**. Merci pour tes conseils avisés, ta passion débordante et surtout de m'avoir incitée à rejoindre ton club d'athlétisme, ces moments ont été d'un grand secours !

Je remercie également les personnes de l'**équipe administrative** pour leur bonne humeur et leur aide, et particulièrement **Nadège** qui s'est démenée dans tous les démarches administratives auxquelles j'ai pu faire face.

Je tiens également à remercier les personnes que j'ai eu l'occasion de rencontrer lors de mon séjour à Nantes. Merci à **Adrien** et **Jézabel** pour l'accueil, les discussions aussi bien scientifiques que

personnelles et pour toute l'aide apportée ! Je remercie également **Caroline, Juliette** et **Lorine** qui ont encore plus enrichies ce séjour. Et bien entendu, je ne peux oublier cette rencontre amicale : Je te remercie **Nicolas**, d'abord pour ton aide alors même que tu n'étais pas de l'équipe de Jacques, et bien sûr je te remercie pour toutes les bourdes, les blagues et les fous rires incontrôlables qu'on a pu vivre ! On se souviendra tous de ces soirées crêpes et jeux ! Et surtout, on se souviendra qu'il faut faire attention à avoir sa carte de porte d'entrée sur soi avant de claquer sa porte !

Merci à ma famille. Merci à **mes parents** d'avoir toujours cru en moi et d'avoir fait tous leur possible pour assurer ma réussite. Merci à **ma grand-mère** qui, depuis quelques années maintenant, fait encore plus partie de ma vie et qui est toujours de bons conseils dans nos passions communes ! Bien évidemment, merci à **mes sœurs, Emilie** et **Estelle**, pour tous ces moments agréables et complices.

Et puis il y a les amis qui sont toujours là pour écouter nos mésaventures de thèse alors même qu'ils ne sont pas du domaine ! Merci à **Amélie, Antoine, VE, Gery, Habiba** et **Julien** pour tous ces bons moments.

Pour finir, je tiens à remercier la personne la plus importante de ma vie qui m'accompagne depuis maintenant presque 10 ans. Merci **Mathieu** d'être là et de me soutenir en toute circonstance. Ça a été une aventure exceptionnelle qu'on a eu la chance de vivre ensemble : c'est tellement enrichissant de pouvoir parler le même langage ! Ça a également été une période pleine de doutes et où on est confronté à ses propres limites. La thèse, et en particulier cette fin de thèse, aura été une épreuve que je n'aurais su surmonter sans toi. Merci pour tout, merci d'être toi et j'espère que ce n'est que le début d'une plus grande aventure encore !

# RESUME

---

Les Norovirus (NoVs) sont la principale cause de gastro-entérite dans le monde. Les contaminations se produisent principalement de personne à personne ou par la consommation d'aliments contaminés comme les fruits, les légumes ou les crustacés. Dans ce dernier cas, l'huître est connue pour être un excellent vecteur, notamment en raison de sa consommation crue, de sa proximité avec les effluents humains, de sa capacité de filtration et de sa rétention spécifique des NoVs. En utilisant différentes lectines, anticorps et virus-like particles (VLPs), il a été démontré que la rétention spécifique des NoVs repose sur la présence d'une molécule similaire à l'antigène A (Ag A) chez l'huître.

Afin de mettre en évidence la présence et la distribution d'antigènes de groupes sanguins tissulaires sur les glycoconjugués des espèces d'huîtres *Crassostrea gigas* et *Ostrea edulis*, nous avons établi les *N*- et *O*-glycomes du manteau, des branchies et du tissu digestif. Les *N*-glycannes ont été séquentiellement libérés par la Peptide:*N*-glycosidase F (PNGase F) et A. Après perméthylation, les analyses structurales par MALDI-QIT-TOF ont révélé la présence des Ags A et A Lewis b (ALe<sup>b</sup>) pour les deux espèces ainsi que des antigènes H et HLe<sup>b</sup> uniquement pour *O. edulis*. La deutéroperméthylation à la place de la perméthylation suivie d'analyses MALDI-MS<sup>n</sup> ont permis de montrer la présence de *N*-glycannes naturellement méthylés comme cela a été démontré chez l'huître *Crassostrea virginica*. Pour identifier les monosaccharides porteurs de groupements méthyles et préciser leur position, les monosaccharides des *N*-glycannes ont été transformés en leur dérivé itol acétate et analysés par chromatographie en phase gazeuse couplée à la spectrométrie de masse EI-MS. Nous avons ainsi pu révéler la présence d'un groupement méthyle sur le carbone 3 des résidus de *N*-acétylgalactosamine des Ags A pour les deux espèces. Des résidus de fucose sont aussi apparus méthylés sur leur carbone 3 ou 4 pour les huîtres *C. gigas* et *O. edulis*, respectivement. Ces résidus de méthyl-fucose ont été identifiés dans l'antigène Le<sup>b</sup> pour les deux espèces et dans l'Ag A uniquement pour *O. edulis*.

Dans le but d'étendre ces analyses aux *O*-glycannes, un protocole de préparation des mucines a été mis en œuvre et les *O*-glycannes ont été libérés par  $\beta$ -élimination réductive, perméthylés ou perdeutérométhylés et analysés par spectrométrie de masse MS<sup>n</sup>. Nos résultats indiquent également la

présence d'Ags A mais également d'Ags H pour les deux espèces. La présence d'Ags Le<sup>b</sup> est également suspectée. De plus, la présence de groupements méthyles au sein des *O*-glycannes natifs a été mise en évidence.

En parallèle, des expériences ELISA et de Far-Western Blotting (FarWB) utilisant des VLPs dérivées des NoVs GI.1 et GII.4 ont été réalisées pour déterminer la ou les macromolécules des deux espèces impliquées dans l'interaction hôte-virus. Les résultats ont démontré une reconnaissance différentielle des tissus d'huîtres par les deux génotypes. Les données préliminaires obtenues par FarWBs indiquent sans ambiguïté que les ligands du virus sont associés à des glycoprotéines de très haut poids moléculaire. L'élimination des *N*-glycannes par la PNGase F n'affecte pas la reconnaissance par les VLPs. L'ensemble de ces résultats suggère fortement que seuls les *O*-glycannes seraient impliqués dans la reconnaissance NoV-huître et pourraient appartenir à des glycoprotéines de type mucine. La présence de mucines chez les mollusques n'a pas encore été formellement démontrée.

Mots clés : HBGA, Glycomique, Interaction hôte-pathogène, Norovirus, Huître.

# ABSTRACT

---

Noroviruses (NoVs) are the most common cause of gastroenteritis worldwide. Contamination occurs mainly from person to person or through the consumption of contaminated foods such as fruits, vegetables or shellfish. In the latter case, oysters are known to be excellent vectors, particularly due to their raw consumption, proximity to human sewage, filtration capacity and specific retention of NoVs. Using various lectins, antibodies and virus-like particles (VLPs), it has been demonstrated that specific retention of NoVs relies on the presence of an A-like antigen (A Ag) in oysters.

To highlight the presence and distribution of Histo-blood group antigens on glycoconjugates from the oyster species *Crassostrea gigas* and *Ostrea edulis*, we established *N*- and *O*-glycomes from mantle, gills and digestive tissue. *N*-glycans were sequentially released by peptide:*N*-glycosidase F (PNGase F) and A. After permethylation, structural analyses by MALDI-QIT-TOF revealed the presence of A and A Lewis b (ALe<sup>b</sup>) Ags for both species, as well as H and HLeb Ags for *O. edulis* only. Deutero-permethylation instead of permethylation, followed by MALDI-MS<sup>n</sup> analyses revealed the presence of naturally methylated *N*-glycans, as demonstrated in the oyster *Crassostrea virginica*. To identify the monosaccharides carrying methyl groups and specify their position, the *N*-glycan monosaccharides were converted into their itol acetate derivative and analyzed by gas chromatography coupled with EI-MS mass spectrometry. This revealed the presence of a methyl group on carbon 3 of *N*-acetylgalactosamine residues in A Ags for both species. Fucose residues also appeared methylated on their carbon 3 or 4 for *C. gigas* and *O. edulis* oysters, respectively. These methyl-fucose residues were identified in Le<sup>b</sup> Ag for both species, and in A Ag only for *O. edulis*. To extend these analyses to *O*-glycans, a mucin preparation protocol was implemented and *O*-glycans were released by reductive  $\beta$ -elimination, permethylated or perdeuteromethylated and analysed by MS<sup>n</sup> mass spectrometry. Our results also indicate the presence of A Ags but also H Ags for both species. The presence of Le<sup>b</sup> Ags is also suspected. The presence of methyl groups within native *O*-glycans was also demonstrated.

In parallel, ELISA and Far-Western Blotting (FarWB) experiments using VLPs derived from NoVs GI.1 and GII.4 were carried out to determine the macromolecule(s) of the two species involved



in the host-virus interaction. The results demonstrated differential recognition of oyster tissues by the two genotypes. Preliminary data obtained by FarWBs unambiguously indicate that the virus ligands are associated with very high molecular weight glycoproteins. Removal of *N*-glycans by PNGase F does not affect recognition by VLPs. Taken together, these results strongly suggest that only *O*-glycans are involved in NoV-oyster recognition and may belong to mucin-type glycoproteins. The presence of mucins in mollusks has not yet been formally demonstrated.

Key words : HBGA, Glycomic, Host-pathogen interaction, Norovirus, Oyster.

# ABREVIATIONS

---

AA : acide aminé	HexNAc : <i>N</i> -Acetylhexosamine
AAL : <i>Aleuria aurantia</i> , : <i>Aleuria aurantia</i>	HPA : <i>Helix pomatia</i>
ACN : acétonitrile	HRP : HorseRadish Peroxydase
Ag : antigène	HSV : virus de l'herpès
Asn : asparagine	IRM : Ion Routing Multipole
ATCA : acide trichloroacétique	Kdn : acide 3-déoxy-D-glycéro-D-galacto- non-2-ulosonique
ATFA : acide trifluoroacétique	LacNAc : <i>N</i> -acétyllactosamine
BSA : serum albumin bovine	LSEM : Laboratoire Santé Environnement et Microbiologie
C : carbone	MALDI : Matrix-Assisted Laser Desorption/Ionization
CID : Collision Induced Dissociation	Man : mannose
CNX : calnexine	MW : masse moléculaire
Con A : Concanavaline A	nanoLC : chromatographie liquide nano haute performance
CoV : coronavirus	Neu5Ac : acide 5- <i>N</i> -acétylneuraminique
CPG : Chromatographie en Phase Gazeuse	Neu5Gc : acide 5- <i>N</i> -glycolylneuraminique
CRT : calréticuline	NoV : Norovirus
DHB : acide 2,5-dihydroxybenzoïque	ORF : Open Reading Frame
dHex : déoxyhexose	OST : oligosaccharyltransférase
DMSO : diméthylsulfoxyde	PAS : Acide Périodique Schiff
Dol-P : dolichol phosphate	PBS : Phosphate Buffered Saline Solution
DTT : dithiothréitol	Pent : pentose
ECL : électrochimiluminescence	PMAA : partially methylated alditol acetates
EDEM : ERAD-enhancing- $\alpha$ -mannosidases degradation	PNGase : Peptide- <i>N</i> -glycosidase
EI : source d'ionisation électronique	ppGalNAcT : polypeptide GalNAc-transférase
EI-MS : Electron Impact-Mass Spectrometry	QIT : Quadrupole Ion Trap
ELISA : enzyme-linked immunosorbent assay	RdRp : RNA-dependent RNA polymerase
ERAD : Endoplasmic Reticulum Associated Degradation	RE : réticulum endoplasmique
ESI : ElectroSpray Ionization	RIPA : RadioImmunoPrecipitation Assay
ETD : Electron Transfer Dissociation	RPM : rotations par minute
Fuc : fucose	SDS : sodium dodecyl sulfate
GAG : glycosaminoglycane	Ser : sérine
Gal : galactose	SNFG : Symbol Nomenclature for Glycans
GalNAc : <i>N</i> -acétylgalactosamine	SPE : solid phase extraction
GBP : Glycan binding proteins	Tamb : température ambiante
GCNT : $\beta$ 1-6 <i>N</i> -acétylglucosaminyltransférase	Thr : thréonine
GI : $\alpha$ -glucosidase I	TMB : tétraméthylbenzidine
GII : $\alpha$ -glucosidase II	TOF : Time Of Flight
Glc : glucose	UEA-I : <i>Ulex europaeus</i> I
GlcNAc : <i>N</i> -acétylglucosamine	UGGT : UDP glucose:glycoprotein glucosyltransférase
GlcNAc-T : <i>N</i> -acétylglucosaminyltransférase	UV : ultraviolet
GP : glycoprotéine	VLP : Virus Like-Particle
GuCl : chlorure de guanidine	VPg : Viral Protein linked to Genome
HBGA : antigène de groupes sanguins tissulaires	WB : western-blot
HCD : cellule de collision à haute énergie	Xyl : xylose
Hex : hexose	
HexA : acide hexuronique	

# PUBLICATION ET COMMUNICATIONS

---

## I. Publication

Auger A, Yu SY, Guu SY, Quéméner A, Euller-Nicolas G, Ando H, Desdouits M, Le Guyader FS, Khoo KH, Le Pendu J, Chirat F, Guérardel Y. **Species-Specific N-Glycomes and Methylation Patterns of Oysters *Crassostrea gigas* and *Ostrea edulis* and Their Possible Consequences for the Norovirus-HBGA Interaction**. Mar Drugs. 2023 Jun 2;21(6):342.

## II. Communications orales

Auger A, Euller G, Yu S-Y, Desdouits M, Jazat J, Le Moullac-Vaidye B, Le Pendu J, Le Guyader S.F, Biot C, Chirat F and Guérardel Y. **Structural and functional analyses of *Crassostrea gigas* glycans highlight new norovirus-oysters recognition**. 28e journées du Groupe Français des Glycosciences, 30 Mai – 3 Juin 2022, Branvilles, FR.

Prix : meilleure communication orale

Auger A, Euller G, Yu S-Y, Desdouits M, Jazat J, Le Moullac-Vaidye B, Le Pendu J, Le Guyader S.F, Chirat F and Guérardel Y. **A histo-blood group antigen appeared methylated on N-glycans from oysters *Crassostrea gigas* and *Ostrea edulis***. 21e Edition Journée André Verbert des Doctorants, 22 septembre 2021 (en ligne).

Auger A, Yu S-Y, Jazat J, Le Moullac-Vaidye B, Le Pendu J, Le Guyader S.F, Biot C, Chirat F and Guérardel Y. **A histo-blood group antigen appeared methylated on N-glycans from oysters *Crassostrea gigas* and *Ostrea edulis***. 12th edition: Young Researchers in Life sciences, 16 – 18 juin 2021 (en ligne).

## III. Communications par affiche

Auger A, Yu S-Y, Jazat J, Le Moullac-Vaidye B, Le Pendu J, Le Guyader S.F, Chirat F and Guérardel Y. **A histo-blood group antigen appeared methylated on N-glycans from oysters *Crassostrea gigas* and *Ostrea edulis***. Congrès Annuel de la Société Française de Biochimie et Biologie Moléculaire, 1– 2 juillet 2021, Paris, FR.

Auger A, Euller G, Yu S-Y, Desdouits M, Jazat J, Le Moullac-Vaidye B, Le Pendu J, Le Guyader S.F, Chirat F and Guérardel Y. **A histo-blood group antigen appeared methylated on N-glycans from oysters *Crassostrea gigas* and *Ostrea edulis***. Joint Warren and Beilstein Symposium on Glycosciences, 23 – 25 juin 2021 (en ligne).

# TABLE DES MATIERES

---

INTRODUCTION .....	30
<b>CHAPITRE 1 :      PROCESSUS DE GLYCOSYLATION .....</b>	<b>31</b>
<b>I. Rappel sur la structure des monosaccharides et oligosaccharides .....</b>	<b>31</b>
A. Structure et représentation spatiale des monosaccharides .....	31
B. Monosaccharides naturels .....	32
C. Oligosaccharides .....	34
<b>II. N-glycosylation chez les mammifères .....</b>	<b>35</b>
A. Généralités .....	35
B. Biosynthèse des N-glycannes .....	36
1. Phase réticulaire.....	36
2. Phase Golgienne .....	39
a. Cis-Golgi .....	39
b. Médian-Golgi .....	39
c. Trans-Golgi et trans-Golgi Network .....	41
<b>III. O-glycosylation de type mucine chez les mammifères .....</b>	<b>42</b>
A. Généralités .....	42
B. Biosynthèse des O-glycannes de type mucine.....	42
<b>IV. Glycosylation terminale chez les mammifères .....</b>	<b>43</b>
A. Fucosylation et antigènes de groupes sanguins ABH et Lewis .....	43
B. Sialylation .....	47
C. Sulfatation .....	47
<b>V. Particularités de la glycosylation chez les invertébrés .....</b>	<b>48</b>
A. N-glycosylation chez les invertébrés.....	48
1. Types de N-glycannes.....	48
2. Substitutions du noyau pentasaccharidique .....	49
3. Antennes .....	50
4. Modification des N-glycannes .....	51
a. Groupements zwitterioniques .....	51
b. Groupements sulfates et phosphates .....	52
c. Groupements méthyles.....	52

B. <i>O</i> -glycosylation de type mucine chez les invertébrés .....	52
<b>VI. Rôles biologiques des glycannes.....</b>	<b>53</b>
A. Rôles physico-chimiques .....	53
B. Reconnaissances spécifiques .....	54
<b>VII. Interactions hôte-virus médiées par les glycannes.....</b>	<b>55</b>
A. Acides sialiques.....	55
B. Glycosaminoglycannes.....	56
C. HBGAs.....	57
<b>CHAPITRE 2 :      NOROVIRUS.....</b>	<b>58</b>
<b>I. Présentation .....</b>	<b>58</b>
<b>II. Structure des Norovirus .....</b>	<b>58</b>
A. Génome.....	58
B. Capside virale protéique.....	59
<b>III. Classification des Norovirus .....</b>	<b>60</b>
<b>IV. Reconnaissance des HBGAs par les Norovirus .....</b>	<b>61</b>
A. Influence du statut « sécréteur ».....	61
B. Reconnaissance d'antigènes spécifiques .....	63
C. Evolution des Norovirus.....	64
1. Variants des norovirus GII.4.....	64
2. Reconnaissance fine des HBGAs.....	65
<b>V. Infections à Norovirus chez l'Homme.....</b>	<b>66</b>
A. Pathologie .....	66
B. Voies de transmission.....	67
<b>VI. Contamination des huîtres par les Norovirus.....</b>	<b>68</b>
A. Espèces d'huîtres en France .....	68
B. Biologie de l'huître.....	68
C. Vecteur de Norovirus .....	69
1. Bioaccumulation et rétention spécifique des Norovirus.....	70
2. Ligands des Norovirus chez l'huître .....	72
3. Contrôle des parcs à huître.....	73

## CHAPITRE 3 : ANALYSE DES GLYCANNES PAR SPECTROMETRIE DE MASSE ... 75

I. Libération des glycanes.....	75
II. Composition d'un spectromètre de masse .....	76
III. Analyse des <i>N</i> - et <i>O</i> -glycannes par MALDI-QIT-TOF.....	77
A. Fonctionnement du MALDI-QIT-TOF .....	77
1. Description .....	77
2. Source d'ionisation MALDI .....	78
3. Analyseur Quadrupole Ion Trap .....	79
4. Analyseur à temps de vol TOF .....	82
B. Dérivation des glycanes : perméthylation.....	83
C. Interprétation des spectres de masse.....	84
1. Spectres de masse MS .....	84
2. Spectres de masse MS <sup>n</sup> .....	88
IV. Analyse des <i>N</i> - et <i>O</i> -glycannes par GC-EI-MS .....	92
A. Fonctionnement du GC-EI-MS .....	92
1. Présentation .....	92
2. Chromatographie en phase gaz .....	92
3. Source d'ionisation électronique EI .....	93
4. Analyseur quadripôle.....	94
B. Dérivation des monosaccharides .....	96
C. Interprétation des résultats.....	98
V. Analyse des <i>N</i> -glycannes par nanoLC-ESI-MS/MS .....	100
A. Présentation du spectromètre de masse nanoLC Orbitrap Fusion™ Tribid™ .....	101
1. Description .....	101
2. Source d'ionisation ESI .....	102
3. Analyseur orbitrap .....	102
4. Fonctionnement général du spectromètre de masse nanoLC Orbitrap Fusion™ Tribid™ .....	103
B. Interprétation des résultats.....	104
RESULTATS.....	105

## CHAPITRE 1 : ANALYSE STRUCTURALE DES *N*- ET *O*-GLYCANNES OBTENUS A PARTIR DE TISSUS DES ESPECES D'HUITRE *C. gigas* ET *O. edulis*..... 106

<b>I. Méthodologie employée pour la préparation des échantillons .....</b>	<b>106</b>
A. Libération et purification des <i>N</i> - et <i>O</i> -glycannes .....	106
B. Traitements chimiques effectués sur les <i>N</i> - et <i>O</i> -glycannes .....	108
1. Analyse par MALDI-MS pour l'établissement des <i>N</i> - et <i>O</i> -glycomes des tissus des huîtres <i>C. gigas</i> et <i>O. edulis</i> .....	108
2. Analyse par nanoLC-MS/MS des <i>N</i> -glycannes perméthylés des tissus des huîtres <i>C. gigas</i> et <i>O. edulis</i> .....	109
3. Analyses par GC-EI-MS des différents dérivés monosaccharidiques issus des <i>N</i> -glycannes des tissus des huîtres <i>C. gigas</i> et <i>O. edulis</i> .....	110
<b>II. Analyse des <i>N</i>-glycannes .....</b>	<b>111</b>
A. Publication .....	111
B. Analyses complémentaires .....	142
1. Etablissement des <i>N</i> -glycomes de palpes labiaux, de l'hémolymphe, du plasma et des hémocytes de <i>C. gigas</i> .....	142
1. Essais de mise en évidence de groupements sulfates .....	146
<b>III. Analyse des <i>O</i>-glycannes.....</b>	<b>147</b>

## **CHAPITRE 2 :      ÉTUDE DES INTERACTIONS ENTRE LES NOROVIRUS ET LES GLYCANNES ISSUS DES HUITRES *C. gigas* ET *O. edulis*.....**

<b>I. Stratégies développées .....</b>	<b>154</b>
<b>II. Mise au point préalable des conditions opératoires des expériences de FarWB et d'ELISA.....</b>	<b>155</b>
A. Préparation des échantillons.....	155
1. Choix du tampon d'extraction des protéines et glycoprotéines.....	155
2. Contrôle de l'efficacité d'action de la PNGase F.....	156
B. Mise au point des différentes étapes des expériences de FarWB.....	157
1. Concentrations en lectines .....	157
2. Détermination des quantités optimales des protéines et glycoprotéines.....	158
3. Choix de l'agent de saturation des membranes de nitrocellulose.....	158
4. Résultats obtenus avec le protocole optimisé.....	158
a. Révélation avec les lectines .....	158
b. Révélation avec les VLPs.....	159

C. Mises au point pour les tests ELISA .....	160
<b>III. Reconnaissance des ligands glycaniques de l'huître par les Norovirus : utilisation de VLPs.....</b>	<b>161</b>
A. Etudes des reconnaissances par les VLPs de Norovirus GI.1.....	161
B. Etudes des reconnaissances par les VLPs de Norovirus GII.4 Dijon.....	163
C. Etudes des reconnaissances par les VLPs de Norovirus GII.4 MD145 .....	164
<b>DISCUSSION.....</b>	<b>166</b>
<b>I. La glycosylation des huîtres <i>C. gigas</i> et <i>O. edulis</i> parmi les invertébrés .....</b>	<b>167</b>
A. La <i>N</i> - et <i>O</i> -glycosylation.....	167
B. La méthylation .....	169
<b>II. Interactions Norovirus-huître .....</b>	<b>171</b>
A. Interaction médiée par des ligands glycaniques .....	171
B. Influence des HBGAs méthylés sur les Norovirus .....	173
1. « Tropicité » d'espèce .....	174
2. Influence selon les souches virales : modélisation moléculaire .....	175
C. « Tropicité » tissulaire .....	176
<b>CONCLUSION &amp; PERSPECTIVES .....</b>	<b>177</b>
<b>MATERIEL &amp; METHODES.....</b>	<b>181</b>
I. Echantillons .....	182
II. Préparation des <i>N</i> -glycannes.....	182
III. Préparation des <i>O</i> -glycannes.....	183
IV. Dérivation des <i>N</i> - et <i>O</i> -glycannes pour l'analyse en MALDI-QIT-TOF .....	184
V. Purification des <i>N</i> -glycannes sulfatés.....	184
VI. Dérivation des <i>N</i> -glycannes à la DAP.....	185
VII. Analyse par MALDI-QIT-TOF .....	185
VIII. Préparation des itols.....	185
IX. Peracétylation et perméthylation des monosaccharides sous forme itols.....	186
X. Analyse des itols-acétates et itols-perméthylés par GC-MS.....	186
XI. Quantification des antigènes A et ALe <sup>b</sup> par nanoLC-MS/MS .....	186



XII. Préparation des échantillons pour WB et ELISA .....	187
XIII. SDS-PAGE et transfert sur membrane de nitrocellulose .....	187
XIV. Lectine blot .....	188
XV. Far-Western blot.....	188
XVI. ELISA.....	188
<b>BIBLIOGRAPHIE.....</b>	<b>190</b>
<b>ANNEXE.....</b>	<b>209</b>
<b>PERMISSION DE REPRODUCTION DES FIGURES.....</b>	<b>226</b>

# TABLE DES FIGURES

---

Figure 1 : Représentations de Fischer et de Haworth du glucose.....	31
Figure 2 : Monosaccharides communément retrouvés dans les <i>N</i> - et <i>O</i> -glycannes des vertébrés .....	33
Figure 3 : Types de <i>N</i> -glycannes retrouvés chez les mammifères .....	35
Figure 4 : Synthèse du précurseur Dolichol-P-P-GlcNAc <sub>2</sub> Man <sub>9</sub> Glc <sub>3</sub> des <i>N</i> -glycannes dans le réticulum endoplasmique.....	36
Figure 5 : Contrôle qualité du repliement des glycoprotéines nouvellement synthétisées.....	38
Figure 6 : Maturation golgienne des <i>N</i> -glycannes .....	40
Figure 7 : Initiation des antennes et substitution du noyau des <i>N</i> -glycannes de type complexe.....	41
Figure 8 : Unités composant les antennes des <i>N</i> -glycannes.....	41
Figure 9 : Noyaux majoritaires des <i>O</i> -glycannes de type mucine retrouvés chez les mammifères .....	42
Figure 10 : Antigènes de groupes sanguins.....	46
Figure 11 : Représentation des acides 5- <i>N</i> -acétylneuraminique (Neu5Ac), 5- <i>N</i> -glycolylneuraminique (Neu5Gc) et 3-déoxy-D-glycero-D-galacto-non-2-ulosonique (Kdn).....	47
Figure 12 : Biosynthèse des <i>N</i> -glycannes pseudohybride, paucimannose et paucimannose tronqué ...	49
Figure 13 : Formules chimiques développées des groupements zwitterioniques phosphocholine, phosphoéthanolamine, aminoéthylphosphonate et méthylaminoéthylphosphonate.....	51
Figure 14 : Noyaux des <i>O</i> -glycannes de type mucine retrouvés chez les invertébrés.....	53
Figure 15: Organisation du génome des norovirus humains .....	59
Figure 16 : Organisation de la protéine VP1 dans la capsidie protéique des norovirus .....	60
Figure 17 : Classification des norovirus basée sur la phylogénie de la protéine VP1 et leurs cibles animales.....	61
Figure 18: Emergence des variants GII.4 de 1995 à 2017 .....	64
Figure 19: Evolution des reconnaissances des HBGAs par les variants GII.4 de 1987 à 2011 .....	65
Figure 20 : Illustration des voies de transmission des Norovirus à l'homme.....	67
Figure 21: Anatomie et mécanisme de nutrition de l'huître.....	69

Figure 22 : Bioaccumulation des norovirus GI.1, GII.3 et GII.4 dans différents tissus de l’huître <i>Crassostrea gigas</i> .....	70
Figure 23 : Inhibition de liaison des VLPs de norovirus sur des extraits tissulaires de l’huître <i>Crassostrea gigas</i> .....	71
Figure 24 : Libération des <i>N</i> - et <i>O</i> -glycannes .....	76
Figure 25 : Fonctionnement de la source d’ionisation MALDI .....	78
Figure 26 : Acide 2,5-dihydroxybenzoïque.....	79
Figure 27 : Fonctionnement d'une trappe à ions (QIT) .....	80
Figure 28 : Zones de stabilité selon les courants U et V dans le Quadrupole Ion Trap (QIT) .....	81
Figure 29 : Fonctionnement d'un analyseur TOF en mode réflectron.....	83
Figure 30: Schéma explicatif du calcul des masses moléculaires des <i>N</i> -glycannes non réduits et <i>O</i> -glycannes réduits perméthylés observés sur un spectre de masse.....	86
Figure 31 : Nomenclature des différents fragments générés par MS/MS selon Domon et Costello (1988) .....	88
Figure 32 : Mécanismes de fragmentation des oligosaccharides perméthylés conduisant aux fragments B et Y, C et Z et A et X par spectrométrie de masse MS <sup>n</sup> .....	89
Figure 33 : Localisation des fragmentations majoritaires observées pour les <i>N</i> - et <i>O</i> -glycannes perméthylés pour des analyses MS <sup>2</sup> par MALDI-QIT-TOF .....	92
Figure 34 : Schéma d’une source d’ionisation électronique.....	94
Figure 35 : Schéma d’un quadrupôle .....	94
Figure 36 : Zones de stabilité selon les courants U et V dans le quadrupôle .....	95
Figure 37 : Traitement chimique des oligosaccharides pour l’obtention d’alditols peracétylés, d’alditols perméthylés et d’alditols acétates partiellement méthylé.....	97
Figure 38 : Réduction des monosaccharides en leur alditol correspondant par le borodutérure de sodium (NaBD <sub>4</sub> ) ou le borohydrure de sodium (NaBH <sub>4</sub> ) .....	99
Figure 39 : Mécanismes de fragmentation des itols acétates, itols perméthylés et itols acétates partiellement méthylés en EI-MS. des itols perméthylés .....	100
Figure 40 : Composition d’un spectromètre de masse Orbitrap Fusion Tribrid.....	101

Figure 41 : Composition d'un orbitrap.....	103
Figure 42 : Méthodologie employée pour les analyses structurales des <i>N</i> - et <i>O</i> -glycannes obtenus à partir des branchies, tubes digestif et manteaux des huîtres <i>C. gigas</i> et <i>O. edulis</i> .....	107
Figure 43 : Analyse MALDI-MS des <i>N</i> -glycannes perméthylés isolés des palpes labiaux de <i>Crassostrea gigas</i> .....	143
Figure 44 : Analyse MALDI-MS des <i>N</i> -glycannes libérés par la peptidyl- <i>N</i> -glycosidase F puis perméthylés à partir de l'hémolymphe, du plasma et des hémocytes de <i>Crassostrea gigas</i> .....	144
Figure 45 : Proportions des familles et des antennes de <i>N</i> -glycannes retrouvés sur les glycoprotéines des branchies, du tube digestif, du manteau, des palpes labiaux, de l'hémolymphe, du plasma et des hémocytes de <i>Crassostrea gigas</i> .....	145
Figure 46 : Analyse MALDI-MS des <i>N</i> -glycannes libérés par la peptidyl- <i>N</i> -glycosidase F puis dérivés à la diaminopyridine à partir du manteau de <i>Crassostrea gigas</i> .....	147
Figure 47 : Coloration et Slot-blot réalisés sur les fractions des glycoprotéines du tube digestif de l'huître <i>Crassostrea gigas</i> obtenu après ultracentrifugation.....	149
Figure 48 : Spectres de masse MALDI-MS, MS <sup>2</sup> et MS <sup>3</sup> des <i>O</i> -glycannes réduits et perméthylés obtenus à partir du tube digestif de l'huître <i>Crassostrea gigas</i> .....	153
Figure 49 : Stratégie employée pour l'étude des interactions entre les Norovirus (NoV) GI.1, GII.4 MD145 et GII.4 Dijon et les glycoprotéines des 3 tissus des huîtres <i>Crassostrea gigas</i> et <i>Ostrea edulis</i> .....	154
Figure 50 : Evaluation du protocole d'élimination des <i>N</i> -glycannes par la PNGase F.....	157
Figure 51 : Lectine blots des glycoprotéines des huîtres <i>Crassostrea gigas</i> et <i>Ostrea edulis</i> .....	159
Figure 52 : Tests ELISA utilisant des VLPs à des concentrations croissantes en protéines du tube digestif des huîtres <i>Crassostrea gigas</i> et <i>Ostrea edulis</i> .....	161
Figure 53 : Far-western blot des protéines et glycoprotéines des huîtres <i>Crassostrea gigas</i> et <i>Ostrea edulis</i> avec les VLPs GI.1 comme sonde.....	162
Figure 54 : Reconnaissance des glycoprotéines et protéines des tissus de <i>Crassostrea gigas</i> et <i>Ostrea edulis</i> par les VLPs GI.1 avant et après traitement par la PNGase F.....	163

Figure 55 : Far-western blot des protéines et glycoprotéines des huîtres <i>Crassostrea gigas</i> et <i>Ostrea edulis</i> avec les VLPs GII.4 Dijon comme sonde .....	164
Figure 56 : Reconnaissance des glycoprotéines et protéines des tissus de <i>Crassostrea gigas</i> par les VLPs GII.4 MD145 avant et après traitement par la PNGase F.....	165

# TABLE DES TABLEAUX

---

Tableau 1 : Nomenclature pour les glycannes (Symbol Nomenclature for Glycans : SNFG) .....	34
Tableau 2 : Inhibition de la liaison de VLPs GI.1 de Norwalk sur le tissu digestif par de la salive .....	73
Tableau 3 : Masse en Dalton des monosaccharides et des extrémités réductrice et non réductrice des oligosaccharides pour le calcul de la masse moléculaire des oligosaccharides natifs, perméthylés et perdeutérométhylés. ....	86
Tableau 4 : Liste des structures <i>O</i> -glycanniques putatives identifiées par spectrométrie de masse MALDI-QIT-TOF dans les branchies, le tube digestif et le manteau des huîtres <i>Crassostrea gigas</i> et <i>Ostrea edulis</i> .....	150

# AVANT-PROPOS

---

Les glucides font partie des quatre biomolécules essentielles pour la vie avec les protéines, les lipides et les acides nucléiques. Les unités de base, nommés les monosaccharides, oses ou encore les hydrates de carbone, s'associent entre elles pour former des oligosaccharides et polysaccharides. Portés par des protéines, ils définissent les glycoprotéines. Selon le mode de liaison entre l'oligosaccharide (que l'on nomme glycane sous leur forme liée) et la protéine, on distingue essentiellement les *N*- et les *O*-glycannes. Leur biosynthèse a lieu dans le réticulum endoplasmique et l'appareil de Golgi pour les *N*-glycannes et l'appareil de Golgi pour les *O*-glycannes. Elle fait intervenir une multitude d'enzymes, glycosidases et glycosyltransférases, pour aboutir à des structures très variées dépendantes de l'espèce. Cette variété de structures dont les différences reposent sur le type de monosaccharides, le type de liaison mais aussi au niveau de modifications supplémentaires associées à ces monosaccharides va définir des antigènes impliqués dans des interactions entre les cellules mais également avec les pathogènes. Pour ces dernières, les Norovirus sont bien connus pour reconnaître les antigènes de groupes sanguins tissulaires (HBGAs) au niveau de la cellule hôte et sont impliqués dans de nombreuses épidémies de gastro-entérites qui causent jusqu'à 200 000 morts par an dans le monde (World Health Organization, 2022). La transmission est féco-orale entre individus mais est également liée à la consommation d'aliments contaminés dont les huîtres sont de très bons vecteurs de contamination humaine. Bien que cette contamination puisse être expliquée par leur consommation crue, leur culture à proximité des effluents humains ainsi que leur mode d'alimentation par filtration, cela n'explique pas que l'on retrouve encore des particules virales chez les huîtres contaminées malgré plusieurs étapes d'épuration dans de l'eau propre avant leur commercialisation. Il est apparu qu'elles bioaccumulent spécifiquement les Norovirus et les retiennent sur plusieurs semaines mais ne développent aucune pathologie liée à cette contamination. L'explication avancée dans la littérature par des études de reconnaissance utilisant des anticorps et lectines comme sondes est que les Norovirus se fixent au niveau de motifs HBGAs, notamment les antigènes A et H. La caractérisation structurale fine de ces antigènes est une étape obligatoire pour déterminer précisément les ligands glycaniques des Norovirus chez l'huître et comprendre cette contamination difficilement évitable actuellement.

C'est dans le cadre de la tâche « Characterization and localization of virus binding sites in oyster tissues » de l'ANR GOyAVE (« Glycan and Oysters Attachment to enteric Viruses in the coastal Environment ») pilotée par la Dr Soizick Le Guyader (Ifremer, Nantes) (ANR AAPG2019) que l'étude de ces glycanes chez les huîtres *Crassostrea gigas* et *Ostrea edulis* a été planifiée pour être effectuée par notre laboratoire. Cette tâche a fait l'objet de mon principal objectif de thèse qui a concerné l'identification des antigènes glycaniques présents dans les branchies, manteau et tissus digestif des huîtres *C. gigas* et *O. edulis* dans le but d'expliquer les différences de reconnaissance entre les souches de Norovirus, les espèces et les tissus selon les ligands identifiés. Pour cela, nous avons cherché à établir précisément les *N*- et *O*-glycomes de ces tissus des deux espèces d'huître. En parallèle, pour aller au-delà de cette caractérisation structurale, mon deuxième objectif a été d'identifier plus précisément les macromolécules impliquées dans les reconnaissances avec les Norovirus en essayant d'identifier une glycoprotéine et/ou un type de glycanne par l'utilisation de méthodes ELISA et Far-Western blot.

Ainsi, ce manuscrit est composé d'une première partie intitulée « Introduction » qui se décompose en 3 chapitres qui visent à présenter l'état de l'art concernant :

- La biosynthèse des *N*- et *O*-glycanes chez les mammifères et les invertébrés dans le but de mettre en lumière la grande variété de glycanes retrouvés en fonction des espèces et leur implication dans les interactions avec les virus.
- Une présentation des Norovirus et de leurs spécificités de reconnaissance des HBGAs chez l'homme et l'huître.
- Les méthodes physico-chimiques employées durant ma thèse pour établir la structure des *N*- et *O*-glycanes.

La deuxième partie du manuscrit intitulée « Résultats » présente les résultats obtenus durant mes travaux de thèse et est divisée en deux chapitres. Le premier concerne l'analyse structurale des *N*- et *O*-glycomes et qui a conduit à la publication « Species-Specific N-Glycomes and Methylation Patterns of Oysters *Crassostrea gigas* and *Ostrea edulis* and Their Possible Consequences for the Norovirus-HBGA Interaction » parue cette année. Le deuxième chapitre concerne les études d'interaction que nous



avons effectuées sur les différents tissus d'huître à l'aide des virus-like particles de Norovirus dans le but d'identifier plus précisément les ligands impliqués dans la reconnaissance.

Enfin, les parties « Discussion » et « Conclusion et perspectives » visent à mettre en perspective nos résultats avec les connaissances actuelles et établir des perspectives.

# INTRODUCTION

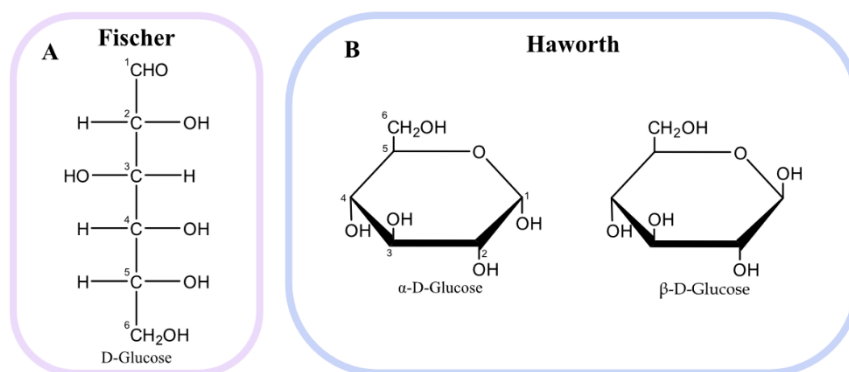
---

# Chapitre 1 : PROCESSUS DE GLYCOSYLATION

## I. Rappel sur la structure des monosaccharides et oligosaccharides

### A. Structure et représentation spatiale des monosaccharides

Les monosaccharides sont des composés de formule brute  $C_n(H_2O)_n$  (avec  $n$  entre 3 et 9) d'où leur ancienne appellation d'hydrates de carbone. Ils sont caractérisés par la présence d'une fonction aldéhyde pour les aldoses ou cétone pour les cétooses parmi plusieurs fonctions alcool. Parmi eux, les monosaccharides à 5 carbones, appelés pentoses et à 6 carbones, appelés hexoses, sont quantitativement majoritaires. On les rencontre sous deux formes : linéaire et cyclisée. La première est représentée par la projection de Fischer (**Figure 1A**) avec la fonction aldéhyde sur le carbone 1 (C1) et la fonction cétone sur le C2 selon les règles de Fischer. De plus, l'hydroxyle de l'avant dernier carbone est représenté à droite pour la série D et à gauche pour la série L. Dans la nature, la majorité des monosaccharides sont de la série D avec comme exception courante le L-fucose (L-Fuc). En milieu aqueux, les monosaccharides forment un cycle à 5 ou 6 sommets que l'on nomme respectivement furanose (en référence à la structure du furane) et pyranose (en référence à la structure du pyrane) qui sont en équilibre avec la forme linéaire qui est très minoritaire (environ 1 %). Le groupement carbonyle s'hydrate pour former un acétal dans le cas d'un aldose ou un cétal dans le cas d'une cétoose. Un hémiacétal est formé lors de la cyclisation via la réaction d'un hydroxyle alcoolique avec un hydroxyle acétalique du carbone 1 ou cétalique du carbone 2. La formation de l'hémiacétal ou de l'hémicétal fait apparaître un carbone asymétrique supplémentaire qui conduira à 2 énantiomères appelés anomères et nommés  $\alpha$  ou  $\beta$ . On le



**Figure 1** : Représentations de (A) Fischer et de (B) Haworth du glucose.

représente par la projection de Haworth (**Figure 1B**). Dans la série D, l'hydroxyle anomérique s'écrit en dessous du plan formé par les carbones 2, 3, 5 et l'oxygène du pont oxhydrique pour l'anomère  $\alpha$  ou au-dessus pour celui  $\beta$ .

## B. Monosaccharides naturels

Dans la nature, neuf monosaccharides composent couramment les *N*- et *O*-glycannes des animaux que l'on classe en oses neutres, en osamines, en oses acides, désoxyose et acides sialiques (**Figure 2**).

Pour les oses neutres qui répondent strictement à la formule brute  $C_n(H_2O)_n$  donnée en début de chapitre, il s'agit du glucose (Glc), du galactose (Gal), du mannose (Man) et du xylose (Xyl).

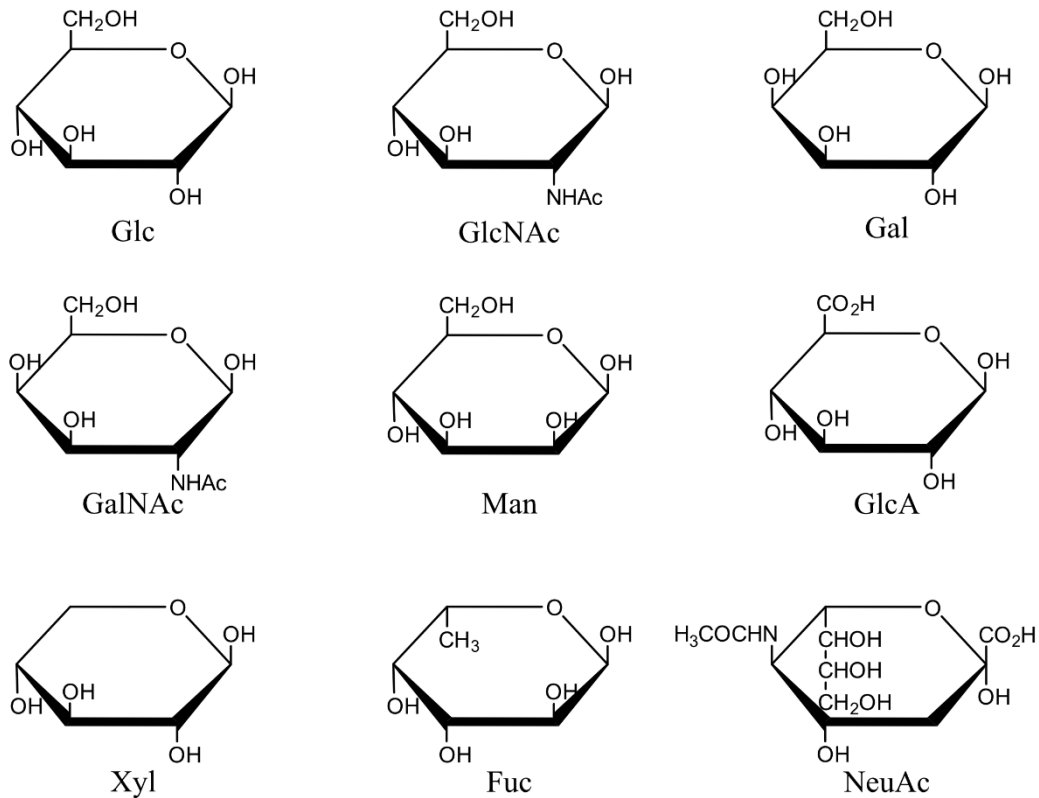
Pour les osamines, le groupement hydroxyle du C2 est remplacé par un groupement amine. Ce groupement amine apparaît très souvent sous forme substituée par un groupement acétyle. Les 2 osamines *N*-acétylées communément retrouvées sont la *N*-acétylglucosamine (GlcNAc) et la *N*-acétylgalactosamine (GalNAc).

Pour les oses acides, on en distingue 3 types. Les acides uroniques pour lesquels la fonction alcool primaire est oxydée en groupement carboxylique, les acides oniques pour lesquels la fonction aldéhyde est oxydée en acide carboxylique et les acides ariques pour lesquels les fonctions alcool primaire et aldéhyde sont oxydées en groupement carboxylique. Parmi ces 3 types d'oses acides, on retrouve très souvent les acides uroniques avec l'acide glucuronique comme représentant majoritaire.

Pour les désoxyoses, il s'agit d'oses pour lesquels un groupement hydroxyle est remplacé par un hydrogène. Pour les *N*- et *O*-glycannes, le représentant majeur est le L-Fuc (6-désoxy-L-Gal).

Enfin, les acides sialiques qui constituent une vaste famille de monosaccharides à 9 carbones possèdent un groupement carboxylique en C1, une fonction cétone en C2, une fonction amine ou hydroxyle en C5 et des groupements hydroxyles sur les C4 et C6 à C9. Une cinquantaine d'acides sialiques ont été décrits selon la nature des groupements qui peuvent substituer les hydroxyles alcooliques ou sur la fonction amine du C5. L'acide 5-*N*-acétylneuraminique et l'acide 5-*N*-glycolylneuraminique sont les 2 acides sialiques les plus fréquemment retrouvés.

Tout comme les acides aminés qui peuvent être modifiés par ajout de divers groupements (processus connu sous le nom de modification post-traductionnelle), d'autres groupements peuvent être ajoutés sur les monosaccharides par des enzymes comme des groupements sulfates et phosphates ou encore méthyles chez les invertébrés.



**Figure 2 :** Monosaccharides communément retrouvés dans les *N*- et *O*-glycannes des vertébrés. Glc : D-Glucose ; GlcNAc : *N*-Acétyl-D-glucosamine ; Gal : D-Galactose ; GalNAc : *N*-acétyl-D-galactosamine ; Man : D-Mannose ; GlcA : Acide D-glucuronique ; Xyl : D-Xylose ; Fuc : L-Fucose ; NeuAc : Acide *N*-acétylneuraminique.

La nomenclature « Symbol Nomenclature for Glycans » (SNFG) a été établie par Varki *et al.*, (1999) pour simplifier et rendre homogène la représentation des monosaccharides et a été améliorée au fil des années (**Tableau 1**) (Neelamegham *et al.*, 2019 ; Varki *et al.*, 2015).

**Tableau 1** : Nomenclature pour les glycannes (Symbol Nomenclature for Glycans : SNFG). Issue de Neelamegham *et al.* (2019).

SHAPE	White	Blue	Green	Yellow	Orange	Pink	Purple	Light Blue	Brown	Red
Filled Circle	Hexose ○	Glc ●	Man ●	Gal ●	Gul ●	Alt ●	All ●	Tal ●	Ido ●	
Filled Square	HexNAc □	GlcNAc ■	ManNAc ■	GalNAc ■	GulNAc ■	AltNAc ■	AllNAc ■	TalNAc ■	IdoNAc ■	
Crossed Square	Hexosamine ◻	GlcN ◻	ManN ◻	GalN ◻	GulN ◻	AltN ◻	AllN ◻	TalN ◻	IdoN ◻	
Divided Diamond	Hexuronate ◊	GlcA ◊	ManA ◊	GalA ◊	GulA ◊	AltA ◊	AllA ◊	TalA ◊	IdoA ◊	
Filled Triangle	Deoxyhexose △	Qui ▲	Rha ▲		6dGul ▲	6dAlt ▲		6dTal ▲		Fuc ▲
Divided Triangle	DeoxyhexNAc ◡	QuiNAc ◡	RhaNAc ◡			6dAltNAc ◡		6dTalNAc ◡		FucNAc ◡
Flat Rectangle	Di-deoxyhexose ▭	Oli ▭	Tyv ▭		Abe ▭	Par ▭	Dig ▭	Col ▭		
Filled Star	Pentose ☆		Ara ★	Lyx ★	Xyl ★	Rib ★				
Filled Diamond	Deoxynonulosonate ◊		Kdn ◆				Neu5Ac ◆	Neu5Gc ◆	Neu ◆	Sia ◆
Flat Diamond	Di-deoxynonulosonate ◊		Pse ◆	Leg ◆		Aci ◆		4eLeg ◆		
Flat Hexagon	Unknown ⬡	Bac ⬡	LDManHep ⬡	Kdo ⬡	Dha ⬡	DDManHep ⬡	MurNAc ⬡	MurNGc ⬡	Mur ⬡	
Pentagon	Assigned ⬠	Api ⬠	Fruc ⬠	Tag ⬠	Sor ⬠	Psi ⬠				

### C. Oligosaccharides

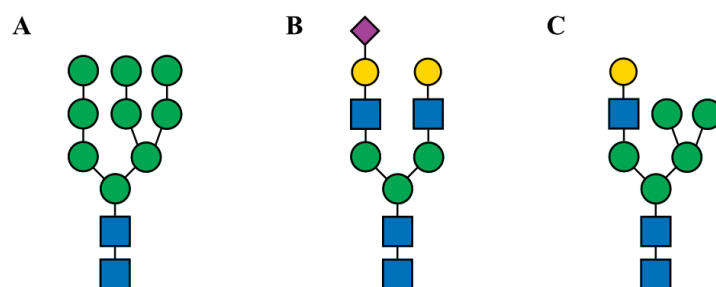
Les monosaccharides se lient entre eux par des liaisons glycosidiques pour former des oligosaccharides (de 2 à 20 unités) et des polysaccharides (> 20 unités). Cette liaison s'effectue entre l'hydroxyle hémiacétalique ou cétalique d'un monosaccharide et un hydroxyle alcoolique ou hémiacétalique ou hémicétalique d'un autre monosaccharide pour former un acétal. Une fois liée, l'anométrie  $\alpha$  ou  $\beta$  de la liaison est fixée. Si l'on considère la série D ou L de 2 monosaccharides qui se lient, le cycle furane ou pyrane qu'ils adoptent, la position de liaison et l'anométrie  $\alpha$  ou  $\beta$  de cette dernière, il est possible d'écrire 632 disaccharides différents. A l'inverse des acides nucléiques et des protéines qui sont des molécules linéaires, un même monosaccharide peut établir plusieurs liaisons osidiques et ainsi créer des ramifications contribuant à enrichir davantage la diversité structurale des chaînes oligosaccharidiques. Enfin, comme pour les acides nucléiques et les protéines, les oligosaccharides ont une polarité définie par une extrémité réductrice (hydroxyle anomérique libre) et une extrémité non réductrice. En plus de se lier entre eux, les monosaccharides et donc aussi les oligosaccharides, peuvent se lier à des biomolécules comme les lipides ou les protéines via leur

hydroxyle anomérique : ce processus est connu sous le nom de glycosylation. Sous sa forme liée, l'oligosaccharide est nommé glycanne et la molécule résultante est un glycoconjugué. La glycosylation constitue la modification co/post-traductionnelle des protéines la plus répandue : on estime que plus de 50 % des protéines sont glycosylées et elle est générée sur des sites particuliers de la séquence protéique (Apweiler, 1999). Parmi les multiples types de glycosylations, les 2 plus communément répandus sont la *N*- et la *O*-glycosylation.

## II. *N*-glycosylation chez les mammifères

### A. Généralités

La *N*-glycosylation est une modification co-traductionnelle des protéines caractérisée par la présence d'un glycanne sur un résidu d'asparagine (Asn) (Johansen *et al.*, 1961 ; Marks *et al.*, 1963 ; Marshall et Neuberger, 1964) inclus dans une séquence consensus Asn-X-Ser/Thr (Marshall, 1972) (avec X différent d'une Proline) nommée séquon. On distingue trois types de *N*-glycannes tous bâtis autour d'une structure pentasaccharidique commune composée de 2 résidus de GlcNAc et 3 résidus de Man, également appelée noyau. Pour le type oligomannosidique, le noyau est prolongé uniquement par des résidus de Man (**Figure 3A**). En revanche, le type complexe dispose du noyau substitué par des résidus de GlcNAc qui pourront initier des antennes et notamment des antennes *N*-acétyllactosaminiques (LacNAc) composées d'un résidu de Gal lié en  $\beta$ 1-4 d'un résidu de GlcNAc (**Figure 3B**). Le dernier type de *N*-glycanne, appelé hybride, partage des caractéristiques des deux types précédents. En effet, pour ces glycannes le Man terminal du noyau lié en  $\alpha$ 1-3 contient un résidu de GlcNAc caractéristique des *N*-glycannes complexes tandis que le Man lié en  $\alpha$ 1-6 est substitué par des résidus de Man caractéristique du type oligomannosidique (**Figure 3C**).



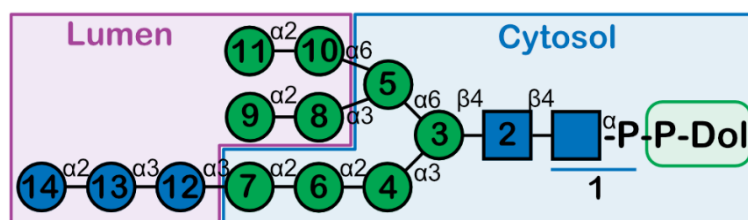
**Figure 3 :** Types de *N*-glycannes retrouvés chez les mammifères. (A) Type oligomannosidique, (B) type complexe et (C) type hybride. Les *N*-glycannes sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019).

Quel que soit le type de *N*-glycane, la biosynthèse débute dans le réticulum endoplasmique (RE) et se poursuit dans l'appareil de Golgi par l'action successive de glycosyltransférases utilisant des monosaccharides activés sous forme de nucléotide-sucres et par l'action de glycosidases (Kornfeld et Kornfeld, 1985). Si la synthèse du précurseur apparaît relativement conservée chez les animaux, les champignons et les plantes, la maturation golgienne est plus variable selon les espèces. Ce qui est décrit ci-dessous concerne le processus général de synthèse chez les mammifères.

## B. Biosynthèse des *N*-glycannes

### 1. Phase réticulaire

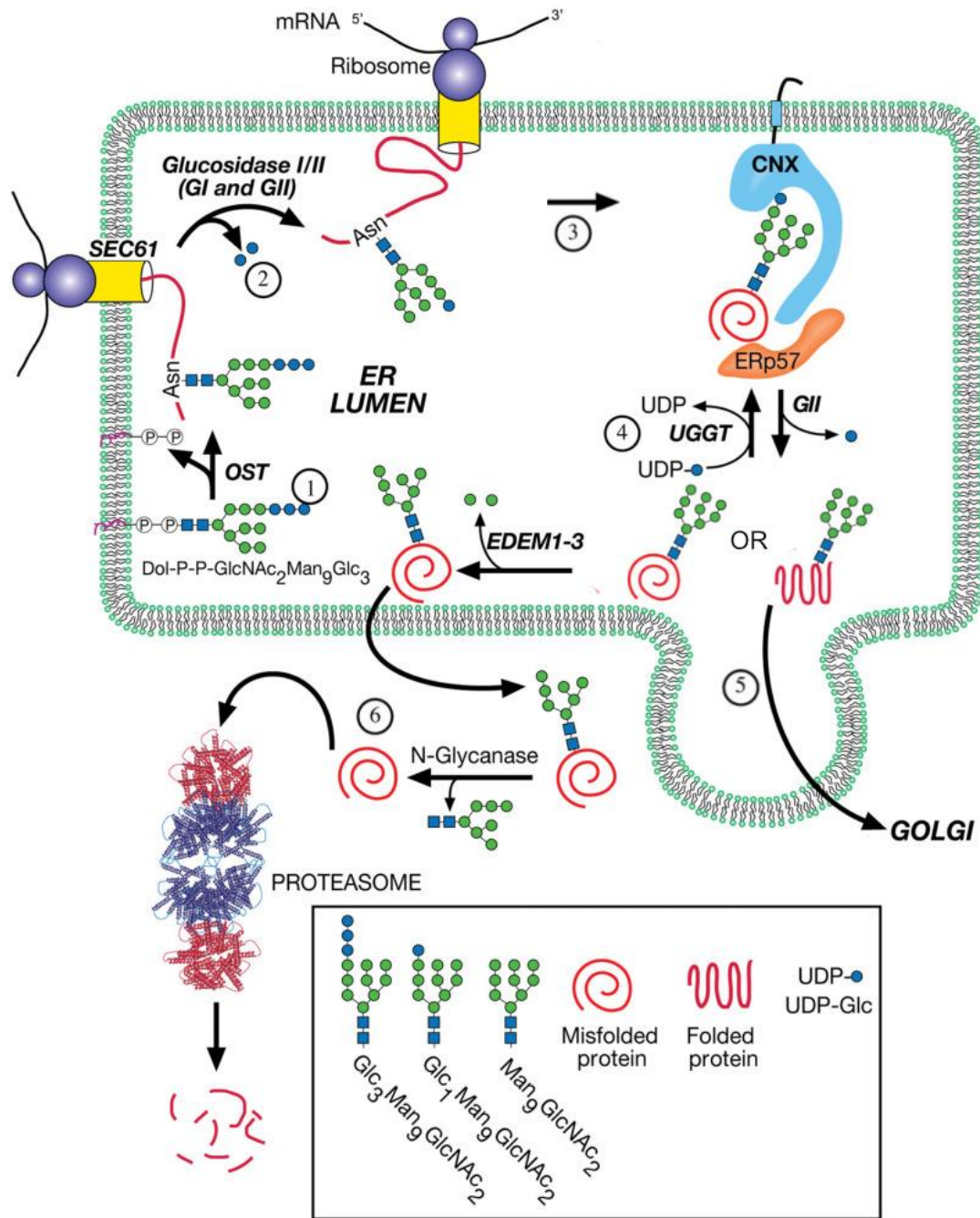
La *N*-glycosylation est obtenue par le transfert *en bloc* du tétradéca-saccharide  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$  porté par un précurseur sur le séquen d'une protéine. La synthèse de ce précurseur débute dans le RE, côté cytosolique, par l'ajout d'un GlcNAc-Phosphate sur le dolichol phosphate (Dol-P) pour former le GlcNAc-P-P-Dol (Schenk *et al.*, 2001) (**Figure 4**). Ce dernier est séquentiellement étendu par un autre résidu de GlcNAc et 5 résidus de Man pour conduire au  $\text{Man}_5\text{GlcNAc}_2\text{-P-P-Dol}$  qui sera transloqué dans la lumière du RE par une flippase. L'ajout de 4 autres résidus de Man et de 3 résidus de Glc permet l'obtention du précurseur final : le  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-P-P-Dol}$ . Les monosaccharides ajoutés le sont sous une forme activée à savoir l'UDP-GlcNAc, le GDP-Man et l'UDP-Glc. L'oligosaccharide nouvellement synthétisé est ensuite transféré *en bloc* du dolichol à la protéine en cours de synthèse par l'oligosaccharyltransférase (OST) (Burda et Aebi, 1999).



**Figure 4** : Synthèse du précurseur Dolichol-P-P-GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>3</sub> des *N*-glycannes dans le réticulum endoplasmique. Au niveau de la face cytoplasmique de la membrane du réticulum endoplasmique, un GlcNAc-Phosphate est transféré sur le Dolichol-P (Dol-P). Après ajout d'autres monosaccharides (2-7), le Dol-P-P-GlcNAc devient le Dol-P-P-GlcNAc<sub>2</sub>Man<sub>5</sub> puis est transloqué dans la lumière du RE par une flippase. La molécule est ensuite étendue par 4 résidus de Man (8-11) et 3 résidus de Glc (12-14). Le précurseur est transféré sur l'Asn d'un séquen Asn-X-Ser/Thr d'une protéine en cours de traduction par oligosaccharyltransférase (OST). Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019).



Puis, les  $\alpha$ -glucosidases I et II (GI et GII nommées MOGS et GANAB) clivent séquentiellement les 2 résidus de Glc terminaux liés respectivement en  $\alpha$ 1-2 et  $\alpha$ 1-3 pour aboutir au *N*-glycane  $\text{Glc}_1\text{Man}_9\text{GlcNAc}_2$ . Le glycane qui en résulte sera à la base d'un contrôle qualité exercé par la cellule pour s'assurer que seules les protéines correctement conformées accèdent à l'appareil de Golgi. Ce processus est connu sous le nom d'ERAD pour Endoplasmic Reticulum Associated Degradation (**Figure 5**) (Bonifacino et Weissman, 1998 ; Lederkremer, 2009). La glycoprotéine (GP) est prise en charge par la calnexine (CNX) et/ou la calréticuline (CRT) qui sont 2 lectines résidentes du RE impliquées dans l'ERAD (Booth et Koch, 1989). Ces 2 lectines, une fois liée à une GP mal conformée, vont permettre le recrutement de la protéine chaperonne Erp 57 qui tentera de corriger la conformation défectueuse de la GP néosynthétisée (Suzuki *et al.*, 2022). Cette dernière est ensuite libérée de ce complexe grâce à la GII qui élimine le dernier résidu de Glc lié en  $\alpha$ 1-3. Si la GP qui porte le glycane  $\text{Man}_9\text{GlcNAc}_2$  n'est toujours pas correctement repliée, elle est prise en charge par l'UDP-glucose:glycoprotein glucosyltransférase (UGGT) qui ajoutera un résidu de Glc sur le *N*-glycane  $\text{Man}_9\text{GlcNAc}$ . Cela permettra ainsi une nouvelle interaction de la GP néo-synthétisée avec la CNX et/ou la CRT. Ce cycle a lieu jusqu'au repliement correct de la GP. Néanmoins, après plusieurs tentatives infructueuses de remise en conformation de la GP, le *N*-glycane des GPs mal conformées subira l'action de mannosidases du système EDEM (ERAD-enhancing- $\alpha$ -mannosidases degradation) qui élimine 1 ou 2 résidus de Man et empêchera ainsi toute nouvelle action de l'UGGT. La GP mal conformée sera alors envoyée au protéasome qui se trouve dans le cytoplasme pour y être détruite. Quant aux GPs correctement conformées, elles subissent l'action de l'ER mannosidase I (MAN1B1) qui clive la liaison  $\alpha$ 1-2Man du bras central et sont ensuite envoyées dans l'appareil de Golgi (Moremen *et al.*, 1994 ; Suzuki *et al.*, 2022). La majorité des GPs qui aboutissent à l'appareil de Golgi arborent un *N*-glycane  $\text{Man}_{8,9}\text{GlcNAc}_2$ .



**Figure 5 :** Contrôle qualité du repliement des glycoprotéines nouvellement synthétisées. (1) Les protéines en cours de synthèse sont *N*-glycosylées par l'oligosaccharyltransférase (OST) dans le réticulum endoplasmique en entrant par le translocon SEC61. (2) Après synthèse complète de la glycoprotéine, les  $\alpha$ -glucosidases I et II (GI et GII) clivent les 2 résidus de Glc terminaux puis (3) les *N*-glycannes monoglucosylés sont reconnus par la calnexine (CNX) et/ou la calréticuline (CRT) associés à ERp57. (4) La glycoprotéine (GP) est libérée du complexe par la GII en éliminant le dernier résidu de Glc et repris en charge par l'action de l'UDP-glucose:glycoprotein glucosyltransférase (UGGT) qui ajoutera, à nouveau, un résidu de Glc. Ce cycle a lieu jusqu'au repliement correct de la GP et sa libération par GII. (5) la GP est ensuite transportée vers l'appareil de Golgi. (6) Les GPs mal conformées sont prises en charge par des mannosidases de EDEM (ERAD-enhancing- $\alpha$ -mannosidases dégradation), déglycosylés dans le cytoplasme puis envoyées au protéasome pour être dégradées. Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019). Adaptée de Suzuki *et al.* (2022).

## 2. Phase Golgienne

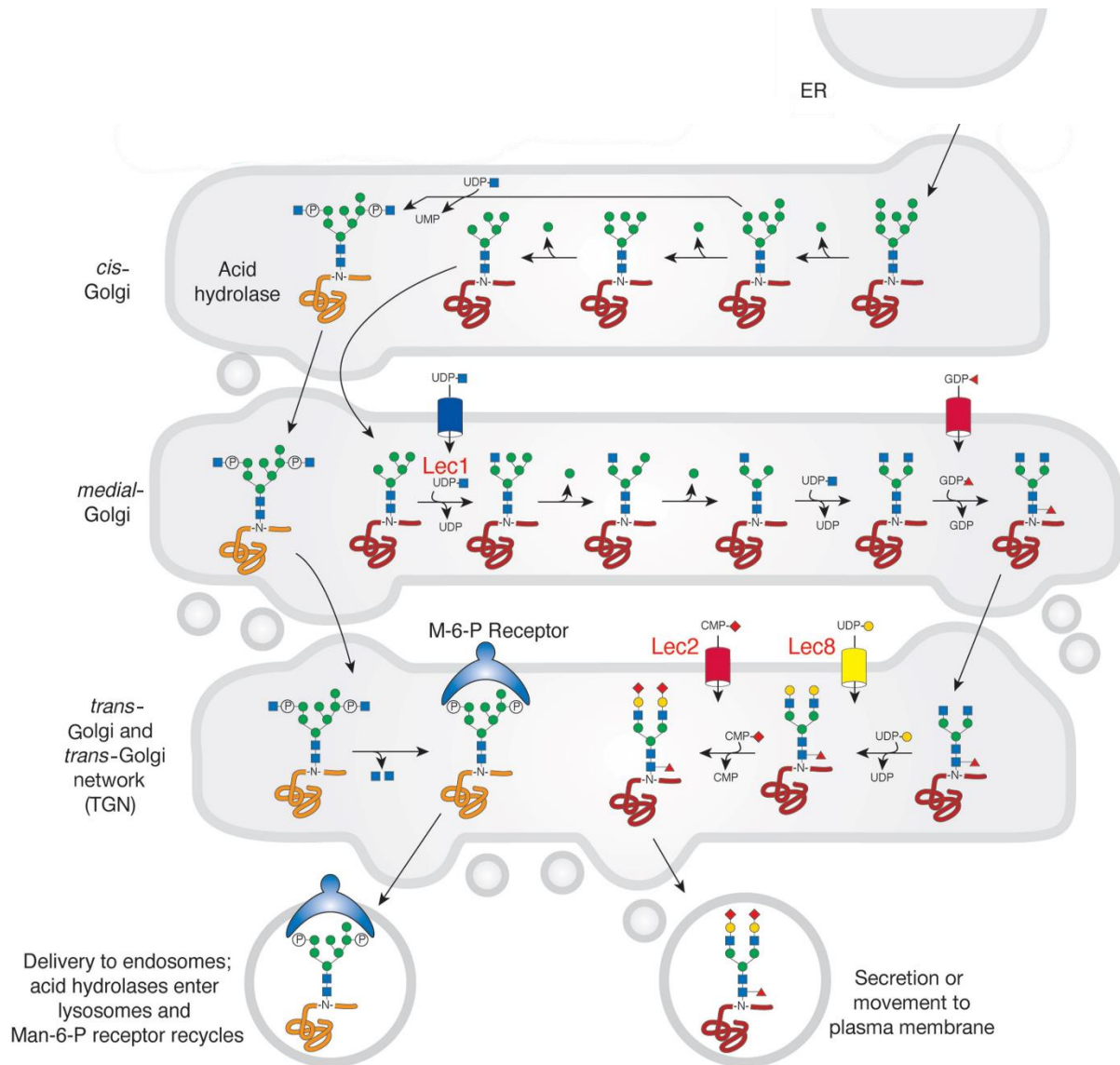
### a. *Cis-Golgi*

A leur arrivée dans le cis-Golgi, les *N*-glycannes des GPs subissent l'action totale ou partielle des  $\alpha$ 1-2 mannosidases IA, IB, et IC golgiennes (MAN1A1, MAN1A2, MAN1C1). Dans le 1<sup>er</sup> cas, on aboutit à des glycannes de type  $\text{Man}_5\text{GlcNAc}_2$  branchés. Dans le 2<sup>nd</sup> cas (absence ou action partielle de ces enzymes), les *N*-glycannes seront de type  $\text{Man}_{6-9}\text{GlcNAc}_2$  (**Figure 6**). Ces derniers constituent le type oligomannosidique à la fin du processus. A noter que certaines GPs correctement repliées arrivent dans le cis golgi en arborant des *N*-glycannes de type  $\text{Glc}_1\text{Man}_6\text{GlcNAc}_2$  empêchant l'action des  $\alpha$ 1-2 mannosidases. Dans ce cas, il existe une enzyme, l'endo- $\alpha$ -mannosidase golgienne, qui clive la liaison  $\alpha$ 1-2 entre les 2 résidus de Man de l'oligosaccharide  $\text{Glc}\alpha 1-3\text{Man}\alpha 1-2\text{Man}$  pour conduire au  $\text{Man}_5\text{GlcNAc}_2$  (Stanley, 2011 ; Stanley *et al.*, 2022a).

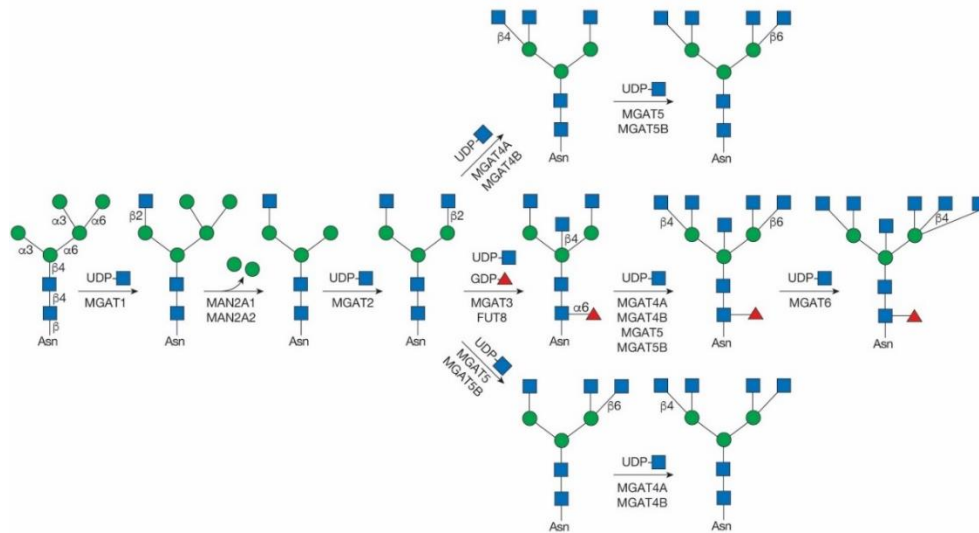
### b. *Médian-Golgi*

Les GPs progressent ensuite dans le médian-Golgi. Celles possédant un *N*-glycanne de type  $\text{Man}_5\text{GlcNAc}_2$  pourront subir l'action de la *N*-acétylglucosaminyltransférase I (GlcNAc-TI ou MGAT1) qui permet d'ajouter un résidu de GlcNAc sur le C2 du Man lié en  $\alpha$ 1-3 sur le noyau et qui initie la synthèse de la première antenne (Stanley *et al.*, 2022a). Puis, l' $\alpha$ -mannosidase II ou IIx (MAN2A1 ou MAN2A2) libère les 2 résidus de Man liés au Man lié en  $\alpha$ 1-6 du noyau pour obtenir la forme  $\text{GlcNAcMan}_3\text{GlcNAc}_2$ . Qu'il y ait action ou non de ces dernières, la GP arbore à ce stade un *N*-glycanne de type hybride avec une antenne qui pourra être allongée. Pour aboutir au précurseur du *N*-glycanne de type complexe, après l'action de la GlcNAc-TI et des mannosidases, la GlcNAc-TII (MGAT2) ajoute un résidu de GlcNAc sur le C2 du Man du noyau lié en  $\alpha$ 1-6 initiant ainsi la deuxième antenne. Sur le précurseur de type complexe ( $\text{GlcNAc}_2$ -noyau), deux autres antennes peuvent être initiées grâce aux actions de la GlcNAc-TIV (MGAT4A, MGAT4B ou MGAT4C) qui ajoute un résidu de GlcNAc sur le C4 du Man  $\alpha$ 1-3 du noyau et de la GlcNAc-TV qui ajoute un résidu de GlcNAc le C6 du Man  $\alpha$ 1-6 du noyau. Le *N*-glycanne sera alors tri ou tetra-antenné (Stanley *et al.*, 2022a). De plus, la GlcNAc-TIII peut ajouter un GlcNAc « intercalaire » sur le Man central du noyau. L'ensemble des GlcNAc-T utilise

comme substrat donneur l'UDP-GlcNAc (Schachter, 2000). Lors de l'initiation des antennes, un résidu de Fuc peut être ajouté par l' $\alpha$ 1-6 fucosyltransférase (FUT8) en  $\alpha$ 1-6 de la GlcNAc réductrice à partir du GDP-Fuc (**Figure 6** et Figure 7). L'enzyme nécessite généralement l'action préalable de la GlcNAc-TI (**Figure 7**) (Stanley *et al.*, 2022a).



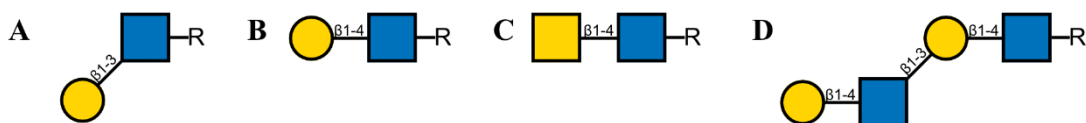
**Figure 6 :** Maturation golgienne des *N*-glycannes. La glycoprotéine correctement conformée va progresser ensuite dans l'appareil de Golgi. Dans le *cis*-Golgi, 3 résidus de Man sont libérés par les mannosidases MAN1A1, MAN1A2 et MAN1C1 pour obtenir le Man<sub>3</sub>GlcNAc<sub>2</sub>. Dans le médian-Golgi, la MGAT initie la première antenne des *N*-glycannes complexes et hybrides. L'action de l' $\alpha$ -mannosidase II permet l'action de MGAT2 qui initie la deuxième antenne. Ces antennes sont ensuite étendues par des résidus de Fuc, Gal, et Sia. D'autres antennes peuvent être générées. Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019). Adaptée de Stanley *et al.* (2022).



**Figure 7** : Initiation des antennes et substitution du noyau des *N*-glycannes de type complexe. L'action de MGAT1 sur les *N*-glycannes  $\text{Man}_5\text{GlcNAc}_2$  aboutit aux *N*-glycannes de type hybride  $\text{GlcNAcMan}_5\text{GlcNAc}_2$ . Les *N*-glycannes de type complexe sont obtenus à partir du type hybride par l'action des  $\alpha$ -mannosidases II, MAN2A1 ou MAN2A2, puis de MGAT2 pour obtenir les *N*-glycannes biantennés  $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ . Les *N*-glycannes peuvent ensuite être substitués par un GlcNAc intercalaire par MAGT3, un Fuc par FUT8, ou d'autres GlcNAc par MGAT4A, MGAT4B, MGAT5 ou MGAT5B. MGAT6 est potentiellement présent chez les mammifères. Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019). Issue de Stanley *et al.* (2022).

### c. *Trans-Golgi et trans-Golgi Network*

Dans la plupart des cas, les résidus de GlcNAc sont allongés par un résidu de Gal pour définir selon la liaison entre le Gal et le GlcNAc, des unités de type 1 ( $\text{Gal}\beta 1-3\text{GlcNAc}$ -) ou des unités de type 2 ( $\text{Gal}\beta 1-4\text{GlcNAc}$ -) également nommées unités LacNAc. La première est catalysée par une  $\beta 1-3$  galactosyltransférase (B3GALT1, 2, 4, ou 5) tandis que la seconde l'est par une  $\beta 1-4$  galactosyltransférase (B4GALT1 à B4GALT6) en utilisant l'UDP-Gal. Sur une unité LacNAc, d'autres unités LacNAc pourront être ajoutées via une liaison  $\beta 1-3$  pour générer des chaînes polyLacNAc par l'action d'une  $\beta 1-3\text{GlcNAc-T}$  (B3GNT2 à B3GNT8). Plus rarement, l'ajout d'un résidu de GalNAc au résidu de Gal conduit à une unité LacdiNAc  $\text{GalNAc}\beta 1-4\text{GlcNAc}$  (**Figure 8**) (Stanley *et al.*, 2022b).



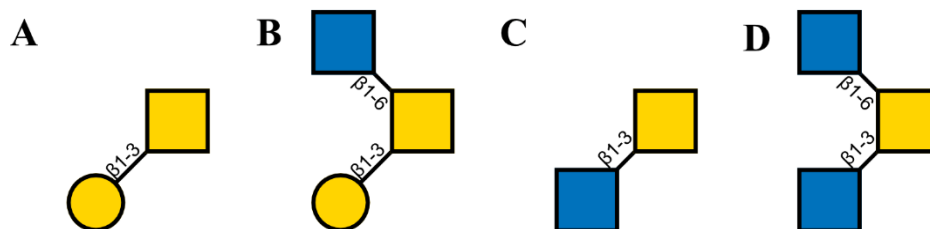
**Figure 8** : Unités composant les antennes des *N*-glycannes. (A) Unité de type 1, (B) de type 2 et (C) LacdiNAc. (D) Les chaînes polyLacNAc se composent d'unités LacNAc liées entre elles par une liaison  $\beta 1-3$ . Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019).

Ces unités pourront servir de squelette pour l'ajout d'autres monosaccharides tels que des résidus de Fuc, de Gal, de GlcNAc ou encore d'acides sialiques qui constitue la glycosylation terminale (partie IV p43).

### III. *O*-glycosylation de type mucine chez les mammifères

#### A. Généralités

La *O*-glycosylation est un processus post-traductionnel caractérisé par l'ajout d'un oligosaccharide sur un acide aminé hydroxylé, Ser ou Thr. A l'inverse des *N*-glycannes, aucune séquence peptidique consensus stricte de *O*-glycosylation n'a encore été décrite (Spiro, 2002). En fonction du monosaccharide du point d'attache et de l'acide aminé sur lequel s'attache les monosaccharides, on distingue plusieurs types de *O*-glycoprotéines dont le type mucine. Dans ce type, un résidu d' $\alpha$ -GalNAc constitue le point d'attache avec un résidu de Ser ou de Thr. Le résidu de GalNAc sera ensuite séquentiellement allongé (ou non) par d'autres monosaccharides. A ce jour, huit noyaux ont été décrits chez les mammifères dont 4 sont retrouvés majoritairement (**Figure 9**) (Brockhausen *et al.*, 2022).



