



HAL
open science

Ingénierie métabolique de la voie de N-Glycosylation chez la diatomée *Phaeodactylum tricornutum*

Rodolphe Dumontier

► **To cite this version:**

Rodolphe Dumontier. Ingénierie métabolique de la voie de N-Glycosylation chez la diatomée *Phaeodactylum tricornutum*. Sciences agricoles. Normandie Université, 2020. Français. NNT : 2020NORMR021 . tel-03348479

HAL Id: tel-03348479

<https://theses.hal.science/tel-03348479v1>

Submitted on 19 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité ASPECTS MOLECULAIRES ET CELLULAIRES DE LA BIOLOGIE

Préparée au sein de l'Université de Rouen Normandie
Laboratoire Glyco-MEV EA 4358

Ingénierie métabolique de la voie de *N*-glycosylation chez la diatomée *Phaeodactylum tricornutum*

Présentée et soutenue par
Rodolphe DUMONTIER

Thèse soutenue publiquement le 18 Septembre 2020
devant le jury composé de

M. Jean-Paul CADORET	Directeur scientifique, ALGAMA, Paris	Rapporteur
M. Vincent FERRIÈRES	Professeur, Rennes institute of Chemical Sciences (UMR CNRS 6226, ISCR), Université de Rennes	Rapporteur
Mme Fayza DABOUSSI	Directeur de recherche, Laboratoire Toulouse Biotechnology Institute (TBI), INSA/INRA 792, INSA de Toulouse	Examineur
M. Bruno SAINT-JEAN	Chargé de recherche, Laboratoire Physiologie et Biotechnologie des Algues (PDG- RBE-BRM-PBA), Centre Atlantique, Nantes	Examineur
Mme Muriel BARDOR	Professeur, Laboratoire de Glycobiologie et Matrice Extracellulaire végétale EA 4358, Université de Rouen Normandie	Directeur de thèse
M. Patrice LEROUGE	Professeur, Laboratoire de Glycobiologie et Matrice Extracellulaire végétale EA 4358, Université de Rouen Normandie	Codirecteur de thèse
M. Alain MARECK	Maître de conférences, Laboratoire de Glycobiologie et Matrice Extracellulaire végétale EA 4358, Université de Rouen Normandie	Co-encadrant de thèse

Thèse dirigée par Pr. Muriel BARDOR, codirigée par Pr. Patrice LEROUGE et co-encadrée
par Dr. Alain Mareck. Laboratoire de Glycobiologie et matrice Extracellulaire végétale EA
4358, Université de Rouen Normandie



Normande de Biologie Intégrative,
Santé, Environnement



Remerciements

Je souhaite, tout d'abord, remercier M. Jean Paul CADORET responsable scientifique de l'entreprise ALGAMA, ainsi que M. Vincent FERRIÈRES, professeur à l'université de Rennes de m'avoir fait l'honneur d'être les rapporteurs de cette thèse.

Je souhaite également remercier Mme Fayza DABOUSSI, directrice de recherche de Toulouse Biotechnology Institute et M. Bruno SAINT-JEAN, chercheur en biologie à l'IFREMER de m'avoir fait l'honneur d'être examinatrice et examinateur de cette thèse et d'avoir été les membres de mon comité de suivi individuel.

Je remercie infiniment ma directrice de thèse Mme Muriel BARDOR, Professeure à l'Université de Rouen Normandie, et mon co-directeur de thèse M. Patrice LEROUGE, Professeur à l'Université de Rouen Normandie. Merci pour votre confiance et votre grande aide dans l'élaboration de ce projet de thèse. J'ai eu l'immense plaisir de pouvoir travailler à vos côtés durant ces nombreuses années depuis le Master 1 jusqu'à la fin de cette thèse. Merci de m'avoir associé à ce projet palpitant, qu'est le « projet microalgue ». Merci pour tous vos précieux conseils qui m'ont permis d'évoluer et qui me seront grandement utiles durant toute ma carrière. Je vous dois beaucoup...

Je remercie infiniment M. Alain MARECK, Maître de conférences à l'Université de Rouen Normandie, d'avoir co-encadré cette thèse mais également d'avoir encadré mes stages de Master 1 et Master 2. J'ai beaucoup appris à vos côtés, pas seulement pour nos heures de discussion sur ce projet de thèse, mais également nos heures de discussions sur la physique quantique, l'astronomie. Merci pour ces magnifiques objets que vous avez pu me montrer après vos inspections à la « poêle à frire ». Gardez précieusement ces magnifiques trésors. Merci pour tous vos conseils sur ce projet de thèse et votre grande aide sur l'écriture de ce manuscrit. Vos précieux conseils m'ont permis d'améliorer mes capacités de rédaction et je vous en serai éternellement reconnaissant. Merci pour votre bonne humeur constante. Restez tel que vous êtes...

Je remercie le Dr Philippe CHAN de la plate-forme PISSARO et Mme Corinne LOUTELIER-BOURHIS, Maître de conférences au laboratoire COBRA pour leurs précieuses aides sur les aspects de spectrométrie de masse lors des analyses des glycanes.

Je remercie chaleureusement Mme Fayza DABOUSSI de m'avoir accueilli à Toulouse dans son laboratoire afin de réaliser les transformations génétiques par biolistique indispensables à cette thèse.

Je remercie Marie-Christine KIEFER-MEYER, Ingénieure de recherche à l'Université de Rouen Normandie, pour son aide précieuse en biologie moléculaire et en bio-informatique durant l'élaboration de ce projet de thèse.

Je remercie Carole BUREL et Carole PLASSON d'avoir fourni une aide précieuse dans de nombreuses expérimentations en biologie moléculaire et en biochimie.

Je remercie tous les membres du laboratoire Glyco-MEV pour leur bonne humeur durant ces nombreuses années à leurs côtés. Je n'oublierai jamais les nombreuses discussions autour du football lors des pauses repas avec Arnaud LEHNER, Patrice LEROUGE, Pierre-Louis LUCAS, Azzedine DRIOUICH et Jean-Claude MOLLET. D'ailleurs, allez l'OL !

Merci à tous mes collègues doctorants et post-doctorants du laboratoire, les anciens comme les nouveaux. Sans vous, les journées au laboratoire n'auraient pas été les mêmes.

Un merci tout particulier à Pierre-Louis LUCAS, pour sa spéciale Pavard (eh oui ! tu y as le droit aussi =)) et ces soirées au stade Michel-D'Ornano et à Jeremy DEHORS pour ces nombreuses heures de parties de jeu d'échec réalisées après nos longues journées de manipulation. Merci à vous deux pour ces séances de sports intensives (et des fois non) et pour ces soirées passées ensemble. Ce fut un grand plaisir de passer ces bons moments avec vous deux.

Merci à Aline PLANCHON, Marc ROPITAUX, Mathilde CAMIADE, Marie CHAMBARD, Coralie CHUBERRE, Quentin ARNAUDIN, Maxime GOTTE et Jesmilien DILUBENZI pour ces soirées passées autour d'un verre au COLBERT ou à domicile. Merci pour votre bonne humeur, restez comme vous êtes. Une petite pensée à tous les poissons de Coralie ayant péri durant ces années de thèse.

Merci à Charlotte TOUSTOU, fan de BMTB et de Rammstein. Heureux que tu aies pu me supporter pendant cette dernière année de thèse dans le même bureau. Tu auras au moins pu découvrir le morceau « Petit poney » =).

Merci à Émilie, ma chérie, de m'avoir supporté pendant ces trois ans de thèse, dans les bons moments comme dans les mauvais moments. Tu m'as été d'une grande aide psychologique dans ce projet de thèse et je te remercie infiniment. A présent, à nous l'avenir. Je t'aime.

Enfin, merci à ma famille et à mes amis qui ont su me soutenir et qui m'ont permis de devenir la personne que je suis. Merci pour vos encouragements dans ces nombreuses années d'études et de vie.

Sommaire

Remerciements	4
Sommaire	8
Liste des abréviations	14
Liste des figures	18
Liste des tableaux	20
Présentation du candidat	22
Chapitre 1 Synthèse bibliographique	26
I. Les biomédicaments	27
A. Définition d'un biomédicament	27
B. Les biomédicaments : cas particuliers des anticorps	27
II. Les anticorps : importance des modifications co- et post-traductionnelles	30
A. Généralités	30
B. Structure d'un anticorps monoclonal	31
C. Rôle des <i>N</i>-glycannes portés par les anticorps	32
1. Rôle des <i>N</i> -glycannes sur la conformation des anticorps et la perte des fonctions effectrices.....	34
III. La voie de la <i>N</i>-glycosylation chez l'Homme	39
A. Généralités	39
B. La synthèse du précurseur dans le réticulum endoplasmique	40
1. Phase cytosolique	40
2. Phase luminale.....	41
C. Transfert du glycanne sur une protéine en cours de synthèse	42
D. Cycles de repliement et de contrôle qualité	43
E. Maturation des <i>N</i>-glycannes des protéines humaines dans l'appareil de Golgi	45
IV. L'ingénierie métabolique de la voie de la <i>N</i>-glycosylation	48
A. Chez les mammifères	49
1. Inactivation du gène codant pour l' $\alpha(1,3)$ -galactosyltransférase.....	49
2. Humanisation des processus de sialylation	50
B. Chez les levures	52
1. Inhibition de l'hypermannosylation des <i>N</i> -glycannes des glycoprotéines de levure.....	53
2. Humanisation des <i>N</i> -glycannes complexes des glycoprotéines chez la levure.....	54
3. Production d'anticorps monoclonaux chez les levures.....	56
C. Chez les plantes	57
1. Élimination des épitopes antigéniques des <i>N</i> -glycannes des glycoprotéines produites chez les plantes	58
2. Humanisation des <i>N</i> -glycannes complexes des glycoprotéines produites chez les plantes	61
3. Les anticorps monoclonaux à usage thérapeutique produits chez les plantes.....	64

V. Les microalgues, systèmes d'expression alternatifs pour la production d'anticorps monoclonaux.....	67
A. La microalgue verte <i>Chlamydomonas reinhardtii</i>	68
B. La diatomée <i>Phaeodactylum tricornutum</i>	69
VI. La N-glycosylation des microalgues et les techniques d'ingénierie	71
A. La N-glycosylation chez la microalgue verte <i>C. reinhardtii</i>	89
B. La N-glycosylation chez <i>P. tricornutum</i>	92
Objectifs de la thèse	95
A. Analyse du précurseur oligosaccharidique chez <i>P. tricornutum</i>	96
B. Analyse structurale détaillée des N-glycannes oligomannosidiques chez <i>P. tricornutum</i>	97
C. Tentative d'humanisation de la voie de la N-glycosylation chez la diatomée <i>P. tricornutum</i>	98
Chapitre 2 Matériel et méthode	100
I. Matériels	101
A. Matériel biologique	101
1. <i>Phaeodactylum tricornutum</i>	101
2. <i>Chlamydomonas reinhardtii</i>	101
3. Souche bactérienne	101
B. Vecteur plasmidique.....	101
C. Amorces utilisées pour le criblage des clones transformés	102
II. Méthodes de culture.....	103
A. <i>Phaeodactylum tricornutum</i>	103
1. Sur des boîtes gélosées.....	103
2. En milieu liquide.....	104
B. <i>Chlamydomonas reinhardtii</i>	105
C. Cultures cellulaires de bactéries.....	105
III. Méthodes de biologie moléculaire	107
A. Préparation plasmidique.....	107
1. Transformation dans des cellules d' <i>E. coli</i>	107
2. Amplification et purification plasmidique.....	107
3. Dosage de l'ADN plasmidique.....	108
4. Digestion enzymatique des plasmides.....	108
B. Transformation génétique de <i>Phaeodactylum tricornutum</i> par biolistique.....	109
1. Culture de la diatomée et étalement sur boîtes gélosées	109
2. Préparation de la biolistique et transformations des cellules de <i>P. tricornutum</i>	110
3. Étape de sélection des clones après la transformation	111
C. Criblage génique des clones isolés	112
1. Extraction de l'ADN des clones isolés	112
2. Amplification génique par PCR (Polymerase Chain Reaction)	112
3. Électrophorèse des ADN	112
D. Criblage des transcrits des clones isolés	113
1. Extraction de l'ARN des clones isolés.....	113
2. Éluion des ARN de la colonne	114
3. Élimination d'ADN résiduels par une réaction de DNase	114
4. Synthèse d'ADNc à partir d'ARN et amplification d'ADNc par PCR.....	114

IV. Méthodes de biochimie 115

A.	Extraction des oligosaccharides-PP-Dolichol (LLO).....	115
1.	Extraction microsomale	115
2.	Extraction des LLO.....	115
3.	Hydrolyse acide.....	116
4.	Purification des oligosaccharides sur colonne carbograph.....	116
5.	Perméthylation des oligosaccharides	117
6.	Purification des oligosaccharides perméthylés sur la colonne C18	117
B.	Criblage protéique des clones isolés.....	118
1.	Préparation rapide de la fraction microsomale	118
2.	Extraction des protéines totales	118
3.	Migration sur gel d'électrophorèse	119
4.	Coloration des protéines au bleu de Coomassie.....	120
5.	Coloration des protéines au nitrate d'argent	120
6.	Transfert électrophorétique sur une membrane de nitrocellulose et révélation en western blot	121
7.	Vérification de la qualité du transfert par coloration au rouge Ponceau	122
8.	Incubation de la membrane avec l'anticorps primaire	122
9.	Incubation de la membrane avec l'anticorps secondaire	122
10.	Immunodétection révélée par chimioluminescence avec le réactif PICO Plus	123
11.	Révélation par Imager Fusion-FX	123
C.	Eastern Blot et galactosylation <i>in vitro</i>.....	123
1.	Galactosylation <i>in vitro</i> des glycoprotéines de <i>P. tricornutum</i>	123
2.	Migration sur gel électrophorétique et transfert sur une membrane de nitrocellulose	124
3.	Incubation avec la lectine RCA120 marquée à la FITC	124
4.	Révélation de la fluorescence sur Typhoon	124
D.	Extraction glycanique du Man₅GlcNAc₂	124
1.	Dénaturation des protéines et digestion à la PNGase F	124
2.	Précipitation à l'éthanol des protéines déglycosylées	125
3.	Purification des oligosaccharides sur une colonne carbograph.....	125
4.	Marquage au 2-AB	125
5.	Dessalage des glycanes marqués au 2-AB.....	126
6.	Perméthylation et purification sur colonne C18.....	126
E.	Extraction glycanique par la PNGase A.....	126
1.	Dénaturation des protéines et digestion protéique par la trypsine et la chymotrypsine.....	126
2.	Déglycosylation des glycoprotéines par la PNGase A	127
3.	Purification des <i>N</i> -glycanes par les colonnes C18 et carbograph	127
4.	Perméthylation et purification des <i>N</i> -glycanes perméthylés par la colonne C18.....	127
F.	Purification et analyse biochimique de la GnT I.....	128
1.	Extraction microsomale de la GnT I	128
2.	Immunoprécipitation de la GnT I	128
3.	Vérification de l'immunoprécipitation par Western Blot	129
4.	Digestion trypsique et chymotrypsique de la GnT I	129

V. Méthodes de spectrométrie de masse 130

A.	Analyse par MALDI-TOF	130
1.	Analyse du précurseur oligosaccharidique et du <i>N</i> -glycane Man ₅ GlcNAc ₂	130
2.	Analyse des <i>N</i> -glycanes issus de la fraction protéique totale	130
B.	Analyse par ESI-MSⁿ.....	131
C.	Analyse par ESI-IMS	131
D.	Analyse des glycanes perméthylés par nanoLC-ESI-MS.....	132
E.	Analyse de la GnT I par QTOF XT	133
F.	Analyse de la GnT I par Qexactive Plus.....	133

Chapitre 3 Résultats 136

A.	Analyse des précurseurs oligosaccharidiques liés au dolichol par spectrométrie de masse chez la diatomée <i>P. tricornutum</i>	137
	Figure supplémentaire du chapitre I	154
B.	Analyse structurale des <i>N</i> -glycannes oligomannosidiques synthétisés chez la diatomée <i>P. tricornutum</i>	160
	Figure supplémentaire du chapitre II	183
C.	Tentative d'humanisation de la voie de la <i>N</i> -glycosylation chez la diatomée <i>P. tricornutum</i>	188
1.	Surexpression de la GnT I de <i>P. tricornutum</i> (PtGnT I).....	191
2.	Séquençage protéique de la PtGnT I- V5 par chromatographie liquide couplée à de la spectrométrie de masse 193	
3.	Analyses des transformants sélectionnés	194
4.	Analyse de la <i>N</i> -glycosylation des protéines des clones sélectionnés	196
5.	Conclusion.....	203
Chapitre 4 Conclusion-discussion-perspectives.....		206
Conclusion		207
A.	Le précurseur oligosaccharidique est tronqué d'un résidu glucose chez <i>P. tricornutum</i>	207
B.	Analyse structurale détaillée des <i>N</i> -glycannes oligomannosidiques chez <i>P. tricornutum</i>	208
C.	Tentative d'humanisation de la voie de la <i>N</i> -glycosylation des protéines chez la diatomée <i>P. tricornutum</i>	209
Discussion perspective		210
A.	Le cycle de contrôle qualité de la diatomée	210
B.	Hypothèses expliquant pourquoi l'activité GnT I ne conduit pas à l'accumulation de <i>N</i> -glycannes complexes chez la diatomée ?	212
C.	Ingénierie corrective des étapes golgiennes de maturation des <i>N</i> -glycannes de <i>P. tricornutum</i> en vue de leur humanisation	216
1.	Les fucosyltransférases	217
2.	L' α -mannosidase II et la <i>N</i> -acétylglucosaminyltransférase II.....	218
3.	Les hexosaminidases.....	221
Chapitre 5 Bibliographie.....		228
Résumé / Abstract		253

Liste des abréviations

% : Pourcent
 °C : Degrés Celsius
 $\alpha(1,3)$ FuT : $\alpha(1,3)$ -fucosyltransférase
 $\alpha(1,4)$ FuT : $\alpha(1,4)$ -fucosyltransférase
 2-AB : 2-aminobenzamide
 A : Ampère
 aa : Acide aminé
 ACTH : Adenocorticotropin hormone
 ADCC : Antibody-dependent cell-mediated cytotoxicity
 ADN : Acide désoxyribonucléique
 ADNc : Acide désoxyribonucléique complémentaire
 ALG : Asparagine Linked Glycosylation
 ARN : Acide ribonucléique
 ASGPR : Antibodies specific for asialoglycoprotein receptor
 Asn : Asparagine
 CaCl_2 : Chlorure de calcium
 CDC : Complement-Dependent Cytotoxicity
 CDG : Congénital disorder of glycosylation
 CE : Energie de collision
 CH_3COOH : Acide acétique
 CHO : Chinese hamster ovaries
 CMAH : CMP-Neu5Ac Hydroxylase
 CMAS : CMP-acide sialique synthétase
 CMP : Cytidine monophosphate
 Cys : Cystéine
 D-ManNAc : D-N-acétylmannosamine
 DHB : Acide 2,5 - dehydroxybenzoïque
 DMSO : Diméthylsulfoxyde
 Dol-P : Dolichol phosphate
E. Coli : *Escherichia coli*
 EDEM1 : ER-degradation enhancing mannosidase-like protein 1
 EDM : Eau de mer
 EDTA : Ethylènediaminetétraacétique
 EMA : European Medicines Agency
 EPO : Erythropoïétine
 ERAD : Endoplasmic Reticulum Associated Degradation
 ERQC : ER-derived Quality Control Compartment
 ESI-MSⁿ : Electrospray ionization mass spectrometry
 eV : Electronvolt
 Fab : Fragment antigen-binding
 Fc : Fragment cristallisable
 FDA : Food and Drug Administration
 Fuc : Fucose

Fuc : Fucose
FuT : Fucosyltransférase
g : Gramme
Gal : Galactose
GalT : $\beta(1,4)$ -galactosyltransférase
GDP : Guanosine diphosphate
Glc : Glucose
GlcNAc : *N*-acétylglucosamine
GNE : UDP-GlcNAc 2-épimérase/*N*-acétylmannosamine kinase
GnT I : *N*-acétylglucosaminyltransférase I
GnT II : *N*-acétylglucosaminyltransférase II
GnT III : *N*-acétylglucosaminyltransférase III
GnT IV : *N*-acétylglucosaminyltransférase IV
GnT V : *N*-acétylglucosaminyltransférase V
GSI : α -glucosidases I
GSII : α -glucosidases II
h : Heure
H₂O : Eau
HCl : Acide chlorhydrique
HEK : Human Embryonic Kidney 293
ICC : Charge ionique
IgA : Immunoglobine A
IgG : Immunoglobuline G
IgM : Immunoglobine M
KCl : Chlorure de potassium
kDa : Kilodalton
Kdo : Acide 3-désoxy-D-manno-oct-2-ulosonique
kV : Kilovolt
L : Litre
LB : Lysogeny Broth
LMCO : Low mass cut-off
m : Mètre
M : Molaire
m/z : Rapport masse sur charge
MAC : Membrane Attack Complex
Man : Mannose
Man II : α -mannosidase II
mbar : Millibar
MgCl₂ : Chlorure de magnésium
min : Minute
mL : Millilitre
mM : Millimolaire
mmHg : Millimètre de mercure
Mn²⁺ : Ion manganèse
mU : Microunité
N₂ : Diazote
NaBH₃CN : Cyanobohydrure de sodium

NaCl : Chlorure de sodium
NANS : Acide *N*-acétylneuraminique phosphate synthase
NaOH : Soude
Neu5Ac : acide *N*-acétylneuraminique
Neu5Gc : acide *N*-glycolylneuraminique
ng : Nanogramme
NK : Natural killer
nL : Nanolitre
nm : nanomètre
NR : Nitrate réductase
OCH1P : α (1,6)-mannosyltransférase
OST : Oligosaccharyltransférase
P : Phosphate
p : Poids
pH : Potentiel hydrogène
PNGase A : Peptide: *N*-glycosidase A
PNGase F : Peptide: *N*-glycosidase F
PVPP : Polyvinylpolypyrrolidone
QIT : Piège à ions quadripolaire
RE : Réticulum endoplasmique
RNAi : RNA interférence
rpm : Tour par minute
s : Seconde
SDS : Dodécylsulfate de sodium
Ser : Sérine
SiaT : sialyltransférases
 β (1,2)XylT : β (1,2)-xylosyltransférase
 β (1,3)GalT : β (1,3)-galactosyltransférase
 β (1,4)GalT : β (1,4)-galactosyltransférase
TAE : Tris acétate EDTA
TAP : Tris acétate phosphate
TBS : Tris buffer saline
TBST : TBS tween
TFA : Acide trifluoroacétique
Thr : Thréonine
TMD : Transmembrane domain
UDP : Uridine diphosphate
UGGT : UDP-glucose: glycoprotéine glucosyltransférase
UV : Ultra violet
V : Volt
v : Volume
Xyl : Xylose
XylT : Xylosyltransférase
 Zn^{2+} : Ion zinc
 μ E : Microeinstein
 μ g : Microgramme
 μ L : Microlitre

μm : Micromètre

Liste des figures

Figure 1 : Nombre de biomédicaments approuvés par l'agence réglementaire américaine FDA et l'agence de l'Union Européenne EMA (Walsh, 2018)	28
Figure 2 : Représentation statistique des modifications post-traductionnelles putatives des protéines les plus abondantes chez les eucaryotes	30
Figure 3 : Structure protéique d'un anticorps monoclonal de type IgG	32
Figure 4 : Représentation des mécanismes ADCC (Antibody-Dependent Cell-mediated Cytotoxicity) (à gauche) et CDC (Complement-Dependent Cytotoxicity) (à droite) mis en jeu lors d'une thérapie par injection d'anticorps thérapeutiques (figure adaptée de Bernardo <i>et al.</i> , 2010)	34
Figure 5 : Structures majoritaires des N-glycannes portés par les anticorps monoclonaux produits dans les cellules de mammifères	35
Figure 6 : Voie de la biosynthèse du précurseur oligosaccharidique lié au dolichol puis de son transfert sur une protéine en cours de synthèse chez l'Homme (figure adaptée de Aebi, 2013)	41
Figure 7 : Cycle de contrôle qualité des glycoprotéines humaines impliquant les chaperonnes calnexine et calréticuline (figure adaptée de Ferris <i>et al.</i> , 2014)	44
Figure 8 : Processus de la maturation des N-glycannes en N-glycannes complexes au sein de l'appareil de Golgi chez l'Homme	46
Figure 9 : Représentation tridimensionnelle de la structure de la GnT I de lapin déterminée par cristallographie aux rayons X (Ünligil <i>et al.</i> , 2000)	47
Figure 10 : Structures glycaniques portant l'épitope $\alpha(1,3)$ -galactose résultant de l'action de l' $\alpha(1,3)$ GalT	50
Figure 11 : Voie de la biosynthèse du Neu5Gc à partir du Neu5Ac (Yin and Huang, 2012)	51
Figure 12 : Voie de la biosynthèse des N-glycannes hypermannosylés dans l'appareil de Golgi chez <i>Saccharomyces cerevisiae</i> (De Pourcq <i>et al.</i> , 2010)	53
Figure 13 : Stratégie utilisée pour l'humanisation des voies de la N-glycosylation chez les levures	56
Figure 14 : Voie de la biosynthèse des N-glycannes complexes chez les plantes (figure adaptée de Lerouge <i>et al.</i> , 1998)	58
Figure 15 : Représentation des stratégies d'ingénierie de la voie de la N-glycosylation chez des plantes déficientes en activités $\beta(1,2)$ XylIT et $\alpha(1,3)$ FuT	60
Figure 16 : Représentation de la voie de biosynthèse des acides sialiques des mammifères reconstruite chez les plantes d' <i>A. thaliana</i> par l'expression transitoire de gènes impliqués dans la biosynthèse, le transport puis le transfert du CMP-Neu5Ac (Castilho <i>et al.</i> , 2008)	64
Figure 17 : Schéma représentant l'arbre phylogénétique des espèces (Burki <i>et al.</i> , 2020)	67
Figure 18 : Photos en microscopie optique des trois morphotypes principaux de <i>P. tricornutum</i> (d'après Ovide <i>et al.</i> , 2018)	69
Figure 19 : Voie de la synthèse du précurseur glycanique effectuée dans le réticulum endoplasmique chez <i>C. reinhardtii</i> suggérée par les analyses bio-informatiques et biochimiques (Mathieu-Rivet <i>et al.</i> , 2013 ; Mathieu-Rivet <i>et al.</i> , 2014 ; Vanier <i>et al.</i> , 2017)	90
Figure 20 : Voies de la xylosylation et de la fucosylation proposées chez <i>C. reinhardtii</i> (figure construite d'après Lucas <i>et al.</i> , 2020 et Oltmanns <i>et al.</i> , 2020)	92
Figure 21 : Voie de la biosynthèse des N-glycannes complexes chez <i>P. tricornutum</i> (Baïet <i>et al.</i> , 2011 ; Mathieu-Rivet <i>et al.</i> , 2014 ; Zhang <i>et al.</i> , 2019)	95
Figure 22 : Isomère du Man ₅ GlcNAc ₂ dit « canonique » retrouvé chez la majorité des eucaryotes (Aebi, 2013) et linéaire comme retrouvé chez <i>C. reinhardtii</i>	98
Figure 23 : Carte du vecteur plasmidique Pha-NR utilisé pour les transformations génétiques de <i>Phaeodactylum tricornutum</i>	102
Figure 24 : Photos représentant le matériel nécessaire à une transformation génétique chez <i>P. tricornutum</i> par biolistique	111
Figure 25 : Masses moléculaires des protéines de calibration utilisées pour l'électrophorèse SDS-PAGE (PageRuler™ Prestained Protein Ladder, Fisher Scientific)	120
Figure 26 : Schéma représentant le montage nécessaire pour réaliser un transfert des protéines du gel d'électrophorèse sur une membrane de nitrocellulose	121
Figure 27 : Voie de la biosynthèse du précurseur glycanique des N-glycannes des protéines chez <i>P. tricornutum</i>	139
Figure 28 : Stratégie d'humanisation des voies de la N-glycosylation chez <i>P. tricornutum</i>	189

Figure 29 : Criblages génique et protéique des clones ayant inséré le gène codant la PtGnT I-V5	192
Figure 30 : Séquençage peptidique de la protéine PtGnT I-V5 immunoprécipitée à partir de la fraction microsomale du clone GX5	194
Figure 31 : Analyses par RT-PCR validant la dépendance de la transcription du gène codant la PtGnT I -V5 de l'induction des cellules par le nitrate.	195
Figure 32 : Analyse par Western blot montrant que l'expression de la protéine PtGnT I-V5 est dépendante de l'induction des cellules de <i>P. tricornutum</i> par le nitrate.	196
Figure 33 : Principe de l'affinodétection des épitopes glucidiques obtenus après la galactosylation enzymatique in vitro des <i>N</i> -glycannes. Cette affinodétection des épitopes glucidiques a été effectuée sur une empreinte protéique obtenue après un transfert sur une membrane de nitrocellulose.	197
Figure 34 : Immunodétection des épitopes core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose portés par les <i>N</i> -glycannes des protéines totales de <i>P. tricornutum</i> isolées des clones 65, 105 et 107.	199
Figure 35 : Tests de galactosylation in vitro réalisés sur les protéines totales extraites à partir du clone 105 de <i>P. tricornutum</i> surexprimant la PtGnT I-V5	201
Figure 36 : Analyse des <i>N</i> -glycannes de <i>P. tricornutum</i> obtenue par chromatographie liquide couplée à la spectrométrie de masse de la souche sauvage (WT) et des transformants surexprimant PtGnT I-V5 (clones 2 et GX5)	204
Figure 37 : Bilan des analyses structurales rapportées dans les chapitres I et II	209
Figure 38 : Liste des transporteurs putatifs de l'UDP-GlcNAc obtenue par la recherche dans le génome de <i>P. tricornutum</i> de séquences homologues du gène de l'isoforme 1 du transporteur golgien de l'UDP-GlcNAc humain (Uniprot : Q9Y2D2)	214
Figure 39 : Structure schématique de la protéine de fusion Man II/GnT II.	218
Figure 40 : Analyse comparée des séquences protéiques correspondant aux sites catalytiques de la Man II de drosophile et de <i>P. tricornutum</i>	220
Figure 41 : Analyse comparée des séquences protéiques correspondant aux sites catalytiques de la GnT II humaine et de <i>P. tricornutum</i>	221
Figure 42 : Criblage des transformants de <i>P. tricornutum</i> exprimant le gène humain codant la $\beta(1,4)$ GalT-HA	223
Figure 43 : Criblage protéique des doubles transformants ayant inséré simultanément les gènes codant la $\beta(1,4)$ GalT-HA et la PtGnT I-V5	224

Liste des tableaux

Tableau 1 : Exemples de projets utilisant les plantes pour la production d'anticorps monoclonaux à intérêt thérapeutique (A) et d'anticorps monoclonaux produits chez les plantes actuellement en phase d'essai clinique (B) (Yao et al., 2015)	66
Tableau 2 : Nom et séquence des amorces utilisées lors du criblage par amplification génique	103
Tableau 3 : Composition des milieux Conway et F/2 utilisés pour cultiver les différentes souches de <i>P. tricornutum</i>	104
Tableau 4 : Composition du milieu Tris Acétate Phosphate (TAP) (selon les conditions décrites dans la thèse de Pierre-Louis Lucas, 2019)	106
Tableau 5 : Composition des milieux utilisés lors des digestions enzymatiques effectuées lors du contrôle de la purification plasmidique	109
Tableau 6 : Composition du mix PCR utilisé pour l'amplification du gène d'intérêt	113
Tableau 7 : Composition d'un gel Bis-Tris 4-12%	119
Tableau 8 : Récapitulatif des anticorps primaires utilisés durant ce projet de thèse	122
Tableau 9 : Bilan du criblage des transformants PtGnTI-V5	193

Présentation du candidat

Présentation du candidat

Rodolphe DUMONTIER



PARCOURS UNIVERSITAIRE

2017 - 2020

UNIVERSITE DE ROUEN NORMANDIE
DOCTORAT EN BIOCHIMIE ET BIOLOGIE MOLECULAIRE

2016 - 2017

UNIVERSITE DE ROUEN NORMANDIE
MASTER 2 ECOPRODUCTION ET BIOVALORISATION
✓ Mention Très Bien (16/20)

2015 - 2016

UNIVERSITE DE ROUEN NORMANDIE
MASTER 1 ANALYSE, CONTRÔLE, EXPERTISE DANS L'AGROCHIMIE ET LES BIO-INDUSTRIES
✓ Mention Assez Bien (12/20)

2011 - 2015

UNIVERSITE DE ROUEN NORMANDIE
LICENCE EN BIOCHIMIE ET BIOLOGIE MOLECULAIRE



ACTIVITES D'ENSEIGNEMENT

2018 - 2020

MONITORAT DE BIOCHIMIE
✓ 64h d'enseignement réalisées
✓ Heures réalisées en TP de L2



PRODUCTIONS SCIENTIFIQUES

Publications / Communications affichées / Communications orales

3

4

2



PUBLICATIONS

2018

Rodolphe Dumontier, Alain Mareck, Narimane Mati-Baouche, Patrice Lerouge, Muriel Bardor.