**Figure 9** : Noyaux majoritaires des *O*-glycannes de type mucine retrouvés chez les mammifères. (A) noyau 1, (B) 2, (C) 3 et (D) 4. Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019).

#### B. Biosynthèse des *O*-glycannes de type mucine

En comparaison à la biosynthèse des *N*-glycannes, celle des *O*-glycannes de type mucine révèle 3 différences majeures. La 1<sup>ère</sup> est qu'elle est exclusivement golgienne ce qui signifie, 2<sup>ème</sup> différence, qu'elle est post-traductionnelle. La 3<sup>ème</sup> différence réside dans l'absence d'un précurseur oligosaccharidique, type Dol-P-P-oligosaccharide. La biosynthèse des *O*-glycannes débute dans l'appareil de Golgi par l'ajout d'un résidu de GalNAc sur une Ser ou une Thr par une GalNAc-transférase de la famille des polypeptide GalNAc-transférases (ppGalNAcTs) (Clausen et Bennett, 1996 ; Hagen *et al.*, 2001) dont 20 gènes ont été répertoriés chez l'Homme, la plupart codant une enzyme



active. Par ailleurs, les enzymes ppGalNAcT 1, 2 et 4 sont ubiquitaires tandis que l'expression des autres est tissu dépendante (Bennett *et al.*, 2012).

Le transfert du résidu de GalNAc s'effectue à partir de l'UDP-GalNAc sur le groupement hydroxyle de la Ser ou de la Thr. Si le résidu de GalNAc n'est pas étendu, il définit l'antigène Tn. En fonction des monosaccharides qui lui seront ajoutés, 8 noyaux différents seront définis. L'ajout d'un résidu de Gal grâce à l'enzyme C1GALT1 définit le noyau 1. Le noyau 2 s'obtient à partir du noyau 1 par l'action de la  $\beta$ 1-6 *N*-acétylglucosaminyltransférase 1, 2 ou 3 (GCNT1, GCNT3 et GCNT4) qui va transférer un résidu de GlcNAc sur le résidu de GalNAc via une liaison  $\beta$ 1-6. Pour le noyau 3, le résidu de GalNAc du point d'attache sera substitué par un résidu de GlcNAc par la  $\beta$ 1-3 *N*-acétylglucosaminyltransférase 6 (B3GNT6). Tout comme le noyau 2, le noyau 4 est obtenu par l'ajout d'un résidu de GlcNAc sur la GalNAc du noyau 3 grâce à l'action de la GCNT3 (Brockhausen *et al.*, 2022).

De la même manière que les *N*-glycannes, les noyaux des *O*-glycannes vont être étendus par des unités de type 1, LacNAc et LacdiNAc comme décrit précédemment (partie II.B.2.c p41) sur lesquelles d'autres monosaccharides pourront être ajoutés (partie suivante).

## IV. Glycosylation terminale chez les mammifères

La glycosylation terminale correspond à l'ajout de monosaccharides sur les unités de type 1, LacNAc et/ou LacdiNAc des *N*- et *O*-glycannes mais aussi sur leur noyau. Ce processus est particulièrement important car les motifs glycaniques résultants sont le siège d'importantes fonctions biologiques.

### A. Fucosylation et antigènes de groupes sanguins ABH et Lewis

Comme nous l'avons mentionné précédemment, le Fuc peut être retrouvé sur le noyau des *N*-glycannes mais peut également l'être sur les antennes des *N*- et *O*-glycannes et conduire, dans ce dernier cas, à la formation d'antigènes (Ags).

Chez l'homme, les Ags de groupes sanguins A, B et H qui ont été les plus documentés sont des épitopes glycaniques bâtis à partir de précurseurs disaccharidiques de type 1 et 2, pour les *N*- et

O-glycannes, et de type 3 (Gal $\beta$ 1-3GalNAc $\alpha$ 1-R) exclusivement pour les O-glycannes puisque caractéristique du noyau 1 (**Figure 10** p46). La synthèse commence par le transfert d'un résidu de Fuc en  $\alpha$ 1-2 sur un résidu de Gal par une  $\alpha$ 1-2 fucosyltransférase issue du locus *FUT1* ou *FUT2* (Kelly *et al.*, 1995 ; Larsen *et al.*, 1990). Le transfert s'effectue sur des unités de type 1 ou 2 des précurseurs érythrocytaires pour l'enzyme FUT1 ce qui va conduire à la formation d'Ags H sur les érythrocytes (**Figure 10B**) (Kyprianou *et al.*, 1990 ; Sarnesto *et al.*, 1990). En revanche, l'enzyme FUT2 est exprimée dans les cellules épithéliales et va agir sur les unités de type 1 ou 3 et ainsi générer des Ags H fixes à la surface des épithéliums (muqueuses gastro-intestinales, respiratoires et reproductives) et solubles dans des fluides tels que la salive et le lait maternel (**Figure 10A**) (Sarnesto *et al.*, 1990). Les Ags A et B sont construits à partir de l'Ag H par les  $\alpha$ 1-3 N-acétylgalactosaminyltransférase (A3GALNT) et  $\alpha$ 1-3 galactosyltransférase (A3GALT1) du locus *ABO* qui vont générer respectivement les Ags A et B définissant les groupes sanguins A et B associés (Yamamoto *et al.*, 1990). Chez les individus porteurs du groupe sanguin O, les glycosyltransférases A/B sont inactives et laissent donc les Ags H intacts. Lorsque l'individu synthétise à la fois les Ags A et B, il est de groupe sanguin AB.

Les individus dont le gène *FUT2* est inactif sont qualifiés de « non sécréteur » et ne disposent donc pas d'Ags A, B ou H au niveau des épithéliums et sous forme solubles. Ce phénotype concerne 20 % des populations d'Europe, d'Amérique du Nord, d'Asie et certaines d'Afrique (Kindberg *et al.*, 2006 ; Watkins, 1980) et est plus rare dans les populations mésoaméricaines (5 %) (Bucardo *et al.*, 2009 ; Currier *et al.*, 2015). Au contraire, l'inactivation de *FUT1* est extrêmement rare et conduit au phénotype de Bombay qui se manifeste par l'absence de groupes sanguins sur les érythrocytes (Marionneau *et al.*, 2001).

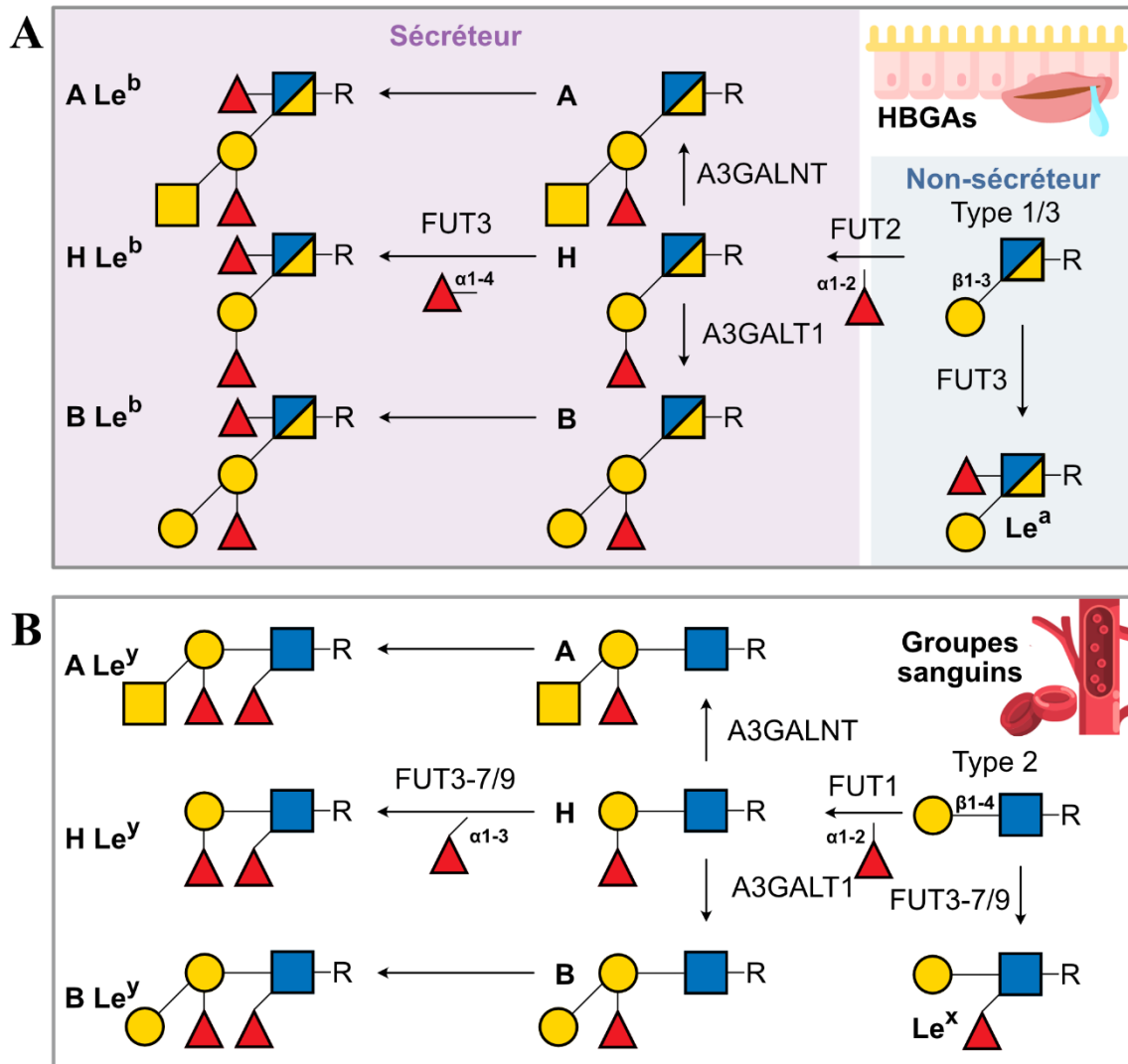
Les Ags de groupe sanguin de Lewis concernent l'ajout de résidus  $\alpha$ 1-3/ $\alpha$ 1-4Fuc sur les unités de type 1 et 2 natives ou définissant un épitope de groupe sanguin A, B ou H (**Figure 10**). L'Ag a de Lewis (Le<sup>a</sup>) est généré par l'ajout d'un résidu de Fuc en  $\alpha$ 1-4 sur un résidu de GlcNAc d'une unité de type 1. L'association de l'Ag H à l'Ag Le<sup>a</sup> définit l'Ag Le<sup>b</sup>. De la même façon, on retrouvera des Ags A et ALe<sup>b</sup> et des Ags B et BLe<sup>b</sup>. L'ajout de ce Fuc sur le résidu de GlcNAc est catalysé par une  $\alpha$ 1-3/ $\alpha$ 1-4 fucosyltransférase du locus *FUT3* exprimée dans les mêmes épithéliums que l'enzyme *FUT2*.



Les Ags Le<sup>b</sup>, ALe<sup>b</sup> et BLe<sup>b</sup> sont donc uniquement retrouvés chez les individus sécréteurs. Les individus non sécréteurs n'exprimeront que des Ags Le<sup>a</sup> avec l'unique action de FUT3. Il existe des individus Lewis-négatif dont le gène *FUT3* est inactif. Cela concerne environ 6 à 10 % en Europe et peut atteindre 20 à 33 % dans certains pays d'Afrique (Bucardo *et al.*, 2009 ; Larsson *et al.*, 2006 ; Nordgren *et al.*, 2013 ; Torrado *et al.*, 1997). Les individus sécréteurs et Lewis négatif pourront uniquement exprimer des Ags ABH et des individus non sécréteur et Lewis négatif présenteront uniquement des antennes de type 1/3 sans résidu. Les Ags x et y de Lewis (Le<sup>x</sup> et Le<sup>y</sup>) qui sont bâtis sur les unités de type 2 sont les équivalents des Ags Le<sup>a</sup> et Le<sup>b</sup> bâtis sur le type 1 et 3. Du fait de l'engagement du C4 de la GlcNAc dans la liaison avec le Gal pour former l'unité LacNAc de type 2, le résidu de Fuc se lie sur le carbone 3 via une liaison en  $\alpha$ 1-3 de ce même résidu de GlcNAc. Cette réaction est catalysée par l'une des fucosyltransférases suivantes, FUT3, FUT4, FUT5, FUT6, FUT7 ou encore FUT9. De la même manière que l'on peut avoir des Ags ALe<sup>b</sup> ou BLe<sup>b</sup>, on peut retrouver des Ags ALe<sup>y</sup> et BLe<sup>y</sup>.

On définit comme antigènes de groupes sanguins tissulaires (HBGAs), les antigènes ABH et Lewis obtenus par les loci *FUT2* et *FUT3* que l'on retrouve à la surface des cellules épithéliales et qui sont sécrétés sous forme libre ou complexés à des protéines ou lipides dans les fluides biologiques.

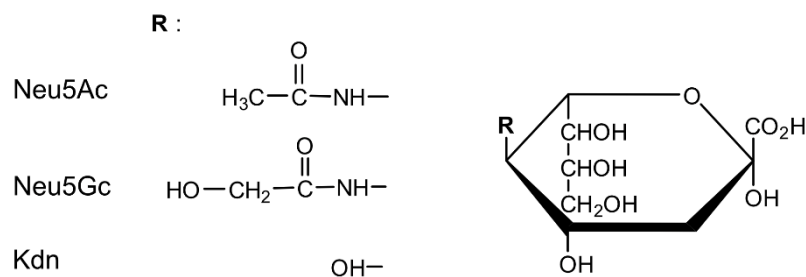
Chez les animaux, l'étude des Ags de groupe sanguin a beaucoup concerné les singes et le porc dans un objectif de pouvoir effectuer des xénogreffes. Chez les singes de l'ancien monde comprenant les grands singes (chimpanzés, gorilles...), les Ags des deux systèmes sont présents (groupes sanguins et HBGAs) avec une faible occurrence de l'Ag H (Blancher et Socha, 1997 ; Oriol *et al.*, 1999). De plus, l'Ag Le<sup>a</sup> a été observé dans la salive (Blancher et Socha, 1997). Pour les singes du nouveau monde (capucins, ouistiti...) et les autres mammifères y compris le porc on ne trouve pas les Ags ABH dans les compartiments vasculaires mais on les retrouve au niveau des épithéliums et des fluides avec principalement les Ags A et H (Oriol *et al.*, 1993, 1994, 1999).



**Figure 10 :** Antigènes de groupes sanguins. **(A)** Les antigènes de groupes sanguins tissulaires (HBGAs) sont synthétisés à partir des unités de type 1 et 3. Le transfert d'un résidu de fucose (Fuc) en  $\alpha$ 1-2 du résidu de galactose (Gal) par l'enzyme FUT2 aboutit à l'antigène (Ag) H. L'ajout supplémentaire d'un résidu de *N*-acétylgalactosamine (GalNAc) ou Gal en  $\alpha$ 1-3 du Gal donne les Ags A et B. Les transferts sont catalysés par les enzymes A3GALNT et A3GALT1, respectivement. L'ajout d'un résidu de Fuc supplémentaire sur les Ags A, B et H en  $\alpha$ 1-4 du résidu de GalNAc ou *N*-acetylglucosamine (GlcNAc) des unités de type 1 et 3 vont définir les Ags Lewis b ( $Le^b$ ). L'ajout de ce Fuc sur les unités de type 1 et 3 natives donnent l'Ag  $Le^a$ . **(B)** Les Ags de groupes sanguins sont obtenus à partir de l'unité de type 2. Le transfert d'un résidu de Fuc en  $\alpha$ 1-2 du résidu de galactose (Gal) par l'enzyme FUT1 aboutit à l'Ag H qui définit le groupe sanguin O. L'ajout supplémentaire d'un résidu de GalNAc ou Gal en  $\alpha$ 1-3 du Gal donne les Ags A et B définissant les groupes sanguins A et B. Les transferts sont catalysés par les mêmes enzymes que pour les HBGAs. L'ajout d'un résidu de Fuc supplémentaire sur les Ags A, B et H en  $\alpha$ 1-3 du résidu de GlcNAc de l'unité de type 2 vont définir les Ags  $Le^y$ . L'ajout de ce Fuc sur l'unité de type 2 native donne l'Ag  $Le^x$ . Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019).

## B. Sialylation

Les acides sialiques sont fréquemment retrouvés chez les animaux (Tiralongo et Martinez-Duncker, 2013). La fonction amine est couramment substituée par un groupement acétyle pour définir l'acide 5-*N*-acétylneuraminique (Neu5Ac) constituant le composé principal des acides sialiques. Juste derrière, on retrouve l'acide 5-*N*-glycolylneuraminique (Neu5Gc) et l'acide 3-déoxy-D-glycero-D-galacto-non-2-ulosonique (Kdn) qui diffèrent par leur substitution du C5. Le premier possède un groupement *N*-glycolyl sur la fonction amine du C5 tandis que le second possède un groupement hydroxyle à la place de cette fonction amine (**Figure 11**). La substitution des autres groupements hydroxyles va permettre de définir plus de 50 composés (Angata et Varki, 2002).



**Figure 11** : Représentation des acides 5-*N*-acétylneuraminique (Neu5Ac), 5-*N*-glycolylneuraminique (Neu5Gc) et 3-déoxy-D-glycero-D-galacto-non-2-ulosonique (Kdn).

Les acides sialiques sont liés en  $\alpha$ 2-3 ou  $\alpha$ 2-6 d'un résidu de Gal et en  $\alpha$ 2-6 d'un résidu de GalNAc des *N*- et *O*-glycannes. Lorsque 2 acides sialiques se lient ensemble, ils le font via une liaison  $\alpha$ 2-8 pour générer des queues polysialyles plus ou moins longues principalement retrouvées sur les *N*-glycannes.

## C. Sulfatation

Les *N*- et *O*-glycannes de type mucine, à la manière des modifications post-traductionnelles des protéines, peuvent être modifiés par l'ajout de groupements sulfates. Chez les mammifères, ils sont retrouvés sur des résidus de Gal, GlcNAc, GalNAc et GlcA (Lei *et al.*, 2010 ; Spiro et Bhoyroo, 1988) et notamment sur le résidu de Gal ou GlcNAc des Ags de Le<sup>x</sup> et Le<sup>a</sup>. Pour ces derniers, on retrouve les groupements sulfates sur le C3 ou 6 du résidu de Gal et sur le C6 du résidu de GlcNAc.

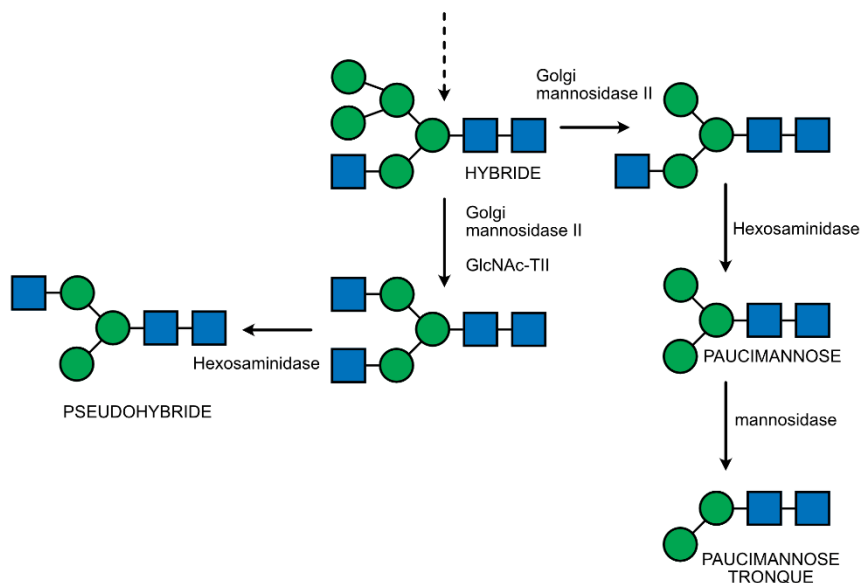
## V. Particularités de la glycosylation chez les invertébrés

### A. *N*-glycosylation chez les invertébrés

Chez les invertébrés, la phase réticulaire de la *N*-glycosylation semble partagée avec celle des mammifères. En revanche, la phase golgienne des invertébrés révèle des différences notables. On ne connaît en effet qu'une faible proportion de l'équipement enzymatique des invertébrés avec des études concernant surtout *Caenorhabditis elegans* mais de nombreuses analyses structurales ont montré certaines caractéristiques absentes chez les mammifères. La partie suivante les décrit.

#### 1. Types de *N*-glycannes

En plus des 3 types de *N*-glycannes (oligomannose, hybride et complexe) retrouvés chez les mammifères, deux autres types de *N*-glycannes ont été décrits chez les invertébrés : le type paucimannose et le type pseudohybride. Ces 2 types de *N*-glycannes s'expliquent par la présence d'une hexosaminidase dans l'appareil de Golgi qui permet l'élimination d'un résidu de GlcNAc en position terminale non réductrice. Le type paucimannose, très répandu chez les invertébrés, correspond à des *N*-glycannes de composition  $\text{Man}_{3-4}\text{GlcNAc}_2$ . Ces *N*-glycannes sont obtenus par l'action d'une hexosaminidase golgienne après l'initiation de la première antenne par la GlcNAc-T I et le clivage par la mannosidase II golgienne d'un ou deux Man de la branche  $\alpha$ 1-6 du noyau (**Figure 12**) (Altmann *et al.*, 1995 ; Gutternigg *et al.*, 2007b ; Léonard *et al.*, 2006 ; Vitale et Chrispeels, 1984 ; Zhang *et al.*, 2003). L'action supplémentaire d'une  $\alpha$ 6-mannosidase qui enlève le Man  $\alpha$ 1-6 du noyau conduit à un noyau tronqué  $\text{Man}_2\text{GlcNAc}_2$  (Cipollo *et al.*, 2002 ; Eckmair *et al.*, 2016 ; Natsuka *et al.*, 2002 ; Wilson et Paschinger, 2016 ; Wuhler *et al.*, 2004 ; Yan *et al.*, 2018). Lorsque l'action de l'hexosaminidase a lieu après l'action de la GlcNAc-TII, on obtient une structure *N*-glycannique que l'on nomme « pseudohybride » et qui ne pourra disposer que d'une seule antenne LacNAc sur le Man  $\alpha$ 1-6 du noyau (Paschinger et Wilson, 2019). Ce type de *N*-glycanne a été identifié chez des insectes et des nématodes (Geisler et Jarvis, 2012 ; Gutternigg *et al.*, 2007b).



**Figure 12 :** Biosynthèse des *N*-glycannes pseudohybride, paucimannose et paucimannose tronqué. Après l'action de la GlcNAc-TII, le résidu de GlcNAc de la première antenne est clivé par une hexosaminidase aboutissant à un *N*-glycanne pseudohybride. Le *N*-glycanne paucimannose est obtenu par l'action successive de la Golgi mannosidase II et d'une hexosaminidase. L'action supplémentaire d'une mannosidase aboutit à un *N*-glycanne paucimannose tronqué. Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019). Inspirée de Paschinger et Wilson (2019).

## 2. Substitutions du noyau pentasaccharidique

Alors que les *N*-glycannes de type mammifère ne peuvent arborer sur leur GlcNAc en position terminale réductrice qu'un seul Fuc lié en  $\alpha$ 1-6, la présence de Fuc en  $\alpha$ 1-6 mais aussi en  $\alpha$ 1-3 et même les 2 simultanément est très courante chez les invertébrés comme chez les nématodes *C. elegans* et *Haemonchus contortus* (Haslam *et al.*, 1996). Cette substitution peut également être retrouvée sur le 2<sup>ème</sup> résidu de GlcNAc du noyau. Chez de nombreux invertébrés comme l'escargot de mer, le calmar, les planaires et les nématodes, le Fuc peut être galactosylé par une  $\beta$ 4-galactosyltransférase et ainsi former le motif oligosaccharidique Gal $\beta$ 1-4Fuc également nommé « GalFuc » (Subramanian *et al.*, 2018 ; Takahashi *et al.*, 2003 ; Titz *et al.*, 2009 ; Wuhrer *et al.*, 2004). Ce motif peut lui-même être étendu par un résidu de Gal ou de Fuc supplémentaire (Wuhrer *et al.*, 2004) voire même par de multiples résidus de Gal comme décrit pour le planaire *Schmidtea mediterranea* (Subramanian *et al.*, 2018). Enfin, sur la GlcNAc de l'extrémité réductrice, la substitution par un Man a également été décrite chez l'escargot de mer *Volvarina rubella* (Eckmair *et al.*, 2016).

Trois  $\alpha$ -fucosyltransférases (FUT1, FUT6 et FUT8) et une galactosyltransférase (GALT-1) permettant le transfert d'un Gal sur un Fuc lié en  $\alpha$ 1-6 d'une des GlcNAc ont été identifiées chez *C. elegans* (Paschinger *et al.*, 2005 ; Titz *et al.*, 2009 ; Yan *et al.*, 2013). Les  $\alpha$ 1-3 fucosyltransférases ont aussi été identifiées chez *Drosophila melanogaster*, l'abeille *Apis mellifera* et le moustique *Anopheles gambiae*. Pour ce dernier, l' $\alpha$ 1-6 fucosyltransférase a également été identifiée (Fabini *et al.*, 2001 ; Kurz *et al.*, 2016 ; Rendić *et al.*, 2007).

Au niveau du résidu de Man central du noyau, la seule substitution connue chez les mammifères est celle par un résidu de GlcNAc nommé intercalaire. Chez les invertébrés, les *N*-glycannes peuvent arborer un résidu de Xyl lié au Man central par une liaison  $\beta$ 1-2. D'abord caractérisée chez les plantes, cette substitution a ensuite été décrite chez des escargots (*Helix pomatia*, *Lymnea stagnalis* et *Biomphalaria glabrata*) et certains trématodes de type *Schistosoma* (Ishihara *et al.*, 1979 ; Khoo *et al.*, 1997 ; Lehr *et al.*, 2007 ; Van Kuik *et al.*, 1985, 1986, 1987). L'action de la xylosyltransférase responsable du transfert du Xyl nécessite celle préalable de la MGAT1. A noter que le motif GalFuc décrit ci-dessus a également été observé sur le Man centrale du noyau chez *C. elegans* (Yan *et al.*, 2015).

### 3. Antennes

D'une manière similaire aux mammifères, il est retrouvé chez les invertébrés des antennes de type 1, LacNAc et LacdiNAc. Chez ces derniers, des unités LacdiNAc prolongées ou non par un résidu de  $\beta$ 1-3Gal ont été caractérisées. De même, des séquences GlcNAc $\beta$ 1-4GlcNAc ont été identifiées chez divers invertébrés (Paschinger et Wilson, 2019). Quelques enzymes impliquées dans la synthèse de ces motifs ont été identifiées comme des  $\beta$ 1-4 *N*-acétylgalactosaminyltransférases de *C. elegans* et *Trichoplusia ni*,  $\beta$ 1-3 galactosyltransférase de l'abeille et la  $\alpha$ 1-4 *N*-acétylgalactosaminyltransférase de la drosophile (Ichimiya *et al.*, 2015 ; Kowar *et al.*, 2002 ; Mucha *et al.*, 2004).

Comme décrit chez les mammifères, les antennes des invertébrés décrites ci-dessus peuvent être substituées par des résidus de Fuc qui définiront les mêmes Ags (Kubelka *et al.*, 1993 ; Wilson et Paschinger, 2016 ; Wuhler *et al.*, 2006). Ainsi, chez le nématode *Dictyocaulus viviparus*, le trématode *Schistosoma mansoni*, et le moustique *Anopheles gambiae*, il a été détecté la présence de Le<sup>x</sup>. Chez ces 2 derniers, les fucosyltransférases impliquées dans la synthèse de l'Ag Le<sup>x</sup> ont été identifiées. Chez

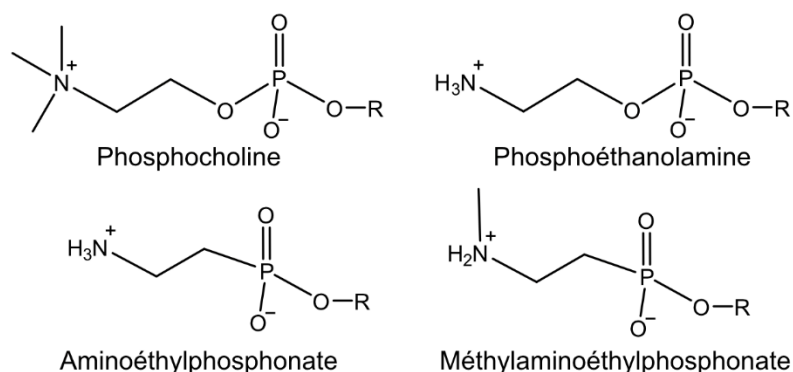
l'huître *Crassostrea virginica*, c'est la présence d'Ag A qui a été révélée (Haslam *et al.*, 2000 ; Kurz *et al.*, 2013, 2016 ; Smit *et al.*, 2015). Plus rarement, on a retrouvé des séquences Fuc $\alpha$ 1-2Fuc $\alpha$ 1-3 comme chez *S. mansoni* (Smit *et al.*, 2015). A l'extrémité des antennes des *N*-glycannes, l'acide glucuronique a été retrouvé chez un certain nombre d'insectes ainsi que chez l'escargot *Volvarina rubella* (Aoki et Tiemeyer, 2010 ; Eckmair *et al.*, 2016 ; Hykollari *et al.*, 2018 ; Kurz *et al.*, 2015). En revanche, à l'inverse des mammifères, la sialylation est peu retrouvée chez les invertébrés.

#### 4. Modification des *N*-glycannes

Et pour complexifier davantage encore ces structures, certains monosaccharides seront le siège de modifications par des groupements zwitterioniques comme la phosphocholine, la phosphoéthanolamine, l'ainoéthylphosphonate et le méthylainoéthylphosphonate mais également par des groupements sulfates, phosphates et/ou méthyles.

##### a. Groupements zwitterioniques

La phosphocholine a surtout été repérée chez des nématodes comme le parasite porcin *Oesophagostomum dentatum* ou encore *H. contortus* (Jiménez-Castells *et al.*, 2017 ; Paschinger et Wilson, 2015). Les autres groupements zwitterioniques sont moins courants et ont été observés dans la gelée royale pour la phosphoéthanolamine, chez un criquet pour l'ainoéthylphosphonate et chez l'escargot marin *Volvarina rubella* pour la méthylainoéthylphosphonate (**Figure 13**) (Eckmair *et al.*, 2016 ; Hård *et al.*, 1993 ; Hykollari *et al.*, 2018).



**Figure 13** : Formules chimiques développées des groupements zwitterioniques phosphocholine, phosphoéthanolamine, aminoéthylphosphonate et méthylainoéthylphosphonate.

### b. Groupements sulfates et phosphates

Les groupements sulfates sont très répandus chez les mollusques marins et les arthropodes et ont été localisés sur différents résidus comme le Man, le Fuc et le Gal (Hykollari *et al.*, 2018 ; Kurz *et al.*, 2013, 2015, 2016 ; van Kuik *et al.*, 1987). Concernant la phosphorylation, on ne l'a identifiée que sur le mannose de *Fasciola hepatica* (Ravidà *et al.*, 2016). A noter que chez les mammifères, des résidus de phospho-Man sont également présents et constituent des séquences signal qui destinent au lysosome les GPs qui les portent (Reitman *et al.*, 1981 ; Waheed *et al.*, 1982).

### c. Groupements méthyles

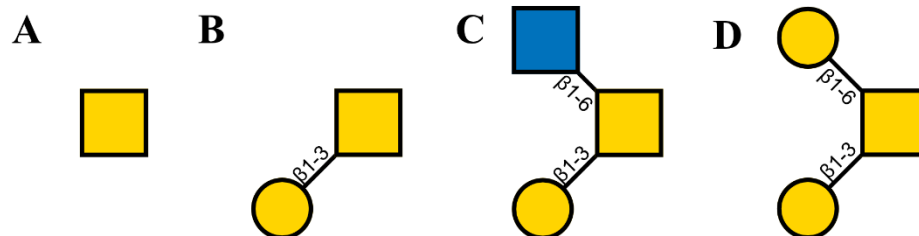
La méthylation est assez répandue chez les invertébrés. Elle a été détectée pour la première fois pour des GPs chez les gastéropodes *Helix pomatia* et *Lymnea stagnalis*. Ces groupements ont été positionnés sur le C3 à la fois sur des résidus de Gal pour les deux escargots et sur des résidus de Man pour la deuxième espèce (Hall *et al.*, 1977 ; Van Kuik *et al.*, 1986, 1987). D'autres gastéropodes ont ensuite montré la présence de ces modifications (Dolashka-Angelova *et al.*, 2003 ; Gutternigg *et al.*, 2004, 2007a) ainsi que d'autres mollusques, un planaire et des nématodes libres ou parasites (Guérardel *et al.*, 2001 ; Jiménez-Castells *et al.*, 2017 ; Khoo *et al.*, 1991 ; Kurz *et al.*, 2013 ; Natsuka *et al.*, 2011 ; Paschinger *et al.*, 2011 ; Yan *et al.*, 2018). Pour les *N*-glycannes, le plus souvent les groupements méthyles sont associés aux Man sur le noyau ou sur les structures oligomannosidiques. Ils ont aussi été observés sur des résidus de Fuc seul du noyau ou dans les séquences substituants la GlcNAc réductrice comme FucGalFuc et GalFuc (Jiménez-Castells *et al.*, 2017 ; Paschinger *et al.*, 2011 ; Yan *et al.*, 2018). De manière plus intéressante pour nos travaux, ces substituants ont été observés sur le résidu de GalNAc de l'Ag A de l'huître *C. virginica* (Kurz *et al.*, 2013).

## B. O-glycosylation de type mucine chez les invertébrés

Comme pour les *N*-glycannes des invertébrés, la biosynthèse des *O*-glycannes de type mucine est beaucoup moins documentée que chez les mammifères. Dans la plupart des cas, seules des structures ont été décrites sans que les enzymes impliquées dans leur biosynthèse ne l'aient forcément été. A cet égard, *C. elegans* apparaît comme le modèle le mieux caractérisé avec au moins 9 GalNAc-T identifiées (Hagen et Nehrke, 1998).



Chez les nématodes, arthropodes, plathelminthes et mollusques étudiés, la présence d'Ag Tn et des noyaux de type 1 et 2 ont été détectés. De plus, un noyau spécifique aux invertébrés Gal $\beta$ 1-6[Gal $\beta$ 1-3]GalNAc a pu être identifié chez des nématodes, schistosomes et mollusques (**Figure 14**) (Staudacher, 2015).



**Figure 14** : Noyaux des *O*-glycannes de type mucine retrouvés chez les invertébrés. (A) Antigène Tn, (B) noyau 1, (C) noyau 2 et (D) noyau des invertébrés. Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019). Inspirée de Staudacher (2015).

L'élongation de ces noyaux est assez variée et reprend un certain nombre des caractéristiques identifiées pour les *N*-glycannes avec néanmoins un taux de substitution par le Fuc, l'acide glucuronique et les groupements méthyles plus important (partie précédente).

## VI. Rôles biologiques des glycannes

Les rôles biologiques joués par les glycannes sont très variés mais peuvent être classés en deux grandes catégories : ceux physico-chimiques et ceux impliqués dans des reconnaissances spécifiques dont quelques exemples seront cités ci-dessous (Varki, 2017).

### A. Rôles physico-chimiques

Concernant les rôles physicochimiques, les glycannes sont présents majoritairement à la surface des cellules pour former le glycocalyx et constitue ainsi une barrière physique pour protéger la membrane cellulaire (Möckl, 2020). Ils sont également retrouvés dans la matrice extracellulaire sous la forme de protéoglycannes (kératanes sulfates, chondroïtines sulfates, héparanes sulfates et dermatanes sulfates) qui jouent des rôles mécaniques importants (étirement, compression, lubrification) et dans le même temps, crée un environnement hydrophile (Yanagishita, 1993).

A la surface des épithéliums des voies aériennes et digestives, on retrouve des mucines qui sont des protéines fortement glycosylées. Elles sont liées à la membrane des cellules mais sont également

secrétées pour former le mucus. Il s'agit de molécules de très haut poids moléculaire qui se caractérisent par une grande viscosité. Les fonctions biologiques qu'elles jouent sont multiples et vont de la simple protection physique des épithéliums (barrière contre les bactéries, hydratation, lubrification) jusqu'à des fonctions de reconnaissance de molécules ou de cellules (Bansil et Turner, 2006).

A la surface de certaines protéines, le caractère très hydrophile des glycanes contribue à leur solubilité. Dans d'autres cas, le volume spatial occupé par les chaînes glycaniques permettra de protéger la protéine contre la protéolyse. A titre d'exemple, citons le contrôle du temps de demi-vie des GPs, notamment via l'acide sialique. En masquant les résidus de galactose terminaux, ils empêchent la reconnaissance de ces derniers par les récepteurs d'Ashwell-Morell au niveau hépatique qui se traduit par l'endocytose puis la destruction de la GP (Grewal, 2010 ; Morell *et al.*, 1971).

## B. Reconnaissances spécifiques

Pour les glycanes impliqués dans les reconnaissances spécifiques on retrouve des épitopes reconnus par des protéines de liaison aux glycanes (Glycan binding proteins GBP aussi appelées lectines). Ces GBPs sont soit intrinsèques et vont reconnaître les glycanes du même organisme soit extrinsèques et reconnaître les glycanes d'un autre organisme (Varki, 2017).

La reconnaissance par les GBPs intrinsèques médie la clairance, le renouvellement et le trafic intracellulaire des GPs circulantes par la reconnaissance de certains glycanes terminaux ou subterminaux. On peut citer pour cela le rôle du Man-6-phosphate ajouté aux *N*-glycanes des enzymes lysosomales lors de leur synthèse et permettant leur adressage au lysosome (Coutinho *et al.*, 2012). Ces GBPs intrinsèques permettent également la reconnaissance cellule-cellule. Par exemple, les sélectines qui sont des GPs membranaires retrouvées au niveau des plaquettes, cellules endothéliales et leucocytes médient l'adhésion des leucocytes et des plaquettes à la surface des tissus vasculaires et entraîne leur roulement dans les vaisseaux par la formation et dissociation rapide des interactions. Cette interaction est importante pour le recrutement des leucocytes vers les sites d'inflammation (McEver, 2002). Les GBPs intrinsèques sont aussi impliquées dans les interactions entre les cellules et la matrice extracellulaire.

La glycosylation peut également moduler les interactions protéine-protéine. La glycosylation peut moduler la fonction primaire d'une protéine plutôt que de l'activer ou de la désactiver. On peut prendre l'exemple d'anticorps identiques qui vont avoir des effets différents selon la structure précise du *N*-glycane retrouvé sur le domaine Fc (Alter *et al.*, 2018).

Dans le cas des GBPs extrinsèques, les glycannes reconnus sont portés par un autre organisme et notamment des pathogènes comme les bactéries ou encore les virus. Ces pathogènes peuvent également être porteur de glycannes et ainsi pratiquer le "mimétisme moléculaire" (Varki, 2017). Dans ce cas, ils arrivent à échapper aux réactions immunitaires grâce aux glycannes typiques de l'hôte qui les décore.

## VII. Interactions hôte-virus médiées par les glycannes

Parmi les glycannes qui ont un rôle dans la reconnaissance avec les pathogènes, on va retrouver ceux qui constituent des récepteurs ou des facteurs d'attachement chez la cellule hôte pour les virus (Ströh et Stehle, 2014). Ces récepteurs vont permettre l'adhésion du virus à certaines cellules cibles déterminant ainsi un tropisme d'hôte et de tissu. Cette interaction spécifique constitue la première étape du processus d'attachement du virus et son entrée dans la cellule hôte. Elle est souvent de faible affinité et peut être suivie d'une liaison avec un corécepteur qui fixera plus fortement le virus à la surface de la cellule hôte. Il est à noter que dans bons nombres de cas, les glycannes impliqués dans la reconnaissance spécifique avec une cellule hôte peuvent être retrouvés à la surface des virus. Néanmoins, nous ne parlerons que des reconnaissances impliquant les glycannes portés par l'hôte, puisqu'ils sont au cœur de cette étude.

On peut scinder les reconnaissances glycannes-virus chez l'hôte en trois groupes, celles médiées par les acides sialiques, celles médiées par les glycosaminoglycannes (GAGs) et enfin, celles médiées par des glycannes non chargées comme ceux des HBGAs (Ströh et Stehle, 2014).

### A. Acides sialiques

Les acides sialiques sont impliqués dans les interactions avec les virus puisqu'ils sont situés à l'extrémité terminale des glycannes, sont ubiquitaires des vertébrés supérieurs et sont retrouvés à la fois

sur les *N*-, *O*-glycannes ainsi que les glycolipides. En outre, les multiples modifications que peuvent subir les acides sialiques peuvent conduire à plus de 50 variantes dépendantes de l'espèce (Angata et Varki, 2002).

La *O*-acétylation de l'acide sialique apparaît dans des tropismes d'hôte, d'organe et de cellule notamment dans la famille des *Coronaviridae* avec les torovirus et les bêtacoronavirus. Ces virus possèdent des hémagglutinines estérases qui se lient au Neu5Ac porteur d'un groupement 9-*O*-acétyle (Neu5,9Ac<sub>2</sub>) ou 4-*O*-acétyle (Neu4,5Ac<sub>2</sub>) voire les deux (Neu4,5,9Ac<sub>3</sub>). La plupart des hémagglutinines estérases de coronavirus (CoV) se lient au Neu5,9Ac<sub>2</sub> mais certains CoVs murins se lient au Neu4,5Ac<sub>2</sub>. Pour le CoV bovin BCoV et le murin MHV-S, des changements mineurs dans le site de liaison changent la spécificité de liaison entre ces deux acides sialiques (Kim, 2020).

En plus des modifications, la liaison engagée peut influencer le tropisme d'hôte et tissulaire. Le virus de la grippe Influenza A est impacté par le type de liaison  $\alpha$ 2-3 et  $\alpha$ 2-6 du résidu de Neu5Ac sur celui de Gal retrouvé chez l'hôte qui est expliqué par la conformation spatiale différente qu'occupent chacun des ligands (Byrd-Leotis *et al.*, 2017). Ainsi, les gripes aviaires sont causées par des souches virales qui reconnaissent majoritairement l' $\alpha$ 2-3Neu5Ac alors que celles humaines reconnaissent majoritairement l' $\alpha$ 2-6Neu5Ac. De plus, dans la plupart des contaminations, le virus est retrouvé spécifiquement dans un tissu du fait de la présence d'un de ces ligands (tractus respiratoire, intestin...). Cependant, cela est à nuancer puisque d'autres phénomènes encore mal compris sont à prendre en considération pour expliquer le tropisme d'hôte et d'espèces (valence de l'interaction, accessibilité des glycannes...) (Byrd-Leotis *et al.*, 2017 ; Liu *et al.*, 2023).

## B. Glycosaminoglycannes

Les GAGs sont des protéoglycannes polyanioniques constitués de disaccharides liés en  $\beta$ 1-4 dont les charges sont portées par des groupements sulfates et des acides uroniques. On les retrouve au niveau de la matrice extracellulaire ainsi qu'à la surface des cellules de la plupart des tissus et organes eucaryotes. A côté des rôles physico-chimiques cités précédemment, les GAGs sont également le support d'une multitude de reconnaissance par les virus. A titre d'exemples, pour le virus de l'herpès (HSV), les chaînes d'héparane sulfate permettent l'attachement de HSV-1 à la surface cellulaire *via* les

GPs B et C du virion. Il y a ensuite une implication des chaînes d'héparane sulfate 3-O-sulfatées dans l'interaction avec la GP D pour faciliter l'entrée virale. Les héparanes sulfates jouent également un rôle important dans l'attachement et l'entrée du virus de l'hépatite C dans la cellule cible notamment grâce à la *N*- et la 6-*O*-sulfatation de certains de ses monosaccharides (Koganti *et al.*, 2021).

### C. HBGAs

Les Norovirus ont été largement étudiés pour leur reconnaissance des HBGAs qui est dépendante de la souche. On peut également citer les Rotavirus dont la plupart des souches reconnaissent les acides sialiques mais dont certaines chez l'Homme reconnaissent les HBGAs. Par exemple, la sous unité VP8\* de la protéine de capsid VP4 (impliquée dans l'interaction avec les glycanes) du rotavirus humain HAL1166 présente des modifications subtiles dans son site de liaison qui l'empêchent de se lier à l'acide sialique mais lui permettent d'engager des liaisons avec les HBGAs (Hu *et al.*, 2012).

La reconnaissance des HBGAs par les Norovirus a été le cœur du sujet durant ma thèse et sera largement détaillé dans le chapitre suivant.

## Chapitre 2 : NOROVIRUS

### I. Présentation

Les épidémies de gastroentérites aiguës surviennent régulièrement dans le monde et peuvent s'avérer mortelles dans les pays en voie de développement. La maladie cause principalement des diarrhées et vomissements chez l'individu contaminé et entraîne une déshydratation importante chez les sujets jeunes et âgés pouvant conduire à la mort. Sa transmission est interhumaine ou causée par la consommation d'eau ou d'aliments contaminés. Alors que les sources causant ces épidémies dans les pays en voie de développement sont très variées (virus, bactéries, parasites), dans les pays développés, elles sont majoritairement virales et surviennent essentiellement l'hiver. Tout âge confondu, les Norovirus (NoVs) représentent la majorité de ces épidémies dans ces derniers (84 % recensés en France pour la saison hivernale 2021-2022 (Santé Publique France, 2023)) et sont responsables d'environ 200 000 morts par an dans le monde (World Health Organization, 2022).

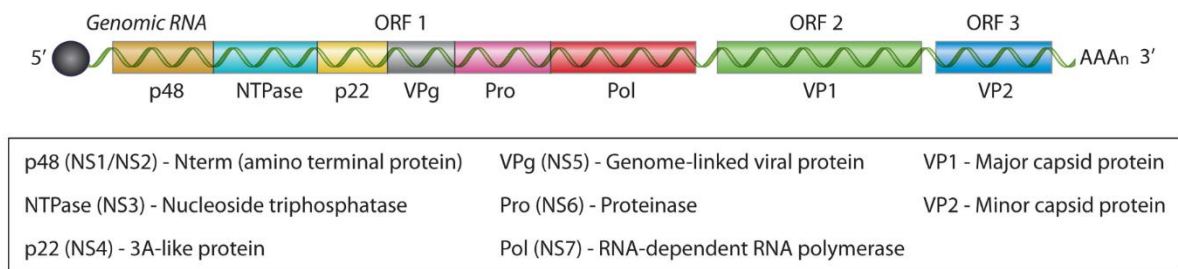
### II. Structure des Norovirus

Les NoVs appartiennent à la famille des *Caliciviridae*. Ils sont constitués d'un simple brin d'ARN de 7,5-7,7 Kb renfermé dans une capsidie icosaédrique non enveloppée de 27 à 40 nm de diamètre. Cette capsidie est composée de deux protéines : la protéine VP1 qui est dimérique et majoritaire et la protéine VP2, minoritaire (Bányai *et al.*, 2018).

#### A. Génome

Le génome des NoVs est composé d'une extrémité 5' liée de manière covalente à une protéine virale (VPg : Viral Protein linked to Genome) essentielle pour la réplication virale et d'une extrémité 3' polyadénylée (Robilotti *et al.*, 2015). Entre celles-ci, 3 cadres de lecture ouverts (Open Reading Frame, ORF) ont été décrits hormis pour le NoV murin où un quatrième a été observé. L'ORF1, situé à proximité de l'extrémité 5', code une polyprotéine qui après clivage par la protéase NS6 va générer six protéines non structurales (NS1/2 à NS7 incluant NS6). Parmi ces protéines, la NS7 correspond à la polymérase RdRp (RNA-dependent RNA polymerase). D'autre part, les ORF2 et 3 sont impliqués dans

la capside protéique et vont coder les protéines structurales VP1 et VP2 respectivement (**Figure 15**) (Robilotti *et al.*, 2015).

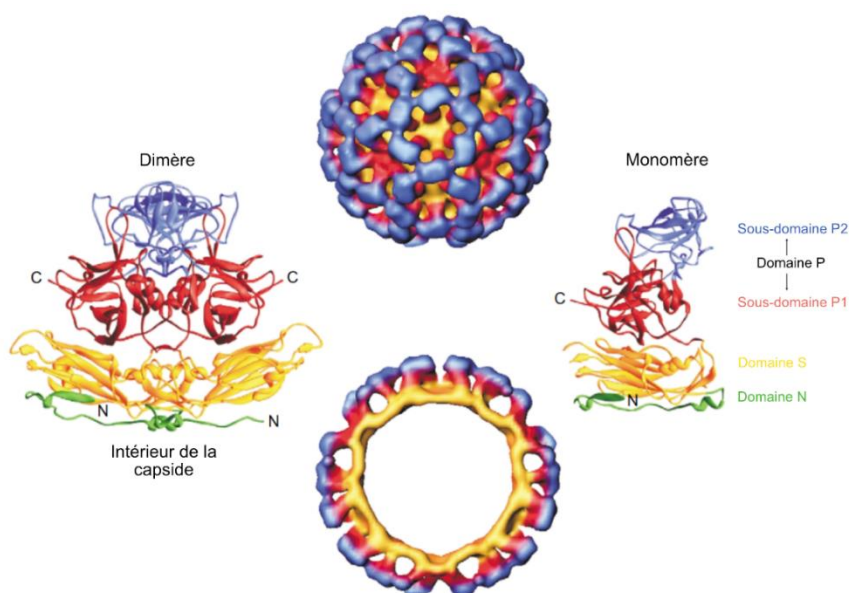


**Figure 15:** Organisation du génome des norovirus humains. Le génome des norovirus se compose de 3 cadres de lecture ouverts (ORF). L'ORF 1 code pour une polyprotéine composée des protéines NS1 à NS7 qui seront libérées par la protéine NS6. Issue de Robilotti *et al.*, 2015.

## B. Capside virale protéique

La capside des NoVs est majoritairement composée de la protéine VP1 assemblée en 90 dimères et conduisant à la formation d'une capside de symétrie icosaédrique (Robilotti *et al.*, 2015). Cette protéine dispose d'une capacité d'autoassemblage qui a été exploitée pour la création de Virus Like-Particles (VLPs) qui ressemblent au virion natif mais sans leur génome et donc dépourvu de toute infectiosité. Ainsi, ces VLPs constituent un outil précieux pour l'étude des NoVs dans des systèmes biologiques (Jiang *et al.*, 1992). La protéine VP1 est composée de trois domaines : le domaine N à l'extrémité N-terminale et les domaines S (Shell) et P (Protruding) (Robilotti *et al.*, 2015). Le domaine S est la région la plus conservée de la protéine VP1 qui est essentielle pour la formation de la capside et qui entoure l'ARN. En revanche, le domaine P est une région variable exposée à la surface de la capside et subdivisé en 2 sous-domaines nommés P1 et P2 (**Figure 16**). Le sous-domaine P2 constitue le site de liaison à la cellule hôte. Il présente une grande variabilité de séquence à l'origine de la diversité de fixation des NoVs aux HBGAs de la cellule hôte (Almand *et al.*, 2017).

La protéine de capside VP2 est retrouvée sur la face interne de la capside. Son rôle n'est pas encore clairement établi mais il est postulé qu'elle serait impliquée dans la stabilisation de la protéine VP1 et dans l'encapsulation du génome (Vongpunsawad *et al.*, 2013).



**Figure 16 :** Organisation de la protéine VP1 dans la capsidie protéique des norovirus. La protéine VP1 est constituée des domaines Protruding (P), Shell (S) et *N*-terminal (N). Le domaine P se décompose en sous-domaines P1 et P2. Pour la formation de la capsidie, la protéine VP1 s'organise en 90 dimères. Issue de Hutson *et al.* (2004).

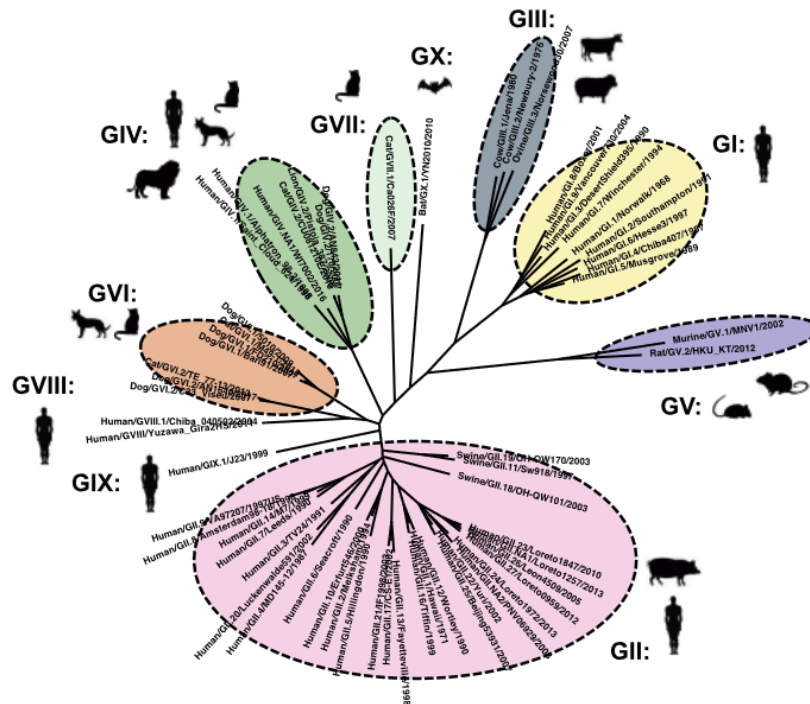
### III. Classification des Norovirus

A l'origine, la nomenclature des NoVs était basée sur la séquence en acides aminés (AA) de la protéine VP1. Les NoVs sont sujets à de nombreuses recombinaisons génétiques dont le point de recombinaison se situe le plus souvent au niveau de la jonction entre l'ORF1 et l'ORF2. Cela crée de nouveaux virus avec une polymérase et une capsidie qui n'était pas associées précédemment. Pour tenir compte de ce phénomène, un double typage a été mis en place à partir de la séquence en AA de la protéine VP1 et de la séquence en nucléotides de l'ORF1 qui code la polymérase RdRp (Chhabra *et al.*, 2019). Nous nous focaliserons uniquement sur le typage relatif à la protéine VP1 puisqu'elle est directement impliquée dans les reconnaissances des HBGAs par les NoVs.

Le typage basé sur la structure primaire de la protéine VP1 conduit à 10 génogroupes (GI à X) qui sont eux-mêmes subdivisés en 27 génotypes (de 1 à 27) en fonction du taux d'homologie de séquence entre les souches (Parra, 2019). Les génogroupes GI, GII, GIV, GVIII et GIX sont impliqués dans des infections chez l'homme. Les génogroupes GI et GII sont les plus communs. En particulier les souches GII.4 qui sont fréquemment retrouvées dans les épidémies humaines (dans environ 50 % des cas) (Kendra *et al.*, 2022). Ces souches évoluent rapidement conduisant à une grande variabilité de variants



qui sont nommés du nom de la ville où l'épidémie est apparue. Les autres génogroupes sont impliqués dans des infections qui touchent une multitude d'animaux (Figure 17) (Parra, 2019).



**Figure 17 :** Classification des norovirus basée sur la phylogénie de la protéine VP1 et leurs cibles animales. A noter que les génogroupes GI, GII, GIV, GVIII et GIX infectent l'homme. Issue de Parra (2019).

## IV. Reconnaissance des HBGAs par les Norovirus

Les différents génogroupes et génotypes de NoVs humains présentent des reconnaissances spécifiques aux HBGAs. Ainsi, le statut sécréteur et la diversité des Ags vont contribuer à la susceptibilité ou à la résistance de certains individus vis-à-vis d'une infection par les NoVs.

### A. Influence du statut « sécréteur »

C'est en 1977 qu'on a suspecté pour la première fois qu'un autre facteur que le développement d'anticorps anti-NoVs découlant d'une précédente infection était impliqué dans la résistance à l'infection par les NoVs. En effet, parmi des volontaires infectés par la souche Norwalk GI.1, seulement la moitié a développé les symptômes caractéristiques d'une infection à NoVs. Pour l'autre moitié, asymptomatique, aucune trace d'infection et notamment aucune augmentation du taux d'anticorps n'a été constatée. Les auteurs ont émis l'hypothèse de l'absence d'un récepteur spécifique chez les individus asymptomatiques en faisant un parallèle avec le paludisme dont on venait de découvrir que l'infection dépendait des Ags de groupe sanguin retrouvés à la surface des érythrocytes (Parrino *et al.*, 1977).

Plus tard, plusieurs études ont rapporté également une telle résistance à l'infection pour les souches GI.1 et GII.4 et il a été établi une corrélation entre le statut « sécréteur » des individus infectés et leur sensibilité à ces virus (Huang *et al.*, 2003 ; Hutson *et al.*, 2005 ; Lindesmith *et al.*, 2003 ; Marionneau *et al.*, 2002). Cette observation a été étendue à d'autres souches de NoVs comme la souche GII.5 (MOH) ou encore GII.17 (Huang *et al.*, 2003 ; Zhang *et al.*, 2015). A l'inverse, le statut « sécréteur » des individus ne semblait pas impliqué dans la sensibilité ou la résistance des individus vis-à-vis d'infection par certains NoVs tels que ceux des génotypes GII.2 SMV ou encore GII.9 VA207 (Huang *et al.*, 2003 ; Lindesmith *et al.*, 2005). Pour les souches dont l'infection est indépendante du statut « sécréteur », il a été montré un tropisme d'hôte lié à la reconnaissance des motifs Le<sup>a</sup> mais aussi à d'autres récepteurs cellulaires (Almand *et al.*, 2017 ; Tarris *et al.*, 2022).

La résistance ou non des individus selon le statut sécréteur n'est donc pas un facteur binaire. On peut retrouver des individus non sécréteurs infectés par une souche qui infecte habituellement des individus sécréteurs et inversement des individus sécréteurs non-infectés par cette même souche (Nordgren et Svensson, 2019). Cela peut être expliqué par des mutations au niveau du gène *FUT2* qui induit une activité plus faible de la fucosyltransférase ou encore par des facteurs ayant un impact sur la glycosylation de l'épithélium intestinal. Une étude a montré que certains individus non sécréteurs du fait de l'inactivité de l'enzyme FUT2, présentaient néanmoins des motifs  $\alpha$ 1-2Fuc à la surface de leur muqueuse intestinale sur lesquels pouvaient se lier des NoVs GII.4. Ce phénomène a été corrélé à l'infection par une souche d'*Helicobacter pylori* exprimant la cytotoxine CagA qui est connue pour créer des dysfonctionnements cellulaires (Ruvoën-Clouet *et al.*, 2014). Les auteurs suggèrent que la bactérie entraîne l'expression de l'enzyme FUT1, dans les cellules intestinales conduisant à l'apparitions de motifs  $\alpha$ 1-2fuc qui constituent des ligands pour les NoVs GII.4. Pour rappel, FUT1 est normalement exprimée par les érythrocytes. Un autre aspect qui pourrait expliquer que l'infection ne corrèle pas avec le statut « sécréteur » ou « non sécréteur » des individus est lié au microbiote intestinal (Miura *et al.*, 2013). Une étude a révélé la présence d'HBGAs à la surface de certaines bactéries du microbiote et leur capacité à interagir avec des NoVs. Cette interaction pourrait donc influencer l'infection des individus

soit en conduisant à une rétention des NoVs dans l'intestin en favorisant ainsi l'infection, soit en ayant un effet protecteur par l'inhibition des sites de liaisons des NoVs aux HBGAs (Miura *et al.*, 2013).

## B. Reconnaissance d'antigènes spécifiques

En plus de l'influence du statut sécréteur, la reconnaissance des souches de NoVs va être influencée par la diversité des HBGAs à la surface de la cellule hôte. Les modalités de liaisons des norovirus des génogroupes I et II aux HBGAs apparaissent différents. Tous deux possèdent 2 sites de fixation, un premier qui est conservé au sein de chaque génogroupe et un second plus variable (Shanker *et al.*, 2017). Ces 2 sites ne sont ni localisés au même endroit dans le sous-domaine P2 de la protéine de capsid VP1 ni structurellement identiques. Pour les NoVs du génogroupe GI, le site primaire correspond au résidu de Gal tandis que pour le génogroupe GII il s'agit du Fuc lié en  $\alpha$ 1-2 du Gal de l'HBGA. Ces préférences conduisent donc les souches GI à reconnaître majoritairement des Ags H et/ou Lewis et pour les NoVs GII à reconnaître des Ags ABH. Leur second site de liaison est plus variable et permet l'évolution de reconnaissance aux HBGAs des souches émergentes (Shanker *et al.*, 2017).

Dans le génogroupe I, la liaison des NoVs des génotypes GI.1, GI.2, GI.3, GI.5 et GI.7 se fait majoritairement sur l'Ag H. A noter que les génotypes GI.2 et GI.3 reconnaissent également fortement les Ags de Lewis tout comme les génotypes GI.4, GI.6, GI.8 et GI.9 (Kimura-Someya *et al.*, 2022 ; Someya, 2022). Tous les NoVs GI, quel que soit le génotype considéré, se fixent à l'Ag Le<sup>b</sup>. Au contraire, ils ne se fixent jamais aux Ags B à l'exception du génotype GI.6 qui reconnaît fortement l'Ag B de la même manière que l'Ag A (Someya, 2022). Parmi tous ces génotypes, la souche GI.1 Norwalk qui a été le premier NoV identifié, est une des plus étudiées. Cette souche reconnaît préférentiellement les Ags H et Le<sup>b</sup> et plus faiblement les Ags A et Le<sup>y</sup> (Huang *et al.*, 2003, 2005 ; Hutson *et al.*, 2003, 2005 ; Marionneau *et al.*, 2002 ; Rockx *et al.*, 2005 ; Someya, 2022 ; Tan et Jiang, 2005).

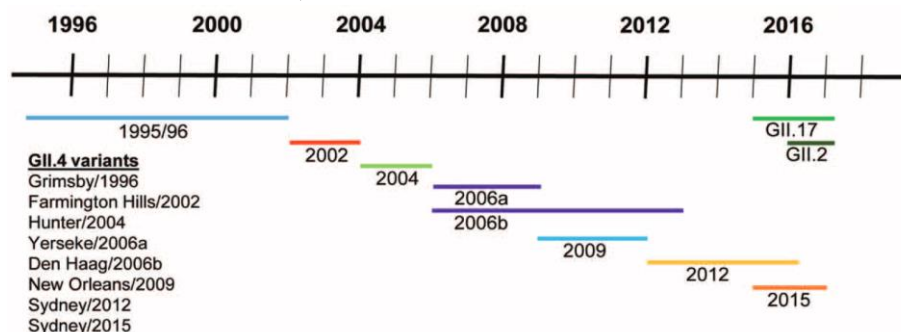
Les NoVs du génogroupe II sont moins sélectifs dans la reconnaissance des Ags due à leur premier site de liaison au Fuc $\alpha$ 1-2 qui de fait sélectionne tous les Ags du statut sécréteur. Le deuxième site de liaison, plus variable que le premier, serait impliqué dans la reconnaissance des Ags Le<sup>a</sup> par certains génotypes ce qui expliquerait les infections de population d'individus non sécréteurs. Les NoVs GII.4, GII.9, GII.10, et GII.17 ont une affinité de liaison aux HBGAs plus importante sur le premier

site alors que les NoVs GII.13 l'ont pour le second ce qui suggère une préférence des Ags des individus sécréteurs (Ags ABH et Le<sup>b</sup>) pour les premiers et non sécréteur (Ag Le<sup>a</sup>) pour GII.13 (Kocak, 2019). Autre exemple, les NoVs GII.17 reconnaissent les Ags sécréteur ABH/Le<sup>b</sup> (Zhang *et al.*, 2015) alors que la souche GII.9 VA207 les Ags A et O mais aussi l'Ag Le<sup>a</sup> (Huang *et al.*, 2003). Même si la reconnaissance est préférentielle sur le premier site de liaison, elle est quand même possible pour le deuxième ce qui élargit davantage le type d'Ags reconnus des génotypes II. Cependant, le génotype qui reconnaît la plus grande diversité d'Ags est celui GII.4 qui est capable d'infecter tous les individus sécréteurs, peu importe leur phénotype ABH et Lewis. De plus, au fur et à mesure des mutations, il est capable de reconnaître les Ags Le<sup>a</sup> (Huang *et al.*, 2003, 2005 ; Jin *et al.*, 2013 ; Larsson *et al.*, 2006 ; Rockx *et al.*, 2005 ; Tan et Jiang, 2005).

## C. Evolution des Norovirus

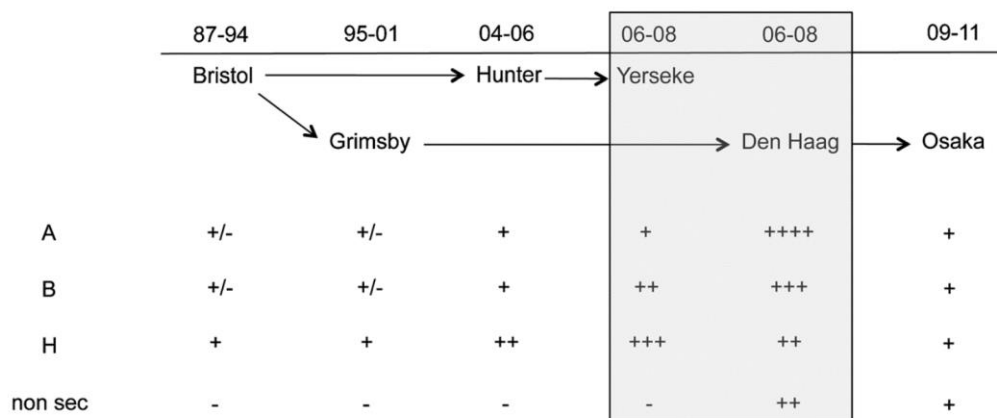
### 1. Variants des norovirus GII.4

Dans le monde, les NoVs GII.4 sont dominants et persistent chez l'homme mais évoluent continuellement. Entre 2002 et 2016, un nouveau variant GII.4 apparaît environ tous les 2-3 ans environ et remplace le précédent en devenant à son tour le plus prédominant et causant une nouvelle pandémie dont la durée peut être variable. Pour illustrer cette évolution constante des NoVs, 6 variants du génotype GII.4 se sont succédés entre 1995 et 2012 (**Figure 18**). Les épidémies causées par les variants GII.4 Grimsby avant 2002 et Sydney en 2012 ont duré plus longtemps avec une durée de plus de 10 ans pour ce dernier. Depuis 2012, le NoV GII.4\_Sydney\_2012 est le variant dominant. En 2015, il a subi un changement polymérase, il est alors passé de Sydney 2012 [P31] à Sydney 2012 [P16] (Atmar *et al.*, 2018 ; Parra, 2019 ; Parra *et al.*, 2023).



**Figure 18:** Emergence des variants GII.4 de 1995 à 2017. Issu de Atmar *et al.* (2018).

Une des raisons qui explique la persistance des NoVs du génotype GII.4 réside dans la capacité des différents variants à reconnaître un large panel d’HBGAs et ainsi à infecter une grande proportion de la population (80-95%). Quant à l’émergence rapide de nouveaux variants, elle s’explique par le taux d’erreur plus important de la RNA polymérase des GII.4 par rapport aux souches de NoVs non-pandémiques (5 à 36 fois plus en comparaison à des NoVs GII.3 et GII.7) et s’explique également par la pression immunitaire (Bull *et al.*, 2010). L’apparition d’un nouveau variant s’accompagne souvent de changements pour la reconnaissance des motifs HBGAs comme l’illustre la **Figure 19**. On peut noter qu’au gré de l’apparition de nouveaux mutants, ces derniers reconnaissent davantage d’HBGAs avec des affinités variables mais qui tendent à être plus importantes avec les souches récentes (Sydney, GII.4[P16]) (Liang *et al.*, 2021 ; Ruvoën-Clouet *et al.*, 2013). On notera également la capacité des variants apparus en 2006 et en 2009 à infecter des individus Lewis-positif non sécréteur.



**Figure 19:** Evolution des reconnaissances des HBGAs par les variants GII.4 de 1987 à 2011. Issu de Ruvoën-Clouet *et al.* (2013).

En plus des variants pandémiques qui apparaissent au sein du génotype GII.4, on peut ajouter qu’occasionnellement, d’autres génotypes peuvent engendrer des épidémies dans certains pays comme le génotype GII.17 qui a frappé l’Asie entre 2014 et 2016 ou encore le génotype GII.2[P16] qui a sévi entre 2016 et 2017 dans de nombreux pays tels que la Chine, le Japon ou encore l’Allemagne (Atmar *et al.*, 2018 ; de Graaf *et al.*, 2015 ; Tohma *et al.*, 2017).

## 2. Reconnaissance fine des HBGAs

La reconnaissance des HBGAs par les NoVs est tellement fine qu’elle dépend parfois du changement d’un seul AA. Ce phénomène peut être observable entre génotypes. C’est notamment le cas

du génotype GI.9 qui reconnaît préférentiellement les Ags de Lewis (Kimura-Someya *et al.*, 2022). Son mode de liaison implique le résidu de Gal (comme tous les GI) et une interaction avec un résidu  $\alpha$ 1-3/4Fuc (Lewis). Son mode de liaison est similaire à celui de GI.8. Cependant, le génotype GI.8 ne reconnaît pas l'Ag Le<sup>x</sup>. Cela s'explique par la présence d'un résidu d'arginine en position 403 pour GI.9 impliquée dans la liaison à l' $\alpha$ 1-3Fuc qui correspond à un résidu de thréonine en position 400 pour GI.8. Le résidu de thréonine ne permet pas cette liaison et donc ne permet pas la reconnaissance de l'Ag Le<sup>x</sup> (Kimura-Someya *et al.*, 2022).

Un phénomène similaire s'observe également entre certains variants de GII.4 et le génotype GII.13. Ces virus contiennent une boucle dont la séquence est Arg<sub>339</sub>-Ala-Asp-Gly<sub>342</sub> pour les variants GII.4 et Val<sub>339</sub>-Ser-Glu-Gly<sub>342</sub> pour GII.13 (Kocak, 2019). Le simple remplacement de l'Arg par la Val change la stabilité de cette boucle qui est flexible avec l'Arg et rigide avec la Val. Alors que la flexibilité de la boucle permet aux NoVs GII.4 de reconnaître des HBGAs sur leur 2 sites de fixation, la rigidité de la boucle pour GII.13 se traduit par une fixation préférentielle des HBGAs dans le deuxième site de liaison due à l'encombrement stérique créé dans le premier site par la boucle (Kocak, 2019).

Enfin, au sein même du génotype GII.4 qui, comme évoqué précédemment, évolue très rapidement, il a été observé de grands changements de reconnaissance à cause du changement d'un seul AA entre certains variants. Alors que les AA impliqués dans l'interaction avec les HBGAs sont conservés, d'autres AA proches du site d'interaction vont changer et ainsi moduler l'affinité (Liang *et al.*, 2021). Ainsi, le variant GII.4 de 1997 reconnaissait les Ags A, B, H de type 3 et Le<sup>y</sup> alors que ce n'était pas le cas du variant précédent GII.4 1987 qui reconnaissait uniquement les Ags H de type 3 et Le<sup>y</sup>. Ce changement de reconnaissance a été attribué au remplacement du résidu de glycine en position 393 par un résidu d'acide aspartique (Lindesmith *et al.*, 2008).

## V. Infections à Norovirus chez l'Homme

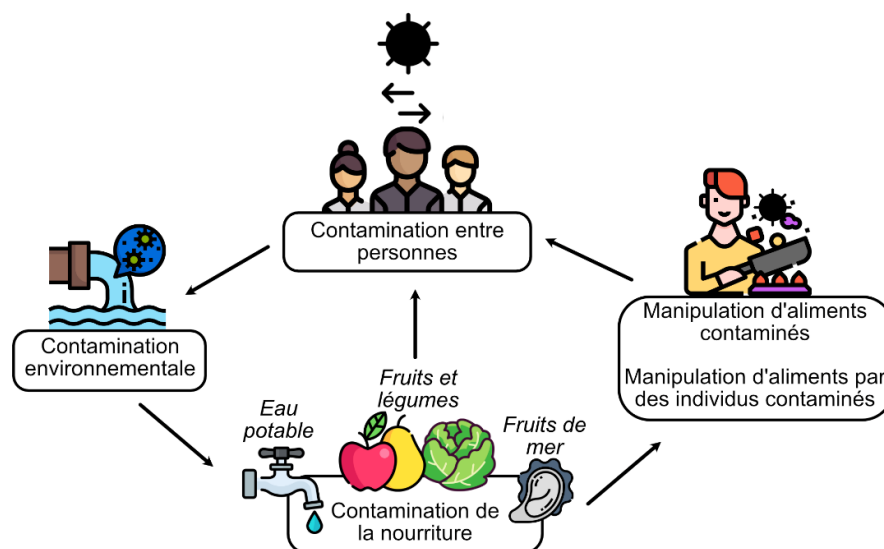
### A. Pathologie

Les NoVs causent une gastro-entérite aigüe dont les symptômes apparaissent après une période d'incubation de 24 à 72h et perdurent jusqu'à 5 jours (Bányai *et al.*, 2018). Cette phase symptomatique est principalement caractérisée par des diarrhées et des vomissements qui peuvent s'accompagner de

nausées, de crampes abdominales ou encore de fièvre. La complication la plus sévère est la déshydratation qui peut aboutir à un choc hypovolémique, au coma et à la mort. La sévérité de la maladie est liée à la pathogénicité de la souche en question ainsi qu'à l'individu infecté (génétique, comorbidités et état immunitaire). Elle apparaît plus sévère chez les jeunes enfants (moins de 5 ans) et les personnes âgées (plus de 65 ans) (Bányai *et al.*, 2018).

## B. Voies de transmission

Les particules virales sont présentes dans les fèces des individus malades et sont transmises entre individus par voie féco-orale *via* les mains ou *via* l'ingestion d'aliments contaminés. La nourriture se retrouve souillée après manipulation par un individu malade ou, pour les fruits et légumes, après lavage par de l'eau contaminée (Gao *et al.*, 2016). Elle peut aussi être contaminée directement à la source : c'est le cas des fruits de mer. En effet, les NoVs peuvent être retrouvés dans les eaux traitées, les eaux côtières et les rivières environnantes lorsque le traitement des eaux usées n'est pas suffisamment efficace, lors de fortes précipitations qui entraînent le débordement des stations d'épuration ou encore à cause de fosses septiques défectueuses. Les fruits de mer cultivés à proximité des effluents humains peuvent alors être contaminés et en particulier les huîtres (Figure 20) (Maalouf *et al.*, 2010a ; Yang *et al.*, 2021).



**Figure 20** : Illustration des voies de transmission des Norovirus à l'homme. Crédits icones : Flaticon.com

Il est difficile de limiter la transmission de la maladie à cause de sa haute contagiosité liée à plusieurs facteurs : (1) le faible niveau de particules nécessaire pour être infecté (18 à 1000 particules

virales) (Teunis *et al.*, 2008), (2) la sécrétion des particules virales par certains individus avant et après l'apparition des symptômes, (3) la stabilité des NoVs dans l'environnement (inactivation après 15 min à 60°C) (Ettayebi *et al.* 2016, Desdouits *et al.* 2022) et (4) l'absence de protection immunologique face à la diversité et la constante évolution des souches de NoVs (Bányai *et al.*, 2018).

## VI. Contamination des huîtres par les Norovirus

### A. Espèces d'huîtres en France

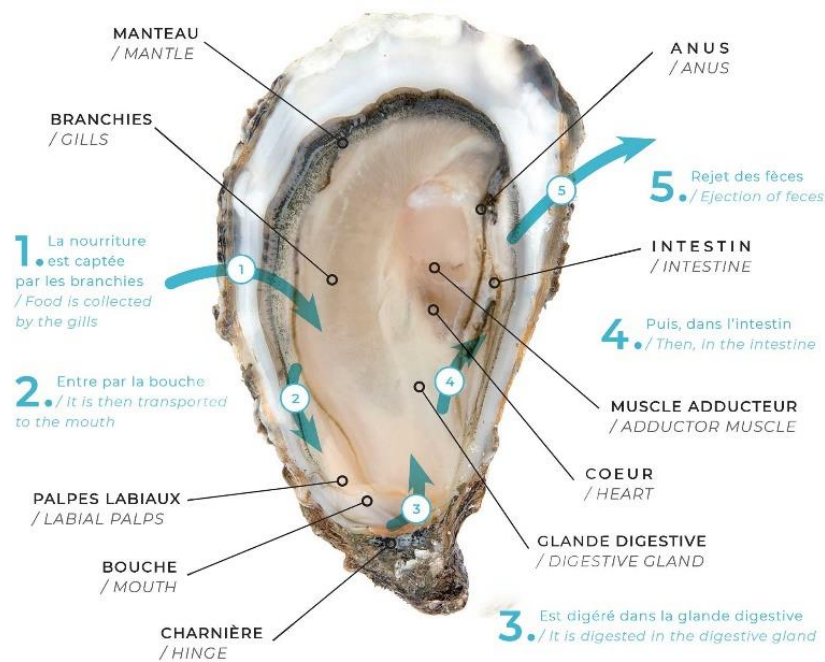
Au 19e siècle, l'huître plate indigène d'Europe, *Ostrea edulis*, a presque été décimée à cause de sa surpêche et d'infections par des parasites comme *Marteilia refringens* et *Bonamia ostrae*. Pour relancer son ostréiculture, la France a importé d'autres espèces, en particulier l'huître portugaise *Crassostrea angulata* qui a à son tour été décimée dans les années 70. Pendant le déclin de cette dernière, l'huître creuse *Crassostrea gigas* avait été importée du Japon et du Canada afin d'augmenter la production d'huîtres en France. Cette espèce est aujourd'hui la plus consommée en France mais aussi dans le monde (Botta *et al.*, 2020). La France continue de cultiver en parallèle l'huître *O. edulis*. Cette dernière, avec l'huître *C. gigas*, ont fait l'objet des travaux présentés dans cette thèse. L'huître *C. gigas* est l'huître la plus consommée dans le monde et est pour cela l'une des espèces les plus étudiées avec l'espèce *Crassostrea virginica* (cultivée sur la côte Est des Etats-Unis).

### B. Biologie de l'huître

L'huître est un mollusque bivalve hermaphrodite qui devient successivement mâle et femelle tout au long de sa croissance. Sa coquille est constituée d'une valve gauche inférieure et d'une valve droite supérieure articulées au niveau d'une charnière constituée d'un ligament (Marteil, 1976). La première est généralement bombée alors que la seconde est plate. Contrairement à la charnière, le muscle adducteur permet le rapprochement des deux valves et ainsi la fermeture de la coquille. A l'intérieur, un tissu conjonctif appelé manteau compose la partie externe du corps mou et participe à la croissance et au développement de la coquille. Ses branchies en font un organisme filtrant qui lui permettent de respirer mais aussi de s'alimenter. L'eau traverse les branchies et l'oxygène est capté pour contribuer à la respiration. Le phytoplancton qui est capté à cette occasion est redirigé vers la bouche après enrobage de mucus par l'action des cils à la surface de l'épithélium branchiale. Avant d'arriver à la bouche, les



palpes labiaux effectuent un premier tri des particules nutritives qui sont ensuite transmises au système digestif (**Figure 21**) (Marteil, 1976). Une partie des aliments ingérés est phagocytée par des hémocytes mobiles qui constituent le système immunitaire de l'huître tandis que l'autre est éliminée sous forme de fèces. Les virus présents dans l'environnement sont généralement adsorbés sur des particules leur permettant d'être suffisamment gros pour être filtrés par les huîtres. Les hémocytes vont les phagocyter dans leurs vacuoles acides avant de les transporter autour de l'intestin. Les NoVs, étant stables en milieu acide, arriveront donc sous forme intacts au niveau de ce tissu (Yang *et al.*, 2021).



**Figure 21:** Anatomie et mécanisme de nutrition de l'huître. Crédit : huîtresmareennesoleron.info

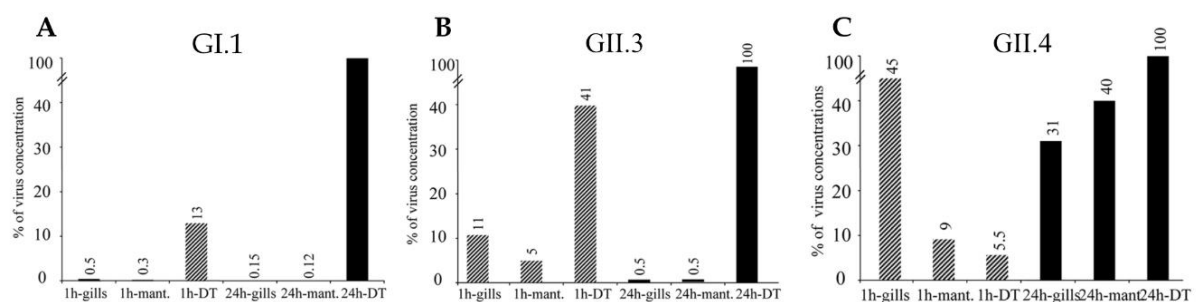
### C. Vecteur de Norovirus

L'huître apparaît comme un excellent vecteur de contamination des hommes par les NoVs (Yu *et al.*, 2015). Elle est cultivée dans les zones côtières à proximité des effluents et renferme parfois des polluants ainsi que des microbes pathogènes pour l'homme (Griffin *et al.*, 2003). Pour sa respiration et son alimentation, l'huître filtre des volumes d'eau très importants (jusqu'à 5L/h) ce qui l'amène à concentrer ces agents toxiques. De surcroît, sa consommation crue par l'homme ne permet pas de détruire ces pathogènes. Au-delà de cette accumulation « mécanique », des études ont démontré une bioaccumulation des NoVs médiée par des interactions spécifiques conduisant à une rétention sur une longue période. L'huître n'est pas infectée par les NoVs qu'elle bioaccumule, ces virus ne s'y répliquent

donc pas mais y sont stockés et peuvent y persister plusieurs semaines malgré une épuration dans de l'eau non contaminée ou un reparaillage dans une nouvelle zone (Le Guyader *et al.*, 2006, 2012).

### 1. Bioaccumulation et rétention spécifique des Norovirus

Des expériences de bioaccumulation artificielle ont été réalisées dans des bassins contenant des huîtres *C. gigas* immergées dans de l'eau de mer contaminée par une souche virale (Le Guyader *et al.*, 2006 ; Maalouf *et al.*, 2011 ; Su *et al.*, 2018). Après des temps de bioaccumulation variables, les huîtres ont été disséquées et les NoVs ont été quantifiés par RT-qPCR dans différents tissus (branchies, le tissu digestif et le manteau). Ces études ont montré que les NoVs GI.1 et GII.3 se concentrent majoritairement dans le tissu digestif. Toutefois, pour les NoVs GII.3, il a été montré qu'ils sont d'abord retrouvés dans le manteau et les branchies avant d'être observés uniquement dans le tissu digestif 24h plus tard. A l'inverse, les NoVs GII.4 sont plus faiblement bioaccumulés. Leur bioaccumulation est d'abord observée au niveau des branchies, avant d'être retrouvés dans les 2 autres tissus avec une préférence pour le tissu digestif après 24h de bioaccumulation (**Figure 22**) (Le Guyader *et al.*, 2006 ; Maalouf *et al.*, 2011 ; Su *et al.*, 2018). Une autre étude de bioaccumulation dans le tissu digestif de l'huître *C. gigas* utilisant les NoVs GII.17 et GII.2 révèle uniquement la bioaccumulation des NoVs GII.17 (Ueki *et al.*, 2021).

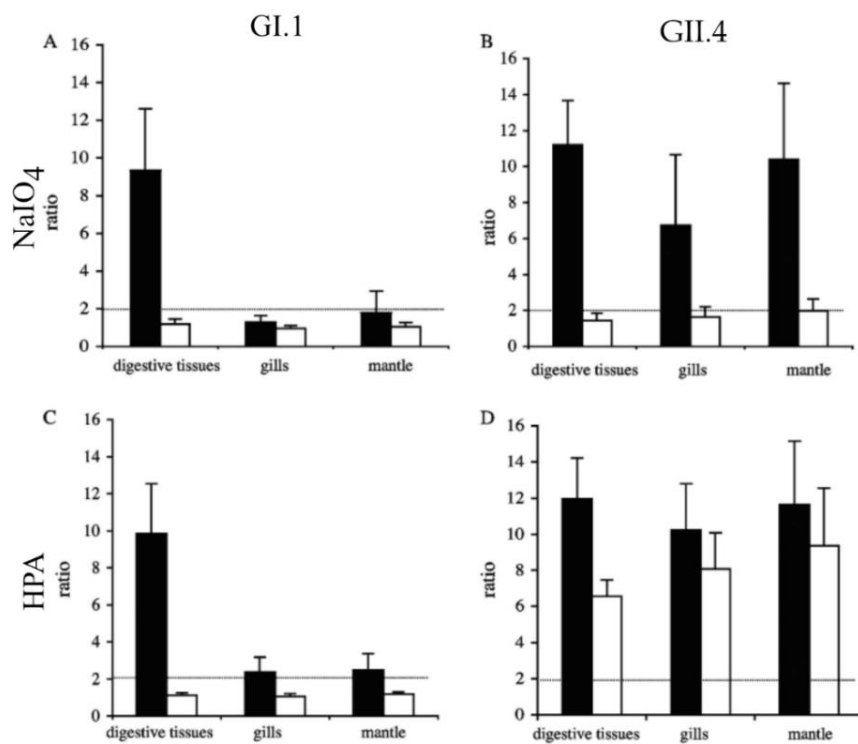


**Figure 22 :** Bioaccumulation des norovirus GI.1, GII.3 et GII.4 dans différents tissus de l'huître *Crassostrea gigas*. La contamination par les NoVs (A) GI.1, (B) GII.3 et (C) GII.4 est mesurée dans les différents tissus après bioaccumulation. Les résultats sont exprimés en pourcentage après 1h (barres grises) et 24h (barres noires). Gills : branchies ; mant : manteau ; DT : tube digestif. Issues de Maalouf *et al.* (2011).

Les huîtres à l'origine d'une épidémie de gastro-entérite à l'hiver 2006 en France ont été analysées sur 4 semaines. Les résultats ont montré une involution du nombre d'huîtres contaminées par le génogroupe GI de 59 % à 41 % et par le génogroupe GII de 70 % à 17 %. Pourtant, ces huîtres avaient été préalablement épurées pendant 2 semaines (Le Guyader *et al.*, 2008). Cette étude montre la

persistance des NoVs dans les tissus d’huître et est en accord avec les expériences de bioaccumulation qui montrent une contamination plus importante des huîtres par le génogroupe GI que celui GII. Les auteurs suggèrent que le génogroupe GII est moins stable dans l’environnement ce qui expliquerait qu’il s’élimine plus rapidement des tissus d’huître.

Des tests ELISA réalisés sur des homogénats de tissus de *C. gigas* ont montré qu’il existe bien une reconnaissance spécifique par les NoVs et qu’elle diffère selon les génotypes. Ainsi, le génotype GI.1 a montré des reconnaissances pour le tissu digestif alors que le génotype GII.4 reconnaît à la fois le tube digestif, le manteau et les branchies (**Figure 23**) (Maalouf *et al.*, 2010b).



**Figure 23** : Inhibition de liaison des VLPs de norovirus sur des extraits tissulaires de l’huître *Crassostrea gigas*. Liaison des VLPs (A et C) GI.1 et (B et D) GII.4 sur les tissus de *Crassostrea gigas* après un (A et B) prétraitement au périodate de sodium ou une (C et D) préincubation avec la lectine *Helix pomatia*. Les barres noires représentent les homogénats sans traitement et les barres blanches avec les traitements inhibiteurs. Digestive tissues : tube digestif ; gills : branchies ; mantle : manteau. Issue de Maalouf *et al.* (2010b).

L’ensemble de ces études indiquent que les huîtres exposeraient des ligands différents selon les tissus dont la reconnaissance dépend de la souche de NoVs et qui conduit à leur rétention.

## 2. Ligands des Norovirus chez l'huître

Les NoVs étant connus pour reconnaître les HBGAs chez l'homme, ces épitopes ont été recherchés et mis en évidence chez l'huître essentiellement grâce à des expériences d'immunohistochimie et d'ELISA.

Pour démontrer l'implication des glycanes dans la reconnaissance des NoVs chez l'huître *C. gigas*, Maalouf *et al.*, (2010b) ont effectué un prétraitement des extraits du tube digestif, des branchies et du manteau avec du périodate de sodium connu pour altérer la structure des oligosaccharides. Les tissus traités ont ensuite été incubés avec des VLPs GI.1 et GII.4 et il a été observé une abolition des reconnaissances démontrant l'implication des glycanes (**Figure 23**). Cette observation a été confirmée par d'autres études semblables (Le Guyader *et al.*, 2006 ; Tian *et al.*, 2007).

Pour aller plus loin dans la détermination des ligands glycaniques, des études utilisant différents anticorps dirigés contre les HBGAs ont démontré la présence majoritaire d'Ags A et à un moindre degré d'Ags H chez l'huître *C. gigas* (Le Guyader *et al.*, 2006 ; Morozov *et al.*, 2018 ; Su *et al.*, 2018 ; Tian *et al.*, 2007). Les études ont essentiellement concerné le tissu digestif. Pour celles qui ont porté sur le manteau et les branchies, des résultats légèrement différents ont été obtenus. Ainsi, les travaux de Su *et al.* (2018) montrent la présence d'Ags A et H dans les 3 tissus et d'Ags Le<sup>b</sup> uniquement pour les branchies tandis que ceux de Morozov *et al.* (2018) montrent uniquement la présence d'Ag H pour les branchies et aucun Ag pour le manteau.

Dans le but de vérifier l'implication de l'Ag A voire H dans le mécanisme de rétention des NoVs par les huîtres, des tests d'inhibition de reconnaissance ont été effectués. Ainsi, les travaux de Le Guyader *et al.* (2006) montrent qu'une préincubation d'extraits du tube digestif de *C. gigas* avec de la salive d'individus sécréteurs A ou d'individus sécréteurs O entraîne une inhibition complète de la reconnaissance par les VLPs GI.1 pour la salive de sujet de groupe A sécréteur et une diminution partielle pour la salive des individus de groupe O sécréteurs. En revanche, aucun changement n'est observé en utilisant de la salive d'individus sécréteur B ou non sécréteur (**Tableau 2**). La pré-incubation des tissus d'huître avec la lectine *Helix pomatia* (HPA) qui est spécifique des résidus de GalNAc et par extension de l'Ag A, entraîne une inhibition totale de la reconnaissance par les VLP GI.1 (**Figure 23**)

(Maalouf *et al.*, 2010b). Enfin, les travaux de Tian *et al.* (2007), utilisant des anticorps anti-A sur le tissu digestif de *C. gigas* avant leur incubation avec un mélange de VLPs de différentes souches de NoVs diminue significativement la reconnaissance ce qui n'est pas le cas des anticorps anti-B ou anti-O.

Ces différentes expériences d'inhibition suggèrent fortement une implication des Ags A dans la fixation spécifique des NoVs sur différents tissus d'huître (Le Guyader *et al.*, 2006 ; Maalouf *et al.*, 2010b ; Tian *et al.*, 2007). En revanche, d'autres Ags glycaniques pourraient être impliqués. Dans les travaux de Maalouf *et al.* (2010b), l'inhibition de reconnaissance de GII.4 par HPA n'atteint que 50 % pour le tissu digestif et n'est pas significative pour les autres tissus (**Figure 23**). Cette observation suggère donc la présence d'un ou plusieurs autre(s) ligand(s) glycanique(s) que l'Ag A pour cette souche.

**Tableau 2** : Inhibition de la liaison de VLPs GI.1 de Norwalk sur le tissu digestif par de la salive. Adapté de Le Guyader *et al.* (2006).

Salive	Nb. échantillons testés	Liaison des VLPs*
A	14	–
O	18	+
B	4	+++
Non sécréteur	5	+++
Contrôle	5	+++

\*Après incubation avec de la salive, aucune (-), peu (+), ou beaucoup (+++) de reconnaissance par les VLPs au tissu digestif de l'huître a été observé. Pour les contrôles, les VLPs ont été ajoutées aux tissus d'huîtres dans du tampon salin phosphate/ 1% Serum Albumine Bovine à la place de l'échantillon de salive dilué dans le même tampon.

D'un point de vue structural, aucune étude à notre connaissance des glycomes, *N*- et *O*-, chez *C. gigas* n'a été réalisée. Les seuls travaux dans ce sens sont ceux de Kurz *et al.* en 2013 portant sur les *N*-glycannes de l'hémolymphe et du plasma de l'huître *C. virginica* et qui ont confirmé la présence d'Ag A chez cette dernière. De manière plus intéressante, cette étude a révélé la présence d'Ags A substitués par des groupements méthyles et sulfates. Si les modifications aglycones identifiées chez *C. virginica* sont également présentes chez *C. gigas*, elles pourraient expliquer les différences de reconnaissances tissulaires de manière qualitatives et quantitatives par les NoVs.

### 3. Contrôle des parcs à huître

La culture des coquillages, dont l'ostréiculture, est soumise à des normes européennes très strictes, notamment microbiologiques, qui consiste à suivre leur contamination par la bactérie *Escherichia coli* qui est un indicateur de contamination fécale (Le Guyader *et al.*, 2019). Selon la quantité

d'*E. coli*, les cultures sont classées en 3 catégories. Celles de la catégorie A sont directement commercialisables, celles de la catégorie B doivent être préalablement purifiées en eau non contaminée avant d'être commercialisées et celles de la catégorie C doivent être reparquées dans une zone de bonne qualité microbiologique ou doivent subir un traitement thermique (Le Guyader *et al.*, 2019).

Pour les huîtres, ces mesures permettent d'éliminer efficacement les bactéries mais ne sont pas suffisantes pour l'élimination des virus et particulièrement les NoVs. Une étude réalisée sur l'huître *C. virginica* a montré une élimination de 95 % de *E. coli* après 48h de dépuración mais seulement 7 % des NoVs au niveau génomique (Schwab *et al.*, 1998). Néanmoins, cette étape peut être améliorée en changeant la température. Les travaux de Choi et Kingsley (2016) réalisés sur l'huître *C. virginica* artificiellement contaminée montrent que la dépuración réalisée à 25 °C au lieu de 7 et 15 °C réduit la persistance du NoV GI.1 de 6 à 4 semaines.

Lorsque des gastroentérites associées à la consommation d'huîtres sont détectées, les zones de production peuvent faire l'objet de fermeture temporaire et les huîtres sont alors retirées de la vente. Des événements climatiques particuliers ou des défauts de traitement des eaux usées etc... peuvent conduire aux mêmes mesures (Le Guyader *et al.*, 2019).

Pour des individus sécréteurs, qui constituent la majorité de la population française (80 %), la dose infectieuse médiane (ID50) est de 1,6 à 7,51 copies de génome de NoV par huître consommée (Thebault *et al.*, 2013). Ces quantités sont très faibles et difficilement détectables. A l'heure actuelle, leur détection se fait par RT-qPCR mais cette méthode ne permet pas de différencier les NoVs infectieux des non-infectieux.

Trouver le moyen de prévenir les épidémies causées par la contamination des huîtres par les NoVs est donc un enjeu de santé publique et économique important. Les moyens actuels ne permettent pas d'éviter les infections virales par des huîtres contaminés par les NoVs. Il est donc essentiel de comprendre comment les NoVs sont retenus dans les tissus d'huître en identifiant très précisément la structure moléculaire du ou des ligand(s) impliqué(s).

## Chapitre 3 : ANALYSE DES GLYCANNES PAR SPECTROMETRIE DE MASSE

### I. Libération des glycannes

Pour analyser les chaînes oligosaccharidiques associées à une protéine, il faut au préalable les libérer. Elles peuvent l'être de manière enzymatique ou chimique.

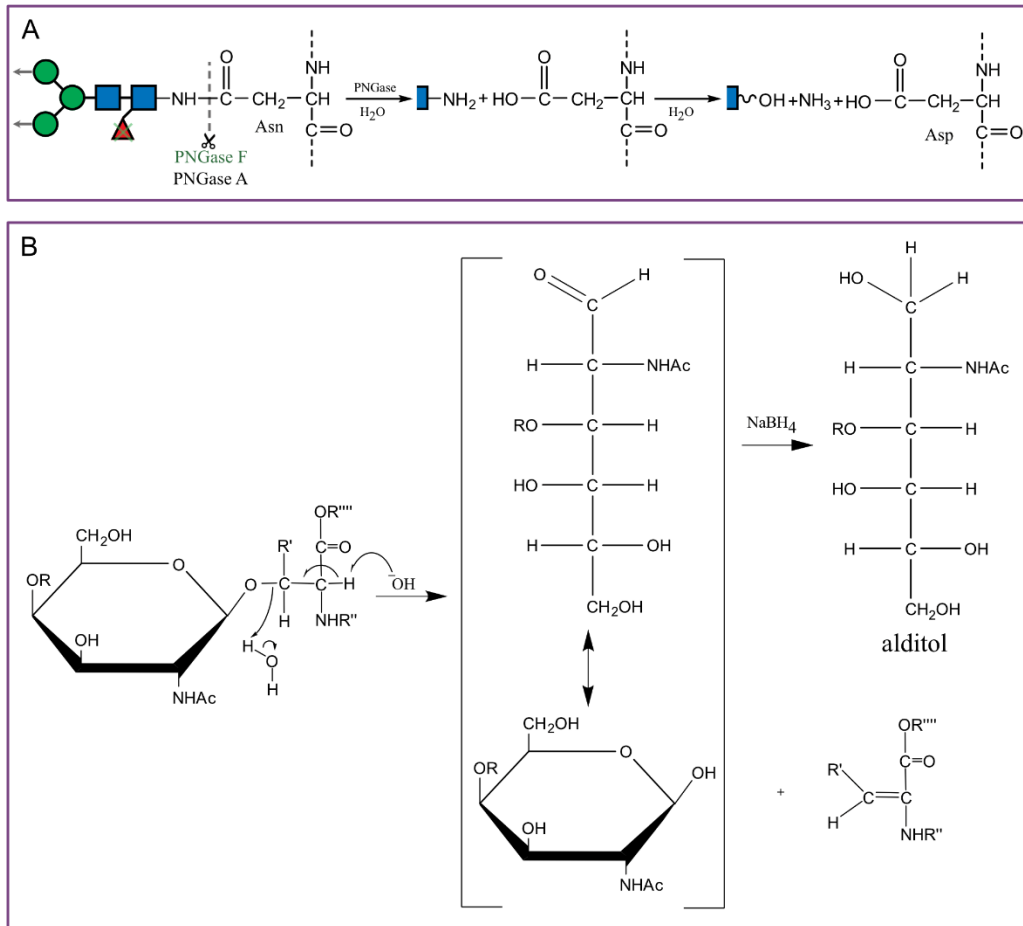
Plusieurs enzymes sont utilisables pour la libération des *N*-glycannes : (1) la Peptide-*N*-glycosidase A (PNGase A) qui hydrolyse la liaison entre l'Asn d'un séquon et la GlcNAc terminale de l'oligosaccharide, (2) la PNGase F qui a la même action mais ne peut pas agir sur un *N*-glycanne dont la GlcNAc du point d'attache est substituée par un résidu de Fuc sur l'hydroxyle de son C3 et (3) l'endoglycosidase H (Endo H) qui va hydrolyser la liaison osidique entre les 2 résidus de GlcNAc du chitobiose des *N*-glycannes de type hybride et oligomannose (Maley *et al.*, 1989 ; Tretter *et al.*, 1991). Pour la libération enzymatique des *O*-glycannes, il existe une Endo- $\alpha$ -*N*-acetylgalactosaminidase (*O*-glycanase) mais son action est fortement dépendante de la structure du *O*-glycanne de type mucine et est donc très peu utilisée (Iwase et Hotta, 1993).

La libération chimique quant à elle peut être réalisée par hydrazinolyse et va libérer les *N*- et *O*-glycannes en fonction des conditions expérimentales. Les *O*-glycannes peuvent aussi être libérés en conditions basiques par  $\beta$ -élimination suivie d'une réduction au borohydrure et va aboutir à la formation d'alditols. Ces méthodes chimiques ne vont pas permettre de garder les protéines intactes et/ou résultent en l'obtention de glycannes modifiés au niveau du résidu de GlcNAc réducteur.

Durant ma thèse, les *N*-glycannes ont été libérés par les PNGases F et A et les *O*-glycannes l'ont été par  $\beta$ -élimination en milieu basique (**Figure 24**).

Il existe 2 grandes méthodes physico-chimiques pour déterminer la structure des oligosaccharides libérés qui sont la spectrométrie de masse et la RMN. Si la sensibilité de cette dernière s'est considérablement améliorée, elle reste néanmoins insuffisante dès lors qu'il s'agit de travailler sur de faibles quantités. La spectrométrie de masse, quoique beaucoup moins précise dans l'établissement

de la structure primaire des oligosaccharides que la RMN (nature des monosaccharides, de leur cyclisation, nature de leur liaison comprenant l'anométrie et la position) reste la méthode de choix pour établir un *N*- ou *O*-glycome.



**Figure 24 :** Libération des *N*- et *O*-glycannes. (A) Libération enzymatique des *N*-glycannes par les PNGase F et A. Le PNGase F agit sur une *N*-glycanne non substitué sur le carbone 3 du résidu de GlcNAc réducteur. (B) Libération chimique des *O*-glycannes par  $\beta$ -élimination en milieu basique (NaOH) suivi d'une réduction par le borohydrure de sodium (NaBH<sub>4</sub>).

## II. Composition d'un spectromètre de masse

Un spectromètre de masse est un appareil qui permet de mesurer très précisément le rapport masse sur charge (noté  $m/z$ ) d'un ion et donc de déterminer la masse moléculaire de la molécule qui lui a donné naissance. Appliqué à un mélange de molécules, il permet de les discriminer sur la base de leur rapport  $m/z$ . Un spectromètre de masse est toujours constitué de la même manière : une source d'ionisation capable de générer des ions, un analyseur qui va permettre la séparation des ions formés dans la source selon leur rapport  $m/z$  et un détecteur pour convertir les ions en signal électrique qui sera amplifié et digitalisé.



De plus en plus de spectromètres de masse permettent d'obtenir, en plus du rapport  $m/z$  des ions, des informations structurales les concernant. Un ion parent de  $m/z$  donné est isolé et fragmenté puis les ions fils générés sont analysés par un autre analyseur ou le même. On parle alors de spectrométrie de masse en tandem MS/MS ou MS<sup>2</sup>. Si ce processus de fragmentation est répété  $n$  fois, on parle alors de spectrométrie de masse MS <sup>$n$</sup> .

Un des objectifs de ma thèse était d'établir précisément les *N*- et *O*-glycomes de différents tissus des huîtres *C. gigas* et *O. edulis*. En outre, nous voulions déterminer les types de monosaccharides, leur enchaînement, leur liaison ainsi que leurs possibles modifications. Pour y parvenir, nous avons utilisé un spectromètre de masse MALDI-QIT-TOF, une chromatographie en phase gazeuse couplée à la spectrométrie de masse GC-EI-MS et un spectromètre de masse ESI-MS/MS couplé à un dispositif de séparation de type nanoLC : l'Orbitrap Fusion™ Tribrid™. Le fonctionnement de chacun des appareils sera décrit ci-dessous et sera accompagné du traitement subi par les échantillons ainsi que de la manière dont ont été interprétés les résultats obtenus.

### III. Analyse des *N*- et *O*-glycannes par MALDI-QIT-TOF

#### A. Fonctionnement du MALDI-QIT-TOF

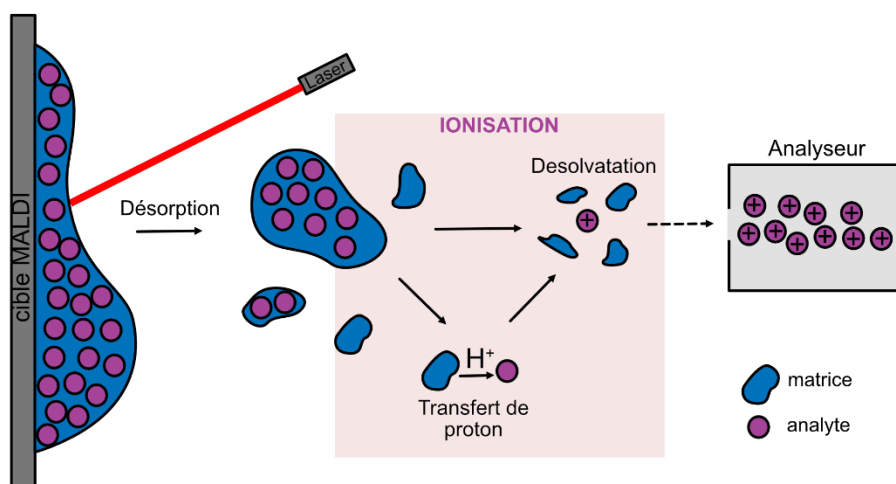
##### 1. Description

Le spectromètre de masse MALDI-QIT-TOF se compose d'une source d'ionisation MALDI (Matrix-Assisted Laser Desorption/Ionization) et de 2 analyseurs : une trappe à ion QIT (Quadrupole Ion Trap) et un analyseur à temps de vol TOF (Time Of Flight). Pour cet appareil, c'est l'analyseur TOF qui permet de mesurer les différents  $m/z$ . L'association des deux analyseurs QIT et TOF permet :

- d'augmenter la sensibilité due à l'accumulation d'ions formés dans le QIT pour ensuite séparer les ions dans l'analyseur TOF,
- de mesurer avec une meilleure résolution les  $m/z$  avec l'analyseur TOF plutôt que l'analyseur QIT,
- permet d'effectuer des analyses MS <sup>$n$</sup>  en utilisant l'analyseur QIT comme cellule de collision (appelée Collision Induced Dissociation (CID)).

## 2. Source d'ionisation MALDI

La source d'ionisation MALDI fait partie des méthodes d'ionisation douce n'entraînant pas la destruction des molécules analysées. Elle consiste en la désorption, par un tir laser, de co-cristaux d'un mélange matrice/analytes permettant la formation d'analytes ionisés en phase gaz (Israr *et al.*, 2020 ; Karas et Hillenkamp, 1988). L'échantillon d'intérêt est mélangé à une grande quantité de matrice puis un micro-dépôt est réalisé sur une cible métallique dont l'évaporation du solvant va permettre la cristallisation. La cible est ensuite introduite dans la source d'ionisation MALDI sous vide et le dépôt est irradié par un laser qui va causer la désorption des molécules d'analytes et de matrice (**Figure 25**). Lors de cette phase, les analytes seront ionisés puis envoyés à l'élément suivant de l'appareil.



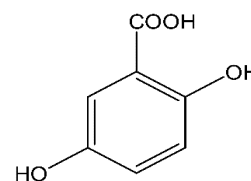
**Figure 25** : Fonctionnement de la source d'ionisation MALDI. Les analytes cocrystallisés avec des molécules d'une matrice sont désorbés par un laser qui va permettre l'ionisation dans analytes avant leur envoi à l'analyseur.

Les processus d'ionisation engendrés par le MALDI sont complexes et encore assez mal compris mais les travaux réalisés jusqu'à présent suggèrent la coexistence de deux voies d'ionisation (**Figure 25**) (Hillenkamp *et al.*, 2013 ; Jaskolla et Karas, 2011). Pour la première, les analytes sont neutres dans le co-cristal et vont obtenir leur charge lors de la désorption consécutive au laser. En outre, les photons émis par le laser vont exciter les molécules de matrice. Les échanges d'énergies qui vont avoir lieu au sein des molécules de matrice vont aboutir à la création de molécules de matrices donneur  $[M + H]^+$  et accepteur de proton  $[M + H]^-$  (ou donneur de cation) et leurs interactions avec les analytes vont engendrer des analytes ionisés. Pour la seconde voie d'ionisation, les analytes sont déjà chargés au sein du co-cristal mais neutralisés par des contre-ions. Lors de l'irradiation, les ions sont simplement désorbés (Hillenkamp *et al.*, 2013 ; Jaskolla et Karas, 2011).

La matrice va avoir un rôle d'intermédiaire entre l'énergie primaire très élevée du laser et la molécule d'intérêt. On la choisit en fonction du type de molécule à analyser mais elle doit remplir certaines conditions décrites ci-dessous (Fuchs et Schiller, 2009) :

- (1) Les molécules de matrice doivent absorber à la longueur d'onde d'émission du laser qui émet dans l'infrarouge (IR) ou dans l'ultraviolet (UV). L'émission UV est la plus utilisée pour les biomolécules avec l'utilisation d'un laser à azote (337nm) ou Nd:YAG (neodymium-doped yttrium aluminum garnet, 355nm). Les matrices possèdent donc souvent un cycle aromatique permettant la délocalisation d'électrons.
- (2) La matrice doit être donneur ou accepteur de proton pour générer des analytes ionisés. Des molécules disposant d'une fonction carboxylique sont beaucoup utilisées ce qui contribue aussi à la solubilité de la matrice dans des solvants polaires.
- (3) La matrice doit être sublimable.
- (4) La matrice doit être stable sous vide poussé et donc jusqu'à sa désorption par le laser.
- (5) La matrice doit permettre l'isolement des ions générés les uns des autres pour éviter la formation de clusters d'analytes ce qui explique l'importance d'utiliser la matrice en excès (rapport matrice/analyte compris entre 5 000 et 10 000).
- (6) Enfin, la matrice doit permettre la formation de cristaux homogènes avec les analytes.

La molécule de matrice la plus utilisée pour l'analyse des glycanes est l'acide 2,5-dihydroxybenzoïque (DHB) qui remplit tous ces critères et qui est celle qui a été utilisée lors de ma thèse (Figure 26).

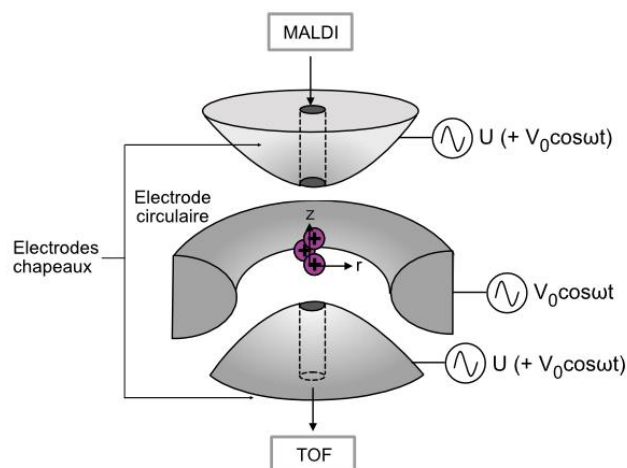


**Figure 26 :** Acide 2,5-dihydroxybenzoïque

### 3. Analyseur Quadrupole Ion Trap

L'analyseur QIT est un dispositif qui permet, dans un premier temps, d'emprisonner les ions contenus dans une certaine gamme de masse  $m/z$  en provenance d'une source d'ions et, dans un second temps, de les expulser un par un en fonction de leur  $m/z$ . Comme nous le verrons plus loin, il permet également de sélectionner un ion de  $m/z$  donné (dit ion parent) afin de le fragmenter.

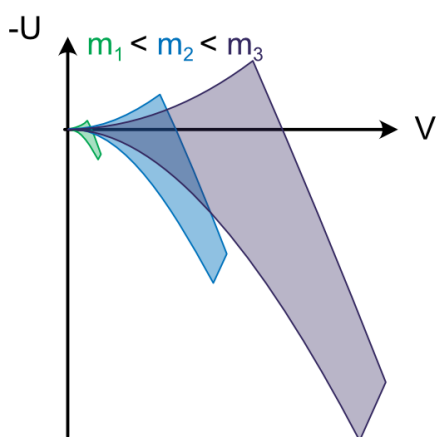
Le QIT se compose de deux électrodes hyperboliques l'une à l'entrée et l'autre à la sortie de la cage appelées électrodes chapeaux (ou parfois calottes) ainsi qu'une électrode circulaire ou annulaire située entre les deux (March, 2009) (**Figure 27**).



**Figure 27** : Fonctionnement d'une trappe à ions (QIT). Les ions en provenance de la source entrent dans la trappe par une électrode chapeau d'entrée puis sont emprisonnés dans la cage grâce au champ électrique créé par les électrodes chapeaux et circulaires. Les ions sont éjectés par l'électrode chapeau de sortie grâce à l'application d'une radiofréquence sur les électrodes chapeaux.  $U$  : le courant continu ;  $V_0$  : amplitude de la radiofréquence ;  $\omega$  : la fréquence ;  $t$  : le temps.

L'arrêt momentané du potentiel électrique de l'électrode chapeau d'entrée permet l'entrée des ions dans la cage tandis que celle de sortie dispose d'un potentiel de charge contraire aux ions. Pour limiter la répulsion entre les différents ions trappés qui ont la même charge, de l'hélium à la pression de  $10^{-3}$  Torr est injecté dans la cage ce qui va permettre de les confiner dans un petit espace au centre de la trappe. Un courant continu  $U$  de même signe que celui des ions que l'on souhaite analyser est appliqué sur les électrodes chapeaux de sorte à maintenir ces ions dans la cage par le jeu de la répulsion entre particules de même signe tandis qu'une radiofréquence  $V = V_0 \cos \omega t$  dont  $V_0$  est l'amplitude,  $\omega$  la fréquence et  $t$  le temps est appliquée sur l'électrode circulaire. Il résulte de ces différents courants, la création d'un champ électrique  $\phi$  ( $\phi = U + V_0 \cos \omega t$ ) hyperbolique à l'intérieur de la cage qui va imprimer aux ions un mouvement en forme de 8 tri-dimensionnel selon les axes  $r$  (rayon à l'intérieur de la cage) et  $z$  (de l'entrée à la sortie) et qui définit leur zone de stabilité (**Figure 28**) (Hoffmann et Stroobant, 2007b). Cela explique la raison pour laquelle seuls les ions compris dans une gamme de masse  $m/z$  peuvent être analysés simultanément. Ce fait va limiter la zone de  $m/z$  que l'on peut analyser, des ions de  $m/z$  très éloignés ne seront pas stables ensemble. Pour analyser l'ensemble des ions formés à partir

d'un échantillon complexe et dont les rapports  $m/z$  peuvent s'étaler de 300 à plusieurs milliers, il est nécessaire de pratiquer plusieurs analyses à différents potentiels (Hoffmann et Stroobant, 2007b).



**Figure 28 :** Zones de stabilité selon les courants U et V dans le Quadrupole Ion Trap (QIT). A  $U=0$ , l'augmentation de V déstabilise les ions dans un ordre suivant leur masse croissante.

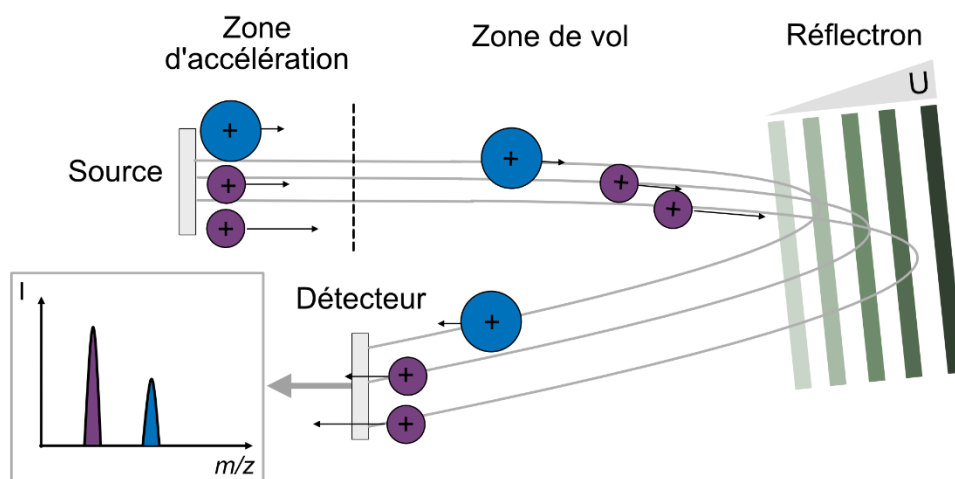
A ce stade, les différents ions oscillent dans leur zone de stabilité avec une fréquence qui leur est propre. Lorsque l'on veut éjecter les ions, une radiofréquence V est appliquée sur les 2 électrodes chapeaux dont la fréquence est égale à la fréquence d'oscillation selon l'axe z d'un ion de masse  $m/z$ . Ainsi, cet ion va absorber l'énergie correspondante ce qui va conduire à une augmentation de son oscillation sur l'axe z qui aura pour effet de l'expulser. Ce mode d'éjection est nommée éjection par résonance.

Le QIT peut également servir de CID pour fragmenter un ion de  $m/z$  donné. Cette fragmentation est réalisée en injectant un gaz comme l'Argon dans la cage à ions de sorte à provoquer des collisions entre ces molécules de gaz et les ions. Consécutivement à une collision, les ions vont voir leur énergie interne augmenter jusqu'au point où ces ions vont se fragmenter. En pratique, dans un mélange d'ions de différents  $m/z$ , un ion  $m/z$  donné sera retenu dans la cage en appliquant aux autres ions la radiofréquence permettant leur éjection. L'ion sélectionné, qu'on appellera « ion parent », va être fragmenté pour conduire à des « ions fils » qui seront ensuite éjectés un par un de la cage pour mesurer leur  $m/z$ . Ce mode d'analyse correspond à la MS/MS ou encore MS<sup>2</sup>. L'intérêt de la cage à ion est que l'on va pouvoir répéter cette opération en sélectionnant un « ion fils » de  $m/z$  donné, le fragmenter à nouveau pour conduire à des ions « petits fils », puis « petits-petits fils ». Ce processus peut être répété

jusqu'à 5 fois, la limite de sensibilité en signant souvent la fin. Les « n » répétitions constituent la spectrométrie de masse MS<sup>n</sup>.

#### 4. Analyseur à temps de vol TOF

Les ions formés sont accélérés par un champ électrique qui va leur conférer une énergie cinétique  $E_c$  ( $E_c = 1/2 \times m \times v^2$  avec  $m$ , la masse de l'ion et  $v$  sa vitesse) avant leur entrée dans le tube de vol. L'analyseur TOF va mesurer le temps que vont mettre ces ions pour parcourir une distance donnée dans ce tube (comprise en général entre 1 à 2 m) (Hoffmann et Stroobant, 2007b). Ce temps de parcours est proportionnel à leur rapport  $m/z$ , selon la formule :  $t = \sqrt{\frac{m}{z}} \frac{L}{\sqrt{2eV}}$  qui dépend de la distance parcourue par les ions ( $L$ ), du potentiel électrique appliqué ( $V$ ), de la charge de l'électron ( $e$ ), de la masse moléculaire de l'ion ( $m$ ) et de son nombre de charge ( $z$ ). Autrement dit, deux ions de même rapport  $m/z$  vont atteindre le détecteur en même temps tandis qu'un ion de rapport  $m/z$  plus important qu'un autre aura un temps de parcours plus important pour atteindre le détecteur. Utilisé seul, ce mode de détection dit linéaire souffre d'un manque de résolution du fait que deux ions ayant le même rapport  $m/z$  mais placés à des endroits différents dans l'épaisseur de la matrice ne vont pas arriver exactement en même temps au détecteur du fait d'énergies cinétiques différentes. Pour corriger cette dispersion dans l'espace, on associe au TOF un réflectron qui est un miroir électrostatique constitué de nombreuses plaques portées à des potentiels croissants depuis l'entrée jusqu'à son extrémité et de même signe que les ions à analyser. Ainsi, les ions pénétreront plus ou moins profondément dans le réflectron avant d'être repoussés en direction opposée. Ainsi, pour deux ions de même rapport  $m/z$  mais d'énergie cinétique différente, celui qui aura l'énergie cinétique la plus élevée pénétrera plus profondément dans le réflectron avant d'être repoussé en direction du détecteur. La différence de distance à parcourir entre deux ions de même  $m/z$  permet de les « réaligner » et de les faire arriver quasiment en même temps au détecteur. Le réflectron permet donc une augmentation significative de la résolution. Cette détection est appelée détection en mode réflectron (**Figure 29**) (Hoffmann et Stroobant, 2007b).



**Figure 29** : Fonctionnement d'un analyseur TOF en mode réflectron. Après une zone d'accélération, les ions parcourent la zone de vol pour être envoyés au détecteur via le réflectron qui permettra de réaligner des ions de même  $m/z$  mais d'énergies cinétiques différentes.

## B. Dérivation des glycanes : perméthylation

Dans le domaine des oligosaccharides, les analyses par spectrométrie de masse MALDI-MS sont rarement réalisées sur les molécules natives. En effet, le caractère très hydrophile de ces macromolécules et de fait, la présence d'autres molécules hydrophiles comme les ions, ne favorise pas le processus d'ionisation/désorption qui a lieu avec la matrice. Pour pallier cet inconvénient, il est souvent nécessaire de les dériver. Il existe 2 grandes stratégies pour cela, soit greffer un composé hydrophobe à l'extrémité *N*-terminale des *N*-glycanes (impossible pour les *O*-glycanes libérés par  $\beta$ -élimination réductive) qui pourra s'ioniser facilement et ainsi améliorer le processus de désorption, soit remplacer les protons mobiles des groupements hydroxyles, amines et acides par des chaînes alkyles afin de diminuer l'hydrophilie. Pour ce dernier point, nous avons mis en œuvre durant ma thèse une des dérivations les plus couramment utilisées qui est la perméthylation. Cette réaction s'effectue en général dans le diméthylsulfoxyde (DMSO) en présence d'hydroxyde de sodium (NaOH) et d'iodure de méthyle ( $\text{ICH}_3$ ) (Yu *et al.*, 2006).

Cette dérivation a 3 avantages majeurs :

- L'augmentation de la désorption des molécules rendues hydrophobes ce qui se traduit par une augmentation du rapport signal sur bruit (S/N, Signal/Noise en anglais) et donc de la sensibilité du MALDI-MS. A noter qu'après la perméthylation, les oligosaccharides perméthylés deviennent solubles dans les solvants organiques comme chloroforme ce qui permet de fait

d'éliminer les sels présents dans l'échantillon initial, contribuant ainsi à faciliter davantage encore la désorption.

- La stabilisation de la liaison sialyle en transformant le groupement carboxylique COOH des acides sialiques en ester méthylique COO-CH<sub>3</sub>.
- Une plus grande rapidité dans l'interprétation des spectres de fragmentation MS<sup>n</sup> car les voies de fragmentation sont très largement documentées.

Ce qui a été décrit ci-dessus reste valable pour une alkylation utilisant de l'iodure de deutérométhyle qui renferme des groupements méthyles deutérés (CD<sub>3</sub>). Dans ce cas on parle de perdeutérométhylation. Cette dérivation a également été utilisée durant ma thèse. Les raisons de son utilisation seront expliquées ultérieurement dans la partie Résultats.

## C. Interprétation des spectres de masse

### 1. Spectres de masse MS

Pour déterminer si un pic de masse  $m/z$  correspond à une structure glycanique ou non sur un spectre de masse MS, on doit au préalable considérer les points suivants :

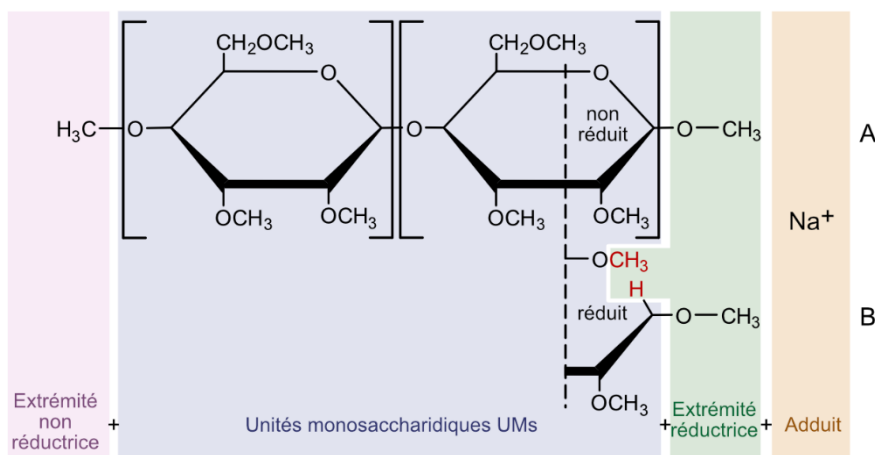
- L'état des oligosaccharides analysés (natif, réduit, perméthylé, perdeutérométhylé).
- La source d'ionisation MALDI crée très majoritairement des ions monochargés dont les signaux observés  $m/z$  sont proches de la masse  $m$  en Da.
- La nature des ions formés ( $[M+H]^+$ ,  $[M+Na]^+$ , ...) pour le calcul de la masse moléculaire  $m$  exacte de l'oligosaccharide. Il n'est pas rare qu'un même oligosaccharide soit détecté sous ses formes  $[M+H]^+$  et  $[M+Na]^+$ , même si cette dernière apparaît très majoritaire en MALDI-MS en présence de DHB comme matrice. Ce phénomène conduit à diminuer la sensibilité du spectromètre de masse. Pour pallier cette perte de sensibilité, on utilise une matrice (ex : DHB) dopée en sodium de sorte à induire exclusivement la formation d'ions adduits sodium  $[M + Na]^+$ .
- Les voies de biosynthèse des *N*- et *O*-glycannes. Ainsi, pour les *N*-glycannes on part de la masse du noyau pentasaccharidique commun à tous les *N*-glycannes et pour les *O*-glycannes, des différents noyaux recensés à ce jour.



On considérera qu'un signal  $m/z$  peut correspondre à un oligosaccharide dès lors qu'on arrive à établir une composition en monosaccharides qui rend compte à l'unité près de la valeur de ce  $m/z$ . Pour cela, on somme la masse moléculaire des différents monosaccharides, natifs ou perméthylés, présentée dans le **Tableau 3**. Il est à noter que ces masses moléculaires correspondent aux monosaccharides lorsqu'ils sont liés (qu'on appellera unités monosaccharidiques UMs) c'est-à-dire ayant perdu, pour l'un des 2 monosaccharides impliqués dans la liaison, son OH anomérique (-17 Da) et pour le monosaccharide auquel il est lié, le proton (-1 Da) d'un de ses OH alcooliques. Le calcul ne s'arrête pas à la simple somme de la masse moléculaire des UMs qui compose un oligosaccharide (**Figure 30**). En effet, il faut ajouter la masse des groupements présents aux extrémités réductrice et non réductrice de l'oligosaccharide. Pour la masse du groupement de l'extrémité réductrice, il faudra ajouter +17 ou +31 ou encore +34 selon qu'il s'agira respectivement d'un groupement hydroxyle -OH (forme native), d'un groupement méthoxy -O-CH<sub>3</sub> (forme méthylée) ou encore d'un groupement méthoxydeutééré -O-CD<sub>3</sub> (forme deutérométhylée). Pour l'extrémité non réductrice, il faudra ajouter +1 ou +15 ou +18 selon qu'il s'agira respectivement d'un hydrogène, d'un groupement méthyle ou d'un groupement deutérométhyle, comme ci-dessus. Enfin, comme les oligosaccharides se cationisent très facilement sous forme d'un adduit [M+Na]<sup>+</sup>, il faudra ajouter 23 Da qui est la masse atomique du sodium. Pour les *O*-glycannes libérés par β-élimination réductrice, il faudra prendre en compte la réduction de l'extrémité réductrice qui a pour conséquence de linéariser le monosaccharide qui occupe cette position et donc de faire apparaître un hydroxyle supplémentaire du fait de l'ouverture du cycle pyranique, sans oublier l'ajout d'un hydrogène sur le carbone anomérique. Cet hydroxyle sera méthylé comme tous les autres soit un supplément de 15 Da (groupement méthyle) et de 1 Da pour l'hydrogène ajouté, soit +16 Da.

**Tableau 3 :** Masse en Dalton des monosaccharides et des extrémités réductrice et non réductrice des oligosaccharides pour le calcul de la masse moléculaire (MW) des oligosaccharides natifs, perméthylés et perdeutérométhylés.

		Natifs	Perméthylés	Perdeutérométhylés
Monosaccharides	Pentose (Pent)	132	160	166
	Deoxyhexose (dHex)	146	174	180
	Hexose (Hex)	162	204	213
	Acide hexuronique (HexA)	176	218	227
	<i>N</i> -Acetylhexosamine (HexNAc)	203	245	254
	Acide 5- <i>N</i> -acetylneuraminique	291	361	376
Extrémités	Non réductrice	1	15	18
	Réductrice libre	17	31	34
	Réductrice réduite	19	47	53



**Figure 30:** Schéma explicatif du calcul des masses moléculaires des (A) *N*-glycannes non réduits et (B) *O*-glycannes réduits perméthylés observés sur un spectre de masse.

La formule ci-dessous résume la méthode de calcul énoncée précédemment. Pour un *N*-glycane libéré par la PNGase F ou A et perméthylé, on pose :

$$\begin{aligned}
 (1)[M + Na]^+ &= 15 + [(a \times 160) + (b \times 174) + (c \times 204) + (d \times 218) + (e \times 245) \\
 &\quad + (f \times 361)] + 31 + 23 \\
 &= \sum MW_{UM} + 69
 \end{aligned}$$

Avec a, b, c, d, e et f le nombre de résidus de Pent, dHex, Hex, HexA, HexNAc et Neu5Ac respectivement, les plus retrouvés dans les glycanes.

Pour un *O*-glycane libéré par  $\beta$ -élimination réductrice et perméthylé, on obtient :

$$\begin{aligned}
 (2)[M + Na]^+ &= 15 + [(a \times 160) + (b \times 174) + (c \times 204) + (d \times 218) + (e \times 245) \\
 &\quad + (f \times 361)] + 16 + 31 + 23 \\
 &= \sum MW_{UM} + 85
 \end{aligned}$$

Avec a, b, c, d, e et f le nombre de résidus de Pent, dHex, Hex, HexA, HexNAc et Neu5Ac respectivement.

Lorsqu'on a affecté à un signal *m/z* une structure oligosaccharide potentielle, il est plus facile de repérer les autres signaux *m/z* pouvant correspondre à des oligosaccharides. En effet, les écarts entre 2 signaux *m/z* proches correspondent le plus souvent à des UMs. Par exemple, deux signaux espacés d'un écart de 204 Da signifient que l'ion le plus lourd possède un résidu d'Hex de plus par rapport au plus léger.

Ce travail d'interprétation d'un spectre de masse est un travail long et fastidieux car il n'est pas rare de retrouver sur un même spectre une cinquantaine d'oligosaccharides différents, voire plus. Le principal défaut de cette méthode est l'obtention de structures hypothétiques qu'il convient de valider ou non par des expériences de spectrométrie de masse MS<sup>2</sup>, voire MS<sup>n</sup>. Enfin, lorsqu'une structure brute en monosaccharides a été établie, il n'y a pas d'information sur la nature de chacun des monosaccharides (par exemple la nature de l'Hex : Gal, Glc, Man...).

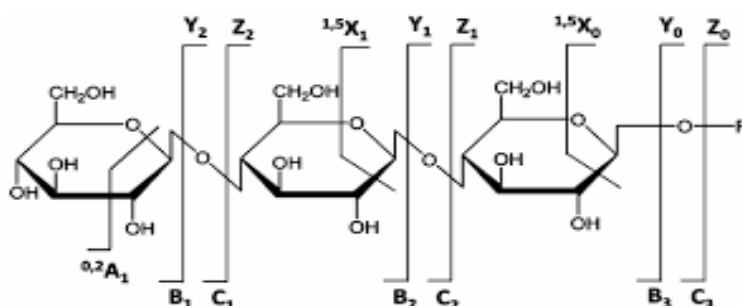
## 2. Spectres de masse MS<sup>n</sup>

Lors d'analyse MS<sup>n</sup> des oligosaccharides, les clivages ont lieu au niveau des liaisons glycosidiques et/ou à l'intérieur des cycles pyraniques (**Figure 31**). Ces clivages conduisent à 6 types de fragments dont les désignations B/Y, C/Z et A/X reposent sur la nomenclature de Domon et Costello (1988).

Pour les fragments observés entre les monosaccharides de la chaîne, ils interviennent :

- soit en amont de l'oxygène pour former 2 types d'ions complémentaires, les ions B pour lesquels la charge est retenue sur le fragment correspondant à l'extrémité non réductrice et les ions Y pour lesquels la charge positive est retenue sur le fragment correspondant à l'extrémité réductrice,
- soit en aval de l'oxygène de la liaison osidique pour former à nouveau 2 ions complémentaires nommés C et Z, les ions C pour lesquels la charge est conservée sur le fragment correspondant à l'extrémité non réductrice et les ions Z pour lesquels la charge est conservée sur le fragment correspondant à l'extrémité réductrice.

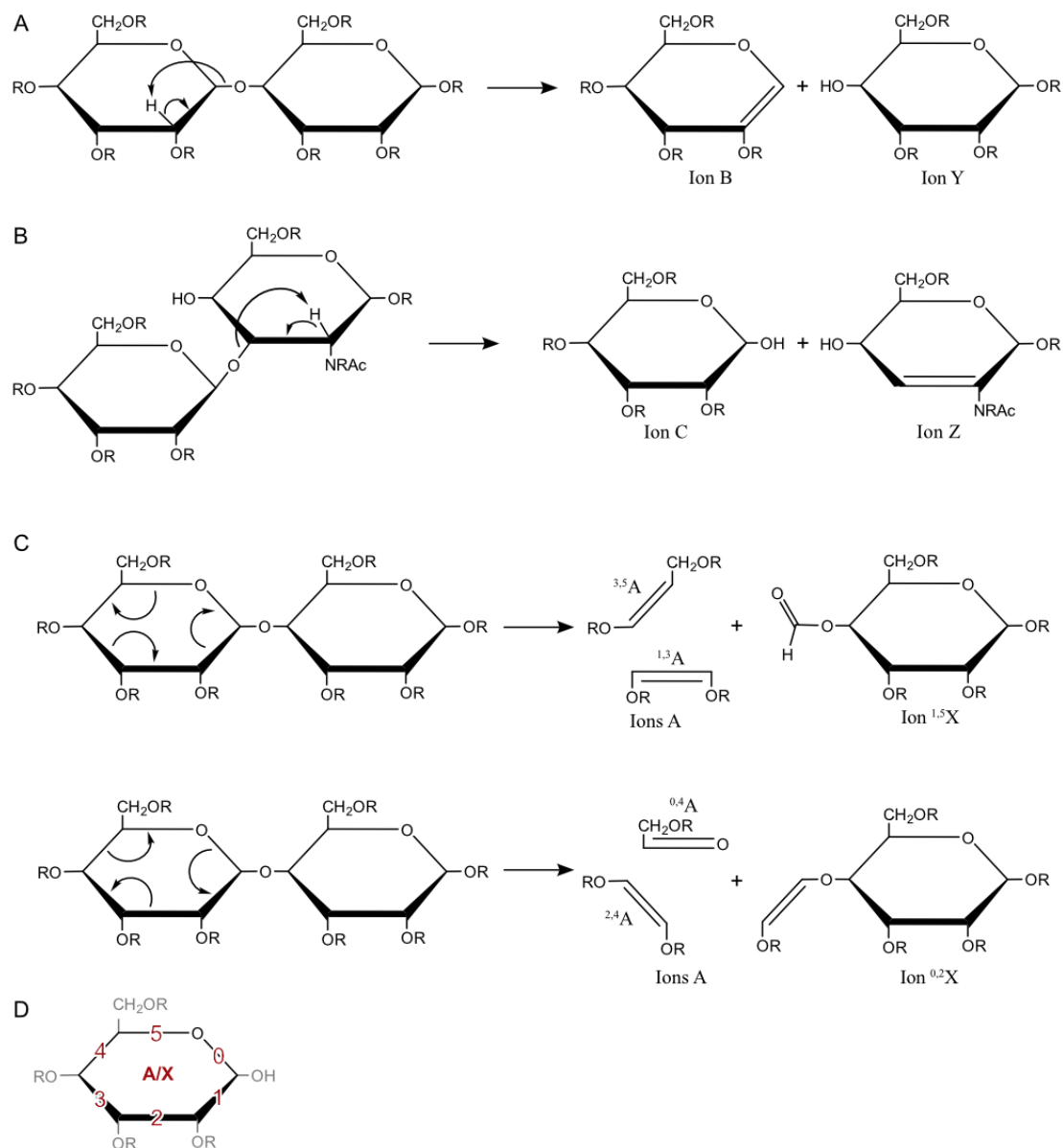
Lorsque la fragmentation intervient à l'intérieur d'un cycle, les ions formés sont nommés A et X.



**Figure 31** : Nomenclature des différents fragments générés par MS/MS selon Domon et Costello (1988). Les fragments résultants de clivage au niveau des liaisons glycosidiques sont nommés Y/B et Z/C selon qu'ils sont créés avant ou après l'oxygène de la liaison glycosidique. Les fragments au sein des cycles sont nommés A et X.

La formation de ces fragments dépend de la nature des oligosaccharides (natifs ou dérivés) et du mode de fragmentation choisi pour fragmenter les oligosaccharides. Dans le cas de l'utilisation du MALDI-QIT-TOF que nous avons utilisé tout au long de ma thèse, il s'agit d'un mécanisme de

fragmentation CID basse énergie qui conduit très majoritairement à des ruptures localisées au niveau des liaisons osidiques pour générer des ions B, Y, C et Z et très rarement à des fragments intra-cycles A et X pour des glycanes per(deutéro)méthylés. Les mécanismes de fragmentation qui conduisent à ces différents ions sont ceux décrits par A. Dell (1987) à propos de la spectrométrie de masse FAB-MS (Figure 32).



**Figure 32 :** Mécanismes de fragmentation des oligosaccharides perméthylés conduisant aux fragments (A) B et Y, (B) C et Z et (C) A et X par spectrométrie de masse MS<sup>n</sup>. (D) Numérotation des liaisons d'un cycle pyranique pour caractériser les fragments A et X. Selon la nomenclature de Domon et Costello (1985).

Les formules permettant de calculer la masse moléculaire de ces ions fragments issus de glycanes perméthylés sont présentées ci-dessous :

- Pour un ion B :  $[M + Na]^+ = \sum MW_{UM} + 15 - 1 + 23 = \sum MW_{UM} + 37$

Avec +15 pour le groupement méthyle (CH<sub>3</sub>) qui coiffe le monosaccharide de l'extrémité non réductrice du fragment, -1 correspondant à la perte de l'hydrogène du C2 et +23 pour la masse de l'ion sodium Na<sup>+</sup>.

- Pour un ion Y :  $[M + Na]^+ = \sum MW_{UM} + 31 + 1 + 23 = \sum MW_{UM} + 55$

Avec +31 pour la masse du groupement méthoxy (OCH<sub>3</sub>) qui coiffe le monosaccharide de l'extrémité réductrice du fragment, +1 correspondant à l'hydrogène qui se fixe sur l'oxygène de la liaison glycosidique rompue et +23 pour la masse de l'ion sodium Na<sup>+</sup>.

- Pour un ion C :  $[M + Na]^+ = \sum MW_{UM} + 15 + 17 + 23 = \sum MW_{UM} + 55$

Avec +15 pour le groupement méthyle (CH<sub>3</sub>) qui coiffe le monosaccharide de l'extrémité non réductrice du fragment, +17 correspondant à la masse de la fonction hydroxyle (OH) récupérée sur le C1 du monosaccharide en position terminale réductrice post-clivage et +23 pour la masse de l'ion sodium Na<sup>+</sup>.

- Pour un ion Z :  $[M + Na]^+ = \sum MW_{UM} + 31 - 16 - 1 + 23 = \sum MW_{UM} + 37$

Avec +31 pour la masse du groupement méthoxy (OCH<sub>3</sub>) qui coiffe le monosaccharide de l'extrémité réductrice du fragment, -16 pour la perte de l'oxygène du monosaccharide en position terminale non réductrice post-clivage, -1 pour la perte de l'hydrogène du C2 et +23 pour la masse de l'ion sodium Na<sup>+</sup>.

Il conviendra d'ajouter +16 Da pour les ions Y et Z issus de *O*-glycanes réduits (comme vu précédemment). On peut noter que les formules des ions Y et C ainsi que des ions B et Z sont les mêmes ce qui pourrait conduire à des ambiguïtés quant à la nature des ions formés. Dans ce cas, des analyses MS<sup>3</sup> supplémentaires permettront de les lever.

Les analyses MS<sup>2</sup> réalisées sur le MALDI-QIT-TOF avec des oligosaccharides perméthylés conduisent à 3 clivages préférentiels qui ont lieu :

- du côté réducteur des résidus d'HexNAc selon une fragmentation B/Y,

- au niveau de la liaison sialyle qui est une liaison fragile ce qui conduit à la libération des acides sialiques,
- au niveau de la liaison entre un monosaccharide et le C3 d'un HexNAc selon un clivage C/Z (Yu *et al.*, 2006).

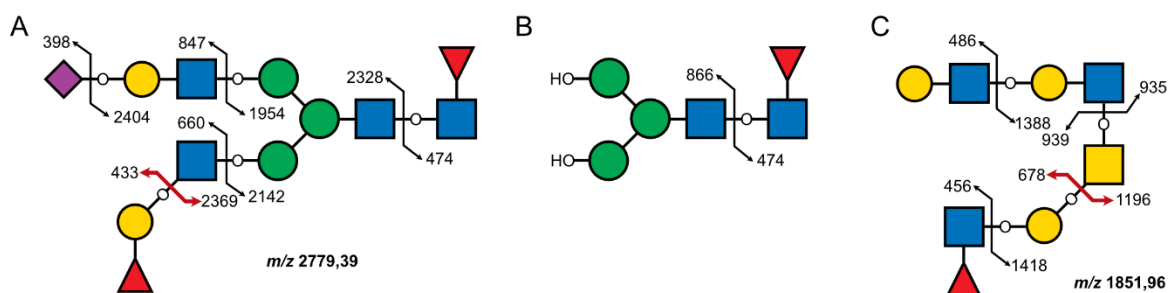
Pour les *N*-glycannes perméthylés, cela se traduit par un clivage entre les deux résidus de GlcNAc du noyau (B/Y) ce qui permettra d'identifier la nature des substituants que peut porter le résidu de GlcNAc en position terminale réductrice comme par exemple un ou 2 résidu(s) de Fuc (**Figure 33A**). Dans le cas de l'utilisation du MALDI-QIT-TOF qui ne détecte que les ions dont le rapport  $m/z$  est compris dans la gamme de masse sélectionnée, cet ion Y généré n'est pas toujours observé sur le spectre du fait d'un rapport  $m/z$  trop faible. Dans ce cas, la différence de masse entre l'ion parent et l'ion B complémentaire permet de l'identifier. Toutefois, cet ion B est souvent peu abondant car il est le siège d'autres clivages.

Un autre clivage important est celui qui a lieu entre les 2 résidus de Man du noyau et les antennes (**Figure 33B**). D'une part, il va permettre d'identifier le nombre d'antennes que porte le *N*-glycanne grâce au nombre de groupements hydroxyles présents sur le noyau. D'autre part, les fragments correspondants aux antennes permettent de repérer et identifier les épitopes glycaniques lorsqu'ils sont présents. Cependant, il n'est pas possible de connaître sur quel résidu de Man est portée une antenne particulière. Enfin, il est possible de faire la distinction entre une unité de type 1 (Gal $\beta$ 1-3GlcNAc) et de type 2 (Gal $\beta$ 1-4GlcNAc). En effet, le clivage de la liaison osidique entre le Gal et le GlcNAc d'une unité de type 1 conduit majoritairement à des ions C/Z alors que celui entre le résidu de Gal et de GlcNAc d'une unité de type 2 conduit à des ions B/Y (Terada *et al.*, 2005 ; Yu *et al.*, 2006).

Pour les *O*-glycannes, le processus d'élimination des substituants sur le carbone 3 des résidus d'HexNAc permet d'identifier le noyau ainsi que la nature des unités de type 1 et 2 (**Figure 33C**) (Guérardel *et al.*, 2006).

La fragmentation d'oligosaccharides perméthylés par CID basse énergie qui aboutit essentiellement à des ions B/Y ne permet pas toujours l'identification précise des épitopes portés par les

antennes. Néanmoins, ce problème peut être corrigé en effectuant des analyses de fragmentation additionnelles de type MS<sup>n</sup> (comme le permet le MALDI-QIT-TOF utilisé durant ces travaux) à partir d'ions fragments et ainsi obtenir des informations supplémentaires. Par exemple, pour un ion B FucGalGlcNAc identifié par une analyse MS<sup>2</sup> correspondant à une antenne substituée par un résidu de Fuc, on ne connaît pas la position du Fuc. Il peut être situé sur le résidu de Gal et définir un Ag H ou être situé sur le résidu de GlcNAc et définir un Ag Le<sup>a/x</sup>. Une analyse par MALDI-MS<sup>3</sup> permettra de cliver la liaison entre les résidus de Gal et GlcNAc et selon les ions formés, de déterminer la localisation du résidu de Fuc.



**Figure 33** : Localisation des fragmentations majoritaires observées pour les (A-B) *N*- et (C) *O*-glycannes perméthylés pour des analyses MS<sup>2</sup> par MALDI-QIT-TOF. Les doubles flèches noires indiquent des clivages B/Y et celles rouges des clivages C/Z. (B) Lorsqu'un substituant (*e.g.* une antenne LacNAc) est éliminé au niveau d'un des 2 Man du noyau, il apparaît un -OH sur le Man. Ainsi, le nombre de OH sur le fragment indique le nombre de substituants (*e.g.* antennes) sur le noyau qu'il faudra prendre en compte pour le calcul de sa masse. Les oligosaccharides sont représentés selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019).

## IV. Analyse des *N*- et *O*-glycannes par GC-EI-MS

### A. Fonctionnement du GC-EI-MS

#### 1. Présentation

L'appareil GC-EI-MS se compose d'une chromatographie en phase gazeuse, d'une source d'ionisation électronique (EI) et d'un analyseur quadrupôle. Les analytes de l'échantillon sont séparés par chromatographie avant d'être ionisés et fragmentés par la source EI. Les ions obtenus sont ensuite séparés par le quadrupôle avant d'arriver au détecteur.

#### 2. Chromatographie en phase gazeuse

La chromatographie en phase gazeuse (CPG ou GC en anglais) permet la séparation de molécules gazeuses en fonction de leur température de volatilisation et leur polarité qui leur permettra

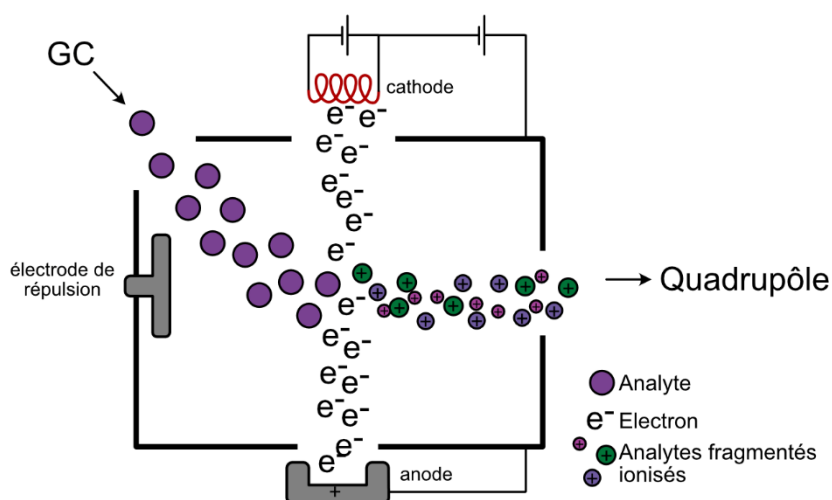


d'interagir plus ou moins fortement avec la phase stationnaire de la colonne. Cette dernière est placée dans un four dans lequel règne une température qu'il sera possible de faire varier très rapidement afin d'optimiser la séparation des molécules. L'échantillon, sous forme liquide, est introduit dans la colonne sous très haute température afin d'assurer sa volatilité immédiate. L'échantillon sous forme vapeur est ensuite entraîné grâce à un gaz vecteur (hélium ou azote) dans la colonne où les molécules subiront une rétention plus ou moins importante avec la phase stationnaire et qu'il sera possible de moduler en augmentant la température. A la sortie de la colonne, les molécules éluées sont détectées, le plus couramment, grâce à un détecteur à ionisation de flamme. Le principe de ce dernier repose sur la pyrolyse des molécules gazeuses dans une flamme à hydrogène qui a pour effet de produire des ions et électrons qui produiront un courant électrique qui sera amplifié, digitalisé et enregistré en vue de retraitements ultérieurs. Ce détecteur, très sensible, ne donne comme seule information que le temps de rétention des molécules qui peut permettre leur identification par rapport à des témoins. Son principal défaut est qu'il ne donne aucune information structurale à l'inverse d'une détection par un spectromètre de masse de type EI-MS. Dans le cadre de nos travaux, l'EI-MS a été utilisé comme détecteur dans le but d'obtenir des caractéristiques structurales des monosaccharides présents dans nos échantillons.

### 3. Source d'ionisation électronique EI

La source EI est constituée d'une enceinte dans laquelle règne une température élevée et qui est traversée par un faisceau d'électrons produit par un filament de Tungstène ou de Rhénium (cathode) parcouru par un courant électrique (**Figure 34**) (Hoffmann et Stroobant, 2007a). Après leur émission, les électrons sont accélérés grâce à une anode placée à l'opposé du filament de tungstène et portée à un potentiel de +70 eV. Les analytes sous forme vapeur qui sortent de la colonne de CPG et arrivent dans cette enceinte seront percutés par ces électrons. Sous ce potentiel élevé, ils seront non seulement ionisés mais également fragmentés. Cette fragmentation est telle qu'elle ne permet que très rarement d'observer l'ion moléculaire. En mode positif, les cations radicalaires créés sont envoyés vers l'analyseur grâce à une électrode de répulsion portée à un potentiel positif et placée à l'opposé de l'entrée de l'analyseur d'ions du spectromètre de masse. Pour augmenter le rendement d'ionisation, le trajet des électrons va

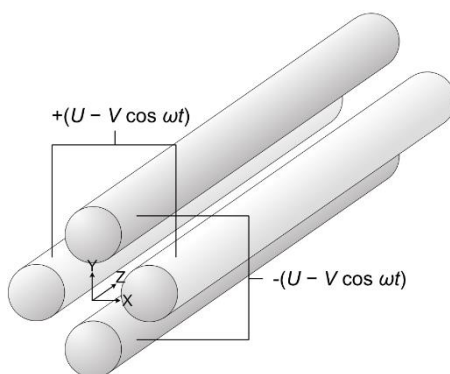
être augmenté par des aimants à l'extérieur de la source en leur induisant une trajectoire hélicoïdale (Hoffmann et Stroobant, 2007a).



**Figure 34** : Schéma d'une source d'ionisation électronique. Une cathode constituée d'un filament produit des électrons qui vont rencontrer les analytes de l'échantillon. Les analytes reçoivent une quantité d'énergie qui conduit à la perte d'électrons et à leur fragmentation. Les cations créés sont ensuite envoyés à l'analyseur grâce à une électrode de répulsion positive.

#### 4. Analyseur quadripôle

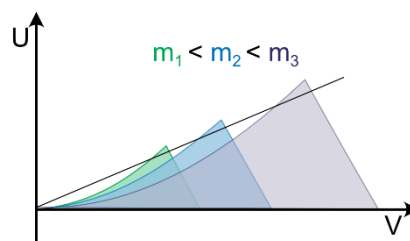
La source EI-MS que nous avons utilisée durant ma thèse était couplée à un quadripôle. Ce dernier est composé de 4 barres métalliques cylindriques parallèles d'une vingtaine de cm, opposées 2 à 2 perpendiculairement (**Figure 35**). Ces 4 barres sont parcourues par 2 courants électriques, l'un continu (U) et l'autre alternatif (V), et dont le rapport U/V reste constant tout au long de l'analyse. Deux barres, diamétralement opposées sont parcourues par un champ électrique positif et les 2 autres, par un champ électrique négatif.



**Figure 35** : Schéma d'un quadripôle. Le quadripôle se compose de 4 électrodes cylindriques traversées 2 à 2 par un potentiel de charge opposée.

Ces 4 barres définissent un cercle de rayon  $r$ . A l'entrée du quadropôle, les ions qui se propagent selon une trajectoire  $z$ , vont subir un mouvement d'oscillation dans les axes  $x$  et  $y$  du fait de l'application sur les barres d'un courant continu et alternatif (radiofréquence) (Hoffmann et Stroobant, 2007b). Dans le cas où l'amplitude de l'oscillation adoptée en  $x$  et  $y$  ne dépasse pas le rayon  $r$  du cercle défini par les 4 barres, l'ion parcourra toute la longueur des 4 barres sans les heurter et en sortira pour aller percuter le détecteur. Cette oscillation stable dépend de la tension du courant continu et de la radiofréquence appliquée à l'instant  $t$  et surtout du rapport  $m/z$  de l'ion considéré. A cette valeur de champ électrique, les autres ions qui ne disposent pas du bon rapport  $m/z$  verront l'amplitude de leurs oscillations en  $x$  et  $y$  dépasser le rayon  $r$  et percuteront l'une des 4 barres du quadropôle. Opérant un balayage continu du champ électrique avec un rapport  $U/V$  constant, chaque ion adoptera à un moment  $t$  en fonction de son rapport  $m/z$  un mouvement oscillant stable lui permettant d'être détecté (Hoffmann et Stroobant, 2007b).

Les zones de stabilité et d'instabilité d'un ion sont définies par les équations de Mathieu et sont représentées dans la **Figure 36**. Un ion sera successivement attiré puis repoussé par chaque paire d'électrodes tout en se déplaçant selon l'axe  $z$ . Un ion arrive au détecteur seulement s'il reste stable (l'amplitude de son oscillation) selon  $x$  et  $y$ . Les potentiels varient et augmentent au cours du temps en conservant un rapport  $U/V$  constant ce qui permet la stabilité d'un ion à la fois et ainsi la détection successive des différentes masses comme le montre les zones de stabilité des différentes masses selon  $U$  et  $V$  en **Figure 36**. Plus la pente est élevée, plus la résolution sera élevée mais moins la sensibilité sera grande. En générale, la pente employée est celle nécessaire pour obtenir une résolution unitaire qui permet la séparation de deux pics à une unité de masse d'écart.



**Figure 36** : Zones de stabilité selon les courants  $U$  et  $V$  dans le quadropôle. Pour un rapport  $U/V$  constant dont les valeurs de  $U$  et  $V$  augmentent, un ion à la fois est stable dans l'analyseur.

## B. Dérivation des monosaccharides

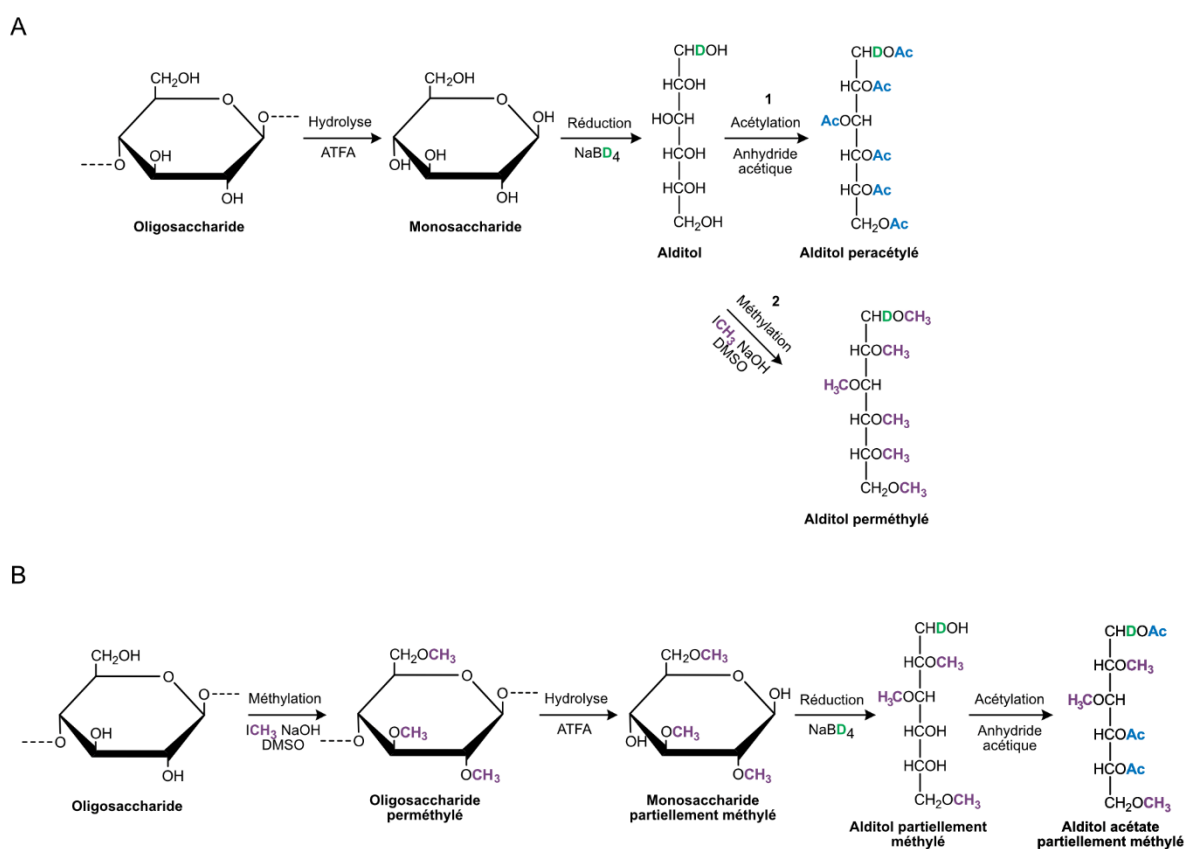
Dans le cadre d'une étude structurale précise des oligosaccharides, la CPG est utilisée pour séparer, identifier et quantifier les monosaccharides qui les composent. En effet, tandis que l'analyse de standards permet d'identifier les monosaccharides présents par la comparaison des temps de rétention, l'utilisation d'un standard interne dont la quantité ajoutée est connue (comme le mésoinositol) permet de les quantifier. En sortie de colonne, le couplage de la CPG avec un spectromètre de masse EI-MS permet de confirmer l'identification des monosaccharides mais surtout d'obtenir des informations structurales supplémentaires grâce aux fragments générés. En effet, les fragments de monosaccharides sur les spectres MS peuvent permettre d'identifier la structure d'un monosaccharide inconnu et selon les traitements chimiques employés permettre de déterminer les liaisons qui liaient les monosaccharides.

La CPG nécessite que les composés à analyser soient volatils ce qui n'est pas le cas intrinsèquement des monosaccharides qui devront être dérivés chimiquement pour le devenir. Différents traitements peuvent être utilisés et sont choisis en fonction des informations que l'on souhaite obtenir. Dans un souci de concision, nous ne détaillerons que les traitements relatifs à nos travaux.