Toward Future Engineering of the N-Glycosylation Pathways in Microalgae for Optimizing the Production of Biopharmaceuticals. IntechOpen, DOI: 10.5772/intechopen.73401

2018

Pierre-Louis Lucas*, **Rodolphe Dumontier***, Corinne Loutelier-Bourhis, Alain Mareck, Carlos Afonso, Patrice Lerouge, Narimane Mati-Baouche, Muriel Bardor. **User-friendly extraction and multistage tandem mass spectrometry based analysis of lipid-linked oligosaccharides in microalgae.** Plant Methods, BioMed Central, 2018, 14 (1), pp.e107. 10.1186/s13007-018-0374-8

* Equal contribution of these authors

2020

Rodolphe Dumontier*, Corinne Loutelier-Bourhis*, Marie-Laure Walet-Balieu,, Carole Burel, Alain Mareck, Carlos Afonso, Patrice Lerouge*, Muriel Bardor*. **Towards deciphering structural features of the isomers of oligomannoside N-glycans in the diatom, *Phaeodactylum tricornutum* using a combination of cutting edge mass spectrometry techniques.** Publication en préparation

* Equal contribution of these authors



COMMUNICATIONS PAR AFFICHE

Le nom de la personne ayant réalisé la communication est souligné

Mars 2018

Rodolphe Dumontier, Alain Mareck, Patrice Lerouge, Muriel Bardor. Metabolic engineering of the N-glycosylation pathway in the model diatom, *Phaeodactylum tricornutum*.

21ème journées de l'école doctorale NBISE, Mar 2018, Mont-Saint-Aignan, France.

Mai 2018

Rodolphe Dumontier, Pierre-Louis Lucas, Corinne Loutelier-Bourhis, Alain Mareck, Carlos Afonso, Patrice Lerouge, Narimane mati-Baouche, Muriel Bardor. A simple and non-radioactive method to characterize lipid-linked oligosaccharides in microalgae. Prix du 3ème poster

Journées de la Structure fédérative de recherche Normandie Végétal May 2018, Deauville, France

Mars 2019

Rodolphe Dumontier, Pierre-Louis Lucas, Corinne Loutelier-Bourhis, Alain Mareck, Carlos Afonso, Patrice Lerouge, Narimane mati-Baouche, Muriel Bardor. User friendly extraction and multistage tandem mass spectrometry based analysis of lipid linked oligosaccharides in microalgae.

22ème journées de l'Ecole Doctorale NBISE, Mar 2019, Caen, France.

Août 2019

Rodolphe Dumontier, Pierre-Louis Lucas, Corinne Loutelier-Bourhis, Alain Mareck, Carlos Afonso, Patrice Lerouge, Narimane mati-Baouche, Muriel Bardor. User-friendly extraction and multistage tandem mass spectrometry based analysis of lipid-linked oligosaccharides in microalgae.

International Symposium of Glycoconjugates, Aug 2019, Milan, Italy.



COMMUNICATIONS ORALES

Le nom de la personne ayant réalisé la communication est souligné

Juin 2018

Pierre-Louis Lucas, **Rodolphe Dumontier**, Corinne Loutelier-Bourhis, Carole Plasson, Carole Burel, Alain Mareck, Carlos Afonso, Patrice Lerouge, Marie-Chrsitine Kiefer-Meyer, Narimane Mati-Baouche, **Muriel Bardor**. Understanding the regulation of the N-glycosylation pathway is a prerequisite to optimize microalgae as a cell factory for the production of biopharmaceuticals.

Conference of the International Society for Plant Molecular Farming, Jun 2018, Helsinki, Finland.

Août 2019

Rodolphe Dumontier, Pierre-Louis Lucas, Corinne Loutelier-Bourhis, Alain Mareck, Carlos Afonso, Patrice Lerouge, Narimane mati-Baouche, Muriel Bardor. ESI-MSn and IMS-MS to analyse N-Glycan isomers.

25th International Symposium on Glycoconjugates, Aug 2019, Milan, Italy.



FORMATIONS SUIVIES

Formation de l'EdNBISE

CV et lettre de motivation
Connaître le marché de l'emploi
Formation des « missions enseignement »
Intégrité scientifique

Responsabilités

Représentant élu des doctorants Rouennais de l'EdNBISE
Membre du bureau de l'EdNBISE

Organisation

Membre du comité d'organisation des 21^{ème} « journées de l'école doctorale » 2018 à Rouen

Chapitre 1

Synthèse bibliographique

I. Les biomédicaments

A. Définition d'un biomédicament

Un médicament est défini par le code de la Santé publique ([site internet : article L.5111-1](#)) de la façon suivante : « une substance ou composition présentée comme possédant des propriétés curatives ou préventives à l'égard des maladies humaines ou animales ». Il est également défini comme « toute substance ou composition pouvant être administrée, en vue d'établir un diagnostic médical ou de restaurer, corriger ou modifier les fonctions physiologiques en exerçant une action pharmacologique, immunologique ou métabolique » (**Cavazzana-Calvo and Debais, 2011**). Au début du 20^{ème} siècle, ces substances étaient principalement extraites de sources naturelles. De nos jours, ce sont majoritairement des molécules de synthèse. Depuis les années 80, une nouvelle catégorie de produits à visée thérapeutique a vu le jour : le « biomédicament ». Celui-ci se définit comme étant « une protéine, un acide nucléique ou une cellule utilisés à des fins thérapeutiques et synthétisés à partir d'organismes vivants à l'aide de méthodes biotechnologiques » (**Walsh, 2002 ; Walsh, 2014**).

Le terme « biotechnologie » correspond ici à l'utilisation d'organismes vivants sous leurs formes naturelles ou modifiées. La biotechnologie moderne, développée dans les années 70 à la suite de la découverte des techniques de clonage de gènes ou de fragments d'ADN, utilise les techniques de génie génétique et de fusion cellulaire (**Volkin et al., 2015**).

B. Les biomédicaments : cas particuliers des anticorps

Aujourd'hui, le marché du biomédicament est en constante progression. Les grands groupes pharmaceutiques internationaux investissent de plus en plus dans la commercialisation des biomédicaments. Ce marché génère des revenus globaux de 163 milliards de dollars US représentant 20% du marché pharmaceutique (**Otto et al., 2014**). Sur la période 2014-2018, l'agence américaine du médicament (Food and Drug Administration : FDA) a approuvé la commercialisation de 155 nouveaux biomédicaments contre 147 sur la période 2010-2014 (**Figure 1 ; Walsh, 2014 ; Walsh, 2018**). En Europe, 88 ont été approuvés

par l'agence européenne du médicament (European Medicines Agency : EMA) sur la même période.

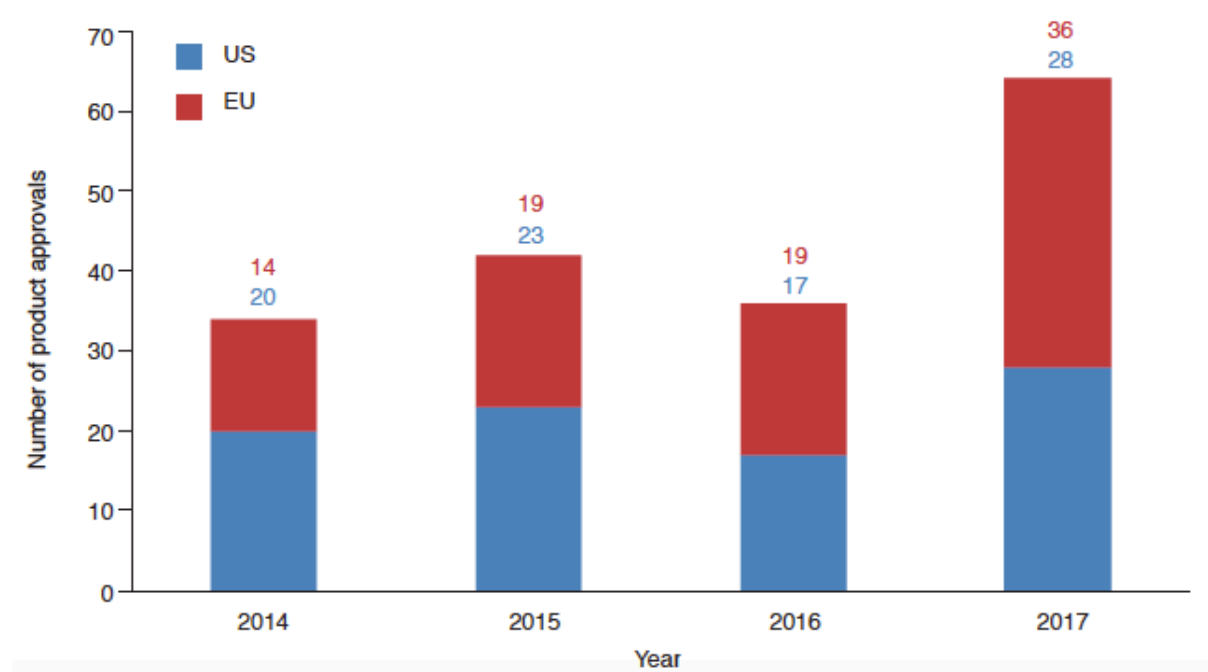


Figure 1 : Nombre de biomédicaments approuvés par l'agence réglementaire américaine FDA et l'agence de l'Union Européenne EMA (Walsh, 2018)

Le nombre de biomédicaments approuvés par la FDA est représenté en bleu et celui des biomédicaments approuvés par l'EMA en rouge.

EU : Union européenne ; US : Etats-Unis ; FDA : Food and Drug Administration ; EMA : European Medicines Agency.

La production de biomédicaments est impossible par synthèse chimique. En effet, les médicaments obtenus par synthèse chimique sont majoritairement composés de molécules de petites tailles et de structures moléculaires simples alors que les biomédicaments sont essentiellement des molécules complexes et de grandes tailles comme les protéines. C'est notamment le cas des anticorps monoclonaux dont la masse moléculaire avoisine les 170 kDa (**Rader, 2008**). Ces anticorps monoclonaux ne peuvent être produits que par des organismes vivants eucaryotes parfaitement adaptés pour synthétiser des molécules complexes de hautes masses moléculaires.

Les anticorps monoclonaux constituent un secteur majeur sur le marché des biomédicaments représentant à eux seuls une valeur d'environ 80 milliards de dollars US en 2016 avec une projection de 114,6 milliards de dollars US en 2022. Le marché des anticorps

monoclonaux est en constant développement (**Grilo and Mantalaris, 2018**). Actuellement, la production d'anticorps monoclonaux est réalisée majoritairement dans des cellules animales et plus précisément dans les cellules CHO (Chinese Hamster Ovary) pour deux raisons principales : (i) les anticorps monoclonaux produits possèdent une glycosylation très similaire à celle retrouvée sur les anticorps humains ; par conséquent, les modifications génétiques nécessaires pour éliminer des épitopes allergènes sont limitées ; (ii) les anticorps monoclonaux produits par ces cellules sont sécrétés directement dans le milieu de culture ce qui facilite leur purification (**Schillberg et al., 2019**). De plus, la production d'anticorps monoclonaux dans les cellules CHO a été optimisée et permet aujourd'hui d'atteindre des rendements très importants de l'ordre de 5-10g/L et pouvant monter dans certains cas jusqu'à plus de 20g/L (**Pujar et al., 2017**). En 2014, 47 biomédicaments correspondant à des anticorps monoclonaux étaient disponibles sur les marchés européen et américain (**Ecker et al., 2015**). Quarante-quatre étaient produits dans des cellules de mammifères. Parmi ces anticorps monoclonaux, l'Abciximab® a été commercialisé dès 1997 par le groupe Janssen Biotech sous le nom de ReoPro™. Il est composé d'anticorps monoclonaux chimériques (humain-murin) capables de se lier aux récepteurs de glycoprotéines présentes sur les plaquettes humaines et ainsi d'inhiber l'agrégation plaquettaire (**Miller et al., 2003**).

Pour distinguer l'origine et la cible thérapeutique des anticorps monoclonaux utilisés, une nomenclature internationale a été adoptée (**Scheen, 2009**). Le nom d'un anticorps monoclonal se termine par le suffixe « mab », précédé par une syllabe caractéristique de son origine. On parle d'anticorps «o-mab» lorsque celui-ci est de nature murine, de «xi-mab» lorsque l'anticorps monoclonal provient d'une version chimérique murin-humain, de «zu-mab» lorsque celui-ci est de nature murine et qui a été par la suite humanisé, et enfin «u-mab» pour les anticorps monoclonaux entièrement humains.

Quel que soit l'organisme choisi pour la production d'anticorps monoclonaux, ceux-ci doivent être capables de réaliser des modifications co- et post-traductionnelles sur les anticorps monoclonaux qu'ils produisent. Ces modifications sont propres à l'espèce dans laquelle l'anticorps est produit et jouent un rôle essentiel dans la fonction, la stabilité et la demi-vie des anticorps monoclonaux obtenus (**Lingg et al., 2012**).

II. Les anticorps : importance des modifications co- et post-traductionnelles

A. Généralités

Pendant et après sa traduction, une protéine subit des modifications co- et post-traductionnelles afin que celle-ci soit correctement repliée et active. De nombreuses modifications co- et post-traductionnelles des protéines ont pu être caractérisées (**Figure 2 ; Kannicht and Fuchs, 2008 ; Walsh and Jefferis, 2006**). Parmi ces modifications des protéines, on peut retrouver notamment, la phosphorylation, la méthylation, l'acétylation et les glycosylations. Il existe 2 types de glycosylation majoritaires sur les protéines : la O-glycosylation (**Brockhausen, 1999 ; Brockhausen, 2006 ; Gerken, 2012**) et la N-glycosylation (**Lerouge *et al.*, 1998 ; Breitling and Aebi, 2013**). Des analyses bio-informatiques ont montré que la N-glycosylation est la modification post-traductionnelle la plus abondante chez les protéines (**Figure 2 ; Khoury *et al.*, 2011 ; Apweiler *et al.*, 1999**). Ce type de glycosylation est présent également sur les anticorps.

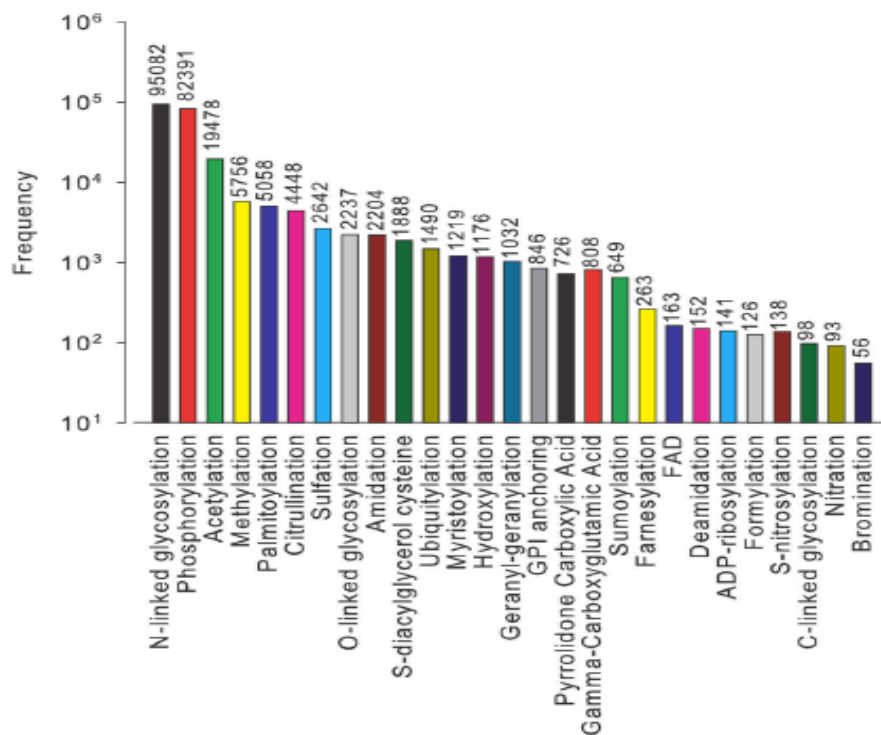


Figure 2 : Représentation statistique des modifications post-traductionnelles putatives des protéines les plus abondantes chez les eucaryotes

Cet histogramme représente les prédictions de modifications post-traductionnelles putatives présentes sur les protéines synthétisées par les cellules eucaryotes à partir de la base de donnée Swiss-Prot (**Khoury et al., 2011**). Les valeurs annotées sur l'axe des ordonnées correspondent à la fréquence de prédiction de ces modifications post-traductionnelles sur les protéines.

B. Structure d'un anticorps monoclonal

Les anticorps monoclonaux sont des glycoprotéines complexes constituées de deux chaînes lourdes et de deux chaînes légères reliées entre elles par des ponts disulfures (**Schroeder and Cavacini, 2010**). Ils possèdent deux parties, la première est une partie appelée Fab pour *Fragment antigen-binding* constituée des domaines de la chaîne légère : V_L et C_L et d'une partie des domaines de la chaîne lourde : V_H et C_{H1} (**Figure 3**). La seconde partie est une partie constante appelée le Fragment cristallisable (Fc) constituée des domaines C_{H2} et C_{H3} de la chaîne lourde (**Figure 3**). Il existe différents types d'immunoglobuline : les immunoglobulines de type G, de type A, de type M et de type E. Les anticorps jouent un rôle majeur dans le système immunitaire des mammifères. Ils reconnaissent et éliminent les agents pathogènes de manière spécifique. Dans le cadre de cette synthèse bibliographique, seuls les anticorps monoclonaux de type G (IgG) seront décrits.

Les IgG monoclonaux possèdent au minimum un site de *N*-glycosylation localisé sur l'asparagine située en position 297 (N297) présente sur la partie constante des chaînes lourdes au niveau du domaine C_{H2} (**Figure 3**). Environ 20% des anticorps monoclonaux possèdent un autre site de *N*-glycosylation situé sur la partie variable présente au niveau du Fab (**Holland et al., 2006 ; Arnold et al., 2007 ; Stadlmann et al., 2010**)

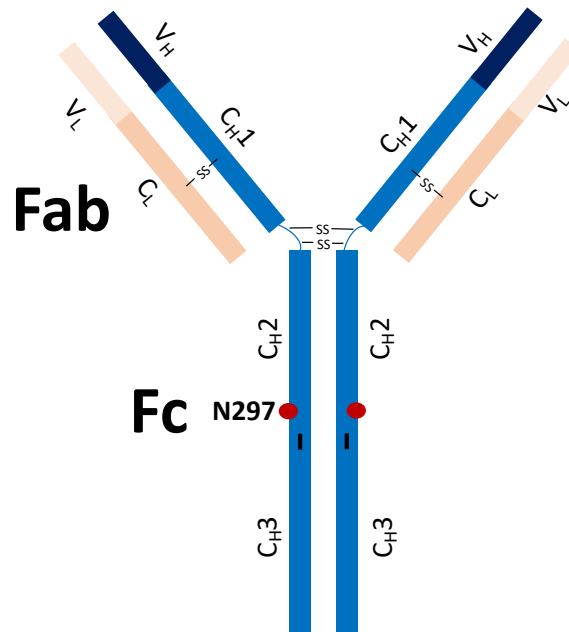


Figure 3 : Structure protéique d'un anticorps monoclonal de type IgG

On peut voir en nuances de bleu la chaîne lourde et en nuances de beige la chaîne légère. La chaîne lourde est divisée en quatre domaines : V_H , C_{H1} , C_{H2} et C_{H3} . Le fragment Fab (Fragment antigen-binding) de l'anticorps est composé de la chaîne légère ($V_L + C_L$) et d'une partie de la chaîne lourde ($V_H + C_{H1}$). La reconnaissance avec l'antigène se fait par la partie Fab. Le fragment Fc (Fragment cristallisable) est le domaine de reconnaissance de l'anticorps monoclonal avec un récepteur permettant notamment l'activation du complément lors d'une réaction immunitaire. Les sites de N-glycosylation du fragment Fc sont représentés par les points rouges en position 297 (N297). Les ponts disulfures reliant les chaînes lourdes entre elles et les chaînes lourdes avec la chaîne légère sont représentés par -SS-.

Dans le cas d'une production d'anticorps monoclonaux de type IgG dans un système d'expression donné, la N-glycosylation des anticorps synthétisés jouera un rôle primordial sur leur fonction biologique (Zheng *et al.*, 2011 ; Lingg *et al.*, 2012). En effet, la N-glycosylation assure la stabilité, la conformation et la participation aux fonctions effectrices des anticorps monoclonaux (Lingg *et al.*, 2012 ; Higel *et al.*, 2016). De plus, elle sera responsable d'une potentielle immunogénicité chez l'Homme en cas de présence de certains épitopes glucidiques spécifiques sur les N-glycannes des anticorps recombinants utilisés pour une thérapie. Outre la glycosylation, l'anticorps monoclonal agrégé ou modifié (Ex : PEGylation, oxydation–des méthionines ...) peut être responsable d'immunogénicité (van Beers and Bardor, 2012).

Un anticorps monoclonal reconnaît son antigène par sa partie Fab (**Figure 3**). Cette fonction de reconnaissance de l'antigène est indépendante de sa glycosylation (**Ko et al., 2003 ; Liu et al., 2019**). A l'inverse, la fonctionnalité de la partie Fc de l'anticorps dépend de la N-glycosylation présente sur l'asparagine N297 (**Cobb, 2020**). La partie Fc se lie aux récepteurs Fc (**Figure 3 ; Liu, 2015**). Il existe chez l'Homme, 4 types de récepteurs Fc pour les anticorps de type IgG : FcγRI (CD64), FcγRII(a,b) (CD32), FcγRIII(a, b) (CD16) et FcRn (récepteur néonatal) (**Liu, 2015**). Les récepteurs FcγRI ont une forte affinité pour les IgGs alors que les autres ont plutôt une faible affinité pour ceux-ci (**Hayes et al., 2016**). L'affinité de ces récepteurs pour l'IgG est également dépendante de la structure des N-glycannes portés par les anticorps monoclonaux sur l'asparagine N297 (**Subedi and Barb, 2016**). La formation du complexe IgG-récepteur Fc va engendrer, selon la nature du récepteur, soit une réaction de phagocytose, soit une réaction appelée *Antibody-dependent cell-mediated cytotoxicity* (ADCC), ou soit une réaction nommée *Complement-Dependent Cytotoxicity* (CDC) (**Hayes et al., 2016**). Lors du mécanisme ADCC, la liaison de l'anticorps thérapeutique sur l'antigène permet le recrutement des cellules NK (Natural Killer) qui libèrent des molécules cytotoxiques et provoquent l'apoptose de la cellule (**Figure 4**). Lors du mécanisme CDC, la liaison de l'anticorps thérapeutique sur l'antigène permet l'activation de la cascade du complément entraînant la formation d'un complexe nommé *Membrane Attack Complex* (MAC), complexe formant des pores transmembranaires à la surface de la cellule permettant l'entrée d'eau et d'ions provoquant la lyse de la cellule cible (**Figure 4**). Dans un contexte d'une immunothérapie ciblée, le fonctionnement et le rôle du récepteur FcγRIIIa ont été plus spécifiquement étudiés du fait de son implication dans l'activité ADCC, mécanisme primordial pour tuer les cellules cancéreuses ciblées (**Hayes et al., 2014 ; Hayes et al., 2016**).

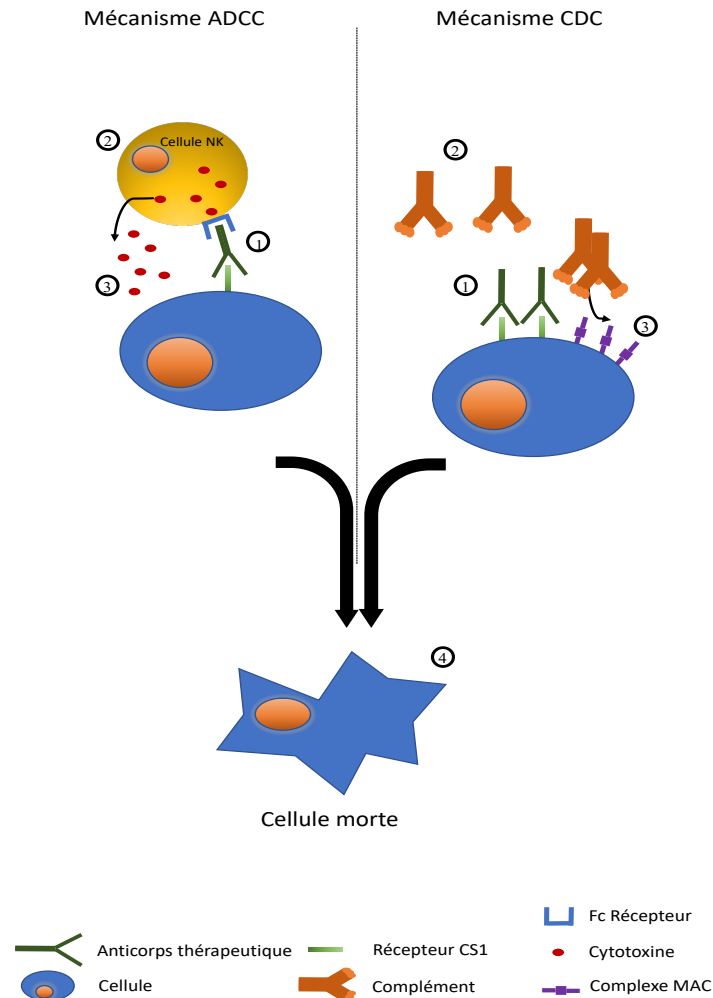


Figure 4 : Représentation des mécanismes ADCC (Antibody-Dependent Cell-mediated Cytotoxicity) (à gauche) et CDC (Complement-Dependent Cytotoxicity) (à droite) mis en jeu lors d'une thérapie par injection d'anticorps thérapeutiques (figure adaptée de Bernardo *et al.*, 2010)

Lors du mécanisme ADCC, la liaison de l'anticorps thérapeutique sur l'antigène (1) permet le recrutement des cellules NK (2), libérant des molécules cytotoxiques (3) qui déclenchent l'apoptose de la cellule (4). Lors du mécanisme CDC, la liaison de l'anticorps thérapeutique sur l'antigène (1) permet l'activation de la cascade du complément (2) entraînant la formation du complexe MAC (3) et provoquant la lyse cellulaire (4).

ADCC : Antibody-Dependent Cell-mediated Cytotoxicity ; CDC : Complement-Dependent Cytotoxicity ; NK : Natural Killer ; MAC : Membrane Attack Complex

1. Rôle des *N*-glycannes sur la conformation des anticorps et la perte des fonctions effectrices

La *N*-glycosylation joue un rôle important dans la stabilité des anticorps (Zheng *et al.*, 2011). L'absence de *N*-glycane sur l'asparagine N297 des anticorps monoclonaux provoque une mauvaise conformation du domaine C_H2 de la chaîne lourde. Ce défaut de la *N*-

glycosylation impacte en conséquence la fonction et la durée de demi-vie des anticorps monoclonaux thérapeutiques (**Krapp et al., 2003 ; Feige et al., 2009**). L'anticorps monoclonal Onartuzumab® produit chez *Escherichia coli* (*E. Coli*) est aglycosylé. Il a montré, lors de la phase pré-clinique, une durée du temps de demi-vie plus courte que la moyenne des anticorps monoclonaux glycosylés (8-12 jours contre 20-32 jours par exemple pour un anticorps monoclonal glycosylé produit chez les levures) (**Clarke et al., 2009**). L'absence de *N*-glycannes sur les anticorps monoclonaux a également pour conséquence une perte importante de l'affinité pour les récepteurs Fc présentés dans le paragraphe précédent (**Rudd et al., 2001 ; Zheng et al., 2011 ; Lingg et al., 2012**). Chez l'Homme, les *N*-glycannes portés par les anticorps monoclonaux possèdent majoritairement des résidus mannose (Man), *N*-acétylglucosamine (GlcNAc), fucose (Fuc) et galactose (Gal) (**Cobb, 2020**). Un à 13% de ces anticorps monoclonaux peuvent également porter des acides sialiques (**Figure 5 ; Thobhani et al., 2009 ; Gudelj et al., 2018**). Ces divers motifs glucidiques conduisent à une grande hétérogénéité des *N*-glycannes portés par les anticorps monoclonaux dont les structures *N*-glycanniques majoritaires sont présentées sur cette figure (**Figure 5**). L'influence de la structure fine des *N*-glycannes sur la fonction des anticorps sera détaillée dans le paragraphe ci-dessous.