Le 1er traitement que nous avons exécuté a consisté à transformer les monosaccharides constitutifs des *N*-glycannes en itols peracétylés dans le but de les identifier et les quantifier. Pour ce faire, après leur libération par hydrolyse acide (acide trifluoroacétique 4M à 100°C pendant 4h), ces derniers sont réduits par le borohydrure de sodium ou borodeutérure de sodium suivi d'une peracétylation qui conduit à la formation d'« alditols peracétylés » directement analysables par CPG (**Figure 37A1**). L'utilisation des dérivés alditols acétates présente plusieurs avantages. Cette dérivation produit des dérivés très stables et permet la séparation de monosaccharides dans un mélange complexe. De plus, le traitement réducteur permet d'éliminer la possibilité d'une cyclisation en pyranne ou furanne des monosaccharides avec pour chacune d'elles une anomérie  $\alpha$  ou  $\beta$  qui entraîne une démultiplication du nombre de pics lors des analyses CPG. Ainsi sous la forme alditol, chaque monosaccharide apparaît sous la forme d'un pic unique. En pratique, le borohydrure de sodium est souvent remplacé par du borodeutérure de sodium afin de distinguer le C1 du C6 des dérivés. De plus amples informations sur ce traitement seront donnés dans la partie suivante (Haslam *et al.*, 2022 ; Ruiz-Matute *et al.*, 2011).

Un autre traitement, quoique beaucoup moins courant, consiste à réaliser une perméthylation des alditols à la place de l'acétylation. Cela permet l'obtention d'alditols perméthylés (**Figure 37A2**) (Black *et al.*, 2019 ; Walker *et al.*, 1962).

Un autre traitement très souvent utilisé en chimie des glucides consiste à transformer les monosaccharides d'un oligosaccharide en itols acétates partiellement méthylés. Ce traitement permet de déterminer le ou les carbone(s) d'un monosaccharide engagé(s) dans une liaison glycosidique (**Figure 37B**). Brièvement, après perméthylation d'un oligosaccharide, les monosaccharides méthylés sont libérés comme précédemment. Les monosaccharides partiellement méthylés sont ensuite réduits pour former les itols correspondants. Enfin, les hydroxyles libres issus de la rupture des liaisons glycosidiques dans lesquelles ils étaient engagés sont acétylés pour générer des « alditols acétates partiellement méthylés » (PMAA, partially methylated alditol acetates). L'analyse de ces dérivés par CPG couplée à la spectrométrie de masse EI-MS permet de localiser très précisément le ou les groupement(s) méthoxy au sein de chacun des dérivés générés et ainsi de connaître les points de branchement des



**Figure 37 :** Traitement chimique des oligosaccharides pour l'obtention (A) (1) d'alditols peracétylés, (2) d'alditols perméthylés et (B) d'alditols acétates partiellement méthylés.

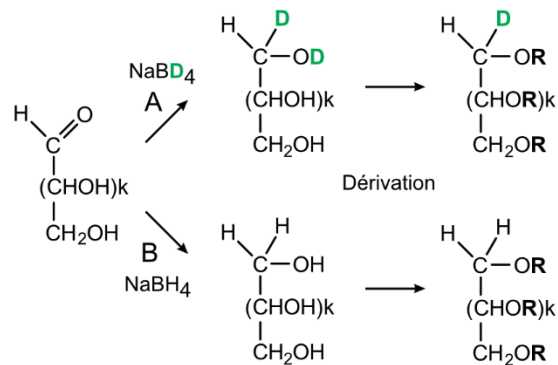
monosaccharides au sein de l'oligosaccharide (Fournet *et al.*, 1981 ; Haslam *et al.*, 2022). Cette dernière dérivation n'a pas été utilisée dans les travaux qui seront présentés. Néanmoins, la compréhension des mécanismes de fragmentation des dérivés PMAAs en EI-MS est nécessaire pour comprendre comment la structure des monosaccharides porteur d'un groupement méthyle chez l'huître ont été établies.

### C. Interprétation des résultats

L'identification d'un monosaccharide en GC-EI-MS est basée sur son temps de rétention en CPG ainsi que sur son spectre de fragmentation EI-MS. Au-delà de l'identification des monosaccharides, la spectrométrie de masse EI-MS des dérivés PMAAs est une méthode extrêmement puissante pour déterminer les points de liaison entre les monosaccharides d'un oligosaccharide en déterminant la position du/des groupement(s) méthyle(s) au sein des monosaccharides. Les mécanismes de fragmentation des différents dérivés ont beaucoup été étudiés ce qui facilite l'interprétation des spectres EI-MS (Lönngrén et Svensson, 1974).

Un alditol issu de la réduction d'un monosaccharide neutre possède à ses 2 extrémités une fonction alcool primaire ( $\text{CH}_2\text{OH}$ ), l'une correspondant à la réduction de la fonction aldéhyde du carbone 1 et l'autre à la fonction alcool primaire présente sur le carbone 6. Ainsi, il est impossible de distinguer les carbones 1 et 6 et donc les carbones 2 et 5, et 3 et 4 du fait de la symétrie de la molécule. Pour pallier cet inconvénient, l'utilisation du borodeutériure de sodium (ou de potassium) comme agent réducteur à la place du borohydrure de sodium (ou de potassium) permet de générer un alcool primaire deutéré ( $\text{CD}_2\text{OH}$ ) au niveau du C1 en introduisant 2 atomes de deutérium, l'un sur le carbone du carbonyle et l'autre sur l'oxygène du carbonyle. Le deutérium de ce dernier est éliminé lors de la réaction d'estérification ou encore de perméthylation. A l'arrivée il ne reste que le deutérium du C1 ce qui a pour effet de l'alourdir d'1 Da par rapport au C6. Il devient dès lors possible de numéroter chacun des carbones de la molécule à partir des ions fragments (**Figure 38**).

Lorsqu'un itol peracétylé, perméthylé ou PMAA rencontre des électrons avec une énergie cinétique de 70 eV, un électron périphérique d'un oxygène ou d'un azote va être arraché pour conduire à un cation. La stabilisation de ce cation s'effectue en récupérant un électron d'une liaison et conduit



**Figure 38** : Réduction des monosaccharides en leur alditol correspondant par (A) le borodeutérure de sodium ( $\text{NaBD}_4$ ) ou (B) le borohydrure de sodium ( $\text{NaBH}_4$ ). La présence d'un atome de deutérium sur le carbone 1 d'un monosaccharide permet de le distinguer du carbone n (issu de la formule brute  $\text{C}_n(\text{H}_2\text{O})_n$ ).

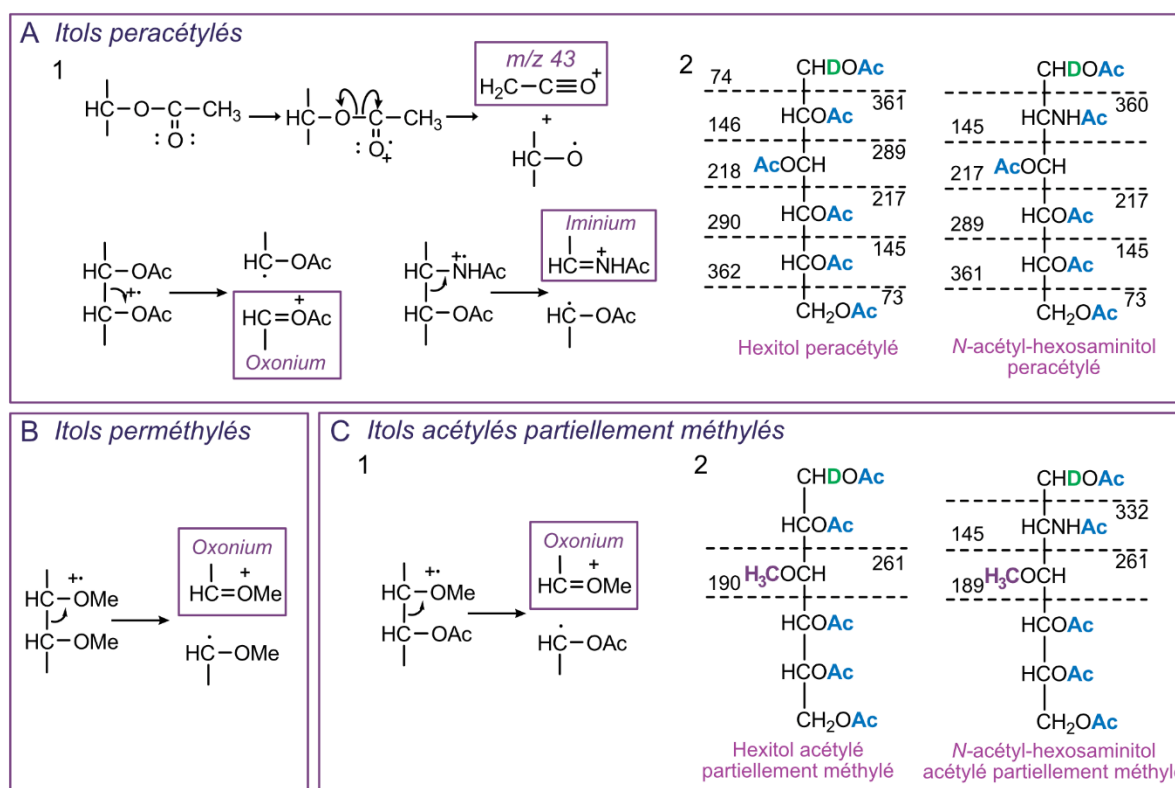
ainsi à 2 fragments dont l'un portera une charge positive et l'autre sera radicalaire. Si le potentiel appliqué au répulseur d'ion de la source est positif, seuls les ions positifs seront envoyés vers l'analyseur.

Dans le cas de l'analyse par EI-MS des dérivés alditols peracétylés, l'ion majeur détecté correspond à l'élimination du groupement acétyle  $m/z$  43 selon le mécanisme présenté en **Figure 39A1**. Au-delà de cette réaction de stabilisation, une autre voie conduit à la rupture de la liaison C-C de la chaîne glucidique à proximité de l'oxonium formé (**Figure 39A1**). Ainsi, sur un hexose ou une hexosamine peracétylé, l'arrachement d'un électron peut avoir lieu entre n'importe quel groupement acétyle de la molécule ce qui signifie que différents fragments peuvent être générés comme présenté dans la **Figure 39A2**. A noter que le clivage survient plus abondamment en faveur de la formation d'ions iminiums que d'ions oxoniums car la charge est davantage stabilisée sur une fonction iminium.

Dans le cas de l'analyse par EI-MS des dérivés alditols perméthylés, les clivages au sein de la chaîne glucidique auront lieu entre les groupements méthoxy pour conduire à la formation d'ions oxoniums (et iminiums pour un HexNAc) (**Figure 39B**).

Dans le cas de l'analyse par EI-MS des dérivés PMAAs, le clivage se fait entre une fonction acétyle et méthoxy et le cation formé se stabilisera uniquement sur le groupe méthoxy (**Figure 39C1**). Ainsi, lorsqu'un seul groupement méthyle est présent dans un dérivé PMAA (Hexose ou hexosamine), 2 fragments majeurs avec l'ion oxonium du méthoxy sont observés selon la liaison rompue, au-dessus ou en-dessous de l'oxonium méthylé (**Figure 39A2**). Pour ces dérivés, on observe également la

formation des ions iminiums issus des HexNAc. En conclusion, grâce à l'EI-MS, il est possible de connaître quel carbone d'un PMAA est porteur d'un groupement méthyle.



**Figure 39** : Mécanismes de fragmentation des (A) itols acétates, (B) itols perméthylés et (C) itols acétates partiellement méthylés en EI-MS. (A)(1) Mécanismes conduisant à la formation de l'ion  $m/z$  43, et des ions oxonium et iminium et (2) fragmentations primaires des itols acétates. (B) Mécanismes conduisant à la formation de l'ion oxonium des itols perméthylés. (C)(1) Mécanismes conduisant à la formation de l'ion oxonium et (2) fragmentations primaires des itols acétates partiellement méthylés.

## V. Analyse des N-glycannes par nanoLC-ESI-MS/MS

Certains résultats présentés dans cette thèse ont été obtenus en utilisant un spectromètre de masse ESI-MS/MS couplé à un dispositif de séparation de type nanoLC : l'Orbitrap Fusion™ Tribrid™ développé par la société Thermo Fischer Scientific. Nous ne disposons pas au laboratoire de ce dispositif et avons établi une collaboration avec l'équipe du Pr. Kay-Hooi Khoo de l'Institute of Biochemical Sciences à l'université nationale de Taiwan.



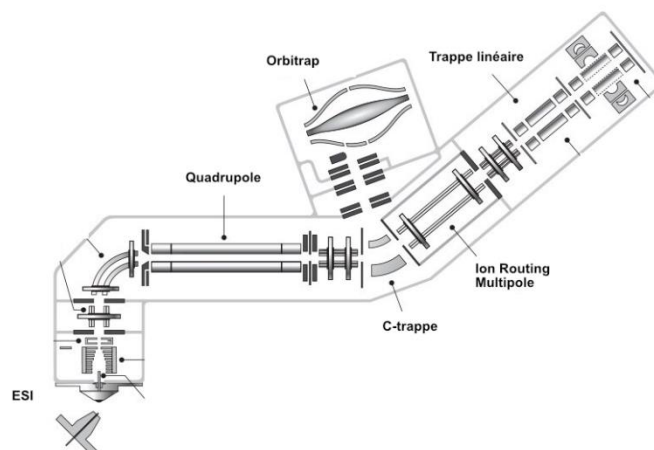
## A. Présentation du spectromètre de masse nanoLC Orbitrap Fusion™ Tribrid™

### 1. Description

Le spectromètre de masse Thermo Orbitrap Fusion de chez Thermo Scientific™ (**Figure 40**) comprend :

- Une source ESI (ElectroSpray Ionization) couplée à une chromatographie liquide nano haute performance (nanoLC),
- un quadrupôle qui permet la sélection des ions précurseurs et permet aux 2 analyseurs de masse (la trappe à ions et l'orbitrap) de fonctionner en parallèle,
- un orbitrap qui est un analyseur de masse très haute résolution (supérieure à 450 000) qui permet l'analyse d'ions isobares,
- une trappe à ion linéaire comprenant 2 étages de pression différentes. Le 1<sup>er</sup> étage, haute pression, permet d'effectuer des fragmentations CID et ETD (Electron Transfer Dissociation). Le second étage, basse pression, permet de mesurer le rapport  $m/z$  des ions fragments.

Le dispositif Ion Routine Multiple (IRM) placé entre l'orbitrap et la trappe à ions, facilite les analyses parallèles grâce à une accumulation des ions et leur routage vers l'un des 2 analyseurs de masse, l'orbitrap et la trappe à ion linéaire double étage. Il permet également de réaliser des fragmentations HCD (High Energy Collision Dissociation) à n'importe quelle étape de fragmentation ( $MS^2$ ,  $MS^3$ ,  $MS^n$ ).



**Figure 40 :** Composition d'un spectromètre de masse Orbitrap Fusion Tribrid (<https://www.thermofisher.com>).

## 2. Source d'ionisation ESI

Cette source est équipée d'un capillaire placé dans une enceinte où règne une température élevée (environ 250 °C) et dans lequel se trouve l'échantillon. A l'extrémité de ce capillaire est appliqué un courant électrique élevé de quelques milliers de volts, soit positif, soit négatif selon que l'on souhaite analyser les molécules en mode positif ou négatif. Lorsque l'échantillon arrive à l'extrémité du capillaire, il y a création d'un spray (d'où le nom de cette source) composé de fines gouttelettes dans lesquelles se trouvent les analytes chargés. Sous l'effet conjugué d'un courant de séchage par un gaz inerte et de la température élevée, le solvant s'évapore totalement et les molécules chargées vont se retrouver à l'état gazeux. C'est sous cette forme qu'elles quittent la source pour entrer dans l'analyseur (Hoffmann et Stroobant, 2007a). La source ESI du spectromètre de masse peut être couplée à une chaîne de chromatographie en phase liquide à haute performance (HPLC) équipée d'une colonne nanobore qui permet de séparer les molécules d'un mélange complexe avant leur introduction dans la source du spectromètre de masse.

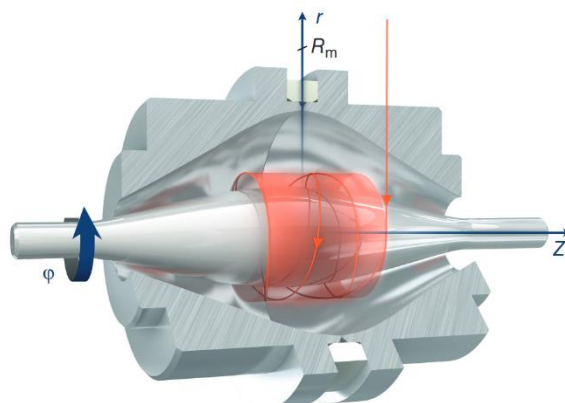
Pour une analyse en mode MS où le but est de mesurer le rapport  $m/z$  des ions moléculaires, le quadripôle, placé immédiatement en sortie de source, est utilisé comme un guide pour acheminer les ions dans une C-trappe qui va permettre de concentrer les ions, les ralentir et les envoyer dans l'orbitrap avec une dispersion minimale.

## 3. Analyseur orbitrap

L'orbitrap est composé d'une électrode centrale en forme de fuseau placée dans 2 demi-électrodes creuses externes (**Figure 41**).

Après injection des ions dans l'orbitrap en quelques 50  $\mu$ sec, un potentiel négatif est appliqué sur l'électrode centrale dans le cas d'une analyse en mode positif. A leur arrivée dans l'orbitrap, les ions vont adopter 2 types d'oscillation, l'une dite radiale autour du fuseau central et l'autre axiale selon l'axe z. Le passage successif des ions d'une demi-électrode à l'autre selon l'axe axial va créer un courant induit qui sera capté, amplifié et enregistré. La fréquence du courant induit généré par un ion est directement corrélé à son rapport  $m/z$ . Dans le cas où de multiples ions différents sont analysés en même temps, un traitement du signal global par transformée de Fourier permet d'obtenir le spectre de fréquence

de chacun des ions et ainsi leur rapport  $m/z$ . Cet analyseur brille par sa très grande résolution (env. 500 000) (Hoffmann et Stroobant, 2007b).



**Figure 41** : Composition d'un orbitrap. Cet analyseur se compose d'une électrode en forme de fuseau placée dans une électrode creuse. Les ions vont osciller autour d'électrode en fuseau selon les axes  $r$  et  $z$ . Issue de Hecht *et al.* (2019).

Concernant le principe de fonctionnement du quadrupôle et de la trappe à ion qui équipent l'Orbitrap Fusion™ Tribrid™, ils ont été vus dans les parties précédentes (III.A.3 p79 et IV.A.4 p94).

#### 4. Fonctionnement général du spectromètre de masse nanoLC Orbitrap Fusion™ Tribrid™

l'Orbitrap Fusion™ Tribrid™ permet d'effectuer des analyses MS, MS<sup>2</sup> et MS<sup>3</sup> sur échantillon entier avec la particularité de pouvoir déclencher automatiquement une analyse MS<sup>3</sup> à partir d'ions fragments issus de la MS<sup>2</sup> préalablement sélectionnés. Pour ce faire, une analyse MS permet de déterminer le rapport  $m/z$  de toutes les molécules présentes dans l'échantillon grâce à l'orbitrap. Dans ce cas, le quadrupôle qui est en amont de l'orbitrap (**Figure 40**) est utilisé comme guide d'ions. Pour les analyses MS<sup>2</sup>, le quadrupôle va permettre de sélectionner la molécule que l'on souhaite fragmenter. Cette dernière arrive au niveau du dispositif IRM pour y subir une fragmentation à haute énergie HCD. Les ions fils sont ensuite transmis à l'orbitrap où ils sont analysés pour, *in fine*, établir le spectre de masse MS<sup>2</sup> correspondant. Les ions fils présentant un intérêt particulier déclencheront une analyse MS<sup>3</sup> dès lors qu'ils seront détectés en MS<sup>2</sup>. A cette fin, après fragmentation dans l'IRM, ces ions sont transmis à la trappe à ions linéaire au niveau de l'étage haute pression où ils subiront un nouveau processus de fragmentation, soit par CID ou par ETD. Les ions petits-fils générés seront ensuite analysés dans l'étage basse pression de la trappe à ions linéaire pour déterminer leur rapport  $m/z$ . Cette méthode d'analyse est nommée nanoLC-HCD-MS<sup>2</sup>-pd-CID-MS<sup>3</sup> (Hsiao *et al.*, 2017).

## B. Interprétation des résultats

Appliquée à l'analyse des glycanes, la nanoLC-HCD-MS<sup>2</sup>-pd-CID-MS<sup>3</sup> permet de quantifier avec une grande fiabilité les épitopes glycaniques encore nommés glycotopes. En effet, comme je l'ai rappelé au début de ma thèse, le nombre d'oligosaccharides pouvant être formé à partir de 2 monosaccharides est de 632 et de quelques 10<sup>12</sup> à partir de seulement 6 monosaccharides différents (Morelle *et al.*, 2004). Ainsi, la quantification qui s'effectue sur la base de la masse moléculaire d'un glycanne peut conduire à une sur-évaluation de la quantité du glycotope visé. Néanmoins, la structure des ions fils obtenus lors d'une analyse MS<sup>2</sup> d'un ion parent dépend largement de la structure de ce dernier. Ainsi, en ciblant un ion-fils particulier, on restreint le nombre de structures pouvant le générer rendant ainsi plus précise la quantification. Cette dernière sera encore plus précise si elle est basée, non pas sur l'intensité du signal de l'ion fils mais sur l'intensité du signal d'un ion petits-fils provenant de cet ion fils lors d'une analyse MS<sup>3</sup>. Cela restreint encore davantage le nombre d'ions parents pouvant le générer. A la fin de l'analyse, les données *m/z* acquises lors des analyses MS<sup>n</sup> sont compilées dans le logiciel Glypick qui effectue une quantification des glycotopes à partir de la somme des intensités de l'ion cible détecté et ce, pour l'intégralité des scans réalisés durant l'analyse LC d'un mélange complexe de glycanes.

# RESULTATS

---

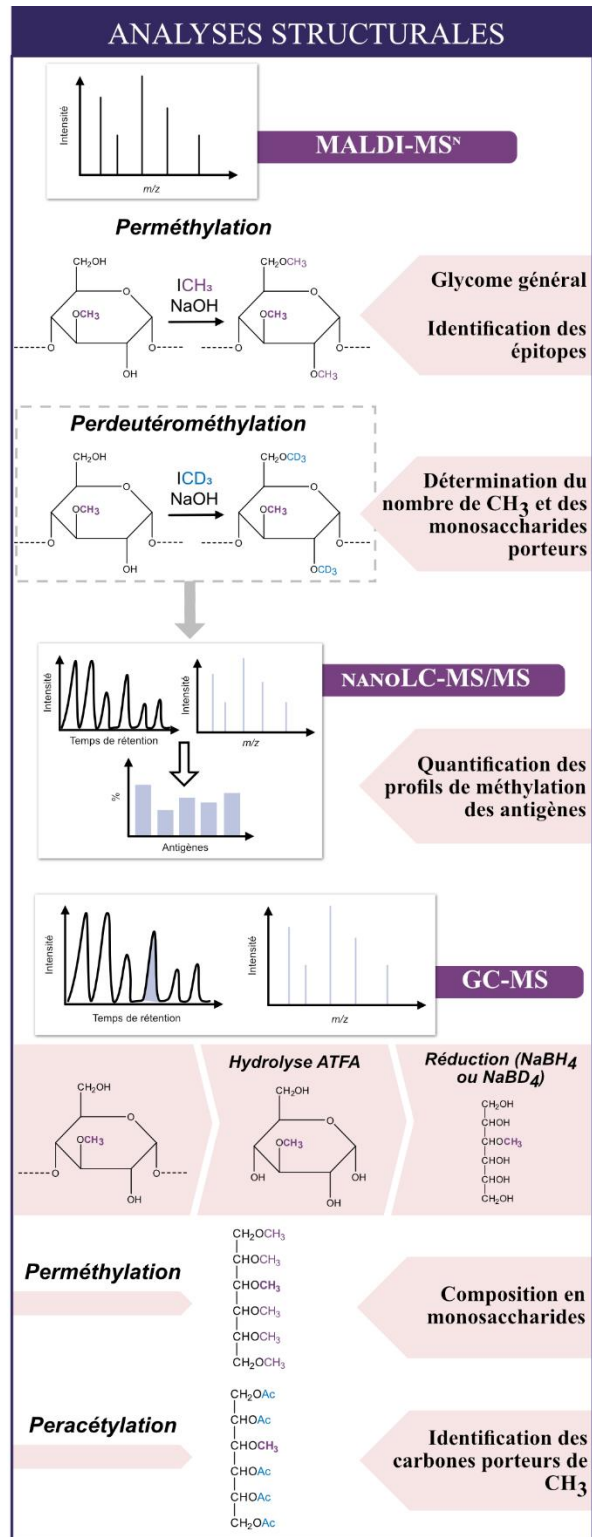
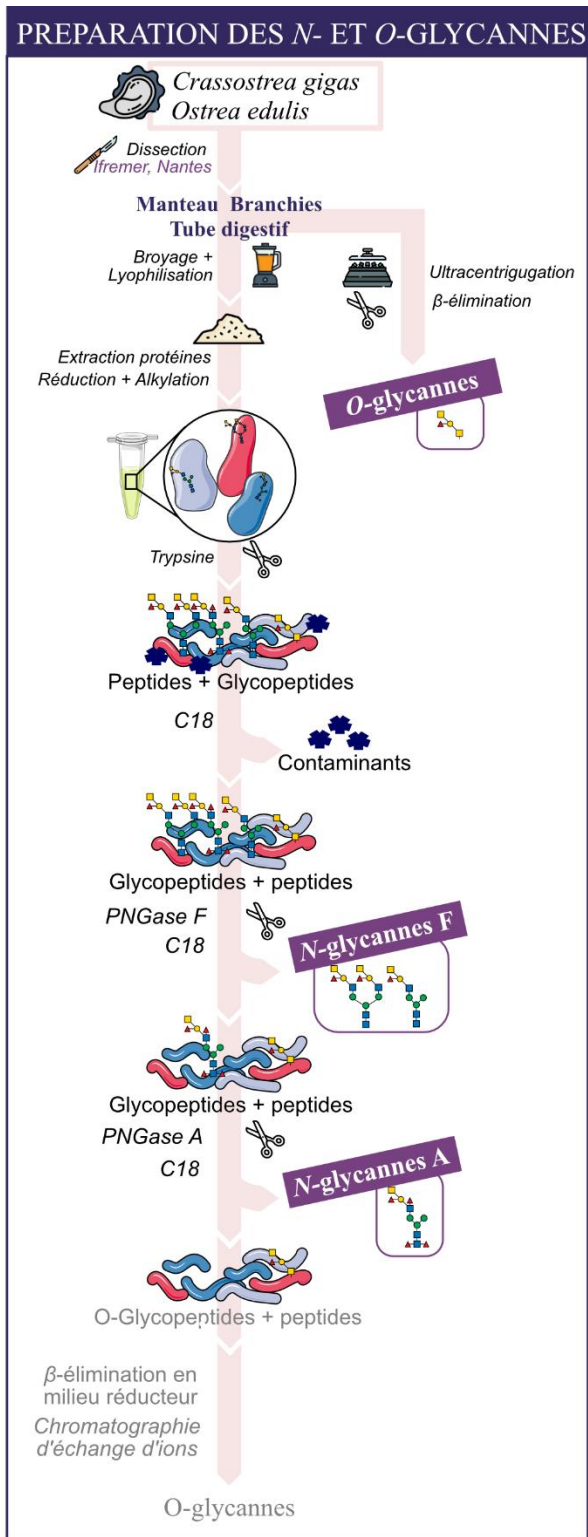
# Chapitre 1 : ANALYSE STRUCTURALE DES *N*-ET *O*-GLYCANNES OBTENUS A PARTIR DE TISSUS DES ESPECES D'HUITRE *C. gigas* ET *O. edulis*

## I. Méthodologie employée pour la préparation des échantillons

La méthodologie complète employée pour la préparation des *N*- et *O*-glycannes en vue de leurs analyses structurales est présentée à la **Figure 42**.

### A. Libération et purification des *N*- et *O*-glycannes

Les branchies, manteaux et tube digestif des huîtres *C. gigas* et *O. edulis* ont été disséqués par nos collaborateurs du Laboratoire Ifremer « Santé Environnement et Microbiologie » (LSEM) à Nantes puis broyés et lyophilisés après réception. Le protocole d'extraction des protéines et GPs a été réalisé en soumettant les tissus à un traitement au Triton X100, un détergent neutre, couplé à un traitement mécanique aux ultrasons. (Delannoy *et al.*, 2020). Après extraction des protéines et GPs, les ponts disulfures ont été réduits et alkylés et les protéines digérées par la trypsine. Ces traitements ont pour but de faciliter l'action ultérieure des PNGases qui montrent une piètre activité sur des GPs natives. Pour éliminer un polymère supposé être du glycogène du fait de sa quantité très élevée chez l'huître, nous avons ensuite réalisé une extraction en phase solide (SPE pour Solid Phase Extraction) sur une cartouche contenant de la silice greffée avec des chaînons C<sub>18</sub>. Les peptides et glycopeptides ainsi débarrassés de ce polymère ont été soumis à l'action de la PNGase F et les *N*-glycannes libérés ont été séparés des peptides et glycopeptides résistants à la PNGase F grâce à une étape de SPE sur une cartouche C<sub>18</sub>. Pour rappel, le clivage de la PNGase F est restreint aux *N*-glycannes ne possédant pas de substituants sur le C3 du résidu de GlcNAc en position réductrice. Afin de libérer la totalité des *N*-glycannes, nous avons ensuite utilisé la PNGase A sur les glycopeptides restants. Nous aurions pu utiliser directement la PNGase A qui ne présente aucune restriction de clivage de la liaison glycanne-peptide. Néanmoins, nous avons choisi de l'utiliser après la PNGase F afin de générer deux pools distincts de *N*-glycannes, nommés respectivement *N*-glycannes F et *N*-glycannes A rendant ainsi plus facile et plus complet leur



**Figure 42 :** Méthodologie employée pour les analyses structurales des *N*-et *O*-glycannes obtenus à partir des branchies, tubes digestif et manteaux des huîtres *C. gigas* et *O. edulis*. Crédits icones : Flaticon.com

analyse par spectrométrie de masse MALDI-QIT-TOF. Après action de la PNGase A, les *N*-glycannes A ont été séparés des peptides et *O*-glycopeptides potentiels par SPE sur cartouche C<sub>18</sub>. En effet, dans le cas de la présence de *O*-glycosylprotéines dans les tissus testés, nous devons obtenir à l'issue de la SPE

une fraction composée de peptides et *O*-glycopeptides. Dans le but d'étudier les *O*-glycannes contenus dans cette fraction, nous l'avons soumise à une  $\beta$ -élimination en milieu réducteur. Après purification et perméthylation, les *O*-glycannes perméthylés ont été analysés par MALDI-MS. Malheureusement, aucun signal *m/z* présents sur les spectres de masse n'a pu être corrélé à une structure *O*-glycannique. L'explication que nous en avons donnée est que la succession des étapes SPE sur cartouche C<sub>18</sub> appauvrissait grandement les échantillons en *O*-glycopeptides rendant les *O*-glycannes indétectables en spectrométrie de masse. Pour pallier ce problème, nous avons décidé d'effectuer directement une  $\beta$ -élimination en milieu réducteur sur l'échantillon avant même la libération préalable des *N*-glycannes. Si cet essai n'a pas été concluant à cause de la très grande hétérogénéité des échantillons du fait de la présence de *N*-glycannes sous leurs formes réduites et non réduites, il nous a été possible d'identifier des *O*-glycannes de type mucine. Forts de ce résultat, nous avons décidé de mettre en œuvre un protocole adapté à la préparation des mucines par ultracentrifugation afin d'enrichir la population *O*-glycannique. Ce protocole sera détaillé dans la partie III p147.

## B. Traitements chimiques effectués sur les *N*- et *O*-glycannes

La stratégie employée pour l'établissement des structures glycanniques a été déployée sur 12 échantillons pour les *N*-glycannes (2 espèces d'huître, 3 tissus (manteau, branchie et tube digestif) et 2 PNGases) et 6 pour les *O*-glycannes (2 espèces et 3 tissus). Plusieurs traitements chimiques ont été réalisées en vue de les analyser par MALDI-MS, MALDI-MS<sup>n</sup>, nanoLC-MS/MS et GC-EI-MS.

### 1. Analyse par MALDI-MS pour l'établissement des *N*- et *O*-glycomes des tissus des huîtres *C. gigas* et *O. edulis*

Les travaux de Kurz *et al.* (2013) sur les *N*-glycannes dérivés à la 2-aminopyridine des hémocytes et du plasma de l'huître *C. virginica* ont révélé la présence de groupements méthyles (-CH<sub>3</sub>) sur certains monosaccharides. Dans le but de vérifier que de tels groupements étaient également présents dans les *N*-glycannes des 2 huîtres étudiées, nous avons décidé d'effectuer deux dérivations distinctes des glycannes avant leur analyse par MALDI-QIT-TOF (MS et MS<sup>n</sup>): la perméthylation (CH<sub>3</sub>) et la perdeutérométhylation (CD<sub>3</sub>).



La perméthylation nous a permis d'établir les différents *N*- et *O*-glycomes, c'est-à-dire l'ensemble des *N*- et *O*-glycannes présents au sein d'un tissu ou d'une espèce. A partir des masses moléculaires déduites des rapports  $m/z$  des adduits sodium  $[M+Na]^+$  de chacun des glycannes, nous avons pu établir une composition brute en monosaccharides de ces glycannes en combinant la masse moléculaire des différents monosaccharides. Quant à la structure proposée pour chacun des *N*-glycannes, elle repose sur cette composition brute, l'analyse des fragments en MALDI-MS<sup>n</sup>, les analyses GC-MS (expliquées ultérieurement) ainsi que la connaissance que nous avons des voies de biosynthèse des *N*-glycannes qui apparaît particulièrement bien conservée dans le règne animal et des *O*-glycannes.

Le second traitement nous a permis de mettre en évidence les groupements méthyles naturellement présents au sein des *N*- et *O*-glycannes. En effet, lors de la deutérométhylation, l'ajout d'un groupement CD<sub>3</sub> se traduit par une augmentation de masse de 18 Da contre 15 Da dans le cas d'un groupement méthyle. Ainsi, toute diminution de 3 Da par rapport à la masse attendue indique la présence naturelle d'un groupement CH<sub>3</sub> (+15 Da) au sein du glycanne qui a empêché la fixation d'un groupement CD<sub>3</sub> (+18 Da).

## 2. Analyse par nanoLC-MS/MS des *N*-glycannes perméthylés des tissus des huîtres *C. gigas* et *O. edulis*

La quantification de chaînes oligosaccharidiques perméthylées ou de motifs oligosaccharides provenant d'un processus de fragmentation par un spectromètre de masse MALDI-QIT-TOF reste relativement imprécise. Il y a plusieurs raisons à cela. Tout d'abord, l'utilisation d'une cage à ions ne permet pas d'avoir un spectre global de l'ensemble des chaînes oligosaccharidiques présentes dans un échantillon dès lors que leur rapport  $m/z$  s'étale de 900 à 5 000 Da. Cela tient au fait que le piégeage des ions ne peut se faire que sur une gamme restreinte de  $m/z$ . Ainsi, on a coutume de sous fractionner les *N*-glycomes en 3 sous-*N*-glycomes, un pour les signaux  $m/z$  compris entre 350 et 1 600 Da, un pour les signaux  $m/z$  compris entre 800 et 3 000 Da et le dernier pour les signaux  $m/z$  compris entre 1 700 et 4 500 Da. De fait, la comparaison d'une intensité de signal dans une zone avec celle d'un signal acquis dans une autre zone conduit à des distorsions parfois très importantes. Une autre cause de distorsion est liée à la co-cristallisation matrice-échantillon qui ne conduit pas à une distribution uniforme des analytes

dans le cristal final. D'autres contraintes que celles précédemment citées existent. Pour pallier ces inconvénients, les *N*-glycannes perdeutérométhylés porteurs de motifs antigéniques ont été analysés par nanoLC-MS<sup>3</sup> qui permet de fractionner les chaînes oligosaccharidiques en amont de la source du spectromètre de masse (réduction de l'hétérogénéité du glycome) et effectue leur quantification via leurs ions petits-fils issus d'analyses MS<sup>3</sup>. Ces analyses ont été réalisées par l'équipe du Dr. Kay-Hooi Khoo de l'Institute of Biochemical Sciences à l'université nationale de Taiwan. Elles nous ont permis de quantifier les profils de méthylation des différents Ags identifiés par la quantification d'ions caractéristiques propres à ces Ags.

### 3. Analyses par GC-EI-MS des différents dérivés monosaccharidiques issus des *N*-glycannes des tissus des huîtres *C. gigas* et *O. edulis*

L'identification des différents monosaccharides entrant dans la composition des *N*-glycannes a été faite en libérant les monosaccharides grâce à un traitement à l'acide trifluoroacétique (ATFA) (4h à 80 °C) suivie de leur réduction par le borodeutérure ou borohydrure de sodium qui vise à les transformer en leuritol correspondant. Deux traitements chimiques ont été effectués sur les itols générés, d'une part, une perméthylation pour identifier les monosaccharides présents dans les *N*-glycannes, et d'autre part, une peracétylation pour identifier les monosaccharides renfermant des groupements méthyles et les localiser précisément au sein de ces monosaccharides.

La perméthylation a été réalisée du fait de l'existence de groupements méthyles au sein de certains *N*-glycannes apparus lors des analyses MALDI-MS des *N*-glycannes deutérométhylés (partiellement méthylés et deutérométhylés de ce fait). Or, il n'existe pas de standards de monosaccharides commerciaux monométhylés sur d'autres carbones que le carbone 1. La perméthylation sur les dérivés itols a pour effet négatif de masquer les groupements méthyles naturellement présents sur certains monosaccharides mais comme effet bénéfique de pouvoir identifier ces monosaccharides en comparant leurs comportement chromatographique et spectre de fragmentation EI-MS à ceux dérivés itols perméthylés générés à partir de monosaccharides standards.

Quant à la peracétylation des dérivés itols suivie de leur analyse par GC-EI-MS, elle a eu pour but, d'une part, d'identifier les monosaccharides porteurs d'un groupement méthyle à l'état natif, et

d'autre part, de préciser le numéro du carbone porteur de ce groupement. En effet, la peracétylation a conduit à des dérivés semblables à ceux obtenus lors de l'analyse de liaisons grâce aux dérivés PMAAs, hormis pour les résidus d'HexNAc. En effet, dans le cas de notre traitement, le groupement amide n'est pas sous la forme NCH<sub>3</sub>Ac comme pour les PMAAs mais sous la forme NHAc. Néanmoins, les mêmes règles de fragmentation que celles explicitées dans l'introduction (IV.C p98) pour les PMAAs s'appliquent à nos dérivés.

## II. Analyse des *N*-glycannes

### A. Publication

L'ensemble des résultats obtenus pour l'analyse des *N*-glycannes issus du manteau, des branchies, et du tube digestif des huîtres *C. gigas* et *O. edulis* selon les approches expérimentales décrites dans les pages précédentes ont fait l'objet de la rédaction d'une publication intitulée « Species-Specific N-Glycomes and Methylation Patterns of Oysters *Crassostrea gigas* and *Ostrea edulis* and Their Possible Consequences for the Norovirus–HBGA Interaction » et publiée dans *Marine drugs* (*Mar. Drugs* 2023, 21(6), 342; <https://doi.org/10.3390/md21060342>) que vous trouverez dans les pages suivantes. Vous trouverez également les données supplémentaires associées à la publication en Annexe (p209).

Article

# Species-specific *N*-glycomes and methylation patterns of oysters *Crassostrea gigas* and *Ostrea edulis* and their possible consequences for norovirus-HBGA interaction

Audrey Auger<sup>1</sup>, Shin-Yi Yu<sup>1</sup>, ShihYun Guu<sup>2</sup>, Agnès Quéméner<sup>3</sup>, Gabriel Euller-Nicolas<sup>4</sup>, Hiromune Ando<sup>5</sup>, Marion Desdouits<sup>4</sup>, Françoise S. Le Guyader<sup>4</sup>, Kay-Hooi Khoo<sup>2,6</sup>, Jacques Le Pendu<sup>7</sup>, Frederic Chirat<sup>1</sup> and Yann Guerardel<sup>1,5\*</sup>

**Citation:** Auger, A.; Yu, S.-Y.; Guu, S.-Y.; Quéméner, A.; Euller-Nicolas, G.; Ando, H.; Desdouits, M.; Le Guyader, F.S.; Khoo, K.-H.; Le Pendu, J.; et al. Species-Specific *N*-Glycomes and Methylation Patterns of Oysters *Crassostrea gigas* and *Ostrea edulis* and Their Possible Consequences for the Norovirus–HBGA Interaction. *Mar. Drugs* 2023, 21, 342. <https://doi.org/10.3390/md21060342>.

Academic Editor: Firstname  
Lastname

Received: 8 May 2023

Revised: 29 May 2023

Accepted: 29 May 2023

Published: 2 June 2023



**Copyright:** © 2023 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

<sup>1</sup> Univ. Lille, CNRS, UMR 8576-UGSF-Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France; audrey.auger.etu@univ-lille.fr (A.A.); shinyi.yu@gmail.com (S.-Y.Y.); frederic.chirat@univ-lille.fr (F.C.)

<sup>2</sup> Institute of Biological Chemistry, Academia Sinica, Nangang, Taipei 11529, Taiwan; ufogirl100@gmail.com (S.-Y.G.); kkhoo@gate.sinica.edu.tw (K.-H.K.)

<sup>3</sup> Nantes Université, Inserm UMR 1307, CNRS UMR 6075, Université d'Angers, CRCI2NA, F-44000 Nantes, France; agnes.quemener@univ-nantes.fr

<sup>4</sup> MASAE Microbiologie Aliment Santé Environnement, Ifremer, BP 21105, 44311 Nantes, France; gabriel.euller@ifremer.fr (G.E.-N.); marion.desdouits@ifremer.fr (M.D.); soizick.le.guyader@ifremer.fr (F.S.L.G.)

<sup>5</sup> Institute of Glyco-core Research (iGCORE), Gifu University, Gifu 501-1193, Japan; ando.hiromune.i0@f.gifu-u.ac.jp

<sup>6</sup> Institute of Biochemical Sciences, National Taiwan University, Taipei 10617, Taiwan

<sup>7</sup> Immunology and New Concepts in ImmunoTherapy, Nantes Université, Inserm, CNRS, UMR 1302/EMR6001, 44200 Nantes, France; jacques.le-pendu@univ-nantes.fr

\* Correspondence: yann.guerardel@univ-lille.fr

**Abstract:** Noroviruses, the major cause of acute viral gastroenteritis, are known to bind to Histo-Blood Group Antigens (HBGA), including ABH groups and Lewis-type epitopes, which decorate the surface of erythrocytes and epithelial cells of their host tissues. The biosynthesis of these antigens is controlled by several glycosyltransferases, the distribution and expression of which varies between tissues and individuals. The use of HBGAs as ligands by viruses is not limited to humans, as many animal species, including oysters, which synthesize similar glycan epitopes that act as a gateways for viruses, become vectors for viral infection of humans. Here we show that different oyster species synthesize a wide range of *N*-glycans that share histo-blood A-antigens but differ in the expression of other terminal antigens and in their modification by *O*-methyl groups. In particular, we have shown that the *N*-glycans isolated from *Crassostrea gigas* and *Ostrea edulis* exhibit exquisite methylation patterns of their terminal *N*-acetylgalactosamine and fucose residues in terms of position and number, adding another layer of complexity to the post-translational glycosylation modifications of glycoproteins. Furthermore, modelling of the interactions between norovirus capsid proteins and carbohydrate ligands strongly suggests that methylation has the potential to fine-tune the recognition events of oysters by virus particles.

**Keywords:** glycomics; norovirus ligands; oysters; methylation

## 1. Introduction

Acute gastroenteritis is the second leading cause of morbidity and mortality in infants and young children worldwide [1]. These infections are very short (about 2 days) and cause diarrhea, vomiting and fever that lead to significant dehydration in young and elderly subjects and may lead to death with more than 500,000 victims per year. In addition to its high social cost, infectious gastroenteritis has a very high economic cost due to care and lost work days. Among the many causative agents, viruses are responsible for nearly three out of four infections during seasonal gastroenteritis episodes, particularly noroviruses (NoV). [1–4]. NoVs are non-enveloped, single-stranded RNA viruses belonging to the *Caliciviridae* family [5]. Among the ten described genogroups, which can be broken down into numerous genotypes, only five have been described to be involved in human pathologies (GI, GII, GIV, GVIII and GIX) with a high prevalence (80 to 90% of cases) for the GII strain [6]. NoV genotypes responsible of Human infections are known to recognize Histo-Blood Group Antigen (HBGA) glycan motifs including ABH and Lewis (Le) antigens decorating the surface glycans of human erythrocytes and epithelial cells of different tissues [7]. They are also present in body fluids (serum, saliva, milk,...) either associated to proteins or as free oligosaccharides [8]. The biosynthesis of these antigens is under the control of several glycosyltransferases which are differentially distributed and expressed in different tissues and individuals [9] and which act on multiple carbohydrate substrates [10]. For glycoproteins, the synthesis is initiated by the addition of a  $\alpha$ 1-2Fuc residue on the Gal residue of type-1 or type-2 LacNAc motif to yield H antigen. This step is catalyzed by a  $\alpha$ 1,2 fucosyltransferase encoded by the genes *FUT1* in erythrocytes and *FUT2* in epithelial cells. About 20% of European individuals, described as non-secretor, present an inactive *FUT2* gene and therefore cannot synthesize ABH antigen in epithelia. As a result, they carry only Le<sup>a</sup> and Le<sup>x</sup> antigens on the surface of their epithelial cells. Three other glycosyltransferases, namely the  $\alpha$ 1,3/4 fucosyltransferase encoded by the *FUT3* gene, the  $\beta$ 1-3 galactosyltransferase B-enzyme and/or the  $\beta$ 1-3 N-acetylgalactosyltransferase A-enzyme can sequentially or differentially act to generate multiple HBGA including, A-type 1/2, B-type 1/2, ALe<sup>b/y</sup> or BLe<sup>b/y</sup> Ag [10]. Of particular importance, these so-called Histo-blood group antigens are also shared with many other animal species, including dogs, pigs and cattle and therefore could contribute to cross-species transmission and zoonotic emergence of NoV [11–13]. Importantly, recognition of HBGA motifs by enteric viruses is strain-specific, leading to differential susceptibility according to the expression of individual HBGAs [14].

Among the multiple routes of contamination of human, food, and in particular oysters, are a common vector, due to multiple factors that can be summarized as follows [15]. First, oyster production sites, located in coastal waters potentially close to contaminated effluents, are particularly exposed to circular contamination by humans and livestock sewage [6,16,17]. They are also often consumed raw, so unlike most other potentially contaminated seafood, they do not undergo any possible heat inactivation. Then, the oyster's feeding method of filtering large volumes of water allows it to concentrate and accumulate viruses and other pathogens from the environment, even in seawater. [18,19]. Finally, oysters are capable to retain viruses over a very long period time, up to 4

weeks, despite depuration procedure before consumption [20–23], which strongly suggests the occurrence of a specific and high affinity retention mechanism. Importantly, oyster tissues, including the mantle, digestive tract and gills, were shown to express HBGA-like motifs that may be similar to some of those present in human tissues to which viruses could attach. This would provide a strong rationale for the specific and high affinity binding of norovirus to oyster tissues. [24–26]. Consistent with this hypothesis, it was observed that the recognition of these motifs by different virus strains was tissue dependent. [27,28]. Thus, GI strains bind more strongly to oyster digestive tissues via HBGA-like structures than GII strains, which also bind to the gills and mantle. While these glycans are recognized by HBGA-specific lectins or anti-HBGA [24,26], their exact structure remains unknown, hampering our understanding of the possible selection operated by oysters among the very diverse strains of noroviruses.

Invertebrate *N*-glycans significantly differ from canonical *N*-glycans observed in vertebrates [29]. First, they tend to be shorter with a high proportion of paucimannosylated or pseudohybrid types *N*-glycans. Second, they are substituted by monosaccharides rarely observed in vertebrates such as pentoses (e.g. xylose). Third, reducing terminal GlcNAc residue is commonly substituted by Fuc residues at both C-3 and C-6 positions or by a Gal( $\beta$ 1-4)Fuc disaccharide at C-6. Finally, substituents such as methyl, *N*-methyl-aminoethylphosphonate or phosphoethanolamine groups may also be present. Very little information is currently available on the *N*-glycome of oysters, except for a comprehensive analysis of the *N*-glycans associated with total plasma and haemocytes of the Pacific oyster *Crassostrea virginica*. [30]. In this report, the authors have successfully identified a highly variable *N*-glycome characterized by the presence of human histo-blood group A antigen, most of which were methylated on the terminal GalNAc residue, and by a high degree of sulfation on Gal residues. In parallel, detailed analysis of the binding specificity of *Crassostrea virginica* Galectin 1 (CvGal1) showed that it recognized with high affinity type 2, then type 1, blood group A oligosaccharides and could bind to endogenous hemocyte glycoproteins ( $\beta$ -integrin and dominin) in a carbohydrate-dependent manner [31].

In the present study, we have deciphered the tissue-specific *N*-glycomes of two oyster species, *Crassostrea gigas* and *Ostrea edulis* and showed that they synthesize related but significantly different sets of *N*-glycans. Although they partially shared similar terminal HBGA-like antigens, these were substituted on different monosaccharides and at different positions by one to three methyl-groups, dramatically increasing the number of exposed antigens. Finally, modelling of the interactions of differentially methylated HBGA identified in the two species with P-domains of prototypical noroviruses strongly suggests that methylation of carbohydrate ligands may represent a new paradigm for the fine-tuning of carbohydrate-protein interactions in nature and a new molecular player in the diversification of glycans involved the evolution of long-term host-pathogen recognition [32].

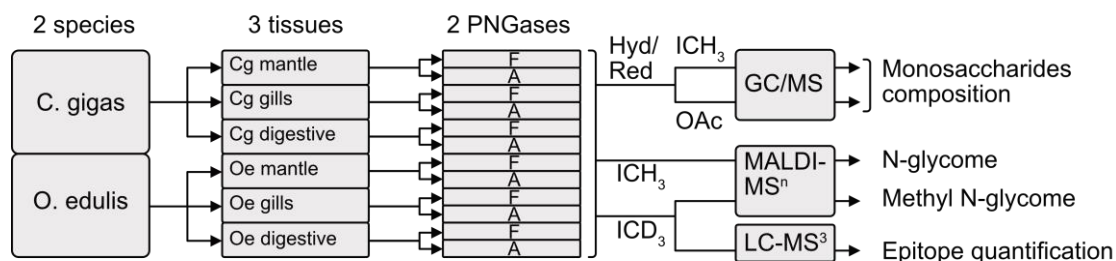
## 2. Results

### 2.1 Release and analysis of *N*-glycans

Three organs, namely gills, digestive tract and mantle were collected from three separate batches of 10-15 individuals from two different species of

oysters *C. gigas* and *O. edulis*, and immediately processed according to a standard procedure of *N*-linked glycans isolation [33]. Briefly, whole proteins and glycoproteins were solubilized from tissues by ultrasound treatment in presence of triton X100. After removing insoluble materials by centrifugation, glycoproteins were reduced and alkylated prior to enzymatic cleavage by trypsin in order to facilitate the action of peptide-*N*-glycosidases (PNGases) used to release *N*-glycans from asparagine residues. *N*-glycans were sequentially released using two PNGases, PNGase F [34] then PNGase A [35], generating two pools of *N*-glycans according to the presence or not of C3 substitution, generally by a Fuc residue, of the reducing GlcNAc of the chitobiose core. The two subglycomes, PNGase F and PNGase A, were analysed individually in order to obtain a more complete interpretation of the spectra.

The analytical workflow of purified glycans is summarized in Figure 1. In brief, the organ specific *N*-glycomes of the two oyster species were first analyzed by GC/MS in order to determine their monosaccharide compositions. The *N*-glycan diversity was then evaluated by mass spectrometry in permethylated forms using MALDI-MS and MS<sup>n</sup> in order to achieve a higher sensitivity and generate more informative MS<sup>n</sup> spectra in terms of sequence and monosaccharide linkages. The in depth mapping and interspecies comparison of the natural methylation profiles of *N*-glycans was performed by MS<sup>2</sup> and MS<sup>3</sup> sequencing following I-CD<sub>3</sub> derivation. Finally, individual epitopes were relatively quantified by LC-MS<sup>3</sup> analyses.



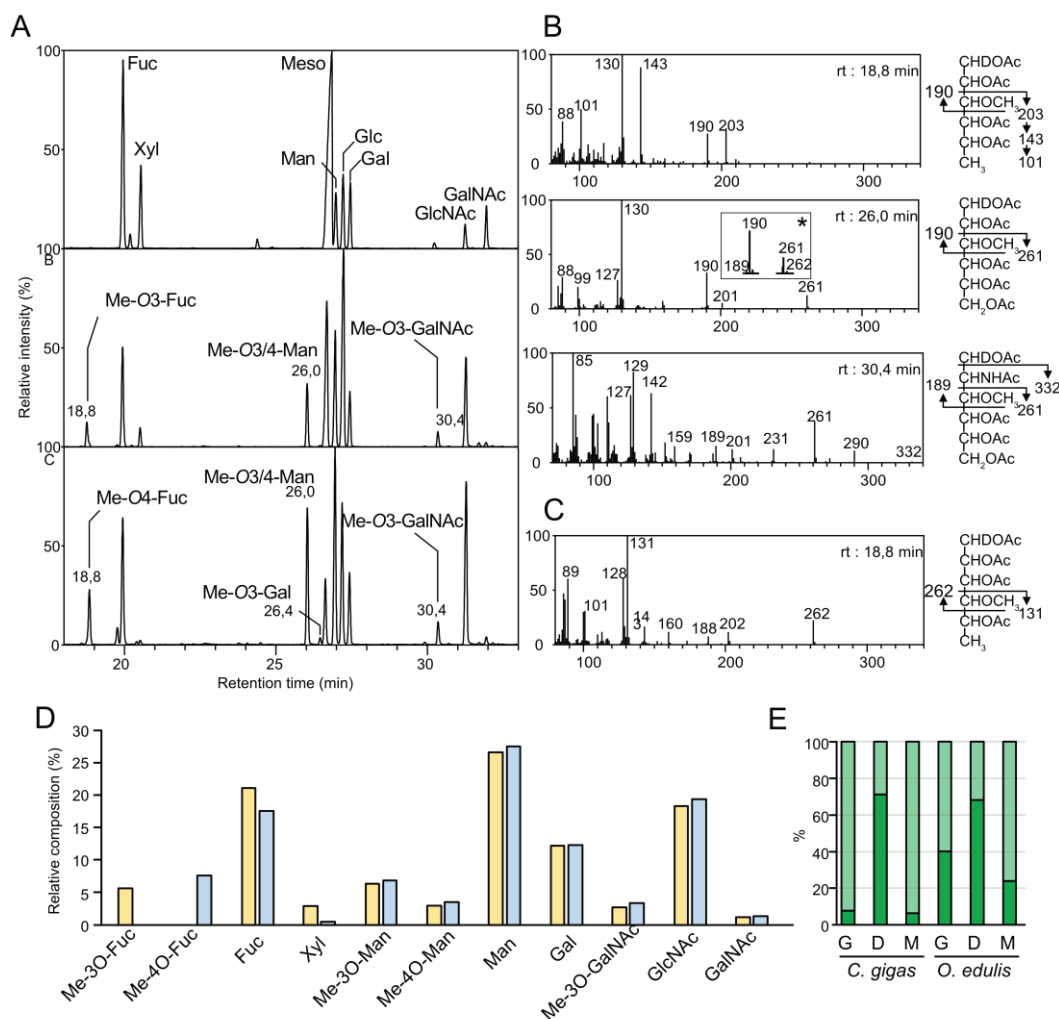
**Figure 1.** Analytical workflow used to characterize *N*-glycome of gills, digestive tract and mantle of two oyster species, *Crassostrea gigas* and *Ostrea edulis*. Hyd/Red, hydrolysis and reduction; ICH<sub>3</sub>, permethylation; ICD<sub>3</sub>, perdeuteromethylation; OAc, peracetylation, F, PNGase F; A, PNGase A.

## 2.2. Monosaccharides composition of *N*-glycans

The monosaccharide composition of PNGase F and PNGase A released *N*-glycans from both oyster species were established as acetylated alditol residues by GC/MS, in comparison with authentic standards of common monosaccharides. As expected, chromatograms showed the presence of Fuc, Xyl, Man, Gal, GalNAc and GlcNAc for glycans isolated from both species, albeit with a much lower proportion of Xyl in *O. edulis* in comparison with *C. gigas* (Fig. 2A and Table S1). Glc was also observed but considered as a ubiquitous contaminant that may originate from remnant glycogen that is synthesized in large amounts by oysters [36]. The chromatograms revealed three additional signals at RT 18.8, 26.0 and 30.4min that did not match with authentic standards (Fig. 2A). EI-MS analysis of these three signals typified them as *O*-methylated monosaccharides Me-*O*-deHex (RT 18.8 min) Me-*O*-Hex (TR 26.0 min)

and Me-O-HexNAc (RT 30.4 min) as demonstrated below. The nature of these monosaccharides and the localization of methyl groups were established by GC/EI-MS following reduction with sodium borodeuteride ( $\text{NaBD}_4$ ) that allows to discriminate C1 from C6. Under these conditions, cleavage between skeletal carbons preferentially occurs between a carbon carrying a methoxy group and a carbon carrying an acetyl group generating two intense fragments which allows an easy positioning of the unique methyl group. In particular, EI-MS analysis of the GC/MS signal at RT 18.8 min identified in *C. gigas* generated two intense primary fragments  $[\text{CDHOAc-CHOAc-CHOMe}]^+$  at  $m/z$  190 and  $[\text{CHOMe-CHOAc-CHOAc-CH}_3]^+$  at  $m/z$  203, confirmed by the observation of secondary fragment ions generated by Mc-Lafferty rearrangements at 130 and 141-101 respectively, which localized the methyl group at the C3 position of deHex residue and established the presence of Me-3O-deHex in *C. gigas* (Fig. 2B). EI-MS fragmentation of the signal at the same retention time from *O. edulis* samples generated a very different pattern characterized by two intense fragment ions  $[\text{CDHOAc-(CHOAc)}_2\text{-CHOMe}]^+$  at  $m/z$  262 and  $[\text{CHOMe-CHOAc-CH}_3]^+$  at  $m/z$  131, which indicated that deHex residue was methylated on C4 in *O. edulis* (Me-4O-deHex) rather than on C3 as in *C. gigas* (Fig. 2C). Then, EI-MS analysis of the second signal at RT 26 min showed two major signals at  $m/z$  190 and  $m/z$  261 identified as  $[\text{CHOMe-CHOAc-CDHOAc}]^+$  and  $[\text{CHOAc-CHOAc-CHOAc-CHOMe}]^+$  fragments respectively, indicative of the presence of a Hex residue substituted by a methyl group in C3 position (Me-3O-Hex) (Fig. 2B). A set of signals at  $m/z$  189 and  $m/z$  262 attributed to  $[\text{CHOAc-CHO-Ac}]^+$  and  $[\text{CHOMe-CHOAc-CHOAc-CDHOAc}]^+$  fragments demonstrated the presence of a very minor methylated hexose at C4 position (Me-4O-Hex) co-eluting with the first one. Finally, the signal at RT 30,4 min was characterized as a Me-3O-HexNAc residue based on the observation of primary fragment  $[\text{CDHOAc-CHNH}_2\text{OAc-CHOMe}]^+$  at  $m/z$  189 and  $[\text{CHOMe-(CHOAc)}_2\text{-CH}_2\text{OAc}]^+$  at  $m/z$  261. Any possible ambiguity between a methylation on C3 or C4 position due to the combined presence of deuterium atom on C1 position and *N*-acetyl group on C2 position was lifted by running a similar experiment after  $\text{NaBH}_4$  reduction rather than  $\text{NaBD}_4$  (data not shown).





**Figure 2.** Monosaccharide composition analysis of *N*-glycans. (A) Total Ion Count generated from GC-MS analysis of itol-acetate derivatives of (top) authentic standards, (middle) *N*-glycans released from the mantle of *C. gigas* by PNGase F, (bottom) *N*-glycans released from the mantle of *O. edulis* by PNGase F. (B) EI-MS fragmentation patterns of itol-acetates signals from *C. gigas* at (top) 18.8 min identified as Me-3O-Fuc; (middle) 26.0 min identified as Me-3O-Man, inset (\*) corresponds to a minor signal identified as Me-4O-Man; (bottom) 30.4 min identified as Me-3O-GalNAc. (C) EI-MS fragmentation patterns of itol-acetates signal at 18.8 min from *O. edulis* identified as Me-4O-Fuc, all other signals were identical to *C. gigas*. (D) Relative quantifications (in %) of monosaccharides calculated as the sum of individual signals of *N*-glycans released by PNGase F and A from the three tissues of *C. gigas* (yellow) and *O. edulis* (blue). The percentage of Me-3O-Man and Me-4O-Man was determined by computing the ratios of signals at  $m/z$  190/261 vs  $m/z$  189/262 recorded on the EI-MS spectra. (E) Ratios (in %) of Me-3O-Man (light green) and Me-4O-Man (dark green) in *N*-glycans isolated from individual tissues of *C. gigas* and *O. edulis* (G, gills; D, digestive tissue; M, mantle).

In the absence of authentic standards to identify the nature of methylated deHex, Hex and HexNAc residues, samples were hydrolyzed, reduced and permethylated prior to analysis in order to negate the effect of natural methylation on monosaccharide migration and compared with authentic reduced-permethylated (Fig. S1). Identification of Fuc as the only deHex signal in all samples established the presence of Me-3O-Fuc and Me-4O-Fuc differentially expressed in *C. gigas* and *O. edulis*. Then, the modification of the relative intensities of permethylated monosaccharides

compared to peracetylated ones established the presence of Me-3O-GalNAc as well as a mixture of Me-4O-Man and Me-3O-Man in both oyster species. Altogether, GC-MS analyses of permethylated and peracetylated reduced monosaccharides of *N*-glycans isolated from *C. gigas* and *O. edulis* permitted identification of Me-3O-Fuc, Me-4O-Fuc, Fuc, Xyl, Me-3O-Man, Me-4O-Man, Gal, Me-3O-GalNAc, GalNAc and GlcNAc as the major components of *N*-glycans (Fig. 2A). A very small amount of Me-3O-Gal was also found in all of the samples. The identification of methylated monosaccharides in both species of oysters is in total agreement with numerous studies that have reported the occurrence of methylated monosaccharides in invertebrates glycans, including methylated GalNAc and Man residues substituting *N*-glycans isolated from the hemolymph of the *C. virginica* oyster [30]. Comparison of the relative quantities of individual monosaccharides in all twelve *N*-glycan fractions showed striking differences that were more pronounced between species than between tissues (Fig. 2D). The more obvious qualitative difference concerns the exclusive presence of Me-3O-Fuc in *C. gigas* compared to the presence of Me-4O-Fuc in *O. edulis*. Another dramatic difference concerns the much higher prevalence of Xyl in *C. gigas* *N*-glycans (up to 3.4 % in PNGaseF and 9.4% in PNGaseA) compared to *O. edulis* (up to 0.4 % in PNGase F and 2.2% in PNGase A) (Table S1). In contrast, Me-3O-Man and Me-4O-Man have been identified in both species, but with very different ratios depending on the tissue, as illustrated by the large prevalence of Me-3O-Man in gills and mantle compared to the prevalence of Me-4O-Man in digestive tissues (Fig. 2E). Although the figures are different, similar trend is observed in both species, except for *O. edulis* PNGase-A fraction from gills in which the proportion of Me-4O-Man is slightly higher (Fig. S2). Finally, quantitative composition analysis suggests that *O. edulis* tissues are substituted by a much higher quantity of *N*-glycans than *C. gigas*, although a wide variability was observed between individual organs and types of *N*-glycans (Table S1). It should however be noted that digestive track in *C. gigas* is noticeably richer in *N*-glycans than the two other organs, whereas it is lower in *O. edulis*.

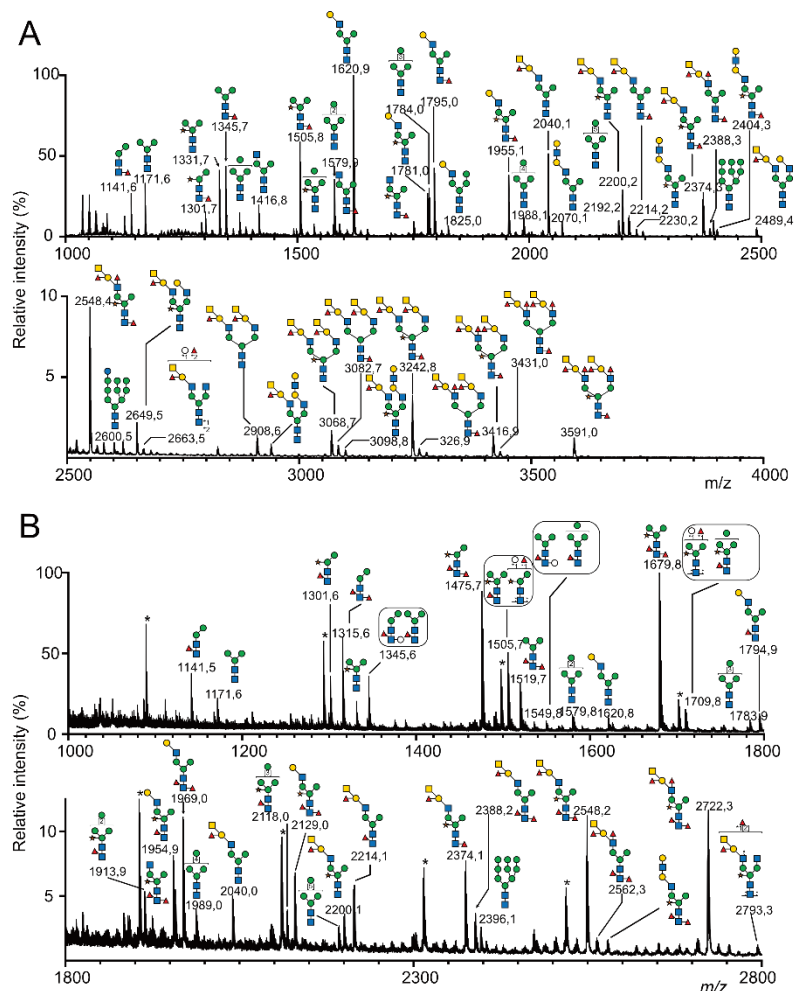
### 2.3. *N*-glycome analysis

Permethylated *N*-glycans were detected as their  $[M+Na]^+$  adducts and their monosaccharides compositions were inferred by calculating the numbers of Hex, HexNAc, deHex and Pent residues. The nature of monosaccharides was postulated from the known conserved *N*-glycans biosynthetic pathways and further confirmed by composition analysis. Structural assignment of all glycans was mostly based on the specific CID MS/MS fragmentation patterns afforded by the  $[M+Na]^+$  molecular ions of permethylated glycans in positive mode, as previously documented [37]. MALDI-QIT-TOF mainly generated sodiated B-fragment ions from the non-reducing end of permethylated *N*-glycans [38]. These B-ion are particularly useful to reveal the presence of antigen motif such as HBGA or Lewis antigens. Other relevant B/Y fragmentation events also occurred between the two GlcNAc residues of the chitobiose moiety that generated B/Y-ions used to identify substituents on the terminal reducing GlcNAc such as common core-fucosylation or other substitutions. Moreover, the number of free -OH groups on terminal Man residues of the core generated by the release of antennae informed the number of antennae.

Along the B/Y cleavage events, C/Z ions between Hex or Fuc  $\alpha/\beta$ 1-3 HexNAc from the non-reducing end of antennae permitted to unambiguously distinguish type-1 from type-2 LacNAc antennae and consequently Le<sup>a</sup>/Le<sup>b</sup> from Le<sup>x</sup>/Le<sup>y</sup> units. It should be noted that low energy collisions allowed by this mass spectrometer did not generate intense cross ring cleavages preventing further information on the bonds.

### 2.3.1 N-glycome of *C. gigas*

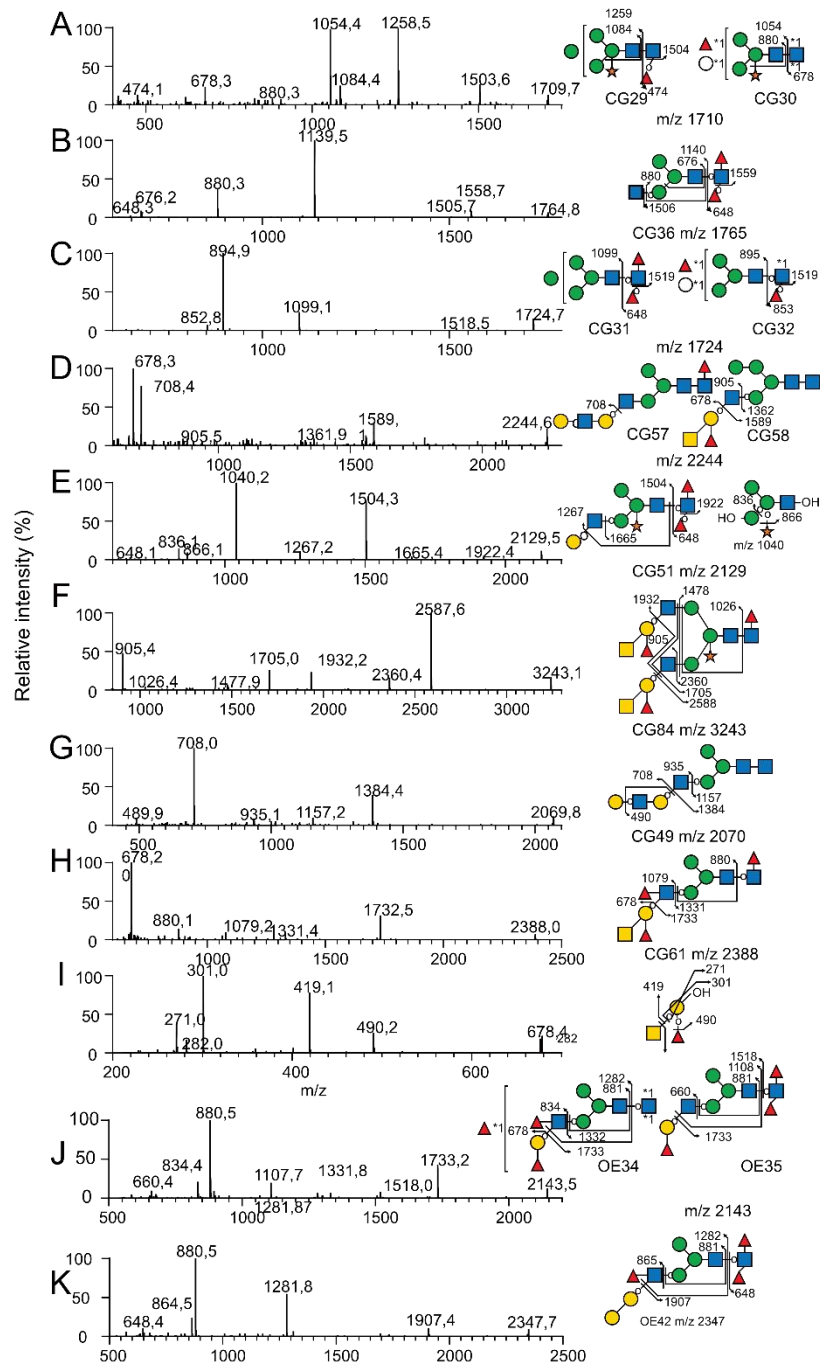
Altogether, we have confidently identified by mass spectrometry 88 N-glycans from the *C.gigas* two sub-N-glycomes, as illustrated in Figure 3, that were labeled CG1 to CG88 and compiled in Table S2. These were assigned to four different families namely paucimannose (26 NG), oligomannose (9 NG), hybrid (2 NG) and complex (51 NG). Out of the 51 complex N-glycans, it is noteworthy that the majority (34 glycans) were substituted by a single complex antenna that are referred to as pseudo-hybrid according to the recent review of Paschinger and Wilson, whereas the 17 remaining were complex bi-antennary N-glycans [29].



**Figure 3.** MALDI-MS analysis of permethylated N-glycans isolated from the mantle of *Crassostrea gigas*. Prior to their permethylation, N-glycans were sequentially released from mantle of *C. gigas* by (A) peptidyl-N-glycosidase F and (B) peptidyl-N-glycosidase A. Glycosyl composition assignment was based on the detected mass values for the  $[M+Na]^+$  molecular ions and subsequent MS<sup>n</sup> sequencing. Putative N-glycan structures were represented according to the

standard Symbol Nomenclature for Glycan system (Varki *et al.*, 2015). Complete list of identified *N*-glycans is provided in Table S2. The presence of an asterisk (\*) above an *m/z* signal indicates that no structure could be deduced.

Independently of the *N*-glycan types, 69 out of the 88 *N*-glycans isolated from *C. gigas* were substituted on their Man<sub>3</sub>GlcNAc<sub>2</sub> core either by a Xyl residue (37 *N*-glycans) on the internal Man residue, or by one Fuc residue (36 *N*-glycans), two Fuc residues (14 *N*-glycans), one Fuc and one Hex residue (5 *N*-glycans) or yet a disaccharide Hex<sub>1</sub>Fuc<sub>1</sub> (4 *N*-glycans) on the reducing GlcNAc residue. Substitution of the terminal reducing GlcNAc residue by a fucose residue was characterized by B-ions at *m/z* M+Na-451 resulting from the chitobiose fragmentation and resulting multiple fragmentations. Core α1-3Fuc and α1-6Fuc were differentiated according to their susceptibility to PNGase F and A. Some *N*-glycans showed a reducing GlcNAc simultaneously substituted by a Fuc and a Hex residues (CG10, 16, 22, 30 and 66), as characterized by Y-ion at *m/z* 678 and its corresponding B-ion (Fig. 4A). Substitution of the terminal reducing GlcNAc residue by two fucose residues (CG 6, 13, 17, 28, 31, 36, 43, 46, 51, 62, 71, 72, 73, 77) was also characterized by the Y-ion [Fuc<sub>2</sub>HexNAc] detected at *m/z* 648 and its corresponding B-ions at M+Na-625 (Fig 4B). Finally, three glycans (CG20, 32, 65 and 76) with a terminal Fuc<sub>2</sub>HexGlcNAc tetrasaccharide were identified based on the observation of B/Y fragment ions at *m/z* 895/852. The observation of an additional Z-ion at parent ion-206 strongly suggested that reducing GlcNAc was substituted by a Fuc residue on the C3 position and by a Fuc-Hex disaccharide on the C6 position, as exemplified for CG32 (Fig. 4C). Since CG 20, 65 and 76 were identified from the pool of *N*-glycans released by the PNGase F, it was concluded that this disaccharide was bound to the terminal reducing GlcNAc via an α1-6 linkage.



**Figure 4.** Repertory of MS<sup>2</sup> and MS<sup>3</sup> spectra of representative permethylated *N*-glycans. MS<sup>2</sup> spectra of *N*-glycans (A) CG29 and CG30, (B) CG36, (C) CG31 and CG32, (D) CG57 and CG58, (E) CG51, (F) CG84, (G) CG49 and (H) CG61 released from *C. gigas* by PNGase F. MS<sup>2</sup> spectra of *N*-glycans (J) OE34 and OE35 and (K) OE42 released by PNGase A from *O. edulis*. All glycans were permethylated prior their analysis. (I) MS<sup>3</sup> spectrum of the *m/z* 678 ion generated during MS<sup>2</sup> analysis of the *N*-glycan CG61. *N*-glycan structures were represented according to the standard Symbol Nomenclature for Glycan system (Varki *et al.*, 2015).

Paucimannose *N*-glycans corresponding to the minimal structure Man<sub>2-4</sub>-HexNAc<sub>2</sub> were differentiated by the number of their Man residues, the presence of Fuc (0, 1, 2) or Hex (0, 1) residues on the reducing GlcNAc

residue and Xyl on the internal Man residue. The presence and positioning of Fuc and Hex residues on the internal GlcNAc residue was easily assessed owing to MS<sup>2</sup> B-ion fragments resulting from the cleavage of the  $\beta$ 1-4 GlcNAc bond of the chitobiose as for oligomannosylated *N*-glycans. Depending on the presence of these residues on the Man<sub>2</sub>-GlcNAc<sub>1</sub> core fragmentation generated B-ions at *m/z* 690 for Man<sub>2</sub> (CG1, 2, 6, 10), *m/z* 850 for Man<sub>2</sub>Xyl (CG4, 5, 13,16), *m/z* 894 for Man<sub>3</sub> (CG3, 8, 9, 17, 20, 22 and 32), *m/z* 1054 for Man<sub>3</sub>Xyl (CG7, 14, 15, 28, 29 and 30), *m/z* 1098 for Man<sub>4</sub> (CG11, 19, 21 and 31) or *m/z* 1258 (CG18) Man<sub>4</sub>Xyl (Fig. 4A).

Out of the nine identified oligomannosidic type *N*-glycans, five signals at *m/z* 1580 (CG24), 1784 (CG38), 1988 (CG47), 2192 (CG52) and 2396 (CG63) were assigned to the so-called Man<sub>5</sub>–Man<sub>9</sub> high mannose (Man) type NGs. Signal at *m/z* 2600 (CG74) calculated as Hex<sub>10</sub>HexNAc<sub>2</sub> was also tentatively assigned to the biosynthetic intermediate Glc<sub>1</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> *N*-glycan. The MS/MS fragmentation patterns of these six compounds did not reveal any obvious deviation from the canonical structures and thus these six NGs will not be discussed further. Finally, signals at *m/z* 1754 (CG35), 1914 (CG42) and 2118 (CG50) were assigned to Fuc<sub>1</sub>Man<sub>5</sub>GlcNAc<sub>2</sub>, Fuc<sub>1</sub>Xyl<sub>1</sub>Man<sub>5</sub>GlcNAc<sub>2</sub> and Fuc<sub>1</sub>Man<sub>6</sub>GlcNAc<sub>2</sub> in which Fuc and Xyl are substituting internal GlcNAc and internal Man residues owing to core B-fragments at *m/z* 1303, 1463 and 1667, respectively (data not shown).

Two *N*-glycans were classified as hybrid structure (CG41 and 58) for which Man $\alpha$ 1-6 and Man $\alpha$ 1-3 of the core were substituted by at least 1 Man residue and 1 GlcNAc residue, respectively. Their MS/MS patterns were characterized by the release of the LacNAc antennae as a Y-ion at *m/z* 1362 identified as [HO-Hex<sub>4</sub>HexNAc<sub>2</sub>+Na]<sup>+</sup> (Fig. 4D). Single antennae complex-type *N*-glycans was the most abundant family in *C. gigas* with 34 identified *N*-glycans members (CG12, 23, 25-27, 33-34, 36-37, 39-40, 43-46, 48-49, 51, 53-57, 60-62, 64-66, 70-73 and 77). They were characterized by the absence of substitution on one of the two terminal Man residues of the core whereas the second one was substituted by at least one GlcNAc residue. The canonical glycosylation pathways suggests that the antennae are associated with the Man $\alpha$ 1-3 of the penta-saccharide core but their exact localization on either Man $\alpha$ 1-3 or Man $\alpha$ 1-6 was not confirmed [40]. The GlcNAc $\alpha$ 1-2 Man linkages were easily fragmented during MS<sup>2</sup> process to yield B- and Y-ions which corresponded to the terminal non-reducing fragment and the core fragment respectively. Y-ions generated from pseudohybrid *N*-glycans demonstrated the presence of a single OH-Man confirming that only one Man residue of the core was substituted. Primary Y-ions derived were detected at either *m/z* 1157 for unsubstituted core (CG 12, 27, 48 and 49), *m/z* 1317 for xylosylated core (CG23, 37, 53 and 56), *m/z* 1332 for mono-fucosylated core (CG25, 26, 39, 40, 54, 55, 57 and 61), *m/z* 1492 for xylosylated and mono-fucosylated core (CG33, 34, 44, 45, 60, 64 and 70), *m/z* 1506 for di-fucosylated (CG36, 46, 62 and 72) (Fig. 4B), *m/z* 1536 for Fuc and Hex containing core (CG65 and 66), and finally *m/z* 1666 for xylosylated and di-fucosylated core (CG43, 51, 71 and 73) (Fig 4E). When present, secondary fragments from Y-ions at *m/z* 880 or at *m/z* 1040 confirmed un-substituted (Hex<sub>3</sub>GlcNAc) (Fig. 4B) or Xylose substituted core (XylHex<sub>3</sub>GlcNAc) (Fig. 4E), respectively.