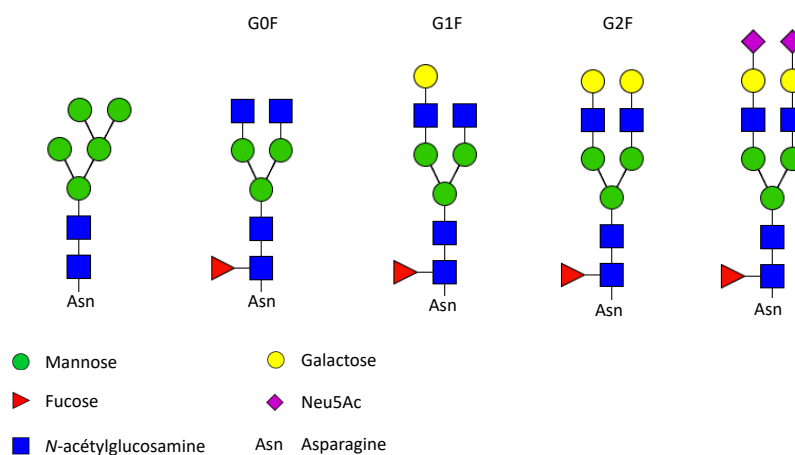


Figure 5 : Structures majoritaires des *N*-glycannes portés par les anticorps monoclonaux produits dans les cellules de mammifères

Les anticorps monoclonaux portent des *N*-glycannes de type $\text{Man}_5\text{GlcNAc}_2$ et des *N*-glycannes complexes bi-antennés (**Arnold et al., 2007 ; Spearman et al., 2011**). Les structures des *N*-glycannes sont dessinées en accord avec la nomenclature internationale (**Varki et al., 2015**).

G0F : $\text{GlcNAc}_2\text{Man}_3\text{FucGlcNAc}_2$; G1F : $\text{Gal}_1\text{GlcNAc}_2\text{Man}_3\text{FucGlcNAc}_2$; G2F : $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{FucGlcNAc}_2$.

a. Rôle du fucose sur le core glycanique

Le fucose présent sur le core du *N*-glycane semble important pour la fonction effectrice de l'anticorps. En effet, l'absence de ce fucose provoque une augmentation de la cytotoxicité des anticorps (**Figure 5**). La présence d'un *N*-glycane dépourvu de fucose sur le core glycanique montre, de plus, une augmentation de l'affinité de la liaison avec le récepteur FcγRIIIa et donc une meilleure activité ADCC sans altérer l'activité CDC (**Figure 4 ; Subedi and Barb, 2016 ; Yamaguchi and Barb, 2020**). La présence de ce fucose sur l'anticorps engendrerait une affinité 4 fois plus faible pour le récepteur FcγRIIIa (**Subedi and Barb, 2016**).

La production d'anticorps monoclonaux à intérêt thérapeutique dépourvus de ces résidus fucose est intéressante, notamment pour une immunothérapie ciblée contre les cancers. L'industrie pharmaceutique a donc généré des souches de CHO mutantes pour le gène *FUT8* codant l' $\alpha(1,6)$ -fucosyltransférase (**Yamane-Ohnuki et al., 2004 ; Malphettes et al., 2010**), enzyme responsable du transfert du résidu $\alpha(1,6)$ -fucose sur le core des *N*-glycannes chez les mammifères. Ces travaux ont montré que l'anticorps Rituximab® produit dans les cellules CHO dénuées d'activité *FUT8* possède une activité ADCC 100 fois supérieure à celle observée avec le même anticorps produit dans les cellules CHO sauvages (**Yamane-Ohnuki et al., 2004**). L'augmentation de l'activité ADCC provoquée par l'absence du résidu fucose a été confirmée par les travaux réalisés sur l'anticorps monoclonal utilisé dans le traitement contre le cancer du sein et commercialisé sous le nom de Pertuzumab® (**Luo et al., 2017**). D'autres stratégies expérimentales ont été entreprises afin de produire des anticorps monoclonaux dépourvus de résidus $\alpha(1,6)$ -fucose dans les cellules CHO. Parmi ces stratégies, la surexpression de la *N*-acétylglucosaminyltransférase III (GnT III) (**Umaña et al., 1999 ; Popp et al., 2018**) ou encore l'expression de la GDP-6-désoxy-D-lyxo-4-hexulose réductase bactérienne (**Von Horsten et al., 2010**). La surexpression de la GnT III dans les cellules CHO permet d'insérer un résidu GlcNAc sur le β -mannose du core, ce qui conduit à un oligosaccharide non reconnu par l' $\alpha(1,6)$ -fucosyltransférase permettant ainsi de générer des anticorps monoclonaux dépourvus de fucose. Dans la seconde approche, l'expression de la GDP-6-désoxy-D-lyxo-4-hexulose réductase bactérienne dans les cellules CHO permet de convertir le GDP-4-céto-6-désoxymannose en GDP-rhamnose. Par conséquent, cet intermédiaire n'est plus transformé en GDP-fucose par l'enzyme effectuant cette synthèse. Le

GDP-fucose, substrat donneur de l' $\alpha(1,6)$ -fucosyltransférase, n'est plus disponible dans ce cas, empêchant alors la fucosylation des glycoprotéines, y compris des anticorps monoclonaux qui seraient produits dans les cellules exprimant la réductase bactérienne.

b. Rôle des résidus galactose et acide sialique localisés en position terminale des N-glycannes

La majorité des anticorps monoclonaux produits chez les mammifères possèdent un ou deux résidus galactose liés en $\beta(1,4)$ en position terminale de leurs N-glycannes (**Figure 5**). Ces résidus galactose jouent un rôle important dans la fonctionnalité de ces anticorps. Les travaux réalisés sur les anticorps Rituximab® ont montré que la présence des résidus galactose en position terminale augmente l'activité CDC (**Figure 4 ; Boyd et al., 1995 ; Hodoniczky et al., 2005 ; Raju, 2008**). Par contre, la présence de ces résidus galactose en position terminale des N-glycannes semble peu affecter l'activité ADCC que peuvent déclencher ces anticorps (**Figure 4 ; Hodoniczky et al., 2005 ; Raju, 2008**). Des études complémentaires suggèrent également que la galactosylation des N-glycannes des anticorps monoclonaux permettrait dans certains cas une augmentation légère (environ 1,7 fois) de l'activité ADCC (**Yamaguchi et al., 2006 ; Subedi and Barb, 2016**).

La présence de résidus acide sialique en position terminale des N-glycannes de l'anticorps réduit leur affinité avec le récepteur Fc γ RIIIa, diminuant ainsi leur activité ADCC (**Scallon et al., 2007 ; Kaneko et al., 2006**). Par exemple, l'absence des résidus acide sialique sur l'anticorps Pertuzumab® provoque une augmentation de 2 à 4 fois de l'activité ADCC et d'environ 4 fois de l'activité CDC (**Luo et al., 2017**).

Un autre rôle essentiel des acides sialiques dans la N-glycosylation concerne le temps de demi-vie des anticorps. En effet, il a été montré que la présence d'acides sialiques augmente fortement le temps de demi-vie des anticorps monoclonaux (**Wright and Morrison, 1997 ; Liu et al., 2011**). La présence de ce résidu osidique protège les anticorps de leur dégradation en diminuant l'affinité de la liaison entre les anticorps et les récepteurs ASGPR (asialoglycoprotein receptors) reconnaissant les glycoprotéines non sialylées (**Hu et al., 2014 ; Luo et al., 2017**). En effet, les anticorps monoclonaux possédant un résidu galactose en position terminale sur leurs N-glycannes sont rapidement internalisés par les récepteurs ASGPR présents sur les hépatocytes. Les anticorps thérapeutiques galactosylés sont donc

rapidement éliminés de l'organisme après leur injection du fait de ce mécanisme d'action (**Hu et al., 2014 ; Luo et al., 2017**). En complément, il a été montré que la sialylation pouvait augmenter l'activité anti-inflammatoire des IgGs d'environ 10 fois (**Anthony et al., 2008 ; Anthony et al., 2012 ; Kaneko et al., 2006 ; Cobb, 2020**).

c. Présence de résidus mannose et GlcNAc en position terminale des N-glycannes des anticorps monoclonaux

Les anticorps monoclonaux produits peuvent également porter des N-glycannes de type Man₅GlcNAc₂ (**Figure 5**). La présence de ce type de N-glycane n'est pas sans conséquence pour l'anticorps. En effet, cette structure glycanique provoque une augmentation de l'affinité avec le récepteur FcγRIIIa ce qui contribue donc à une augmentation de l'activité ADCC de l'anticorps (**Zhou et al., 2008 ; Yu et al., 2012 ; Liu, 2015**). Cependant, cette augmentation de l'activité ADCC pourrait ne pas être liée directement à la présence de résidus mannose mais plutôt à l'absence de résidu fucose sur le résidu GlcNAc proximal des N-glycannes (**Zhou et al., 2008 ; Konno et al., 2012 ; Yu et al., 2012**). De façon similaire, une augmentation de 10 à 20 fois de l'activité ADCC a également été observée sur les anticorps portant des résidus GlcNAc en position terminale (**Shinkawa et al., 2003 ; Hodoniczky et al., 2005 ; Liu, 2015**).

La présence de N-glycannes de type Man₅GlcNAc₂ sur les anticorps monoclonaux aurait également pour conséquence de diminuer leur temps de demi-vie (**Jones et al., 2007**). Ce type de N-glycane est en effet capable de se lier à des récepteurs affins pour les résidus mannose présents au niveau des membranes des hépatocytes provoquant ainsi l'internalisation puis l'élimination de l'anticorps (**Jones et al., 2007**). De même, les anticorps possédant des résidus GlcNAc en position terminale seraient capables de se lier aux récepteurs affins pour les résidus mannose des N-glycannes, diminuant ainsi leur temps de demi-vie (**Figure 5 ; Raju et al., 2001 ; Jones et al., 2007 ; Raju, 2008**).

d. Le rôle des N-glycannes dans l'immunogénicité des anticorps

Une N-glycosylation d'anticorps monoclonaux thérapeutiques trop éloignée de la N-glycosylation humaine peut être responsable d'une immunogénicité chez les patients lors des traitements (**Lingg et al., 2012 ; Van Beers and Bardor, 2012**). Cette immunogénicité est provoquée par certains motifs osidiques reconnus comme appartenant au non-soi par le

système immunitaire humain. On peut citer le motif osidique $\alpha(1,3)$ -galactose et le motif acide *N*-glycolylneuraminique (Neu5Gc) présents sur les glycoprotéines de certains mammifères ou encore les motifs core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose présents sur les *N*-glycannes des protéines de plantes. Ces épitopes immunogènes sont introduits par le système d'expression utilisé ou les conditions de culture utilisées pour la production d'IgG recombinantes. Les cas documentés d'immunogénicité apportés par les *N*-glycannes des IgG recombinantes et les solutions apportées seront développés dans le paragraphe IV. de cette introduction bibliographique.

III. La voie de la *N*-glycosylation chez l'Homme

A. Généralités

Comme évoqué précédemment, la *N*-glycosylation des protéines est la modification co- et post-traductionnelle la plus commune chez toutes les espèces eucaryotes (**Figure 2 ; Khoury *et al.*, 2011**). Ce processus consiste en l'ajout d'un polymère osidique, appelé glycanne sur une protéine en cours de synthèse. Les sites de fixation de ces glycannes sur la protéine sont dépendants de séquences consensus particulières de type Asn-X-Ser/Thr/Cys où X peut être n'importe quel acide aminé excepté la proline. La *N*-glycosylation chez les protéines eucaryotes débute au niveau du réticulum endoplasmique (RE) avec une première phase cytosolique suivie par une phase luminale. Par la suite, ce processus se poursuit dans l'appareil de Golgi par des étapes de maturation permettant l'obtention d'un *N*-glycanne complexe. Dans ce paragraphe, nous détaillerons la voie de la *N*-glycosylation chez l'Homme. Puis, nous introduirons dans les paragraphes suivants, les différentes stratégies d'ingénierie réalisées dans différents systèmes d'expression dans le cadre d'une production optimisée d'anticorps monoclonaux portant une *N*-glycosylation adaptée pour un traitement thérapeutique chez l'Homme.

B. La synthèse du précurseur dans le réticulum endoplasmique

1. Phase cytosolique

L'assemblage du précurseur dans le RE se déroule en deux phases : une phase cytosolique avec le début de synthèse d'un précurseur oligosaccharidique lié à une ancre lipidique appelée dolichol et une phase luminale conduisant à la fin de la synthèse du précurseur oligosaccharidique et au transfert de l'oligosaccharide précurseur du dolichol vers une protéine en cours de synthèse donnant ainsi naissance à la glycoprotéine (**Figure 6 ; Breitling and Aebi, 2013**). La biosynthèse de l'oligosaccharide précurseur est réalisée par l'action successive de glycosyltransférases appelées Asparagine Linked Glycosylation (ALG) (**Figure 6**) utilisant pour substrats donneurs des nucléotides-sucres activés et des accepteurs spécifiques (**Figure 6 ; Snider and Rogers, 1984**). Une première glycosyltransférase, appelée ALG7, va ainsi initier la synthèse de l'oligosaccharide précurseur en transférant un résidu GlcNAc depuis un substrat donneur phosphorylé (GlcNAc-P) sur une ancre lipidique, le dolichol phosphate (Dol-P), formant le dolichol pyrophosphate-GlcNAc (GlcNAc-PP-Dol). Cette étape est suivie par l'ajout d'un second résidu GlcNAc catalysé par le complexe enzymatique formé par l'association de l'ALG13 et de l'ALG14.

Ensuite, 3 mannosyltransférases (ALG1, ALG2 et ALG11, respectivement) vont permettre le transfert de cinq résidus mannose (Man) à partir du GDP-Man, substrat donneur nucléotidique, sur l'accepteur. Ces glycosyltransférases forment, comme les précédentes, un complexe multi-enzymatique (**Gao et al., 2004**). Le premier mannose est transféré par la $\beta(1,4)$ -mannosyltransférase codée par *ALG1*, le second et le troisième résidu mannose sont transférés en $\alpha(1,3)$ et en $\alpha(1,6)$ par l'ALG2. Les deux derniers résidus mannose, transférés par l'ALG11, permettent la formation d'un intermédiaire oligosaccharidique de type $\text{Man}_5\text{GlcNAc}_2\text{PP-Dol}$ (**Figure 6**).

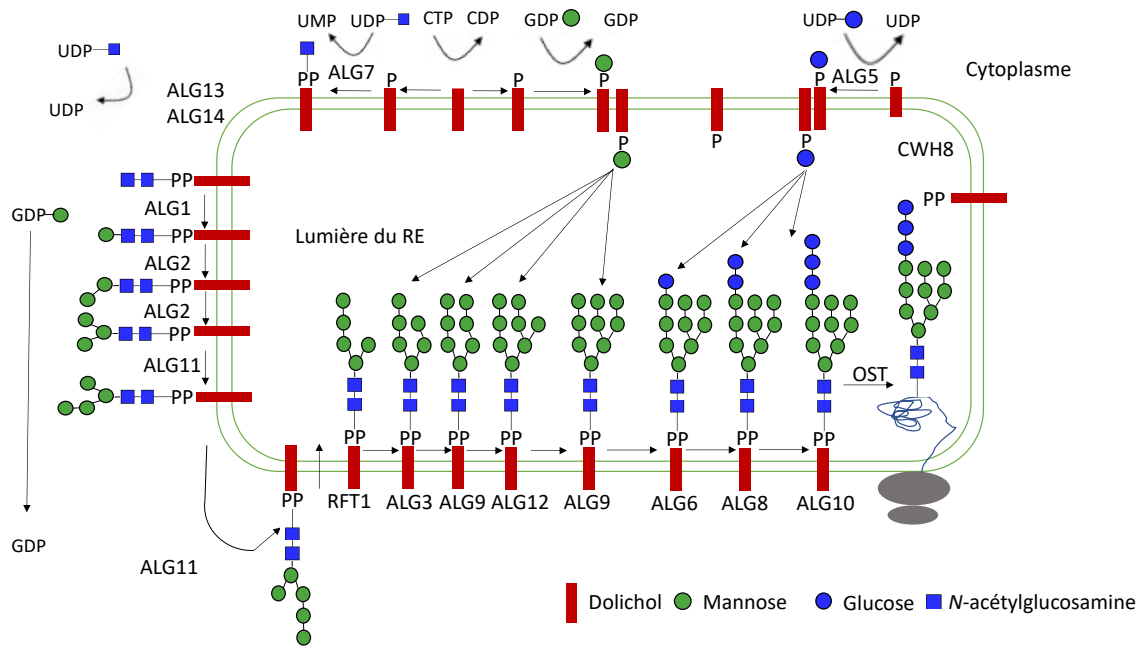


Figure 6 : Voie de la biosynthèse du précurseur oligosaccharidique lié au dolichol puis de son transfert sur une protéine en cours de synthèse chez l'Homme (figure adaptée de Aebi, 2013)

Ce schéma résume les étapes initiales de la *N*-glycosylation se déroulant tout d'abord sur la phase cytosolique puis sur la phase luminale du réticulum endoplasmique. Le changement d'orientation est opéré par l'enzyme flippase RFT1. Les étapes suivantes consistent en l'ajout de différents monosaccharides de façon séquentielle. Elles sont catalysées par une succession d'enzymes appelées ALG (Asparagine linked Glycosylation). Ces étapes conduisent à la formation d'un précurseur oligosaccharidique final lié au dolichol composé de 2 résidus *N*-acétylglucosamine, 9 résidus mannose et 3 résidus glucose. Ce précurseur oligosaccharidique est ensuite transféré sur une protéine en cours de synthèse par l'oligosaccharyltransférase (OST). Les structures des *N*-glycannes sont dessinées en accord avec la nomenclature internationale (Varki *et al.*, 2015).

UDP : uridine diphosphate ; UMP uridine monophosphate ; GDP : guanosine diphosphate ; ALG : asparagine linked glycosylation ; DPM : dolichol-phosphate mannosyltransférase I ; OST : oligosaccharyltransférase ; CTP : cytosine triphosphate ; CDP : cytosine diphosphate ; PP : pyrophosphate

2. Phase luminale

La suite du processus de la *N*-glycosylation des protéines humaines se déroule dans la lumière du RE. L'intermédiaire oligosaccharidique subit une translocation par une enzyme de type flippase (Rush, 2016). Celle-ci est appelée RFT1 chez les mammifères (Alaimo *et al.*, 2006 ; Rush and Waechter, 1995). Cette flippase est primordiale dans le processus de *N*-glycosylation. Son absence est en effet responsable d'une maladie chez l'Homme caractérisée par une quasi-absence de protéines glycosylées (Vleugels *et al.*, 2009). Ce syndrome provoque

des symptômes tels qu'un retard de croissance, une perte auditive neurosensorielle ou encore des troubles épileptiques (**Aeby et al., 2016**). Une fois l'oligosaccharide précurseur orienté vers la lumière du RE, l'action séquentielle des glycosyltransférases ALG3, ALG6, ALG8, ALG9, ALG12, et ALG10 va permettre le transfert de 4 résidus Man et de 3 résidus glucose (Glc) pour former le précurseur oligosaccharidique de structure finale, $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{PP-Dol}$.

C. Transfert du glycanne sur une protéine en cours de synthèse

L'oligosaccharyltransférase (OST) est le complexe enzymatique qui transfère le précurseur oligosaccharidique précédemment synthétisé depuis le dolichol vers une protéine en cours de synthèse sur la séquence consensus Asn-X-Ser/Thr/Cys (**Silberstein and Gilmore, 1996 ; Lizak et al., 2011 ; Breitling and Aebi, 2013**). Plus précisément, celle-ci catalyse la formation d'une liaison *N*-osidique entre un azote de l'asparagine et le carbone anomérique de la GlcNAc proximale du glycanne. L'OST humaine est composée de cinq chaînes polypeptidiques différentes pouvant être associées à la sous-unité catalytique STT3a (on parlera alors de OST-A) ou STT3b (OST-B) (**Mohorko et al., 2011 ; Hamieh et al., 2017 ; Wild et al., 2018 ; Ramirez et al., 2019**). En supplément, une sous-unité appelée DC2 peut être associée à l'OST-A et une sous-unité MAGT1 peut s'associer à l'OST-B. Une nouvelle sous-unité appelée SélénO_T a été mise en évidence récemment (**Hamieh et al., 2017**). Cette dernière sous-unité interagit avec la sous-unité KCP2 et est impliquée dans la régulation de la sécrétion hormonale. Ces études montrent que les deux sous-unités KCP2 et SélénO_T ne peuvent être associées qu'à l'OST-A (**Hamieh et al., 2017**). Récemment, une analyse structurale par cryo-microscopie électronique de l'OST humaine a été réalisée (**Ramirez et al., 2019**). Ces données structurales fournissent des informations essentielles sur la fonction de l'OST-A et celle de l'OST-B. Tout d'abord, les complexes OST-A et OST-B interagissent de manière identique avec le précurseur oligosaccharidique. Cependant, l'interaction avec la partie protéique diffère : l'OST-A se lie majoritairement à des polypeptides non conformés alors que l'OST-B se lie préférentiellement aux protéines partiellement conformées (**Ramirez et al., 2019**). Ces résultats permettent donc de conclure que l'OST-A, la plus conservée, est responsable du transfert co-translationnel du précurseur oligosaccharidique sur la protéine naissante lorsque celle-ci rentre dans la lumière du RE alors que l'OST-B est responsable du transfert post-

traductionnel du précurseur oligosaccharidique sur la glycoprotéine. L'activité enzymatique de ces complexes est portée par les sous-unités STT3a ou b retrouvées à ce jour chez tous les eucaryotes (**Kelleher and Gilmore, 2006**). Il a déjà été montré que l'inactivation d'une des deux sous-unités provoque une réduction du taux de glycoprotéines formées (**Ruiz-Canada et al., 2009**).

D. Cycles de repliement et de contrôle qualité

À la suite du transfert du glycanne sur la protéine, l'action de deux α -glucosidases (GSI et GSII) permet à la glycoprotéine naissante d'entrer dans les cycles de repliement et de contrôle qualité (**Roth et al., 2010**). En effet, des études suggèrent que l'addition et la modification du glycanne *N*-lié dépendent étroitement de la conformation de la protéine (**Ferris et al., 2014**). Ce contrôle de la conformation protéique joue un rôle fondamental dans la voie de biosynthèse des *N*-glycannes. Deux lectines chaperonnes, la calnexine et la calréticuline, vont intervenir lors de ce processus (**Williams, 2006**). Ce cycle de contrôle qualité a pour objectif d'identifier et d'éliminer les glycoprotéines mal conformées. Ce processus commence par l'élimination du premier résidu glucose catalysée par la GSI, formant le glycanne intermédiaire $\text{Glc}_2\text{Man}_9\text{GlcNAc}_2$ (**Hubbard and Robbins, 1979**). Ensuite, ce glycanne intermédiaire s'associe à la malectine, une lectine chaperonne, induite par le stress du RE. Ce stress est induit par l'accumulation de protéines mal conformées dans la lumière du RE. Cette association permettrait (i) d'éviter une agrégation des polypeptides naissants lors de la synthèse précoce et (ii) de recruter la GSII (**Figure 7 ; Helenius and Aebi, 2004 ; Schallus et al., 2008 ; Galli et al., 2011**). Par la suite, la GSII catalyse une seconde déglucosylation éliminant le premier résidu glucose lié en $\alpha(1,3)$. Le $\text{Glc}_1\text{Man}_9\text{GlcNAc}_2$ ainsi formé, va ensuite être reconnu soit par la chaperonne calnexine, soit par la chaperonne calréticuline (**Figure 7 ; Hammond et al., 1994**). La glycoprotéine va ainsi s'associer à l'une ou à l'autre des deux lectines présentes dans le RE : soit à la calnexine pour les glycoprotéines membranaires, soit à la calréticuline pour les glycoprotéines solubles (**Leach et al., 2002**). Par la suite, ERp57, la thiodisulfide oxydoréductase et la cyclophiline B, une isomérase, se lieront ensemble à la calréticuline ou à la calnexine pour participer à la mise en conformation finale de la glycoprotéine (**Ferris et al., 2014**).

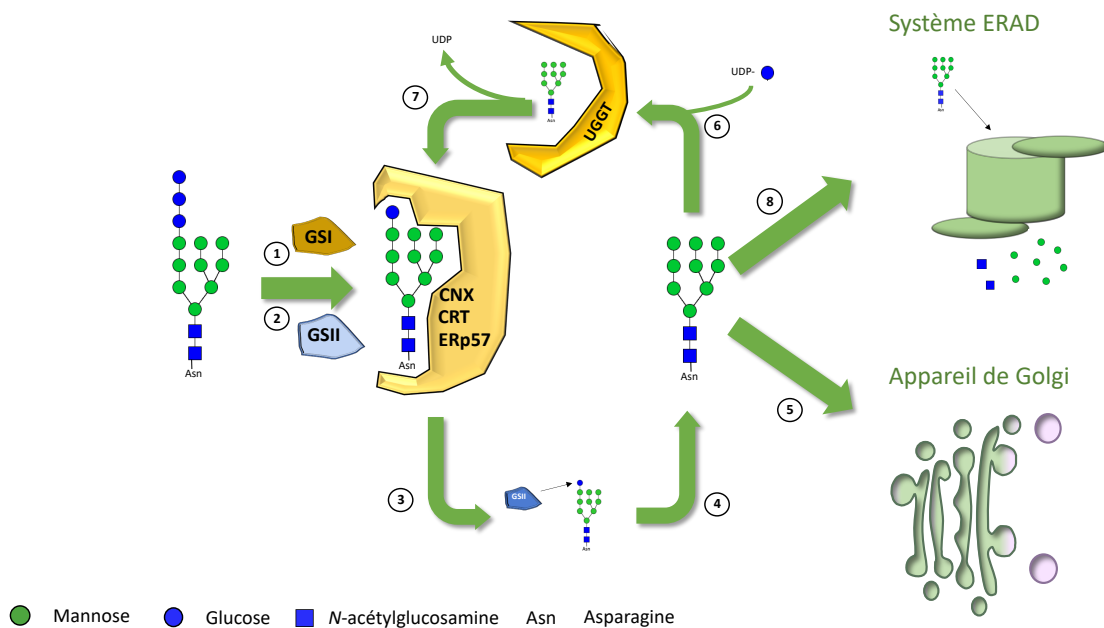


Figure 7 : Cycle de contrôle qualité des glycoprotéines humaines impliquant les chaperonnes calnexine et calréticuline (figure adaptée de Ferris *et al.*, 2014)

Etapes 1 et 2 : la GSI et la GSII éliminent les deux premiers résidus glucose du précurseur oligosaccharidique $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$.

Etape 3 : la glycoprotéine est prise en charge par le complexe CNX/CRT/ERp57 puis la GSII élimine le dernier résidu glucose.

Etapes 4 et 5 : si la glycoprotéine est correctement conformée, celle-ci sera envoyée dans l'appareil de Golgi afin que le N-glycane subisse des maturations supplémentaires.

Etapes 4, 6 et 7 : si la glycoprotéine n'est pas correctement conformée, celle-ci subira de nouveau le cycle de contrôle qualité jusqu'à ce qu'elle soit correctement repliée. Le signal d'une nouvelle entrée dans le cycle de contrôle qualité est orchestré par l'UGGT, enzyme capable d'ajouter un résidu glucose précédemment hydrolysé. La glycoprotéine subit de nouveau les étapes de contrôle qualité jusqu'à obtenir une bonne conformation

Etape 8 : la glycoprotéine est dégradée par le système ERAD si elle subit plusieurs fois le cycle de contrôle qualité sans succès. Les structures des N-glycannes sont dessinées en accord avec la nomenclature internationale (Varki *et al.*, 2015).

GSI : α -glucosidase I ; GSII : α -glucosidase II ; CNX : calnexine ; CRT : calréticuline ; UGGT : UDP-glucose: glycoprotéine glucosyltransférase ; ERAD : Endoplasmic Reticulum Associated Degradation

Par la suite, une troisième déglucosylation aura lieu. Celle-ci est orchestrée là encore par l'alpha-glucosidase II. Deux possibilités s'offrent alors à la protéine : soit la glycoprotéine est bien conformée et sort du cycle de contrôle qualité, soit elle ne l'est pas et l'enzyme UDP-glucose: glycoprotéine glucosyltransférase (UGGT) transférera un nouveau résidu glucose lié

en $\alpha(1,3)$ sur le *N*-glycane de la glycoprotéine mal conformée. Cette dernière modification permet à la glycoprotéine de rentrer à nouveau dans le cycle de contrôle qualité et de lui redonner une chance de se replier correctement. En cas de nouvel échec, celle-ci sera envoyée pour être dégradée dans le système ERAD pour *Endoplasmic Reticulum Associated Degradation* (**Figure 7 ; Ruggiano et al., 2014**). Brièvement, un des résidus mannose du *N*-glycane présent sur la protéine mal conformée sera éliminé par la Mannosidase I du réticulum endoplasmique (ER-Mannosidase I), une $\alpha(1,2)$ -mannosidase permettant l'obtention d'un glycane oligomannosidique composé de 8 résidus mannose (**Benyair et al., 2015**). Cette étape primordiale entraîne la diminution de l'affinité de la glycoprotéine mal conformée avec l'UGGT et les chaperonnes calnexine/calréticuline et donc d'éviter de subir de nouveau le cycle de contrôle qualité (**Molinari et al., 2003 ; Römisch, 2005**). L'élimination de ce résidu mannose permet à l'enzyme EDEM1 (ER-degradation enhancing mannosidase-like protein 1) d'éliminer un second résidu mannose. Cette seconde élimination permet à la glycoprotéine d'être directement reconnue par les lectines OS-9 et XTP3-B déclenchant la retranslocation de la glycoprotéine dans le cytosol puis son élimination par ubiquitination (**Tepedelen and Kirmizibayrak, 2019**).

E. Maturation des *N*-glycannes des protéines humaines dans l'appareil de Golgi

Les glycoprotéines, ayant passé le cycle de contrôle qualité avec succès, vont sortir du RE par des vésicules de sécrétion pour rejoindre le saccule *cis* de l'appareil de Golgi. Dans ce compartiment, des étapes de maturation séquentielle faisant intervenir des glycosyltransférases et des glycosidases vont conduire à la synthèse de structures glycaniques complexes spécifiques pour chaque espèce. Chez l'Homme, la première étape consiste en une démannosylation catalysée par l' α -mannosidase I (Man I) conduisant à une structure de type $\text{Man}_5\text{GlcNAc}_2$ (**Figure 8**). La localisation de cette enzyme est à ce jour controversée. Une étude récente a montré que cette glycosidase serait localisée dans le compartiment ERQC pour *ER-derived Quality Control Compartment* (**Benyair et al., 2015**).

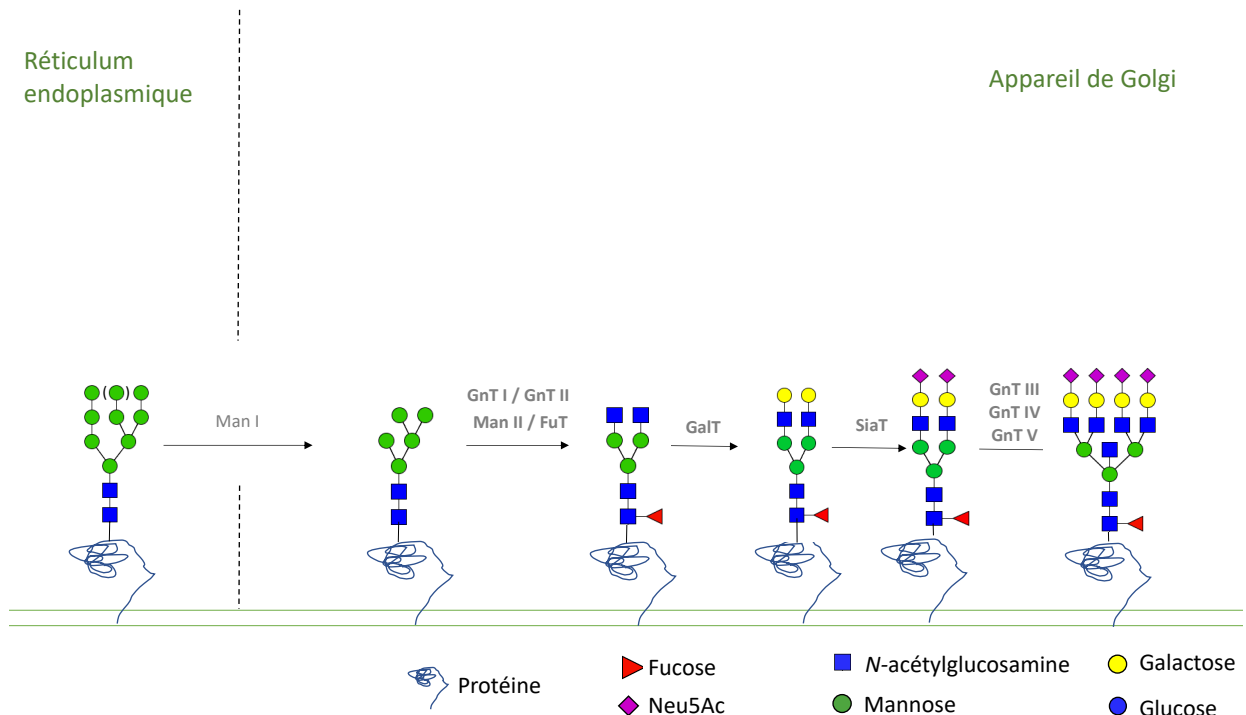


Figure 8 : Processus de la maturation des *N*-glycannes en *N*-glycannes complexes au sein de l'appareil de Golgi chez l'Homme

La Man I initie le processus de la maturation des *N*-glycannes dans l'appareil de Golgi en éliminant des résidus mannose pour former un intermédiaire de type $\text{Man}_5\text{GlcNAc}_2$, structure primordiale à la synthèse des *N*-glycannes complexes. Une succession d'étapes faisant intervenir différentes enzymes telles qu'une GnT I, une Man II, une GnT II, une FuT et une GalT et une SiaT, permet la synthèse d'une structure bi-antennée possédant des résidus acide *N*-acétylneuraminique (Neu5Ac) en position terminale. Les GnT III, GnT IV et GnT V puis des GalT et SiaT supplémentaires peuvent intervenir pour former une structure tétra-antennée sialylée. Les structures des *N*-glycannes sont dessinées en accord avec la nomenclature internationale (Varki *et al.*, 2015).

Man I : α -mannosidase I ; Man II : α -mannosidase II ; GnT I : *N*-acétylglucosaminyltransférase I ; GnT II : *N*-acétylglucosaminyltransférase II ; GnT III : *N*-acétylglucosaminyltransférase III ; GnT IV : *N*-acétylglucosaminyltransférase IV ; GnT V : *N*-acétylglucosaminyltransférase V ; FuT : α (1,6)-fucosyltransférase ; GalT : β (1,4)-galactosyltransférase I ; SiaT : α -sialyltransférase ; Neu5Ac : Acide *N*-acétylneuraminique

L'étape suivante de la maturation est catalysée par la *N*-acétylglucosaminyltransférase I (GnT I), une enzyme primordiale dans la maturation des *N*-glycannes en *N*-glycannes complexes. La GnT I est une protéine transmembranaire de type II (Breton *et al.*, 2006). Elle appartient à la superfamille des glycosyltransférases GT-A, connues pour utiliser un ion divalent comme cofacteur, et à la famille CAZy GT13 ([site internet : Cazy GT13](#) ; Lairson *et al.*, 2008). Elle est située dans les saccules *cis* et *médian* de l'appareil de Golgi. Cette enzyme joue

un rôle fondamental dans le développement embryonnaire des mammifères (**Metzler *et al.*, 1994**). Elle possède deux domaines protéiques liés par une région hydrophile de 23 acides aminés (**Ünlügil *et al.*, 2000**). Le premier domaine, du côté *N*-terminal, est composé de 8 feuillets bêta flanqués par 6 hélices alpha ainsi que de deux feuillets bêta antiparallèles. Le second domaine, du côté *C*-terminal, est quant à lui, composé de 4 feuillets bêta flanqués de 3 hélices alpha. Lors de la réaction enzymatique, la protéine change de conformation. L'UDP-GlcNAc, substrat donneur, est emprisonné au sein de la protéine du côté *N*-terminal et le substrat accepteur interagit avec le domaine catalytique qui forme une poche (**Figure 9 ; Ünlügil *et al.*, 2000 ; Gordon *et al.*, 2006**).

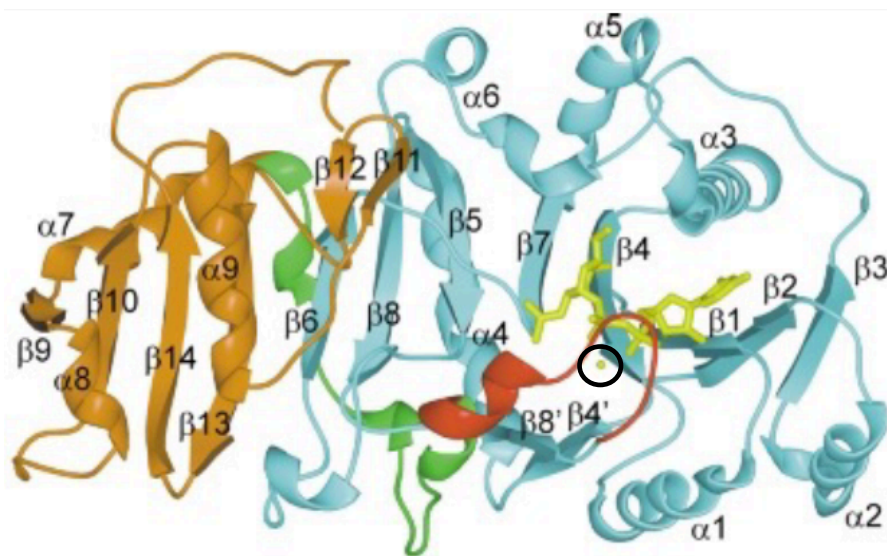


Figure 9 : Représentation tridimensionnelle de la structure de la Gnt I de lapin déterminée par cristallographie aux rayons X (Ünlügil *et al.*, 2000)

La partie en rouge correspond au site actif interagissant avec l'UDP-GlcNAc (en jaune-vert) et l'ion Mn^{2+} représenté par le point en jaune entouré en noir.

UDP-GlcNAc : Uridine diphosphate *N*-acétylglucosamine ; Gnt I : *N*-acétylglucosaminyltransférase I ; Mn^{2+} : Manganèse

L'étape catalysée par cette enzyme conduit ainsi à la formation d'une structure glycanique intermédiaire de type $GlcNAcMan_5GlcNAc_2$. Cet intermédiaire est rapidement modifié par l'action de deux autres enzymes, l' α -mannosidase II (Man II) et la *N*-acétylglucosaminyltransférase II (Gnt II), qui vont successivement hydrolyser les deux résidus mannose liés en $\alpha(1,3)$ et en $\alpha(1,6)$ de la structure glycanique $GlcNAcMan_5GlcNAc_2$ et ajouter un résidu GlcNAc sur le carbone 2 du mannose lié en $\alpha(1,6)$ permettant la synthèse d'une structure glycanique $GlcNAc_2Man_3GlcNAc_2$ (**Figure 8**). Cette structure, appelée « core », constitue le motif de base de la synthèse de *N*-glycannes complexes bi-antennés. Des

structures plus complexes de type tri- ou tétra-antennés et portant un résidu GlcNAc sur le mannose du core peuvent être formées par les actions successives des enzymes *N*-acétylglucosaminyltransférase III (GnT III), *N*-acétylglucosaminyltransférase IV (GnT IV) et *N*-acétylglucosaminyltransférase V (GnT V) (**Figure 8**). Ensuite, des résidus galactose liés en $\beta(1,4)$ vont être ajoutés sur les résidus GlcNAc terminaux de ce « core » glycanique par des $\beta(1,4)$ -galactosyltransférases (GalT) (**Mengle-Gaw et al., 1991**). Des résidus acide *N*-acétylneuraminique (Neu5Ac) se grefferont en $\alpha(2,3)$ ou en $(\alpha 2,6)$ sur les galactoses terminaux grâce à l'action de sialyltransférases (SiaT) (**Figure 8 ; Harduin-Lepers et al., 2001**).

IV. L'ingénierie métabolique de la voie de la *N*-glycosylation

Dans les chapitres précédents, il a été rappelé l'importance de la structure du *N*-glycane et des épitopes glucidiques pour la fonction biologique, la conformation et la stabilité des anticorps monoclonaux. Il a également été décrit l'immunogénicité potentielle due à la présence de certains épitopes glucidiques chez l'Homme. Les processus de *N*-glycosylation des protéines chez les systèmes d'expression d'anticorps recombinants utilisés en industrie (cellules animales) ou en développement (plantes, levures...) sont proches de celui de l'Homme décrit dans le chapitre III dans les étapes initiales mais différents dans les étapes finales de maturation prenant place dans l'appareil de Golgi. Ces voies de *N*-glycosylation doivent donc être éventuellement modifiées pour obtenir une glycosylation optimisée afin de préserver la tolérance immunitaire et les fonctionnalités optimales des anticorps monoclonaux thérapeutiques produits. Dans cette optique, des outils de génie génétique ont été utilisés pour modifier les cellules hôtes afin d'obtenir une voie de *N*-glycosylation similaire à celle retrouvée chez l'Homme. Ainsi, il s'est avéré nécessaire d'insérer ou d'inactiver les gènes codant les différentes glycosyltransférases et glycosidases dans le génome de l'hôte. Cette stratégie permet *in fine* la synthèse de glycoprotéines par la cellule hôte possédant des *N*-glycannes complexes optimisés pour une utilisation thérapeutique chez l'Homme. Dans ce contexte, on parlera d'optimisation ou d'humanisation des voies de *N*-glycosylation du système hôte.

A. Chez les mammifères

Les cellules de mammifères sont les systèmes d'expression majoritairement utilisés pour la production de protéines recombinantes comme les anticorps monoclonaux. Elles représentent 50% des systèmes biologiques de production d'anticorps monoclonaux actuellement sur le marché. Leur utilisation préférentielle est liée premièrement à leur capacité à synthétiser des glycoprotéines ayant une conformation correcte et des modifications co- et post- traductionnelles similaires à celles retrouvées chez l'Homme (**Jacobs and Callewaert, 2009**). Ces systèmes d'expression permettent également de contourner les problèmes de biais de codons que l'on peut rencontrer lors de la production de glycoprotéines recombinantes dans les autres systèmes d'expression hétérologue (**Gustafsson et al., 2004**). Les principales lignées cellulaires utilisées dans la production des biomédicaments sont les cellules CHO, les cellules Human Embryonic Kidney 293 (HEK) et la lignée cellulaire de souris NS0. Lors de l'évolution des espèces, de nouvelles glycosyltransférases sont apparues chez certaines lignées transférant des monosaccharides variés sur les *N*-glycannes des glycoprotéines, monosaccharides absents et allergènes pour l'Homme. Des méthodes d'ingénierie métabolique à visée « corrective » ont en conséquence été développées afin d'éliminer ces résidus indésirables.

1. Inactivation du gène codant pour l' $\alpha(1,3)$ -galactosyltransférase

L' $\alpha(1,3)$ -galactosyltransférase ($\alpha(1,3)$ GalT) est une enzyme de la biosynthèse des *N*-glycannes des glycoprotéines présente chez la plupart des lignées de mammifères sauf chez l'Homme et certains primates (**Jacobs and Callewaert, 2009**). Elle est responsable du transfert d'un résidu galactose lié en $\alpha(1,3)$ sur les galactoses terminaux liés quant à eux en $\beta(1,4)$ du glycanne *N*-lié (**Figure 10**). Environ 1% des anticorps présents dans le sérum humain sont dirigés contre cet épitope (**Jacobs and Callewaert, 2009 ; Hilger et al., 2019**). Il a été montré que le *N*-glycanne portant des épitopes $\alpha(1,3)$ -galactose était retrouvé sur les anticorps recombinants produits dans les cellules de souris NS0, système utilisé à l'échelle industrielle pour la production de protéines recombinantes (**Baker et al., 2001**). Le premier cas de ce « syndrome α -gal » a été découvert à la suite d'une hypersensibilité sévère observée chez 3% des patients ayant reçu une injection de l'anticorps monoclonal Cétuximab[®] produit dans

ces cellules murines (O'Neil *et al.*, 2007 ; Chung *et al.*, 2008). Une analyse structurale de cet anticorps a montré la présence de l'épitope immunogène α -gal, épitope exclusivement présent sur la région Fab de l'anticorps (Qian *et al.*, 2007). A la suite de ces découvertes, une souche de souris déficiente pour l'enzyme $\alpha(1,3)$ -galactosyltransférase a été générée permettant l'élimination complète de l'épitope immunogène $\alpha(1,3)$ -galactose (Basnet *et al.*, 2010).

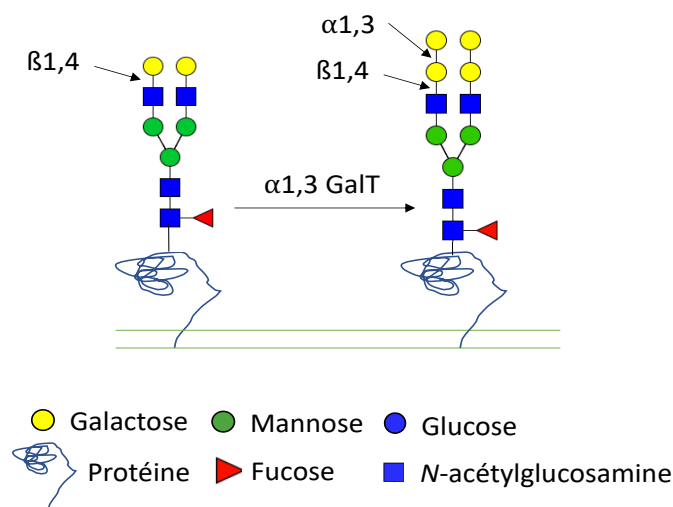


Figure 10 : Structures glycaniques portant l'épitope $\alpha(1,3)$ -galactose résultant de l'action de l' $\alpha(1,3)$ GalT

L'enzyme $\alpha(1,3)$ GalT est capable de transférer des résidus galactose terminaux en $\alpha(1,3)$ sur les résidus $\beta(1,4)$ -galactose présents en position terminale sur les N-glycannes. Cet épitope $\alpha(1,3)$ -galactose est immunogène pour l'Homme. Les structures des N-glycannes sont dessinées en accord avec la nomenclature (Varki *et al.*, 2015).

$\alpha(1,3)$ GalT : $\alpha(1,3)$ galactosyltransférase

2. Humanisation des processus de sialylation

Le Neu5Ac est un acide sialique présent sur les glycannes N-liés de la plupart des mammifères. Ce monosaccharide peut subir une hydroxylation par l'intervention de la CMP-Neu5Ac Hydroxylase (CMAH) pour conduire à la formation de Neu5Gc (Figure 11 ; Varki, 2001). Chez l'Homme, le gène codant cette hydroxylase est présent dans notre génome. Cependant, celui-ci est dépourvu d'un de ses exons, ce qui a pour conséquence que la protéine synthétisée est une protéine tronquée inactive (Larsen *et al.*, 1990).

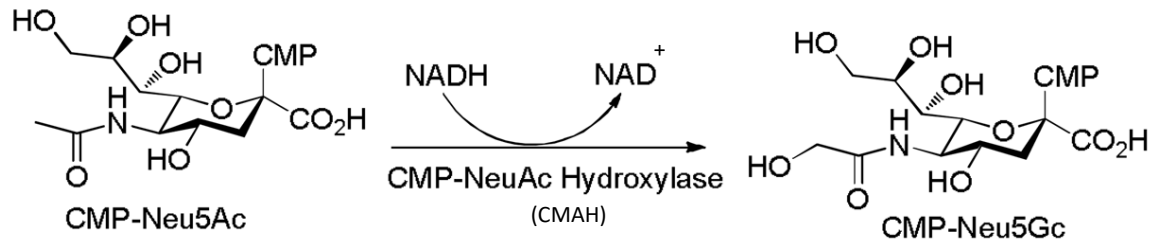


Figure 11 : Voie de la biosynthèse du Neu5Gc à partir du Neu5Ac (Yin and Huang, 2012)

Le CMP-Neu5Ac est hydroxylé en CMP-Neu5Gc par l'action de la CMP-Neu5Ac Hydroxylase.

CMP : cytidine monophosphate ; Neu5Gc : acide *N*-glycolylneuraminique ; Neu5Ac : acide *N*-acétylneuraminique

L'extrémité *N*-terminale de la protéine traduite, essentielle pour l'activité enzymatique, est en effet manquante (Irie and Suzuki, 1998). En conséquence, le Neu5Gc est absent des *N*-glycannes des protéines synthétisées chez l'Homme et donc immunogène pour l'Homme. Certains anticorps monoclonaux thérapeutiques produits dans des cellules animales portent des glycannes *N*-liés possédant du Neu5Gc en position terminale (Maeda et al., 2012 ; Ghaderi et al., 2010 ; Yu et al., 2016). Ces épitopes, absents chez l'Homme, sont cependant présent sur l'anticorps Cétuximab®. Comme évoqué dans le paragraphe précédent, certains patients traités par le Cétuximab® ont développé une hypersensibilité sévère survenant quelques minutes après l'injection de cet anticorps (Chung et al., 2008 ; van Beers and Bardor, 2012 ; Yu et al., 2016). Pour la production de protéines recombinantes thérapeutiques, il est donc nécessaire de réduire fortement le taux de Neu5Gc terminaux sur les *N*-glycannes. Dans cette optique, la réduction de l'activité enzymatique de la CMAH a permis de baisser le taux de Neu5Gc terminal de 4% à 1% dans les cellules CHO (Chenu et al., 2003). L'activité enzymatique a été réduite en utilisant des acides ribonucléiques (ARN) anti-sens dirigés contre le gène codant cette enzyme. De même, une souche de souris déficiente en activité CMAH a été obtenue conduisant à l'élimination complète des épitopes Neu5Gc immunogènes (Basnet et al., 2010). Il est à noter cependant que la présence de résidus Neu5Gc a été détectée « naturellement » dans des cellules humaines. En effet, des travaux ont montré que ce résidu pouvait être incorporé depuis le milieu de culture dans une cellule humaine *via* pinocytose puis être activé sous forme de CMP-Neu5Gc (Varki, 2001 ; Tangvoranuntakul et al., 2003 ; Bardor et al., 2005). Cette forme active peut ensuite être transportée et entrer dans l'appareil de Golgi pour être utilisée comme substrat par les SiaT humaines (Tangvoranuntakul et al., 2003 ; Bardor et al., 2005). Il a ensuite été montré que l'Homme possède la capacité d'incorporer le Neu5Gc à partir de son alimentation (notamment

s'il consomme un régime à base de viande rouge qui est riche en Neu5Gc) (**Samraj et al., 2015 ; Dhar et al., 2019**). Un apport excessif de Neu5Gc est responsable d'une maladie xenoauto-immune appelée « Xenosialitis » (**Dhar et al., 2019**). Dès lors, dans le cadre de la production d'anticorps recombinants à usage thérapeutique, il est primordial de s'assurer que la cellule hôte utilisée n'est pas en mesure de synthétiser le Neu5Gc ou d'en incorporer dans les *N*-glycannes synthétisés à partir du milieu de culture.

Un second problème lié à la sialylation des protéines est rencontré dans les systèmes de production de protéines recombinantes à usage thérapeutique utilisant des cellules de mammifères et humaines. Il s'agit de la position des résidus terminaux d'acides sialiques sur les *N*-glycannes. Chez l'Homme, on retrouve des acides sialiques majoritairement liés en $\alpha(2,6)$ et de façon minoritaire liés en $\alpha(2,3)$. Cependant, les protéines recombinantes produites dans des cellules de mammifères contiennent majoritairement des *N*-glycannes porteurs de résidus d'acide sialique liés en $\alpha(2,3)$ (**Lee et al., 1989**). Pour optimiser leur sialylation, le gène codant l' $\alpha(2,6)$ -sialyltransférase transférant des résidus d'acide sialique terminaux liés en $\alpha(2,6)$ a été inséré dans le génome des cellules CHO productrices de protéines recombinantes. Cette insertion a permis l'obtention de *N*-glycannes portant des acides sialiques terminaux liés en $\alpha(2,6)$ (**Zhang et al., 1998 ; Bragonzi et al., 2000 ; Bork et al., 2009 ; Houeix and Cairns, 2019**).

B. Chez les levures

Les levures ont l'avantage, comme tout eucaryote, de réaliser des modifications co- et post-traductionnelles telles que la *N*-glycosylation. Elles ont également la capacité de se développer sur des milieux nutritifs en absence de facteurs de croissance et procurent un rendement important de protéines sécrétées (**Hamilton and Gerngross, 2007**). C'est pourquoi les levures sont utilisées pour la production de protéines recombinantes. Le processus naturel initial de la *N*-glycosylation chez les levures aboutit à la formation d'un oligosaccharide précurseur de type $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ dans le RE semblable à la structure glycanne retrouvée chez l'Homme. Ce précurseur est ensuite transféré sur la glycoprotéine naissante par le complexe OST puis mûri en $\text{Man}_8\text{GlcNAc}_2$ selon un processus analogue à celui décrit

chez l'Homme. Le $\text{Man}_8\text{GlcNAc}_2$ issu de la maturation du précurseur après transfert sur les glycoprotéines est ensuite hypermannosylé dans l'appareil de Golgi pour conduire à des *N*-glycannes poly- voire hypermannosylés comportant de 15 à 100 résidus mannose (**Figure 12**).

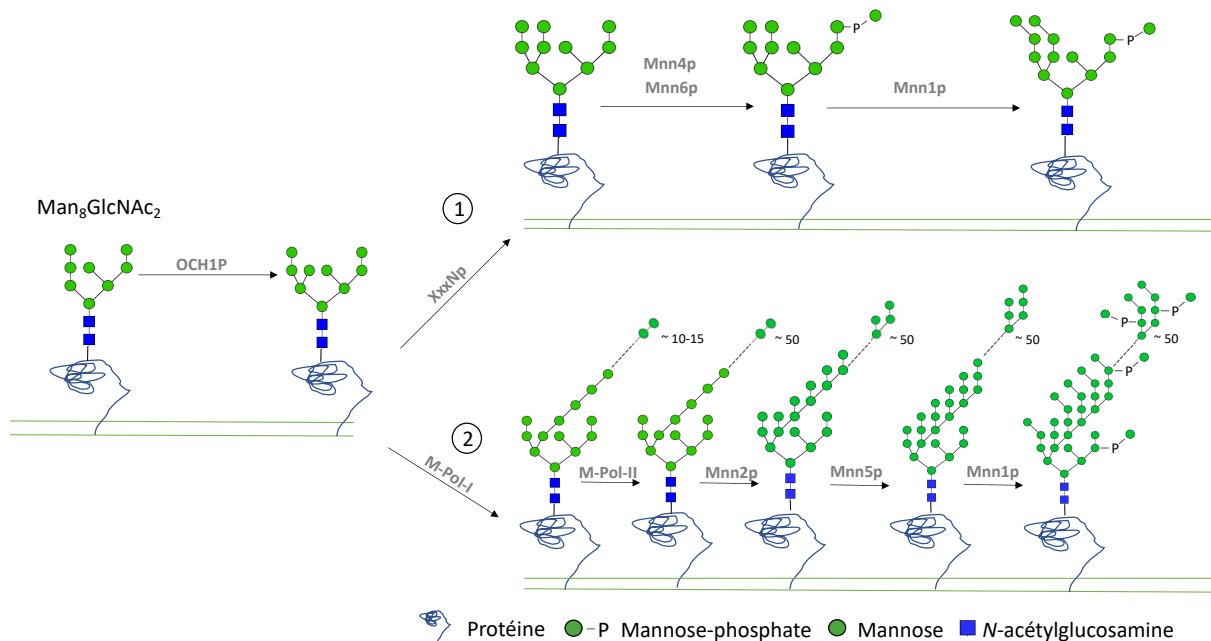


Figure 12 : Voie de la biosynthèse des *N*-glycannes hypermannosylés dans l'appareil de Golgi chez *Saccharomyces cerevisiae* (De Pourcq et al., 2010)

La première étape de la biosynthèse est catalysée par l'enzyme OCH1P transférant un résidu mannose sur le $\text{Man}_8\text{GlcNAc}_2$. Les glycannes formés sont ensuite séparés en deux groupes distincts : (1) le type hyperglycosylé et (2) le type core. Le type hyperglycosylé conduit à la formation de *N*-glycannes comportant jusqu'à une centaine de résidus mannose. L'addition de ces résidus est orchestrée par les glycosyltransférases M-Pol-I, M-Pol-II, Mnn2P, Mnn5p et Mnn1p. Le type core permet la synthèse de *N*-glycannes avec un degré de mannosylation plus faible. Cette mannosylation est formée sous l'action des glycosyltransférases XxxNp, Mnn4p, Mnn6p et Mnn1p. Les structures des *N*-glycannes sont dessinées en accord avec la nomenclature internationale (Varki et al., 2015).

M-Pol : Mannose polymérase ; MnnXp : mannosylphosphate-transférase ; XxxNp : mannosyltransférase de type inconnu ; OCH1P : $\alpha(1,6)$ -mannosyltransférase

1. Inhibition de l'hypermannosylation des *N*-glycannes des glycoprotéines de levure

La *N*-glycosylation des protéines chez les levures est donc caractérisée par une hypermannosylation des *N*-glycannes (De Pourcq et al., 2010). L'enzyme responsable de ce processus est l' $\alpha(1,6)$ -mannosyltransférase (OCH1P), impliquée dans le transfert de résidus mannose sur la branche $\alpha(1,3)$ du core glycanne et conduisant à la formation d'une

extension $\alpha(1,6)$ mannosidique. Cette enzyme initie la formation d'un *N*-glycane comportant des dizaines de résidus mannose (**Gemmill and Trimble, 1999 ; Wildt and Gerngross, 2005**). Dans ce contexte, la première stratégie mise en œuvre chez la levure a consisté à inhiber l'enzyme OCH1P. Une équipe a réussi à générer des souches de *Saccharomyces cerevisiae* mutées dépourvues de cette enzyme inhibant ainsi l'hypermannosylation des *N*-glycannes (**Nakanishi-Shindo et al., 1993**). Cependant, cette stratégie conduit à la formation d'une structure de type $\text{Man}_8\text{GlcNAc}_2$, inadaptée pour entreprendre l'humanisation des *N*-glycannes complexes chez *S. cerevisiae*.

La seconde stratégie mise en place chez la levure a donc visé à limiter l'hypermannosylation des *N*-glycannes par addition d'une activité enzymatique ciblée. Elle consiste à introduire dans le génome de la levure, le gène codant l' $\alpha(1,2)$ -mannosidase. La protéine produite contiendra un signal HDEL en C-terminal de sa séquence permettant sa rétention dans le RE. La rétention dans le RE de l' $\alpha(1,2)$ -mannosidase permet ainsi à l'enzyme d'éliminer de façon précoce les résidus mannosyles des *N*-glycannes en cours de maturation et d'empêcher l'action de l' $\alpha(1,6)$ -mannosyltransférase golgienne (**Figure 13**). Cette stratégie a été développée avec succès chez *S. cerevisiae* par l'insertion du gène codant l' $\alpha(1,2)$ -mannosidase d'*Aspergillus saitoi* dans son génome (**Chiba et al., 1998**). Cette équipe a ainsi démontré la formation d'une structure majoritaire de *N*-glycannes de type $\text{Man}_5\text{GlcNAc}_2$ chez la levure transformée. Cette structure est le substrat oligomannosidique nécessaire chez l'Homme pour la formation des *N*-glycannes complexes sur les protéines dans l'appareil de Golgi (**Hamilton and Gerngross, 2007**).