Seventeen biantennary complex type *N*-glycans were observed in *C. gigas* (CG59, 67, 68, 69, 75, 76, 78-88) and identified according to MS<sup>2</sup>-generated

B- and Y-ions generated by the preferential cleavage of GlcNAc $\beta$ 1-2Man linkage. As mentioned previously, when present, secondary Y-ions generated by the fragmentation of the chitobiose moiety generated signals at  $m/z$  1026 or  $m/z$  866 according to the presence or not of a Xyl residue on the internal Man residue (Fig. 4F). Accordingly, Y-ions were detected either at  $m/z$  1144 for unsubstituted core (CG59, 69, 79 and 80),  $m/z$  1304 for Xyl-core (GC67, 75, 81 and 83),  $m/z$  1318 for Fuc-core (CG68, 82, 85 and 87),  $m/z$  1478 for XylFuc-core (CG84, 86 and 88)(Fig. 4F),  $m/z$  1492 for Fuc<sub>2</sub>-core (CG78) or  $m/z$  1522 for FucHex-Core (CG76).

The hybrid and complex type *N*-glycans were all characterized by the presence of complex antennae made of different epitopes including disaccharide mono-LacNAc, tetrasaccharide di-LacNAc, tetrasaccharide A-group and pentasaccharide Leb/A-group epitopes that were identified by a combination of MS fragmentation, composition and linkage analyses. Twelve *N*-glycans (CG27, 37, 39-41, 44-46, 51, 69 and 75) were shown to be substituted by one, but never two, LacNAc epitope owing to the presence of a B-ion at  $m/z$  486 and of its complementary M-463 Y-ions (Fig. 4E). The systematic observation of single and multiple fragmentation Z-ions resulting from the cleavage between Gal-GlcNAc strongly suggested the presence of  $\beta$ 1-3 linkage and thus type-1 antennae. Type-1 LacNAc antennae were shown to be further substituted by an additional LacNAc repeat (CG49, 56, 57, 64, 80 and 83) as indicated by the presence of HexNAc<sub>2</sub>Hex<sub>2</sub> B-ion at  $m/z$  935 and HexNAc<sub>1</sub>Hex<sub>2</sub> C-ion at  $m/z$  708 on the corresponding spectra (Fig.4G). The absence of a Z-ion resulting from the cleavage between the terminal Gal residue and the GlcNAc strongly suggested that the additional LacNAc motif was of type-2 rather than type-1.

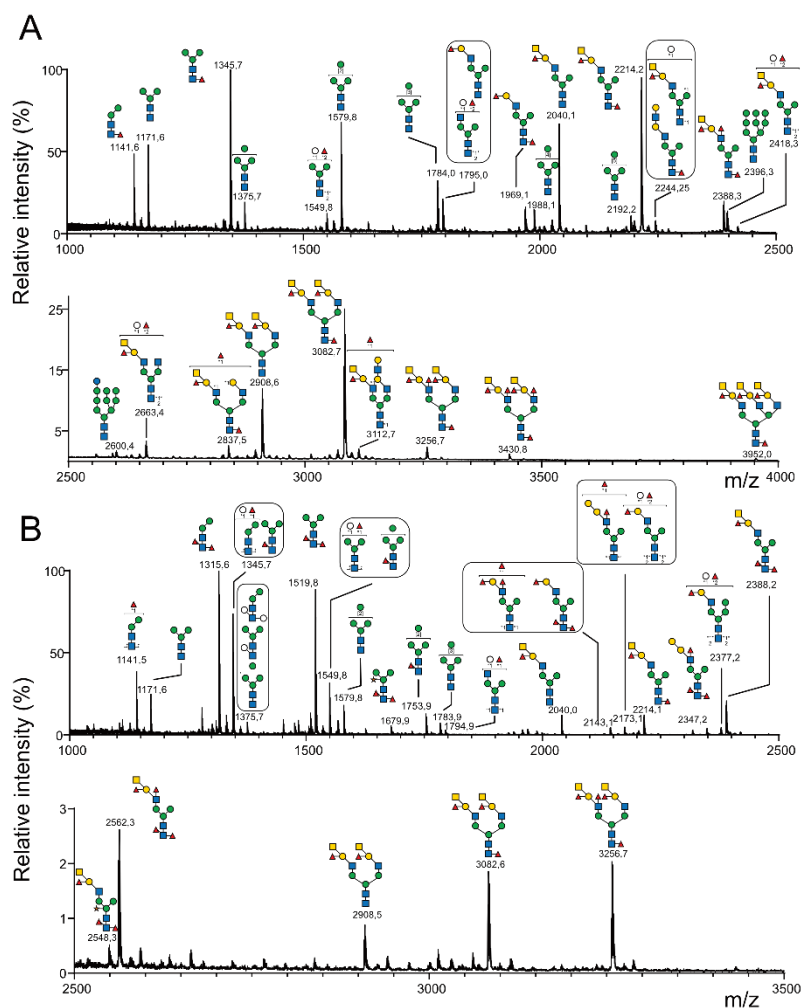
The vast majority of hybrid and complex-type *N*-glycans were substituted by a single (CG CG48, 53-55, 58-62, 64-66, 70-73, 75-78, 80 and 83) or two (CG79, 81-82 and 84-88) terminal HexNAc(deHex)Hex trisaccharides tentatively identified as type-1 A-group epitope GalNAc $\beta$ 1,3(Fuc $\alpha$ 1,2)Gal $\beta$  based on the observation of HexNAc<sub>2</sub>HexdeHex B-ion at  $m/z$  905 (Fig.4F) and HexNAcHexdeHex C-ion at  $m/z$  678 (Fig.4H), as well as multiple Y-ions and Z-ions generated from the loss of terminal tetra- and trisaccharides respectively. MS<sup>3</sup> analyses of the C-ion at  $m/z$  678 observed in all those *N*-glycans showed B-ions at  $m/z$  282, in agreement with the presence of the terminal GalNAc of A-group epitope. Furthermore, the observation of a <sup>0,2</sup>X-ion at  $m/z$  271 typified the internal Fuc(1,2)Gal motif, whereas the <sup>0,3</sup>X-ion signal at  $m/z$  301 typified the GalNAc(1,3)Gal motif of the A-group epitope confirming the attribution of the terminal trisaccharide as A-group epitope (Fig. 4I) in agreement with monosaccharides analysis.

The terminal tetrasaccharide GalNAc $\beta$ 1,3(Fuc $\alpha$ 1,2)Gal $\beta$ 1,3GlcNAc could be further substituted by a Fuc in  $\alpha$ 1-4 position of the GlcNAc residue typifying *N*-glycans elongated with one (CG61, 70, 72, and 77) or two Leb/A-group motifs (CG87 and 88) as established by the presence of a B-ion at  $m/z$  1079 and of a signal at  $m/z$  678 which was attributed to the GalNAc(Fuc)Gal C-ion (Fig.4H).

### 2.3.2. *N*-glycome of *O.edulis*

A similar analytical pipeline applied to *O. edulis* (Fig. 5) led to the identification by MS and MS/MS of 59 different *N*-glycans isolated from

gills, digestive tract and mantle, versus 88 in *C. gigas*. The decreased number of *N*-glycans in *O. edulis* compared to *C. gigas* was mainly due to the lower occurrence of xylose substituted *N*-glycans on internal Man (11 versus 37 *N*-glycans in *C. gigas*). Even when identified, the corresponding *N*-glycans were detected by MS as very low intensity signals, whereas these were observed as major signals in *C. gigas*. (Fig. 3A and 3B). The lower occurrence of Xyl substitution was further confirmed by GC/MS composition analysis, as shown later in the present study. Conversely, the occurrence of *N*-glycans with Hex core-substitution was higher in *O. edulis* compared to *C. gigas* (16 vs 9).



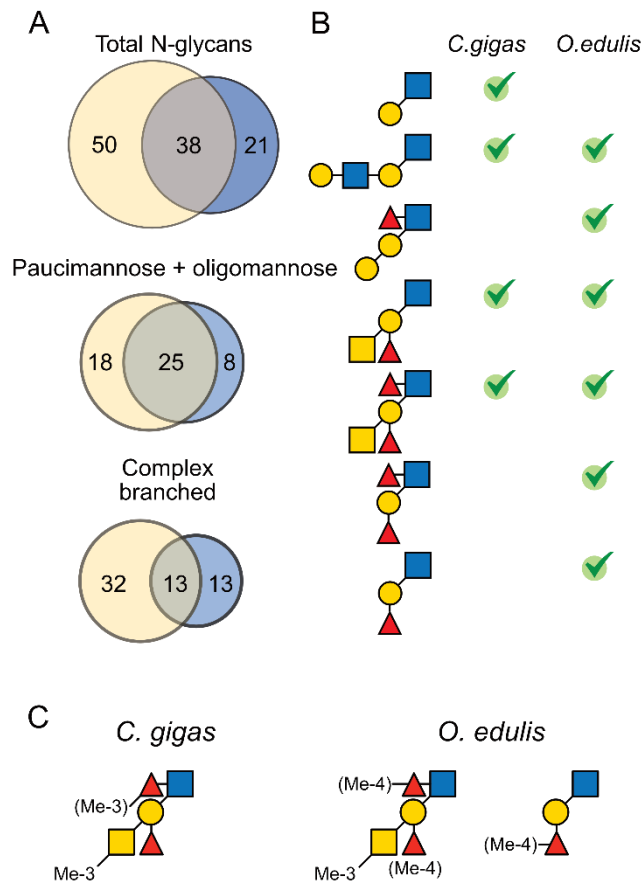
**Figure 5.** MALDI-MS analysis of permethylated *N*-glycans isolated from the mantle of *Ostrea edulis*. Prior their permethylation, *N*-glycans were sequentially released from mantle of *O. edulis* by (A) peptidyl-*N*-glycosidase F and (B) peptidyl-*N*-glycosidase A. Glycosyl composition assignment was based on the detected mass values for the  $[M+Na]^+$  molecular ions and subsequent MS<sup>n</sup> sequencing. Putative *N*-glycan structures were represented according to the standard Symbol Nomenclature for Glycan system (Varki *et al.*, 2015). Complete list of identified *N*-glycans is provided in Table S1. The presence of an asterisk (\*) above an *m/z* signal indicates that no structure could be deduced.

As for *C. gigas*, *N*-glycans were distributed in paucimannose (24), oligomannose (7), hybrid (1) and complex families (27) (Table S2),



according to the substitution pattern of the terminal Man residues of Man<sub>3</sub>GlcNAc<sub>2</sub> pentasaccharide core with a combination of single GlcNAc, and LacNAc-based complex antennae. According to MS/MS fragmentation analyses, the vast majority of LacNAc-substituted *N*-glycans comprised one and two antennae, except for a single tri-antennary complex *N*-glycan (OE59) observed at *m/z* 3952. Three *N*-glycans (OE40, 55 and 57) were substituted with LacNAc<sub>2</sub> repeat but the vast majority of LacNAc motifs were substituted by both Fuc and GalNAc to define A-group antigen (OE33, 39, 41, 44, 45, 47–49, 51–59) and for some of them an additional Fuc on the GlcNAc that defines an ALe<sup>b</sup> antigen (OE44, 49 56–58). In contrast to *C. gigas*, 7 *N*-glycans (OE29, 31, 34–36, 43, 52) were characterized by the presence of terminal Fuc $\alpha$ 1,2Gal1-3GlcNAc motif owing to intense Z-ion at M+Na-410 resulting from the fragmentation of the Gal $\beta$ 1-3GlcNAc linkage, coupled to B-fragment ion of Fuc $\alpha$ 1,2Gal1-3GlcNAc at *m/z* 660, as exemplified for OE35 (Fig. 4J), which strongly suggests the presence of terminal type-1 H-antigen in those *N*-glycans. In OE34 the observation of B-ion at *m/z* 834 rather than 660 the presence of a Le<sup>b</sup> antigen. In addition to the presence of H and Le<sup>b</sup> Ag, MS<sup>2</sup> analysis of *N*-glycans OE37 and OE42 revealed the presence of a terminal disaccharide GalGal though the C/Z cleavage which occurred between Gal and GlcNAc of the LacNAc antennae. The C/Z cleavage confirmed also the existence of the GalGal sequence at the terminal position moreover the B-ion at *m/z* 864 revealed the presence of a Fuc residue on the GlcNAc of the antennae (Fig. 4K for OE 42). Comparative analysis between organs established that all three tissues from *O. edulis*, gills, digestive tract and mantle, expressed essentially similar *N*-glycan profiles, although the digestive tract expressed a slightly lower structural diversity (Table 1).

Altogether comparison of the *N*-glycomes of both species revealed not only a large variability in terms of diversity, *C. gigas* *N*-glycome being much more diverse than that of *O. edulis*, but most importantly in terms of antigenic presentation. Out of the 109 different *N*-glycans identified by mass spectrometry analysis of permethylated derivatives, only 38 were commonly expressed in both species (Fig. 6A). However, most of these common *N*-glycans were part of oligomannosylated and paucimannosylated families. In contrast, *N*-glycans substituted by complex branches, exhibited very low similarities between species, with only 13 identified common *N*-glycans out of 65 (Fig. 6A). (Leb)/A-group appeared as the only common antigenic hallmarks in both species which otherwise differed widely in the expression of terminal epitopes: type-1 and type-2 LacNAc for *C. gigas* and Gal-Gal and (Leb)/H-group for *O. edulis* (Fig. 6B).



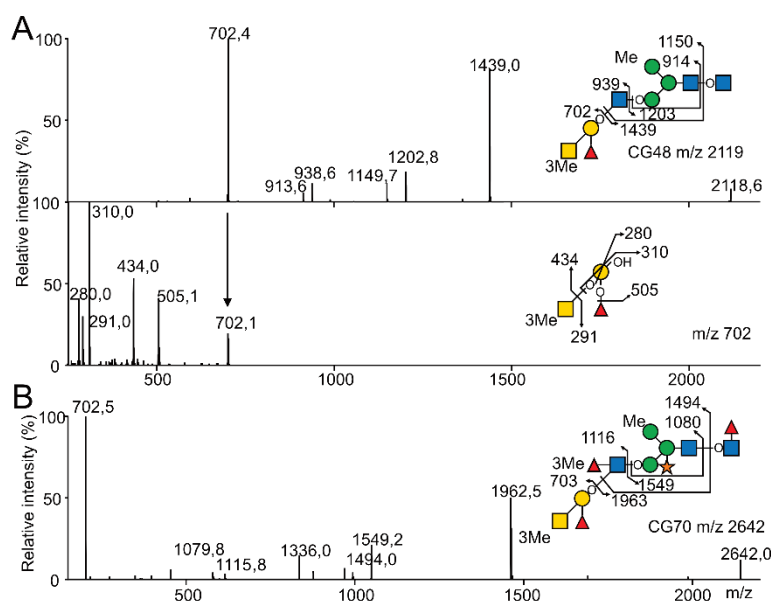
**Figure 6.** Distribution of common and specific terminal glycan epitopes in *C. gigas* and *O. edulis*. **(A)** Venn diagram showing the distribution of common epitopes in *C. gigas* (in yellow) vs *O. edulis* (in blue). The numbers indicate the number of glycans identified in each species and the sizes of the circles and intersections are proportional to the number of *N*-glycans. Diagrams were generated by BioVenn (Hulsen *et al.*, 2008). Top diagram cumulates all identified *N*-glycans; other diagrams correspond to individual family of *N*-glycans as depicted. **(B)** Distribution of the different oligosaccharide motifs (LacNAc, LacNAc repeat, GalGal[Fuc]GlcNAc-, A Ag, ALe<sup>b</sup> Ag, Le<sup>b</sup> and H-type 1 Ag from top to bottom) identified on the pentasaccharide core in the two oyster species. Ticks indicate the presence of the motif in each species. **(C)** Summary of the methylation status of terminal HBGA in *C. gigas* and *O. edulis*.

## 2.4. Methyl *N*-glycome analysis

### 2.4.1. Localization of methyl groups on HBGA

Based on sequence and composition analyses, the methylated monosaccharides Me-3/4O-Fuc, Me-3/4O-Man and Me-O3-GalNAc were further localized on the *N*-glycans isolated from the different samples by mass spectrometry following per-deuteromethylation thanks to the 3 m.u. difference that it generates compared to its predicted per-deuteromethylated form. Consequently, mono, di, tri and tetra methylated *N*-glycans isomers were differentiated owing to the observation of -3, -6, -9 and -12 m.u. deviations from MS analyses and the positions of methylated monosaccharides within individual glycans were established by MS<sup>2</sup> and MS<sup>3</sup> fragmentation (Table S3). In total, MS analyses of per-deuteromethylated *N*-glycans allowed the identification

of 104 compounds substituted by 0 to 4 methyl groups in *C. gigas* and 49 in *O. edulis*, all derived from the basic structures identified after permethylation. It should be noted that all the *N*-glycans identified as permethylated derivatives could not be retrieved as perdeuteromethylated derivatives due to lower sensitivity, either due to lower efficiency of substitution by I-CD<sub>3</sub> and/or the isomeric dispersion for individual *N*-glycans caused by multiple methylation states. Perdeuteromethylated *N*-glycans were further analyzed by MS<sup>2</sup> and MS<sup>3</sup> in order to localize individual methyl groups on previously identified epitopes. Irrespective of the origin, Me-3/4O-Man residues were systematically localized on the Man<sub>3</sub>GlcNAc<sub>2</sub> core of *N*-glycans from oligomannosylated, paucimannosylated and (pseudo)hybrid families owing to intense Y- and Y/B-fragmentation ions generated by the cleavage of GlcNAc linkages which values depend on its substitution by Xyl and Fuc residues, as exemplified by the observation of Y-ion at *m/z* 1203 and B/Y secondary ion at *m/z* 914 (Fig. 7A) or Y-ion at *m/z* 1549 and B/Y ion at *m/z* 1079 (Fig. 7B). In accordance with the literature that reports the presence of methylated mannose on the Man<sub>3</sub>- and/or Man<sub>6</sub> of the trimannosylated core in numerous invertebrates, we postulated that the latter methyl group was localized on the terminal Man residue [42].

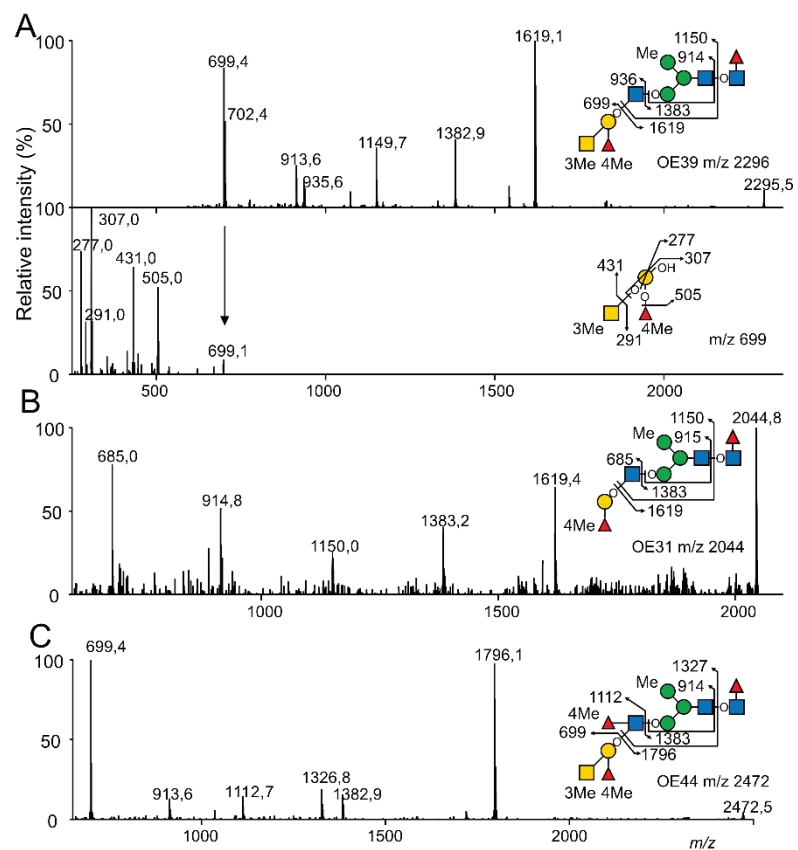


**Figure 7.** MALDI MS sequencing of the partially methylated and deuteromethylated *N*-glycans from *C. gigas* for the precise location of methyl group. (A) MALDI-MS<sup>2</sup> and MS<sup>3</sup> spectrum of *N*-glycan CG48 and (B) MALDI-MS<sup>2</sup> of *N*-glycan CG70. *N*-glycans were released by peptidyl-*N*-glycosidase F from *C. gigas* and permethylated with I-CD<sub>3</sub> prior their analysis. For CG48, MS<sup>2</sup> signal at *m/z* 702 corresponding to mono-methylated terminal A-Ag was selected for MS<sup>3</sup> fragmentation.

Then, Me-3O-GalNAc and Me-3/4O-Fuc residues could be localized on the different HBGA-like terminal epitopes identified above in the two species of oysters (Fig. 7A). In *C. gigas*, as exemplified by MS<sup>2</sup> analysis of the di-methylated *N*-glycan CG48 at *m/z* 2118, one methyl group was localized on the terminal A-group epitope owing to the intense C/Z couple ions at *m/z* 702/1439 whereas another one was localized on the

trimannosylated (Fig.7A). MS<sup>3</sup> analysis of the C-ions at *m/z* 702 demonstrated that the methyl group on the A-antigen was exclusively substituting the terminal GalNAc monosaccharide owing to the observation of the C/Z couple ions at *m/z* 291/434, and in agreement with the absence of CH<sub>3</sub>- group on Fuc residue, as demonstrated by the Y ion at *m/z* 505 and X ions at *m/z* 280 and 310 (Fig.7A). Similar analysis on trimethylated isomer of CG70 at *m/z* 2642 showed the presence of two methyl groups on the terminal Leb A-group pentasaccharide owing to the B/Y couple ions at *m/z* 1116/1549, out of which one was substituting the terminal A-group trisaccharide owing to the intense C/Z couple ions at *m/z* 702/1962 (Fig. 7B). MS<sup>3</sup> analysis of the C-ions at *m/z* 702 also demonstrated the presence of a terminal Me-O-GalNAc residue, as in CG48 (data not shown). Considering that the composition analysis did not show the presence of Me-O-GlcNAc, the second methyl group on the antenna was localized on the Fucα1,4 residue of the Leb epitope. Coupled with monosaccharide composition analysis (shown in Figure 2) with systematic MS<sup>2</sup> and MS<sup>3</sup> analyses on per-deuteromethylated *N*-glycans indicated that HBGA of complex *N*-glycans isolated from *C. gigas* were partially methylated on the C3 position of terminal GalNAc residue of the A-group antigen and on the C3 position Fucα1-4 residue of the Leb antigen to generate a mixture of unmethylated, mono- and di-methylated A- and ALeb-antigens as depicted on Figure 6C.

A similar analytical approach on *N*-glycans from *O.edulis* showed that complex *N*-glycans from this species could be similarly substituted by a methyl group on terminal GalNAc residue of A-group Ag as well on Fuc residue of the disaccharide Fucα1-4GlcNAc of the A-Leb epitope. In addition, MS<sup>2</sup> and MS<sup>3</sup> analysis of multimethylated complex *N*-glycans showed that Fucα1,2 that substitutes the Gal residue of the A antigen could also be methylated, contrarily to *C. gigas* where it is not. As exemplified by the fragmentation of di-methylated OE39 at *m/z* 2296, the MS<sup>2</sup> C/Z couple ions at *m/z* 699/1619 demonstrated the presence of two CH<sub>3</sub>- groups on the terminal GalNAc<sub>1</sub>Fuc<sub>1</sub>Gal<sub>1</sub> trisaccharide that were localized on terminal GalNAc and H-type Fuc residues by MS<sup>3</sup> analysis of C ion at *m/z* 699 owing to Y-ions at *m/z* 505 and 431 (Fig. 8A). The GalNAc residue as well as the two Fuc residues could also be simultaneously methylated to form a trimethylated ALeb epitope, as observed for the tetra-methylated OE44 (Fig. 8C). Finally, MS<sup>2</sup> fragmentation analysis of H-antigen containing *N*-glycans confirmed that the terminal Fucα1,2 residue of the H-disaccharide could also be methylated, owing to the observation of B/Y couple ions at *m/z* 685/1383 and the Z ion at *m/z* 1619 (Fig 8B). Altogether, analyses showed that *N*-glycans from *O. edulis* were substituted by a combination of A- and ALeb-antigens methylated on C3 position of terminal GalNAc residue and on C4 positions of H- and Lewis Fuc residues, as well as on C4 positions of Fuc residue of terminal H-antigen as depicted on Figure 6C.

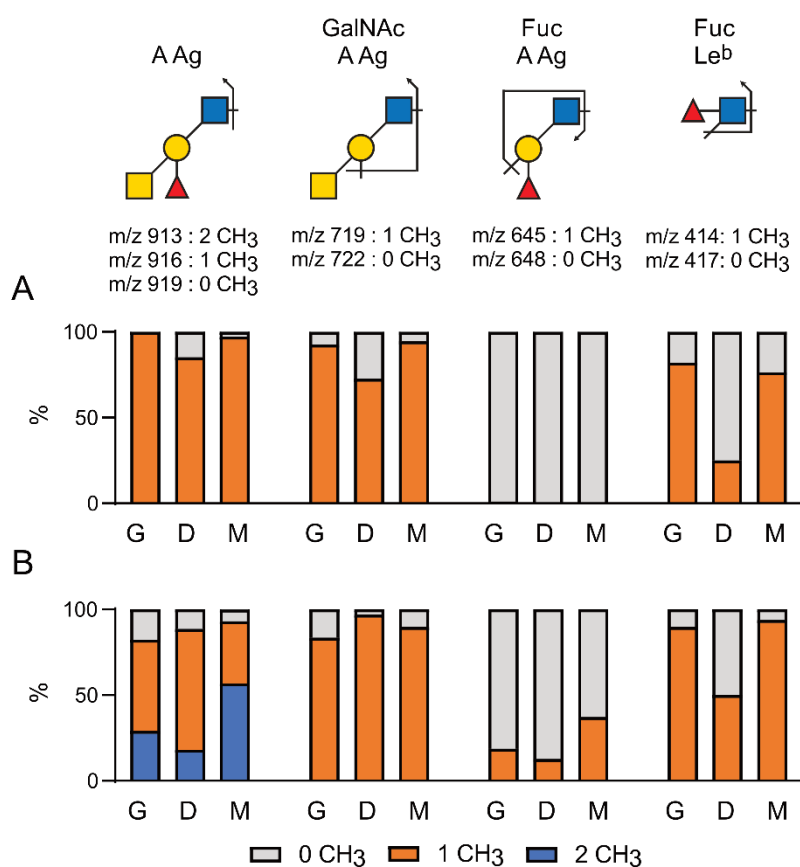


**Figure 8.** MALDI MS sequencing of the partially methylated and deuteromethylated *N*-glycans from *O. edulis* for the precise location of methyl group. (A) MALDI-MS<sup>2</sup> and MS<sup>3</sup> spectrum of *N*-glycan OE44 and (B) MALDI-MS<sup>2</sup> of *N*-glycan OE31. *N*-glycans were released by peptidyl-*N*-glycosidase F from *O. edulis* and permethylated with I-CD<sub>3</sub> prior their analysis. For OE44, MS<sup>2</sup> signal at *m/z* 702 corresponding to mono-methylated terminal A-Ag was selected for MS<sup>3</sup> fragmentation.

#### 2.4.2. Relative quantification of methylated antigens

To evaluate the proportion of differentially methylated A and ALe<sup>b</sup> antigens in the gills, mantle, and digestive tract of *C. gigas* and *O. edulis*, deuteromethylated *N*-glycans were further subjected to comprehensive data-dependent HCD-MS<sup>2</sup> acquisition by reverse phase C18 nano-LC-MS<sup>2</sup> analysis [43]. Under the acidic solvent conditions used, the parent glycans were mostly protonated, which favored primary MS<sup>2</sup> cleavages at HexNAc to produce nonreducing terminal oxonium ions, often accompanied by loss of terminal residue via secondary glycosidic cleavages. Each of the distinctive MS<sup>2</sup> fragment ions produced and measured at high mass accuracy (< 5 ppm) would identify a specific terminal glycotope or its substructure. The summed intensity of each of these target ions extracted from all the MS<sup>2</sup> spectra acquired across the entire LC range where the glycans eluted is indicative of their respective glycomic abundance in comparative analyses. This allowed the relative abundance of the A and ALe<sup>b</sup> antigens differing only in their methylation status and distributed over a diverse range of parent *N*-glycans to be relatively quantified. As shown in Fig. 9, glycans carrying differentially methylated A antigen would afford the characteristic MS<sup>2</sup> ions at *m/z* 919

(no CH<sub>3</sub>), 916 (1 CH<sub>3</sub>) and 913 (2 CH<sub>3</sub>). Additional loss of Fuc produced the ions at *m/z* 722 (no CH<sub>3</sub>) and 719 (1 CH<sub>3</sub>), whereas loss of the terminal GalNAc yielded the ions at *m/z* 648 (no CH<sub>3</sub>) and 645 (1 CH<sub>3</sub>). The overall pattern is consistent with both the terminal GalNAc and Fuc of A antigen could be methylated but more completely so for the GalNAc. Similarly, glycans carrying ALe<sup>b</sup> antigen with up to 3 methylations would afford the ions at *m/z* 1099 (no CH<sub>3</sub>), 1096 (1 CH<sub>3</sub>), 1093 (2 CH<sub>3</sub>) and 1090 (3 CH<sub>3</sub>), while further loss of the mostly methylated terminal GalNAc produced the ions at *m/z* 828 (no CH<sub>3</sub>), 825 (1 CH<sub>3</sub>) and 822 (2 CH<sub>3</sub>). The methylation status of the extra Fuc on the GlcNAc of ALe<sup>b</sup> could be relatively quantified by the ions at *m/z* 417 (no CH<sub>3</sub>) and 414 (1 CH<sub>3</sub>) for the internal Fuc $\alpha$ 1,4GlcNAc disaccharide moiety.



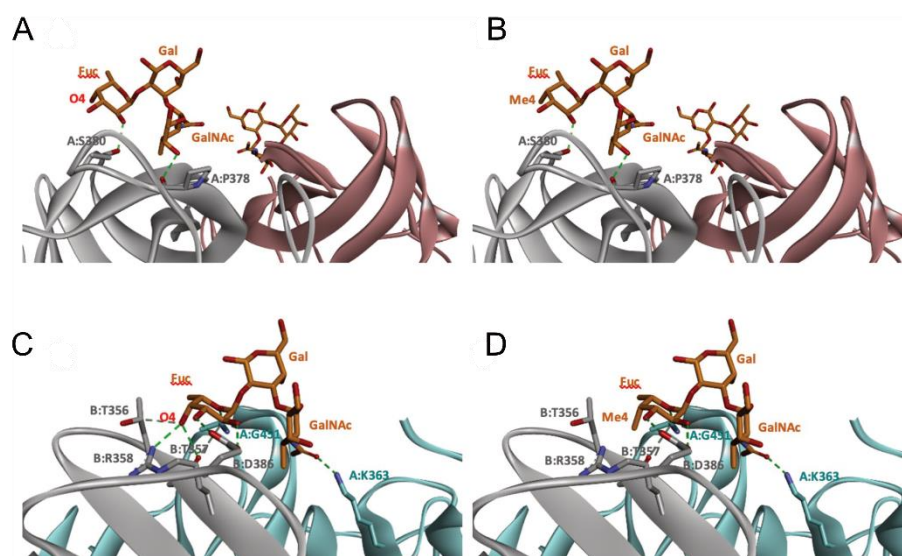
**Figure 9.** Relative quantification of native and methylated form of A Ag and Leb Ag. Deuteromethylated *N*-glycans of *C. gigas* and *O. edulis* were analyzed by LC-MS<sup>2</sup>. Characteristics MS<sup>2</sup> ions for individual antigens and their respective methylated forms were identified as depicted on top. Proportions of methylated individual glycan epitopes were quantified in *N*-glycans isolated from the three tissues of (A) *C. gigas* and (B) *O. edulis*. Light gray, 0 CH<sub>3</sub> group; orange, 1 CH<sub>3</sub> group; blue, 2 CH<sub>3</sub> groups. G, gills; D, digestive tract; M, mantle.

Identification and relative quantification of each of the target MS<sup>2</sup> ions in the PNGase F-released *N*-glycans confirmed that the A antigen from both *C. gigas* and *O. edulis* was almost completely methylated at the terminal GalNAc, whereas the Fuc $\alpha$ 1,2 was methylated at a significant level only in *O. edulis*, as shown by the detection of di-methylated A antigen (*m/z* 913) and mono-methylated internal Fuc $\alpha$ 1,2Gal1,3 $\beta$ GlcNAc (*m/z* 645)

fragments only in *O. edulis*. Relative quantification of these ions further indicated that the A antigen was almost completely mono-methylated in the three organs of *C. gigas*, although with less than 10% unmethylated form in the digestive tissue. (Fig 9A). Analysis of *O. edulis* tissues, on the other hand, showed that the terminal A antigen in all three tissues was more than 80% substituted by at least one methyl group while a significant proportion was di-methylated (Fig. 9B). The MS<sup>2</sup> ions at *m/z* 722/719 and *m/z* 648/645 further indicated that more methylated Fuc $\alpha$ 1,2 was found in the mantle of *O. edulis* than in the other two tissues, consistent with a much higher proportion of di-methylated A antigen in this tissue (>50%) compared to the other two (<25%) (Fig 9B). For ALe<sup>b</sup> antigen, the ion intensity ratio of signals at *m/z* 414 vs 417 showed that the extra Fuc $\alpha$ 1,4 residues could be methylated in both species, with much higher level in gills and mantle relative to digestive tissue. Due to this extra methyl, the ALe<sup>b</sup> antigen can carry up to 3 methylation, particularly in *O. edulis*, since it also has a significantly higher level of methylation on the other Fuc compared to *C. gigas*. However, due to the preferential MS<sup>2</sup> cleavage of the terminal GalNAc residue, the overall signal intensities of the signals associated with the methylated forms of the ALe<sup>b</sup> antigen were very low and could not be reliably quantified.

### 2.5. Molecular modeling of Norovirus P domain interaction with methylated A antigens

In order to determine the potential influence of methyl groups decorating the terminal A antigen trisaccharide on the interactions of oyster saccharides with human noroviruses from different genogroups, the interaction maps of a GI.1 (Norwalk virus) and GII.3 (TV24) with non-methylated Human A-antigen were compared with those of mono- and di-methylated epitopes identified in *C. gigas* and *O. edulis*. To do that, the X-ray crystallographic structures of the P domain of GI-1 (PDB code 3D26) and GII.3 (PDB code 6IS5) viruses in complex with A-trisaccharide or A-tetrasaccharide respectively were used as patterns. The glucose of the A-tetrasaccharide was removed and hydroxyls at position 3 of GalNAc and 4 of Fuc in the trisaccharides were substituted by a methyl group to model saccharides exposed on the surface by oyster species *Crassostrea gigas* and *Ostrea edulis*. Non-bond interactions between the atoms of viruses and those of saccharides were monitored using BIOVIA Discovery Studio visualizer (Dassault Systèmes, San Diego). First, the recognition between GI-1 and oyster epitopes appeared to differ by the loss of two hydrogen bonds (HB) between the oxygen (O) at position 3 of GalNAc and the residues Asp327 and Ser377 of GI-1 P domain due to the methylation of this position in both *C. gigas* and *O. edulis*, thus potentially weakening the interaction (Fig. 10).



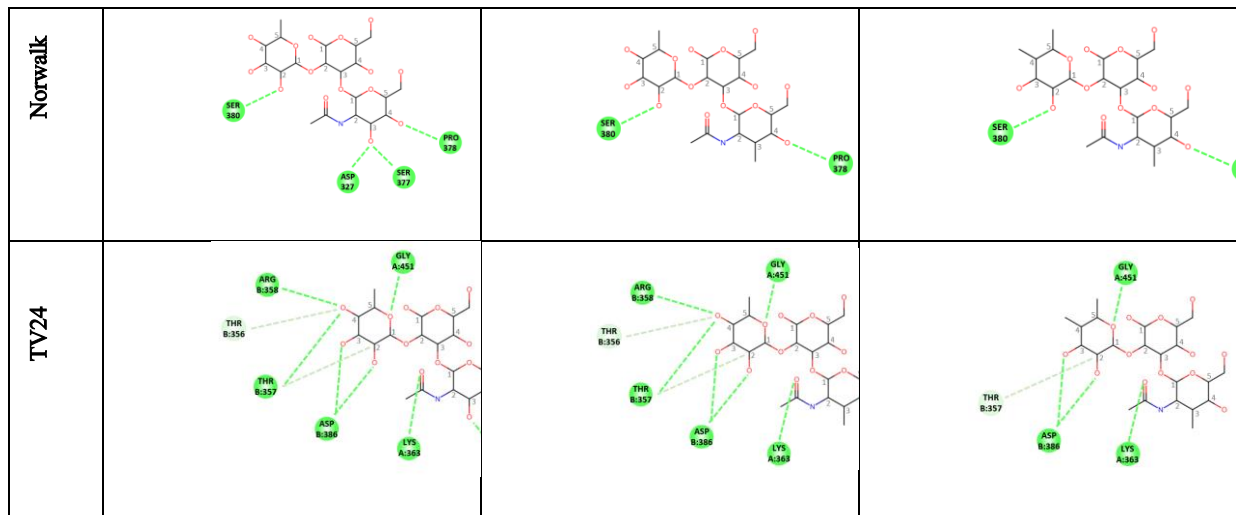
**Figure 10.** Structural models of GI.1 (Norwalk) and GII.3 (TV24) P dimers in complex with oyster trisaccharides. Structure of Norwalk (A and B) and TV24 (C and D) P dimer in complex with Me3GalNAc $\alpha$ 1,3(Fuc $\alpha$ 1,2)Gal $\beta$  (A and C) and Me3GalNAc $\alpha$ 1,3(Me4Fuc $\alpha$ 1,2)Gal $\beta$  (B and D). HB and CHB between atoms of virus P domain and atoms of saccharides are shown in green and pale green dashed lines respectively.

However, GI.1 strain seems to recognize in a similar manner both saccharides Me3GalNAc $\alpha$ 1,3(Fuc $\alpha$ 1,2)Gal $\beta$  and Me3GalNAc $\alpha$ 1,3(Me4Fuc $\alpha$ 1,2)Gal $\beta$ . The oxygen (O) at position 4 of GalNAc makes hydrogen bond (HB) with Pro378 whereas O2 of Fuc makes HB with Ser380 of GI-1 (Fig. 10 A). In comparison, recognition by GII.3 seems stronger as it involves a larger number of HB (Fig. 10 B). The Fuc residue makes 5 HB with Gly451, Asp386, Thr357 and Arg358 and 2 carbon HB (CHB) with Thr356 and Thr357 of TV24 virus and the GalNAc residue makes 1 HB with Lys363 of TV24 when only GalNAc is methylated at position 3. The methylation of Fuc at position 2 causes the loss of 2 HB and 1 CHB. These observations suggest that the attachment of the GII.3 strain to glycans of the *Crassostrea gigas* oyster species is stronger than that of the GI.1 strain. However, attachment of the latter strain should be similar between the two oyster species, whilst binding of the GII.3 virus by *Ostrea edulis* may certainly be weaker than that by *Crassostrea gigas* depending on the quantity and proportion of Me3GalNAc $\alpha$ 1,3(Fuc $\alpha$ 1,2)Gal $\beta$  and Me3GalNAc $\alpha$ 1,3(Me4Fuc $\alpha$ 1,2)Gal $\beta$  in *Ostrea edulis*.

**Table 1.** 2D diagram of the detailed HB (green dashed lines) and CHB (pale green dashed lines) network between amino acids of the P domain of viruses and individual saccharides of the A-trisaccharide.

irus	Trisaccharides		
	$\alpha$ -L-Fuc-(1 $\rightarrow$ 2)-[ $\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)]- $\beta$ -D-Gal	$\alpha$ -L-Fuc-(1 $\rightarrow$ 2)-[ $\alpha$ -D-Me3GalNAc-(1 $\rightarrow$ 3)]- $\beta$ -D-Gal	$\alpha$ -L-Me4-Fuc-(1 $\rightarrow$ 2)-[ $\alpha$ -D-Me3GalNAc-(1 $\rightarrow$ 3)]- $\beta$ -D-Gal





### 3. Discussion

Although belonging to the same Ostreoidae superfamily, the two oyster species *Crassostrea gigas* and *Ostrea edulis* showed very different *N*-glycan profiles in terms of variability, modification of the ManGlcNAc<sub>2</sub> core, expression of terminal antigens and methylation patterns. Of the 109 unique *N*-glycans identified, 88 in *C. gigas* and 59 in *O. edulis*, excluding the variability generated by methylation modification, only 38 were common to both species (Fig.6). Most of these common glycans are oligomannosylated and pauci-mannosylated *N*-glycans, leaving only 13 that present complex type branches commonly expressed in both species. These *N*-glycans exhibit many structural features of Man<sub>3</sub>GlcNAc<sub>2</sub> core that have been previously described in invertebrates [29]. Among those, complex type *N*-glycans showed a high proportion of core fucosylation, either in  $\alpha$ 1-6 linkage or  $\alpha$ 1-3 linkage, out of which some simultaneously present both  $\alpha$ 1-6 and  $\alpha$ 1-3 fucose residues (about 12 % in *C. gigas* and 20% in *O. edulis*) already described in invertebrates [44]. Fuc $\alpha$ 1-3 has been widely detected outside Human's glycoproteins in plants [45], insects [46] and nematode [47] and shown to be transferred by  $\alpha$ -1,3-fucosyltransferase (FUT-1) [48]. Based on data collected in *Caenorhabditis elegans* and *Drosophila melanogaster*, it was suggested that  $\alpha$ 1,3 and  $\alpha$ 1,6 core fucoses are sequentially added to chitobiose core by FUT1 and  $\alpha$ -1,6-fucosyltransferase (FUT-6) [48,49]. Among invertebrate-specific core modifications, we also observed the presence of a disaccharide Fuc-Hex on the reducing terminal GlcNAc residue. Here, MALDI-MS/MS did not allow us to discriminate between Fuc-Hex or Hex-Fuc disaccharides, nor to identify the nature of hexose as in Gal-Fuc motif previously reported to results from the galactosylation of the core  $\alpha$ 1,6-fucose by GALT-1 galactosyltransferase [29,50]. This unusual galactosylation has also been reported for  $\alpha$ 1,3Fuc linked to either the reducing terminal GlcNAc and/or the non-reducing terminal GlcNAc of the core [47]. We also observed the presence of a single Hex residue on the terminal reducing GlcNAc of numerous *N*-glycans in both species (26% of *N*-glycans in *O. edulis* and 11% in *Crassostrea gigas*). A similar modification was observed in *Volvarina rubella* a margin snail of the clade Neogastropoda [51] and in the oyster *Crassostrea virginica* [30] and was described as a Man residue based on its sensitivity to  $\beta$ -mannosidase [30,51]. When present, this Hex was

observed either alone or in combination with a single Fuc residue or more rarely associated with a second Hex on the terminal reducing GlcNAc in *Ostrea edulis*. The nature of this core modification was not further inquired as the main object of the present report was the terminal complex epitopes that may serve as ligands for virus. Finally, xylosylation of internal Man residue was shown to be a common feature of *N*-glycans isolated from *C. gigas* and *O. edulis*, although in various amount in each species (about 42% of *N*-glycans are substituted by Xyl residues in *C. gigas* and 19% in *O. edulis*), as observed in numerous species of parasites and gasteropodes [52–54]. Noticeably, no Xyl residue was found on *N*-glycans isolated from the pacific oyster *Crassostrea virginica*, which strongly suggests that this modification is highly species specific [30].

More than half of the complex type *N*-glycans (51% in *C. gigas* and 62% in *O. edulis*) were substituted by a terminal A-antigen motif, associated or not with an internal Leb motif to form the ALeb motif, and a minority by a terminal H-antigen in *O. edulis* (14%), but none in *C. gigas*. All GalNAc residues identified in *N*-glycans of both *C. gigas* and *O. edulis*, irrespective of the tissue, are part of either A- and ALeb-Ag, suggesting that the amount of GalNAc residues is a direct reflection of the amount of terminal A-antigen present in the tissues of the two oyster species. With the exception of a single *N*-glycan, the same conclusion can be drawn in *C. virginica* [30]. Quantification of GalNAc in *N*-glycans (Table S1) demonstrates that tissues from *O. edulis* express from two (in digestive track) to five times (in the gills and mantle) more A- and ALeb-Ag than tissues from *C. gigas*, which results in an increased exposure of potential glycan-ligands of noroviruses on *O. edulis* compared to *C. gigas*.

The two species also differ in their methylation patterns, as demonstrated by the detailed localization of methyl groups on monosaccharides on individual *N*-glycans, in a tissue-specific manner. In both species, most of the GalNAc residues were shown to be methylated at their C3 position, although in different proportions according to the tissues : about 55% in gill (CG 54% and OE 59%), 66% in digestive track (67% in CG and 65% in OE) and 78% in mantle (79% in CG and 76% in OE). In comparison, about 37% of the GalNAc was reported to be methylated in the Pacific oyster *C. virginica* [30]. About 25% of the fucose was also shown to be methylated in *C. gigas* and *O. edulis* (19% in CG and 29% in OE) whereas no methylated fucose was observed in *C. virginica* [30]. However, in contrast to GalNAc which was methylated at the C3 position in both species, Fuc residues were methylated in a species-specific manner at the C3 and C4 positions in *C. gigas* and *O. edulis* respectively, demonstrating the existence of an exquisitely specific methylation process. It should be noted that both 3-*O*-Me-GalNAc and 3-*O*-Me-Fuc have been previously identified in glycolipids isolated from the freshwater mussel *Hyriopsis schlegelii*, but to our knowledge the presence of 4-*O*-Me-Fuc has never been reported in invertebrates [55]. Similarly to GalNAc and Fuc, about 25% of the Man residues in *N*-glycans isolated from the three tissues of the two species were methylated. However, 3-*O*-Me-Man and 4-*O*-Me-Man, were simultaneously identified in both species but in very different ratios according to the tissues (Fig. 2).

Surprisingly, our analyses did not reveal any sulfation on the *N*-glycans in *C. gigas*, in contrast to the Pacific oyster the glycans of which were shown to be heavily sulfated on galactose residues, either in terminal position or within A antigenic motif [30]. Sulfation is indeed a ubiquitous and well known modification affecting various monosaccharides such as

the galactose or glucuronic acid residues in humans or sialic acid in sea urchins [56,57]. Sulfated *N*-glycans are however notoriously difficult to detect by mass spectrometry-based glycomic analysis due to ion suppression, which occurs regardless of the positive or negative mode chosen. Furthermore, during the classical methylation protocol, permethylated sulfated *N*-glycans are totally or partially lost during the chloroform/water cleaning procedure due to their solubility in water. To overcome this problem, solid phase extraction using a weak anion exchange cartridge was used to recover putative sulfated *N*-glycans, as previously demonstrated [58,59]. Several attempts using this optimized procedure coupled to MALDI-MS analysis operating in negative and positive modes failed to identify any sulfated compound, strongly suggesting that *N*-glycans of *C. gigas* are devoid of sulfate substitution, in contrast to *Crassostrea virginica*.

Although the two oyster species studied here present HBGA motifs, these were shown to be, for a large part, modified by methyl groups in contrast with human HBGAs. These modifications may affect the binding and bioaccumulation of norovirus strains. To begin exploring this possibility, we modeled the interaction of two viral strains with different versions of A-group epitope. The Norwalk and TV24 strains were used for the analysis because their binding modes to HBGAs are representative of the GI, and GII binding modes of the majority of circulating human strains [60–62]. The modeling results should be considered with caution since they do not take into account dynamic aspects of the virus-glycan interactions nor the effect of glycans clustering. They nonetheless show relevant potential differences between the GI and GII attachment to the two oyster species. Because the GII.3 strain interaction with *Crassostrea gigas* involves more HB than that of the GI.1 strain, it suggests a higher binding of GII.3 strains to tissues of that oyster species. This might explain the previously reported binding and bioaccumulation of GII to a larger set of oyster tissues [19,25,26,28,63]. Interestingly, interaction of the GI strain does not appear to be different between the two oyster species *Crassostrea gigas* and *Ostrea edulis*, whilst that of the GII strain is expected to be much lower in the latter species, according to the amount and tissue distribution of methylated  $\alpha$ 1,2-linked fucose. These observations suggest that the attachment of norovirus to oyster tissues may not only be virus strain-dependent, but also dependent on species, *Ostrea edulis* being possibly less prone to bioaccumulate epidemiologically dominant viral strains (personal observation). These hypotheses should be experimentally demonstrated, ideally by the use of saturation transfer difference (STD) NMR spectroscopy that provides a direct mapping of the glycan epitope recognized by VLP. This approach allowed us to precisely map the glycan epitopes recognized by norovirus epitopes out of synthetic type 1, type 2, type 3, type 5, and type 6 blood group A- and B-tetrasaccharides without relying on X-ray crystallography analysis of carbohydrate-protein complexes could be directly applied to differentially methylated A-antigens either from natural source or generated by chemical synthesis [64,65].

#### 4. Materials and Methods

##### Preparation of anatomic pieces

*Crassostrea gigas* and *Ostrea edulis* adult oysters of commercial size were bought live in November 2018 from a local shellfish producer. Upon arrival at the laboratory, oysters were shucked, their body removed from

the shell and dissected on ice using sterile scalpels and pliers. Three tissues (the gills, the mantle, and the digestive tract) were collected from 10 to 15 individuals, pooled per tissue type, and stored at  $-80^{\circ}\text{C}$ . Before use, tissues were homogenized for 2 min in water using gentleMACS™ Dissociator (Miltenyi Biotec, France), freeze dried and stored at  $-20^{\circ}\text{C}$  until further use.

### Preparation of N-glycans

Digestive tract, gills or mantle (100mg) were suspended in 1mL of PBS containing 1% of triton X-100 and sonicated for 2 min in an ice bath. The mixture was stirred at  $37^{\circ}\text{C}$  overnight. After centrifugation at 13000rpm for 10min at room temperature, proteins and glycoproteins contained in supernatants were reduced by 100 $\mu\text{L}$  of a 0.1M DTT in PBS solution at  $37^{\circ}\text{C}$  for 1h under stirring and then alkylated by adding 110 $\mu\text{L}$  of a 0,5M iodoacetamide in PBS solution and incubate at  $37^{\circ}\text{C}$  for 1h under stirring. Proteins and glycoproteins were precipitated by adding 130 $\mu\text{L}$  of a saturated trichloroacetic acid solution and placed at  $-20^{\circ}\text{C}$  for 30min. After centrifugation at 13000rpm for 10 min at  $4^{\circ}\text{C}$ , the pellet was washed 3 times by 1mL of cold acetone ( $-20^{\circ}\text{C}$ ). After drying with a vacuum Speed Vac (5min.), 500 $\mu\text{L}$  of a 50mM  $\text{NH}_4\text{HCO}_3$  solution were added followed by 40 $\mu\text{L}$  of a 2mgmL<sup>-1</sup> trypsin in a 50mM  $\text{NH}_4\text{HCO}_3$  solution and incubate at  $37^{\circ}\text{C}$  for overnight. Contaminating glycogen fragments were removed on a C18 SPE cartridge equilibrated in a 5% acetic acid solution (10mL). Peptides and glycopeptides were recovered by passing successively 10mL of a 20%, 40% and 60% 2-propanol in 5% acetic acid. After freeze drying, they were resuspended in 500 $\mu\text{L}$  of a 50mM solution and 1 $\mu\text{L}$  of a PNGase F (62,5 IU) was added to release N-glycans which were recovered in the unretained fraction of a C18 SPE cartridge operated as previously. Peptides and insensitive PNGase F glycopeptides were recovered by pooling the 3 different isopropanol fractions (20, 40 and 60%) and freeze dried before digestion by PNGase A according to the manufacturer recommendations.

### Protein assay

Protein concentrations were measured in triplicate by Micro BCA™ Kit (Thermo scientific) on 96-well microplate following the manufacturer recommendations and using Bovine Serum Albumin 1mg/mL as standard.

### Permethylation of N-glycans

N-glycans were dissolved in 400 $\mu\text{L}$  of a slurry base extemporaneously prepared as follows. 400 $\mu\text{L}$  of a 50% sodium hydroxide in water were vigorously mixed to 800 $\mu\text{L}$  of methanol and then 4mL of DMSO. After centrifugation at 1500rpm for 1min at room temperature, insoluble reagent was removed and the resulting clear solution washed again with 4mL of DMSO. The operation was repeated until no insoluble compound was observed. Iodomethane or deuterated iodomethane (200 $\mu\text{L}$ ) was added to the slurry base and the methylation was achieved in 30 min at room temperature under stirring. The reaction was stopped by adding 2mL of a cold 5% acetic acid solution and 1mL of chloroform. Aqueous phase was removed and the chloroform phase was washed 5 times with 2 mL of water before drying under a stream of nitrogen

### **MALDI-MS analysis**

The permethylated or deuteromethylated *N*-glycans were solubilized in 100  $\mu$ L of ACN and 1  $\mu$ L was mixed with 1  $\mu$ L of a matrix solution prepared by dissolving 10mg.mL<sup>-1</sup> of 2,5-dihydroxybenzoic acid in 1mL NaOH 1mM and spotted on a MALDI plate. MS and MS<sup>n</sup> spectra were acquired on a 4800 Proteomics Analyzer mass spectrometer (Applied Biosystems, Framingham, MA, USA) and on an Axima Resonance spectrometer (Shimadzu Kyoto, Japan) respectively, in positive reflectron modes. For MS<sup>2</sup> experiments, collision energy was tuned from 300 to 600eV and argon was used as collision gas with a CID control manually adjusted (between 80 ~ 130) to achieve optimum degree of fragmentation. For MS<sup>3</sup>, ion selection occurred with standard resolution with the same collision energy.

### **Quantification of A and ALe<sup>b</sup> epitopes by nanoLC-MS/MS**

Deuteromethylated glycan samples were analyzed by reverse phase C18 nanoLC-MS/MS on an Orbitrap Fusion Tribrid Mass Spectrometer system (ThermoFisher Scientific) fitted with a ReproSil-Pur 120 C18-AQ column (120  $\text{\AA}$ , 1.9  $\mu$ m, 75  $\mu$ m  $\times$  200 mm, Dr. Maisch), using the same LC conditions and data acquisition method as described previously (Tseng et al, 2021). The HCD-MS2 data acquired were processed by in-house developed LC-MS2/MS3 glycomic data mining tool, Glypick, as described [43]. The intensities of MS2 ions representing the diagnostic fragment ions of differentially methylated terminal A and ALe<sup>b</sup> glycotopes that passed the preset 5 ppm filtering criteria were extracted, summed and output in Excel format. The summed ion intensities for each of the target ions differing by 3 mass units representing different degrees of methylation on the same deuteromethylated glycotope were calculated as % total of that structure present in each sample analyzed.

### **Preparation of itol-acetate derivatives**

Monosaccharides of *N*-glycans were released with 4M TFA for 4h at 100°C. After drying under a nitrogen stream, traces of TFA were removed by 4 co-distillations with 1mL of methanol. Monosaccharides were reduced 4h in 500 $\mu$ L 50mM ammonia in the presence of either sodium borohydride 1M or sodium borodeuteride 1M. Borates were removed by 9 co-distillations in methanol/acetic acid (9/1; v/v). Reduced monosaccharides were peracetylated in 500 $\mu$ L acetic anhydride and 20 $\mu$ L pyridine at 100°C for 4h.

### **GC-MS Analysis ItoI-acetates and permethylated derivatives.**

ItoI acetates derivatives were dissolved in 200 $\mu$ L chloroform and injected in splitless mode on a Solgel 1MS, 30 m  $\times$  0.25 mm  $\times$  0.25  $\mu$ m capillary column with the following gradient temperature: 120 to 230°C, 3°C/min and then to 270 °C, 10°C/min. Compounds were detected at 70 eV on a HP-7820 gas chromatograph coupled to a 5977B single quad (Agilent Technologies, Santa Clara, US) in full scan mode from 45 to 500 Da. Permethylated monosaccharides were resuspended in 200 $\mu$ L of chloroform and injected using the same GC set with following gradient : 80°C to 180°C, 2°C/min and then to 250 °C, 30°C/min.

### **Molecular modeling :**

To determine the potential interactions of oyster saccharides with human noroviruses from different genogroups, the X-ray crystallographic

structures of the P domain of Norwalk (GI-1, PDB code 3D26) and TV24 (GII.3, PDB code 6IS5) viruses in complex with A-trisaccharide or A-tetrasaccharide respectively were used. The glucose of the A-tetrasaccharide was removed and hydroxyls at position 3 of GalNAc and 4 of Fuc in the trisaccharides were substituted by a methyl group to model saccharides exposed on the surface by oyster species *Crassostrea gigas* and *Ostrea edulis*. Non-bond interactions between the atoms of viruses and those of saccharides were monitored using BIOVIA Discovery Studio visualizer (Dassault Systèmes, San Diego).

**Supplementary Materials:** The following supporting information can be downloaded at: [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), **Figure S1.** Monosaccharide composition analysis of *N*-glycans; **Figure S2.** Ratios (in %) of Me-3O-Man and Me-4O-Man in *N*-glycans isolated from individual tissues of *C. gigas* and *O. edulis*; **Table S1.** Quantification of monosaccharides in *N*-glycans; **Table S2.** List of *N*-glycans identified as permethylated derivatives; **Table S3.** List of methylated *N*-glycans identified from the mantle of *Crassostrea gigas* and *Ostrea edulis*.

**Supplementary Materials:** The following supporting information can be downloaded at: [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), Figure S1: title; Table S1: title; Video S1: title.

**Author Contributions:** Conceptualization, Y.G., S.F.L-G. and J.L-P.; methodology, A.A., F.C., S.Y.G, S.Y.Y., A.Q., G.E-N.; validation, Y.G., S.F.L-G., M.D., K.K. and J.L-P.; investigation, A.A., F.C., S.Y.G, S.Y.Y., A.Q., G.E-N.; resources, Y.G., S.F.L.G., H.A., K-H.K. and J.L.P.; data curation, Y.G., F.C., K-H.K.; writing—original draft preparation, F.C. and Y.G.; writing—review and editing, F.C., S.F.L-G., M.D., K-H.K. and J.L-P.Y.G.; funding acquisition, Y.G., S.F.L-G., H.A. and J.L-P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded supported by the French National Research Agency ANR- ANR-19-CE35-0014 (GOyAVE) to YG, JLP and SLG; Academia Sinica AS-IR-111-01 (to K-H.K.); JSPS Core-to-Core Program Grant No. JRMJCR18H2 to H.A.; International Research Network UGSF-iGCORE Glyco-Network (CNRS) to YG.

**Acknowledgments:** We are grateful to UAR 2014 - US 41 - Plateformes Lilloises en Biologie et Santé and PAGés-P3M for providing the scientific and technical environment conducive to achieving this work. We also thank the Academia Sinica Common Mass Spectrometry Facilities for Proteomics and Protein Modification Analysis (AS-CFII-108-107) for nanoLC-MS/MS data acquisition.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Farahmand, M.; Moghoofei, M.; Dorost, A.; Shoja, Z.; Ghorbani, S.; Kiani, S.J.; Khales, P.; Esteghamati, A.; Sayyahfar, S.; Jafarzadeh, M.; et al. Global Prevalence and Genotype Distribution of Norovirus Infection in Children with Gastroenteritis: A Meta-analysis on 6 Years of Research from 2015 to 2020. *Reviews in Medical Virology* **2022**, *32*, doi:10.1002/rmv.2237.
2. Glass, R.I.; Parashar, U.D.; Estes, M.K. Norovirus Gastroenteritis. *N Engl J Med* **2009**, *361*, 1776–1785, doi:10.1056/NEJMra0804575.
3. Bányai, K.; Estes, M.K.; Martella, V.; Parashar, U.D. Viral Gastroenteritis. *The Lancet* **2018**, *392*, 175–186, doi:10.1016/S0140-6736(18)31128-0.
4. Caddy, S.; Papa, G.; Borodavka, A.; Desselberger, U. Rotavirus Research: 2014–2020. *Virus Research* **2021**, *304*, 198499, doi:10.1016/j.virusres.2021.198499.
5. Chhabra, P.; de Graaf, M.; Parra, G.I.; Chan, M.C.-W.; Green, K.; Martella, V.; Wang, Q.; White, P.A.; Katayama, K.; Vennema, H.; et al. Updated Classification of Norovirus Genogroups and Genotypes. *Journal of General Virology* **2019**, *100*, 1393–1406, doi:10.1099/jgv.0.001318.
6. Li, Y.; Xue, L.; Gao, J.; Cai, W.; Zhang, Z.; Meng, L.; Miao, S.; Hong, X.; Xu, M.; Wu, Q.; et al. A Systematic

Review and Meta-Analysis Indicates a Substantial Burden of Human Noroviruses in Shellfish Worldwide, with GII.4 and GII.2 Being the Predominant Genotypes. *Food Microbiology* **2023**, *109*, 104140, doi:10.1016/j.fm.2022.104140.

7. King, M.-J. Blood Group Antigens on Human Erythrocytes-Distribution, Structure and Possible Functions. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1994**, *1197*, 15–44, doi:10.1016/0304-4157(94)90017-5.
8. Ewald, D.R.; Sumner, S.C.J. Blood Type Biochemistry and Human Disease. *WIREs Mechanisms of Disease* **2016**, *8*, 517–535, doi:10.1002/wsbm.1355.
9. Ravn, V.; Dabelsteen, E. Tissue Distribution of Histo-Blood Group Antigens. *APMIS* **2000**, *108*, 1–28, doi:10.1034/j.1600-0463.2000.d01-1.x.
10. Marionneau, S.; Cailleau-Thomas, A.; Rocher, J.; Le Moullac-Vaidye, B.; Ruvoën, N.; Clément, M.; Le Pendu, J. ABH and Lewis Histo-Blood Group Antigens, a Model for the Meaning of Oligosaccharide Diversity in the Face of a Changing World. *Biochimie* **2001**, *83*, 565–573, doi:10.1016/S0300-9084(01)01321-9.
11. Ruvoën-Clouet, N.; Belliot, G.; Le Pendu, J. Noroviruses and Histo-Blood Groups: The Impact of Common Host Genetic Polymorphisms on Virus Transmission and Evolution: Noroviruses and Herd Innate Protection. *Rev. Med. Virol.* **2013**, *23*, 355–366, doi:10.1002/rmv.1757.
12. Caddy, S.; Breiman, A.; le Pendu, J.; Goodfellow, I. Genogroup IV and VI Canine Noroviruses Interact with Histo-Blood Group Antigens. *J Virol* **2014**, *88*, 10377–10391, doi:10.1128/JVI.01008-14.
13. Jiang, X.; Liu, Y.; Tan, M. Histo-Blood Group Antigens as Receptors for Rotavirus, New Understanding on Rotavirus Epidemiology and Vaccine Strategy: Rotavirus Host Receptor and Vaccine Strategy. *Emerging Microbes & Infections* **2017**, *6*, 1–8, doi:10.1038/emi.2017.30.
14. Le Pendu, J.; Ruvoën-Clouet, N. Fondness for Sugars of Enteric Viruses Confronts Them with Human Glycans Genetic Diversity. *Hum Genet* **2020**, *139*, 903–910, doi:10.1007/s00439-019-02090-w.
15. Yu, Y.; Cai, H.; Hu, L.; Lei, R.; Pan, Y.; Yan, S.; Wang, Y. Molecular Epidemiology of Oyster-Related Human Noroviruses and Their Global Genetic Diversity and Temporal-Geographical Distribution from 1983 to 2014. *Appl Environ Microbiol* **2015**, *81*, 7615–7624, doi:10.1128/AEM.01729-15.
16. Metcalf, T.G.; Melnick, J.L.; Estes, M.K. Environmental Virology: From Detection of Virus in Sewage and Water by Isolation to Identification by Molecular Biology—a Trip of over 50 Years. *Annu Rev Microbiol* **1995**, *49*, 461–487, doi:10.1146/annurev.mi.49.100195.002333.
17. Lowther, J.A.; Gustar, N.E.; Powell, A.L.; O’Brien, S.; Lees, D.N. A One-Year Survey of Norovirus in UK Oysters Collected at the Point of Sale. *Food Environ Virol* **2018**, *10*, 278–287, doi:10.1007/s12560-018-9338-4.
18. Jørgensen, C. Bivalve Filter Feeding Revisited. *Mar. Ecol. Prog. Ser.* **1996**, *142*, 287–302, doi:10.3354/meps142287.
19. Maalouf, H.; Schaeffer, J.; Parnaudeau, S.; Le Pendu, J.; Atmar, R.L.; Crawford, S.E.; Le Guyader, F.S. Strain-Dependent Norovirus Bioaccumulation in Oysters. *Appl. Environ. Microbiol.* **2011**, *77*, 3189–3196, doi:10.1128/AEM.03010-10.
20. Loisy, F.; Atmar, R.L.; Le Saux, J.-C.; Cohen, J.; Caprais, M.-P.; Pommepuy, M.; Le Guyader, F.S. Use of Rotavirus Virus-Like Particles as Surrogates To Evaluate Virus Persistence in Shellfish. *Appl Environ Microbiol* **2005**, *71*, 6049–6053, doi:10.1128/AEM.71.10.6049-6053.2005.
21. Polo, D.; Schaeffer, J.; Fournet, N.; Le Saux, J.-C.; Parnaudeau, S.; McLeod, C.; Le Guyader, F.S. Digital PCR for Quantifying Norovirus in Oysters Implicated in Outbreaks, France. *Emerg. Infect. Dis.* **2016**, *22*, 2189–2191, doi:10.3201/eid2212.160841.
22. Rupnik, A.; Doré, W.; Devilly, L.; Fahy, J.; Fitzpatrick, A.; Schmidt, W.; Hunt, K.; Butler, F.; Keaveney, S. Evaluation of Norovirus Reduction in Environmentally Contaminated Pacific Oysters During Laboratory Controlled and Commercial Depuration. *Food Environ Virol* **2021**, *13*, 229–240, doi:10.1007/s12560-021-09464-2.
23. Battistini, R.; Masotti, C.; Maurella, C.; Costa, E.; Orlandi, M.; Feletti, M.; Ercolini, C.; Serracca, L. Study on the Effect of Relaying on Norovirus Reduction from *Crassostrea Gigas* Oysters. *Microorganisms* **2022**, *10*, 2389, doi:10.3390/microorganisms10122389.
24. Le Guyader, F.S.; Loisy, F.; Atmar, R.L.; Hutson, A.M.; Estes, M.K.; Ruvoën-Clouet, N.; Pommepuy, M.; Le Pendu, J. Norwalk Virus-Specific Binding to Oyster Digestive Tissues. *Emerg. Infect. Dis.* **2006**, *12*, 931–936, doi:10.3201/eid1206.051519.
25. Tian, P.; Engelbrektson, A.L.; Jiang, X.; Zhong, W.; Mandrell, R.E. Norovirus Recognizes Histo-Blood Group Antigens on Gastrointestinal Cells of Clams, Mussels, and Oysters: A Possible Mechanism of Bioaccumulation. *Journal of Food Protection* **2007**, *70*, 2140–2147, doi:10.4315/0362-028X-70.9.2140.
26. Maalouf, H.; Zakhour, M.; Le Pendu, J.; Le Saux, J.-C.; Atmar, R.L.; Le Guyader, F.S. Distribution in Tissue and Seasonal Variation of Norovirus Genogroup I and II Ligands in Oysters. *Applied and Environmental Microbiology* **2010**, *76*, 5621–5630, doi:10.1128/AEM.00148-10.

27. Lowmoung, T.; Pombubpa, K.; Duangdee, T.; Tipayamongkolgul, M.; Kittigul, L. Distribution of Naturally Occurring Norovirus Genogroups I, II, and IV in Oyster Tissues. *Food Environ Virol* **2017**, *9*, 415–422, doi:10.1007/s12560-017-9305-5.
28. Su, L.; Ma, L.; Liu, H.; Zhao, F.; Su, Z.; Zhou, D. Presence and Distribution of Histo-Blood Group Antigens in Pacific Oysters and the Effects of Exposure to Noroviruses GI.3 and GI.4 on Their Expression. *Journal of Food Protection* **2018**, *81*, 1783–1790, doi:10.4315/0362-028X.JFP-18-074.
29. Paschinger, K.; Wilson, I.B.H. Comparisons of N-Glycans across Invertebrate Phyla. *Parasitology* **2019**, *146*, 1733–1742, doi:10.1017/S0031182019000398.
30. Kurz, S.; Jin, C.; 金春生; Hykollari, A.; Gregorich, D.; Giomarelli, B.; Vasta, G.R.; Wilson, I.B.H.; Paschinger, K. Hemocytes and Plasma of the Eastern Oyster (*Crassostrea Virginica*) Display a Diverse Repertoire of Sulfated and Blood Group A-Modified N-Glycans\*. *Journal of Biological Chemistry* **2013**, *288*, 24410–24428, doi:10.1074/jbc.M113.478933.
31. Feng, C.; Ghosh, A.; Amin, M.N.; Giomarelli, B.; Shridhar, S.; Banerjee, A.; Fernández-Robledo, J.A.; Bianchet, M.A.; Wang, L.-X.; Wilson, I.B.H.; et al. The Galectin CvGal1 from the Eastern Oyster (*Crassostrea Virginica*) Binds to Blood Group A Oligosaccharides on the Hemocyte Surface. *J. Biol. Chem.* **2013**, *288*, 24394–24409, doi:10.1074/jbc.M113.476531.
32. Varki, A. Nothing in Glycobiology Makes Sense, except in the Light of Evolution. *Cell* **2006**, *126*, 841–845, doi:10.1016/j.cell.2006.08.022.
33. Delannoy, C.; Huang, C.; Coddeville, B.; Chen, J.-Y.; Mouajjah, D.; Groux-Degroote, S.; Harduin-Lepers, A.; Khoo, K.-H.; Guerardel, Y.; Ellass-Rochard, E. Mycobacterium Bovis BCG Infection Alters the Macrophage N -Glycome. *Molecular Omics* **2020**, *16*, 345–354, doi:10.1039/C9MO00173E.
34. Lemp, D.; Haselbeck, A.; Klebl, F. Molecular Cloning and Heterologous Expression of N-Glycosidase F from *Flavobacterium Meningosepticum*. *J Biol Chem* **1990**, *265*, 15606–15610.
35. Takahashi, N. Demonstration of a New Amidase Acting on Glycopeptides. *Biochemical and Biophysical Research Communications* **1977**, *76*, 1194–1201, doi:10.1016/0006-291X(77)90982-2.
36. Qin, Y.; Li, R.; Liao, Q.; Shi, G.; Zhou, Y.; Wan, W.; Li, J.; Ma, H.; Zhang, Y.; Yu, Z. Comparison of Biochemical Composition, Nutritional Quality, and Metals Concentrations between Males and Females of Three Different *Crassostrea* Sp. *Food Chemistry* **2023**, *398*, 133868, doi:10.1016/j.foodchem.2022.133868.
37. Yu, S.-Y.; Wu, S.-W.; Khoo, K.-H. Distinctive Characteristics of MALDI-Q/TOF and TOF/TOF Tandem Mass Spectrometry for Sequencing of Permethylated Complex Type N-Glycans. *Glycoconj J* **2006**, *23*, 355–369, doi:10.1007/s10719-006-8492-3.
38. Domon, B.; Costello, C.E. A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconjugate Journal* **1988**, *5*, 397–409.
39. Varki, A.; Cummings, R.D.; Aebi, M.; Packer, N.H.; Seeberger, P.H.; Esko, J.D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25*, 1323–1324, doi:10.1093/glycob/cwv091.
40. Schachter, H. The ‘Yellow Brick Road’ to Branched Complex N -Glycans. *Glycobiology* **1991**, *1*, 453–461, doi:10.1093/glycob/1.5.453.
41. Hulsén, T.; de Vlieg, J.; Alkema, W. BioVenn – a Web Application for the Comparison and Visualization of Biological Lists Using Area-Proportional Venn Diagrams. *BMC Genomics* **2008**, *9*, 488, doi:10.1186/1471-2164-9-488.
42. Staudacher, E. Methylation – an Uncommon Modification of Glycans. *Biological Chemistry* **2012**, *393*, 675–685, doi:10.1515/hsz-2012-0132.
43. Hsiao, C.-T.; Wang, P.-W.; Chang, H.-C.; Chen, Y.-Y.; Wang, S.-H.; Chern, Y.; Khoo, K.-H. Advancing a High Throughput Glycotope-Centric Glycomics Workflow Based on NAnoLC-MS2-Product Dependent-MS3 ANALYSIS of Permethylated Glycans\*. *Molecular & Cellular Proteomics* **2017**, *16*, 2268–2280, doi:10.1074/mcp.TIR117.000156.
44. Takahashi, N.; Masuda, K.; Hiraki, K.; Yoshihara, K.; Huang, H.-H.; Khoo, K.-H.; Kato, K. N-Glycan Structures of Squid Rhodopsin. Existence of the Alpha1-3 and Alpha1-6 Difucosylated Innermost GlcNAc Residue in a Molluscan Glycoprotein. *Eur J Biochem* **2003**, *270*, 2627–2632, doi:10.1046/j.1432-1033.2003.03636.x.
45. Oltmanns, A.; Hoepfner, L.; Scholz, M.; Zinzus, K.; Schulze, S.; Hippler, M. Novel Insights Into N-Glycan Fucosylation and Core Xylosylation in *C. Reinhardtii*. *Front. Plant Sci.* **2020**, *10*, 1686, doi:10.3389/fpls.2019.01686.
46. Minagawa, S.; Sekiguchi, S.; Nakaso, Y.; Tomita, M.; Takahisa, M.; Yasuda, H. Identification of Core Alpha 1,3-Fucosyltransferase Gene From Silkworm: An Insect Popularly Used to Express Mammalian Proteins. *J Insect Sci* **2015**, *15*, 110, doi:10.1093/jisesa/iev088.
47. Yan, S.; Vanbeselaere, J.; Jin, C.; Blaukopf, M.; Wöls, F.; Wilson, I.B.H.; Paschinger, K. Core Richness of N-Glycans of *Caenorhabditis Elegans*: A Case Study on Chemical and Enzymatic Release. *Anal. Chem.* **2018**, *90*, 928–935, doi:10.1021/acs.analchem.7b03898.



48. Fabini, G.; Freilinger, A.; Altmann, F.; Wilson, I.B.H. Identification of Core A1,3-Fucosylated Glycans and Cloning of the Requisite Fucosyltransferase CDNA from *Drosophila Melanogaster*: POTENTIAL BASIS OF THE NEURAL ANTI-HORSERADISH PEROXIDASE EPI TOPE \*. *Journal of Biological Chemistry* **2001**, *276*, 28058–28067, doi:10.1074/jbc.M100573200.
49. Paschinger, K.; Staudacher, E.; Stemmer, U.; Fabini, G.; Wilson, I.B.H. Fucosyltransferase Substrate Specificity and the Order of Fucosylation in Invertebrates. *Glycobiology* **2005**, *15*, 463–474, doi:10.1093/glycob/cwi028.
50. Titz, A.; Butschi, A.; Henrissat, B.; Fan, Y.-Y.; Hennet, T.; Razzazi-Fazeli, E.; Hengartner, M.O.; Wilson, I.B.H.; Künzler, M.; Aebi, M. Molecular Basis for Galactosylation of Core Fucose Residues in Invertebrates. *J Biol Chem* **2009**, *284*, 36223–36233, doi:10.1074/jbc.M109.058354.
51. Eckmair, B.; Jin, C.; Abed-Navandi, D.; Paschinger, K. Multistep Fractionation and Mass Spectrometry Reveal Zwitterionic and Anionic Modifications of the N- and O-Glycans of a Marine Snail. *Mol Cell Proteomics* **2016**, *15*, 573–597, doi:10.1074/mcp.M115.051573.
52. Stoeva, S.; Rachev, R.; Severov, S.; Voelter, W.; Genov, N. Carbohydrate Content and Monosaccharide Composition of *Rapana Thomasiana* Grosse (Gastropoda) Hemocyanin and Its Structural Subunits. Comparison with Gastropodan Hemocyanins. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **1995**, *110*, 761–765, doi:10.1016/0305-0491(94)00201-5.
53. Khoo K.-H. Structural Variations in Schistosomal Glycans. *Trends in Glycoscience and Glycotechnology* **2001**, *13*, 493–506, doi:10.4052/tigg.13.493.
54. Smit, C.H.; Diepen, A. van; Nguyen, D.L.; Wuhler, M.; Hoffmann, K.F.; Deelder, A.M.; Hokke, C.H. Glycomic Analysis of Life Stages of the Human Parasite *Schistosoma Mansoni* Reveals Developmental Expression Profiles of Functional and Antigenic Glycan Motifs \*[S]. *Molecular & Cellular Proteomics* **2015**, *14*, 1750–1769, doi:10.1074/mcp.M115.048280.
55. Hori, T.; Sugita, M.; Ando, S.; Kuwahara, M.; Kumauchi, K.; Sugie, E.; Itasaka, O. Characterization of a Novel Glycosphingolipid, Ceramide Nonasaccharide, Isolated from Spermatozoa of the Fresh Water Bivalve, *Hyriopsis Schlegelii*. *J Biol Chem* **1981**, *256*, 10979–10985.
56. Spiro, R.G.; Bhojroo, V.D. Occurrence of Sulfate in the Asparagine-Linked Complex Carbohydrate Units of Thyroglobulin. Identification and Localization of Galactose 3-Sulfate and N-Acetylglucosamine 6-Sulfate Residues in the Human and Calf Proteins. *Journal of Biological Chemistry* **1988**, *263*, 14351–14358, doi:10.1016/S0021-9258(18)68227-4.
57. Miyata, S.; Sato, C.; Kumita, H.; Toriyama, M.; Vacquier, V.D.; Kitajima, K. Flagelliasialin: A Novel Sulfated 2,9-Linked Polysialic Acid Glycoprotein of Sea Urchin Sperm Flagella. *Glycobiology* **2006**, *16*, 1229–1241, doi:10.1093/glycob/cwl036.
58. Yu, S.-Y.; Wu, S.-W.; Hsiao, H.-H.; Khoo, K.-H. Enabling Techniques and Strategic Workflow for Sulfoglycomics Based on Mass Spectrometry Mapping and Sequencing of Permethylated Sulfated Glycans. *Glycobiology* **2009**, *19*, 1136–1149, doi:10.1093/glycob/cwp113.
59. Yu, S.-Y.; Snovidia, S.; Khoo, K.-H. Permethylated and Microfractionation of Sulfated Glycans for MS Analysis. *BIO-PROTOCOL* **2020**, *10*, doi:10.21769/BioProtoc.3617.
60. Tan, M.; Jiang, X. Histo-Blood Group Antigens: A Common Niche for Norovirus and Rotavirus. *Expert Rev Mol Med* **2014**, *16*, e5, doi:10.1017/erm.2014.2.
61. Taube, S.; Mallagaray, A.; Peters, T. Norovirus, Glycans and Attachment. *Curr Opin Virol* **2018**, *31*, 33–42, doi:10.1016/j.coviro.2018.04.007.
62. Tenge, V.R.; Hu, L.; Prasad, B.V.V.; Larson, G.; Atmar, R.L.; Estes, M.K.; Ramani, S. Glycan Recognition in Human Norovirus Infections. *Viruses* **2021**, *13*, 2066, doi:10.3390/v13102066.
63. Le Guyader, F.S.; Atmar, R.L.; Le Pendu, J. Transmission of Viruses through Shellfish: When Specific Ligands Come into Play. *Current Opinion in Virology* **2012**, *2*, 103–110, doi:10.1016/j.coviro.2011.10.029.
64. Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P.I.; Parra, F.; Peters, T. Molecular Details of the Recognition of Blood Group Antigens by a Human Norovirus as Determined by STD NMR Spectroscopy. *Angew Chem Int Ed Engl* **2012**, *51*, 928–932, doi:10.1002/anie.201105719.
65. Fiege, B.; Leuthold, M.; Parra, F.; Dalton, K.P.; Meloncelli, P.J.; Lowary, T.L.; Peters, T. Epitope Mapping of Histo Blood Group Antigens Bound to Norovirus VLPs Using STD NMR Experiments Reveals Fine Details of Molecular Recognition. *Glycoconj J* **2017**, *34*, 679–689, doi:10.1007/s10719-017-9792-5.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

## B. Analyses complémentaires

### 1. Etablissement des *N*-glycomes de palpes labiaux, de l'hémolymphe, du plasma et des hémocytes de *C. gigas*

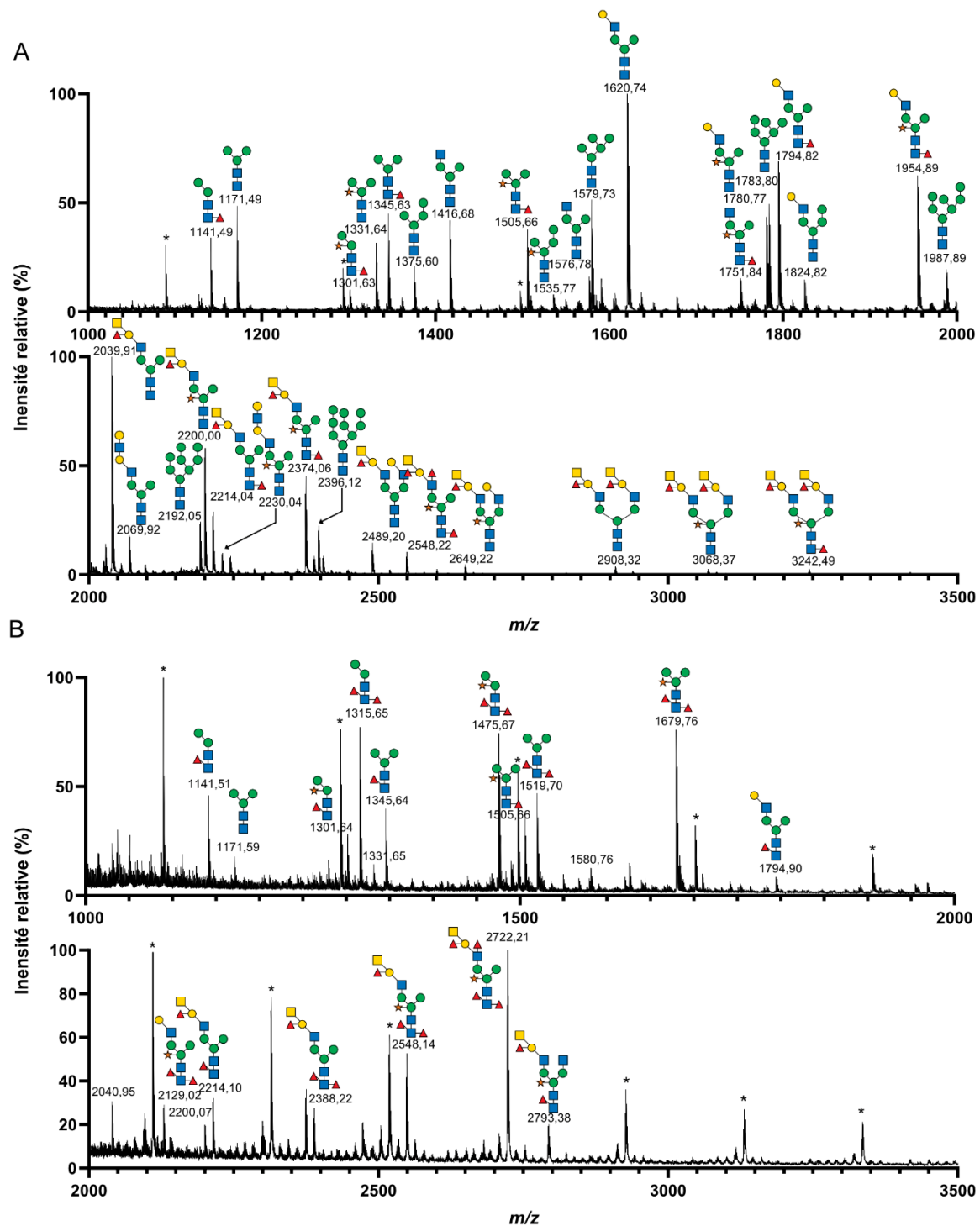
En parallèle des analyses des *N*-glycannes des branchies, manteau et tube digestif, nous avons analysé les *N*-glycannes de quelques tissus, liquides ou cellules supplémentaires isolés de l'huître *C. gigas* : les palpes labiaux, l'hémolymphe, du plasma et des hémocytes isolés de l'hémolymphe (**Figure 43** p143 et Figure 44 p144).

Nous avons identifié 47, 20, 18 et 6 *N*-glycannes dans les palpes labiaux, l'hémolymphe, le plasma et les hémocytes de *C. gigas* respectivement. Qualitativement, les proportions de chaque type de structures *N*-glycannique des palpes labiaux (paucimannose 32 %, oligomannose 11 %, complexe 55 % et hybride 2 %) sont très similaires avec celles retrouvées pour les *N*-glycannes totaux des branchies, manteau et tube digestif de *C. gigas* (**Figure 45A** p145). L'hémolymphe et le plasma ont des proportions de chaque type de *N*-glycannes moins semblables aux tissus mais en ayant tout de même une majorité de *N*-glycannes de type complexe (85 et 83 %) puis de type paucimannose (10 et 11 %) et oligomannose (5 et 6 %). Pour ces échantillons, nous n'avons pas identifié de type hybride. Par rapport aux tissus, ces derniers semblent disposer de moins de *N*-glycannes de type paucimannose. Pour les hémocytes dont les structures *N*-glycanniques s'élève à 6, les structures sont réparties de manière équivalentes entre les types paucimannose, oligomannose et complexe. Il est peu judicieux de faire un comparatif avec les autres échantillons au vu de la faible diversité de *N*-glycannes et du spectre de faible sensibilité.

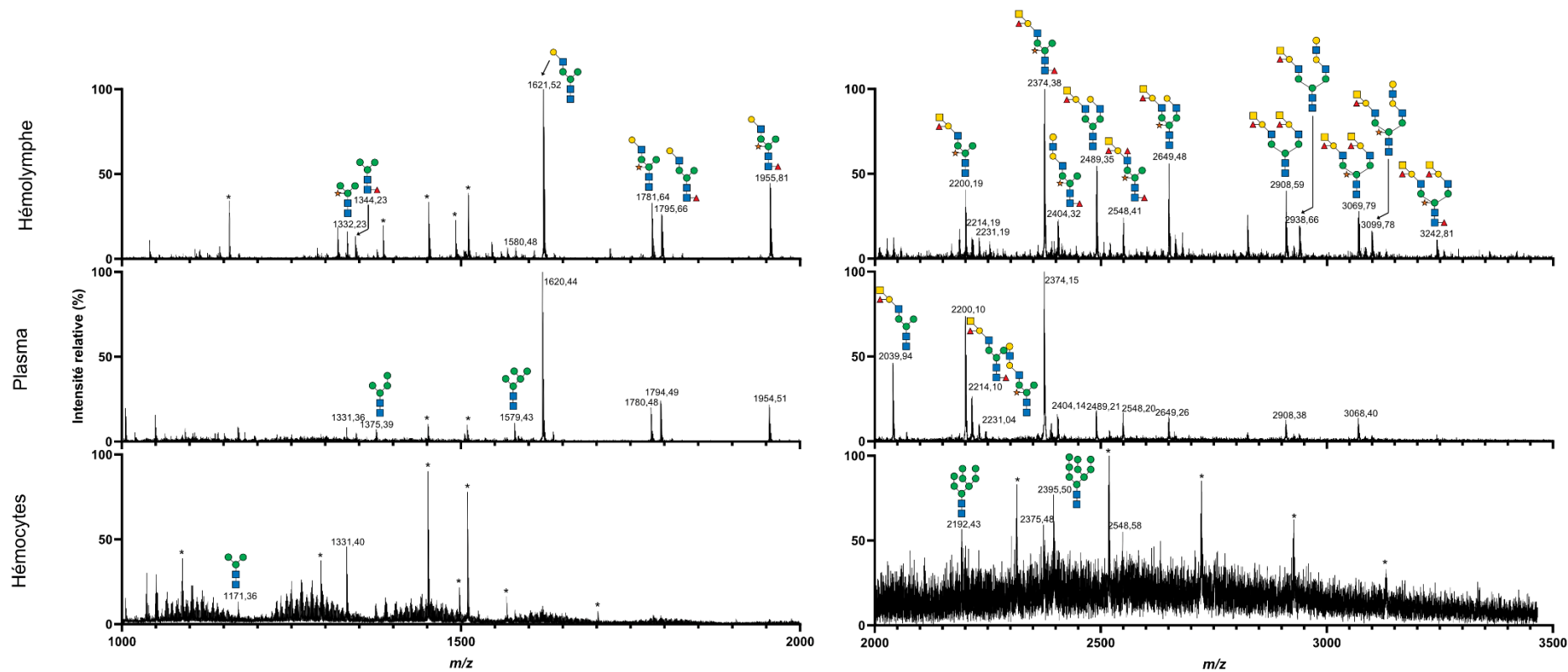
Les palpes labiaux, l'hémolymphe et le plasma ont des proportions d'Ag A, d'Ag ALe<sup>b</sup>, d'antenne de Type 1 nue et de répétition LacNAc très similaires aux branchies, tube digestif et manteau (**Figure 45B**). On peut noter une proportion plus importante d'antennes de type 1 dans ces 3 échantillons par rapport aux 3 tissus. Pour les hémocytes, uniquement un *N*-glycanne porteur d'un Ag A et un *N*-glycanne porteur d'un Ag ALe<sup>b</sup> ont été identifiés.

Ces échantillons n'ont pas fait apparaître d'autres structures *N*-glycanniques que celles décrites exhaustivement dans notre article pour les branchies, le tube digestif et le manteau. Le fait que ces

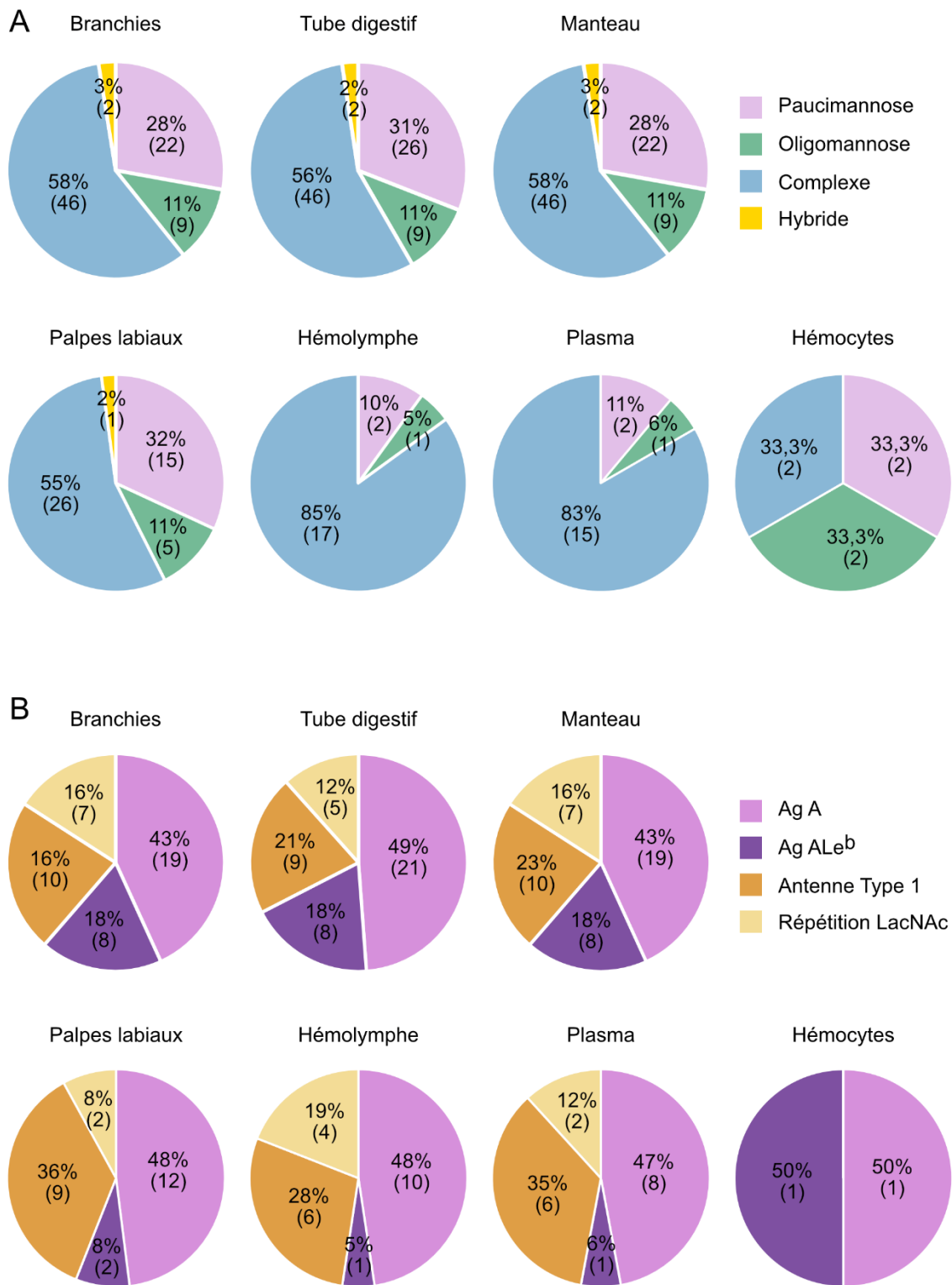
échantillons soient, d'une part, pauvres en quantité de matériels glycaniques et, d'autre part, difficile à préparer, nous avons décidé de ne pas aller plus loin dans leurs analyses.



**Figure 43 :** Analyse MALDI-MS des *N*-glycannes perméthylés isolés des palpes labiaux de *Crassostrea gigas*. Avant leur perméthylation, les *N*-glycannes ont été libérés séquentiellement par (A) la peptidyl-*N*-glycosidase F et (B) la peptidyl-*N*-glycosidase A. L'enchaînement des monosaccharides est basé sur la masse détectée pour l'ion moléculaire  $[M+Na]^+$  et les analyses MS<sup>n</sup>. Les structures *N*-glycanniques sont représentées selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019). La présence d'un astérisque (\*) sur un signal *m/z* indique la présence d'un contaminant de nature polymérique dont les unités de répétitions ont une masse de 204 Da.



**Figure 44 :** Analyse MALDI-MS des *N*-glycannes libérés par la peptidyl-*N*-glycosidase F puis perméthylés à partir de (A) l'hémolymphe, (B) du plasma et (C) des hémocytes de *Crassostrea gigas*. L'enchaînement des monosaccharides est basé sur la masse détectée pour l'ion moléculaire  $[M+Na]^+$  et les analyses  $MS^n$ . Les structures *N*-glycanniques sont représentées selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019). La présence d'un astérisque (\*) sur un signal *m/z* indique la présence de contaminants dont l'un est de nature polymérique avec des unités de répétitions de 204 Da.



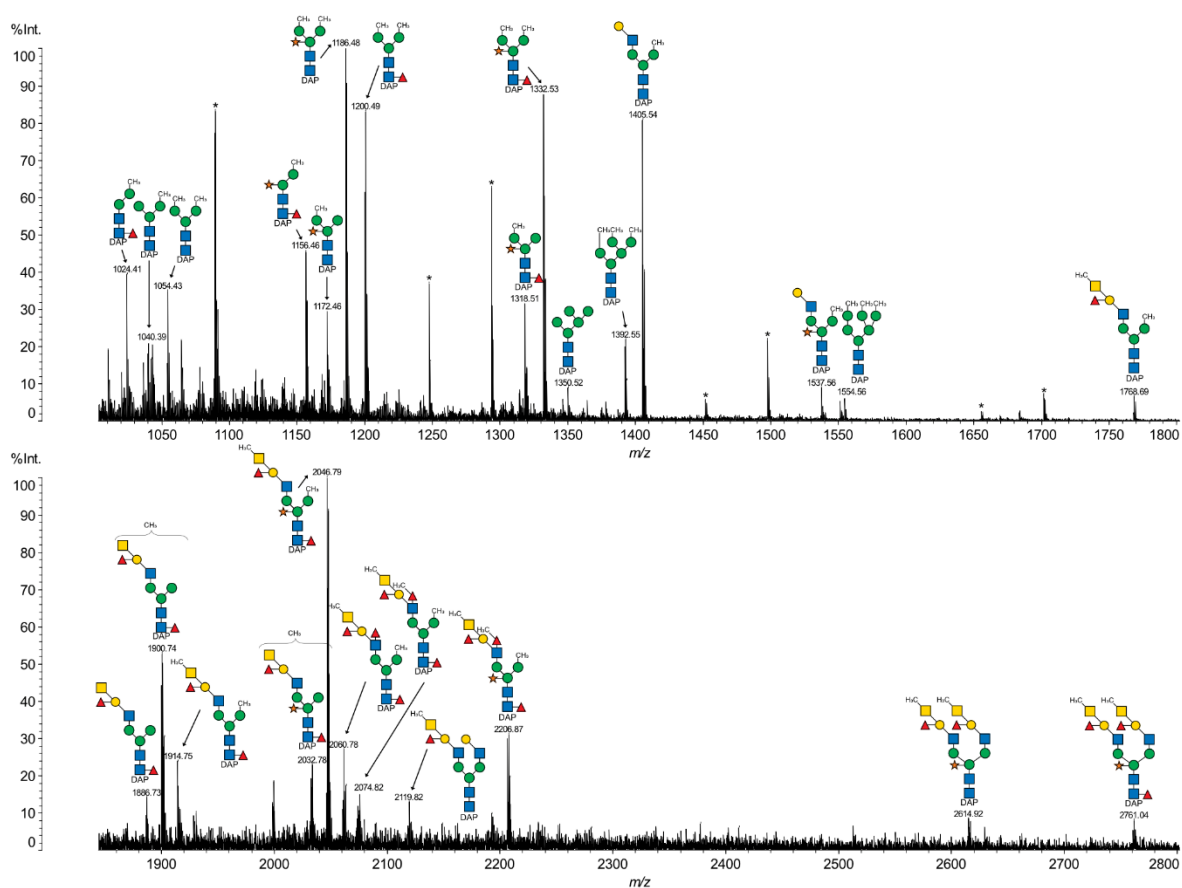
**Figure 45** : Proportions des familles et des antennes de *N*-glycannes retrouvés sur les glycoprotéines des branchies, du tube digestif, du manteau, des palpes labiaux, de l'hémolymphe, du plasma et des hémocytes de *Crassostrea gigas*. **(A)** Proportions en pourcentage des différentes familles (paucimannose, oligomannose, complexe et hybride) de *N*-glycannes. **(B)** Proportion en pourcentage des antigènes A (Ag A), antigènes A Lewis b (ALe<sup>b</sup>), antennes de type 1 et répétition LacNAc. Les nombres indiqués entre parenthèses correspondent au nombre de structures *N*-glycanniques par catégorie.

## 1. Essais de mise en évidence de groupements sulfates

L'étude des *N*-glycannes de l'hémolymphe de l'huître *Crassostrea virginica* réalisée par Kurz *et al.* (2013) a montré la présence de groupements méthyles mais aussi de groupements sulfates. L'analyse des oligosaccharides sulfatés est rendue difficile car l'extraction classique des *N*-glycannes perméthylés à l'issue de la perméthylation par les solvants chloroforme et eau entraîne une élimination des oligosaccharides sulfatés dans la phase aqueuse du fait de la présence des groupements sulfates qui leur confèrent un caractère hydrophile suffisant. Pour éviter la perte de ces oligosaccharides, Yu *et al.* (2020) ont mis en place un protocole utilisant une colonne Oasis MAX dont le fractionnement des molécules repose à la fois sur leur polarité et leur état de charge. Ce support permet, d'une part, d'éliminer les réactifs en excès et les sels produits durant l'étape de perméthylation en fixant les glycannes perméthylés et, d'autre part, de fractionner en 3 populations distinctes les *N*-glycannes sulfatés et perméthylés selon leur degré de sulfatation, d'abord les oligosaccharides non-sulfatés puis mono-sulfatés et enfin poly-sulfatés. Les essais entrepris avec ce protocole sur les *N*-glycannes du manteau de *C. gigas*, ont permis de détecter par MALDI-MS en mode positif les *N*-glycannes perméthylés neutres élués dans la première fraction. Cependant, l'analyse des deux autres fractions (mono et poly-sulfatés) par MALDI-MS en mode négatif n'a pas montré la présence de signaux *m/z* correspondant à des *N*-glycannes sulfatés. En effet, les spectres auraient dû révéler la présence de signaux à un rapport *m/z* correspondant à la masse théorique de l'oligosaccharide (extrémité réductrice, non réductrice et composition en monosaccharides) incrémenté de 65 Da pour un seul groupement sulfate ( $\text{OSO}_3^- - \text{OCH}_3$ ) et de 88 Da à partir du deuxième groupement sulfate, à cause de la fixation d'un ion sodium  $\text{Na}^+$  (23 Da) à chaque groupement sulfate supplémentaire.

En parallèle, en prenant exemple sur les travaux de Kurz *et al.* (2013), nous avons dérivé par amination réductive les *N*-glycannes du manteau de *C. gigas* en utilisant la diaminopyridine (DAP) plutôt que la 2-aminopyridine (2AP). De façon similaire à l'analyse des chaînes oligosaccharidiques natives, cette dérivation préserve tous les substituants pouvant être portés par les monosaccharides et confère de surcroît une plus grande sensibilité de détection des oligosaccharides dérivés comparée à celle des oligosaccharides non dérivés grâce au groupement amide introduit lors de la dérivation qui se

protone facilement. Le spectre MALDI-MS obtenu en mode positif est présenté en Figure 46. Cette analyse confirme les caractéristiques structurales des *N*-glycannes identifiées précédemment mais ne fait pas apparaître d'autres signaux pouvant correspondre à des *N*-glycannes sulfatés. De plus, l'analyse MALDI-MS en mode négatif du même échantillon n'a pas permis d'obtenir de spectre exploitable ce qui conforte l'idée qu'il n'y ait pas de groupements sulfates sur les *N*-glycannes de l'huître *C. gigas*.



**Figure 46 :** Analyse MALDI-MS des *N*-glycannes libérés par la peptidyl-*N*-glycosidase F puis dérivés à la diaminopyridine à partir du manteau de *Crassostrea gigas*. L'enchaînement des monosaccharides est basé sur la masse détectée pour l'ion moléculaire  $[M+Na]^+$  et les analyses  $MS^n$  ultérieure. Les structures *N*-glycanniques sont représentées selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019). La présence d'un astérisque (\*) sur un signal *m/z* indique la présence d'un contaminant de nature polymérique dont les unités de répétitions ont une masse de 204Da.

### III. Analyse des *O*-glycannes

En plus de l'établissement du *N*-glycome, nous avons cherché à établir le *O*-glycome des 3 mêmes tissus des 2 espèces d'huîtres. La libération des *O*-glycannes des glycopeptides a été réalisée par voie chimique en milieu alcalin modéré en présence de borohydrure de sodium (NaOH 50 mM NaBH<sub>4</sub> 1 M) à 37°C pendant 72h.

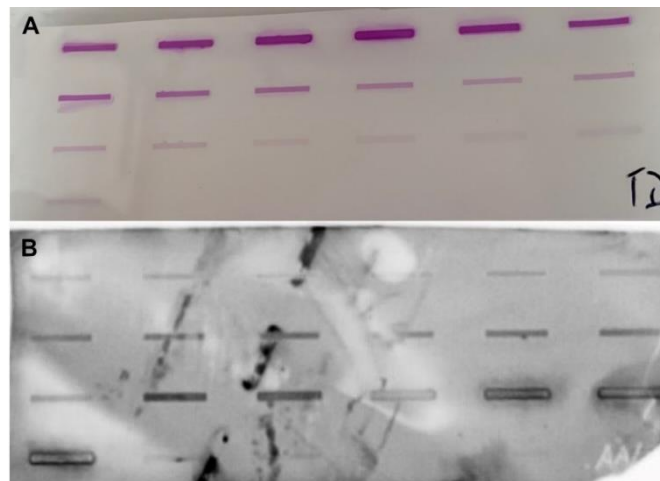
Appliqué aux *O*-glycopeptides résiduels après libération des *N*-glycannes par les PNGases F et A, les analyses MALDI-MS ont révélé la présence de quelques signaux *m/z* pouvant correspondre à des *O*-glycannes. Néanmoins, l'intensité de ces signaux était trop faible pour assurer avec certitude leur existence. Compte tenu du nombre d'étapes effectués depuis les tissus de départ jusqu'à l'obtention de ces fractions (3 extractions en milieu solide sur cartouche C18 et 2 traitements enzymatiques), nous avons postulé que la quantité résiduelle de *O*-glycopeptides résiduels était trop faible pour espérer visualiser les *O*-glycannes (Figure 42 p107). Pour vérifier notre hypothèse, nous avons effectué une  $\beta$ -élimination réductive directement sur le mélange peptides/glycopeptides obtenu après action de la trypsine. L'analyse par spectrométrie de masse MALDI-MS des oligosaccharides perméthylés a fait apparaître de très nombreux signaux *m/z* correspondants aux *N*-glycannes qui sont également libérés par ce traitement, et quelques signaux pouvant correspondre à des *O*-glycannes (spectre non montré). Malheureusement, du fait de ces très nombreux signaux *m/z*, l'interprétation des spectres s'est avérée extrêmement difficile et peu fiable.

On sait depuis très longtemps que les *O*-glycannes sont surtout retrouvés en abondance au niveau des mucines qui tapissent les épithéliums ou du mucus sécrété par les cellules. Ces molécules se caractérisent par un très haut poids moléculaire et une densité élevée qui est mise à profit lorsque l'on veut préparer les mucines (Rossez *et al.*, 2012). Dans ce cas, les tissus sont broyés dans un milieu contenant du chlorure de guanidine 4 M auquel est ajouté une quantité suffisante de chlorure de césium pour amener la densité de la solution finale à 1,4. Après ultracentrifugation et formation d'un gradient de densité auto-formé, les mucines sont retrouvées dans une zone de densité voisine de 1,2-1,3. Nous avons donc décidé de mettre en œuvre ce protocole, couramment appliqué à l'étude des mucines chez les mammifères.

A l'issue de l'ultracentrifugation, des fractions de 0,5 mL ont été collectées et déposées par une unité de transfert Slot blot (50  $\mu$ L) sur 2 membranes de nitrocellulose. Une des membranes a ensuite été révélée par coloration au réactif de Schiff après oxydation ménagée à l'acide périodique (PAS) et l'autre membrane par la lectine *Aleuria aurantia* (AAL) marquée à la biotine (**Figure 47**). Alors que la coloration PAS révèle la présence de molécules renfermant du matériel glucidique, la seconde reconnaît



spécifiquement les résidus de Fuc quelle que soit la façon dont ils sont liés. Comme le montre la membrane de nitrocellulose après révélation, la coloration PAS révèle très fortement les 12 premières fractions les plus denses que l'on pense être du glycogène alors que pour ces mêmes fractions, la lectine AAL marquée à la biotine ne conduit à aucun signal. En revanche, elle reconnaît les dernières fractions collectées de plus faible densité. Ainsi, les fractions révélées au PAS ont été écartées tandis que celles révélées par AAL ont été rassemblées et dessalées pour les traitements ultérieurs.



**Figure 47** : Coloration et Slot-blot réalisés sur les fractions des glycoprotéines du tube digestif de l'huître *Crassostrea gigas* obtenu après ultracentrifugation. (A) Les fractions déposées sur membrane de nitrocellulose ont été colorées à l'acide périodique de Schiff puis (B) un lectine blot utilisant la lectine AAL comme sonde a été réalisé. Les fractions ont été déposées de gauche à droite et de haut en bas. La première fraction déposée correspond à celle la plus dense et la dernière, à celle la moins dense.

Après dessalage, les GPs ont été réduites et alkylées puis soumises à une  $\beta$ -élimination en milieu réducteur en vue de libérer les *O*-glycannes. Après 3 étapes de dessalage pour éliminer les sels, peptides et acides aminés générés durant ce traitement, une fraction de l'échantillon a été perméthylée puis analysée par spectrométrie de masse MALDI-MS afin de vérifier la présence de *O*-glycannes. En parallèle, une autre fraction de l'échantillon a été perdeutérométhylée puis analysée par spectrométrie de masse MALDI-MS et MALDI-MS<sup>n</sup> afin de visualiser ou non la présence de groupements méthyles au sein de ces *O*-glycannes.

L'analyse exhaustive des spectres de masse MALDI-MS des *O*-glycannes perméthylés obtenus à partir des 3 tissus des huîtres *C. gigas* et *O. edulis* a montré la présence potentielle de 24 *O*-glycannes dont la composition en monosaccharides est présentée dans le **Tableau 4**. Les signaux  $m/z$  534, 708,

738, 912, 953, 1157, 1198, 1331, 1443 et 2067 ont été analysés en MALDI-MS<sup>2</sup> voire MS<sup>3</sup> sur le tissu digestif de *C. gigas* (indiqués par \* dans la **Figure 48** p153).

**Tableau 4 :** Liste des structures *O*-glycanniques putatives identifiées par spectrométrie de masse MALDI-QIT-TOF dans les branchies (B), le tube digestif (D) et le manteau (M) des huîtres *Crassostrea gigas* et *Ostrea edulis*. Les *O*-glycannes ont été libérés par β-élimination, dessalés puis perméthylés. Les signaux sont ordonnés suivant leur rapport *m/z* croissant. La présence d'un astérisque (\*) indique que le signal a été analysé par MALDI-MS<sup>2</sup> dans le tissu digestif de *C. gigas*. Les compositions brutes en monosaccharides ont été déduites à partir des analyses MALDI-MS des dérivés perméthylés. La présence des différents signaux dans chacun des tissus analysés est indiqué par un rectangle bleu. La quantité de groupement méthyle naturel a été déterminée à partir des analyses MALDI-MS des dérivés perdeutérométhylés. A titre d'exemple, 012 signifie que l'oligosaccharide existe sous les formes non-, mono- et di-méthylées.

<i>O</i> -glycannes perméthylés <i>m/z</i>	Composition en monosaccharides	<i>C. gigas</i> <i>O. edulis</i>						Nombre de groupements méthyles					
		<i>C. gigas</i>			<i>O. edulis</i>			<i>C. gigas</i>		<i>O. edulis</i>			
		B	D	M	B	D	M	D	M	D	M		
534*	Hex <sub>1</sub> HexNAC <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
575	HexNAC <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
678	HexNAC <sub>1</sub> dHex <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
708*	Hex <sub>1</sub> HexNAC <sub>1</sub> dHex <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
738*	Hex <sub>2</sub> HexNAC <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
779	Hex <sub>1</sub> HexNAC <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
882	Hex <sub>1</sub> HexNAC <sub>1</sub> dHex <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
912*	Hex <sub>2</sub> HexNAC <sub>1</sub> dHex <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
923	HexNAC <sub>2</sub> dHex <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
942	Hex <sub>3</sub> HexNAC <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
953*	Hex <sub>1</sub> HexNAC <sub>2</sub> dHex <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
983	Hex <sub>2</sub> HexNAC <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1127	Hex <sub>1</sub> HexNAC <sub>2</sub> dHex <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1157*	Hex <sub>2</sub> HexNAC <sub>2</sub> dHex <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1198*	Hex <sub>1</sub> HexNAC <sub>3</sub> dHex <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1301	Hex <sub>1</sub> HexNAC <sub>2</sub> dHex <sub>3</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1331*	Hex <sub>2</sub> HexNAC <sub>2</sub> dHex <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1372	Hex <sub>1</sub> HexNAC <sub>3</sub> dHex <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1402	Hex <sub>2</sub> HexNAC <sub>3</sub> dHex <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1443*	Hex <sub>1</sub> HexNAC <sub>4</sub> dHex <sub>1</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1505	Hex <sub>2</sub> HexNAC <sub>2</sub> dHex <sub>3</sub>	■	■	■	■	■	■	■	■	■	■	■	■
1576	Hex <sub>2</sub> HexNAC <sub>3</sub> dHex <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
2067*	Hex <sub>2</sub> HexNAC <sub>5</sub> dHex <sub>2</sub>	■	■	■	■	■	■	■	■	■	■	■	■
2240	Hex <sub>2</sub> HexNAC <sub>5</sub> dHex <sub>3</sub>	■	■	■	■	■	■	■	■	■	■	■	■

Les quelques analyses MS<sup>2</sup> effectuées sur les *O*-glycannes perméthylés isolés du tube digestif de *C. gigas* ont permis d'identifier des *O*-glycannes de type mucine bâtis autour des noyaux 1 (*m/z* 534, 708, 738 et 953), 3 (*m/z* 1199) et 4 (*m/z* 1444 et 2067) ainsi que du noyau des invertébrés (*m/z* 738, 912, 1158 et 1332) (**Figure 48**).

De manière très intéressante, ces analyses révèlent la présence d'AgS A de type 1 pour les signaux  $m/z$  1199 et 2067 ainsi que des Ags A de type 3 pour les signaux  $m/z$  953, 1158 et 1332. En effet, comme pour les *N*-glycannes, l'ion C  $m/z$  679 correspondant au fragment trisaccharidique HexNAc[dHex]Hex-OH a été identifié sur les spectres de masse MS<sup>2</sup> de ces *O*-glycannes (**Figure 48B**). De plus, le spectre MS<sup>3</sup> de ce fragment dispose du couple d'ions B/Y  $m/z$  282/419 caractéristique de la perte du résidu d'HexNAc terminal et de l'ion Y à  $m/z$  490 correspondant à la perte du résidu de dHex, comme identifié précédemment pour les *N*-glycannes (**Figure 48B**). A cela s'ajoute la présence des fragments intracycles <sup>0,2</sup>X à  $m/z$  271 et <sup>0,3</sup>X à  $m/z$  301 qui démontre les liaisons  $\alpha$ 1-2 du résidu Fuc et  $\alpha$ 1-3 du résidu de GalNAc sur le résidu de Gal, respectivement. Parmi les *O*-glycannes identifiés, certains présentent deux Ags A ( $m/z$  2067) (**Figure 48C**). De plus, les signaux analysés en MS<sup>2</sup> font aussi apparaître des *O*-glycannes disposant de l'Ag H de type 3 pour les ions  $m/z$  708 et 912 (**Figure 48A**). Certains signaux qui n'ont pas été analysés en MS<sup>2</sup> et listés dans le **Tableau 4** ( $m/z$  1127, 1301, 1372, 1505, 1576 et 2240) laissent à penser que des Ag ALe<sup>b</sup> construits sur les noyaux de type 1 et 3 sont également présents sur les *O*-glycannes. Il conviendra de confirmer cela par les analyses MALDI-MS<sup>n</sup> adéquates.

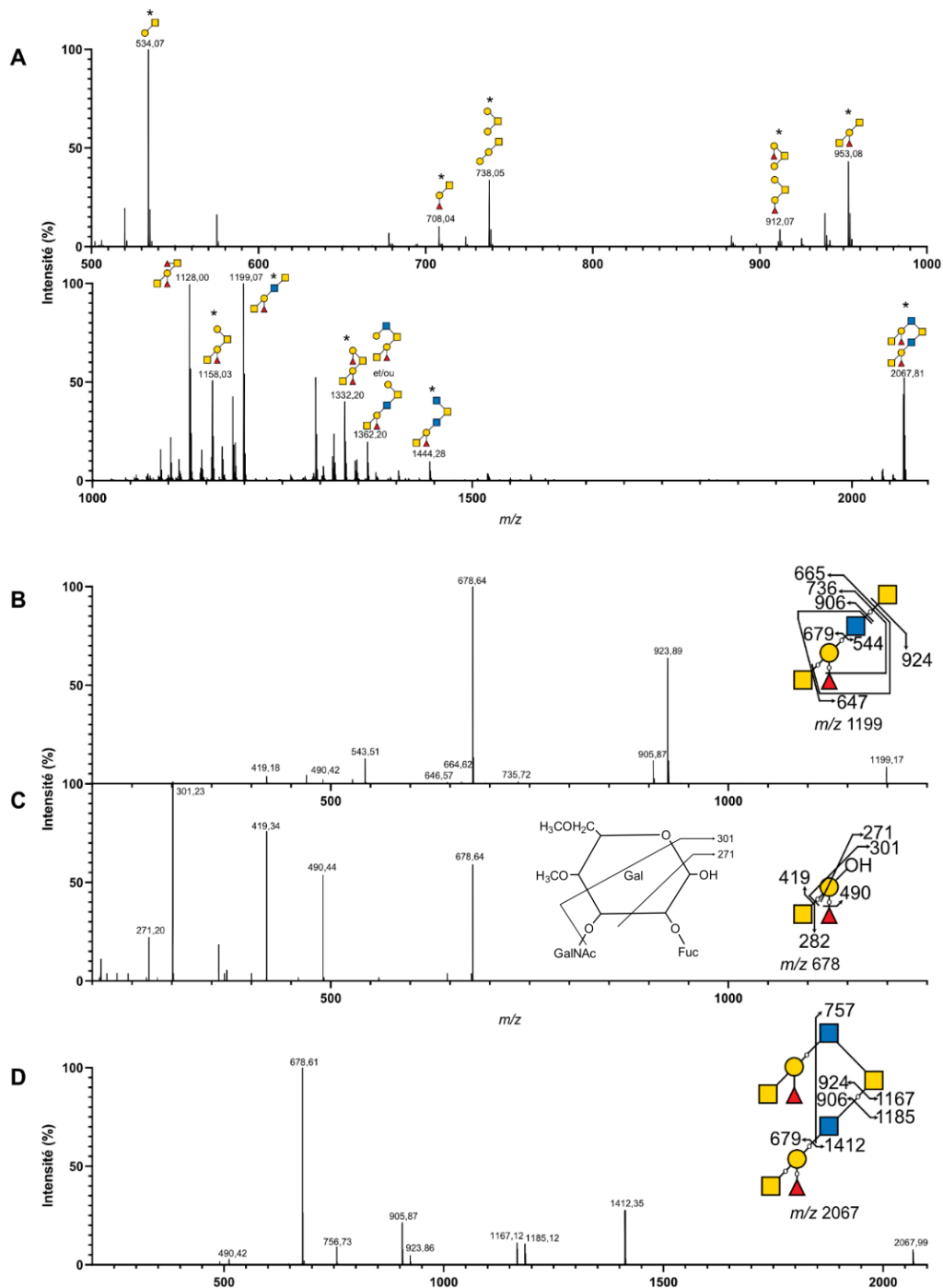
Il est à noter que d'autres signaux correspondants à du matériel oligosaccharidique (non répertoriés dans le tableau) ont été relevés. Une partie d'entre eux semble correspondre à des *N*-glycannes dont on sait qu'ils peuvent également être libérés durant le processus de  $\beta$ -élimination réductive. Enfin, d'autres signaux ont été détectés avec un  $m/z$  inférieur de 14 Da par rapport à la masse moléculaire théorique. Il pourrait s'agir soit de *O*-glycannes sous-méthylés (détectés à un  $m/z$  -14 par rapport au glycanne perméthylé) soit à un glycanne perméthylé renfermant un résidu de pentose ( $m/z$  160, *e.g.* le Xyl) à la place d'un résidu de Fuc ( $m/z$  174), soit une différence de 14 Da également.

Quant à l'analyse des spectres de masse MALDI-QIT-TOF obtenus après perdeutérométhylation de la fraction dessalée post  $\beta$ -élimination réductive elle révèle, d'une façon similaire aux *N*-glycannes, que bon nombre de *O*-glycannes portent des groupements méthyles, jusqu'à 4 pour certains *O*-glycannes (**Tableau 4**).

Pour compléter le travail structural entrepris sur les *O*-glycannes, il conviendra :

- d'analyser par MALDI-QIT-TOF les *O*-glycannes perdeutérométhylés des branchies des deux espèces d'huître,
- d'effectuer une analyse MALDI-MS<sup>2</sup> systématique de tous les signaux *m/z* pour lesquels nous avons proposé une composition brute afin de confirmer ou d'infirmer la structure proposée,
- d'analyser par MALDI-MS<sup>3</sup> tous les *O*-glycannes qui révéleront la présence d'un motif antigénique (Ags A, H, Le, etc.) afin de voir si ces motifs sont porteurs de groupements méthyles et dans l'affirmatif, de préciser sur quel(s) monosaccharide(s) se trouve(nt) le ou les groupement(s) méthyle(s).
- d'effectuer comme pour notre étude sur les *N*-glycannes, les mêmes analyses GC-EI-MS afin de vérifier les compositions en monosaccharides déduites des analyses MALDI-MS, identifier les monosaccharides méthylés et pour chacun d'eux, localiser le carbone porteur du groupement méthyle.

Ainsi, l'ensemble des résultats pourront permettre d'effectuer des comparaisons à la fois entre tissus mais aussi entre espèces d'huîtres.

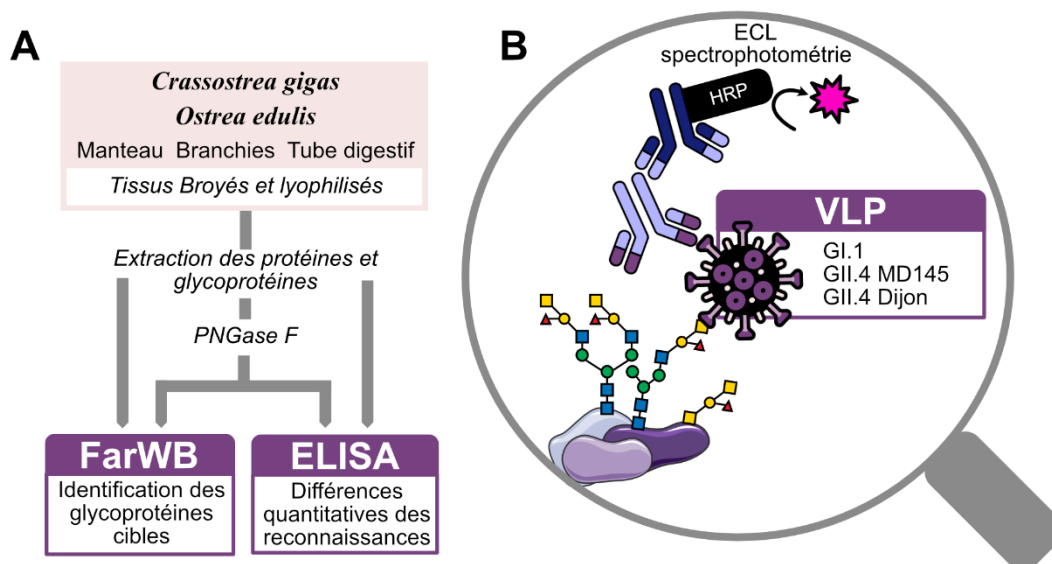


**Figure 48** : Spectres de masse MALDI-MS, MS<sup>2</sup> et MS<sup>3</sup> des *O*-glycannes réduits et perméthylés obtenus à partir du tube digestif de l’huître *Crassostrea gigas*. **(A)** Spectre MS du *O*-glycome du tube digestif. La présence d’un astérisque (\*) sur un signal *m/z* indique que la structure *O*-glycannique a été déduite à partir des analyses MALDI-MS et MS<sup>2</sup>. **(B)** Spectre MS<sup>2</sup> du *O*-glycane *m/z* 1199. **(C)** spectre MS<sup>3</sup> de l’ion *m/z* 678 généré durant l’analyse MS<sup>2</sup> du *O*-glycane *m/z* 1199. **(D)** Spectre MS<sup>2</sup> du *O*-glycane *m/z* 2067. Les structures *O*-glycanniques sont représentées selon la Symbol Nomenclature for Glycans (Neelamegham *et al.* 2019).

## Chapitre 2 : ETUDE DES INTERACTIONS ENTRE LES NOROVIRUS ET LES GLYCANNES ISSUS DES HUITRES *C. gigas* ET *O. edulis*

### I. Stratégies développées

En parallèle des études structurales des *N*- et *O*-glycannes des différents tissus des huîtres *C. gigas* et *O. edulis*, nous avons initié une étude visant à établir le tropisme tissulaire et d'espèce de 2 génotypes de NoVs, GI.1 et GII.4, en élaborant la stratégie présentée dans la **Figure 49**. Des expériences de Far-Western blot (FarWB) et d'ELISA (Enzyme-Linked ImmunoSorbent Assay) ont été mises en œuvre sur les protéines et GPs extraites des 3 tissus des 2 espèces d'huître.



**Figure 49** : Stratégie employée pour l'étude des interactions entre les Norovirus (NoV) GI.1, GII.4 MD145 et GII.4 Dijon et les glycoprotéines des 3 tissus des huîtres *Crassostrea gigas* et *Ostrea edulis*. (A) Les protéines et glycoprotéines ont été extraites des 3 tissus des huîtres *C. gigas* et *O. edulis* puis une partie a été soumise à une déglycosylation partielle par l'action de la PNGase F. Les protéines et glycoprotéines natives et ayant subi l'action de la PNGase F ont ensuite été analysées par des expériences de FarWB et d'ELISA. (B) Les VLPs sont détectées par un anticorps primaire spécifique de la souche de NoV puis un anticorps secondaire anti-anticorps marqué à la Horse Radish Peroxidase (HRP) est employé. La révélation est réalisée par électrochimiluminescence (ECL) pour les expériences de FarWB et par précipitation par le tétraméthylbenzidine (TMB) pour les tests ELISA.

Pour les expériences de FarWB, l'objectif était de visualiser les différentes protéines et GPs reconnues par les VLPs et de les comparer entre tissus et espèces d'huître tandis que les tests ELISA avaient pour objectif de quantifier ces reconnaissances et également de les comparer.

Pour ces expériences, les protéines et GPs ont été extraites par un tampon d'extraction contenant du sodium dodecyl sulfate (SDS) et la moitié du lysat cellulaire obtenu a subi une déglycosylation partielle par la PNGase F. Pour les expériences de FarWB, les échantillons natifs et ceux ayant subi la PNGase F ont ensuite été soumis à un SDS-PAGE. Après leur séparation sur gel, les protéines et GPs ont été transférées sur une membrane de nitrocellulose et incubées en présence des VLPs de NoVs GI.1, GII.4 Dijon ou GII.4 MD145. Après plusieurs lavages, les VLPs fixées ont été visualisées par électrochimiluminescence (ECL) ou par spectrophotométrie d'absorption en utilisant un anticorps spécifique primaire anti-VLP et un anticorps secondaire anti-anticorps marqué à l'HRP. Pour les tests ELISA, les mêmes échantillons ont été déposés dans les puits d'une plaque ELISA 96 puits à différentes dilutions. Pour le reste, les étapes sont les mêmes que celles décrites ci-dessus pour le FarWB.

L'ensemble de ces manipulations a été effectuée lors d'un stage de 6 semaines au sein de l'équipe du Dr Jacques Le Pendu au Centre de Recherche en Cancérologie et Immunologie Nantes-Angers (CRCINA, UMR Inserm 1232) en utilisant les VLPs qu'ils avaient produites.

## II. Mise au point préalable des conditions opératoires des expériences de FarWB et d'ELISA

Avant d'appliquer la stratégie décrite ci-dessus, nous avons cherché à optimiser chacune des étapes en commençant par la préparation des échantillons comprenant le choix du tampon d'extraction des protéines et GPs ainsi que les conditions de déglycosylation par la PNGase F. Puis pour les expériences de FarWB et d'ELISA, nous avons déterminé les quantités de protéines et GPs optimales à utiliser, les conditions de saturation des membranes de nitrocellulose ou des puits des plaques ELISA ainsi que les concentrations optimales de lectines et de VLPs à employer.

### A. Préparation des échantillons

#### 1. Choix du tampon d'extraction des protéines et glycoprotéines

Le choix du meilleur tampon d'extraction des protéines des 3 tissus utilisés a été notre premier point d'attention. Deux protocoles couramment utilisés au laboratoire ont ainsi été testés. Le premier utilise le tampon commercial RIPA (RadioImmunoPrecipitation Assay) et le second un tampon bicarbonate d'ammonium 50 mM contenant 0,25 % de SDS, 0,25 % de  $\beta$ -mercaptoéthanol et des

inhibiteurs de protéases et que nous nommerons « tampon SDS » dans la suite du manuscrit. Les résultats obtenus par lectine blots avec la lectine AAL se sont avérés similaires quel que soit le tampon d'extraction utilisé. Néanmoins, au regard des tests d'inhibition de reconnaissance que nous avons mis en œuvre en utilisant la PNGase F pour éliminer les *N*-glycannes (voir ci-après), il nous est apparu que le tampon SDS était plus adéquat (pH similaire à celui du tampon habituellement utilisé).

## 2. Contrôle de l'efficacité d'action de la PNGase F

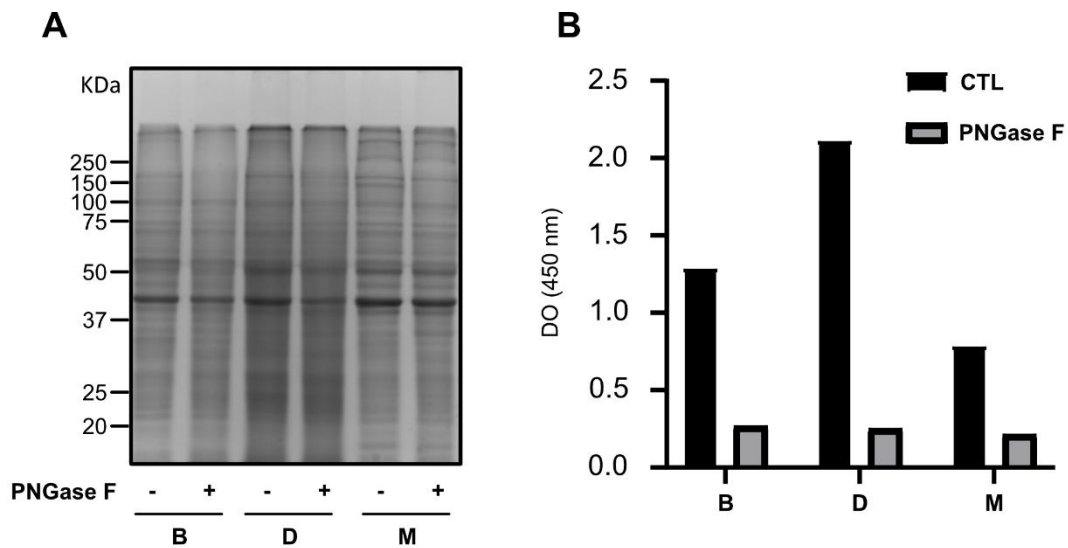
Dans le but de mesurer l'implication des *N*-glycannes dans la reconnaissance des différents tissus de *C.gigas* et *O.edulis* par les NoVs, nous avons entrepris d'établir un protocole de libération des *N*-glycannes avec la PNGase F pour les manipulations de FarWB et d'ELISA.

Les protéines et GPs ayant été extraites par du SDS, nous avons utilisé le Nonidet P40, un détergent anionique, pour neutraliser l'excès de SDS qui, sans cela, aurait inhibé la PNGase F (Morelle et Michalski, 2007). Nous avons d'abord fait agir la PNGase F dans les mêmes conditions que celles utilisées pour l'établissement du *N*-glycome à savoir 62,5 UI de PNGase F durant une nuit entière à 37 °C. La coloration au bleu de Coomassie du gel SDS-PAGE effectué à l'issue de cette étape a montré une forte dégradation protéique malgré la présence d'un cocktail d'inhibiteurs de protéases. Pour diminuer cette dégradation, nous avons donc décidé d'utiliser une quantité plus importante de PNGase F (500 UI plutôt que 62,5 UI) et de ramener le temps d'incubation à 37 °C de 16h à 1h. Comme précédemment, un gel SDS-PAGE coloré au bleu de coomassie a été effectué. Comme le montre la **Figure 50A**, une très faible dégradation protéique entre les tissus natifs et ceux ayant subi l'action de la PNGase F est observée. Les quelques bandes protéiques qui disparaissent après traitement à la PNGase F pourraient être le résultat de l'élimination de leurs *N*-glycannes qui se traduit par une modification de leur poids moléculaire et donc une accélération de leur migration électrophorétique.

Pour vérifier l'élimination des *N*-glycannes par la PNGase F, un test ELISA a ensuite été effectué en utilisant la Concanavoline A (Con A) comme sonde. Cette lectine qui reconnaît les résidus de Man a été choisie dans le but de cibler les *N*-glycannes de type oligomannosidique. L'histogramme présenté dans la **Figure 50B** démontre sans ambiguïté l'action de la PNGase F dans les conditions que



nous avons établies et qui se traduit par une diminution du signal après action de la PNGase F d'un facteur allant jusqu'à 10.



**Figure 50** : Evaluation du protocole d'élimination des *N*-glycannes par la PNGase F. (A) SDS-PAGE coloré au bleu de Coomassie et (B) test ELISA utilisant la lectine Con A effectué avec 2,5 µg/mL de protéines et glycoprotéines extraites des branchies (B), du tube digestif (D) et du manteau (M) de l'huître *Crassostrea gigas* avant et après traitement à la PNGase F.

## B. Mise au point des différentes étapes des expériences de FarWB

Les quantités de VLPs dont nous disposons étant limitées, nous avons réalisés ces mises au point en utilisant des lectines pour mimer les reconnaissances par les VLPs. Pour ce faire, nous avons utilisé 3 lectines : la lectine AAL connue pour reconnaître le Fuc quelle que soit la liaison qu'il engage, la lectine *Ulex europaeus* (UEA-I) qui reconnaît les résidus de Fuc liés en  $\alpha$ 1-2 comme cela est le cas pour les Ags A et H et la lectine HPA qui reconnaît les résidus de GalNAc et donc l'Ag A ou ALe<sup>b</sup>.

### 1. Concentrations en lectines

Les 3 lectines citées précédemment étaient marquées à la biotine et ont été utilisées selon les concentrations préconisées par le fournisseur à savoir 1 µg/mL à l'exception de l'AAL que nous avons utilisée à une concentration de 0,1 µg/mL du fait de sa très forte réactivité. Puis, des anticorps anti-biotine marqués à la peroxydase de raifort (HorseRadish Peroxydase, HRP) ont été utilisés pour les révéler spécifiquement.

## 2. Détermination des quantités optimales des protéines et glycoprotéines

Concernant les quantités optimales de protéines à déposer sur les gels SDS-PAGE, nous avons déterminé que 50 µg de protéines étaient appropriés pour une coloration au bleu de Coomassie et que 20 µg de protéines suffisaient pour la réalisation de lectine blots. La coloration au bleu de Coomassie avait pour but de vérifier l'intégrité des protéines et le bon déroulement de la migration avant d'effectuer les lectine blots. A titre d'exemple, le gel présenté dans la **Figure 51** montre une bonne séparation des protéines issues des branchies, du tube digestif et du manteau des huîtres *C. gigas* et *O. edulis*, et atteste d'une absence de dégradation protéique lors de la préparation des échantillons. Dans le cas contraire, des bandes mal résolues et une révélation sous forme de trainées auraient été observées.

## 3. Choix de l'agent de saturation des membranes de nitrocellulose

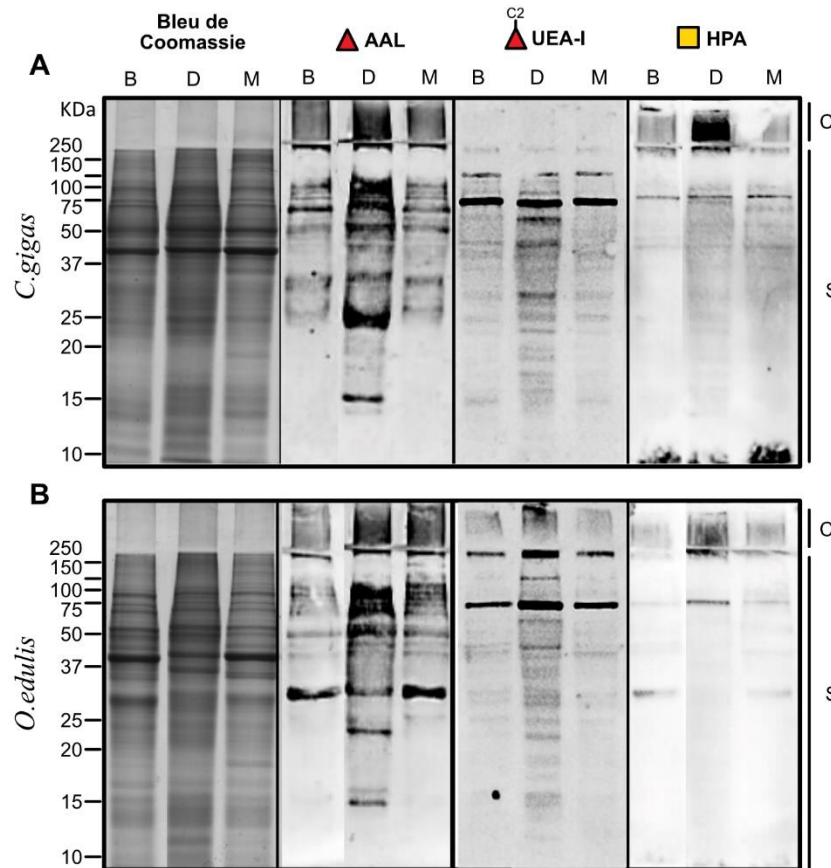
Classiquement, dans ce type d'expérience, le lait en poudre est largement utilisé. Néanmoins, les nombreuses études effectuées sur les oligosaccharides libres et sur les GPs du lait ont révélé la présence de très nombreux motifs antigéniques susceptibles d'entrer en compétition avec les GPs de nos échantillons lors de la reconnaissance par les lectines (Koromyslova *et al.*, 2017 ; Schrotten *et al.*, 2016 ; Weichert *et al.*, 2016). Pour cette raison, nous avons choisi d'utiliser la sérum albumine bovine (BSA) qui est une holoprotéine.

## 4. Résultats obtenus avec le protocole optimisé

### a. Révélation avec les lectines

Appliqué aux protéines et GPs extraites des 3 tissus des 2 espèces d'huître, notre protocole optimisé a conduit aux gels et lectine blots présentés dans la **Figure 51**. Comme attendu, la révélation par la lectine AAL conduit à un très fort signal pour l'ensemble des échantillons. Ce résultat est en accord avec le fort taux de fucosylation des glycanes mis en lumière lors de l'établissement des *N*- et *O*-glycomes. On note également un signal de reconnaissance moindre pour les lectines UEA-I et HPA moindre mais tout à fait acceptable qui confortent l'existence d'Ag A sur les *N*- et *O*-glycanes des huîtres *C. gigas* et *O. edulis*. Malgré les quantités similaires de protéines déposées pour les 3 tissus des 2 espèces d'huîtres, une reconnaissance plus intense pour le tube digestif est observée au niveau du gel de séparation. De manière très surprenante, le même résultat est observé pour des GPs présentes au

niveau du gel de concentration et qui atteste de leur très haut poids moléculaire. En parallèle, dans cette zone, on note que l'intensité de coloration par le bleu de coomassie est très faible. Ce résultat laisse à penser que les macromolécules qui restent dans le gel de concentration pourraient correspondre à des GPs de très haut poids moléculaire et très fortement glycosylées. En effet, on sait qu'un taux de glycosylation très élevé peut empêcher la coloration du squelette protéique par le bleu de Coomassie.



**Figure 51** : Lectine blots des glycoprotéines des huîtres *Crassostrea gigas* et *Ostrea edulis*. (A) Les protéines du gel SDS-PAGE des branchies (B), tube digestif (D) et manteau (M) de *C. gigas* et (B) *O. edulis* ont été (**Panneau 1**) colorées au bleu de Coomassie ou transférées sur une membrane de nitrocellulose suivi d'une révélation utilisant les lectines (**Panneau 2**) AAL, (**Panneau 3**) UEA-I et (**Panneau 4**) HPA marquées à la biotine. Les lettres C et S indiquent les zones de gels de concentration (4 %) et de (10 %) séparation, respectivement.

#### b. Révélation avec les VLPs

Nous avons réalisé l'ensemble des manipulations en utilisant une concentration en VLPs de 3 µg/mL couramment utilisée dans le laboratoire du Dr Jacques Le Pendu. Comparés aux tests de reconnaissances effectués avec les lectines, l'utilisation des VLPs a nécessité un changement concernant l'agent de saturation des membranes de nitrocellulose. En effet, la BSA utilisée comme agent de

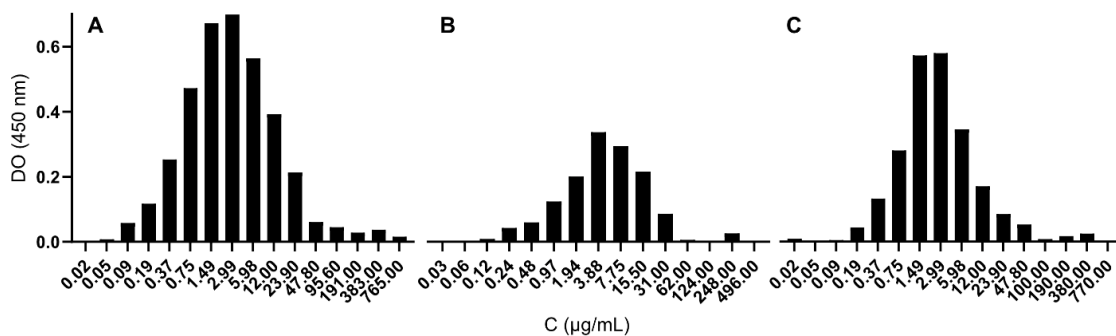
saturation pour l'utilisation avec les lectines n'a pas permis d'obtenir de résultats lorsqu' utilisée avec les VLPs. En dépit des réticences que nous avons vis-à-vis de l'utilisation du lait en poudre comme agent de saturation des membranes ou des puits d'ELISA, nous l'avons testé à la place de la BSA. Contre toute attente, nous avons obtenu de bien meilleurs résultats malgré la présence d'un grand nombre d'oligosaccharides libres et GPs dans le lait qui pouvait laisser craindre un bruit de fond élevé mais cela n'a pas été le cas. Dans le but de vérifier la spécificité de reconnaissance des VLPs vis-à-vis des Ags A, la salive d'individu sécréteur du groupe A et d'individu non sécréteur a été utilisée comme contrôle positif et négatif, respectivement. Les résultats obtenus ont été conformes à notre attente et seront présents dans chaque FarWB présentés dans la partie III.

### C. Mises au point pour les tests ELISA

Pour la mise au point du protocole des tests ELISA, nous avons décidé d'effectuer les manipulations de reconnaissance directement avec les VLPs car nous avons choisi de déployer le protocole utilisé en routine dans l'équipe du Dr Jacques Le Pendu. Ainsi, nous avons effectué les manipulations avec une solution à 1 µg/mL en VLPs et nous avons réalisé la saturation des puits par la BSA.

Pour la quantité de protéines d'échantillons d'huître à utiliser, nous avons effectué des dilutions en série mais les premiers résultats obtenus se sont avérés très surprenants. En effet, au-delà d'une certaine concentration en matériel protéique, les absorbances diminuaient alors même que la concentration en matériel protéique augmentait comme illustré dans les graphes en forme de courbes de Gauss présentés (Figure 52). Cela a été observé quelles que soient les VLPs, GI.1 et GII.4 MD145, utilisés. En théorie, nous aurions dû atteindre un plateau. Lors d'une discussion avec le Dr Jacques le Pendu, ce dernier m'a indiqué qu'il n'y avait pas d'explications rationnelles pour ce phénomène mais que ce résultat était assez classique lorsqu'on s'adresse à des GPs de très haut poids moléculaire (Argüeso et Gipson, 2006). La diminution du signal lors de l'augmentation de la concentration en protéines déposée pourrait s'expliquer par des phénomènes d'agrégation entre GPs lorsque l'on en dépose trop dans les puits et qui aurait pour conséquence de masquer les Ags.

Dans le but d'obtenir des réponses optimales, nous avons déterminé la dilution la plus adéquate des échantillons protéiques et glycoprotéiques à réaliser pour chacun des 3 tissus des 2 espèces d'huître dans le cas d'expériences de reconnaissance par les VLP GI.1 et GII.4 MD145. C'est ainsi que pour les VLPs GI.1 nous avons fixé les concentrations en protéines et GPs à 2,5 µg/mL pour les 3 tissus des 2 huîtres et pour les VLPs GII.4 MD145 nous avons utilisé des concentrations de 4, 1,75 et 3 µg/mL pour les branchies, le tube digestif et le manteau, respectivement, uniquement chez l'huître *C. gigas*.



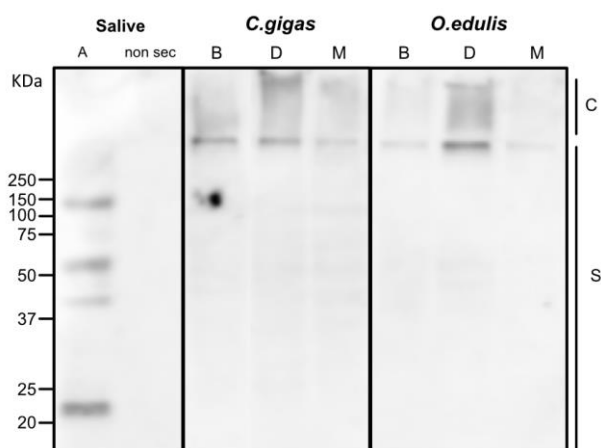
**Figure 52 :** Tests ELISA utilisant des VLPs à des concentrations croissantes en protéines du tube digestif des huîtres *Crassostrea gigas* et *Ostrea edulis*. (A) Les VLPs GI.1 ont été utilisées sur le tube digestif de *C. gigas* et (B) *O. edulis*. (C) Les VLPs GII.4 MD145 ont été utilisées sur le tube digestif de *C. gigas*.

### III. Reconnaissance des ligands glycaniques de l'huître par les Norovirus : utilisation de VLPs

Pour mimer les reconnaissances des GPs des différents tissus de *C. gigas* et *O. edulis* par les NoVs, nous avons utilisés des VLPs de NoVs GI.1 Norwalk et des souches GII.4 Dijon (variant US95/96) et MD145 (variant Bristol). Les deux souches de GII.4 se distinguent par leur année d'émergence : 1987 pour GII.4 MD145 et 1996 pour GII.4 Dijon. Du reste, elles présentent des profils de reconnaissances très similaires, notamment vis-à-vis de l'Ag H (de Rougemont *et al.*, 2011).

#### A. Etudes des reconnaissances par les VLPs de Norovirus GI.1

Le FarWB réalisé avec les VLPs GI.1 sur les protéines et GPs des branchies, du tube digestif et du manteau des huîtres *C. gigas* et *O. edulis* montre une reconnaissance de protéines et GPs de très haut poids moléculaire au niveau du gel de concentration et à l'interface des deux gels pour les deux huîtres (Figure 53). La comparaison des signaux de reconnaissance entre les tissus de chacune des huîtres indique une reconnaissance plus intense pour le tissu digestif. De plus, on note des signaux plus faibles pour les branchies et le manteau de *O. edulis* que ceux de *C. gigas*.

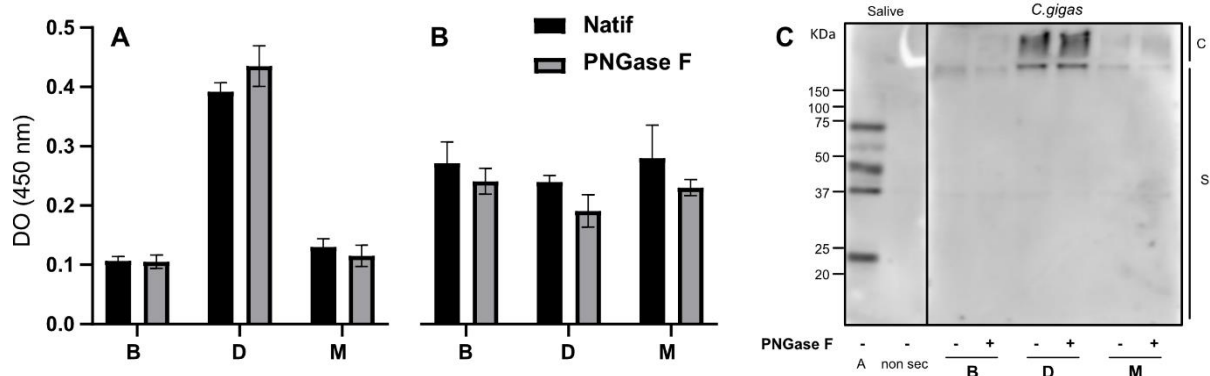


**Figure 53** : Far-western blot des protéines et glycoprotéines des huîtres *Crassostrea gigas* et *Ostrea edulis* avec les VLPs GI.1 comme sonde. (**Panneau 1**) Les contrôles positif et négatif se composent de 10  $\mu$ L de salive d'individu sécréteur de groupe A et non sécréteur. Les protéines et glycoprotéines (20  $\mu$ g) des branchies (B), du tube digestif (D) et du manteau (M) de (**Panneau 2**) *C. gigas* et (**Panneau 3**) *O. edulis* ont été déposées sur un gel SDS-PAGE puis transférées sur une membrane de nitrocellulose avant d'être incubées avec les VLPs GI.1 utilisées comme sonde (voir matériel et méthodes). Les lettres C et S indiquent les zones de gels de concentration (4%) et de séparation (10%), respectivement.

Pour le test ELISA réalisé sur les différents tissus de l'huître *C. gigas*, la reconnaissance par les VLPs GI.1 apparaît 4 fois plus élevée pour le tube digestif (Abs 0,4) comparé aux branchies (Abs 0,1) et au manteau (Abs 0,1) (**Figure 54A**). Ces résultats sont en accord avec ceux obtenus pour les expériences précédentes de FarWB ci-dessus (**Figure 53**) à savoir une reconnaissance majoritaire du tube digestif. En revanche, pour l'huître *O. edulis*, la reconnaissance par les VLPs GI.1 apparaît similaire pour les trois tissus avec une absorbance voisine de 0,3 (**Figure 54B**). Ce résultat n'est pas en accord avec celui obtenu pour l'expérience de FarWB où seul le tube digestif était préférentiellement reconnu. Cela pourrait être expliqué par une présentation différente des Ags des GPs selon l'expérience employée. Une autre explication pourrait être que certaines GPs insolubles ou de trop haut poids moléculaire ne soient pas entrées dans le gel d'électrophorèse pour l'expérience de FarWB alors qu'elles seraient fixées dans les puits pour les tests ELISA. Cela conduirait à des populations de GPs différentes en fonction de l'expérience employée.

Dans le but de voir si les *N*-glycannes des GPs isolées des 3 tissus pour les 2 espèces d'huître étaient impliqués dans la reconnaissance par les VLPs GI.1, nous avons réalisé une déglycosylation par la PNGase F dans les conditions énoncées dans la partie II.A.2 p156. Que ce soit les expériences de FarWB ou d'ELISA, les résultats obtenus ne montrent aucune diminution des reconnaissances par les

VLPs GI.1 après déglycosylation partielle par la PNGase F (Figure 54). Cela suggère que les *N*-glycannes des GPs ne sont pas ou très peu impliqués dans la reconnaissance par les NoVs GI.1.



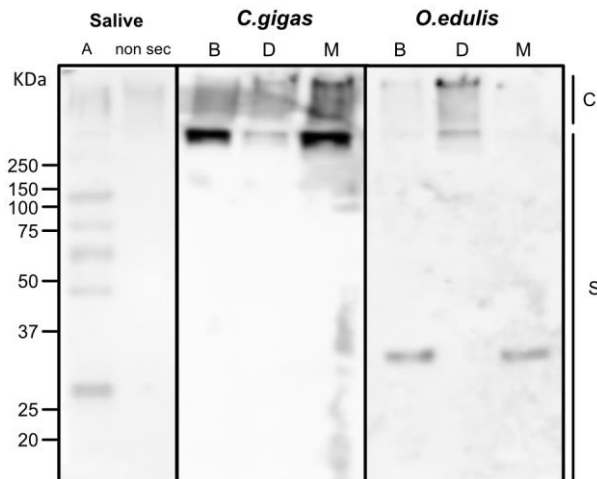
**Figure 54 :** Reconnaissance des glycoprotéines et protéines des tissus de *Crassostrea gigas* et *Ostrea edulis* par les VLPs GI.1 avant et après traitement par la PNGase F. (A) Test ELISA des protéines et glycoprotéines des branchies (B), du tube digestif (D) et du manteau (M) de *C. gigas* et (B) *O. edulis* utilisant les VLPs GI.1 comme sonde. Les barres noires correspondent aux échantillons non traités à la PNGase F et les barres grises, aux échantillons traités. (C) (Panneau 2) Far-western blot des protéines des tissus de *C. gigas*. Les protéines et glycoprotéines des 3 tissus (20 µg) ont été déposées sur un gel SDS-PAGE, transférées sur une membrane de nitrocellulose et incubées avec les VLPs GI.1 comme sonde (voir matériel et méthodes). (Panneau 1) Les contrôles positif et négatif se composent de 10 µL de salive d'individu sécréteur de groupe A et non sécréteur. Les lettres C et S indiquent les zones de gels de concentration (4 %) et de séparation (10 %).

A la lumière de ces résultats, si l'on considère, d'une part, que la reconnaissance des GPs par les VLPs GI.1 concernent très majoritairement des GPs de très haut poids moléculaire et fortement glycosylées et, d'autre part, que les *N*-glycannes sont peu voire pas impliqués dans cette reconnaissance, mon hypothèse est que ces GPs de très haut poids moléculaire pourraient être des mucines ou des GPs mucines-like. Ces molécules étant très riches en *O*-glycannes, cela pourrait signifier que ces derniers soient le siège de la reconnaissance par les NoVs GI.1.

## B. Etudes des reconnaissances par les VLPs de Norovirus GII.4 Dijon

En parallèle des études de reconnaissance des protéines et GPs par les VLPs GI.1, nous avons mené une étude similaire en utilisant des VLPs de NoVs GII.4 Dijon. La Figure 55 présente les résultats du FarWB obtenu avec ces VLPs. D'une façon similaire aux résultats obtenus avec les VLPs GI.1, on remarque que la reconnaissance par les VLPs GII.4 Dijon est essentiellement localisée au niveau du gel de concentration et à l'interface des 2 gels (concentration et séparation). Néanmoins, pour *C. gigas*, les 3 tissus sont reconnus par ces VLPs avec une intensité plus importante pour le manteau et les branchies, notamment à l'interface des 2 gels. Ce résultat contraste avec celui obtenu avec les VLPs GI.1 qui

reconnaissaient très fortement les protéines et GPs du tube digestif. Pour *O. edulis*, on obtient le même modèle de reconnaissance que celui obtenu avec GI.1 à savoir une reconnaissance des protéines et GPs isolées du tube digestif. Il est également à noter qu'une bande d'environ 30 KDa est révélée par les VLPs GII.4 Dijon pour les branchies et le manteau d'*O.edulis*, ce qui n'était pas le cas avec les VLPs GI.1.



**Figure 55 :** Far-western blot des protéines et glycoprotéines des huîtres *Crassostrea gigas* et *Ostrea edulis* avec les VLPs GII.4 Dijon comme sonde. (**Panneau 1**) Les contrôles positif et négatif se composent de 10  $\mu$ L de salive d'individu sécréteur de groupe A et non sécréteur. Les protéines et glycoprotéines (20  $\mu$ g) des branchies (B), du tube digestif (D) et du manteau (M) de (**Panneau 2**) *C. gigas* et (**Panneau 3**) *O. edulis* ont été déposées sur un gel SDS-PAGE puis transférées sur une membrane de nitrocellulose avant d'être incubées avec les VLPs GII.4 Dijon utilisées comme sonde (voir matériel et méthodes). Les lettres C et S indiquent les zones de gels de concentration (4%) et de séparation (10%), respectivement.

Du fait d'une quantité insuffisante de VLPs GII.4 Dijon, nous n'avons pas pu effectuer d'autres manipulations. En revanche, nous disposons de VLPs d'une autre souche de GII.4, les VLPs GII.4 MD145 et dont les résultats sont présentés ci-dessous.

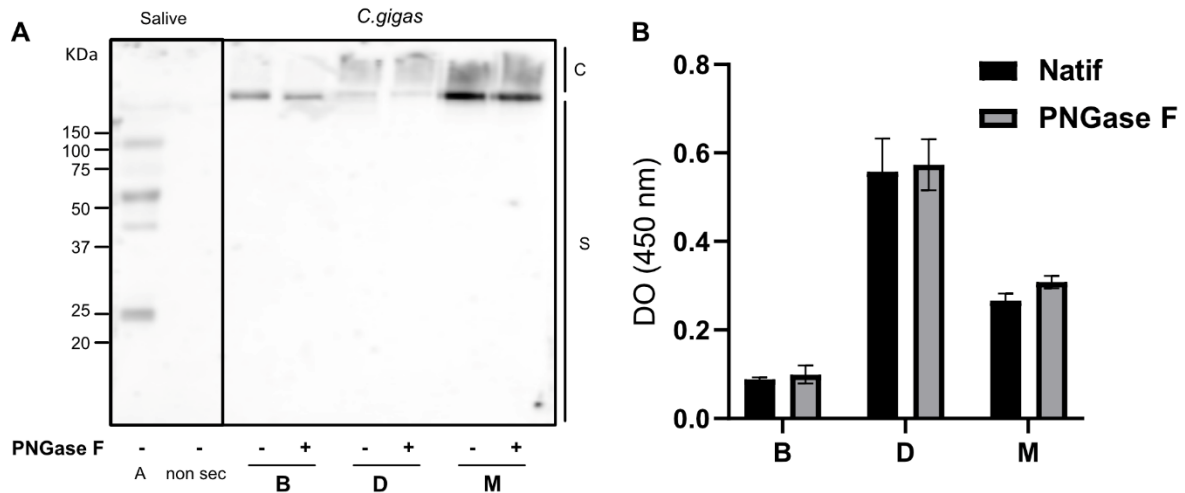
### C. Etudes des reconnaissances par les VLPs de Norovirus GII.4 MD145

Comme pour les 2 types de VLPs précédents, les VLPs du variant GII.4 MD145 ont été utilisées pour révéler les GPs et protéines lors d'expériences de FarWB et d'ELISA sur les 3 tissus de *C. gigas*. Par manque de temps, l'espèce *O. edulis* n'a pas pu être étudiée avec ces VLPs.

Le FarWB révèle une reconnaissance des protéines et GPs localisées au niveau du gel de concentration pour les 3 tissus. Néanmoins, on note que les protéines et GPs obtenues à partir du manteau sont plus fortement reconnues par les VLPs GII.4 MD145 (Figure 56A) que celles des 2 autres tissus.



Les tests ELISA n'ont malheureusement pas confirmés cette préférence de reconnaissance par le VLPs GII.4 MD145. En effet, ils ont révélé une reconnaissance plus importante pour les protéines et GPs présentes dans le tube digestif (Abs 0,56), que celles présentes dans le manteau (Abs 0,27) ou dans les branchies (Abs 0,09).