2. Humanisation des *N*-glycannes complexes des glycoprotéines chez la levure

Une fois le $\text{Man}_5\text{GlcNAc}_2$ obtenu sur les protéines par l'inactivation de l'hypermannosylation, les gènes codant les glucosidases et les glycosyltransférases humaines ont été introduits dans le génome de la levure hôte. Dans ce but, les chercheurs ont inséré le gène humain codant la GnT I et le gène codant un transporteur de l'UDP-GlcNAc de *Kluyveromyces lactis* dans une souche de *Pichia pastoris* préalablement transformée avec l' $\alpha(1,2)$ -mannosidase d'*Aspergillus saitoi* (**Figure 13 ; Choi et al., 2003**). Ensuite, 2 gènes codant respectivement une Man II de *Drosophila melanogaster* et une GnT II de *Rattus norvegicus* ont été insérés. Cette stratégie a été couronnée de succès chez cette levure avec

l'obtention d'un transformant nommé YSH44 capable de synthétiser un *N*-glycane de type GlcNAc₂Man₃GlcNAc₂ (**Figure 13 ; Bobrowicz et al., 2004 ; Hamilton et al., 2003**). L'insertion de la GalT humaine dans le génome de *P. pastoris* a ensuite permis d'obtenir des souches de levure capables de synthétiser des *N*-glycannes galactosylés (**Figure 13 ; Bobrowicz et al., 2004**).

La voie de la sialylation, dernière étape de l'humanisation des *N*-glycannes, a été introduite avec succès chez *P. pastoris* (**Bobrowicz et al., 2004 ; Hamilton et al., 2006**). Cette approche combine l'insertion des 14 gènes différents et permet d'obtenir une voie de la *N*-glycosylation complètement humanisée dans le transformant RDP762 de *P. pastoris* (**Hamilton et al., 2006**). Ces 14 gènes insérés correspondent à la Man I de souris, la Man II de *Drosophila melanogaster* pour l'élimination des résidus mannose du *N*-glycane dans l'appareil de Golgi (**Figure 13**), la GnT I humaine, la GnT II de rat et le transporteur UDP-GlcNAc de souris (**Figure 13**), la GalT humaine ainsi que le transporteur golgien capable de transporter l'UDP-Gal de *D. melanogaster* dans l'appareil de Golgi (**Figure 13**). Pour terminer, la sialylation des *N*-glycannes nécessite – (i) la biosynthèse cytosolique de CMP-Neu5Ac – (ii) le transport de ce nucléotide-sucré depuis le cytosol vers l'appareil de Golgi et – (iii) l'insertion du Neu5Ac sur les *N*-glycannes de *P. pastoris*. Dans ce but, 5 gènes humains, impliqués dans la biosynthèse du nucléotide-sucré CMP-Neu5Ac ont été insérés également dans le génome de la levure. En supplément, l'insertion de l' α (2,6)-sialyltransférase du rat et du transporteur golgien du CMP-Neu5Ac de la souris, ont permis l'accumulation de Neu5Ac sur les *N*-glycannes de *P. pastoris* (**Figure 13 ; Hamilton et al., 2006**).

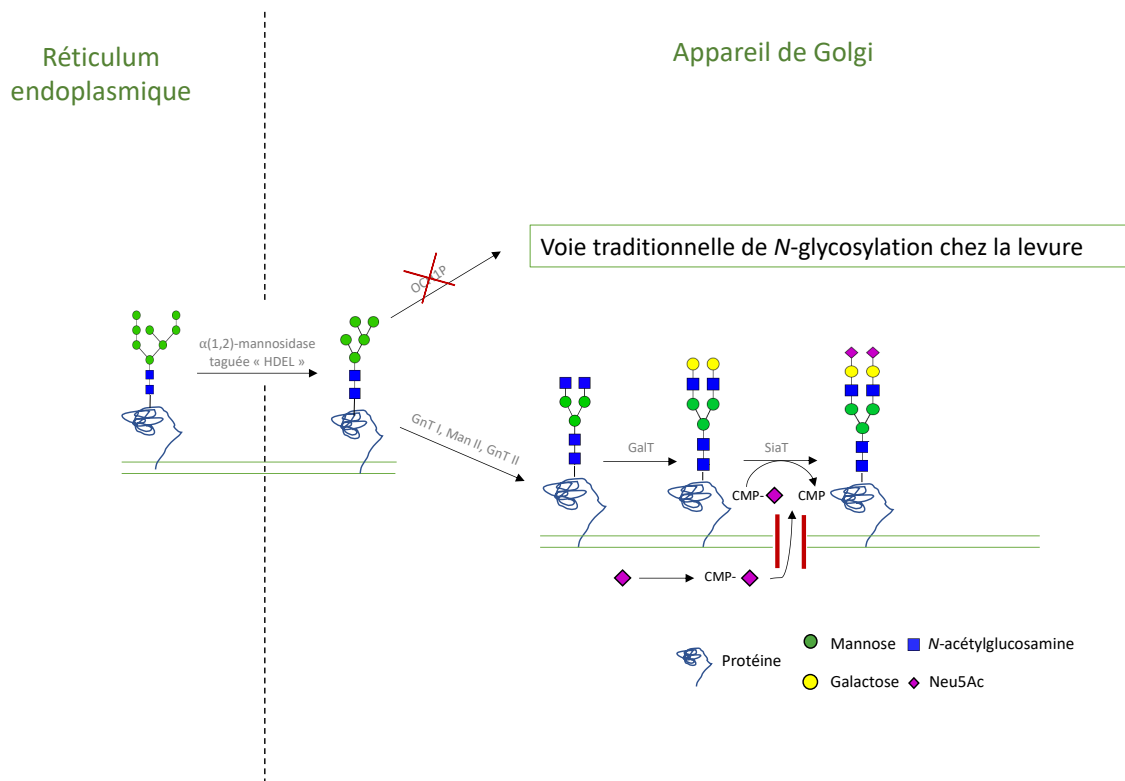


Figure 13 : Stratégie utilisée pour l'humanisation des voies de la N-glycosylation chez les levures

L'insertion de l' $\alpha(1,2)$ -mannosidase taguée par la séquence HDEL, qui retient spécifiquement l'enzyme dans le réticulum endoplasmique, permet (i) d'inhiber l'enzyme OCH1P et donc la synthèse de N-glycannes hypermannosylés, (ii) de synthétiser le N-glycanne $\text{Man}_5\text{GlcNAc}_2$, primordial à l'humanisation des voies de la N-glycosylation chez la levure. Par la suite, l'insertion des gènes codant la Gnt I, la Man II, la Gnt II, la GalT, la SiaT ainsi que ceux codant les enzymes de la voie complète de biosynthèse du CMP-Neu5Ac et de son transporteur permet de synthétiser un N-glycanne optimisé pour une application thérapeutique chez l'Homme. Les structures des N-glycannes sont dessinées en accord avec la nomenclature internationale (**Varki et al., 2015**).

Gnt I : N-acétylglucosaminyltransférase I ; Man II : α -mannosidase II ; Gnt II : N-acétylglucosaminyltransférase II ; GalT : $\beta(1,4)$ -galactosyltransférase ; SiaT : sialyltransférase

3. Production d'anticorps monoclonaux chez les levures

Peu d'articles scientifiques à ce jour relatent la synthèse d'anticorps monoclonaux recombinants chez la levure. La raison première est que l'hypermannosylation produite par la levure altère la stabilité des anticorps (**Olivier and Mehtali, 2009**). Cependant, l'ingénierie des voies de la N-glycosylation de la levure a rendu possible la production d'anticorps monoclonaux humanisés dans ce système. C'est le cas notamment chez *P. pastoris*, où GlycoFi Inc., une filiale du groupe Merck, annonça en 2006 avoir réussi la production de l'anticorps

monoclonal Rituximab® avec une glycosylation humanisée, anticorps utilisé entre autre dans le traitement de la polyarthrite rhumatoïde (**Li et al., 2006**). Cette équipe a inséré dans le génome d'une souche de *P. pastoris* tous les gènes permettant d'obtenir une voie humanisée de la *N*-glycosylation comme celle décrite dans le paragraphe précédent (**Li et al., 2006**), et également les gènes codant les chaînes lourde et légère de l'anticorps monoclonal Rituximab®. Il a été démontré de façon encourageante qu'une des 7 glycoformes identifiées pour cet anticorps possédait des *N*-glycannes galactosylés de type Gal₂GlcNAc₂Man₃GlcNAc₂ (**Li et al., 2006**). Les autres glycoformes étaient des *N*-glycannes de type Man₉GlcNAc₂, Man₈GlcNAc₂, Man₅GlcNAc₂, Man₃GlcNAc₂, GlcNAcMan₅GlcNAc₂ et GlcNAc₂Man₃GlcNAc₂. L'optimisation des conditions de culture (température à 26°C, pH 6,3...) a permis d'atteindre un rendement de 1,6 g/L d'anticorps monoclonaux avec une glycosylation humanisée (**Ye et al., 2011**).

C. Chez les plantes

Depuis les années 90, les plantes sont également apparues comme des systèmes alternatifs de production pour synthétiser des anticorps monoclonaux (**Hiatt et al., 1989 ; Ma et al., 1994**). En effet, les plantes sont considérées comme des systèmes économiques et sûrs pour la synthèse d'anticorps monoclonaux de par leur capacité de production importante et les risques faibles voire inexistantes de transmission de virus ou de prions (**Wycoff, 2005 ; Chargelegue et al., 2005**). En 2012, un premier biomédicament issu de la transgénése végétale a été commercialisé aux États-Unis. Ce biomédicament a pour principe actif la glucocérébrosidase synthétisée dans des cellules de carotte pour traiter les patients souffrant de la maladie de Gaucher (**Aviezer et al., 2009 ; Grabowski et al., 2014**).

La voie de la *N*-glycosylation chez les plantes possède l'avantage de partager les mêmes étapes de synthèse du précurseur oligosaccharidique que la voie de glycosylation humaine. Les plantes synthétisent également un core glycanique analogue à celui synthétisé chez l'Homme de type GlcNAc₂Man₃GlcNAc₂. Cependant, des glycosyltransférases complémentaires vont modifier ce core conduisant à la synthèse de *N*-glycannes complexes spécifiques des plantes (**Figure 14**).

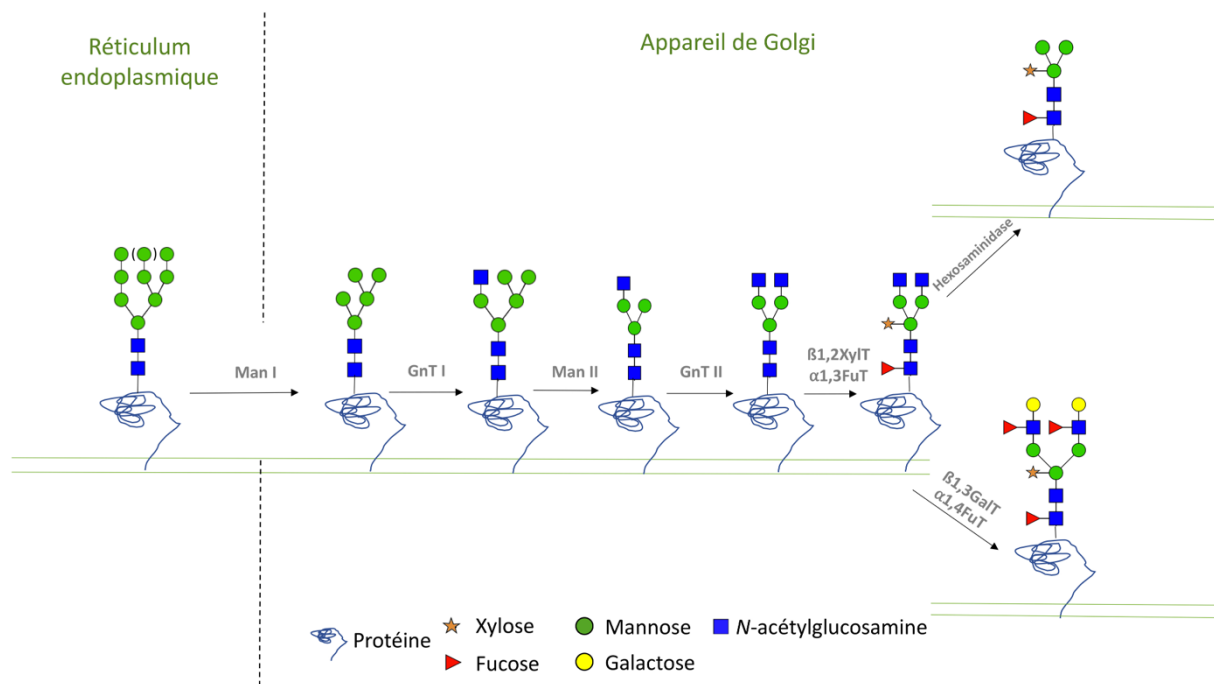


Figure 14 : Voie de la biosynthèse des *N*-glycannes complexes chez les plantes (figure adaptée de Lerouge *et al.*, 1998)

Le *N*-glycane $\text{Man}_{9-8}\text{GlcNAc}_2$, précédemment obtenu à l'issue des étapes localisées dans le RE est transformé en $\text{Man}_5\text{GlcNAc}_2$ par la Man I. Cette structure, substrat accepteur de la GnT I, permet la synthèse de *N*-glycannes complexes de type $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ obtenus par l'action successive de la GnT I, de la Man II et de la GnT II. La $\beta(1,2)$ -xylosyltransférase ($\beta(1,2)\text{XylIT}$) et l' $\alpha(1,3)$ -fucosyltransférase ($\alpha(1,3)\text{FuT}$) permettent les ajouts respectifs des épitopes xylose et fucose, connus pour être potentiellement immunogènes chez l'Homme. Par la suite, l'activité hexosaminidase peut cliver les deux résidus GlcNAc positionnés aux extrémités terminales du *N*-glycane pour conduire à la synthèse d'une structure paucimannosidique de type $\text{Man}_3\text{XylFucGlcNAc}_2$. Alternativement, la $\beta(1,3)$ -galactosyltransférase ($\beta(1,3)\text{GalT}$) et l' $\alpha(1,4)$ -fucosyltransférase ($\alpha(1,4)\text{FuT}$) permettent la synthèse d'épitopes terminaux appelés épitopes de Lewis^a (Fitchette-Lainé *et al.*, 1997). Les structures des *N*-glycannes sont dessinées en accord avec la nomenclature internationale (Varki *et al.*, 2015).

Man I : α -mannosidase I ; GnT I : *N*-acétylglucosaminyltransférase I ; Man II : α -mannosidase II ; GnT II : *N*-acétylglucosaminyltransférase II ; $\beta(1,2)\text{XylIT}$: $\beta(1,2)$ -xylosyltransférase ; $\alpha(1,3)\text{FuT}$: $\alpha(1,3)$ -fucosyltransférase ; $\beta(1,3)\text{GalT}$: $\beta(1,3)$ -galactosyltransférase ; $\alpha(1,4)\text{FuT}$: $\alpha(1,4)$ -fucosyltransférase

1. Élimination des épitopes antigéniques des *N*-glycannes des glycoprotéines produites chez les plantes

Les *N*-glycannes complexes des glycoprotéines synthétisés par les plantes possèdent des résidus spécifiques comme le fucose branché en $\alpha(1,3)$ (core $\alpha(1,3)$ -fucose) et le xylose branché en $\beta(1,2)$ sur le core glycanique (core $\beta(1,2)$ -xylose) (Bardor *et al.*, 2003a ; Gomord and Faye, 2004 ; Bakker *et al.*, 2006). Les sérums humains contiennent naturellement des

anticorps dirigés contre le core $\alpha(1,3)$ -fucose et le core $\beta(1,2)$ -xylose suggérant que ces deux épitopes, absents chez les mammifères, sont capables d'induire une réponse immunitaire chez l'Homme (**Bardor et al., 2003a**). Cependant, le rôle de ces résidus core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose dans l'immunogénicité reste encore débattu (**Altmann, 2007 ; Landry et al., 2010 ; Ward et al., 2014 ; Rup et al., 2017**). C'est dans ce contexte que des approches d'ingénierie métabolique ont été mises en place afin d'éliminer ces deux épitopes immunogènes, permettant de rendre compatibles les glycoprotéines recombinantes produites chez les plantes avec une utilisation thérapeutique chez l'Homme.

a. Rétention des protéines glycosylées dans le RE chez les plantes

Une première stratégie a consisté à éviter le transit de la glycoprotéine d'intérêt dans l'appareil de Golgi pour empêcher la maturation de ces *N*-glycannes en oligosaccharides complexes et accumuler majoritairement sur la protéine recombinante des *N*-glycannes oligomannosidiques. Dans cet objectif, le gène d'intérêt a été fusionné à une séquence codant un signal de rétention dans le réticulum endoplasmique, comme par exemple « KDEL ou HDEL » en 3'. Cette fusion permettra la rétention de la glycoprotéine dans le RE *via* l'étiquette « KDEL ou HDEL » présente en position C-terminale de la protéine. L'utilité de cette stratégie a été démontrée pour la production d'anticorps monoclonaux (**Sriraman et al., 2004 ; Petruccelli et al., 2006**).

b. Inhibition de l'expression chez les plantes des gènes codant les enzymes responsables de l'addition d'épitopes immunogènes sur les *N*-glycannes

En 2006 et 2008, différentes équipes ont respectivement, réussi à diminuer chez les plantes de façon plus ou moins forte, l'expression des gènes codant les $\beta(1,2)$ XylT et les $\alpha(1,3)$ FuT, responsables de l'addition des épitopes immunogènes $\alpha(1,3)$ -fucose et $\beta(1,2)$ -xylose dont les expressions ont été inhibées *via* l'utilisation de RNA interférents (RNAi) (**Cox et al., 2006 ; Sourrouille et al., 2008 ; Strasser et al., 2008**). L'utilisation de ces stratégies RNAi n'a pas permis l'élimination complète des épitopes immunogènes (core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose) mais seulement leur diminution partielle. Une seconde stratégie a consisté à générer des mutants appropriés afin d'inhiber efficacement les $\beta(1,2)$ XylT et les $\alpha(1,3)$ FuT.

Dans ce contexte, Richard Strasser et son équipe de l'Université de Boku en Autriche ont généré un triple mutant d'*Arabidopsis thaliana* dépourvu des activités $\beta(1,2)$ XylIT et $\alpha(1,3)$ FuT en inhibant le gène codant la $\beta(1,2)$ XylIT et les deux gènes codant l'activité $\alpha(1,3)$ FuT. Le triple mutant ainsi obtenu accumule des *N*-glycannes de type $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ dépourvus des épitopes immunogènes (**Figure 15 ; Strasser et al., 2004**). Des travaux plus récents, ont permis de générer des mutants de cellules de tabac BY-2 et de plantes de tabac (*Nicotiana benthamiana*), dépourvus de ces épitopes immunogènes par des techniques d'édition génétique de type TALEn ou encore CRISPR-cas9 (**Mercx et al., 2016 ; Mercx et al., 2017 ; Jansing et al., 2019**). Les travaux réalisés en 2019 ont permis de produire par CRISPR-cas9, des plantes mutantes de *N. benthamiana* dépourvues des gènes codant les deux xylosyltransférases et les quatre fucosyltransférases sur les 5 présentes. Ces plantes mutantes ont permis de produire un anticorps monoclonal dépourvu totalement des épitopes core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose (**Jansing et al., 2019**). En complément, l'analyse fonctionnelle de cet anticorps a montré qu'il possédait une capacité de liaison plus importante avec le récepteur Fc γ RI comparée à celle du même anticorps produit dans des plantes de tabac *N. benthamiana* sauvage (**Jansing et al., 2019**).

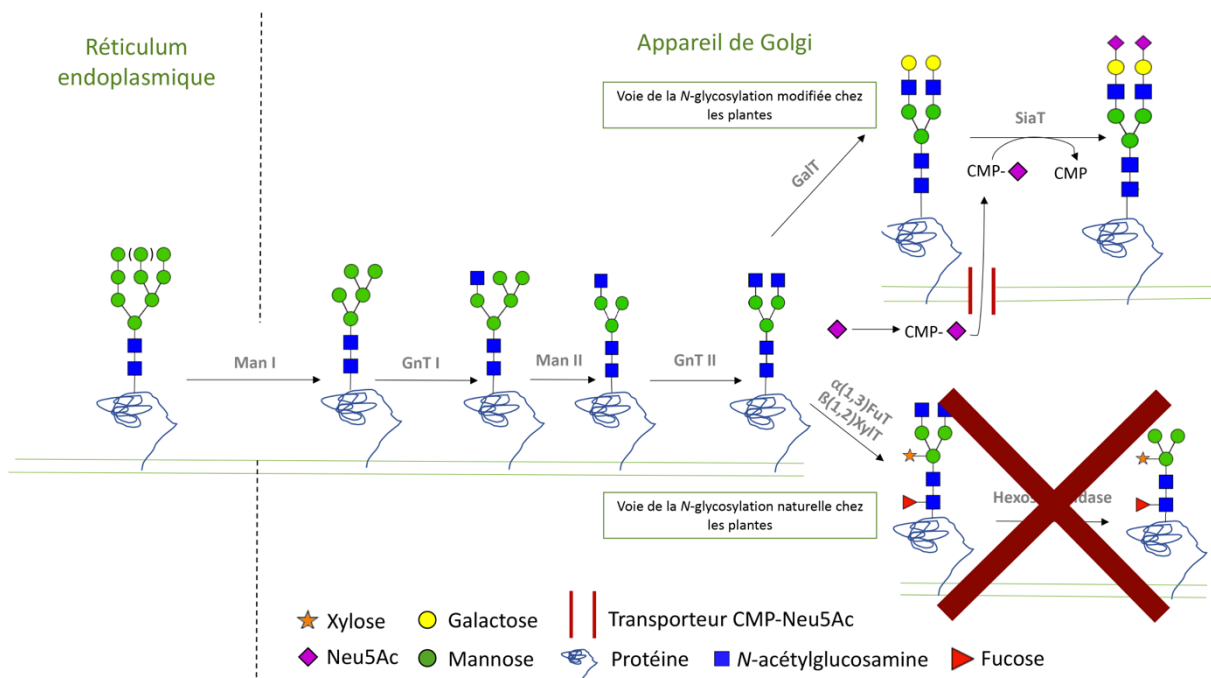


Figure 15 : Représentation des stratégies d'ingénierie de la voie de la N-glycosylation chez des plantes déficientes en activités $\beta(1,2)$ XylIT et $\alpha(1,3)$ FuT

La voie naturelle de la *N*-glycosylation chez les plantes conduit à la synthèse d'un *N*-glycanne paucimannosidique arborant des résidus $\alpha(1,3)$ -fucose et $\beta(1,2)$ -xylose sur le core

glycannique. La voie modifiée permet la synthèse d'une structure *N*-glycannique possédant des résidus Neu5Ac en position terminale. Les structures *N*-glycanniques sont dessinées en accord avec la nomenclature internationale (Varki *et al.*, 2015).

Man I: α -mannosidase I; GnT I: *N*-acétylglucosaminyltransférase I; Man II: α -mannosidase II; GnT II: *N*-acétylglucosaminyltransférase II; SiaT: sialyltransférase; GalT: β (1,4)-galactosyltransférase; α (1,3)FuT: α (1,3)-fucosyltransférase; β (1,2)XylT: β (1,2)-xylosyltransférase

2. Humanisation des *N*-glycannes complexes des glycoprotéines produites chez les plantes

L'inactivation des gènes codant les glycosyltransférases responsables du transfert des épitopes core β (1,2)-xylose et core α (1,3)-fucose conduit à l'accumulation de *N*-glycannes non immunogènes. Si cette stratégie peut s'avérer suffisante pour certaines applications thérapeutiques, l'absence de la maturation du core du *N*-glycanne peut être délétère pour d'autres applications. Les plantes sont capables d'accumuler des *N*-glycannes ayant des résidus GlcNAc terminaux (Figure 14; Figure 15). La présence de ces résidus est nécessaire pour la synthèse de *N*-glycannes complexes bi-antennés. Cependant, ces résidus sont éliminés au cours du transport de la glycoprotéine par des hexosaminidases (Figure 14; Figure 15; Altmann *et al.*, 1995; Shin *et al.*, 2017). Chez *A. thaliana* et *N. benthamiana*, 3 gènes codent des activités hexosaminidases, *HEXO1*, *HEXO2* et *HEXO3*. L'équipe de Richard Strasser a déterminé que l'enzyme codée par le gène candidat *HEXO3* était responsable de la formation de la structure tronquée. Son élimination, en utilisant des RNAi, a permis d'augmenter chez les plantes considérablement la formation de *N*-glycannes complexes portant des résidus GlcNAc terminaux (Shin *et al.*, 2017).

L'insertion du gène de la GalT humaine dans les génomes d'*A. thaliana* et de *N. benthamiana* a été réalisée par plusieurs équipes (Figure 15; Palacpac *et al.*, 1999; Bakker *et al.*, 2001; Bakker *et al.*, 2006; Vézina *et al.*, 2009; Schneider *et al.*, 2015). Cette insertion visait initialement à obtenir non seulement des *N*-glycannes possédant des résidus galactose en position terminale en espérant du même coup inhiber le transfert des résidus β (1,2)-xylose et α (1,3)-fucose. En effet, comme la GalT, les β (1,2)XylT et α (1,3)FuT des plantes utilisent les mêmes substrats accepteurs portant des résidus GlcNAc non substitués. Les auteurs ayant mis en œuvre cette stratégie pensaient pouvoir générer une compétition vis-à-vis du substrat accepteur entre les β (1,2)XylT, α (1,3)FuT et la GalT humaine inhibant ainsi l'addition des résidus xylose et fucose tout en permettant l'ajout des résidus β (1,4)-galactose manquants.

Les travaux de l'équipe de Palacpac et collaborateurs effectués chez les cellules BY-2 ont permis l'obtention d'une population *N*-glycannique totalement dépourvue de $\beta(1,2)$ -xylose et de $\alpha(1,3)$ -fucose de type GalGlcNAcMan₃GlcNAc₂ représentant environ 12% de la population glycannique (**Palacpac et al., 1999**). Cependant, une population hybride de type GalGlcNAcMan₅GlcNAc₂ a également été détectée à hauteur de 35,5% suggérant une maturation incomplète des *N*-glycannes (**Palacpac et al., 1999**). Deux années plus tard, les travaux par Bakker et collaborateurs effectués dans des plantes de tabac *N. benthamiana* ont permis l'obtention d'une population de protéines portant des *N*-glycannes substitués par des résidus galactose en position terminale (**Bakker et al., 2001**). Cette population de *N*-glycannes galactosylés représentait environ 30% des *N*-glycannes synthétisés. Parmi cette population de *N*-glycannes galactosylés, seule une très faible proportion de *N*-glycannes était à la fois dépourvue des épitopes immunogènes et porteurs du galactose en position terminale (**Bakker et al., 2001**).

Par la suite, des plantes ont été transformées avec la GalT humaine fusionnée à des domaines de rétention golgien d'autres glycosyltransférases, comme la GnT I ou la $\beta(1,2)$ XylT, afin de retenir l'enzyme GalT plus précocement dans l'appareil de Golgi. Cette stratégie a permis de réduire considérablement les taux de fucosylation et de xylosylation des *N*-glycannes des protéines produites, mais pas de les supprimer complètement (**Bakker et al., 2001 ; Bakker et al., 2006 ; Strasser et al., 2009 ; Vézina et al., 2009 ; Schneider et al., 2015 ; Navarre et al., 2017**). Par exemple, en 2006, les travaux de Bakker et collaborateurs ont permis l'élimination complète des *N*-glycannes portant l'épitope core $\beta(1,2)$ -xylose chez le tabac grâce à l'expression d'une glycosyltransférase chimérique comportant la partie catalytique de la GalT fusionnée avec le domaine d'insertion aux membranes golgiennes de la $\beta(1,2)$ XylT. Dans cette étude, 15,7% des *N*-glycannes observés sur les glycoprotéines portent l'épitope core $\alpha(1,3)$ -fucose. Parallèlement, cette complémentation a permis d'obtenir environ 20% des glycoprotéines possédant du galactose lié en $\beta(1,4)$ en position terminale des *N*-glycannes. Parmi cette population de *N*-glycannes portant des résidus galactose en position terminale, 2,5% contenait également l'épitope core $\alpha(1,3)$ -fucose (**Bakker et al., 2006**). Cette stratégie de relocalisation de la $\beta(1,4)$ -galactosyltransférase humaine dans l'appareil de Golgi ne semble donc pas être suffisante pour entraîner une élimination complète des épitopes antigéniques. Il est donc nécessaire de générer des plantes caractérisées à la fois par l'inactivation des gènes codant les glycosyltransférases responsables de la synthèse des

épitopes immunogènes et complémentées avec le gène codant la $\beta(1,4)$ -galactosyltransférase pour optimiser la production de protéines recombinantes à intérêt thérapeutique chez les plantes (**Castilho et al., 2010**).

La dernière étape de l'humanisation des voies de la *N*-glycosylation chez les plantes consiste à obtenir des *N*-glycannes contenant des résidus de type acide sialique comme le Neu5Ac en position terminale. Il a été montré, chez le riz, la présence de deux gènes putatifs ayant respectivement 33,2% et 31,2% d'homologie avec une séquence codant un transporteur de CMP-acide sialique humain (**Takashima et al., 2009**). Cependant, aucune trace d'acide sialique n'a pu être détectée dans les *N*-glycannes des protéines synthétisées chez le riz comme chez les autres plantes (**Séveno et al., 2004 ; Zeleny et al., 2006**). Il a été conclu de ces études que la protéine codée par ce gène est capable de transporter du CMP-Kdo (acide 3-désoxy-D-manno-oct-2-ulosonique), un nucléotide-sucré ayant un rôle primordial dans la synthèse du rhamnogalacturonan II (RGII), un polysaccharide pariétal des plantes (**Pérez et al., 2003 ; Delmas et al., 2008 ; Séveno et al., 2010**). Ce nucléotide-sucré est similaire aux nucléotides-sucrés des acides sialiques retrouvés chez les mammifères (**Münster-Kühnel et al., 2004**). Les précurseurs métaboliques nécessaires à la synthèse de CMP-Neu5Ac n'ont pas non plus été identifiés (**Séveno et al., 2004**).