**Figure 56 :** Reconnaissance des glycoprotéines et protéines des tissus de *Crassostrea gigas* par les VLPs GII.4 MD145 avant et après traitement par la PNGase F. **(A) (Panneau 2)** Far-western blot des protéines des tissus de *C. gigas*. Les protéines et glycoprotéines des 3 tissus (20 µg) ont été déposées sur un gel SDS-PAGE, transférées sur une membrane de nitrocellulose et incubées avec les VLPs GII.4 MD145 comme sonde (voir matériel et méthodes). **(Panneau 1)** Les contrôles positif et négatif se composent de 10 µL de salive d'individu sécréteur de groupe A et non sécréteur. Les lettres C et S indiquent les zones de gels de concentration (4 %) et de séparation (10 %). **(B)** Test ELISA des protéines et glycoprotéines des branchies (B), du tube digestif (D) et du manteau (M) de *C. gigas* utilisant les VLPs GII.4 MD145 comme sonde. Les barres noires correspondent aux échantillons non traités à la PNGase F et les barres grises, aux échantillons traités.

Comme pour les protéines et GPs isolés des tissus d'*O. edulis* et révélées par les VLPs GI.1 (Partie A p161), les VLPs GII.4 MD145 ne présentent pas les mêmes reconnaissances tissulaires selon qu'elles soient utilisées en FarWB ou en ELISA.

Enfin, lorsqu'on compare les intensités de reconnaissances avant et après action de la PNGase F, on n'observe pas de changement majeur, tout au plus une très légère diminution des bandes après action de la PNGase F. Comme précédemment, ce résultat semble indiquer que les *N*-glycannes ne constituent pas le siège essentiel des interactions avec les NoVs et que ce rôle pourrait être joué par les *O*-glycannes de GPs de haut poids moléculaire de type mucine.

# DISCUSSION

---

# I. La glycosylation des huîtres *C. gigas* et *O. edulis* parmi les invertébrés

## A. La N- et O-glycosylation

Le point attendu lors de cette étude était la caractérisation structurale de motifs HBGAs chez les huîtres *C. gigas* et *O. edulis* et en particulier d'Ags A et H pour venir conforter les multiples articles relatant leur présence dans les tissus de *C. gigas* sur la base de reconnaissances par des anticorps anti-A et anti-H et leur implication dans la reconnaissance huître-NoV (Le Guyader *et al.*, 2006 ; Maalouf *et al.*, 2010b ; Morozov *et al.*, 2018 ; Su *et al.*, 2018 ; Tian *et al.*, 2007).

Les N- et O-glycomes que nous avons établis pour les 3 tissus étudiés chez les deux espèces d'huître ont définitivement mis en évidence la présence d'Ags A et H. De surcroît, des motifs antigéniques ALe<sup>b</sup> chez les 2 huîtres et HLe<sup>b</sup> uniquement chez *O. edulis* ont été décrits. La présence d'Ag Le<sup>b</sup> est partiellement en accord avec les travaux de Su *et al.* (2018) qui ont révélé la présence de cet Ag mais uniquement dans les branchies alors que nos résultats montrent que cet Ag est présent sous forme ALe<sup>b</sup> dans les 3 tissus. La présence d'Ags A a également été démontrée pour les huîtres *C. virginica* et suggérée chez *Crassostrea sikamea* ce qui semble constituer une caractéristique commune aux huîtres (Kurz *et al.*, 2013 ; Tian *et al.*, 2007). Le séquençage du génome de l'huître *C. gigas* révèle la présence d'un gène codant la galactoside 2- $\alpha$ -L-fucosyltransférase 1 (FUT1) dans les différents tissus et dont l'activité a été démontrée (Gui *et al.*, 2023 ; Zhang *et al.*, 2012). Le gène présente beaucoup de régions très similaires à celui du gène *FUT1* des vertébrés. De plus, l'enzyme est plus proche phylogénétiquement des enzymes FUT1 et FUT2 des vertébrés que d'autres fucosyltransférases comme celles FUT3 à FUT11 de l'Homme impliquées pour la plupart dans les Ags de Lewis (Gui *et al.*, 2023). L'enzyme galactoside 2- $\alpha$ -L-fucosyltransférase 2 (FUT2), quant à elle, n'a pas encore été étudiée mais son gène a été annoté lors de la caractérisation du génome de *C. gigas* (GenBank : XP\_011430557.2). En plus des HBGAs, nous avons également identifié chez *O. edulis* l'antigène de Galili dont la structure est Gal $\alpha$ 1-3Gal. Cet antigène est surtout retrouvé chez les primates du Nouveau Monde et de nombreux mammifères mais il est absent chez l'homme (Macher et Galili, 2008). Ce motif oligosaccharidique est immunogène chez l'homme et entraîne la production d'anticorps

anti-Gal $\alpha$ 1-3Gal. En aparté, sa présence dans les organes porcins et de primates empêche leur utilisation pour les xénotransplantations en provoquant des rejets (Lin *et al.*, 2000 ; Macher et Galili, 2008).

Les *N*-glycannes que nous avons identifiés chez les huîtres *C. gigas* et *O. edulis* partagent plusieurs caractéristiques structurales fréquemment retrouvées chez les invertébrés (présentées dans l'Introduction, Chap. 1, Partie V.A p48). Ainsi, sur le résidu de GlcNAc du point d'attache, nous avons identifié la présence d' $\alpha$ 1-3Fuc qui n'est pas retrouvé chez les vertébrés mais présent chez *C. elegans* tout comme chez l'huître *C. virginica* (Kurz *et al.*, 2013 ; Natsuka *et al.*, 2002), d'un disaccharide HexFuc ou FucHex qui pourrait correspondre au motif GalFuc également identifié chez *C. elegans* (Yan *et al.*, 2018) et d'un Hex qui pourrait correspondre à un résidu de Man comme chez l'escargot *Volvarina rubella* (Eckmair *et al.*, 2016) et comme également suspecté chez *C. virginica* (Kurz *et al.*, 2013). Sur le Man central du noyau, nos analyses ont également montré la présence d'un  $\beta$ 1-2Xyl pour les 2 huîtres mais de façon beaucoup plus marquée pour l'huître *C. gigas*. Ce  $\beta$ 1-2Xyl est également retrouvé sur les *N*-glycannes des escargots (*Helix pomatia*, *Lymnea stagnalis* et *Biomphalaria glabatra*) et des trématodes de type *Schistosoma* (Ishihara *et al.*, 1979 ; Khoo *et al.*, 1997 ; Lehr *et al.*, 2007 ; Van Kuik *et al.*, 1985, 1986, 1987). De façon surprenante, cette dernière caractéristique n'a pas été mise en évidence chez *C. virginica*. En revanche, concernant la sulfatation, nous n'avons pas réussi à la mettre en évidence chez *C. gigas* malgré 2 approches méthodologiques différentes alors que Kurz *et al.* (2013) démontrent la présence de groupements sulfates au sein de quelques *N*-glycannes des hémocytes et du plasma de *C. virginica*. Nos travaux ainsi que ceux de Kurz *et al.* (2013) indiquent que selon l'espèce d'huître considérée, des variations importantes sont observées dans les *N*-glycomes.

A notre connaissance, les seules structures *O*-glycanniques caractérisées à ce jour sont celles décrites par Kurz *et al.* (2013) dans les hémocytes et le plasma de l'huître *C. virginica*. Selon ces auteurs, il s'agit de petits *O*-glycannes constitués du noyau 1, Gal $\beta$ 1-3GalNAc et GlcNAc $\beta$ 1-3Gal $\beta$ 1-3GalNAc. Nos résultats, s'ils confirment l'existence de *O*-glycannes chez les huîtres *C. gigas* et *O. edulis* quel que soit le tissu considéré, révèlent néanmoins qu'il s'agit de structures beaucoup plus longues allant jusqu'à 10 monosaccharides (Hex<sub>2</sub>HexNAc<sub>3</sub>dHex<sub>3</sub>). Ces *O*-glycannes sont bâtis sur 3 noyaux souvent identifiés chez les invertébrés, les noyaux de type 1, 2 et le « noyau des invertébrés » (Gal $\beta$ 1-6[Gal $\beta$ 1-3]GalNAc)

décrit dans la revue de Staudacher (2015) et de manière plus surprenante, sur des noyaux de type 3 et 4 qui sont habituellement retrouvés chez les mammifères. Nos analyses par fragmentation MALDI-MS<sup>n</sup> réalisées par spectrométrie de masse MALDI-QIT-TOF sur la majorité des *O*-glycannes perméthylés du tube digestif de *C. gigas* mettent en évidence, comme pour les *N*-glycannes, la présence d'Ags A et vraisemblablement de ALe<sup>b</sup>. Ce dernier point devra être confirmé.

Même si les analyses MALDI-MS<sup>n</sup> doivent être finalisées et que celles GC-MS des dérivés itols générées à partir des *O*-glycannes n'ont pas encore été faites, la comparaison des *N*- et *O*-glycomes chez les 2 espèces d'huître semble faire apparaître les mêmes caractéristiques structurales.

L'analyse structurale des glycolipides a également été envisagée mais de multiples essais n'ont pas donné de résultats concluants. Nous avons donc décidé de nous focaliser uniquement sur l'analyse des *N*- et *O*-glycannes.

## B. La méthylation

Quels que soient les tissus considérés, nos analyses structurales sur les *N*- et *O*-glycannes ont montré la présence de groupements méthyles chez les deux espèces d'huître. Ce résultat semble être une caractéristique générale des glycannes d'huître puisque les *N*-glycannes de l'huître *C. virginica* sont également apparus porteurs de ces groupements (Kurz *et al.*, 2013).

Ainsi, nous avons identifié des résidus de *O*-MeMan sur les *N*-glycannes de type paucimannose et oligomannose comme pour *C. virginica* ou encore chez des gastéropodes ou des planaires (Gutternigg *et al.*, 2004, 2007a ; Natsuka *et al.*, 2011 ; Van Kuik *et al.*, 1986). L'analyse GC-MS des dérivés itols acétates nous a permis de montrer que ce groupement se situait soit sur le C3 soit sur le C4 d'un résidu de Man. Une proportion plus importante de 4-*O*-MeMan a été identifiée dans le tube digestif de chacune des huîtres en comparaison aux 2 autres tissus. Si l'existence de 3-*O*-MeMan a déjà été rapporté chez des escargots comme *Arion lusitanicus* et *Lymnea stagnalis*, celle de 4-*O*-MeMan n'a, à notre connaissance, encore jamais été rapportée chez des invertébrés (Gutternigg *et al.*, 2004 ; Hall *et al.*, 1977 ; Staudacher, 2012 ; Van Kuik *et al.*, 1986).

Deux autres monosaccharides, la GalNAc et le Fuc, apparaissent également méthylés sur les *N*-glycannes des 3 tissus des deux espèces d'huître. Nos analyses GC-MS ont localisé le groupement méthyle des résidus de GalNAc sur leur C3 (3-*O*-MeGalNAc). Pour rappel, tous les résidus de GalNAc que nous avons identifiés ont été localisés dans des Ags A et ALe<sup>b</sup> et une très forte proportion de ces résidus de GalNAc est apparue méthylée quels que soient les tissus et l'espèce d'huître considérés. Cela pose évidemment la question de l'impact du groupement méthyle sur la reconnaissance de ces Ags non seulement par les lectines et anticorps classiquement utilisés pour leur mise en évidence mais surtout par les NoVs. Concernant le résidu de Fuc, nous avons mis en évidence 2 profils de méthylation, l'un sur le C3 (3-*O*-MeFuc) chez *C. gigas* et l'autre sur le C4 (4-*O*-MeFuc) chez *O. edulis*. Nos analyses MS<sup>n</sup> par spectrométrie de masse MALDI-QIT-TOF ont permis de montrer que ces résidus de MeFuc constituaient l' $\alpha$ 1-4Fuc qui définit l'Ag Le<sup>b</sup> chez les deux espèces d'huître et l' $\alpha$ 1-2Fuc des Ags A et H (comprenant les Ags ALe<sup>b</sup> et HLe<sup>b</sup>) uniquement chez *O. edulis*. Cet élément révèle une nouvelle fois qu'il existe des variations qualitatives de glycosylation entre les 2 espèces d'huître. Si des dérivés 3-*O*-MeGalNAc et 3-*O*-MeFuc ont déjà été rapportés dans la littérature chez le bivalve d'eau douce *Hyriopsis schlegelii* et le ver *Alvinella pompejana*, il n'en est pas de même pour le 4-*O*-MeFuc, sauf erreur de notre part (Hori *et al.*, 1981 ; Staudacher, 2012 ; Talmont et Fournet, 1991).

En dépit du fait que l'identification des méthyl-monosaccharides retrouvés au sein des *O*-glycannes reste encore à faire, il est raisonnable de penser que les 3-*O*-MeGalNAc et 3/4-*O*-MeFuc observées au sein des *N*-glycannes seront retrouvés sur les *O*-glycannes. A noter que la présence de monosaccharides méthylés dans des *O*-glycannes a déjà été rapportée notamment chez les nématodes *C. elegans* et *Toxocara canis* (Guérardel *et al.*, 2001 ; Khoo *et al.*, 1991) mais aussi chez plusieurs escargots comme *Helix pomatia* ou *Arion lusitanicus* (Stepan *et al.*, 2012).

La méthylation de monosaccharides chez les invertébrés n'est pas inhabituelle mais peu d'informations sont disponibles concernant leur biosynthèse et notamment sur les méthyltransférases impliquées. Hormis quelques-unes identifiées chez les bactéries, aucune à notre connaissance ne l'a encore été chez les invertébrés (Staudacher, 2012). Néanmoins, une étude chez *C. elegans* a permis d'identifier le transporteur SAMT-1 de la S-adénosyl-méthionine (utilisée comme donneur de

groupement méthyle), depuis le cytosol jusqu'à l'appareil de Golgi pour la méthylation des *N*- et *O*-glycannes par les méthyltransférases (Wohlschlager *et al.*, 2014). A ce jour, le rôle de la méthylation des glycannes n'est pas clairement établi. Les groupements méthyles sont principalement retrouvés sur les monosaccharides en position terminale des oligosaccharides et notamment le 3-*O*-MeMan postulé pour être un signal de fin d'élongation par les glycosyltransférases (Staudacher, 2012). Il a également été démontré que des glycannes méthylés pouvaient constituer des cibles de l'immunité innée. En effet, l'étude de Wohlschlager *et al.* (2014) qui concernait les Tectonines Tec2 du champignon *Laccaria bicolor* et L6 de la limule japonaise *Tachypleus tridentatus* a permis de mettre en évidence une reconnaissance des résidus de 3-*O*-MeMan et 2-*O*-MeFuc. Ce type de lectine, retrouvé chez les vertébrés et invertébrés, est connu pour reconnaître les lipopolysaccharides bactériens ainsi que d'être associé à l'immunité innée. Cependant, on ne connaît pas le rôle pour l'espèce porteuse de ces modifications.

## II. Interactions Norovirus-huître

### A. Interaction médiée par des ligands glycaniques

En parallèle de nos analyses structurales des *N*- et *O*-glycomes des 3 tissus des 2 espèces d'huître, nous avons cherché à déterminer le ou les ligands impliqués dans les interactions NoVs - huître. L'implication des glycannes dans cette reconnaissance a été démontrée dans de multiples travaux en altérant les chaînes glycaniques par du périodate de sodium (Le Guyader *et al.*, 2006 ; Maalouf *et al.*, 2010b ; Tian *et al.*, 2006, 2007). De manière plus précise, des reconnaissances de motifs glycaniques par des lectines et anticorps et des expériences de compétition avec des substrats naturels (*e.g.* salive d'individu sécréteur de groupe A) ont clairement démontré l'implication des HBGAs (Le Guyader *et al.*, 2006 ; Maalouf *et al.*, 2010b ; Tian *et al.*, 2006, 2007). Les questions en suspens étaient :

- Y-a-t-il un motif glycanique particulier impliqué dans l'interaction de huîtres avec les NoVs ?
- Si oui, est-il porté par un ou des *N*- ou *O*-glycane(s), voire les 2 ?
- Ce ou ces glycannes sont-ils portés par une ou plusieurs protéines ?

Ne disposant pas des *N*- ou *O*-glycannes purs, nous avons cherché à répondre aux questions 2 et 3.

Nos expériences de FarWB et d'ELISA ont confirmé une interaction spécifique des VLPs de NoVs (GI.1 et GII.4) avec les 3 tissus de *C. gigas* et d'*O. edulis*. De manière générale, nos résultats issus des farWBs ont montré que les reconnaissances sont localisées au niveau du gel de concentration et à l'interface des gels de concentration et de séparation ce qui suggère une implication de GPs de très haut poids moléculaires. Si l'on ajoute à cette observation le fait qu'un prétraitement à la PNGase F des échantillons ne conduit à aucun changement significatif, cela laisse à penser que ce sont des *O*-glycannes portés par des GPs de type mucine qui interagissent avec les NoVs.

Chez l'huître, comme chez les invertébrés marins, une couche de mucus est présente à la surface des différents tissus et joue des rôles biologiques extrêmement importants. Pales Espinosa *et al.* (2016) ont caractérisé par LC-MS/MS les constituants du mucus collecté à partir du manteau, des branchies et des palpes labiaux de l'huître *C. virginica*, essentiellement au niveau protéique. Les auteurs n'ont pas identifié de mucines chez *C. virginica* et on ne trouve également pas de gènes codant pour des mucines dans le génome de *C. gigas* (Zhang *et al.*, 2012). Néanmoins, l'existence de mucines chez les huîtres n'est pas complètement tranchée puisque ces auteurs ont également noté la présence d'une halomucine et d'une SCO-spondine dont les séquences sont similaires aux mucines. La première est homologue aux mucines de mammifères et la seconde comporte des unités de répétition peptidique de type mucine-2 avec un haut taux de similarité avec la mucine-5 AC-like (*Aplysia californica*, XP\_005092239). Leurs séquences sont également retrouvées dans le génome de *C. gigas* ce qui pourrait être le signe de la présence de GPs de type mucine chez les huîtres (Pales Espinosa *et al.*, 2016).

Les HBGAs identifiés lors de nos analyses structurales ont été identifiés à la fois sur les *N*- et les *O*-glycannes. Il est donc judicieux de se demander pourquoi la reconnaissance aurait lieu sur les *O*-glycannes plutôt que les *N*-glycannes. Dans les travaux de Ruvoën-Clouet *et al.* (2006), les auteurs suggèrent, à propos de l'interaction de VLPs de NoV GI.1 sur des mucines de lait humain, que la présentation multimérique des glycannes des mucines augmentent l'avidité des interactions. De plus, l'étude de Huang *et al.* (2009) démontre que les ligands glycaniques des NoVs GII.4 (VA387) et GII.9 (VA207) dans le lait et la salive humains sont des GPs de très hauts poids moléculaires suspectés pour être des mucines. Ils démontrent que la présence de l'HBGA spécifique d'une souche de NoV n'est pas



suffisante pour une reconnaissance. L'interaction nécessiterait également un squelette glycanique porteur spécifique. L'importance de la structure des chaînes glycaniques qui portent l'Ag a également été illustrée dans plusieurs travaux concernant le virus Influenza A. Au fil du temps, les souches de virus Influenza A ont une avidité plus forte pour des glycanes biantennés sialylés alors qu'elle diminue pour des trisaccharides, petits glycanes ou glycanes linéaires sialylés. Cela s'expliquerait par la possibilité d'une double liaison des antennes à 2 sites de liaison d'un trimère d'hémagglutinine lorsque les antennes sont suffisamment longues (Ji *et al.*, 2017). Dans une autre étude, les auteurs ont réalisé des puces à glycanes (glycan-array) mimant les mucines du glycocalyx en utilisant des polymères de différentes longueurs et exposant un certain nombre de glycanes (valence) par polymère. Les résultats obtenus avec l'utilisation d'une souche d'Influenza A H1N1 indiquent une augmentation de l'avidité lorsqu'on augmente la valence (Lucas *et al.*, 2021). Si l'on considère que la présentation multimérique des glycanes des mucines augmentent donc l'avidité des interactions également pour les NoVs, cela constitue un argument en faveur d'une reconnaissance des HBGAs portés par des *O*-glycanes plutôt que par des *N*-glycanes.

Nos travaux de reconnaissance par des expériences de farWB des protéines et GPs des 2 espèces d'huître par le NoV GII.4 Dijon, ont également montré qu'une GP plus légère (entre 25 et 37 KDa) présente uniquement dans le manteau et les branchies de l'huître *O. edulis* était reconnue. Si ce résultat est confirmé, cela pourrait signifier l'existence de plusieurs ligands pour cette souche de NoV, mais aussi révéler un certain tropisme tissulaire et d'espèce.

## B. Influence des HBGAs méthylés sur les Norovirus

Comme nous venons de le dire dans le chapitre précédent, les expériences de FarWB et d'ELISA suggèrent fortement une implication des *O*-glycanes dans l'interaction avec les NoVs. Pour ces derniers, nos premiers résultats structuraux indiquent la présence d'Ag<sub>s</sub> A et ALe<sup>b</sup> qui devront néanmoins être vérifiés par des analyses MALDI-MS<sup>n</sup> complémentaires. De surcroît, nous avons montré que ces *O*-glycanes apparaissent méthylés à des degrés divers, jusque 4 groupements pour certains. Si, comme pour les *N*-glycanes, les groupements méthyles sont retrouvés sur des résidus de GalNAc et de Fuc qui définissent les motifs antigénique A et ALe<sup>b</sup>, cela questionne sur l'impact de ces groupements

méthyls en termes de reconnaissance par les NoVs. La présence d'un ou 2 groupements méthyles sur un Ag inhibe-t-elle l'interaction avec le NoV ? diminue-t-elle ou augmente-t-elle l'affinité de la reconnaissance par le NoV comparé au même Ag non méthylé ?

### 1. « Tropisme » d'espèce

Nos travaux ont révélé des différences de profils de méthylation entre les huîtres *C. gigas* et *O. edulis*. Ainsi, les Ags A des *N*-glycannes de *O. edulis* disposent d'un groupement méthyle supplémentaire sur le C4 du résidu de Fuc en comparaison à *C. gigas*. De même, l'Ag Le<sup>b</sup> est constitué d'un 4-*O*-MeFuc pour *O. edulis* mais d'un 3-*O*-MeFuc pour *C. gigas*. Ces différences structurales entraînent-elles des différences d'interaction entre ces 2 espèces d'huître vis-à-vis des NoVs en termes d'affinité ? Un premier élément de réponse pourrait résider dans l'observation qu'ont faite nos collaborateurs d'IFREMER lors de leurs études de contamination in vivo des 2 huîtres par les NoVs. Il semblerait que les huîtres *C. gigas* soient plus sensibles à la contamination par les NoVs que les huîtres *O. edulis* (communication personnelle). Si cette observation était confirmée, en tenant compte notamment des caractéristiques physiologiques des 2 huîtres (poids, taille, volume d'eau de mer filtré, ...), les différences des profils de méthylation pourraient être une explication à ce phénomène.

Tous ces questionnements sont légitimes si l'on se souvient de la très grande spécificité de reconnaissance virus-hôte qui a été illustré dans l'Introduction (Chap. 1. VII p55). On peut donc aisément penser qu'un groupement méthyle puisse changer la reconnaissance par les NoVs puisque la reconnaissance d'un motif monosaccharidique (ou oligosaccharidique) par un virus peut être impacté par la nature de la liaison au reste de la chaîne oligosaccharidique (par exemple Influenza A) ou par la présence d'un groupement particulier (par exemple certains CoVs). Tous ces éléments, qui peuvent apparaître mineurs peuvent être à l'origine du tropisme d'espèce selon qu'ils sont produits chez une espèce et pas chez une autre. La présence de ces groupements et leur position peuvent alors induire un phénomène que l'on peut nommer de « tropisme d'espèce » bien qu'ici les huîtres ne soient pas infectées par les NoVs.

## 2. Influence selon les souches virales : modélisation moléculaire

On ne connaît pas à ce jour l'influence d'un groupement méthyle sur un Ag au regard de sa reconnaissance par des anticorps, des lectines ou encore des virus. Dans notre cas, nous voulions savoir si la présence d'un groupement méthyle sur un Ag A modifiait l'interaction de cet épitope avec les NoVs en comparaison à un Ag non méthylé. Pour cela, nous avons réalisé une étude de modélisation moléculaire à partir des données cristallographiques du domaine P des virus GI.1 et GII.3 en interaction avec le trisaccharide GalNAc $\beta$ 1-3[Fuc $\alpha$ 1-2]Gal qui définit l'Ag A.

La modélisation avec le NoV GI.1 a montré une implication des hydroxyles du C2 du résidu de Fuc et des C3 et C4 du résidu de GalNAc. Ainsi, il est apparu que la présence d'un groupement méthyle sur l'hydroxyle du C3 de la GalNAc supprime la liaison hydrogène du C3 établie par cet hydroxyle avec pour conséquence une diminution de la force d'interaction Ag A-NoV. En revanche, la présence d'un groupement méthyle sur l'hydroxyle du C4 du Fuc ne modifie pas la reconnaissance. Cette dernière observation semble indiquer qu'il n'y aurait donc pas de différence de reconnaissance des Ags A par les NoV GI.1 entre les huîtres *C. gigas* et *O. edulis*.

Pour la souche de GII.3, le résidu de Fuc apparaît beaucoup plus impliqué dans la reconnaissance (interaction avec les hydroxyles des C2, C3 et C4) que le résidu de GalNAc (interaction avec l'hydroxyle du C3). De nouveau, la présence d'un motif 3-*O*-MeGalNAc d'un Ag A engendre la perte d'une liaison hydrogène avec le NoV. Par contre, la présence d'un motif 4-*O*-MeFuc dans un Ag A, ce qui est le cas pour *O. edulis*, fait disparaître 3 liaisons hydrogènes ce qui constitue une différence importante dans la reconnaissance des deux huîtres pour cette souche de NoV. Ce résultat, s'il est confirmé par d'autres approches, pourrait expliquer une plus grande sensibilité de *C. gigas* vis-à-vis de contaminations à ce NoV par rapport à *O. edulis*.

Plusieurs études ont été réalisées pour établir les liaisons qui s'établissent entre l'Ag A et différents variants de GII.4. Une étude d'interaction par RMN STD concernant des VLPs du variant GII.4 Ast6139 rapporte une implication forte du résidu de GalNAc au niveau de son groupement *N*-acétyle et des hydroxyles de ses C3 et C4 (Fiege *et al.*, 2017). Comme pour les souches GI.1 et GII.3, la méthylation de l'hydroxyle du C3 pourrait conduire à une diminution de l'interaction hôte-virus chez

les 2 huîtres. En revanche, l'hydroxyle du C4 du résidu de Fuc de l'Ag A ne semble pas établir de liaison hydrogène avec les VLPs du NoV GII.4 Ast6139. Ces résultats indiquent donc qu'*in fine* il n'y aurait pas de différence de reconnaissance de l'Ag A par les deux espèces d'huîtres étudiées. Néanmoins, il convient de préciser à ce niveau que ces résultats apparaissent extrêmement dépendants de la souche virale utilisée. En effet, des études cristallographiques effectuées avec la protéine P de GII.4 VA387 et du variant 2004, ont révélé une plus forte implication du Fuc avec l'engagement cette fois-ci de l'hydroxyle de son C4 (Cao *et al.*, 2007 ; Shanker *et al.*, 2011).

Il apparaît donc que les mécanismes d'interaction Ag A méthylé-NoV sont non seulement dépendants de la souche de NoV considérée mais aussi des variants au sein d'une même souche virale.

### C. « Tropisme » tissulaire

Nos études structurales sur les *N*- et *O*-glycannes des 3 tissus des 2 huîtres ne montrent pas de différences qualitatives majeures. Cependant, les résultats issus de nos expériences de farWB et ELISA montrent que les NoVs ne se fixent pas en quantité équivalente sur les 3 tissus, manteau, tube digestif et branchies. A ce stade de nos travaux, l'explication est sans doute quantitative.

Pour *C. gigas*, les expériences ELISA et FarWB montrent une fixation préférentielle des VLPs de NoVs GI.1 au niveau du tube digestif. Cette observation est confirmée à la fois par les résultats d'ELISA et les expériences de bioaccumulation (Maalouf *et al.*, 2010b, 2011 ; Tian *et al.*, 2006). Cependant, pour les VLPs de NoVs GII.4 MD145, les résultats de FarWB et d'ELISA apparaissent en contradiction. En effet, par FarWB, le manteau est le tissu le plus reconnu alors que par ELISA, il s'agit essentiellement du tube digestif et plus faiblement le manteau. L'étude de Maalouf *et al.* (2011) montre que lors d'expériences de bioaccumulation, les VLPs de NoV GII.4 Houston se fixent d'abord au niveau des branchies et migrent ensuite vers le tube digestif et le manteau avec une bioaccumulation plus de deux fois supérieure dans le tissu digestif. D'autre part, les expériences ELISA réalisées avec les mêmes VLPs dans leur étude précédente montre une reconnaissance des tous les tissus de manière équivalente (Maalouf *et al.*, 2010b). Au regard de nos résultats, il n'est donc pas possible de conclure sur une éventuelle préférence tissulaire des VLPs GII.4.

# CONCLUSION & PERSPECTIVES

---

Les études structurales des *N*- et *O*-glycomes nous ont permis d'identifier des Ags A, ALe<sup>b</sup> et H chez les huîtres *C. gigas* et *O. edulis* et HLe<sup>b</sup> uniquement pour cette dernière. Nous avons montré que ces Ags sont majoritairement substitués par des groupements méthyles sur des résidus de GalNAc et de Fuc et à des positions différentes selon les deux espèces.

Pour poursuivre les études que nous avons entreprises pour identifier le ou les ligands glycaniques impliqués dans la rétention des NoVs par les 2 espèces d'huître, il faudra, d'un point de vue structural, finaliser la caractérisation structurale du *N*-glycome en étudiant les groupements sulfates de l'huître *O. edulis* et poursuivre l'étude du *O*-glycome des 2 espèces en :

- terminant les analyses de spectrométrie de masse MALDI-MS<sup>n</sup> des oligosaccharides perméthylés et perdeutérométhylés pour confirmer la présence des mêmes Ags (voire d'autres) que ceux retrouvés sur les *N*-glycannes et localiser les groupements méthyles,
- effectuant une analyse qualitative et quantitative des monosaccharides constitutifs sous leur forme réduite et perméthylée par GC-MS,
- identifiant le ou les monosaccharides naturellement méthylé(s) et le positionnement précis des groupements méthyles par une analyse GC-MS de leurs dérivésitol partiellement méthylés et acétylés,
- quantifiant les différentes formes (non-, mono-, di-,...) méthylées des Ags par LC-MS<sup>n</sup> à l'aide de l'Orbitrap Fusion™ Tribrid™,
- étudiant la présence ou non de groupements sulfates par MALDI-QIT-TOF et le cas échéant, comme pour les groupements méthyles, de préciser le ou les monosaccharide(s) sulfaté(s) et le nombre de groupements sulfates par oligosaccharide.

Les études d'interaction VLPs-glycoprotéines que nous avons menées indiquent qu'elles impliquent des protéines de haut poids moléculaire et que les *N*-glycannes ne semblent pas jouer un rôle majeur. Nos expériences ELISA et farWB ont été faites en effectuant une extraction en bloc des protéines à partir de tissu frais. Une première démonstration de la validité de notre hypothèse sur l'implication de glycoprotéines de type mucine, consisterait à préparer les glycoprotéines selon le protocole standard d'extraction des mucines que nous avons déployé pour établir le *O*-glycome. Puis,

les glycoprotéines ainsi préparées seront soumises aux mêmes tests ELISA et farWB que ceux menés sur les glycoprotéines totales. Dans le but d'éliminer un rôle possible des *N*-glycannes, ces tests pourraient être effectués avant et après action des PNGase F et A. Quel que soit les tests de reconnaissance qui seront mis en œuvre, il conviendra de les compléter systématiquement par des expériences d'inhibition en utilisant des anticorps, lectines, les « mucines » elles-mêmes voire les *N*-glycannes comme inhibiteurs, ce que faute de quantité suffisante en VLPs nous n'avons pu faire.

Durant ma thèse, nous avons confirmé la présence d'Ag A, H, ALe<sup>b</sup> et HLe<sup>b</sup> chez les huîtres *C. gigas* et *O. edulis*. Le fait intéressant est que ces Ags existent sous forme native mais aussi sous plusieurs formes méthylées. La question qui se pose est donc l'implication de chacune de ses formes dans les interactions avec les NoVs. Pour apporter une réponse à ces questions, le Dr Yann Guérardel, professeur invité dans l'Institute for Glyco-core Research (iGCORE) associé aux Universités de Nagoya et de Gifu au Japon a établi une collaboration avec l'équipe du Prof. Hiromune Ando de l'université de Gifu afin de synthétiser par voie chimique ces épitopes sous leurs différentes formes méthylées. Ce travail s'inscrit dans le cadre d'un réseau de recherche international entre nos deux laboratoires financés par le CNRS (IRN Glyco-Network). L'utilisation de ces antigènes sera multiple. Tout d'abord, ils seront utilisés pour mesurer grâce à la technique de résonance plasmonique de surface (SPR), l'affinité des ligands glycaniques natifs et modifiés avec les VLPs et permettra ainsi de connaître la structure glycanique la plus affine. Ensuite, et en complémentarité à la mesure des affinités d'interaction, ils serviront à caractériser au plan atomique la nature exacte des interactions Ags-VLP. La technique la plus adéquate pour mener ce travail est la RMN STD (Saturation Transfer Difference) qui permet de déterminer la nature exacte des épitopes glycaniques par irradiation sélective en présence de son récepteur protéique (VLP). Les bases d'une collaboration ont déjà été établies avec l'équipe du Pr. Molinaro à Naples (Di Carluccio *et al.*, 2021).

Cette étude montre que l'identification et la caractérisation structurale des glycannes impliqués dans une interaction hôte-virus peut s'avérer extrêmement longues et fastidieuses. Dans l'environnement et le milieu marin en particulier, il existe de nombreuses espèces qui entrent dans l'alimentation humaine (poissons, mollusques bivalves ou encore crustacées) et qui font l'objet de

contaminations par les bactéries, parasites et virus. Disposer d'un outil diagnostique permettrait de vérifier, pour un pathogène, si le ligand est de nature glucidique et dans l'affirmatif de l'identifier. Cela permettrait également de vérifier rapidement si ce ligand est aussi présent chez d'autres espèces et ainsi déterminer les espèces susceptibles d'être contaminées par ce pathogène. Des glycan-array avec des glycannes issus d'une multitude d'espèces marines et criblées avec des VLPs pourrait constituer cet outil. Ainsi, face à une contamination/infection marine, il sera possible de savoir rapidement si une ou plusieurs espèces marines peuvent être infectées par le virus et ainsi prévenir les contaminations inter-espèces (épizootie) et surtout les contaminations humaines (zoonoses).



# MATERIEL & METHODES

---

## I. Echantillons

Les différents tissus des huîtres *C. gigas* et *O. edulis* (tube digestif, branchies et manteau) d'une douzaine d'huîtres ont été disséqués au LSEM (Ifremer, Nantes) par nos collaborateurs du projet ANR GOyAVE, Marion Desdouits et Gabriel Euler, puis poolés et congelés. Après réception et décongélation, ils ont été broyés dans des cellules de 50 mL « gentleMACS™ Dissociator » pendant 2 min puis lyophilisés.

## II. Préparation des *N*-glycannes

Cent mg de tissus ont été suspendus dans 1 mL d'une solution de triton X-100 à 1 % dans le PBS (Phosphate Buffered Saline Solution) puis soumis aux ultrasons pendant 2 min (Sonoplus, ultrasonic homogenizer mini 20, Brandolin, Allemagne, sonde 1,5 mm de diamètre) à 80 % de la puissance maximale, pulsation toutes les 0,5 s avec un repos de 0,5 s sur glace. L'échantillon a été ensuite incubé toute une nuit à 37 °C sous agitation. Après centrifugation (13 000 rotations par minute (RPM), 10 min à température ambiante (Tamb)), les protéines et GPs ont été récupérées dans le surnageant, réduites par 100 µL de dithiothréitol (DTT) à 0,1 M dans le PBS (37 °C, 1h, sous agitation) puis alkylées par 110 µL d'une solution à 0,5 M de iodoacétamide dans le PBS dans les mêmes conditions. La précipitation des protéines et GPs a été réalisée en ajoutant 130 µL d'acide trichloroacétique (ATCA) saturée suivie d'une incubation à -20 °C pendant 30 min. Le culot protéique a été récupéré après centrifugation (13000 RPM, 10 min, à 4 °C) et lavé 3 fois par 1 mL d'acétone froid (-20 °C) avec une centrifugation de 10 min à 13 000 RPM à 4 °C entre chaque lavage. Après séchage sous vide par un évaporateur centrifugeur sous vide (SpeedVac) (Concentrator 5301, Eppendorf), le culot a été repris dans 500 µL d'une solution de bicarbonate d'ammonium (NH<sub>4</sub>HCO<sub>3</sub>) 50 mM. Les protéines et GPs ont été digérées par 40 µL d'une solution de trypsine à 2 mg/mL dans le NH<sub>4</sub>HCO<sub>3</sub> 50 mM à 37 °C toute la nuit. Après destruction de la trypsine par un traitement thermique à 100 °C pendant 10 min, les polymères de glucose (supposés être des fragments de glycogène intrinsèques) ont été éliminés par extraction en phase solide (solid phase extraction ou SPE en anglais) sur l'automate SPE 6.25ws (Interchim) en utilisant des cartouches de 3 mL remplies de billes de silice greffées avec des chainons C18 (500 mg, cartouche LCC500/3, Interchim). Après activation de la phase stationnaire par 5 mL d'acétonitrile (ACN) et son équilibration par 10 mL d'une solution d'acide acétique 5 %, l'échantillon a été déposé à un débit de 0,75 ml/min. Après lavage de la cartouche par 10 mL d'une solution d'acide acétique à 5% dans l'eau à un débit de 0,75 ml/min (les polymères de glucose sont éliminés durant ce lavage), les peptides et glycopeptides ont été récupérés par une augmentation en palier successif de 20 %, 40 % puis 60 % d'isopropanol dans une solution d'acide acétique à 5 % dans l'eau (10 mL par solution). Après lyophilisation, ils ont été repris dans 500 µL de NH<sub>4</sub>HCO<sub>3</sub> à 50 mM et les *N*-glycannes ont été libérés par 1 µL de PNGase F (62,5 UI, New England BioLabs). Les *N*-glycannes ont été récupérés par extraction en phase solide dans les mêmes conditions que

précédemment. Après dépôt de l'échantillon, la phase de lavage (fraction non retenue) dans laquelle se trouvent les *N*-glycannes a été récupérée puis lyophilisée. Quant aux peptides et glycopeptides résistants à la PNGase F, ils ont été élués en faisant passer successivement à travers la cartouche 10 mL de solutions d'isopropanol à 20, 40 et 60 % dans de l'acide acétique à 5 %. En suivant le protocole employé pour la PNGase F, les glycopeptides restants ont été soumis à une digestion par la PNGase A (10 UI, New England BioLabs) en suivant les recommandations du fournisseur.

### III. Préparation des *O*-glycannes

Une solution de chlorure de guanidine (GuCl) 6 M a été purifiée par du graphite. Après ajustement de la concentration de la solution de GuCl à 4 M par dilution dans l'eau et ajout d'une tablette d'inhibiteur de protéases (cOmplete, EDTA-free, ref 58698000, Roche Diagnostics), 0,5 g de tissu broyé et lyophilisé a été resuspendu dans 8 mL de cette solution. La densité de la solution a été amenée à 1,4 par ajout de CsCl en poudre avant d'être transférée dans un tube de centrifugation pour être y centrifugé à 48 000 g pendant 72 h à 4 °C. Des fractions de 0,5 mL ont été collectées à partir du fond du tube grâce au dispositif Fraction Recovery Systems commercialisé par Beckman Coulter. Cinquante µL de chacune des fractions ont été déposés sur membrane de nitrocellulose (Amersham Protan 0,45 µm NC) disposée sur le support Slot blot (Bio-Dot® SF Apparatus, Bio-rad). Deux révélations du matériel glycanique ont été réalisées, l'une par le PAS et l'autre par la lectine AAL marquée à la biotine (cf. protocole Lectine-blot). Ces révélations permettent de déterminer les fractions contenant du glycogène (matériel de haute densité) qui sont à éliminer et les fractions contenant les mucines (densité voisine de 1,3). Les fractions positives à la lectine AAL ont été poolées puis dialysées sur des membranes de dialyse Spectrapor en boudin (seuil de coupure 6-8 kDa, diamètre 2,04 mm et longueur 10 cm) contre de l'eau désionisée pendant 3 jours en changeant l'eau 2 fois par jour avant d'être lyophilisées. La β-élimination a été réalisée en resuspendant les échantillons dans 5 mL d'une solution de NaOH 50 mM renfermant 1 M de borohydrure de sodium (NaBH<sub>4</sub>). Après 72 h d'incubation à 37 °C, la réaction a été arrêtée par ajout d'acide acétique pur jusqu'à arrêt du dégagement gazeux. Puis, l'excès d'ions sodium a été éliminé par chromatographie d'échange d'ions sur une colonne échangeuse de cations contenant 15 mL d'un gel Dowex 50X8, (200-400Mesh, forme H<sup>+</sup>, Biorad) équilibrée et éluee par une solution d'acide acétique à 5 % dans l'eau. Après lyophilisation de la fraction non retenue, les acides aminés et petits peptides ont été éliminés par chromatographie échangeuse de cations contenant 15 mL d'un gel Dowex 50X2 (200-400 mesh, forme H<sup>+</sup>, Bio-rad) équilibrée et éluee par une solution d'acide acétique 5 %. Après une nouvelle étape de lyophilisation, les *O*-glycannes ont été débarrassés des borates par une dizaine de co-distillation en présence d'un mélange méthanol/acide acétique (9/1 ; V/V) (1 mL).

## IV. Dérivation des *N*- et *O*-glycannes pour l'analyse en MALDI-QIT-TOF

Dix pourcents des *N*-glycannes et *O*-glycannes totaux ont été perméthylés selon un protocole dérivé de la méthode de Ciucanu et Kerek (1984) utilisant une “slurry base” plutôt que de la soude en poudre. Cette slurry base est préparée extemporanément en ajoutant à 400  $\mu$ L d'une solution de NaOH à 50 % dans l'eau, 800  $\mu$ L de méthanol anhydre suivie d'une agitation vigoureuse. Puis 4 mL de diméthylsulfoxyde (DMSO) ont été ajoutés et mélangés vigoureusement. Après centrifugation à 1500 RPM pendant 1 min à Tamb. La phase supérieure contenant un insoluble blanc est retirée à l'aide d'une pipette pasteur jusqu'au culot visqueux (d'où le terme de slurry base). Plusieurs autres lavages du culot par 4 mL de DMSO sont effectués jusqu'à disparition de l'insoluble blanc (6 à 8 lavages). Le culot est ensuite resuspendu dans 4 mL de DMSO par aspiration et refoulement successif dans une pipette Pasteur. Pour la perméthylation, les *N*- et *O*-glycannes ont été resuspendus dans 400  $\mu$ L de slurry base et 200  $\mu$ L d'iodométhane ( $\text{CH}_3\text{I}$ ) ou deutériodométhane ( $\text{CD}_3\text{I}$ ) ont été ajoutés et le mélange a été laissé sous agitation pendant 30 min à Tamb. La réaction a été arrêtée en ajoutant goutte à goutte 2 mL d'acide acétique à 5 % froid et les glycannes per(deutéro)méthylés ont été solubilisés dans 1 mL de chloroforme. La phase aqueuse est jetée après décantation et la phase chloroformique a été lavée 5 fois par 2 mL d'eau puis séchée sous un flux d'azote.

## V. Purification des *N*-glycannes sulfatés

La perméthylation des *N*-glycannes sulfatés nécessite de mettre en place un protocole particulier notamment en ce qui concerne la récupération des glycannes sulfatés perméthylés qui du fait de leur solubilité dans l'eau, sont perdus si l'arrêt de la perméthylation se fait comme dans le protocole précédent par addition d'eau au mélange réactionnel. Dans ce cas particulier, nous avons mis en œuvre le protocole de Yu *et al.* (2020) qui utilise une cartouche Oasis MAX 3 cc (Waters, Oasis).

Brièvement, la perméthylation (décrite ci-dessus) a été arrêtée par ajout de 200  $\mu$ L d'eau et la solution a été neutralisée par quelques gouttes d'acide acétique 5 % pour atteindre un pH entre 5 et 8. Puis, l'excès d'iodométhane a été évaporé sous un flux d'azote. La cartouche Oasis MAX a été activée par 3 mL d'une solution d'ACN à 95 % dans l'eau et équilibrée par 3 mL d'une solution d'acétate d'ammonium ( $\text{NH}_4\text{OAc}$ ) 100 mM. Le mélange réactionnel contenant les *N*-glycannes sulfatés et perméthylés sont resuspendus dans 1 mL de  $\text{NH}_4\text{OAc}$  100 mM puis déposés à la surface de la cartouche. L'éluat a été déposé une seconde fois et la cartouche a été lavée par 1 mL de la solution d' $\text{NH}_4\text{OAc}$  100 mM puis par 6 mL d'eau désionisée pour éliminer les sels. Trois éluutions successives par 6 mL d'ACN à 95 %, puis 6 mL de  $\text{NH}_4\text{OAc}$  1 mM dans l'ACN 80 % et enfin 3 mL de  $\text{NH}_4\text{OAc}$  100 mM dans l'ACN 60 %, MeOH 20 % ont été réalisées pour récupérer successivement les *N*-glycannes perméthylés non sulfatés puis les *N*-glycannes perméthylés monosulfatés et enfin les *N*-glycannes perméthylés polysulfatés. Les échantillons ont ensuite été séchés au SpeedVac. Après dissolution dans

20  $\mu\text{L}$  d'une solution ACN à 10 %, les *N*-glycannes ont été dessalés en utilisant un cône ziptip C18 activé par 10  $\mu\text{L}$  d'ACN 50 % dans une solution aqueuse d'ATFA à 0,1 % puis équilibré dans une solution d'ATFA à 0,1 %. Après dépôt de l'échantillon, 3 lavages par 10  $\mu\text{L}$  d'une solution d'ATFA à 0,1 % ont été effectués puis les *N*-glycannes sulfatés et perméthylés ont été élués par 10  $\mu\text{L}$  d'une solution d'ACN à 50 % dans une solution d'ATFA à 0,1 %.

## VI. Dérivation des *N*-glycannes à la DAP

Les *N*-glycannes (10 nmoles) ont été repris par 50  $\mu\text{L}$  d'une solution de DAP à 0,875 M dans un mélange DMSO/Acide acétique (7/3 ; v/v). Cinquante  $\mu\text{L}$  d'une solution de  $\text{NaBH}_3\text{CN}$  1 M dans le même mélange ont été ajoutés et le tube a été incubé pendant 4 h à 65 °C. L'excès des réactifs a été éliminé par tamassage moléculaire sur une pipette plastique de 5 mL renfermant 4 mL de gel sephadex G10 (40-120mM, VWR). Après 2 lavages à l'eau, le mélange réactionnel a été déposé délicatement à la surface du gel puis la colonne a été lavée à l'eau. Cinquante fractions de 300  $\mu\text{L}$  ont été collectées manuellement et exposées sous une lampe UV (Uvitec LF-206-LS) afin de déterminer la fraction contenant les *N*-glycannes dérivés. Les fractions positives ont été poolées et lyophilisées.

## VII. Analyse par MALDI-QIT-TOF

Les *N*- et *O*-glycannes dérivés (sous forme perméthylé ou deutérométhylé ou DAP) ont été repris par 100  $\mu\text{L}$  d'ACN. L'échantillon (1  $\mu\text{L}$ ) a été déposé sur la cible et mélangé à 1  $\mu\text{L}$  d'une matrice de Super DHB (mélange d'acide 2,5-dihydroxybenzoïque et d'acide 2-hydroxy-5-méthoxybenzoïque, (9/1 ; p/p)) à 10 mg/mL dans une solution d'ACN/NaOH 1 mM (7/3 ; v/v). Après séchage à l'air libre, le dépôt a été recristallisé par ajout d'1  $\mu\text{L}$  d'ACN sans agitation. Les spectres  $\text{MS}^n$  ont été acquis sur le spectromètre de masse MALDI-QIT-TOF (Axima Resonance de Shimadzu). Les analyses des dérivés perméthylés et perdeutérométhylés ont été effectuées en mode positif à l'exception de celles des *N*-glycannes dérivés à la DAP et des *N*-glycannes perméthylés et sulfatés qui l'ont été en mode négatif. Les acquisitions MS ont été réalisées sur 3 zones de  $m/z$  différents (350-1600 Da, 800-3000 Da et 1700-4500 Da) pour visualiser l'ensemble des glycannes. Les spectres résultent de la somme de 300 sous-spectres correspondant chacun à 2 tirs laser. Les analyses  $\text{MS}^2$  ont été réalisées par CID (dissociation induite par collision) en utilisant l'argon comme gaz de collision et une énergie de collision comprise entre 330 et 600 eV en fonction des *N*-glycannes perméthylés ou deutérométhylés. Les analyses  $\text{MS}^3$  ont été réalisées dans les mêmes conditions à partir d'ions fragments issus des analyses de  $\text{MS}^2$ .

## VIII. Préparation des itols

Les monosaccharides des *N*-glycannes natifs ont été libérés par 1 mL d'une solution d'ATFA 4 M pendant 4 h à 100 °C en présence de 2  $\mu\text{g}$  de mésoinositol utilisé comme standard interne. Après séchage sous un flux d'azote, l'ATFA a été éliminé par 4 co-distillations successive avec 1 mL de méthanol. La réduction des monosaccharides libérés a été réalisée à Tamb pendant 4 h en présence de

500  $\mu\text{L}$  d'une solution d'ammoniaque 50 mM contenant du 1 M de  $\text{NaBH}_4$  ou  $\text{NaBD}_4$ . Après séchage, l'échantillon a été stocké sous vide toute la nuit dans un dessiccateur à vide contenant des pastilles d'hydroxyde de sodium. Les borates ont été éliminés par 9 co-distillations avec le mélange méthanol/acide acétique (9/1 ; V/V) (1 mL) puis 1 mL de méthanol pour la dernière. Les standards de monosaccharides (100  $\mu\text{g}$  de Fuc, Xyl, Gal, Glc, Man, GlcNAc et GalNAc) ont été traités dans les mêmes conditions.

## IX. Peracétylation et perméthylation des monosaccharides sous forme itols

La peracétylation a été réalisée en ajoutant 500  $\mu\text{L}$  d'anhydride acétique et 20  $\mu\text{L}$  de pyridine pendant 4 h à 100 °C. L'excès de réactif a été évaporé avec précaution sous flux d'azote afin d'éviter l'évaporation des composés. Les dérivés itols acétates partiellement secs ont été repris par 1,5 mL de chloroforme et lavés 5 fois par 2 mL d'eau. Ils ont ensuite été séchés jusqu'à ce qu'il ne reste qu'une goutte.

La perméthylation a été réalisée sur les dérivés itols comme décrit précédemment pour les *N*- et *O*-glycannes à l'exception de la phase de séchage finale qui a été effectuée jusqu'à ce qu'il ne reste qu'une goutte.

## X. Analyse des itols-acétates et itols-perméthylés par GC-MS

Les itols acétates ont été dissous dans 200  $\mu\text{L}$  de chloroforme pour les échantillons et 1 mL pour les standards. Puis, 1  $\mu\text{L}$  a été injecté en mode splitless sur une colonne capillaire Solgel 1MS (30 m  $\times$  0.25 mm  $\times$  0.25  $\mu\text{m}$ ) en élevant la température de 120 à 230°C avec un pas de 3 °C/min puis jusqu'à 270 °C avec un pas de 10°C/min. Les itols-acétates ont été détectés par impact électronique à 70 eV sur le chromatographe en phase gaz HP-7820 couplée à un simple quadripole 5977B (Agilent Technologies, Santa Clara, US) en mode full scan de 45 à 500 Da.

Quant aux itols perméthylés, ils ont été resuspendus dans 200  $\mu\text{L}$  de chloroforme (1 mL pour les standards) et 1  $\mu\text{L}$  a été injecté sur la même colonne que précédemment. Après 2 minutes d'éluion à 80°C, la température a été élevée jusque 180°C avec un pas de 2°C/min puis jusqu'à 250 °C avec un pas de 30 °C/min (Black *et al.*, 2019).

## XI. Quantification des antigènes A et ALe<sup>b</sup> par nanoLC-MS/MS

La quantification des profils de méthylation des antigènes A et ALe<sup>b</sup> des *N*-glycannes perdeutérométhylés de *C.gigas* et *O.edulis* a été réalisée par nanoLC-MS/MS sur un Orbitrap Fusion Tribrid system (ThermoFisher Scientific) à l'Institute of Biochemical Sciences de l'université de Taiwan dans l'équipe du Pr. Kay-Hooi Khoo dont le protocole a été décrit dans les travaux de Hsiao *et al.* (2017). Les ions *m/z* 414, 417, 645, 648, 719, 722, 913, 916 et 919 ont été ciblés.

## XII. Préparation des échantillons pour WB et ELISA

Les tissus lyophilisés (10 mg) ont été repris par 200 µL du tampon d'extraction des protéines contenant du NH<sub>4</sub>HCO<sub>3</sub> 50 mM, 0,25 % de SDS, 0,25 % de β-mercaptoéthanol et des inhibiteurs de protéases 1X (tablettes cOmplete™, EDTA-free Protease Inhibitor Cocktail, Roche). Le surnageant a été récupéré après centrifugation (10 min à 12000 RPM) puis 200 µL de NH<sub>4</sub>HCO<sub>3</sub> 50 mM et 20 µL de NP-40 à 10 % dans ce tampon ont été ajoutés. Dans le cas du traitement des échantillons à la PNGase F avant leur analyse en SDS-PAGE, 1 µL de PNGase F (62,5 UI) a été ajouté et le tube a été incubé pendant 1 h à 37 °C.

La concentration en protéines de chacun des échantillons a été déterminée en triplicat grâce à un micro dosage BCA (Kit de dosages de protéines Micro BCA™ Thermo scientific) selon les recommandations du fabricant. La mesure des absorbances a été faite à 562 nm grâce au spectrophotomètre SpectroStar Nano (BMG LabTech).

## XIII. SDS-PAGE et transfert sur membrane de nitrocellulose

Les protéines ont été dénaturées à 100 °C pendant 10 min en présence d'un volume *ad hoc* de tampon Laemmli 5X. Ensuite, un volume équivalent à 20µg (pour un WB) ou 50µg (pour une coloration au bleu de coomassie) de protéines a été déposé ainsi que 6 µL du marqueur de poids moléculaire (Precision Plus Protein Dual Color Standards, Bio-rad) sur un gel SDS-PAGE composé d'un gel de concentration à 4 % d'acrylamide (670 µL acrylamide/bisacrylamide 29/1 30 %, 0,5 mL de Tris 0,5 M à pH 6,8, 40 µL de SDS 10 %, 40 µL d'APS à 10 %, 4 µL de TEMED et 2,7 mL d'eau) et de séparation à 10 % d'acrylamide (3,3 mL d'acrylamide/bisacrylamide 29/1 30 %, 2,5 mL de Tris 1,5 M à pH 8,8, 100 µL de SDS 10 %, 100 µL de APS 10 %, 4 µL de TEMED et 4 mL d'eau). La migration a été réalisée sous une tension de 60 V jusqu'à ce que les échantillons arrivent au niveau du gel de séparation, puis 110 V jusqu'à la sortie du bleu de bromophénol en bas du gel). Dans le cas d'une coloration des protéines par le bleu de coomassie, les gels ont été incubés durant une nuit dans du réactif du dosage Pierce™ Coomassie Plus (Thermo Scientific) puis la décoloration a été faite dans de l'eau jusqu'à ce que les bandes soient parfaitement visibles.

Dans le cas où de la mise en œuvre d'un électrotransfert des protéines séparées sur un gel SDS-PAGE sur une membrane de nitrocellulose, les gels ont été immergés dans un tampon de transfert Tris-Glycine 1X méthanol 20 % puis soumis à une tension de 12 V pendant 1h30 à 12 V en condition semi-sec (Trans-Blot® SD Semi-Dry Transfer Cell, Bio-rad). Après transfert, la membrane de nitrocellulose a été colorée quelques minutes au rouge ponceau jusqu'à apparition de bandes afin de vérifier l'efficacité du transfert puis lavée abondamment à l'eau désionisée.



## XIV. Lectine blot

La membrane a été saturée par une solution de PBS contenant 0,05 % de Tween 20 (T) 5 % de BSA pendant 1 h à Tamb ou une nuit à 4 °C puis incubée avec la lectine marquée à la biotine à la concentration finale de 1 µg/mL pour la Con A (BioWorld), UEA-I (BioWorld) et HPA (MyBioSource) et 0,1 µg/mL pour AAL (US Biological life sciences) dans le PBS-T 0,05 % BSA 5 % pendant 2h à 4 °C sous agitation. La membrane a été ensuite lavée 3 fois pendant 5 min dans du PBS-T 0,05 % avant d'être incubée avec un anticorps antibiotine-HRP (fraction monoclonale d'Ig G de souris à 0,8 mg/mL, Jackson Immunoresearch) dilué au 1/25 000<sup>ème</sup> dans le PBS-T 0,05 % BSA 5 % pendant 1h à Tamb et sous agitation. Après 3 lavages au PBS-T 0,05 % et 2 lavages au PBS pendant 5min sous agitation, la membrane a été révélée par un mélange réactionnel (50/50 ; V/V) de substrats ECL Pico (Thermoscientific) préparé extemporanément, pendant 10 min à l'obscurité. La lecture a été effectuée par une caméra de chimiluminescence (Vilber Lourmat) utilisant le logiciel Fusion.

## XV. Far-Western blot

Pour les expériences de FarWB, le protocole suivi est similaire à celui des lectine-blot à l'exception de l'utilisation du lait en poudre comme agent de blocage à la place de la BSA et des lavages qui ont été faits avec du PBS-T 0,05 % après chaque incubation. La membrane est incubée dans une solution de VLPs GI.1 (Norwalk) ou de VLPs GII.4 Dijon (Accession Number AF472623) et MD145 (Accession Number AY032605) produits par l'équipe du Dr Jacques Le Pendu (UMR Inserm 1232 Centre de Recherche en Cancérologie et Immunologie Nantes-Angers CRCINA) à 3 µg/mL dans le PBS-T 0,05 % lait 5 % une nuit à 4 °C sous agitation. Le système de révélation a été composé d'un anticorps primaire anti-GI.1 ou anti-GII.4 (anticorps monoclonal de souris, Abcam) dilué au 1000<sup>ème</sup> dans le PBS-T 0,05 % lait 5 % et d'un anticorps secondaire de chèvre anti-IgG de souris-HRP (Interchim) dans les mêmes conditions que celles décrites pour les lectine-blot.

## XVI. ELISA

Les échantillons ont été dilués dans du tampon carbonate 50 mM (pH 9,6, ThermoFischer Scientific) et 100 µL ont été ajoutés par puit d'une plaque Nunc Maxisorp immunoplates (ThermoFischer Scientific) suivi d'une incubation toute la nuit à 4 °C. La plaque a été lavée 3 fois par du PBS-T 0,05 % puis les puits ont été saturés par du PBS-lait en poudre à 5 % pendant 1 h à 37 °C sous atmosphère humide. Les VLPs diluées à 1 µg/mL dans le PBS-lait en poudre à 5 % ont été ajoutées et les puits sont incubés une nuit à 4 °C. Après 3 lavages au PBS-T 0,05 %, l'anticorps anti-VLP (dilution au 1/1000<sup>ème</sup> dans le PBS lait en poudre à 5%) (anticorps monoclonal de souris, Abcam) a été ajouté et l'incubation a été réalisée pendant 1 h à 37 °C sous atmosphère humide. Après 3 lavages au PBS-T, l'anticorps secondaire anti-IgG souris-HRP (dilué au 1/1000<sup>ème</sup> dans le PBS lait en poudre à 5%)(Interchim) a été ajouté et la plaque a été incubée pendant 1 h à 37 °C sous atmosphère humide.



Cinq lavages (3 au PBS-T 0,05 % et 2 au PBS) ont été réalisés avant la révélation par le TMB (50  $\mu$ L, pendant 10min). La réaction a été arrêtée par l'ajout de 50  $\mu$ L d'acide phosphorique 1 M avant la mesure des absorbances à 450 nm (SPECTROstar® Nano, BMG LABTEC).

# BIBLIOGRAPHIE

---

## A

---

- Almand, E. A., Moore, M. D. et Jaykus, L.-A. (2017). Norovirus Binding to Ligands Beyond Histo-Blood Group Antigens. *Frontiers in Microbiology*, 8. <https://www.frontiersin.org/articles/10.3389/fmicb.2017.02549>
- Alter, G., Ottenhoff, T. H. M. et Joosten, S. A. (2018). Antibody glycosylation in inflammation, disease and vaccination. *Seminars in Immunology*, 39, 102-110. <https://doi.org/10.1016/j.smim.2018.05.003>
- Altmann, F., Schwihla, H., Staudacher, E., Glössl, J. et März, L. (1995). Insect Cells Contain an Unusual, Membrane-bound  $\beta$ -N-Acetylglucosaminidase Probably Involved in the Processing of Protein N-Glycans. *Journal of Biological Chemistry*, 270(29), 17344-17349. <https://doi.org/10.1074/jbc.270.29.17344>
- Angata, T. et Varki, A. (2002). Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chemical Reviews*, 102(2), 439-469. <https://doi.org/10.1021/cr000407m>
- Aoki, K. et Tiemeyer, M. (2010). The glycomics of glycan glucuronylation in *Drosophila melanogaster*. *Methods in Enzymology*, 480, 297-321. [https://doi.org/10.1016/S0076-6879\(10\)80014-X](https://doi.org/10.1016/S0076-6879(10)80014-X)
- Apweiler, R. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1473(1), 4-8. [https://doi.org/10.1016/S0304-4165\(99\)00165-8](https://doi.org/10.1016/S0304-4165(99)00165-8)
- Argüeso, P. et Gipson, I. K. (2006). Quantitative Analysis of Mucins in Mucosal Secretions Using Indirect Enzyme-Linked Immunosorbent Assay. Dans B. Inka, *Glycobiology Protocols* (vol. 347, p. 277-288). Humana Press. <https://doi.org/10.1385/1-59745-167-3:277>
- Atmar, R. L., Ramani, S. et Estes, M. K. (2018). Human noroviruses: recent advances in a 50-year history. *Current Opinion in Infectious Diseases*, 31(5), 422-432. <https://doi.org/10.1097/QCO.0000000000000476>

## B

---

- Bansil, R. et Turner, B. S. (2006). Mucin structure, aggregation, physiological functions and biomedical applications. *Current Opinion in Colloid & Interface Science*, 11(2-3), 164-170. <https://doi.org/10.1016/j.cocis.2005.11.001>
- Bányai, K., Estes, M. K., Martella, V. et Parashar, U. D. (2018). Viral gastroenteritis. *The Lancet*, 392(10142), 175-186. [https://doi.org/10.1016/S0140-6736\(18\)31128-0](https://doi.org/10.1016/S0140-6736(18)31128-0)
- Bennett, E. P., Mandel, U., Clausen, H., Gerken, T. A., Fritz, T. A. et Tabak, L. A. (2012). Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family. *Glycobiology*, 22(6), 736-756. <https://doi.org/10.1093/glycob/cwr182>
- Black, I., Heiss, C. et Azadi, P. (2019). Comprehensive Monosaccharide Composition Analysis of Insoluble Polysaccharides by Permethylated To Produce Methyl Alditol Derivatives for Gas Chromatography/Mass Spectrometry. *Analytical Chemistry*, 91(21), 13787-13793. <https://doi.org/10.1021/acs.analchem.9b03239>
- Blancher, A. et Socha, W. W. (1997). The ABO, Hh and Lewis Blood Group. in Humans and Nonhuman Primates. Dans A. Blancher, J. Klein et W. Socha (dir.), *Molecular Biology and Evolution of Blood Group and MHC Antigens in Primates* (p. 30-92). Springer. [https://doi.org/10.1007/978-3-642-59086-3\\_3](https://doi.org/10.1007/978-3-642-59086-3_3)
- Bonifacino, J. S. et Weissman, A. M. (1998). Ubiquitin and the Control of Protein Fate in the Secretory and Endocytic Pathways. *Annual Review of Cell and Developmental Biology*, 14(1), 19-57. <https://doi.org/10.1146/annurev.cellbio.14.1.19>

- Booth, C. et Koch, G. L. (1989). Perturbation of cellular calcium induces secretion of luminal ER proteins. *Cell*, 59(4), 729-737. [https://doi.org/10.1016/0092-8674\(89\)90019-6](https://doi.org/10.1016/0092-8674(89)90019-6)
- Botta, R., Asche, F., Borsum, J. S. et Camp, E. V. (2020). A review of global oyster aquaculture production and consumption. *Marine Policy*, 117, 103952. <https://doi.org/10.1016/j.marpol.2020.103952>
- Brockhausen, I., Wandall, H. H., Hagen, K. G. T. et Stanley, P. (2022). O-GalNAc Glycans. Dans A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar et P. H. Seeberger (dir.), *Essentials of Glycobiology* (4th éd.). Cold Spring Harbor Laboratory Press. <http://www.ncbi.nlm.nih.gov/books/NBK579921/>
- Bucardo, F., Kindberg, E., Paniagua, M., Vildevall, M. et Svensson, L. (2009). Genetic susceptibility to symptomatic norovirus infection in Nicaragua. *Journal of Medical Virology*, 81(4), 728-735. <https://doi.org/10.1002/jmv.21426>
- Bull, R. A., Eden, J.-S., Rawlinson, W. D. et White, P. A. (2010). Rapid Evolution of Pandemic Noroviruses of the GII.4 Lineage. *PLoS Pathogens*, 6(3), e1000831. <https://doi.org/10.1371/journal.ppat.1000831>
- Burda, P. et Aebi, M. (1999). The dolichol pathway of N-linked glycosylation. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1426(2), 239-257. [https://doi.org/10.1016/S0304-4165\(98\)00127-5](https://doi.org/10.1016/S0304-4165(98)00127-5)
- Byrd-Leotis, L., Cummings, R. D. et Steinhauer, D. A. (2017). The Interplay between the Host Receptor and Influenza Virus Hemagglutinin and Neuraminidase. *International Journal of Molecular Sciences*, 18(7), 1541. <https://doi.org/10.3390/ijms18071541>

## C

---

- Cao, S., Lou, Z., Tan, M., Chen, Y., Liu, Y., Zhang, Z., Zhang, X. C., Jiang, X., Li, X. et Rao, Z. (2007). Structural basis for the recognition of blood group trisaccharides by norovirus. *Journal of Virology*, 81(11), 5949-5957. <https://doi.org/10.1128/JVI.00219-07>
- Chhabra, P., de Graaf, M., Parra, G. I., Chan, M. C.-W., Green, K., Martella, V., Wang, Q., White, P. A., Katayama, K., Vennema, H., Koopmans, M. P. G. et Vinjé, J. (2019). Updated classification of norovirus genogroups and genotypes. *Journal of General Virology*, 100(10), 1393-1406. <https://doi.org/10.1099/jgv.0.001318>
- Choi, C. et Kingsley, D. H. (2016). Temperature-Dependent Persistence of Human Norovirus Within Oysters (*Crassostrea virginica*). *Food and Environmental Virology*, 8(2), 141-147. <https://doi.org/10.1007/s12560-016-9234-8>
- Cipollo, J. F., Costello, C. E. et Hirschberg, C. B. (2002). The Fine Structure of *Caenorhabditis elegans* N-Glycans \* 210. *Journal of Biological Chemistry*, 277(51), 49143-49157. <https://doi.org/10.1074/jbc.M208020200>
- Ciucanu, I. et Kerek, F. (1984). Rapid and simultaneous methylation of fatty and hydroxy fatty acids for gas—liquid chromatographic analysis. *Journal of Chromatography A*, 284, 179-185. [https://doi.org/10.1016/S0021-9673\(01\)87813-4](https://doi.org/10.1016/S0021-9673(01)87813-4)
- Clausen, H. et Bennett, E. P. (1996). A family of UDP-GalNAc: polypeptide *N*-acetylgalactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation. *Glycobiology*, 6(6), 635-646. <https://doi.org/10.1093/glycob/6.6.635>
- Coutinho, M. F., Prata, M. J. et Alves, S. (2012). Mannose-6-phosphate pathway: A review on its role in lysosomal function and dysfunction. *Molecular Genetics and Metabolism*, 105(4), 542-550. <https://doi.org/10.1016/j.ymgme.2011.12.012>
- Currier, R. L., Payne, D. C., Staat, M. A., Selvarangan, R., Shirley, S. H., Halasa, N., Boom, J. A., Englund, J. A., Szilagyi, P. G., Harrison, C. J., Klein, E. J., Weinberg, G. A., Wikswow, M. E., Parashar, U., Vinjé, J. et Morrow, A. L. (2015). Innate Susceptibility to Norovirus Infections Influenced by FUT2 Genotype in a

## D

---

- de Graaf, M., van Beek, J., Vennema, H., Podkolzin, A. T., Hewitt, J., Bucardo, F., Templeton, K., Mans, J., Nordgren, J., Reuter, G., Lynch, M., Rasmussen, L. D., Iritani, N., Chan, M. C., Martella, V., Ambert-Balay, K., Vinjé, J., White, P. A. et Koopmans, M. P. (2015). Emergence of a novel GII.17 norovirus – End of the GII.4 era? *Eurosurveillance*, 20(26). <https://doi.org/10.2807/1560-7917.ES2015.20.26.21178>
- Delannoy, C., Huang, C., Coddeville, B., Chen, J.-Y., Mouajjah, D., Groux-Degroote, S., Harduin-Lepers, A., Khoo, K.-H., Guerardel, Y. et Ellass-Rochard, E. (2020). Mycobacterium bovis BCG infection alters the macrophage N-glycome. *Molecular Omics*, 16(4), 345-354. <https://doi.org/10.1039/C9MO00173E>
- Dell, A. (1987). F.A.B.-mass spectrometry of carbohydrates. *Advances in Carbohydrate Chemistry and Biochemistry*, 45, 19-72. [https://doi.org/10.1016/s0065-2318\(08\)60136-5](https://doi.org/10.1016/s0065-2318(08)60136-5)
- Desdoutis, M., Polo, D., Le Menec, C., Strubbia, S., Zeng, X.-L., Ettayebi, K., Atmar, R. L., Estes, M. K. et Le Guyader, F. S. (2022). Use of Human Intestinal Enteroids to Evaluate Persistence of Infectious Human Norovirus in Seawater. *Emerging Infectious Diseases*, 28(7), 1475-1479. <https://doi.org/10.3201/eid2807.220219>
- de Rougemont, A., Ruvoen-Clouet, N., Simon, B., Estienney, M., Elie-Caille, C., Aho, S., Pothier, P., Le Pendu, J., Boireau, W. et Belliot, G. (2011). Qualitative and Quantitative Analysis of the Binding of GII.4 Norovirus Variants onto Human Blood Group Antigens. *Journal of Virology*, 85(9), 4057-4070. <https://doi.org/10.1128/JVI.02077-10>
- Di Carluccio, C., Forgione, M. C., Martini, S., Berti, F., Molinaro, A., Marchetti, R. et Silipo, A. (2021). Investigation of protein-ligand complexes by ligand-based NMR methods. *Carbohydrate Research*, 503, 108313. <https://doi.org/10.1016/j.carres.2021.108313>
- Dolashka-Angelova, P., Beck, A., Dolashki, A., Beltramini, M., Stevanovic, S., Salvato, B. et Voelter, W. (2003). Characterization of the carbohydrate moieties of the functional unit RvH1-a of Rapana venosa haemocyanin using HPLC/electrospray ionization MS and glycosidase digestion. *Biochemical Journal*, 374(1), 185-192. <https://doi.org/10.1042/bj20030291>
- Domon, B. et Costello, C. E. (1988). A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconjugate Journal*, 5, 397-409.

## E

---

- Eckmair, B., Jin, C., Abed-Navandi, D. et Paschinger, K. (2016). Multistep Fractionation and Mass Spectrometry Reveal Zwitterionic and Anionic Modifications of the N- and O-glycans of a Marine Snail. *Molecular & Cellular Proteomics*, 15(2), 573-597. <https://doi.org/10.1074/mcp.M115.051573>
- Escudero-Abarca, B. I., Rawsthorne, H., Goulter, R. M., Suh, S. H. et Jaykus, L. A. (2014). Molecular methods used to estimate thermal inactivation of a prototype human norovirus: More heat resistant than previously believed? *Food Microbiology*, 41, 91-95. <https://doi.org/10.1016/j.fm.2014.01.009>
- Ettayebi, K., Crawford, S. E., Murakami, K., Broughman, J. R., Karandikar, U., Tenge, V. R., Neill, F. H., Blutt, S. E., Zeng, X.-L., Qu, L., Kou, B., Opekun, A. R., Burrin, D., Graham, D. Y., Ramani, S., Atmar, R. L. et Estes, M. K. (2016). Replication of human noroviruses in stem cell-derived human enteroids. *Science*, 353(6306), 1387-1393. <https://doi.org/10.1126/science.aaf5211>

## F

---

- Fabini, G., Freilinger, A., Altmann, F. et Wilson, I. B. (2001). Identification of core alpha 1,3-fucosylated glycans and cloning of the requisite fucosyltransferase cDNA from *Drosophila melanogaster*. Potential basis of the neural anti-horseadish peroxidase epitope. *The Journal of Biological Chemistry*, *276*(30), 28058-28067. <https://doi.org/10.1074/jbc.M100573200>
- Fiege, B., Leuthold, M., Parra, F., Dalton, K. P., Meloncelli, P. J., Lowary, T. L. et Peters, T. (2017). Epitope mapping of histo blood group antigens bound to norovirus VLPs using STD NMR experiments reveals fine details of molecular recognition. *Glycoconjugate Journal*, *34*(5), 679-689. <https://doi.org/10.1007/s10719-017-9792-5>
- Fournet, B., Strecker, G., Leroy, Y. et Montreuil, J. (1981). Gas-liquid chromatography and mass spectrometry of methylated and acetylated methyl glycosides. Application to the structural analysis of glycoprotein glycans. *Analytical Biochemistry*, *116*(2), 489-502. [https://doi.org/10.1016/0003-2697\(81\)90393-6](https://doi.org/10.1016/0003-2697(81)90393-6)
- Fuchs, B. et Schiller, J. (2009). Recent Developments of Useful MALDI Matrices for the Mass Spectrometric Characterization of Apolar Compounds. *Current Organic Chemistry*, *13*(16), 1664-1681. <https://doi.org/10.2174/138527209789578108>

## G

---

- Gao, X., Esseili, M. A., Lu, Z., Saif, L. J. et Wang, Q. (2016). Recognition of Histo-Blood Group Antigen-Like Carbohydrates in Lettuce by Human GII.4 Norovirus. *Applied and Environmental Microbiology*, *82*(10), 2966-2974. <https://doi.org/10.1128/AEM.04096-15>
- Geisler, C. et Jarvis, D. L. (2012). Substrate specificities and intracellular distributions of three N-glycan processing enzymes functioning at a key branch point in the insect N-glycosylation pathway. *The Journal of Biological Chemistry*, *287*(10), 7084-7097. <https://doi.org/10.1074/jbc.M111.296814>
- Grewal, P. K. (2010). The Ashwell–Morell Receptor. Dans *Methods in Enzymology* (vol. 479, p. 223-241). Elsevier. [https://doi.org/10.1016/S0076-6879\(10\)79013-3](https://doi.org/10.1016/S0076-6879(10)79013-3)
- Griffin, D. W., Donaldson, K. A., Paul, J. H. et Rose, J. B. (2003). Pathogenic Human Viruses in Coastal Waters. *Clinical Microbiology Reviews*, *16*(1), 129-143. <https://doi.org/10.1128/CMR.16.1.129-143.2003>
- Guérardel, Y., Balanzino, L., Maes, E., Leroy, Y., Coddeville, B., Oriol, R. et Strecker, G. (2001). The nematode *Caenorhabditis elegans* synthesizes unusual O-linked glycans: identification of glucose-substituted mucin-type O-glycans and short chondroitin-like oligosaccharides. *Biochemical Journal*, *357*(Pt 1), 167-182.
- Guérardel, Y., Chang, L.-Y., Maes, E., Huang, C.-J. et Khoo, K.-H. (2006). Glycomic survey mapping of zebrafish identifies unique sialylation pattern. *Glycobiology*, *16*(3), 244-257. <https://doi.org/10.1093/glycob/cwj062>
- Gui, B., Yao, L., Qu, M., Zhang, W., Li, M., Jiang, Y. et Wang, L. (2023). Cloning, Expression, and Functional Characterization of FUT1, a Key Gene for Histo-Blood Group Antigens Synthesis in *Crassostrea gigas*. *Current Issues in Molecular Biology*, *45*(5), 4200-4213. <https://doi.org/10.3390/cimb45050267>
- Gutternigg, M., Ahrer, K., Grabher-Meier, H., Bürgmayr, S. et Staudacher, E. (2004). Neutral N-glycans of the gastropod *Arion lusitanicus*. *European Journal of Biochemistry*, *271*(7), 1348-1356. <https://doi.org/10.1111/j.1432-1033.2004.04045.x>
- Gutternigg, M., Bürgmayr, S., Pörtl, G., Rudolf, J. et Staudacher, E. (2007a). Neutral N-glycan patterns of the gastropods *Limax maximus*, *Cepaea hortensis*, *Planorbarius corneus*, *Arianta arbustorum* and *Achatina fulica*. *Glycoconjugate Journal*, *24*(8), 475-489. <https://doi.org/10.1007/s10719-007-9040-5>

Gutternigg, M., Kretschmer-Lubich, D., Paschinger, K., Rendić, D., Hader, J., Geier, P., Ranftl, R., Jantsch, V., Lochnit, G. et Wilson, I. B. H. (2007b). Biosynthesis of Truncated N-Linked Oligosaccharides Results from Non-orthologous Hexosaminidase-mediated Mechanisms in Nematodes, Plants, and Insects. *Journal of Biological Chemistry*, 282(38), 27825-27840. <https://doi.org/10.1074/jbc.M704235200>

## H

---

- Hagen, F. K. et Nehrke, K. (1998). cDNA cloning and expression of a family of UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase sequence homologs from *Caenorhabditis elegans*. *The Journal of Biological Chemistry*, 273(14), 8268-8277. <https://doi.org/10.1074/jbc.273.14.8268>
- Hagen, K. G. T., Bedi, G. S., Tetaert, D., Kingsley, P. D., Hagen, F. K., Balys, M. M., Beres, T. M., Degand, P. et Tabak, L. A. (2001). Cloning and Characterization of a Ninth Member of the UDP-GalNAc:Polypeptide N-Acetylgalactosaminyltransferase Family, ppGalNTase-T9. *Journal of Biological Chemistry*, 276(20), 17395-17404. <https://doi.org/10.1074/jbc.M009638200>
- Hall, R. L., Wood, E. J., Kamberling, J. P., Gerwig, G. J. et Vliegenthart, F. G. (1977). 3-O-methyl sugars as constituents of glycoproteins. Identification of 3-O-methylgalactose and 3-O-methylmannose in pulmonate gastropod haemocyanins. *Biochemical Journal*, 165(1), 173-176. <https://doi.org/10.1042/bj1650173>
- Hård, K., Van Doorn, J. M., Thomas-Oates, J. E., Kamerling, J. P. et Van der Horst, D. J. (1993). Structure of the asn-linked oligosaccharides of apolipoprotein III from the insect *Locusta migratoria*. Carbohydrate-linked 2-aminoethylphosphonate as a constituent of a glycoprotein. *Biochemistry*, 32(3), 766-775. <https://doi.org/10.1021/bi00054a005>
- Haslam, S. M., Coles, G. C., Morris, H. R. et Dell, A. (2000). Structural characterization of the N-glycans of *Dictyocaulus viviparus*: discovery of the Lewisx structure in a nematode. *Glycobiology*, 10(2), 223-229. <https://doi.org/10.1093/glycob/10.2.223>
- Haslam, S. M., Coles, G. C., Munn, E. A., Smith, T. S., Smith, H. F., Morris, H. R. et Dell, A. (1996). Haemonchus contortus Glycoproteins Contain N-Linked Oligosaccharides with Novel Highly Fucosylated Core Structures \*. *Journal of Biological Chemistry*, 271(48), 30561-30570. <https://doi.org/10.1074/jbc.271.48.30561>
- Haslam, S. M., Freedberg, D. I., Mulloy, B., Dell, A., Stanley, P. et Prestegard, J. H. (2022). Structural Analysis of Glycans. Dans A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar et P. H. Seeberger (dir.), *Essentials of Glycobiology* (4th éd.). Cold Spring Harbor Laboratory Press. <http://www.ncbi.nlm.nih.gov/books/NBK579945/>
- Hecht, E. S., Scigelova, M., Eliuk, S. et Makarov, A. (2019). Fundamentals and Advances of Orbitrap Mass Spectrometry. Dans *Encyclopedia of Analytical Chemistry* (p. 1-40). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470027318.a9309.pub2>
- Hillenkamp, F., Jaskolla, T. W. et Karas, M. (2013). The MALDI Process and Method. Dans *MALDI MS* (p. 1-40). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9783527335961.ch1>
- Hoffmann, E. de et Stroobant, V. (2007a). Ion Sources. Dans *Mass Spectrometry: Principles and Applications*. John Wiley & Sons.
- Hoffmann, E. de et Stroobant, V. (2007b). Mass Analysers. Dans *Mass Spectrometry: Principles and Applications*. John Wiley & Sons.
- Hori, T., Sugita, M., Ando, S., Kuwahara, M., Kumauchi, K., Sugie, E. et Itasaka, O. (1981). Characterization of a novel glycosphingolipid, ceramide nonasaccharide, isolated from spermatozoa of the fresh water bivalve, *Hyriopsis schlegelii*. *The Journal of Biological Chemistry*, 256(21), 10979-10985.