La voie de la biosynthèse des acides sialiques comporte des étapes cytosoliques permettant d'aboutir à la formation du CMP-Neu5Ac, substrat donneur des sialyltransférases qui, elles, sont golgiennes. Ainsi, les plantes « humanisées » doivent avoir la capacité de synthétiser du Neu5Ac cytosolique ainsi que sa forme active, le CMP-Neu5Ac, puis de l'importer dans l'appareil de Golgi avant de pouvoir l'utiliser. Le gène de la sialyltransférase, absente des cellules végétales est également à insérer dans le génome des plantes afin de leur permettre de transférer les résidus Neu5Ac sur un *N*-glycane galactosylé de leurs protéines.

La GlcNAc est le seul précurseur métabolique des acides sialiques présent dans le cytosol des plantes pour initier la synthèse de Neu5Ac (**Figure 15**). Les deux premières études décrivant l'obtention de systèmes végétaux synthétisant du Neu5Ac ont porté sur des cellules de tabac et des plantes de luzerne. Ainsi, les cellules de tabac BY-2 et les plantes de luzerne ont été transformées par des gènes bactériens codant soit une Neu5Ac lyase, soit une Neu5Ac synthase, qui sont capables de convertir la D-*N*-acétylmannosamine (D-ManNAc) exogène et le pyruvate cytosolique en Neu5Ac (**Paccalet et al., 2007**). Plus tard, la transformation de plantes d' *A. thaliana* par les gènes codant l'UDP-GlcNAc 2-épimérase/*N*-acétylmannosamine

kinase (GNE), l'acide *N*-acétylneuraminique phosphate synthase (NANS) et la CMP-acide sialique synthétase (CMAS) a été réalisée permettant la synthèse *in planta* de quantités de CMP-Neu5Ac similaires à celles synthétisées dans les cellules CHO (**Figure 16 ; Castilho *et al.*, 2008**). En 2010 et en 2016, la même équipe a pu exprimer de façon transitoire dans le génome de *N. benthamiana*, les gènes nécessaires à la biosynthèse du CMP-Neu5Ac, à son transport du cytosol vers l'appareil de Golgi (CMP-Sialic acid Transporter) et le gène codant l' α (2,6)-sialyltransférase. L'ensemble de ces modifications réalisées dans des plantes exprimant simultanément un anticorps recombinant ont permis d'aboutir à la synthèse d'un anticorps monoclonal portant du Neu5Ac en position terminale sur ses *N*-glycannes (**Figure 15 ; Castilho *et al.*, 2010**).

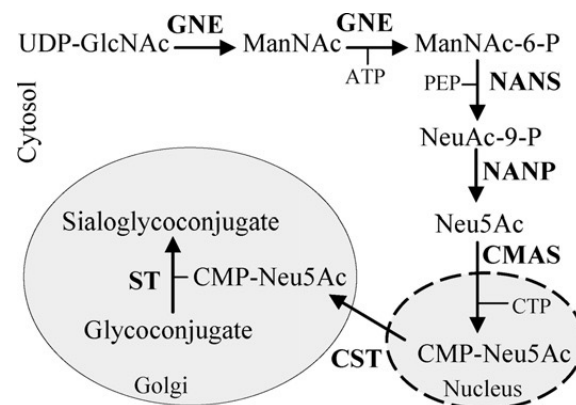


Figure 16 : Représentation de la voie de biosynthèse des acides sialiques des mammifères reconstruite chez les plantes d'*A. thaliana* par l'expression transitoire de gènes impliqués dans la biosynthèse, le transport puis le transfert du CMP-Neu5Ac (Castilho *et al.*, 2008)

L'UDP-GlcNAc est d'abord transformé en ManNAc-6-P par l'enzyme GNE. Ce sucre phosphorylé est ensuite transformé en Neu5Ac par les enzymes NANS, NANP et enfin activé sous forme de nucléotide sucre, CMP-Neu5Ac, par la CMAS. Le transport du CMP-Neu5Ac dans l'appareil de Golgi et l'action de l' α (2,6)-sialyltransférase permettent d'aboutir à la synthèse de *N*-glycannes complexes possédant des acides sialiques sur les glycoprotéines d'*A. thaliana* (**Castilho *et al.*, 2010**).

ManNAc : *N*-acétylmannosamine ; Neu5Ac : acide *N*-acétylneuraminique ; P : phosphate ; CMP : cytidine monophosphate ; GNE : UDP-GlcNAc 2-épimerase/*N*-acétylmannosamine kinase ; NANS : acide *N*-acétylneuraminique phosphate synthase ; CMAS : CMP-acide sialique synthétase ; CST : CMP-Neu5Ac transporteur ; ST : α (2,6)-sialyltransférase

3. Les anticorps monoclonaux à usage thérapeutique produits chez les plantes

A ce jour, plusieurs essais de production d'anticorps monoclonaux chez les plantes ont été conduits. Les essais se consacrent majoritairement à la production d'immunoglobines A

(IgA), d'immunoglobines M (IgM) (Vasilev *et al.*, 2016) et d'IgG (Cabanes-Macheteau *et al.*, 1999 ; Bardor *et al.*, 2003b ; Yao *et al.*, 2015). Le premier cas de production d'un anticorps monoclonal chez les plantes a été obtenu en exprimant une IgG dans des feuilles de tabac (Hiatt *et al.*, 1989). L'anticorps humanisé était fonctionnel et représentait un taux d'expression de 1,4% des protéines totales extraites des feuilles de tabac (Hiatt *et al.*, 1989). Quelques années plus tard, une IgA murine capable de neutraliser *Streptococcus mutans*, la bactérie responsable des caries dentaires, a été produite avec succès (Ma *et al.*, 1994). Cet anticorps a été synthétisé dans des plantes de tabac transgéniques qui ont produit de 200 à 500 µg d'anticorps par g de masse fraîche (Ma *et al.*, 1994). Six traitements oraux ont été réalisés sur des patients infectés par *S. mutans* avec cet anticorps produit dans les plantes de tabac. Les dents des patients ont d'abord été traitées avec de la chlorhexidine gluconate, agent antiseptique, afin d'éliminer la souche bactérienne. Ensuite, ces patients ont été traités à titre préventif avec l'anticorps recombinant produit dans le tabac en l'appliquant directement sur leurs dents. Cette application orale a permis de protéger les patients contre l'apparition de caries dentaires dues à *S. mutans* pendant au moins 4 mois sans qu'aucun effet secondaire ne soit détecté (Ma *et al.*, 1998).

En 2003, les travaux d'une collaboration anglo-américaine ont permis d'insérer les gènes codant les chaînes légère et lourde d'un anticorps humain dirigé contre le virus de la rage dans le génome de *Nicotiana tabacum* (Ko *et al.*, 2003). Ces deux gènes ont été fusionnés avec la séquence codant un signal de rétention dans le RE. Les *N*-glycannes analysés étaient majoritairement de type oligomannosidique (environ 90%) comportant des structures majoritaires de type Man₆₋₉GlcNAc₂. Cependant, les analyses de cet anticorps dirigé contre le virus de la rage ont montré la présence non négligeable de 5,7% de *N*-glycannes portant l'épitope antigénique core β(1,2)-xylose caractéristique des *N*-glycannes complexes produits dans l'appareil de Golgi (Ko *et al.*, 2003).

En 2013, un anticorps chimérique IgG/IgA a été produit chez *A. thaliana* (Nakanishi *et al.*, 2013). Cet anticorps était capable de se lier à la sous-unité B de la shiga toxine produite par la bactérie entérohémorragique *E. coli* O157 : H7, qui peut entraîner la mort du patient contaminé. Plus récemment en 2014, des anticorps de type IgG et de type IgA de trois anticorps thérapeutiques (Infliximab® et Adalumumab®) dirigés contre le facteur de la nécrose tumorale TNFα et Ustekinumab® ciblant l'interleukine 12, une cytokine synthétisée par certaines cellules immunitaires, ont été produites avec succès chez *N. benthamiana*

(Westerhof *et al.*, 2014). Les IgGs produites chez *N. benthamiana* présentaient majoritairement des *N*-glycannes de type GlcNAc₂Man₃XylFucGlcNAc₂ alors que les IgA présentaient majoritairement des *N*-glycannes dépourvus de fucose comme les GlcNAcMan₃XylGlcNAc₂, Man₃XylGlcNAc₂ et des *N*-glycannes de type oligomannosidique (Westerhof *et al.*, 2014).

Plus récemment en 2014, un cocktail médicamenteux appelé « ZMAPP » composé de trois anticorps monoclonaux chimériques capable de cibler des protéines spécifiques du virus EBOLA a été co-développé par l'entreprise américaine Mapp Biopharmaceutical, Inc. et l'entreprise canadienne Defyrus Inc (Zhang *et al.*, 2014 ; Budzianowski, 2015). Ces trois anticorps monoclonaux sont produits dans des plantes de tabac (Zhang *et al.*, 2014 ; Budzianowski, 2015). Ce cocktail médicamenteux est actuellement en phase d'essai clinique chez l'Homme depuis février 2015 (Budzianowski, 2015). Pour rappel, des anticorps dirigés contre le virus HIV avec une glycosylation humanisée ont été produits par le groupe de Castilho (Castilho *et al.*, 2010). D'autres exemples d'anticorps monoclonaux produits chez les plantes sont présentés dans le tableau suivant (Tableau 1).

Tableau 1 : Exemples de projets utilisant les plantes pour la production d'anticorps monoclonaux à intérêt thérapeutique (A) et d'anticorps monoclonaux produits chez les plantes actuellement en phase d'essai clinique (B) (Yao *et al.*, 2015)

A				
Company	Host	Lead product Expression	Technology	Advantage
Synthon, Nijmegen, The Netherlands	Duckweed Leafy Biomass	Antibody for non-Hodgkin's Lymphoma	LEX system Stable expression	Speed Quality
PlanetBiotechnology, Hayward, CA, USA	Tobacco leaves	CaroRx for dental caries; PBI-220 antibody for anthrax; DPP4-Fc for MERS coronavirus infection	Stable Expression	Quality Scale
Fraunhofer IME, Aachen, Germany	Tobacco leaves	HIV Antibody	Stable Nuclear Expression	Scale Cost
B				
Product	Host	Application	Status	Sponsor and Clinical trial
P2G12 Antibody	Tobacco	HIV	Phase 1 Completed (2011)	University of Surrey, Guildford, UK NCT01403792

V. Les microalgues, systèmes d'expression alternatifs pour la production d'anticorps monoclonaux

Les microalgues sont des organismes microscopiques photosynthétiques eucaryotes ou procaryotes. Elles se divisent en 3 classes : les microalgues brunes, rouges et vertes. Elles sont réparties dans l'ensemble des supergroupes de l'arbre phylogénétique à l'exception des Amorphea, supergroupe regroupant les animaux (**Figure 17 ; Burki *et al.*, 2020**).

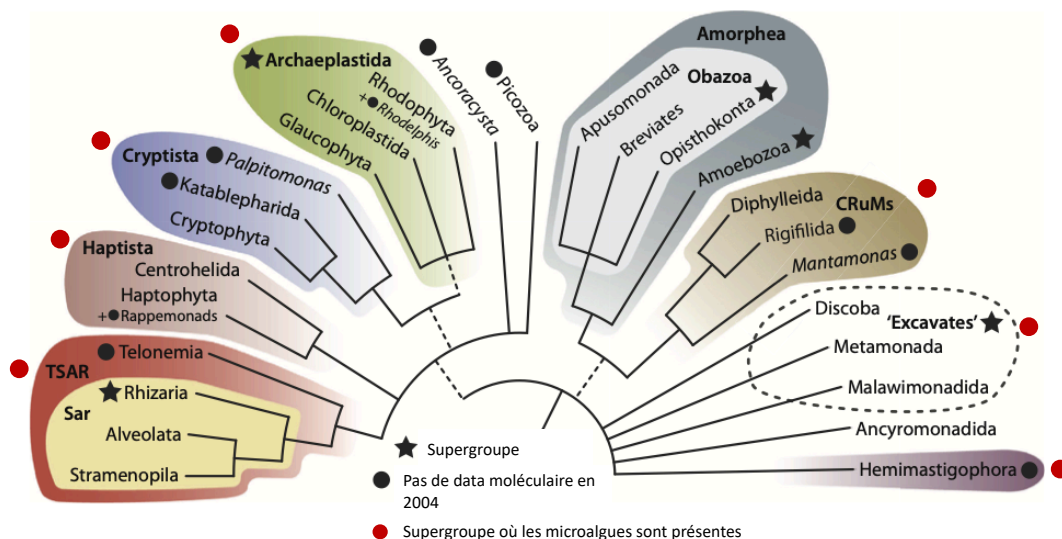


Figure 17 : Schéma représentant l'arbre phylogénétique des espèces (Burki *et al.*, 2020)

Les microalgues sont présentes dans les différents supergroupes TSAR, Haptista, Cryptista, Archaeplastida, CRuMS, Excavates et Hemimastigophora excepté dans le supergroupe Amorphea représenté par les animaux.

TSAR : telomenia-stramenopile-alveolate-rhizaria

Les microalgues intéressent depuis plusieurs années la recherche et l'industrie. Par exemple, l'industrie pétrolière s'intéresse aux microalgues riches en lipides pour les convertir en biocarburant (**Mata *et al.*, 2010 ; Halim *et al.*, 2012**). Dans ce contexte, une équipe de recherche a récemment isolée 6 nouvelles espèces de microalgues potentiellement riches en lipides. Parmi ces candidates potentielles, *Chlorococcum* serait la plus apte pour la production

de biodiesel (**Deshmukh et al., 2019**). Les microalgues sont également utilisées aujourd'hui dans l'industrie agroalimentaire comme compléments alimentaires pour l'Homme. On peut citer le cas de *spirulina* et de *Chlorella vulgaris* utilisés par plus d'une dizaine d'entreprises dans le monde pour leurs propriétés nutritives (**Tokuşoglu and Ünal, 2003 ; Chacón-Lee and González-Mariño, 2010**). La production de biomédicaments dans des microalgues est également envisagée (**Hempel and Maier, 2016**). En effet, la production de protéines recombinantes à intérêt thérapeutique chez ces organismes présente de nombreux avantages (**Barrera and Mayfield, 2013**). Tout d'abord, les microalgues possèdent un fort taux de croissance. De plus, le coût du milieu de culture en photobioréacteur est peu élevé, ce qui permettrait de réduire considérablement les coûts de production (**Barrera and Mayfield, 2013**). En effet, le milieu de culture des microalgues est estimé à environ 0,002\$ par litre de culture (**Mayfield et al., 2003 ; Rosales-Mendoza et al., 2020**). Récemment, des anticorps monoclonaux à intérêt thérapeutique et des immunotoxines ont été produits dans deux microalgues : la microalgue verte *Chlamydomonas reinhardtii* et la diatomée *Phaeodactylum tricornutum* (**Mayfield et al., 2003 ; Tran et al., 2009 ; Hempel et al., 2011 ; Hempel and Maier, 2012 ; Tran et al., 2013 ; Vanier et al., 2015 ; Barrera et al., 2015 ; Hempel et al., 2017 ; Vanier et al., 2018 ; Rosales-Mendoza et al., 2020**). Les immunotoxines sont des anticorps monoclonaux couplés à une toxine. Lorsque l'anticorps reconnaît l'antigène cible présent sur la cellule hôte, la toxine pénètre par endocytose dans la cellule provoquant ainsi la lyse de la cellule cible.

A. La microalgue verte *Chlamydomonas reinhardtii*

C. reinhardtii est une microalgue verte unicellulaire d'eau douce photosynthétique appartenant à la famille des Chlorophycées. Elle possède deux flagelles et un chloroplaste unique. Cette microalgue est utilisée depuis les années 1960 comme un organisme modèle. Ses génomes nucléaire, chloroplastique et mitochondrial ont été entièrement séquencés (**Merchant et al., 2007**). Les techniques de transformation génétique sont également fonctionnelles chez *C. reinhardtii* (**Sizova et al., 2013 ; Gao et al., 2014**). Ces techniques ont permis la création de banques de mutants pour cette microalgue (**Zhang et al., 2014 ; Li et al.,**

2016). Cet organisme est un modèle biotechnologique pour de nombreuses applications (**Merchant et al., 2012 ; Scaife et al., 2015**).

Cette microalgue modèle a été utilisée dès 2003 pour la production d'anticorps monoclonaux. Le premier anticorps monoclonal produit dans cette microalgue fut dirigé contre la glycoprotéine D du virus de l'herpès. Cet anticorps a été exprimé dans le chloroplaste de *C. reinhardtii* (**Mayfield et al., 2003**). Le second anticorps monoclonal, exprimé également dans le chloroplaste, est un anticorps dirigé contre la protéine PA 83, impliquée dans la maladie de l'anthrax, plus connue sous le nom de « maladie du charbon » (**Tran et al., 2009**). Cependant, les anticorps produits ont été obtenus par une transformation du génome chloroplastique de *C. reinhardtii* et par conséquent, ne sont pas glycosylés. A ce jour aucun anticorps glycosylé n'a été produit et caractérisé chez *C. reinhardtii*. Plus récemment, des immunotoxines ont été produites avec succès chez la microalgue verte. Ces immunotoxines sont capables de traiter respectivement le lymphome et le botulinum (**Tran et al., 2013 ; Barrera et al., 2015**).

B. La diatomée *Phaeodactylum tricornutum*

P. tricornutum est une microalgue eucaryote appartenant à la classe des diatomées pennées (*Bacillariophyceae*). Cette microalgue est atypique, de par sa présence dans le milieu naturel sous 3 morphotypes principaux : ovale, fusiforme et triradié (**De Martino et al., 2007 ; Ovide et al., 2018 ; Figure 18**).

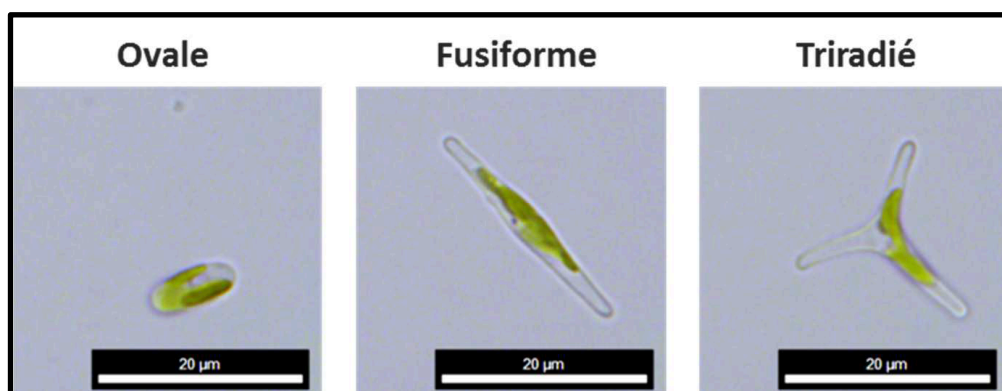


Figure 18 : Photos en microscopie optique des trois morphotypes principaux de *P. tricornutum* (d'après Ovide et al., 2018)

Les cellules de *P. tricornutum* peuvent être ovales, fusiformes ou triradiées. Les cellules mesurent en moyenne 20 µm de long pour les morphotypes fusiformes et triradiés et environ 10 µm pour les cellules ovales.

Un quatrième morphotype a pu être mis en évidence appelé cruciforme (**He et al., 2014**). La composition pariétale de *P. tricornutum* est majoritairement constituée de glucuromannanes sulfatés (**Abdullahi et al., 2006 ; Gügi et al., 2015 ; Le Costaouëc et al., 2017**). Cette composition pariétale a déjà été observée chez d'autres diatomées (**Hoagland et al., 1993 ; Chiovitti et al., 2005 ; Le Costaouëc et al., 2017**). De plus, les diatomées ont la particularité, contrairement aux autres microalgues, de posséder une paroi cellulaire appelée frustule composée de silice. Cependant, chez *P. tricornutum*, seul le morphotype ovale possède une frustule silicatée (**Tesson et al., 2009**).

Le génome de *P. tricornutum* a été entièrement séquencé (**Bowler et al., 2008**). Des données de transcriptomique sont également disponibles (**Cruz de Carvalho et al., 2016 ; Rastogi et al., 2018 ; Remmers et al., 2018 ; Ovide et al., 2018**). De plus, les vecteurs d'expression disponibles, les techniques de transformation génétique (**Zaslavskaja et al., 2000 ; Radakovits et al., 2011 ; Karas et al., 2015**) et les dernières techniques de génie génétique (CRISPR-cas9, TALEn) (**Daboussi et al., 2014 ; Nymark et al., 2016 ; Serif et al., 2018 ; Stukenberg et al., 2018**) font de cette diatomée, un organisme de choix pour la communauté scientifique. De plus, il a été démontré également que *P. tricornutum* était capable de produire des anticorps monoclonaux à intérêt thérapeutique. Par exemple, un anticorps dirigé contre le virus de l'hépatite B et portant des *N*-glycannes oligomannosidiques a pu être produit chez la diatomée (**Hempel et al., 2011 ; Hempel and Maier, 2012 ; Vanier et al., 2015**). Cet anticorps est fonctionnel et capable de se lier efficacement aux récepteurs Fcγ présents sur des monocytes (**Vanier et al., 2018**). Plus récemment, différents variants d'anticorps fonctionnels dirigés contre le virus de Marburg ont également été obtenus à partir de cellules de *P. tricornutum* (**Hempel et al., 2017**).

VI. La *N*-glycosylation des microalgues et les techniques d'ingénierie

L'émergence des microalgues pour la production de protéines recombinantes à intérêt thérapeutique a amené la communauté scientifique à s'intéresser à leurs processus de *N*-glycosylation. A l'heure actuelle, seules quelques microalgues sont étudiées. Le chapitre de livre ci-dessous, publié au cours de ma thèse, résume les données sur les voies de la *N*-glycosylation étudiées chez les microalgues ainsi que les techniques d'ingénierie développées permettant d'envisager à l'avenir la modification de leurs processus de *N*-glycosylation pour la production de glycoprotéines à usage thérapeutique comme les IgG.

Toward Future Engineering of the *N*-Glycosylation Pathways in Microalgae for Optimizing the Production of Biopharmaceuticals

Rodolphe Dumontier, Alain Mareck,
Narimane Mati-Baouche, Patrice Lerouge and
Muriel Bardor

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.73401>

Abstract

Microalgae are eukaryotic and photosynthetic organisms which are commonly used in biotechnology to produce high added value molecules. Recently, biopharmaceuticals such as monoclonal antibodies have been successfully produced in microalgae such as *Chlamydomonas reinhardtii* and *Phaeodactylum tricorutum*. Most of these recombinant proteins are indeed glycosylated proteins, and it is well established that their glycan structures are essential for the bioactivity of the biopharmaceuticals. Therefore, prior to any commercial usage of such algae-made biopharmaceuticals, it is necessary to characterize their glycan structures and erase glycosylation differences that may occur in comparison with their human counterpart. In this context, the chapter summarizes successful attempts to produce biopharmaceuticals in microalgae and underlines current information regarding glycosylation pathways in microalgae. Finally, genome editing strategies that would be essential in the future to optimize the microalgae glycosylation pathways are highlighted.

Keywords: antibodies, biopharmaceuticals, genome editing tools, glycosylation, microalgae

1. Introduction

Microalgae are currently used for a broad spectrum of industrial applications including food and livestock feed industries, bioenergy, cosmetics, healthcare and environment [1–4]. Recently, due to their numerous advantages (high growth rate, easy cultivation, low production cost, etc.), microalgae have emerged as a solar-fueled green alternative cell factories for the production

of recombinant proteins [4–7]. Among different attempts to produce vaccines and biopharmaceuticals in microalgae, the production of monoclonal antibodies (mAbs) represents the most extensive work [7, 8]. Indeed, the first significant effort to produce recombinant mAb fragments was made in the green microalga *Chlamydomonas reinhardtii* with the synthesis and accumulation in its chloroplast of a human single chain antibody directed against the herpes simplex virus glycoprotein D (HSV8-lsc) [8]. Later, a full-length human IgG1 directed against anthrax was produced successfully in the chloroplast of *C. reinhardtii* [9]. The *Chlamydomonas*-made mAb was able to bind the anthrax protective antigen 83 (PA83) [9]. In another study, a series of complex chimeric proteins was expressed in the chloroplast of *C. reinhardtii*. Such chimeric proteins were composed of a single chain antibody fragment (*scFv*) targeting the B-cell surface antigen CD22, genetically fused either to the eukaryotic ribosome inactivating protein, gelonin, from *Gelonium multiflorum* [10] or to *Pseudomonas aeruginosa* exotoxin A domains 2 and 3 [11]. These molecules, termed immunotoxins, were encoded by a single gene that produces an antibody-toxin chimeric protein. Such algae-made immunotoxins are able to bind target B cells and efficiently kill them *in vitro* [11]. Full-length mAbs have also been expressed in the diatom *Phaeodactylum tricornutum* through nuclear transformation [12–14]. Those mAbs correspond respectively to a recombinant mAb directed against the nucleoprotein of Marburg virus, a close relative of Ebola virus [14] and to a human IgG1 directed against the Hepatitis B virus Antigen (HBsAg) [12, 13]. The latter has been biochemically characterized in order to check the quality of the diatom-made mAb as well as its *N*-glycosylation profile [15]. Moreover, it has been demonstrated that this glycosylated antibody is able to bind human Fcγ receptors [16], thus suggesting that it could be efficient in human therapy.

When the production of biopharmaceuticals is considered, their *N*-glycosylation has to be investigated. Indeed, among the biopharmaceuticals that were approved in 2016 and 2017, 96% were glycosylated [17]. The glycosylation of the approved biopharmaceutical represents a critical quality attribute (CQA) that may affect its safety and biological activities [18–20]. In addition, introduction by the expression system of nonhuman epitopes on the recombinant protein may induce immune response after injection to patients [21]. Thus, the *N*-glycosylation of biopharmaceuticals is a real challenge for the commercial production of biopharmaceuticals. The glycosylation state of therapeutic proteins has to be accurately identified and characterized as per the World Health Organization and International Conference on Harmonization Q6B guidelines [17]. Therefore, in the context of developing the microalgae as alternative platforms for the production of biopharmaceuticals, the capability of these unicellular eukaryotic cells to introduce *N*-glycans on their endogenous proteins and on recombinant proteins, as well as their regulation, have to be considered and understood.

2. *N*-glycosylation in microalgae

2.1. General aspects

N-glycosylation is a major post-translational modification of proteins in eukaryotes. Protein *N*-glycosylation first starts by the synthesis of a lipid-linked oligosaccharide formed by transfer of monosaccharides on a dolichol pyrophosphate (PP-Dol) anchored in the membrane of

the endoplasmic reticulum (ER) *via* the action of a set of enzymes named asparagine-linked glycosylation (ALG) [22, 23]. The final $\text{Glc}_3\text{Man}_5\text{GlcNAc}_2$ precursor is transferred *en bloc* by the oligosaccharyltransferase (OST) complex onto the asparagine residues of the consensus Asn-X-Ser/Thr sequences of a protein [22] (Figure 1). Alternative consensus sequences, such as Asn-X-Cys and Asn-X-Val, have also been found to be glycosylated in some proteins [24–26].

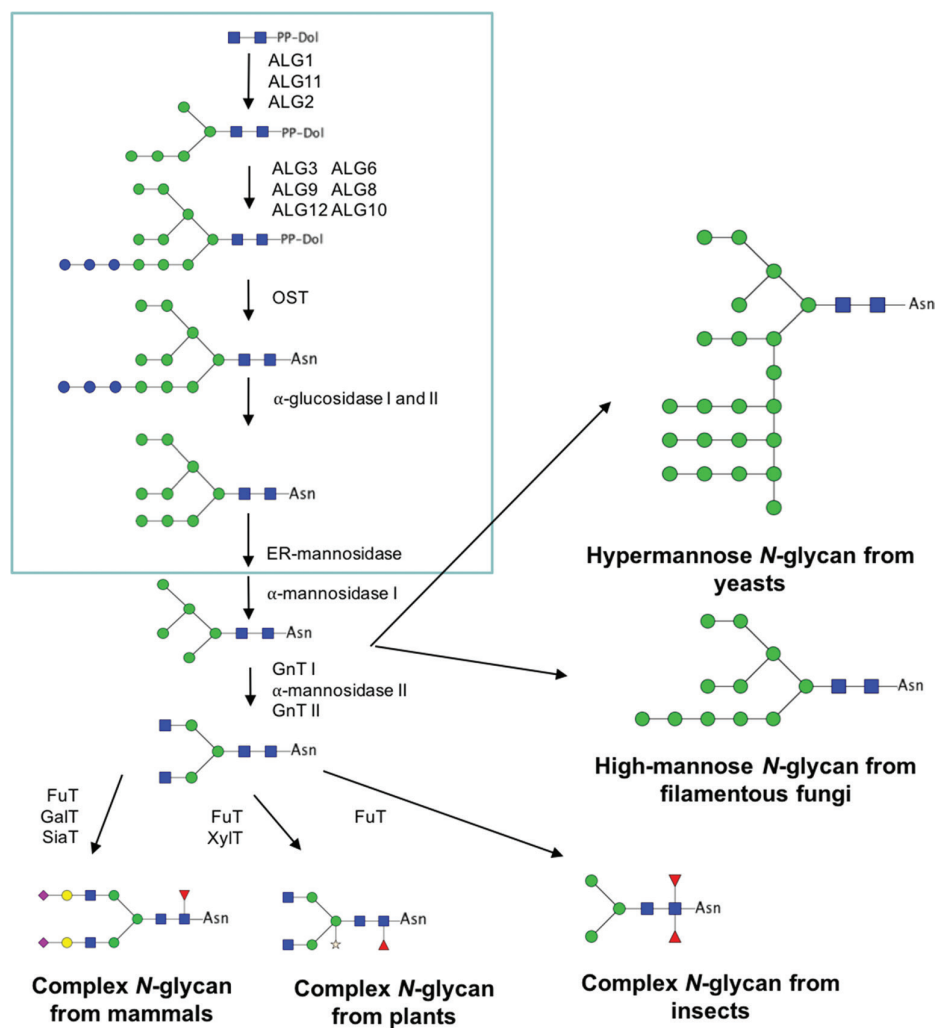


Figure 1. Comparison of protein *N*-glycosylation pathways in eukaryotes. Biosynthesis steps occurring in the ER are gathered in the box. Mature *N*-glycan structures observed in mammals, plants, insects, yeasts and filamentous fungi are drawn according to [33]. ■, N-acetylglucosamine; ☆, xylose; ●, mannose; ▲, fucose; ●, galactose; ◆, sialic acid; Asn, asparagine; PP-Dol, pyroPhosphate dolichol; FuT, fucosyltransferase; GalT, galactosyltransferase; SiaT, sialyltransferase; XylT, xylosyltransferase; ALG, asparagine-linked glycosylation; OST, oligosaccharyltransferase.

In the ER, neo-synthesized glycoproteins are then submitted to a quality control process through the deglycosylation by glucosidases and reglycosylation by an UDP-glucose: glycoprotein glucosyltransferase (UGGT) of the *N*-glycans. This allows the synthesis of monoglucosylated glycan intermediates that interact with ER-resident chaperones, thus ensuring proper folding of the glycoproteins [27]. When the glycoprotein is correctly folded, α -glucosidase II would finally remove the last glucose residue, and ER-mannosidase will eventually remove one mannose residue that leads to the formation of an oligomannoside $\text{Man}_{9/8}\text{GlcNAc}_2$. The quality control events are conserved in eukaryotes because they are crucial for the secretion of well-folded proteins [28]. As a consequence, whatever the expression system used, a recombinant therapeutic protein leaving the ER compartment exhibits a *N*-glycosylation similar to one of the reference proteins with unique oligomannoside $\text{Man}_{9/8}\text{GlcNAc}_2$ attached to the Asn residue of the *N*-glycosylation consensus site.

After transfer to the Golgi apparatus, oligomannosides resulting from the ER processing are modified by the action of specific mannosidases and glycosyltransferases [29]. These Golgi cell-specific repertoires give rise to various organism-specific oligosaccharides. In most eukaryotes, a *N*-acetylglucosaminyltransferase I (GnT I)-dependent *N*-glycan processing occurs (**Figure 1**). In this pathway, the α -mannosidase I converts $\text{Man}_{9/8}\text{GlcNAc}_2$ into the branched isomer of $\text{Man}_5\text{GlcNAc}_2$. Then, actions of GnT I, α -mannosidase II and GnT II, respectively, give rise to the core $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ that is common to most eukaryotes [27–31] (**Figure 1**). This core is then decorated by the action of specific glycosyltransferases that differ from one organism to another. This allows the protein to be decorated by organism-specific *N*-glycans that confer to the mature protein *in vivo* bioactivities [32]. It is worth noting that GnT I-independent *N*-glycan processing also occurs in some eukaryotes such as filamentous fungi and yeasts in which *N*-glycosylation in the Golgi apparatus results in the synthesis of high mannose and hypermannose *N*-glycans, respectively (**Figure 1**). As a consequence, in the context of the production of biopharmaceuticals by genetic engineering, such a diversity of mature *N*-linked glycans is a limitation because the expression system used may introduce inappropriate epitopes and heterogeneous glycosylation on the therapeutics and may also fail in introducing glycan sequences that are required for *in vivo* bioactivity of the biopharmaceuticals.

2.2. Protein *N*-glycosylation in microalgae

Overall, protein *N*-glycosylation in microalgae received little attention. Few studies, published in the 1990s have demonstrated that proteins secreted by green microalgae carry mainly oligomannosides or xylose-containing *N*-glycans based on affinodetection or enzymatic sequencing [34–36]. More recently, analysis by mass spectrometry of glycans *N*-linked to microalgae endogenous proteins has been reported. First, the 66 kDa cell wall glycoprotein from the red microalga *Porphyridium* sp. has been found to carry $\text{Man}_8\text{GlcNAc}_2$ and $\text{Man}_3\text{GlcNAc}_2$ oligomannosides containing 6-*O*-methyl mannose residues and substituted by one or two xylose residues [37, 38] (**Figure 2**). Investigation of *C. reinhardtii* has demonstrated that proteins in this green microalga carry oligomannosides ranging from $\text{Man}_2\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$ as well as $\text{Man}_4\text{GlcNAc}_2$ and $\text{Man}_3\text{GlcNAc}_2$ *N*-glycans containing 6-*O*-methyl mannoses and substituted by one or two xylose residues (**Figure 2**) [39]. Initially reported as branched

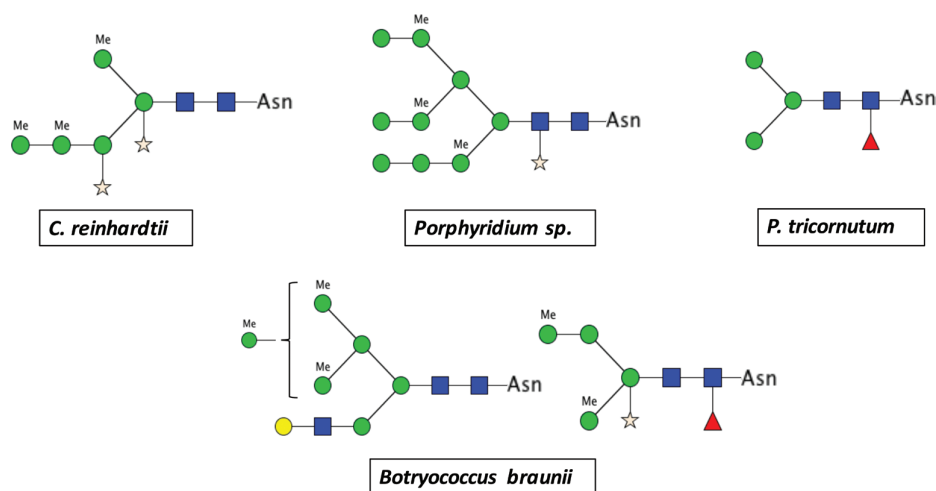


Figure 2. Major mature *N*-linked glycans from the green microalga *Chlamydomonas reinhardtii* and *Botryococcus braunii*, the red microalga *Porphyridium* sp. and the diatom *Phaeodactylum tricornutum*. *N*-glycan structures are drawn according to [33]. ■, *N*-acetylglucosamine; ☆, xylose; ●, mannose; ▲, fucose; ●, galactose; Asn, asparagine; Me, methyl.

oligomannosides, the structure of $\text{Man}_3\text{GlcNAc}_2$ was re-evaluated in 2017 as being linear sequences based on ESI-MSⁿ analyses [40]. Although mature *N*-glycans from *Porphyridium* sp. and *C. reinhardtii* share common structural features, the location of the xylose residues on the *N*-glycan differs between these two microalgae (Figure 2). As mature *N*-glycans do not exhibit any terminal GlcNAc residues, they were proposed to result from Golgi xylosylation and *O*-methylation of oligomannosides deriving from the precursor synthesized in the ER in a GnT I-independent processing, even if this needs to be completely elucidated and that methylation occurring in the ER cannot be ruled out yet [38].

N-glycan profile from *P. tricornutum* has been described to contain $\text{Man}_3\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$ oligomannosides and also minute amount of paucimannosidic fucosylated *N*-glycans (Figure 2) [41]. In contrast to *Porphyridium* sp. and *C. reinhardtii*, these *N*-glycans result from a GnT I-dependent pathway (Figure 2) [41]. As evidence, GnT I gene predicted in the *P. tricornutum* genome encodes an enzyme able to restore the maturation of complex-type *N*-glycans in the CHO Lec1 mutant that lacks endogenous GnT I activity [41]. *N*-glycans arising from a GnT I-dependent pathway have also been recently reported in the green microalga *Botryococcus braunii* through a glycoproteomic approach [42]. In contrast to *P. tricornutum*, these *N*-glycans harbor a GlcNAc residue at the nonreducing end as well as mono- and di-*O*-methylations of the core mannose residue. Moreover, this *N*-glycan bearing a terminal GlcNAc resulting from the GnT I activity could be further elongated with an additional hexose or methyl-hexose residue. In addition, proteins from this green microalga also exhibit methylated *N*-linked oligomannosides carrying core fucose and core xylose residues (Figure 2) [42].

In support to these biochemical data, protein *N*-glycosylation in microalgae can be drawn on the basis of public genomic databases. Microalgae genomes from different phyla are available

to date (<https://genome.jgi.doe.gov/pages/tree-of-life.jsf>) [4, 43]. Since protein *N*-glycosylation occurs in the ER and the Golgi apparatus, bioinformatics analyses of microalgae genomes must be investigated independently for the two compartments: search for gene encoding proteins involved in the precursor biosynthesis and the ER protein quality control on the one hand, and search for Golgi glycosidases and glycosyltransferases involved in the synthesis of mature *N*-glycans on the other hand.

Genes encoding subunits of OST, glucosidases, as well as ER-resident UGGT and chaperones are predicted in microalgae genomes suggesting that the process of ER quality control in these unicellular organisms is similar to the one described in other eukaryotic cells [41, 44, 45]. Among these putative ER candidates, only the activity of the $\alpha(1,3)$ -glucosidase, also called glucosidase II, from the red microalga *Porphyridium* sp. has been biochemically confirmed [44]. Most ALG genes are also predicted in microalgae genomes [39, 41, 44] suggesting that the synthesis of the oligosaccharide precursor is overall conserved. However, some of these ALG, that is ALG3, ALG9 and ALG12, are not predicted in *C. reinhardtii* [39, 45]. These ER enzymes are involved in the completion of the biosynthesis of the precursor $\text{Man}_3\text{GlcNAc}_2\text{-PP-Dol}$, prior to its glucosylation, by addition of mannose residues on the $\alpha(1,6)$ -mannose arm of the core (**Figure 1**). Reinvestigation in *C. reinhardtii* of the structure of oligomannosides and analysis of the ER *N*-glycan precursor [40] confirmed the absence of ALG3, ALG9 and ALG12 activities and the synthesis in this green microalga of linear oligomannoside sequences instead of branched isomer initially proposed in [39]. It is worth noting that in this truncated ER pathway, the presence of the triglycosyl extension is likely sufficient to ensure interaction of the *N*-glycan precursor with chaperones of the ER quality control process. In addition to the lack of the ALG3, ALG9 and ALG12 in *C. reinhardtii*, other microalgae genomes lack genes encoding ALG10 and GCS1, an $\alpha(1,2)$ -glucosidase [44]. Because ALG10 is the $\alpha(1,2)$ -glucosyltransferase responsible for the addition of the $\alpha(1,2)$ -glucose residue on the precursor *N*-glycan and GCS1 is responsible for trimming this residue, we hypothesize that the ER quality control in these microalgae involved only diglycosylated *N*-glycan intermediates.

With regard to Golgi *N*-glycosylation events, the presence of GnT I is predicted in some microalgae including haptophytes and cryptophytes, but not in *C. reinhardtii*, *Volvox* and *Ostreococcus* [41, 42]. As mentioned previously, *P. tricornutum* GnT I activity was confirmed by the complementation of CHO Lec 1 mutant cell line [41]. A recent study of *B. braunii* [42] confirmed the involvement of this transferase in this green microalga *N*-glycosylation pathway. Concerning other Golgi enzymes, α -mannosidases (CAZy GH 47) and $\alpha(1,3)$ -fucosyltransferases (CAZy GT10) are also predicted in microalgae genomes studied so far [41, 44, 45]. These enzymes are respectively involved in the trimming of mannose residue of oligomannosides and the transfer of fucose on the proximal GlcNAc. These sequences exhibit peptide motifs that were demonstrated to be required for activities of such Golgi enzymes, but, in contrast to GnT I, no biochemical data of their activity and specificity are available yet.

As depicted, protein *N*-glycosylation occurring in microalgae is specific and largely differs from the one described in mammals (**Figures 1 and 2**). Therefore, production in microalgae of biopharmaceuticals exhibiting *N*-glycans compatible with their use in human therapy

would be challenging and requiring metabolic engineering of the *N*-glycosylation pathway in microalgae. This will include the inactivation of enzymes that introduce nonhuman glyco-epitopes onto *N*-linked glycans and complementation of microalgae with appropriate glycosyltransferases to introduce missing glycan sequences. These strategies have already been successfully carried out for the engineering of the *N*-glycan pathways in plants and yeasts [46, 47]. In addition, the success of the complementation with human glycosyltransferases requires the availability in the Golgi apparatus of appropriate nucleotide-activated sugars [48]. For instance, sialic acids that terminate bi-antennary *N*-glycans in mammals have not been reported in microalgae such as *P. tricornutum* and *Porphyridium* sp. [38, 41]. As well, there is no evidence for the import of GlcNAc in the Golgi apparatus in microalgae exhibiting a GnT I-independent *N*-glycan pathway, even if putative candidates for UDP-GlcNAc transporter have been identified in microalgae such as *C. reinhardtii* [49]. Indeed, the two GlcNAc of the chitobiose unit of *N*-linked glycans are transferred onto the PP-Dol lipid in the cytosolic face of the ER membrane. Currently, metabolic engineering strategies are now feasible due to the recent development of transgene expression and gene inactivation in microalgae as summarized in Section 3.

3. Genetic engineering tools now available to envision future *N*-glycosylation engineering in microalgae

3.1. Different tools to generate genome-modified organisms

Classical strategies of genetic engineering involve the modulation of gene expression including overexpression and inactivation by RNA interference [50–52]. The most used engineering methods are based on random insertional mutagenesis obtained by various processes such as conjugation, agitation with glass beads, electroporation, biolistic microparticle bombardment, agrobacterium-mediated transformation or multipulse electroporation. The transformation step is followed by phenotypic selection using antibiotics to generate genome-modified organisms [53]. Those processes present the advantage to be simple and reach a high level of transformed cells. For example, *P. tricornutum* transformation reached 1 per 10⁶ cells with biolistic bombardment system [54]. However, cell-wall-less strains are required for almost all the classical methods quoted above [50, 55]. Furthermore, genetic stability of the mutagenesis obtained after transformation by random insertion depends on microalgae species [53]. For example, a high mutagenesis stability has been shown in *C. reinhardtii* [55]. Unlike, mutagenesis was unstable in *Thalassiosira weissflogii* [56]. More recently, new tools have been developed in order to knock in, knock out, modify, replace, or insert genes. These new genetic engineering tools consist of the action of nucleases effecting their molecular scissor activities in specific loci [52]. A break in the DNA causes activation of DNA repair mechanisms, which can be either the homologous-recombination (HR) or the non-homologous end-joining (NHEJ) [52]. The HR results in sequence modification in the target locus [57]. In the NHEJ process, the two ends of the broken chromosome are stuck together causing small deletions or small insertions [57]. These events confer several modifications of the target gene such as gene inactivations or

insertions. Very little is known about these mechanisms in microalgae due to their complexity as reported by Daboussi in 2017 [53].

Several researches have recently contributed to demonstrate that particular nucleases could be used for targeting stable modifications by acting like molecular scissors. Among these nucleases, we can quote meganucleases (MNs), zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and finally, the famous clustered regularly interspaced short palindromic repeats (CRISPR)/nuclease Cas9 system. These four cited nucleases are described in the following paragraphs.

Meganuclease is an engineered endonuclease able to recognize and cleave a long specific DNA sequence from 18 to 30 base pairs. The meganuclease strategy requires to design a homing endonuclease from the LAGLIDADG family especially the I-CreI enzyme from *C. reinhardtii* implied in the targeting of interesting gene sequences that need to be modified [58]. This was tested for the first time in 2014 using *P. tricornutum* as a model [59]. In this study, two engineered meganucleases targeting genes involved in the lipid metabolism are allowed to obtain 29% of targeted mutagenesis [59]. Even successful, this strategy is time-consuming as compared to the other alternatives [52].

Zinc finger nucleases (ZFNs) are hybrid proteins composed of a restriction enzyme *FokI* with a designed zinc-finger DNA-binding domains [60]. These *FokI* enzymes are inactive in a homodimer conformation [61]. Therefore, cleavage of a typical DNA-target sequence requires to design two different ZFNs for binding to adjacent half-sites of a specific locus. Each designed ZFN is able to recognize a sequence of 9–12 nucleotides in the genome [52]. A set of zinc finger nucleases has been recently used to modify by insertion of template DNA, the *Cop3* gene locus encoding a light-activated channel in *C. reinhardtii* [62]. Moreover, in 2017, the genome editing was reliably performed using the ZFN strategy in order to inhibit and modify nuclear photoreceptor genes in this same microalga [63]. Despite these promising results, the ZFN system is barely used because of its low specificity. Indeed, cleavage of DNA requires both ZFN monomers to recognize a homologous target in the genome in the proper spatial orientation to assemble a functional ZFN [64]. Also, ZFN system is time-consuming implementation [64]. Nowadays, other designed nucleases like TALENs or CRISPR/cas9 are emerging in the scientific community to perform genome editing in microalgae.

Transcriptional activator-like effector nuclease (TALEN) system is similar to ZFN because it uses nucleases composed of a restriction enzyme domain fused to a DNA-binding domain (here the TAL effector domain) and a nonspecific DNA cleavage domain *FokI* [65]. TALEN proteins are characterized by a repeated 34-amino acid sequence that recognizes specific DNA sequences [66]. *P. tricornutum* lipid metabolism was recently modified using TALEN [59]. In this study, seven genes involved in this metabolism were modified. Each genome modification had a high frequency reaching up to 56% of colonies with targeted mutagenesis [59]. This genetic engineering allowed creation of a high lipid-producing strain by inactivating a key gene for carbohydrate energy storage [59]. Another team has inactivated successfully the urease gene in *P. tricornutum* with 24% of transformed colonies [67]. In addition, TALEN system has also been used in order to inactivate red/far-red light-sensing *phytochrome* gene of this diatom [68].

The clustered regularly interspaced short palindromic repeats (CRISPR)/cas9 system is the most famous engineered nuclease system of this decade because it is a powerful and precise tool applied in numerous eukaryotic organisms [69]. This system is based on the RNA-guided DNA cleavage defense system from archaea and many bacteria. Indeed, these organisms are able to store bacteriophage DNA fragments along a previous bacteriophage infection in the CRISPR locus, which is formed of DNA repeat sequences spaced by a unique DNA sequence. This system establishes the basis of a bacterial defense as a response to bacteriophage attacks [70]. This defense mechanism has been highlighted for the first time by Pr Emmanuelle Charpentier and her team in 2011 [70, 71]. The CRISPR/Cas9 system has been developed into a simple toolkit based on a custom single guide RNA (sgRNA) that contains a targeting sequence (crRNA sequence) and a cas9 nuclease-recruiting sequence (tracrRNA) [52]. In microalgae, CRISPR/cas9 has been used in *C. reinhardtii* [72]. However, the Cas9 nuclease production seemed to be toxic for the microalga limiting efficiency to obtain genome-modified strains [72]. Two years later, a new assay has been performed in this same microalga using another strategy avoiding toxicity [73]. Indeed, the authors succeeded to generate CRISPR/cas9-induced NHEJ-mediated knock-in mutant strains in three loci [73]. In the same year, CRISPR/cas9 gene knockout technology has been used in *P. tricornutum* to induce mutant for the *CpSRP43* gene, a member of the chloroplast signal recognition particle pathway. Using

	MN system	ZFN system	TALEN system	CRISPR/cas9
Actor(s) of gene targeting	Chimeric endonuclease	Chimeric endonuclease	Chimeric endonuclease	RNA guide and cas9 nuclease
Engineered protein origin	<i>Chlamydomonas reinhardtii</i> <i>Saccharomyces cerevisiae</i> [59, 77]	Xanthomonas [78]	Xanthomonas [78]	Bacteria and Archaea [70]
Nuclease specificity	Low	Moderate	Moderate	High
Mutagenesis frequency in microalgae	Up to 29% [59]	Not reported	Up to 56% [59]	Up to 63% [74]
Toxicity in cells	Low	Moderate ¹	Moderate ¹	Moderate [72]
Time investment	Very high [52]	Very high [64]	Moderate	Low
Possibility of multiple gene targeting	No	No	Yes	Yes
System cost ²	Not reported	Expensive (4000–7000\$)	Expensive (3000–5000\$)	Cheap (500\$)

¹Source: <https://www.news-medical.net/life-sciences/How-Does-CRISPR-Compare-to-Other-Gene-Editing-Techniques.aspx> visited [Accessed: 2017-12-06].

²Source: <http://www.biocompare.com/Editorial-Articles/144186-Genome-Editing-with-CRISPRs-TALENs-and-ZFNs/> [Accessed: 2017-12-06].

Table 1. Comparison of four specific genomic tools based on nuclease systems in order to generate genomic-modified species in microalgae.

this strategy, the authors obtained 31% of mutation efficiency [74]. This team targeted two other genes of the diatom using this technology and obtained from 25 to 63% of mutation level [74]. Adaptability of the CRISPR/Cas9 system has been demonstrated in other diatoms like *Thalassiosira pseudonana* [75] as well as in the heterokont, *Nannochloropsis oceanica* in order to knock out the nitrate reductase activity [76]. In conclusion, CRISPR/cas9 system is a promising technology to generate genome-modified organisms in microalgae. **Table 1** compares this system with the other nuclease systems cited above in terms of their technical characteristics and highlights their advantages and disadvantages.

3.2. Mutant libraries

The study of mutants impaired in a glycosidase or a glycosyltransferase implied in the *N*-glycan pathway is of great interest. Indeed, the synthesis of oligosaccharides is a sequential process. Inactivation of an enzyme usually results in the accumulation of its *N*-glycan substrate which enables the step-by-step dissection of the entire pathway. Moreover, mutant phenotyping of the glycosylation pathway allows to investigate to which extent the protein *N*-glycan processing is required for normal growth and development. An indexed and mapped mutant library has been created in *C. reinhardtii* by single random insertional mutagenesis of gene cassettes in 2016 [79]. This library already envisioned to study the function of genes encoding putative glycosyltransferases, glycosidases or even putative translocators in microalgae and to confirm their physiological role from reverse genetic studies.

4. Conclusion

The production of biopharmaceuticals in microalgae currently requires a better understanding of the *N*-glycosylation pathway mechanism and regulation. Such information can be gained by the use of mutant libraries like the one recently developed for *C. reinhardtii*. Indeed, characterization of each individual mutants will allow an understanding of a specific step of the *N*-glycan processing, and mutant cells could represent interesting cell lines for the production of biopharmaceuticals bearing a chosen *N*-glycan profiling.

Once these pathways would be completely deciphered in the microalgae model intended to be used for the production of biopharmaceuticals, the humanization of the *N*-glycosylation pathway could be initiated using designed engineered nucleases strategies recently developed in microalgae. We can now consider that transformed microalgae by these innovative new genomic tools will constitute in a near future one of the most suitable green cell factories for the production of humanized biopharmaceuticals.

Acknowledgements

The authors are indebted to all coworkers and students at the Glyco-MEV laboratory who are currently contributing to the microalgae research project or did so in the past. They are also

thankful to the University of Rouen Normandy, the region Normandie and the I.U.F. for their financial support.

Conflict of interest

The authors have declared no conflict of interest.

Author details

Rodolphe Dumontier¹, Alain Mareck¹, Narimane Mati-Baouche¹, Patrice Lerouge¹ and Muriel Bardor^{1,2*}

*Address all correspondence to: muriel.bardor@univ-rouen.fr

1 Normandie Univ, UniRouen, Laboratoire de Glycobiologie et Matrice Extracellulaire Végétale, Rouen, France

2 I.U.F. (Institut Universitaire de France), Paris Cedex 05, France

References

- [1] Sasso S, Pohnert G, Lohr M, Mittag M, Hertweck C. Microalgae in the postgenomic era: A blooming reservoir for new natural products. *FEMS Microbiology Reviews*. Jul 1, 2012;**36**(4):761-785
- [2] Spolaore P, Joannis-Cassan C, Duran E, Isambert A. Commercial applications of microalgae. *Journal of Bioscience and Bioengineering*. Feb 1, 2006;**101**(2):87-96
- [3] Mata TM, Martins AA, Caetano NS. Microalgae for biodiesel production and other applications: A review. *Renewable and Sustainable Energy Reviews*. Jan 1, 2010;**14**(1):217-232
- [4] Cadoret J-P, Garnier M, Saint-Jean B. Microalgae, functional genomics and biotechnology. *Advances in Botanical Research*. Jan 1, 2012;**64**:285-341
- [5] León-Bañares R, González-Ballester D, Galván A, Fernández E. Transgenic microalgae as green cell-factories. *Trends in Biotechnology*. Jan 1, 2004;**22**(1):45-52
- [6] Barrera DJ, Mayfield SP. High-value recombinant protein production in microalgae. In: Emeritus ARPD, Hu Q, editors. *Handbook of Microalgal Culture* [Internet]. Oxford, UK: John Wiley & Sons, Ltd; 2013. pp. 532-544. Available from: <http://onlinelibrary.wiley.com/doi/10.1002/9781118567166.ch27/summary>
- [7] Hempel F, Maier UG. Microalgae as solar-powered protein factories. *Advances in Experimental Medicine and Biology*. 2016;**896**:241-262

- [8] Rosales-Mendoza S. Algae-Made Antibodies and Immunotoxins. In: *Algae-Based Biopharmaceuticals* [Internet]. Springer, Cham; 2016 [cited 2017 Nov 17]. pp. 77-93. Available from: https://link.springer.com/chapter/10.1007/978-3-319-32232-2_5
- [9] Tran M, Zhou B, Pettersson PL, Gonzalez MJ, Mayfield SP. Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. *Biotechnology and Bioengineering*. Nov 1, 2009;**104**(4):663-673
- [10] Tran M, Henry RE, Siefker D, Van C, Newkirk G, Kim J, et al. Production of anti-cancer immunotoxins in algae: Ribosome inactivating proteins as fusion partners. *Biotechnology and Bioengineering*. Nov 1, 2013;**110**(11):2826-2835
- [11] Tran M, Van C, Barrera DJ, Pettersson PL, Peinado CD, Bui J, et al. Production of unique immunotoxin cancer therapeutics in algal chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America*. Jan 2, 2013;**110**(1):E15-E22
- [12] Hempel F, Lau J, Klingl A, Maier UG. Algae as protein factories: Expression of a human antibody and the respective antigen in the diatom *Phaeodactylum tricornutum*. *PLoS One*. Dec 2, 2011;**6**(12):e28424
- [13] Hempel F, Maier UG. An engineered diatom acting like a plasma cell secreting human IgG antibodies with high efficiency. *Microbial Cell Factories*. Sep 13, 2012;**11**:126
- [14] Hempel F, Maurer M, Brockmann B, Mayer C, Biedenkopf N, Kelterbaum A, et al. From hybridomas to a robust microalgal-based production platform: Molecular design of a diatom secreting monoclonal antibodies directed against the Marburg virus nucleoprotein. *Microbial Cell Factories*. Jul 27, 2017;**16**:131
- [15] Vanier G, Hempel F, Chan P, Rodamer M, Vaudry D, Maier UG, et al. Biochemical characterization of human anti-hepatitis B monoclonal antibody produced in the microalgae *Phaeodactylum tricornutum*. *PLoS One*. 2015;**10**(10):e0139282
- [16] Vanier G, Stelter S, Vanier J, Hempel F, Maier UG, Lerouge P, et al. Alga-made anti-hepatitis B antibody binds to human Fc γ receptors. *Biotechnology Journal*. 1700496. DOI: 10.1002/biot.201700496
- [17] O'Flaherty R, Trbojević-Akmačić I, Greville G, Rudd PM, Lauc G. The sweet spot for biologics: Recent advances in characterization of biotherapeutic glycoproteins. *Expert Review of Proteomics*. 2018;**15**(1):13-29
- [18] Lingg N, Zhang P, Song Z, Bardor M. The sweet tooth of biopharmaceuticals: Importance of recombinant protein glycosylation analysis. *Biotechnology Journal*. Dec 1, 2012;**7**(12):1462-1472
- [19] Brorson K, Jia AY. Therapeutic monoclonal antibodies and consistent ends: Terminal heterogeneity, detection, and impact on quality. *Current Opinion in Biotechnology*. Dec, 2014;**30**:140-146
- [20] Beyer B, Schuster M, Jungbauer A, Lingg N. Microheterogeneity of recombinant antibodies: Analytics and functional impact. *Biotechnol J*. 2018 Jan;**13**(1). DOI: 10.1002/biot.201700476. PMID: 28862393

- [21] van Beers MMC, Bardor M. Minimizing immunogenicity of biopharmaceuticals by controlling critical quality attributes of proteins. *Biotechnology Journal*. 2012 Dec;7(12):1473-1484
- [22] Burda P, Aebi M. The dolichol pathway of N-linked glycosylation. *Biochimica et Biophysica Acta (BBA)—General Subjects*. Jan 6, 1999;1426(2):239-257
- [23] Weerapana E, Imperiali B. Asparagine-linked protein glycosylation: From eukaryotic to prokaryotic systems. *Glycobiology*. Jun 1, 2006;16(6):91R-101R
- [24] Gil G-C, Velander WH, Van Cott KE. N-glycosylation microheterogeneity and site occupancy of an Asn-X-Cys sequon in plasma-derived and recombinant protein C. *Proteomics*. May, 2009;9(9):2555-2567
- [25] Zielinska DF, Gnad F, Wiśniewski JR, Mann M. Precision mapping of an in vivo N-Glycoproteome reveals rigid topological and sequence constraints. *Cell*. May 28, 2010;141(5):897-907
- [26] Matsui T, Takita E, Sato T, Kinjo S, Aizawa M, Sugiura Y, et al. N-glycosylation at non-canonical Asn-X-Cys sequences in plant cells. *Glycobiology*. Aug 1, 2011;21(8):994-999
- [27] Williams DB. Beyond lectins: The calnexin/calreticulin chaperone system of the endoplasmic reticulum. *Journal of Cell Science*. Feb 15, 2006;119(4):615-623
- [28] Helenius A, Aebi M. Intracellular functions of N-linked glycans. *Science*. Mar 23, 2001;291(5512):2364-2369
- [29] Varki A. Loss of N-glycolylneuraminic acid in humans: Mechanisms, consequences, and implications for hominid evolution. *American Journal of Physical Anthropology*. Jan 1, 2001;116(S33):54-69
- [30] Lerouge P, Cabanes-Macheteau M, Rayon C, Fischette-Lainé A-C, Gomord V, Faye L. N-glycoprotein biosynthesis in plants: Recent developments and future trends. *Plant Molecular Biology*. Sep 1, 1998;38(1-2):31-48
- [31] Wilson IBH, Zeleny R, Kolarich D, Staudacher E, Stroop CJM, Kamerling JP, et al. Analysis of Asn-linked glycans from vegetable foodstuffs: Widespread occurrence of Lewis a, core α 1,3-linked fucose and xylose substitutions. *Glycobiology*. Apr 1, 2001;11(4):261-274
- [32] Gagneux P, Varki A. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology*. Aug 1, 1999;9(8):747-755
- [33] Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, et al., editors. *Essentials of Glycobiology* [Internet]. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK1908/>
- [34] Balshüsemann D, Jaenicke L. The oligosaccharides of the glycoprotein pheromone of *Volvox carteri* f. *nagariensis* Iyengar (Chlorophyceae). *European Journal of Biochemistry*. Aug 1, 1990;192(1):231-237