- Hsiao, C.-T., Wang, P.-W., Chang, H.-C., Chen, Y.-Y., Wang, S.-H., Chern, Y. et Khoo, K.-H. (2017). Advancing a High Throughput Glycotope-centric Glycomics Workflow Based on NAnoLC-MS2-product Dependent-MS3 ANALYSIS of Permethylyated Glycans\*. *Molecular & Cellular Proteomics*, 16(12), 2268-2280. <https://doi.org/10.1074/mcp.TIR117.000156>
- Hu, L., Crawford, S. E., Czako, R., Cortes-Penfield, N. W., Smith, D. F., Le Pendu, J., Estes, M. K. et Prasad, B. V. V. (2012). Cell attachment protein VP8\* of a human rotavirus specifically interacts with A-type histo-blood group antigen. *Nature*, 485(7397), 256-259. <https://doi.org/10.1038/nature10996>
- Huang, P., Farkas, T., Marionneau, S., Zhong, W., Ruvoën-Clouet, N., Morrow, A. L., Altaye, M., Pickering, L. K., Newburg, D. S., LePendou, J. et Jiang, X. (2003). Noroviruses bind to human ABO, Lewis, and secretor histo-blood group antigens: identification of 4 distinct strain-specific patterns. *The Journal of Infectious Diseases*, 188(1), 19-31. <https://doi.org/10.1086/375742>
- Huang, P., Farkas, T., Zhong, W., Tan, M., Thornton, S., Morrow, A. L. et Jiang, X. (2005). Norovirus and histo-blood group antigens: demonstration of a wide spectrum of strain specificities and classification of two major binding groups among multiple binding patterns. *Journal of Virology*, 79(11), 6714-6722. <https://doi.org/10.1128/JVI.79.11.6714-6722.2005>
- Huang, P., Morrow, A. L. et Jiang, X. (2009). The carbohydrate moiety and high molecular weight carrier of histo-blood group antigens are both required for norovirus-receptor recognition. *Glycoconjugate Journal*, 26(8), 1085-1096. <https://doi.org/10.1007/s10719-009-9229-x>
- Hulsen, T., de Vlieg, J. et Alkema, W. (2008). BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics*, 9(1), 488. <https://doi.org/10.1186/1471-2164-9-488>
- Hutson, A. M., Airaud, F., LePendou, J., Estes, M. K. et Atmar, R. L. (2005). Norwalk virus infection associates with secretor status genotyped from sera. *Journal of Medical Virology*, 77(1), 116-120. <https://doi.org/10.1002/jmv.20423>
- Hutson, A. M., Atmar, R. L. et Estes, M. K. (2004). Norovirus disease: changing epidemiology and host susceptibility factors. *Trends in Microbiology*, 12(6), 279-287. <https://doi.org/10.1016/j.tim.2004.04.005>
- Hutson, A. M., Atmar, R. L., Marcus, D. M. et Estes, M. K. (2003). Norwalk Virus-Like Particle Hemagglutination by Binding to H Histo-Blood Group Antigens. *Journal of Virology*, 77(1), 405-415. <https://doi.org/10.1128/JVI.77.1.405-415.2003>
- Hykollari, A., Malzl, D., Eckmair, B., Vanbeselaere, J., Scheidl, P., Karlsson, N. G., Wilson, I. B. H. et Paschinger, K. (2018). Isomeric Separation and Recognition of Anionic and Zwitterionic N-glycans from Royal Jelly Glycoproteins. *Molecular & Cellular Proteomics: MCP*, 17(11), 2177-2196. <https://doi.org/10.1074/mcp.RA117.000462>

## I

---

- Ichimiya, T., Maeda, M., Sakamura, S., Kanazawa, M., Nishihara, S. et Kimura, Y. (2015). Identification of  $\beta$ 1,3-galactosyltransferases responsible for biosynthesis of insect complex-type N-glycans containing a T-antigen unit in the honeybee. *Glycoconjugate Journal*, 32(3-4), 141-151. <https://doi.org/10.1007/s10719-015-9585-7>
- Ishihara, H., Takahashi, N., Oguri, S. et Tejima, S. (1979). Complete structure of the carbohydrate moiety of stem bromelain. An application of the almond glycopeptidase for structural studies of glycopeptides. *Journal of Biological Chemistry*, 254(21), 10715-10719. [https://doi.org/10.1016/S0021-9258\(19\)86580-8](https://doi.org/10.1016/S0021-9258(19)86580-8)
- Israr, M. Z., Bernieh, D., Salzano, A., Cassambai, S., Yazaki, Y. et Suzuki, T. (2020). Matrix-assisted laser desorption ionisation (MALDI) mass spectrometry (MS): basics and clinical applications. *Clinical Chemistry and Laboratory Medicine (CCLM)*, 58(6), 883-896. <https://doi.org/10.1515/cclm-2019-0868>



Iwase, H. et Hotta, K. (1993). Release of O-linked glycoprotein glycans by endo-alpha-N-acetylgalactosaminidase. *Methods in Molecular Biology (Clifton, N.J.)*, 14, 151-159. <https://doi.org/10.1385/0-89603-226-4:151>

## J

---

Jaskolla, T. W. et Karas, M. (2011). Compelling Evidence for Lucky Survivor and Gas Phase Protonation: The Unified MALDI Analyte Protonation Mechanism. *Journal of the American Society for Mass Spectrometry*, 22(6), 976-988. <https://doi.org/10.1007/s13361-011-0093-0>

Ji, Y., White, Y. J., Hadden, J. A., Grant, O. C. et Woods, R. J. (2017). New insights into influenza A specificity: an evolution of paradigms. *Current Opinion in Structural Biology*, 44, 219-231. <https://doi.org/10.1016/j.sbi.2017.06.001>

Jiang, X., Wang, M., Graham, D. Y. et Estes, M. K. (1992). Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. *Journal of Virology*, 66(11), 6527-6532. <https://doi.org/10.1128/jvi.66.11.6527-6532.1992>

Jiménez-Castells, C., Vanbeselaere, J., Kohlhuber, S., Ruttkowski, B., Joachim, A. et Paschinger, K. (2017). Gender and developmental specific N-glycomes of the porcine parasite *Oesophagostomum dentatum*. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1861(2), 418-430. <https://doi.org/10.1016/j.bbagen.2016.10.011>

Jin, M., He, Y., Li, H., Huang, P., Zhong, W., Yang, H., Zhang, H., Tan, M. et Duan, Z. (2013). Two gastroenteritis outbreaks caused by GII Noroviruses: host susceptibility and HBGA phenotypes. *PloS One*, 8(3), e58605. <https://doi.org/10.1371/journal.pone.0058605>

Johansen, P. G., Marshall, R. D. et Neuberger, A. (1961). The preparation and some of the properties of a glycopeptide from Hen's-egg albumin. *Biochemical Journal*, 78, 518-527. <https://doi.org/10.1042/bj0780518>

## K

---

Karas, Michael. et Hillenkamp, Franz. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*, 60(20), 2299-2301. <https://doi.org/10.1021/ac00171a028>

Kawar, Z. S., Die, I. V. et Cummings, R. D. (2002). Molecular Cloning and Enzymatic Characterization of a UDP-GalNAc:GlcNAc $\beta$ -R  $\beta$ 1,4-N-Acetylgalactosaminyltransferase from *Caenorhabditis elegans* \*. *Journal of Biological Chemistry*, 277(38), 34924-34932. <https://doi.org/10.1074/jbc.M206112200>

Kelly, R. J., Rouquier, S., Giorgi, D., Lennon, G. G. et Lowe, J. B. (1995). Sequence and Expression of a Candidate for the Human Secretor Blood Group  $\alpha$ (1,2)Fucosyltransferase Gene (FUT2). *Journal of Biological Chemistry*, 270(9), 4640-4649. <https://doi.org/10.1074/jbc.270.9.4640>

Kendra, J. A., Tohma, K. et Parra, G. I. (2022). Global and regional circulation trends of norovirus genotypes and recombinants, 1995–2019: A comprehensive review of sequences from public databases. *Reviews in Medical Virology*, 32(5), e2354. <https://doi.org/10.1002/rmv.2354>

Khoo, K. H., Maizels, R. M., Page, A. P., Taylor, G. W., Rendell, N. B. et Dell, A. (1991). Characterization of nematode glycoproteins: the major O-glycans of *Toxocara* excretory-secretory antigens are O-methylated trisaccharides. *Glycobiology*, 1(2), 163-171. <https://doi.org/10.1093/glycob/1.2.163>

Khoo, K.-H., Chatterjee, D., Caulfield, J. P., Morris, H. R. et Dell, A. (1997). Structural mapping of the glycans from the egg glycoproteins of *Schistosoma mansoni* and *Schistosoma japonicum*: identification of novel core structures and terminal sequences. *Glycobiology*, 7(5), 663-677. <https://doi.org/10.1093/glycob/7.5.663>

- Kim, C.-H. (2020). SARS-CoV-2 Evolutionary Adaptation toward Host Entry and Recognition of Receptor O-Acetyl Sialylation in Virus–Host Interaction. *International Journal of Molecular Sciences*, *21*(12), 4549. <https://doi.org/10.3390/ijms21124549>
- Kimura-Someya, T., Kato-Murayama, M., Katsura, K., Sakai, N., Murayama, K., Hanada, K., Shirouzu, M. et al. (2022). Lewis fucose is a key moiety for the recognition of histo-blood group antigens by GI.9 norovirus, as revealed by structural analysis. *FEBS Open Bio*, *12*(3), 560-570. <https://doi.org/10.1002/2211-5463.13370>
- Kindberg, E., Hejdeman, B., Bratt, G., Wahren, B., Lindblom, B., Hinkula, J. et al. (2006). A nonsense mutation (428G→A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection. *AIDS*, *20*(5), 685. <https://doi.org/10.1097/01.aids.0000216368.23325.bc>
- Kocak, A. (2019). HBGA binding modes and selectivity in noroviruses upon mutation: a docking and molecular dynamics study. *Journal of Molecular Modeling*, *25*(12), 369. <https://doi.org/10.1007/s00894-019-4261-7>
- Koganti, R., Memon, A. et al. (2021). Emerging Roles of Heparan Sulfate Proteoglycans in Viral Pathogenesis. *Seminars in Thrombosis and Hemostasis*, *47*(03), 283-294. <https://doi.org/10.1055/s-0041-1725068>
- Kornfeld, R. et al. (1985). *Assembly of asparagine-linked oligosaccharides*, *54*, 631-664. <https://doi.org/10.1146/annurev.bi.54.070185.003215>
- Koromyslova, A., Tripathi, S., Morozov, V., Schrotten, H. et al. (2017). Human norovirus inhibition by a human milk oligosaccharide. *Virology*, *508*, 81-89. <https://doi.org/10.1016/j.virol.2017.04.032>
- Kubelka, V., Altmann, F., Staudacher, E., Tretter, V., März, L., Hård, K., Kamerling, J. P. et al. (1993). Primary structures of the N-linked carbohydrate chains from honeybee venom phospholipase A2. *European Journal of Biochemistry*, *213*(3), 1193-1204. <https://doi.org/10.1111/j.1432-1033.1993.tb17870.x>
- Kurz, S., Aoki, K., Jin, C., Karlsson, N. G., Tiemeyer, M., Wilson, I. B. H. et al. (2015). Targeted release and fractionation reveal glucuronylated and sulphated N- and O-glycans in larvae of dipteran insects. *Journal of Proteomics*, *126*, 172-188. <https://doi.org/10.1016/j.jprot.2015.05.030>
- Kurz, S., Jin, C., 金春生, Hykollari, A., Gregorich, D., Giomarelli, B., Vasta, G. R., Wilson, I. B. H. et al. (2013). Hemocytes and Plasma of the Eastern Oyster (*Crassostrea virginica*) Display a Diverse Repertoire of Sulfated and Blood Group A-modified N-Glycans\*. *Journal of Biological Chemistry*, *288*(34), 24410-24428. <https://doi.org/10.1074/jbc.M113.478933>
- Kurz, S., King, J. G., Dinglasan, R. R., Paschinger, K. et al. (2016). The fucosylated N-glycan epitopes and their cognate fucosyltransferases. *Insect Biochemistry and Molecular Biology*, *68*, 52-63. <https://doi.org/10.1016/j.ibmb.2015.11.001>
- Kyprianou, P., Betteridge, A., Donald, A. S. R. et al. (1990). Purification of the blood group H gene associated  $\alpha$ -2-L-fucosyltransferase from human plasma. *Glycoconjugate Journal*, *7*(6), 573-588. <https://doi.org/10.1007/BF01189078>

## L

---

- Larsen, R. D., Ernst, L. K. et al. (1990). *Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase cDNA that can form the H blood group antigen*. <https://www.pnas.org/doi/10.1073/pnas.87.17.6674>
- Larsson, M. M., Rydell, G. E. P., Grahn, A., Rodríguez-Díaz, J., Åkerlind, B., Hutson, A. M., Estes, M. K., Larson, G. et al. (2006). Antibody Prevalence and Titer to Norovirus (Genogroup II) Correlate with

- Secretor (FUT2) but Not with ABO Phenotype or Lewis (FUT3) Genotype. *The Journal of Infectious Diseases*, 194(10), 1422-1427. <https://doi.org/10.1086/508430>
- Le Guyader, F. S., Atmar, R. L. et Le Pendu, J. (2012). Transmission of viruses through shellfish: when specific ligands come into play. *Current Opinion in Virology*, 2(1), 103-110. <https://doi.org/10.1016/j.coviro.2011.10.029>
- Le Guyader, F. S., Le Saux, J.-C., Ambert-Balay, K., Krol, J., Serais, O., Parnaudeau, S., Giraudon, H., Delmas, G., Pommepuy, M., Pothier, P. et Atmar, R. L. (2008). Aichi Virus, Norovirus, Astrovirus, Enterovirus, and Rotavirus Involved in Clinical Cases from a French Oyster-Related Gastroenteritis Outbreak. *Journal of Clinical Microbiology*, 46(12), 4011-4017. <https://doi.org/10.1128/JCM.01044-08>
- Le Guyader, F. S., Loisy, F., Atmar, R. L., Hutson, A. M., Estes, M. K., Ruvoën-Clouet, N., Pommepuy, M. et Le Pendu, J. (2006). Norwalk Virus-specific Binding to Oyster Digestive Tissues. *Emerging Infectious Diseases*, 12(6), 931-936. <https://doi.org/10.3201/eid1206.051519>
- Le Guyader, S. F., Garry, P., Ollivier, J. et Piquet, J.-C. (2019). Microbiologie et coquillages : des enjeux et des développements futurs. *Bulletin de l'Académie Vétérinaire de France*, 172(1), 28-33. <https://doi.org/10.4267/2042/70602>
- Lederkremer, G. Z. (2009). Glycoprotein folding, quality control and ER-associated degradation. *Current Opinion in Structural Biology*, 19(5), 515-523. <https://doi.org/10.1016/j.sbi.2009.06.004>
- Lehr, T., Geyer, H., Maaß, K., Doenhoff, M. J. et Geyer, R. (2007). Structural characterization of N-glycans from the freshwater snail *Biomphalaria glabrata* cross-reacting with *Schistosoma mansoni* glycoconjugates. *Glycobiology*, 17(1), 82-103. <https://doi.org/10.1093/glycob/cwl048>
- Lei, M., Novotny, M. V. et Mechref, Y. (2010). Sequential enrichment of sulfated glycans by strong anion-exchange chromatography prior to mass spectrometric measurements. *Journal of the American Society for Mass Spectrometry*, 21(3), 348-357. <https://doi.org/10.1016/j.jasms.2009.09.017>
- Léonard, R., Rendić, D., Rabouille, C., Wilson, I. B. H., Prêat, T. et Altmann, F. (2006). The *Drosophila* fused lobes Gene Encodes an N-Acetylglucosaminidase Involved in N-Glycan Processing. *Journal of Biological Chemistry*, 281(8), 4867-4875. <https://doi.org/10.1074/jbc.M511023200>
- Liang, Y., Wang, W. B., Zhang, J., Hou, J. W., Tang, F., Zhang, X. F., Du, L. F., Su, J. G. et Li, Q. M. (2021). Evolution of the interactions between GII.4 noroviruses and histo-blood group antigens: Insights from experimental and computational studies. *PLOS Pathogens*, 17(7), e1009745. <https://doi.org/10.1371/journal.ppat.1009745>
- Lin, S. S., Hanaway, M. J., Gonzalez-Stawinski, G. V., Lau, C. L., Parker, W., Davis, R. D., Byrne, G. W., Diamond, L. E., Logan, J. S. et Platt, J. L. (2000). The role of anti-Gal $\alpha$ 1-3Gal antibodies in acute vascular rejection and accommodation of xenografts. *Transplantation*, 70(12), 1667-1674. <https://doi.org/10.1097/00007890-200012270-00002>
- Lindesmith, L. C., Donaldson, E. F., LoBue, A. D., Cannon, J. L., Zheng, D.-P., Vinje, J. et Baric, R. S. (2008). Mechanisms of GII.4 Norovirus Persistence in Human Populations. *PLOS Medicine*, 5(2), e31. <https://doi.org/10.1371/journal.pmed.0050031>
- Lindesmith, L., Moe, C., Lependu, J., Frelinger, J. A., Treanor, J. et Baric, R. S. (2005). Cellular and humoral immunity following Snow Mountain virus challenge. *Journal of Virology*, 79(5), 2900-2909. <https://doi.org/10.1128/JVI.79.5.2900-2909.2005>
- Lindesmith, L., Moe, C., Marionneau, S., Ruvoen, N., Jiang, X., Lindblad, L., Stewart, P., LePendu, J. et Baric, R. (2003). Human susceptibility and resistance to Norwalk virus infection. *Nature Medicine*, 9(5), 548-553. <https://doi.org/10.1038/nm860>
- Liu, M., van Kuppeveld, F. J., de Haan, C. A. et de Vries, E. (2023). Gradual adaptation of animal influenza A viruses to human-type sialic acid receptors. *Current Opinion in Virology*, 60, 101314. <https://doi.org/10.1016/j.coviro.2023.101314>

Lönngren, J. et Svensson, S. (1974). Mass Spectrometry in Structural Analysis of Natural Carbohydrates\*\*This work was supported by Statens Naturvetenskapliga Forskningsråd. Dans R. S. Tipson et D. Horton (dir.), *Advances in Carbohydrate Chemistry and Biochemistry* (vol. 29, p. 41-106). Academic Press. [https://doi.org/10.1016/S0065-2318\(08\)60248-6](https://doi.org/10.1016/S0065-2318(08)60248-6)

Lucas, T. M., Gupta, C., Altman, M. O., Sanchez, E., Naticchia, M. R., Gagneux, P., Singharoy, A. et Godula, K. (2021). Mucin-mimetic glycan arrays integrating machine learning for analyzing receptor pattern recognition by influenza A viruses. *Chem*, 7(12), 3393-3411. <https://doi.org/10.1016/j.chempr.2021.09.015>

## M

---

Maalouf, H., Pommepuy, M. et Le Guyader, F. S. (2010a). Environmental Conditions Leading to Shellfish Contamination and Related Outbreaks. *Food and Environmental Virology*, 2(3), 136-145. <https://doi.org/10.1007/s12560-010-9043-4>

Maalouf, H., Schaeffer, J., Parnaudeau, S., Le Pendu, J., Atmar, R. L., Crawford, S. E. et Le Guyader, F. S. (2011). Strain-Dependent Norovirus Bioaccumulation in Oysters. *Applied and Environmental Microbiology*, 77(10), 3189-3196. <https://doi.org/10.1128/AEM.03010-10>

Maalouf, H., Zakhour, M., Le Pendu, J., Le Saux, J.-C., Atmar, R. L. et Le Guyader, F. S. (2010b). Distribution in Tissue and Seasonal Variation of Norovirus Genogroup I and II Ligands in Oysters. *Applied and Environmental Microbiology*, 76(16), 5621-5630. <https://doi.org/10.1128/AEM.00148-10>

Macher, B. A. et Galili, U. (2008). The Gal $\alpha$ 1,3Gal $\beta$ 1,4GlcNAc-R ( $\alpha$ -Gal) epitope: A carbohydrate of unique evolution and clinical relevance. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1780(2), 75-88. <https://doi.org/10.1016/j.bbagen.2007.11.003>

Maley, F., Trimble, R. B., Tarentino, A. L. et Plummer, T. H. (1989). Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Analytical Biochemistry*, 180(2), 195-204. [https://doi.org/10.1016/0003-2697\(89\)90115-2](https://doi.org/10.1016/0003-2697(89)90115-2)

March, R. E. (2009). Quadrupole ion traps. *Mass Spectrometry Reviews*, 28(6), 961-989. <https://doi.org/10.1002/mas.20250>

Marionneau, S., Cailleau-Thomas, A., Rocher, J., Le Moullac-Vaidye, B., Ruvoën, N., Clément, M. et Le Pendu, J. (2001). ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. *Biochimie*, 83(7), 565-573. [https://doi.org/10.1016/S0300-9084\(01\)01321-9](https://doi.org/10.1016/S0300-9084(01)01321-9)

Marionneau, S., Ruvoën, N., Le Moullac-Vaidye, B., Clément, M., Cailleau-Thomas, A., Ruiz-Palacois, G., Huang, P., Jiang, X. et Le Pendu, J. (2002). Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals. *Gastroenterology*, 122(7), 1967-1977. <https://doi.org/10.1053/gast.2002.33661>

Marks, G. S., Marshall, R. D. et Neuberger, A. (1963). Studies on the carbohydrate-peptide bond in hen's-egg albumin. *Biochemical Journal*, 87(2), 274-281.

Marshall, R. D. (1972). Glycoproteins. *Annual Review of Biochemistry*, 41(1), 673-702. <https://doi.org/10.1146/annurev.bi.41.070172.003325>

Marshall, R. D. et Neuberger, A. (1964). *The Isolation of 2-Acetamido-1-(L- $\beta$ -aspartamido)-1,2-dideoxy- $\beta$ -D-glucose from Hen's Egg Albumin* [research-article]. ACS Publications. American Chemical Society. <https://pubs.acs.org/doi/pdf/10.1021/bi00898a036>

Marteil, L. (1976). La conchyliculture française. 2eme partie. Biologie de l'huitre et de la moule. *Revue des Travaux de l'Institut des Pêches Maritimes*, 40(2), 149-346.

- McEver, R. P. (2002). Selectins: lectins that initiate cell adhesion under flow. *Current Opinion in Cell Biology*, *14*(5), 581-586. [https://doi.org/10.1016/S0955-0674\(02\)00367-8](https://doi.org/10.1016/S0955-0674(02)00367-8)
- Miura, T., Sano, D., Suenaga, A., Yoshimura, T., Fuzawa, M., Nakagomi, T., Nakagomi, O. et Okabe, S. (2013). Histo-blood group antigen-like substances of human enteric bacteria as specific adsorbents for human noroviruses. *Journal of Virology*, *87*(17), 9441-9451. <https://doi.org/10.1128/JVI.01060-13>
- Möckl, L. (2020). The Emerging Role of the Mammalian Glycocalyx in Functional Membrane Organization and Immune System Regulation. *Frontiers in Cell and Developmental Biology*, *8*. <https://www.frontiersin.org/articles/10.3389/fcell.2020.00253>
- Morell, A. G., Gregoriadis, G., Scheinberg, I. H., Hickman, J. et Ashwell, G. (1971). The Role of Sialic Acid in Determining the Survival of Glycoproteins in the Circulation. *Journal of Biological Chemistry*, *246*(5), 1461-1467. [https://doi.org/10.1016/S0021-9258\(19\)76994-4](https://doi.org/10.1016/S0021-9258(19)76994-4)
- Morelle, W., Faid, V. et Michalski, J.-C. (2004). Structural analysis of permethylated oligosaccharides using electrospray ionization quadrupole time-of-flight tandem mass spectrometry and deuterio-reduction. *Rapid Communications in Mass Spectrometry*, *18*(20), 2451-2464. <https://doi.org/10.1002/rcm.1640>
- Morelle, W. et Michalski, J.-C. (2007). Analysis of protein glycosylation by mass spectrometry. *Nature Protocols*, *2*(7), 1585-1602. <https://doi.org/10.1038/nprot.2007.227>
- Moremen, K., Trimble, R. B. et Herscovics, A. (1994). Glycosidases of the asparagine-linked oligosaccharide processing pathway. *Glycobiology*, *4*(2), 113-125. <https://doi.org/10.1093/glycob/4.2.113>
- Morozov, V., Hanisch, F.-G., Wegner, K. M. et Schrotten, H. (2018). Pandemic GII.4 Sydney and Epidemic GII.17 Kawasaki308 Noroviruses Display Distinct Specificities for Histo-Blood Group Antigens Leading to Different Transmission Vector Dynamics in Pacific Oysters. *Frontiers in Microbiology*, *9*, 2826. <https://doi.org/10.3389/fmicb.2018.02826>
- Mucha, J., Domlatil, J., Lochnit, G., Rendić, D., Paschinger, K., HINTERKÖRNER, G., Hofinger, A., Kosma, P. et Wilson, I. B. H. (2004). The Drosophila melanogaster homologue of the human histo-blood group Pk gene encodes a glycolipid-modifying  $\alpha$ 1,4-N-acetylgalactosaminyltransferase. *Biochemical Journal*, *382*(Pt 1), 67-74. <https://doi.org/10.1042/BJ20040535>

## N

---

- Natsuka, S., Adachi, J., Kawaguchi, M., Nakakita, S., Hase, S., Ichikawa, A. et Ikura, K. (2002). Structural Analysis of N-Linked Glycans in *Caenorhabditis elegans*. *The Journal of Biochemistry*, *131*(6), 807-813.
- Natsuka, S., Hirohata, Y., Nakakita, S., Sumiyoshi, W. et Hase, S. (2011). Structural analysis of N-glycans of the planarian *Dugesia japonica*. *The FEBS Journal*, *278*(3), 452-460. <https://doi.org/10.1111/j.1742-4658.2010.07966.x>
- Neelamegham, S., Aoki-Kinoshita, K., Bolton, E., Frank, M., Lisacek, F., Lütteke, T., O'Boyle, N., Packer, N. H., Stanley, P., Toukach, P., Varki, A., Woods, R. J. et The SNFG Discussion Group. (2019). Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology*, *29*(9), 620-624. <https://doi.org/10.1093/glycob/cwz045>
- Nordgren, J., Nitiema, L. W., Ouermi, D., Simpore, J. et Svensson, L. (2013). Host Genetic Factors Affect Susceptibility to Norovirus Infections in Burkina Faso. *PLoS ONE*, *8*(7), e69557. <https://doi.org/10.1371/journal.pone.0069557>
- Nordgren, J. et Svensson, L. (2019). Genetic Susceptibility to Human Norovirus Infection: An Update. *Viruses*, *11*(3), 226. <https://doi.org/10.3390/v11030226>

## O

---

- Oriol, R., Barthod, F., Bergemer, A.-M., Ye, Y., Koren, E. et Cooper, D. K. C. (1994). Monomorphic and polymorphic carbohydrate antigens on pig tissues: implications for organ xenotransplantation in the pig-to-human model. *Transplant International*, *7*(6), 405-413. <https://doi.org/10.1111/j.1432-2277.1994.tb01259.x>
- Oriol, R., Candelier, J., Taniguchi, S., Balanzino, L., Peters, L., Niekrasz, M., Hammer, C. et Cooper, D. K. C. (1999). Major carbohydrate epitopes in tissues of domestic and African wild animals of potential interest for xenotransplantation research. *Xenotransplantation*, *6*(2), 79-89. <https://doi.org/10.1034/j.1399-3089.1999.00008.x>
- Oriol, R., Ye, Y., Koren, E. et Cooper, D. K. (1993). Carbohydrate antigens of pig tissues reacting with human natural antibodies as potential targets for hyperacute vascular rejection in pig-to-man organ xenotransplantation. *Transplantation*, *56*(6), 1433-1442. <https://doi.org/10.1097/00007890-199312000-00031>

## P

---

- Pales Espinosa, E., Koller, A. et Allam, B. (2016). Proteomic characterization of mucosal secretions in the eastern oyster, *Crassostrea virginica*. *Journal of Proteomics*, *132*, 63-76. <https://doi.org/10.1016/j.jprot.2015.11.018>
- Parra, G. I. (2019). Emergence of norovirus strains: A tale of two genes. *Virus Evolution*, *5*(2), vez048. <https://doi.org/10.1093/ve/vez048>
- Parra, G. I., Tohma, K., Ford-Siltz, L. A., Eguino, P., Kendra, J. A., Pilewski, K. A. et Gao, Y. (2023). Minimal Antigenic Evolution after a Decade of Norovirus GII.4 Sydney\_2012 Circulation in Humans. *Journal of Virology*, *97*(2), e01716-22. <https://doi.org/10.1128/jvi.01716-22>
- Parrino, T. A., Schreiber, D. S., Trier, J. S., Kapikian, A. Z. et Blacklow, N. R. (1977). Clinical Immunity in Acute Gastroenteritis Caused by Norwalk Agent. *New England Journal of Medicine*, *297*(2), 86-89. <https://doi.org/10.1056/NEJM197707142970204>
- Paschinger, K., Razzazi-Fazeli, E., Furukawa, K. et Wilson, I. B. H. (2011). Presence of galactosylated core fucose on N-glycans in the planaria *Dugesia japonica*. *Journal of Mass Spectrometry*, *46*(6), 561-567. <https://doi.org/10.1002/jms.1925>
- Paschinger, K., Staudacher, E., Stemmer, U., Fabini, G. et Wilson, I. B. H. (2005). Fucosyltransferase substrate specificity and the order of fucosylation in invertebrates. *Glycobiology*, *15*(5), 463-474. <https://doi.org/10.1093/glycob/cwi028>
- Paschinger, K. et Wilson, I. B. H. (2015). Two types of galactosylated fucose motifs are present on N-glycans of *Haemonchus contortus*. *Glycobiology*, *25*(6), 585-590. <https://doi.org/10.1093/glycob/cwv015>
- Paschinger, K. et Wilson, I. B. H. (2019). Comparisons of N-glycans across invertebrate phyla. *Parasitology*, *146*(14), 1733-1742. <https://doi.org/10.1017/S0031182019000398>

## R

---

- Ravidà, A., Aldridge, A. M., Driessen, N. N., Heus, F. A. H., Hokke, C. H. et O'Neill, S. M. (2016). *Fasciola hepatica* Surface Coat Glycoproteins Contain Mannosylated and Phosphorylated N-glycans and Exhibit Immune Modulatory Properties Independent of the Mannose Receptor. *PLOS Neglected Tropical Diseases*, *10*(4), e0004601. <https://doi.org/10.1371/journal.pntd.0004601>

- Reitman, M. L., Varki, A. et Kornfeld, S. (1981). Fibroblasts from patients with I-cell disease and pseudo-Hurler polydystrophy are deficient in uridine 5'-diphosphate-N-acetylglucosamine: glycoprotein N-acetylglucosaminylphosphotransferase activity. *The Journal of Clinical Investigation*, 67(5), 1574-1579. <https://doi.org/10.1172/jci110189>
- Rendić, D., Kludiny, J., Stemmer, U., Schmidt, J., Paschinger, K. et Wilson, I. B. H. (2007). Towards abolition of immunogenic structures in insect cells: characterization of a honey-bee (*Apis mellifera*) multi-gene family reveals both an allergy-related core  $\alpha$ 1,3-fucosyltransferase and the first insect Lewis-histo-blood-group-related antigen-synthesizing enzyme. *Biochemical Journal*, 402(1), 105-115. <https://doi.org/10.1042/BJ20060964>
- Robilotti, E., Deresinski, S. et Pinsky, B. A. (2015). Norovirus. *Clinical Microbiology Reviews*, 28(1), 134-164. <https://doi.org/10.1128/CMR.00075-14>
- Rockx, B. H. G., Vennema, H., Hoebe, C. J. P. A., Duizer, E. et Koopmans, M. P. G. (2005). Association of histo-blood group antigens and susceptibility to norovirus infections. *The Journal of Infectious Diseases*, 191(5), 749-754. <https://doi.org/10.1086/427779>
- Rossez, Y., Maes, E., Lefebvre Darroman, T., Gosset, P., Ecobichon, C., Joncquel Chevalier Curt, M., Boneca, I. G., Michalski, J.-C. et Robbe-Masselot, C. (2012). Almost all human gastric mucin O-glycans harbor blood group A, B or H antigens and are potential binding sites for *Helicobacter pylori*. *Glycobiology*, 22(9), 1193-1206. <https://doi.org/10.1093/glycob/cws072>
- Ruiz-Matute, A. I., Hernández-Hernández, O., Rodríguez-Sánchez, S., Sanz, M. L. et Martínez-Castro, I. (2011). Derivatization of carbohydrates for GC and GC-MS analyses. *Journal of Chromatography B*, 879(17), 1226-1240. <https://doi.org/10.1016/j.jchromb.2010.11.013>
- Ruvoën-Clouet, N., Belliot, G. et Le Pendu, J. (2013). Noroviruses and histo-blood groups: the impact of common host genetic polymorphisms on virus transmission and evolution: Noroviruses and herd innate protection. *Reviews in Medical Virology*, 23(6), 355-366. <https://doi.org/10.1002/rmv.1757>
- Ruvoën-Clouet, N., Magalhaes, A., Marcos-Silva, L., Breiman, A., Figueiredo, C., David, L. et Le Pendu, J. (2014). Increase in genogroup II.4 norovirus host spectrum by CagA-positive *Helicobacter pylori* infection. *The Journal of Infectious Diseases*, 210(2), 183-191. <https://doi.org/10.1093/infdis/jiu054>
- Ruvoën-Clouet, N., Mas, E., Marionneau, S., Guillon, P., Lombardo, D. et Le Pendu, J. (2006). Bile-salt-stimulated lipase and mucins from milk of 'secretor' mothers inhibit the binding of Norwalk virus capsids to their carbohydrate ligands. *Biochemical Journal*, 393(3), 627-634. <https://doi.org/10.1042/BJ20050898>

## S

- Santé Publique France. (2023). *Gastro-entérites aiguës : bilan des saisons hivernales 2020-2021 et 2021-2022*. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-infectieuses-d-origine-alimentaire/gastro-enterites-aigues/documents/bulletin-national/gastro-enterites-aigues-bilan-des-saisons-hivernales-2020-2021-et-2021-2022>
- Sarnesto, A., Köhlin, T., Thurin, J. et Blaszczyk-Thurin, M. (1990). Purification of H gene-encoded beta-galactoside alpha 1—2 fucosyltransferase from human serum. *Journal of Biological Chemistry*, 265(25), 15067-15075. [https://doi.org/10.1016/S0021-9258\(18\)77224-4](https://doi.org/10.1016/S0021-9258(18)77224-4)
- Schachter, H. (2000). The joys of HexNAc. The synthesis and function of N- and O-glycan branches. *Glycoconjugate Journal*, 17(7-9), 465-483. <https://doi.org/10.1023/a:1011010206774>
- Schenk, B., Fernandez, F. et Waechter, C. J. (2001). The ins(ide) and outs(ide) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum. *Glycobiology*, 11(5), 61R-70R. <https://doi.org/10.1093/glycob/11.5.61R>

- Schroten, H., Hanisch, F.-G. et Hansman, G. S. (2016). Human Norovirus Interactions with Histo-Blood Group Antigens and Human Milk Oligosaccharides. *Journal of Virology*, *90*(13), 5855-5859. <https://doi.org/10.1128/jvi.00317-16>
- Schwab, K. J., Neill, F. H., Estes, M. K., Metcalf, T. G. et Atmar, R. L. (1998). Distribution of Norwalk virus within shellfish following bioaccumulation and subsequent depuration by detection using RT-PCR. *Journal of Food Protection*, *61*(12), 1674-1680. <https://doi.org/10.4315/0362-028x-61.12.1674>
- Shanker, S., Choi, J.-M., Sankaran, B., Atmar, R. L., Estes, M. K. et Prasad, B. V. V. (2011). Structural Analysis of Histo-Blood Group Antigen Binding Specificity in a Norovirus GII.4 Epidemic Variant: Implications for Epochal Evolution. *Journal of Virology*, *85*(17), 8635-8645. <https://doi.org/10.1128/JVI.00848-11>
- Shanker, S., Hu, L., Ramani, S., Atmar, R. L., Estes, M. K. et Venkataram Prasad, B. (2017). Structural features of glycan recognition among viral pathogens. *Current Opinion in Structural Biology*, *44*, 211-218. <https://doi.org/10.1016/j.sbi.2017.05.007>
- Smit, C. H., Diepen, A. van, Nguyen, D. L., Wuhrer, M., Hoffmann, K. F., Deelder, A. M. et Hokke, C. H. (2015). Glycomic Analysis of Life Stages of the Human Parasite *Schistosoma mansoni* Reveals Developmental Expression Profiles of Functional and Antigenic Glycan Motifs \*[S]. *Molecular & Cellular Proteomics*, *14*(7), 1750-1769. <https://doi.org/10.1074/mcp.M115.048280>
- Someya, Y. (2022). Lewis b antigen is a common ligand for genogroup I norovirus strains. *FEBS Open Bio*, *12*(9), 1688-1695. <https://doi.org/10.1002/2211-5463.13455>
- Spiro, R. G. (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, *12*(4), 43R-56R. <https://doi.org/10.1093/glycob/12.4.43R>
- Spiro, R. G. et Bhoyroo, V. D. (1988). Occurrence of sulfate in the asparagine-linked complex carbohydrate units of thyroglobulin. Identification and localization of galactose 3-sulfate and N-acetylglucosamine 6-sulfate residues in the human and calf proteins. *Journal of Biological Chemistry*, *263*(28), 14351-14358. [https://doi.org/10.1016/S0021-9258\(18\)68227-4](https://doi.org/10.1016/S0021-9258(18)68227-4)
- Stanley, P. (2011). Golgi Glycosylation. *Cold Spring Harbor Perspectives in Biology*, *3*(4), a005199-a005199. <https://doi.org/10.1101/cshperspect.a005199>
- Stanley, P., Moremen, K. W., Lewis, N. E., Taniguchi, N. et Aebi, M. (2022a). N-Glycans. Dans A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar et P. H. Seeberger (dir.), *Essentials of Glycobiology* (4th éd.). Cold Spring Harbor Laboratory Press. <http://www.ncbi.nlm.nih.gov/books/NBK579964/>
- Stanley, P., Wuhrer, M., Lauc, G., Stowell, S. R. et Cummings, R. D. (2022b). Structures Common to Different Glycans. Dans A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar et P. H. Seeberger (dir.), *Essentials of Glycobiology* (4th éd.). Cold Spring Harbor Laboratory Press. <http://www.ncbi.nlm.nih.gov/books/NBK579929/>
- Staudacher, E. (2012). Methylation – an uncommon modification of glycans. *bchm*, *393*(8), 675-685. <https://doi.org/10.1515/hsz-2012-0132>
- Staudacher, E. (2015). Mucin-Type O-Glycosylation in Invertebrates. *Molecules*, *20*(6), 10622-10640. <https://doi.org/10.3390/molecules200610622>
- Stepan, H., Pabst, M., Altmann, F., Geyer, H., Geyer, R. et Staudacher, E. (2012). O-Glycosylation of snails. *Glycoconjugate Journal*, *29*(4), 189-198. <https://doi.org/10.1007/s10719-012-9391-4>
- Ströh, L. J. et Stehle, T. (2014). Glycan Engagement by Viruses: Receptor Switches and Specificity. *Annual Review of Virology*, *1*(1), 285-306. <https://doi.org/10.1146/annurev-virology-031413-085417>
- Su, L., Ma, L., Liu, H., Zhao, F., Su, Z. et Zhou, D. (2018). Presence and Distribution of Histo-Blood Group Antigens in Pacific Oysters and the Effects of Exposure to Noroviruses GI.3 and GII.4 on Their



Expression. *Journal of Food Protection*, 81(11), 1783-1790. <https://doi.org/10.4315/0362-028X.JFP-18-074>

Subramanian, S. P., Babu, P., Palakodeti, D. et Subramanian, R. (2018). Identification of multiple isomeric core chitobiose–modified high-mannose and paucimannose N-glycans in the planarian *Schmidtea mediterranea*. *Journal of Biological Chemistry*, 293(18), 6707-6720. <https://doi.org/10.1074/jbc.RA117.000782>

Suzuki, T., Cummings, R. D., Aebi, M. et Parodi, A. (2022). Glycans in Glycoprotein Quality Control. Dans A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar et P. H. Seeberger (dir.), *Essentials of Glycobiology* (4th éd.). Cold Spring Harbor Laboratory Press. <http://www.ncbi.nlm.nih.gov/books/NBK579973/>

## T

---

Takahashi, N., Masuda, K., Hiraki, K., Yoshihara, K., Huang, H.-H., Khoo, K.-H. et Kato, K. (2003). N-Glycan structures of squid rhodopsin. Existence of the alpha1-3 and alpha1-6 difucosylated innermost GlcNAc residue in a molluscan glycoprotein. *European Journal of Biochemistry*, 270(12), 2627-2632. <https://doi.org/10.1046/j.1432-1033.2003.03636.x>

Talmont, F. et Fournet, B. (1991). Isolation and characterization of methylated sugars from the tube of the hydrothermal vent tubicolous annelid worm *Alvinella pompejana*. *FEBS Letters*, 281(1-2), 55-58. [https://doi.org/10.1016/0014-5793\(91\)80357-9](https://doi.org/10.1016/0014-5793(91)80357-9)

Tan, M. et Jiang, X. (2005). Norovirus and its histo-blood group antigen receptors: an answer to a historical puzzle. *Trends in Microbiology*, 13(6), 285-293. <https://doi.org/10.1016/j.tim.2005.04.004>

Tarris, G., Estienney, M., Daval-Frérôt, P., Lariotte, A.-C., Aubignat, D., Sé, K., Michiels, C., Martin, L., de Rougemont, A. et Belliot, G. (2022). Intestinal Norovirus Binding Patterns in Nonsecretor Individuals. *Journal of Virology*, 96(19), e00865-22. <https://doi.org/10.1128/jvi.00865-22>

Terada, M., Khoo, K.-H., Inoue, R., Chen, C.-I., Yamada, K., Sakaguchi, H., Kadowaki, N., Ma, B. Y., Oka, S., Kawasaki, T. et Kawasaki, N. (2005). Characterization of Oligosaccharide Ligands Expressed on SW1116 Cells Recognized by Mannan-binding Protein: A HIGHLY FUCOSYLATED POLYLACTOSAMINE TYPE N-GLYCAN \*. *Journal of Biological Chemistry*, 280(12), 10897-10913. <https://doi.org/10.1074/jbc.M413092200>

Teunis, P. F. M., Moe, C. L., Liu, P., E. Miller, S., Lindesmith, L., Baric, R. S., Le Pendu, J. et Calderon, R. L. (2008). Norwalk virus: How infectious is it? *Journal of Medical Virology*, 80(8), 1468-1476. <https://doi.org/10.1002/jmv.21237>

Thebault, A., Teunis, P. F. M., Le Pendu, J., Le Guyader, F. S. et Denis, J.-B. (2013). Infectivity of GI and GII noroviruses established from oyster related outbreaks. *Epidemics*, 5(2), 98-110. <https://doi.org/10.1016/j.epidem.2012.12.004>

Tian, P., Bates, A. H., Jensen, H. M. et Mandrell, R. E. (2006). Norovirus binds to blood group A-like antigens in oyster gastrointestinal cells. *Letters in Applied Microbiology*, 43(6), 645-651. <https://doi.org/10.1111/j.1472-765X.2006.02010.x>

Tian, P., Engelbrektson, A. L., Jiang, X., Zhong, W. et Mandrell, R. E. (2007). Norovirus Recognizes Histo-Blood Group Antigens on Gastrointestinal Cells of Clams, Mussels, and Oysters: A Possible Mechanism of Bioaccumulation. *Journal of Food Protection*, 70(9), 2140-2147. <https://doi.org/10.4315/0362-028X-70.9.2140>

Tiralongo, J. et Martinez-Duncker, I. (2013). Introduction to Sialic Acid Structure, Occurrence, Biosynthesis and Function. Dans *Sialobiology: Structure, Biosynthesis and Function. Sialic Acid Glycoconjugates in Health and Disease* (p. 3-32). BENTHAM SCIENCE PUBLISHERS. <https://doi.org/10.2174/9781608053865113010004>

- Titz, A., Butschi, A., Henrissat, B., Fan, Y.-Y., Hennet, T., Razzazi-Fazeli, E., Hengartner, M. O., Wilson, I. B. H., Künzler, M. et Aebi, M. (2009). Molecular Basis for Galactosylation of Core Fucose Residues in Invertebrates. *The Journal of Biological Chemistry*, *284*(52), 36223-36233. <https://doi.org/10.1074/jbc.M109.058354>
- Tohma, K., Lepore, C. J., Ford-Siltz, L. A. et Parra, G. I. (2017). Phylogenetic Analyses Suggest that Factors Other Than the Capsid Protein Play a Role in the Epidemic Potential of GII.2 Norovirus. *mSphere*, *2*(3), e00187-17. <https://doi.org/10.1128/mSphereDirect.00187-17>
- Torrado, J., Ruiz, B., Garay, J., Cosme, A., Arenas, J. I., Bravo, J. C., Fontham, E. et Correa, P. (1997). Lewis, secretor, and ABO phenotypes, and sulfomucin expression in gastric intestinal metaplasia. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *6*(4), 287-289.
- Tretter, V., Altmann, F. et März, L. (1991). Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached alpha 1----3 to the asparagine-linked N-acetylglucosamine residue. *European Journal of Biochemistry*, *199*(3), 647-652. <https://doi.org/10.1111/j.1432-1033.1991.tb16166.x>

## U

---

- Ueki, Y., Amarasiri, M., Kamio, S., Sakagami, A., Ito, H., Uprety, S., Umam, A. N., Miura, T., Nguyen, T. H. et Sano, D. (2021). Human norovirus disease burden of consuming *Crassostrea gigas* oysters: A case-study from Japan. *Food Control*, *121*, 107556. <https://doi.org/10.1016/j.foodcont.2020.107556>

## V

---

- Van Kuik, J. A., Sijbesma, R. P., Kamerling, J. P., Vliegthart, J. F. et Wood, E. J. (1986). Primary structure of a low-molecular-mass N-linked oligosaccharide from hemocyanin of *Lymnaea stagnalis*. 3-O-methyl-D-mannose as a constituent of the xylose-containing core structure in an animal glycoprotein. *European Journal of Biochemistry*, *160*(3), 621-625. <https://doi.org/10.1111/j.1432-1033.1986.tb10083.x>
- Van Kuik, J. A., Sijbesma, R. P., Kamerling, J. P., Vliegthart, J. F. et Wood, E. J. (1987). Primary structure determination of seven novel N-linked carbohydrate chains derived from hemocyanin of *Lymnaea stagnalis*. 3-O-methyl-D-galactose and N-acetyl-D-galactosamine as constituents of xylose-containing N-linked oligosaccharides in an animal glycoprotein. *European Journal of Biochemistry*, *169*(2), 399-411. <https://doi.org/10.1111/j.1432-1033.1987.tb13626.x>
- Van Kuik, J. A., van Halbeek, H., Kamerling, J. P. et Vliegthart, J. F. (1985). Primary structure of the low-molecular-weight carbohydrate chains of *Helix pomatia* alpha-hemocyanin. Xylose as a constituent of N-linked oligosaccharides in an animal glycoprotein. *Journal of Biological Chemistry*, *260*(26), 13984-13988. [https://doi.org/10.1016/S0021-9258\(17\)38672-6](https://doi.org/10.1016/S0021-9258(17)38672-6)
- van Kuik, J. A., Breg, J., Kolsteeg, C. E. M., Kamerling, J. P. et Vliegthart, J. F. G. (1987). Primary structure of the acidic carbohydrate chain of hemocyanin from *Panulirus interruptus*. *FEBS Letters*, *221*(1), 150-154. [https://doi.org/10.1016/0014-5793\(87\)80370-8](https://doi.org/10.1016/0014-5793(87)80370-8)
- Varki, A. (2017). Biological roles of glycans. *Glycobiology*, *27*(1), 3-49. <https://doi.org/10.1093/glycob/cww086>
- Varki, A., Cummings, R. D., Aebi, M., Packer, N. H., Seeberger, P. H., Esko, J. D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J. J., Schnaar, R. L., Freeze, H. H., Marth, J. D., Bertozzi, C. R., Etzler, M. E., Frank, M., Vliegthart, J. F., Lütteke, T., ... Kornfeld, S. (2015). Symbol Nomenclature for Graphical Representations of Glycans, First Edition. *Glycobiology*, *25*(12), 1323-1324. <https://doi.org/10.1093/glycob/cwv091>
- Varki, A., Cummings, R. D., Esko, J. D., Hart, G. W. et Marth, J. D. (1999). *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press.

Vitale, A. et Chrispeels, M. J. (1984). Transient N-acetylglucosamine in the biosynthesis of phytohemagglutinin: attachment in the Golgi apparatus and removal in protein bodies. *Journal of Cell Biology*, 99(1), 133-140. <https://doi.org/10.1083/jcb.99.1.133>

Vongpunswad, S., Venkataram Prasad, B. V. et Estes, M. K. (2013). Norwalk Virus Minor Capsid Protein VP2 Associates within the VP1 Shell Domain. *Journal of Virology*, 87(9), 4818-4825. <https://doi.org/10.1128/JVI.03508-12>

## W

---

Waheed, A., Pohlmann, R., Hasilik, A., von Figura, K., van Elsen, A. et Leroy, J. G. (1982). Deficiency of UDP-N-acetylglucosamine: Lysosomal enzyme N-acetylglucosamine-1-phosphotransferase in organs of I-cell patients. *Biochemical and Biophysical Research Communications*, 105(3), 1052-1058. [https://doi.org/10.1016/0006-291X\(82\)91076-2](https://doi.org/10.1016/0006-291X(82)91076-2)

Walker, H. G., Gee, M. et McCready, R. M. (1962). Complete Methylation of Reducing Carbohydrates <sup>2</sup>. *The Journal of Organic Chemistry*, 27(6), 2100-2102. <https://doi.org/10.1021/jo01053a047>

Watkins, W. M. (1980). Biochemistry and Genetics of the ABO, Lewis, and P Blood Group Systems. Dans H. Harris et K. Hirschhorn (dir.), *Advances in Human Genetics 10* (p. 1-136). Springer US. [https://doi.org/10.1007/978-1-4615-8288-5\\_1](https://doi.org/10.1007/978-1-4615-8288-5_1)

Weichert, S., Koromyslova, A., Singh, B. K., Hansman, S., Jennewein, S., Schrotten, H. et Hansman, G. S. (2016). Structural Basis for Norovirus Inhibition by Human Milk Oligosaccharides. *Journal of Virology*, 90(9), 4843-4848. <https://doi.org/10.1128/JVI.03223-15>

Wilson, I. B. H. et Paschinger, K. (2016). Sweet secrets of a therapeutic worm: mass-spectrometric N-glycomic analysis of *Trichuris suis*. *Analytical and Bioanalytical Chemistry*, 408(2), 461-471. <https://doi.org/10.1007/s00216-015-9154-8>

Wohlschlager, T., Butschi, A., Grassi, P., Sutov, G., Gauss, R., Hauck, D., Schmieder, S. S., Knobel, M., Titz, A., Dell, A., Haslam, S. M., Hengartner, M. O., Aebi, M. et Künzler, M. (2014). Methylated glycans as conserved targets of animal and fungal innate defense. *Proceedings of the National Academy of Sciences*, 111(27). <https://doi.org/10.1073/pnas.1401176111>

World Health Organization. (2022). *Norovirus*. <https://www.who.int/teams/immunization-vaccines-and-biologicals/diseases/norovirus>

Wuhrer, M., Koeleman, C. A. M., Deelder, A. M. et Hokke, C. H. (2006). Repeats of LacdiNAc and fucosylated LacdiNAc on N-glycans of the human parasite *Schistosoma mansoni*. *The FEBS Journal*, 273(2), 347-361. <https://doi.org/10.1111/j.1742-4658.2005.05068.x>

Wuhrer, M., Robijn, M. L. M., KOELEMAN, C. A. M., BALOG, C. I. A., GEYER, R., DEELDER, A. M. et HOKKE, C. H. (2004). A novel Gal(beta1-4)Gal(beta1-4)Fuc(alpha1-6)-core modification attached to the proximal N-acetylglucosamine of keyhole limpet haemocyanin (KLH) N-glycans. *Biochemical Journal*, 378(2), 625-632. <https://doi.org/10.1042/bj20031380>

## Y

---

Yamamoto, F., Marken, J., Tsuji, T., White, T., Clausen, H. et Hakomori, S. (1990). Cloning and characterization of DNA complementary to human UDP-GalNAc: Fuc alpha 1----2Gal alpha 1----3GalNAc transferase (histo-blood group A transferase) mRNA. *Journal of Biological Chemistry*, 265(2), 1146-1151. [https://doi.org/10.1016/S0021-9258\(19\)40170-1](https://doi.org/10.1016/S0021-9258(19)40170-1)

Yan, S., Brecker, L., Jin, C., Titz, A., Dragosits, M., Karlsson, N. G., Jantsch, V., Wilson, I. B. H. et Paschinger, K. (2015). Bisecting Galactose as a Feature of N-Glycans of Wild-type and Mutant *Caenorhabditis*

*elegans*\*. *Molecular & Cellular Proteomics*, 14(8), 2111-2125.  
<https://doi.org/10.1074/mcp.M115.049817>

- Yan, S., Sema, S., Reichardt, N.-C., Paschinger, K. et Wilson, I. B. H. (2013). Array-assisted Characterization of a Fucosyltransferase Required for the Biosynthesis of Complex Core Modifications of Nematode N-Glycans. *The Journal of Biological Chemistry*, 288(29), 21015-21028.  
<https://doi.org/10.1074/jbc.M113.479147>
- Yan, S., Vanbeselaere, J., Jin, C., Blaukopf, M., Wöls, F., Wilson, I. B. H. et Paschinger, K. (2018). Core Richness of N-Glycans of *Caenorhabditis elegans*: A Case Study on Chemical and Enzymatic Release. *Analytical Chemistry*, 90(1), 928-935. <https://doi.org/10.1021/acs.analchem.7b03898>
- Yanagishita, M. (1993). Function of proteoglycans in the extracellular matrix. *Pathology International*, 43(6), 283-293. <https://doi.org/10.1111/j.1440-1827.1993.tb02569.x>
- Yang, M., Zhao, F., Tong, L., Wang, S. et Zhou, D. (2021). Contamination, bioaccumulation mechanism, detection, and control of human norovirus in bivalve shellfish: A review. *Critical Reviews in Food Science and Nutrition*, 1-14. <https://doi.org/10.1080/10408398.2021.1937510>
- Yu, S.-Y., Snovida, S. et Khoo, K.-H. (2020). Permethylated and Microfractionation of Sulfated Glycans for MS Analysis. *BIO-PROTOCOL*, 10(10). <https://doi.org/10.21769/BioProtoc.3617>
- Yu, S.-Y., Wu, S.-W. et Khoo, K.-H. (2006). Distinctive characteristics of MALDI-Q/TOF and TOF/TOF tandem mass spectrometry for sequencing of permethylated complex type N-glycans. *Glycoconjugate Journal*, 23(5-6), 355-369. <https://doi.org/10.1007/s10719-006-8492-3>
- Yu, Y., Cai, H., Hu, L., Lei, R., Pan, Y., Yan, S. et Wang, Y. (2015). Molecular epidemiology of oyster-related human noroviruses and their global genetic diversity and temporal-geographical distribution from 1983 to 2014. *Applied and Environmental Microbiology*, 81(21), 7615-7624.  
<https://doi.org/10.1128/AEM.01729-15>

## Z

---

- Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P. W. H., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, J., ... Wang, J. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*, 490(7418), 49-54.  
<https://doi.org/10.1038/nature11413>
- Zhang, W., Cao, P., Chen, S., Spence, A. M., Zhu, S., Staudacher, E. et Schachter, H. (2003). *Synthesis of paucimannose N-glycans by Caenorhabditis elegans requires prior actions of UDP-N-acetyl-D-glucosamine:α-3-D-mannoside β1,2-N-acetylglucosaminyltransferase I, α3,6-mannosidase II and a specific membrane-bound β-N-acetylglucosaminidase*, 12.
- Zhang, X.-F., Huang, Q., Long, Y., Jiang, X., Zhang, T., Tan, M., Zhang, Q.-L., Huang, Z.-Y., Li, Y.-H., Ding, Y.-Q., Hu, G.-F., Tang, S. et Dai, Y.-C. (2015). An outbreak caused by GII.17 norovirus with a wide spectrum of HBGA-associated susceptibility. *Scientific Reports*, 5, 17687.  
<https://doi.org/10.1038/srep17687>

# ANNEXE

---

# Species-specific N-glycomes and methylation patterns of oysters *Crassostea gigas* and *Ostrea edulis* and their possible consequences for norovirus-HBGA interaction

Audrey Auger<sup>1</sup>, Shin-Yi Yu<sup>1</sup>, Shih-Yun Guu<sup>2</sup>, Agnès Quéméner<sup>3</sup>, Gabriel Euler-Nicolas<sup>4</sup>, Hiromune Ando<sup>5</sup>, Marion Desdouits<sup>4</sup>, Françoise S. Le Guyader<sup>4</sup>, Kay-Hooi Khoo<sup>2,6</sup>, Jacques Le Pendu<sup>7</sup>, Frederic Chirat<sup>1</sup> and Yann Guerardel<sup>1,5\*</sup>

<sup>1</sup>Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France

<sup>2</sup>Institute of Biological Chemistry, Academia Sinica

<sup>3</sup>Nantes Université, Inserm UMR 1307, CNRS UMR 6075, Université d'Angers, CRCI2NA, F-44000 Nantes – France

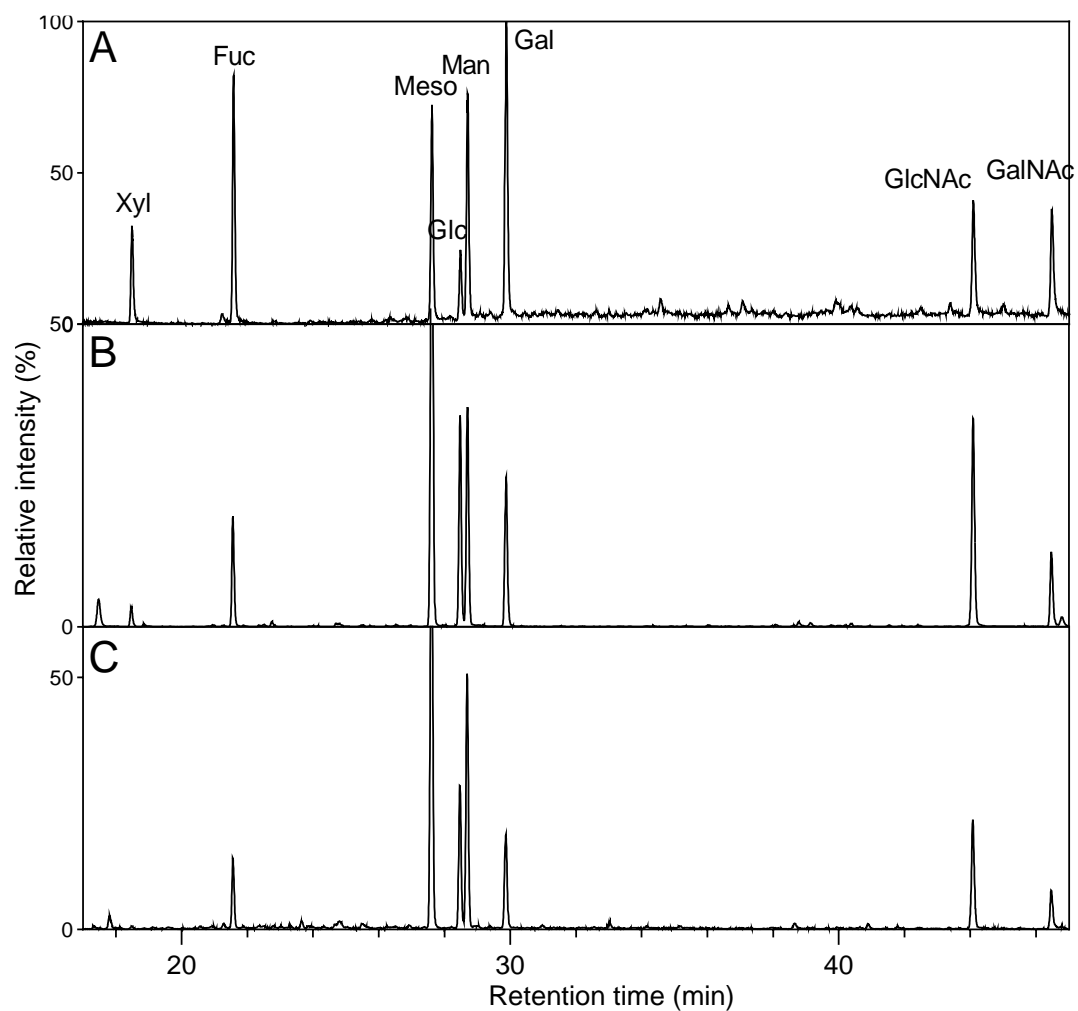
<sup>4</sup>Ifremer, U. Microbiologie Aliment Santé Environnement, LSEM, BP 21105, 44311 Nantes cedex 03 France

<sup>5</sup>Institute for Glyco-core Research (iGCORE), Gifu University, Gifu, Japan

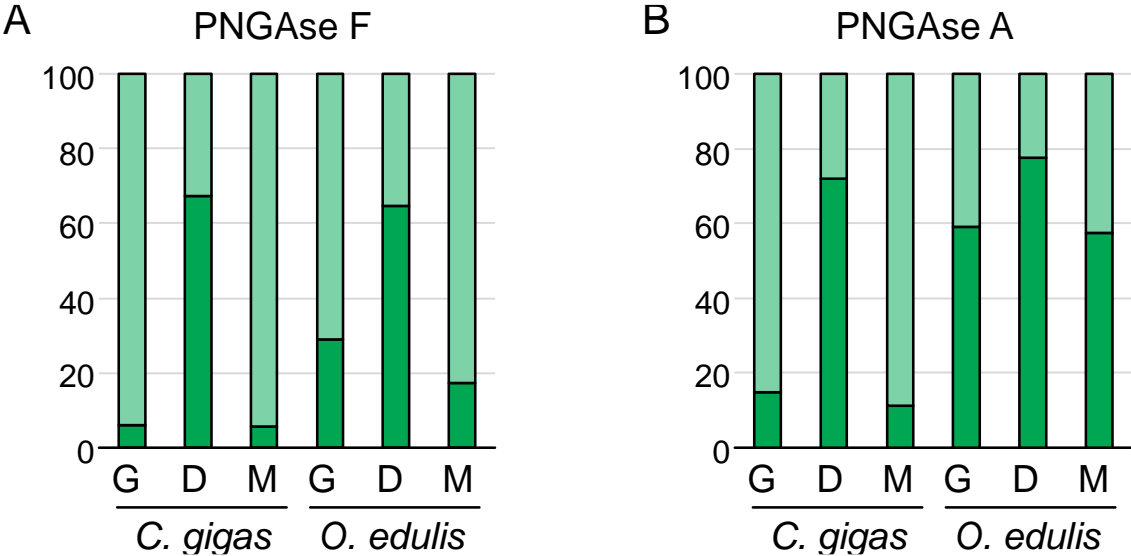
<sup>6</sup>Institute of Biochemical Sciences, National Taiwan University, Taiwan

<sup>7</sup>Nantes Université, Inserm, CNRS, Immunology and New Concepts in ImmunoTherapy, UMR 1302/EMR6001, Nantes, France

**Figure S1.** Monosaccharide composition analysis of *N*-glycans. Total Ion Count generated from GC-MS analysis of reduced and permethylated derivatives of (A) authentic standards, (B) *N*-glycans released from the mantle of *C. gigas* by PNGase F, (C) *N*-glycans released from the mantle of *O. edulis* by PNGase F.



**Figure S2.** Ratios (in %) of Me-3O-Man (light green) and Me-4O-Man (dark green) in N-glycans isolated from individual tissues of *C. gigas* and *O. edulis* (G, gills; D, digestive tissue; M, mantle) using (A) PNGase F and (B) PNGase A.





**Table S1.** Quantification of monosaccharides in *N*-glycans sequentially released by PNGase A and F of (A) *Crassostrea gigas* (Cg) and (B) *Ostrea edulis* (Oe) in the three analysed organs gills (G), digestive track (D) and mantle (M). Values are expressed in ng of monosaccharides per mg of proteins.

**A**

	PNGase F			PNGase A			Total PNGase A + F
	Gg-G	Cg-D	Cg-M	Cg-G	Cg-D	Cg-M	
MeFuc	27,8	28,7	20,0	0,5	6,5	3,6	87,0
Fuc	37,4	115,5	71,2	12,9	52,3	36,2	325,6
Xyl	3,1	15,1	13,1	3,4	2,9	7,8	45,5
MeMan	5,0	78,7	33,9	0,7	15,6	10,7	144,6
MeGal	1,0	1,3	0,6	0,0	0,6	0,3	3,8
Man	20,2	227,1	105,7	3,5	28,9	25,2	410,5
Gal	36,5	64,1	47,5	10,9	15,3	14,5	188,7
MeGalNAc	6,1	14,2	15,9	0,8	1,5	3,6	42,2
GlcNAc	18,8	141,8	71,7	2,9	26,1	21,6	282,9
GalNAc	5,4	6,8	4,2	0,6	1,0	0,9	19,0
Total	161,2	693,2	383,8	36,3	150,8	124,5	1549,8

**B**

	PNGase F			PNGase A			Total PNGase A + F
	Oe-G	Oe-D	Oe-M	Oe-G	Oe-D	Oe-M	
MeFuc	162,0	71,6	133,7	15,7	5,9	18,6	407,6
Fuc	313,2	221,6	256,3	62,1	36,2	53,2	942,6
Xyl	4,8	5,0	6,7	4,2	2,9	3,6	27,2
MeMan	194,3	113,8	196,6	25,8	10,2	18,3	558,9
MeGal	28,4	12,5	10,7	4,4	2,2	5,0	63,3
Man	510,9	365,1	494,6	46,1	21,2	38,0	1475,9
Gal	253,8	150,6	166,9	29,6	23,2	36,4	660,4
MeGalNAc	70,7	25,3	70,6	5,0	2,5	8,5	182,7
GlcNAc	346,6	212,8	366,9	44,5	20,8	46,6	1038,1
GalNAc	31,0	12,9	20,7	2,1	2,0	4,5	73,2
Total	1915,7	1191,1	1723,9	239,4	127,1	232,7	5429,9

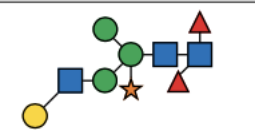
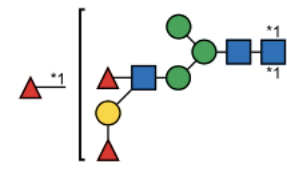
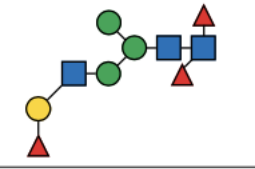
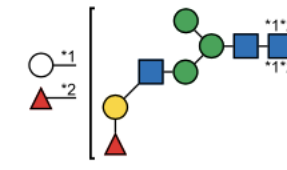
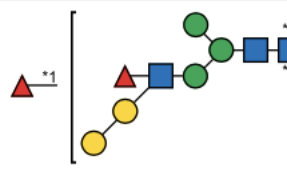
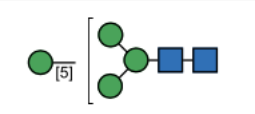
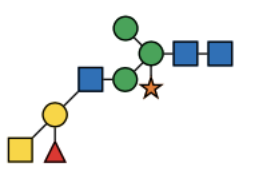
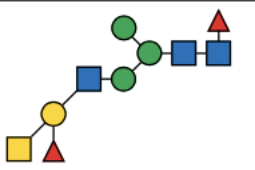
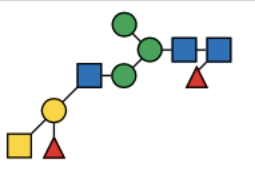
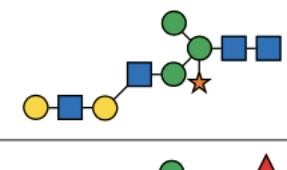
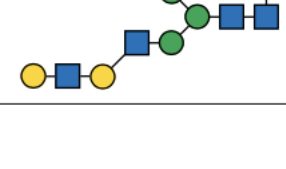
**Table S2.** List of *N*-glycans identified as permethylated derivatives following sequential released by PNGases F and PNGase A from three tissues gills (G), digestive tract (D) and mantle (M) of two oyster species, *Crassostrea gigas* (CG) and *Ostrea edulis* (OE). Structures were deduced from MS and MS<sup>n</sup> analyses of permethyl derivatives. *N*-glycans are ordered in increasing molecular weight. The presence of different *N*-glycans in each organ is indicated by a blue rectangle for *Crassostrea gigas* and green for *Ostrea edulis*, irrespective of the presence of natural methyl groups. Structures were represented according the standard Symbol Nomenclature for Glycan system (Varki *et al.*, 2015).

name	Theoretical m/z	Structure	<i>C.gigas</i>			<i>O.edulis</i>		
			G	D	M	G	D	M
CG1, OE1	1141,57							
CG2, OE2	1141,57							
CG3, OE3	1171,58							
CG4	1301,65							
CG5, OE4	1301,65							
CG6, OE5	1315,66							
CG7, OE6	1331,66							
OE7	1331,66							
CG8, OE8	1345,67							
CG9, OE9	1345,67							
CG10	1345,67							
OE10	1345,67							
CG11, OE11	1375,68							
OE12	1375,68							
OE13	1375,68							
CG12	1416,71							

CG13, OE14	1475,74			
CG14	1505,75			
CG15, OE15	1505,75			
CG16, OE16	1505,75			
CG17, OE17	1519,76			
CG18, OE18	1535,76			
OE19	1535,76			
OE20	1535,76			
CG19	1549,77			
CG20, OE21	1549,77			
CG21, OE22	1549,77			
CG22, OE23	1549,77			
CG23	1576,78			
CG24, OE24	1579,78			
CG25	1590,8			
CG26	1590,8			

CG27	1620,81			
CG28, OE25	1679,84			
CG29	1709,85			
CG30	1709,85			
CG31	1723,86			
CG32	1723,86			
CG33	1750,87			
CG34	1750,87			
CG35, OE26	1753,87			
CG36	1764,89			
CG37	1780,88			
CG38, OE27	1783,88			
CG39	1794,9			
CG40	1794,9			
OE28	1794,9			

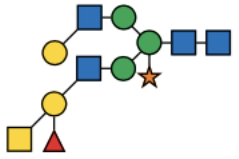
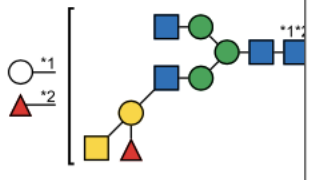
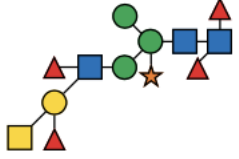
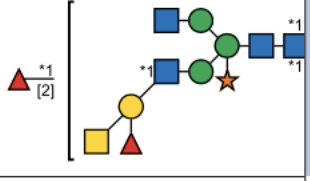
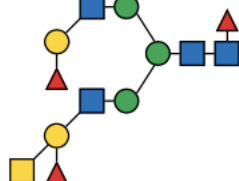
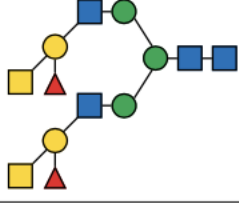
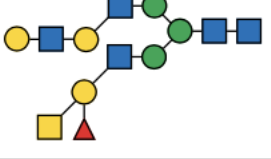
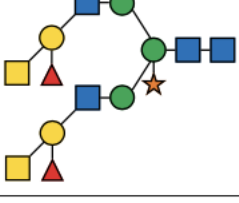
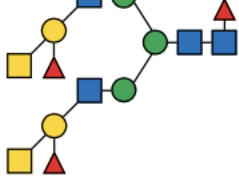
OE29	1794,9			
OE30	1794,9			
CG41	1824,91			
CG42	1913,95			
CG43	1924,96			
CG44	1954,97			
CG45	1954,97			
CG46	1968,99			
OE31	1968,99			
CG47, OE32	1987,98			
CG48, OE33	2040,02			
CG49	2070,03			
CG50	2118,05			

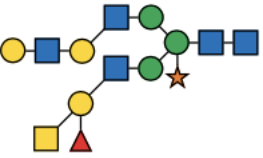
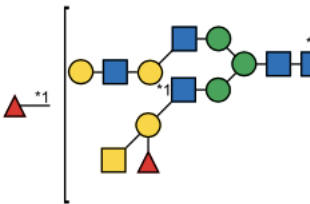
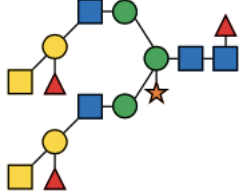
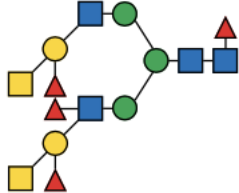
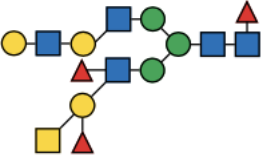
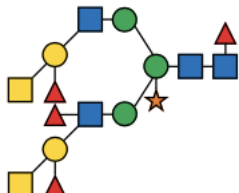
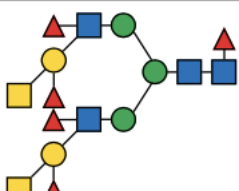
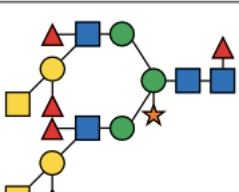
CG51	2129,06			
OE34	2143,08			
OE35	2143,08			
OE36	2173,09			
OE37	2173,09			
CG52, OE38	2192,08			
CG53	2200,1			
CG54, OE39	2214,11			
CG55	2214,11			
CG56	2230,11			
CG57, OE40	2244,12			

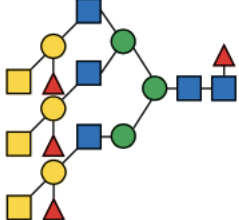
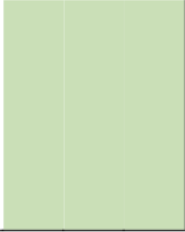
CG58	2244,12			
OE41	2244,12			
CG59	2285,15			
OE42	2347,18			
CG60	2374,19			
OE43	2377,19			
CG61, OE44	2388,2			
CG62, OE45	2388,2			
CG63, OE46	2396,18			
CG64	2404,2			



CG65, OE47	2418,21			
CG66	2418,21			
CG67	2445,22			
CG68	2459,24			
CG69	2489,25			
CG70	2548,28			
CG71, OE48	2548,28			
CG72, OE49	2562,29			
CG73	2578,29			
CG74, OE50	2600,3			

CG75	2649,32				
CG76, OE51	2663,34				
CG77	2722,37				
CG78	2793,4				
OE52	2837,43				
CG79, OE53	2908,47				
CG80	2938,48				
CG81	3068,54				
CG82, OE54	3082,55				

CG83	3098,55			
OE55	3112,57			
CG84	3242,63			
CG85, OE56	3256,64			
OE57	3286,66			
CG86	3416,72			
CG87, OE58	3430,73			
CG88	3590,81			

OE59	3951		
------	------	---	---

**Table S3.** List of methylated N-glycans identified from the mantle of *Crassostrea gigas* and *Ostrea edulis*. Following their sequential release from *C.gigas* and *O.edulis* by PNGase F and A, N-glycans were deuteromethylated and analysed by MALDI-QIT-TOF. Each -3 m.u. difference between the expected mass and the measured mass indicated the presence of a methyl group in the native N-glycan. The number of methyl groups per N-glycan is indicated by the presence of a cross in the corresponding column of the table (yellow for *c.gigas* and blue for *O.edulis*).

N-glycan	0	1	2	3	N-glycan	0	1	2	3	N-glycan	0	1	2	3	N-glycan	0	1	2	3	4
CG1		x			CG21	x	x	x		CG54			x		OE1		x			
CG2	x	x			CG22	x	x	x		CG55			x		OE2	x	x			
CG3	x	x	x		CG24	x	x	x	x	CG56			x		OE3		x	x		
CG4		x			CG25	x	x			CG60		x	x		OE5		x			
CG5		x			CG27	x	x	x		CG61				x	OE6			x		
CG6	x	x			CG28		x	x		CG62			x	x	OE8			x		
CG7	x	x	x		CG33	x	x			CG63	x	x	x		OE9		x	x		
CG8		x	x		CG35			x		CG69		x			OE10	x	x			
CG9	x	x	x		CG37	x	x			CG70			x	x	OE11		x	x		
CG10	x	x			CG38	x	x	x	x	CG71		x	x		OE12		x			
CG11	x	x	x		CG39	x	x			CG74	x				OE13		x			
CG12		x			CG44	x	x			CG77			x	x	OE14		x			
CG13		x			CG45	x	x			CG78			x	x	OE17			x		
CG14	x	x	x		CG46	x									OE21			x		
CG15		x	x		CG47	x	x	x	x						OE22		x	x		
CG16		x			CG48			x							OE23		x	x		
CG17	x	x	x		CG49			x							OE24	x		x	x	
CG18		x	x		CG51	x									OE26			x		
CG19			x		CG52	x		x	x						OE27			x	x	
CG20			x		CG53		x	x							OE29			x		
															OE31			x	x	
															OE33		x	x		
															OE34		x	x	x	
															OE35				x	
															OE36				x	
															OE41	x		x		
															OE42		x	x	x	
															OE52				x	x
															OE54	x	x	x		

# PERMISSION DE REPRODUCTION DES FIGURES

---

Toutes les figures utilisées dans ce manuscrit ont obtenu, si cela était nécessaire, les autorisations de reproduction.

Les figures 5, 6, 7, 12, 14, 17, 41 et les tableaux 1 et 2 ont été obtenus sur des publications en libre accès qui n'ont pas nécessité de demande de permission particulière. Les figures 20 et 42 ont été réalisées avec des icônes issues de Flaticon.com.

Figures avec obtention d'une licence :



Numéro de figure	Numéro de licence
15	1402637-2
16	5640731184662
18	5640740393054
19	5640740571088
22	1402637-1
23	1402632-1
31	5640750880509