- [35] Grunow A, Becker B, Melkonian M. Isolation and characterization of the Golgi apparatus of a flagellate scaly green alga. *European Journal of Cell Biology*. Jun, 1993;**61**(1):10-20
- [36] Gödel S, Becker B, Melkonian M. Flagellar membrane proteins of *Tetraselmis striata* butcher (Chlorophyta). *Protist*. Aug, 2000;**151**(2):147-159
- [37] Levy-Ontman O, Arad SM, Harvey DJ, Parsons TB, Fairbanks A, Tekoah Y. Unique N-glycan moieties of the 66-kDa cell wall glycoprotein from the red microalga *Porphyridium* sp. *The Journal of Biological Chemistry*. Jun 17, 2011;**286**(24):21340-21352
- [38] Levy-Ontman O. N-Glycosylation of the 66-kDa Cell-Wall Glycoprotein of a Red Microalga. 2012 [cited 2017 Nov 17]; Available from: <http://www.intechopen.com/books/glycosylation/n-glycosylation-of-the-66-kda-cell-wall-glycoprotein-of-a-red-microalga>
- [39] Mathieu-Rivet E, Scholz M, Arias C, Dardelle F, Schulze S, Le Mauff F, et al. Exploring the N-glycosylation pathway in *Chlamydomonas reinhardtii* unravels novel complex structures. *Molecular & Cellular Proteomics*. Nov, 2013;**12**(11):3160-3183
- [40] Vanier G, Lucas P-L, Loutelier-Bourhis C, Vanier J, Plasson C, Walet-Balieu M-L, et al. Heterologous expression of the N-acetylglucosaminyltransferase I dictates a reinvestigation of the N-glycosylation pathway in *Chlamydomonas reinhardtii*. *Scientific Reports*. Aug 31, 2017;**7**(1):10156
- [41] Baiet B, Burel C, Saint-Jean B, Louvet R, Menu-Bouaouiche L, Kiefer-Meyer M-C, et al. N-Glycans of *Phaeodactylum tricorutum* diatom and functional characterization of its N-acetylglucosaminyltransferase I enzyme. *The Journal of Biological Chemistry*. Feb 25, 2011;**286**(8):6152-6164
- [42] Schulze S, Urzica E, Reijnders MJMF, van de Geest H, Warris S, Bakker LV, et al. Identification of methylated GnTI-dependent N-glycans in *Botryococcus brauni*. *The New Phytologist*. Sep, 2017;**215**(4):1361-1369
- [43] Casabianca S, Cornetti L, Capellacci S, Vernesi C, Penna A. Genome complexity of harmful microalgae. *Harmful Algae*. Mar 1, 2017;**63**(Supplement C):7-12
- [44] Levy-Ontman O, Fisher M, Shotland Y, Weinstein Y, Tekoah Y, Arad SM. Genes involved in the endoplasmic reticulum N-glycosylation pathway of the red microalga *Porphyridium* sp.: A bioinformatic study. *International Journal of Molecular Sciences*. Feb 7, 2014;**15**(2):2305-2326
- [45] Mathieu-Rivet E, Kiefer-Meyer M-C, Vanier G, Ovide C, Burel C, Lerouge P, et al. Protein N-glycosylation in eukaryotic microalgae and its impact on the production of nuclear expressed biopharmaceuticals. *Frontiers in Plant Science* [Internet]. Jul 28, 2014 [cited 2016 Nov 16];5. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4135232/>
- [46] Kallolimath S, Castilho A, Strasser R, Grünwald-Gruber C, Altmann F, Strubl S, et al. Engineering of complex protein sialylation in plants. *Proceedings of the National Academy of Sciences of the United States of America*. Aug 23, 2016;**113**(34):9498-9503

- [47] Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Rios S, et al. Humanization of yeast to produce complex terminally Sialylated glycoproteins. *Science*. Sep 8, 2006;**313**(5792):1441-1443
- [48] Gügi B, Le Costaouec T, Burel C, Lerouge P, Helbert W, Bardor M. Diatom-specific oligosaccharide and polysaccharide structures help to unravel biosynthetic capabilities in diatoms. *Marine Drugs*. Sep 18, 2015;**13**(9):5993-6018
- [49] Mathieu-Rivet E, Lerouge P, Bardor M. *Chlamydomonas reinhardtii*: Protein glycosylation and production of biopharmaceuticals. In: Hippler M, editor. *Chlamydomonas: Biotechnology and Biomedicine*. Microbiology Monographs. Vol. 31. Cham: Springer; 2017. DOI: https://doi.org/10.1007/978-3-319-66360-9_3
- [50] Galván A, González-Ballester D, Fernández E. Insertional mutagenesis as a tool to study genes/functions in *Chlamydomonas*. *Advances in Experimental Medicine and Biology*. 2007;**616**:77-89
- [51] Doron L, Segal N, Shapira M. Transgene expression in microalgae—from tools to applications. *Frontiers in Plant Science*. 2016;**7**:505
- [52] Daboussi F. Advances in editing microalgae genomes. *Perspectives in Phycology*. 2017;**4**(1):17-23
- [53] Ng I-S, Tan S-I, Kao P-H, Chang Y-K, Chang J-S. Recent developments on genetic engineering of microalgae for biofuels and bio-based chemicals. *Biotechnology Journal*. Oct 2017;**12**(10):1600644. DOI: 10.1002/biot.201600644
- [54] Falcatore A, Casotti R, Leblanc C, Abrescia C, Bowler C. Transformation of nonselectable reporter genes in marine diatoms. *Marine Biotechnology (New York, NY)*. May, 1999;**1**(3):239-251
- [55] Kindle KL. High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences*. Jan 2, 1990;**87**(3):1228-1232
- [56] Armbrust EV. Identification of a new gene family expressed during the onset of sexual reproduction in the centric Diatom *Thalassiosira weissflogii*. *Applied and Environmental Microbiology*. Jan 7, 1999;**65**(7):3121-3128
- [57] Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, et al. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *The EMBO Journal*. Sep 15, 1998;**17**(18):5497-5508
- [58] Maeder ML, Gersbach CA. Genome-editing technologies for gene and cell therapy. *Molecular Therapy*. Mar 1, 2016;**24**(3):430-446
- [59] Daboussi F, Leduc S, Maréchal A, Dubois G, Guyot V, Perez-Michaut C, et al. Genome engineering empowers the diatom *Phaeodactylum tricorutum* for biotechnology. *Nature Communications*. May 29, 2014;**5**:3831

- [60] Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature Biotechnology*. Jul 1, 2007;**25**(7):nbt1319
- [61] Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Research*. Sep 1, 2000;**28**(17):3361-3369
- [62] Sizova I, Greiner A, Awasthi M, Kateriya S, Hegemann P. Nuclear gene targeting in *Chlamydomonas* using engineered zinc-finger nucleases. *The Plant Journal*. 2013 Mar 1; **73**(5):873-882
- [63] Greiner A, Kelterborn S, Evers H, Kreimer G, Sizova I, Hegemann P. Targeting of Photoreceptor Genes in *Chlamydomonas reinhardtii* via Zinc-finger Nucleases and CRISPR/Cas9. *Plant Cell*. Jan 1, 2017;tpc.00659.2017
- [64] Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppl C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nature Biotechnology*. Aug 7, 2011;**29**(9):nbt.1948
- [65] Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*. Oct 1, 2010;**186**(2):757-761
- [66] Kay S, Hahn S, Marois E, Hause G, Bonas U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science*. Oct 26, 2007;**318**(5850):648-651
- [67] Weyman PD, Beeri K, Lefebvre SC, Rivera J, McCarthy JK, Heuberger AL, et al. Inactivation of *Phaeodactylum tricorutum* urease gene using transcription activator-like effector nuclease-based targeted mutagenesis. *Plant Biotechnology Journal*. May, 2015;**13**(4):460-470
- [68] Fortunato AE, Jaubert M, Enomoto G, Bouly J-P, Raniello R, Thaler M, et al. Diatom phytochromes reveal the existence of far-red light based sensing in the ocean. *Plant Cell*. Jan 1, 2016;tpc.00928.2015
- [69] Liao H-K, Hatanaka F, Araoka T, Reddy P, Wu M-Z, Sui Y, et al. In vivo target gene activation via CRISPR/Cas9-mediated trans-epigenetic modulation. *Cell* [Internet]. Dec 7, 2017 [cited 2017 Dec 9]. Available from: [http://www.cell.com/cell/abstract/S0092-8674\(17\)31247-3](http://www.cell.com/cell/abstract/S0092-8674(17)31247-3)
- [70] Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR–Cas systems. *Nature Reviews. Microbiology*. Jun, 2011;**9**(6):467-477
- [71] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. Aug 17, 2012;**337**(6096):816-821
- [72] Jiang W, Brueggeman AJ, Horken KM, Plucinak TM, Weeks DP. Successful transient expression of Cas9 and single guide RNA genes in *Chlamydomonas reinhardtii*. *Eukaryotic Cell*. Jan 11, 2014;**13**(11):1465-1469

- [73] Shin S-E, Lim J-M, Koh HG, Kim EK, Kang NK, Jeon S, et al. CRISPR/Cas9-induced knockout and knock-in mutations in *Chlamydomonas reinhardtii*. *Scientific Reports*. Jun 13, 2016;**6**:27810
- [74] Nymark M, Sharma AK, Sparstad T, Bones AM, Winge P. A CRISPR/Cas9 system adapted for gene editing in marine algae. *Scientific Reports*. Apr 25, 2016;**6**:24951
- [75] Hopes A, Nekrasov V, Kamoun S, Mock T. Editing of the urease gene by CRISPR-Cas in the diatom *Thalassiosira pseudonana*. *Plant Methods*. Nov 24, 2016;**12**:49
- [76] Wang Q, Lu Y, Xin Y, Wei L, Huang S, Xu J. Genome editing of model oleaginous microalgae *Nannochloropsis* spp. by CRISPR/Cas9. *The Plant Journal*. Dec 1, 2016;**88**(6):1071-1081
- [77] Chevalier BS, Stoddard BL. Homing endonucleases: Structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Research*. Sep 15, 2001;**29**(18):3757-3774
- [78] Joung JK, Sander JD. TALENs: A widely applicable technology for targeted genome editing. *Nature Reviews. Molecular Cell Biology*. 2013 Jan;**14**(1):49-55
- [79] Li X, Zhang R, Patena W, Gang SS, Blum SR, Ivanova N, et al. An indexed, mapped mutant library enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. *The Plant Cell*. 2016 Feb;**28**(2):367-387

Depuis la publication de ce chapitre de livre, des travaux complémentaires portant sur les processus de la *N*-glycosylation chez les microalgues ont été publiés. Par exemple, en 2017, des travaux sur la microalgue *Euglena gracilis* ont montré qu'elle était capable d'accumuler des *N*-glycannes de type oligomannosidique sur ses glycoprotéines endogènes (**O'Neill et al., 2017**). Plus récemment, les travaux sur la microalgue verte *Chlorella vulgaris* de l'équipe du Pr Altmann (Université de Boku en Autriche) ont permis d'élucider les profils *N*-glycanniques des protéines de cette microalgue. Les résultats ont montré que cette microalgue produisait majoritairement des *N*-glycannes hautement mannosylés pouvant être plus ou moins méthylés (**Mócsai et al., 2019**). Cette même équipe a travaillé sur d'autres sous-espèces de *Chlorella* montrant la présence sur leurs protéines endogènes de structures *N*-glycanniques plus complexes possédant des résidus galactose. Ces *N*-glycannes peuvent également être substitués par de l'arabinose et du xylose (**Mócsai et al., 2020a ; Mócsai et al., 2020b**).

Le projet de thèse développé dans ce manuscrit se focalise plus spécifiquement sur les deux microalgues utilisées jusqu'alors pour la production d'anticorps monoclonaux à intérêt thérapeutique et pour lesquelles les processus de *N*-glycosylation sont relativement connus à savoir la microalgue verte, *C. reinhardtii* et la diatomée, *P. tricornutum*. En conséquence, les connaissances sur leur voie de *N*-glycosylation sont détaillées dans les paragraphes suivants.

A. La *N*-glycosylation chez la microalgue verte *C. reinhardtii*

La synthèse du précurseur glycannique des glycoprotéines de la microalgue verte *C. reinhardtii* se déroule dans le réticulum endoplasmique, tout comme chez les autres eucaryotes. Cependant, certains gènes candidats pour la synthèse du précurseur glycannique chez les autres eucaryotes semblent absents chez *C. reinhardtii*. En effet, les études bio-informatiques effectuées sur son génome montrent l'absence des gènes codants l'ALG3, ALG9, ALG10 et ALG12 (**Mathieu-Rivet et al., 2013 ; Mathieu-Rivet et al., 2014**). Or, ces gènes sont primordiaux pour la formation d'un *N*-glycanne précurseur de type $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-PP-Dol}$. Ces informations suggèrent la formation d'un précurseur glycannique tronqué de type $\text{Hex}_7\text{HexNAc}_2$ (**Figure 19**). Ce précurseur tronqué serait ensuite transféré sur une protéine en cours de synthèse par l'OST, suivi d'une étape de contrôle qualité avec l'élimination des deux

résidus glucose pour aboutir à une glycoprotéine portant un N -glycane de type $\text{Man}_5\text{GlcNAc}_2$ (Figure 19).

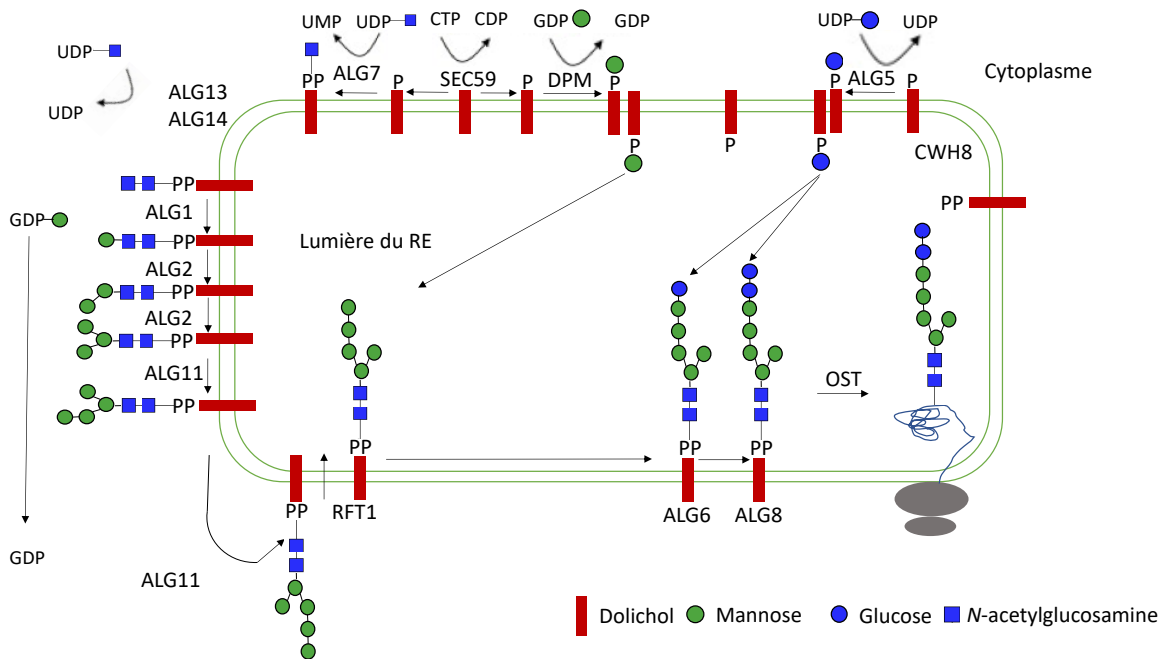


Figure 19 : Voie de la synthèse du précurseur glycanique effectuée dans le réticulum endoplasmique chez *C. reinhardtii* suggérée par les analyses bio-informatiques et biochimiques (Mathieu-Rivet et al., 2013 ; Mathieu-Rivet et al., 2014 ; Vanier et al., 2017)

Ce schéma résume les étapes de la synthèse du précurseur oligosaccharidique chez la microalgue verte *C. reinhardtii* se déroulant dans le réticulum endoplasmique. Le changement d'orientation entre la phase cytosolique et la phase luminale est opéré probablement par une enzyme de type flippase RFT1 dont le gène putatif a été identifié par homologie de séquence dans le génome de *C. reinhardtii*. Les différentes étapes présentées ici consistent en l'ajout de différents monosaccharides. Elles sont catalysées par une succession d'enzymes appelées ALG (Asparagine Linked Glycosylation). Contrairement à la majorité des eucaryotes, ces étapes conduiraient à la formation d'un précurseur oligosaccharidique lié au dolichol composé de 2 résidus N -acétylglucosamine, de 5 résidus mannose et de 2 résidus glucose. Ce précurseur oligosaccharidique serait ensuite transféré sur une protéine en cours de synthèse par l'oligosaccharyltransférase (OST). Les structures N -glycaniques sont dessinées en accord avec la nomenclature internationale (Varki et al., 2015).

UDP : uridine diphosphate ; UMP uridine monophosphate ; GDP : guanosine diphosphate ; ALG : asparagine linked glycosylation ; DPM : dolichyl-phosphate mannosyltransférase I ; OST : oligosaccharyltransférase ; CTP : cytosine triphosphate ; CDP : cytosine diphosphate ; PP : pyrophosphate ; RE : réticulum endoplasmique

Une fois dans l'appareil de Golgi, la glycoprotéine naissante va subir des modifications supplémentaires. Des études glycomiques et glycoprotéomiques réalisées chez cette microalgue ont montré que le N -glycane majoritaire porté par les protéines de l'algue verte

est de type $\text{Man}_5\text{GlcNAc}_2$ (**Mathieu-Rivet et al., 2013**). Cependant, contrairement à la majorité des eucaryotes, aucun gène candidat pour une GnT I ne semble identifiable dans le génome de *C. reinhardtii* (**Mathieu-Rivet et al., 2013 ; Mathieu-Rivet et al., 2014**). L'absence de la GnT I concorde avec la formation d'une structure linéaire de type $\text{Man}_5\text{GlcNAc}_2$ qui n'est pas le substrat accepteur de la GnT I (**Vanier et al., 2017**). Des gènes codant des xylosyltransférases et des méthyltransférases sont présents dans le génome de la microalgue verte. La présence de ces gènes est en accord avec les structures glycaniques identifiées qui sont plus ou moins xylosylées et méthylées (**Mathieu-Rivet et al., 2013 ; Mathieu-Rivet et al., 2014 ; Vanier et al., 2017**). Récemment, deux gènes nommés XTA et XTB, codant des xylosyltransférases ont pu être caractérisés chez *C. reinhardtii* (**Schulze et al., 2018 ; Lucas et al., 2020**). La première XTA a été inactivée par l'équipe du Pr Michaël Hippler de l'Université de Münster en Allemagne en 2018. Cette inactivation a permis de réduire considérablement la xylosylation du core glycanique des protéines sécrétées (**Schulze et al., 2018**). Ces travaux ont également montré que les *N*-glycannes plus ou moins méthylés portant des résidus $\beta(1,2)$ -xylose étaient fortement diminués avec l'absence de l'activité mannosidase I de *C. reinhardtii*, suggérant que la xylosylation et la méthylation sont dépendantes de cette activité mannosidase I (**Schulze et al., 2018**). Un second gène codant une xylosyltransférase appelée XTB a pu être identifié chez la microalgue verte dont la protéine exprimée a pu être caractérisée (**Schulze et al., 2018 ; Lucas et al., 2020**). Les travaux menés par Lucas et collaborateurs en 2020 montrent que la protéine XTA est responsable majoritairement du transfert d'un résidu xylose en $\beta(1,2)$ sur le core du *N*-glycane alors que la protéine XTB est responsable du transfert d'un xylose en $\beta(1,4)$ sur un mannose de la branche linéaire du *N*-glycane des protéines synthétisées et dans une moindre mesure de l'ajout de xylose lié en $\beta(1,2)$ sur le core du *N*-glycane (**Lucas et al., 2020**). Ces travaux ont mis en lumière la présence d'une grande hétérogénéité des structures glycaniques xylosylées chez *C. reinhardtii* aboutissant à l'hypothèse qu'une autre xylosyltransférase serait également impliquée dans la maturation des *N*-glycannes de *C. reinhardtii* (**Figure 20 ; Lucas et al., 2020**). Ainsi, 3 nouveaux gènes candidats ont été identifiés dans le génome de *C. reinhardtii* pour cette activité xylosyltransférase supplémentaire. Une fucosylation potentielle a également été récemment mise en évidence sur les *N*-glycannes de *C. reinhardtii* (**Figure 20 ; Oltmanns et al., 2020**).

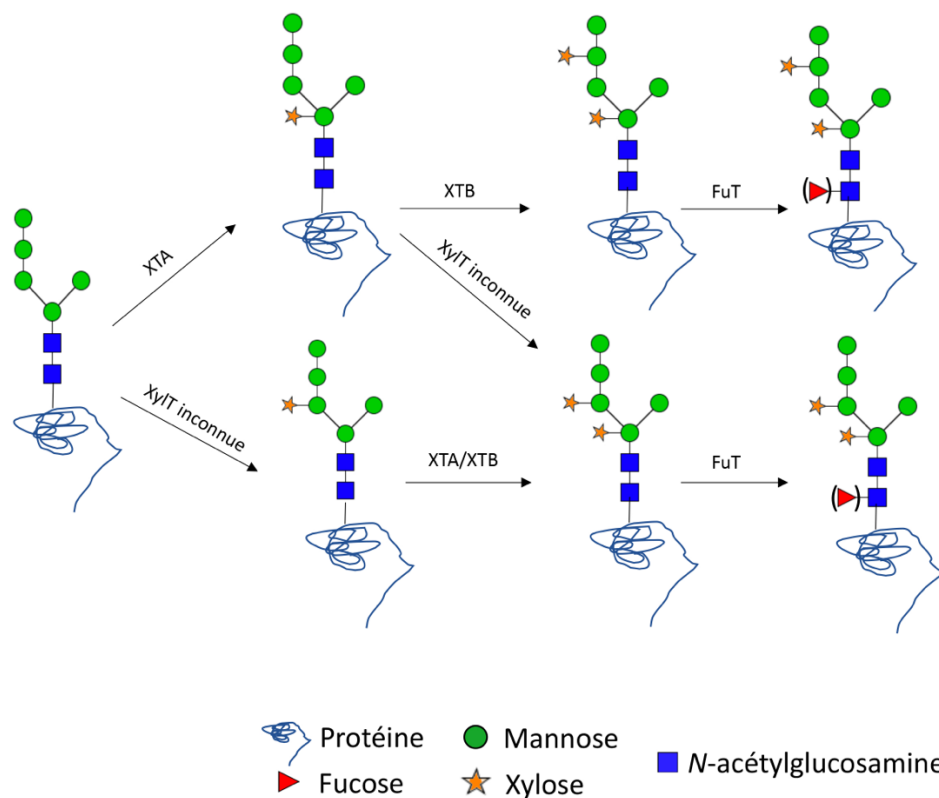


Figure 20 : Voies de la xylosylation et de la fucosylation proposées chez *C. reinhardtii* (figure construite d'après Lucas *et al.*, 2020 et Oltmanns *et al.*, 2020)

Ce schéma résume les étapes de la synthèse des *N*-glycannes complexes portant des résidus xylose et fucose chez la microalgue verte *C. reinhardtii*. Cette étape se déroule dans l'appareil de Golgi. Ces voies de xylosylation sont orchestrées par les deux xylosyltransférases (XylIT) caractérisées (appelées XTA et XTB) ainsi que par une XylIT encore inconnue à ce jour. La fucosylation est orchestrée par une fucosyltransférase (FuT). Les structures *N*-glycanniques sont dessinées en accord avec la nomenclature internationale (Varki *et al.*, 2015).

XylIT : xylosyltransférase ; FuT : fucosyltransférase ; XTA : xylosyltransférase A ; XTB : xylosyltransférase B

B. La *N*-glycosylation chez *P. tricornutum*

Chez la diatomée *P. tricornutum*, le processus de la *N*-glycosylation se déroule tout d'abord dans le réticulum endoplasmique avec la synthèse d'un précurseur oligosaccharidique. D'après les analyses bio-informatiques réalisées, les gènes codant l'ensemble des enzymes ALG sembleraient être présents dans le génome de la diatomée excepté l'ALG10 qui transfère un résidu glucose lié en $\alpha(1,3)$ sur l'oligosaccharide précurseur chez la majorité des eucaryotes (Baïet *et al.*, 2011 ; Levy-Ontman *et al.*, 2014). Ce précurseur oligosaccharidique serait ensuite transféré sur une protéine en cours de synthèse par l'enzyme OST comme chez les eucaryotes. En effet, des analyses bio-informatiques réalisées

sur le génome de *P. tricornutum* ont permis de prédire la présence d'un gène qui coderait la sous-unité catalytique STT3 de l'OST (Baïet *et al.*, 2011). La glycoprotéine naissante atteindra ainsi le cycle de contrôle qualité où des étapes de déglucosylation et l'intervention de lectines surviendront afin d'obtenir une protéine bien conformée. Une fois ce contrôle terminé, la glycoprotéine rejoindra l'appareil de Golgi.

À la suite de l'adressage de la glycoprotéine dans l'appareil de Golgi, le glycanne subira des maturations par l'addition et l'hydrolyse de différents oses. Les analyses bio-informatiques réalisées dans le génome de la diatomée montrent la présence de plusieurs gènes codant des glycosyltransférases et des glycosidases retrouvés aussi dans le génome des mammifères comme le gène de la Man I capable de dégrader le glycanne oligomannosidique de Man₉GlcNAc₂ en Man₅GlcNAc₂. Ces analyses ont aussi détecté la présence du gène codant la GnT I capable de transférer un résidu GlcNAc sur le mannose branché en $\alpha(1,3)$, et finalement, la présence du gène de la Man II, de la GnT II et de trois gènes codant des FuT potentielles permettant respectivement l'hydrolyse de deux mannoses, l'addition éventuelle d'un résidu GlcNAc terminal et enfin l'addition d'un fucose (Baïet *et al.*, 2011 ; Mathieu-Rivet *et al.*, 2014). Les travaux réalisés par Baïet et collaborateurs en 2011 ont permis de confirmer la fonctionnalité de la GnT I de *P. tricornutum* par complémentation fonctionnelle de cellules CHO *lec1* possédant une mutation sur le gène de la GnT I (Baïet *et al.*, 2011). Cette complémentation, par une expression hétérologue du gène *GnT I* de *P. tricornutum* a permis de restaurer l'activité GnT I chez ces mutants CHO (Baïet *et al.*, 2011), confirmant ainsi la localisation golgienne et la fonctionnalité *in vivo* de la GnT I de *P. tricornutum*.

Plus récemment, des travaux ont permis de caractériser un transporteur du GDP-fucose, impliqué dans le transport du GDP-fucose depuis le cytosol vers l'appareil de Golgi. L'activité de ce transporteur du GDP-fucose a été mise en évidence comme précédemment par la complémentation hétérologue de l'expression du gène de *P. tricornutum* dans une cellule mutante CHO-*gmt5* dépourvue de son propre transporteur de GDP-fucose (Zhang *et al.*, 2019). Cette complémentation a permis de restaurer la fucosylation des *N*-glycannes des cellules mutantes CHO-*gmt5* (Zhang *et al.*, 2019). Parallèlement, trois gènes présentant des homologies avec le gène codant l' $\alpha(1,3)$ FuT d'*A. thaliana* ont également été détectés dans le génome de la diatomée (Baïet *et al.*, 2011 ; Ovide, 2016 ; Zhang *et al.*, 2019). L'un de ces 3 gènes putatifs code une $\alpha(1,3)$ FuT fonctionnelle localisée dans l'appareil de Golgi (Ovide, 2016 ; Zhang *et al.*, 2019). En effet, la surexpression de ce gène sous la dépendance du

promoteur de la nitrate réductase inductible par le nitrate, a conduit à une augmentation du taux de core $\alpha(1,3)$ -fucose sur les *N*-glycannes des protéines endogènes de la diatomée **(Zhang *et al.*, 2019)**. L'ensemble de ces données laissent à penser qu'un intermédiaire de type GlcNAc₂Man₃FucGlcNAc₂ pourrait être formé chez *P. tricornutum* **(Figure 21)**. Cependant, cette structure glycanne n'a pas été détectée dans les analyses structurales précédemment réalisées **(Baïet *et al.*, 2011)**. Seule, la forme paucimannosidique de type Man₃FucGlcNAc₂ a été identifiée dans les analyses glycomiques menées sur les protéines totales de la diatomée **(Baïet *et al.*, 2011)**. La présence de ce glycanne paucimannosidique peut s'expliquer par l'identification de deux gènes codant des potentielles candidates à l'activité hexosaminidase qui pourrait être des *N*-acétylglucosaminidases. Ces activités enzymatiques, si leurs présences étaient confirmées, seraient capables d'éliminer le résidu GlcNAc initialement transféré par la GnT I et la GnT II aboutissant à la synthèse d'un *N*-glycanne paucimannosidique sur les glycoprotéines synthétisées **(Figure 21, Baïet *et al.*, 2011)**. Ce type d'activité a déjà été caractérisée chez d'autres espèces, notamment chez *A. thaliana* **(Strasser *et al.*, 2007 ; Liebming *et al.*, 2011)** ou encore chez les insectes **(Geisler *et al.*, 2008 ; Kim *et al.*, 2009)**.

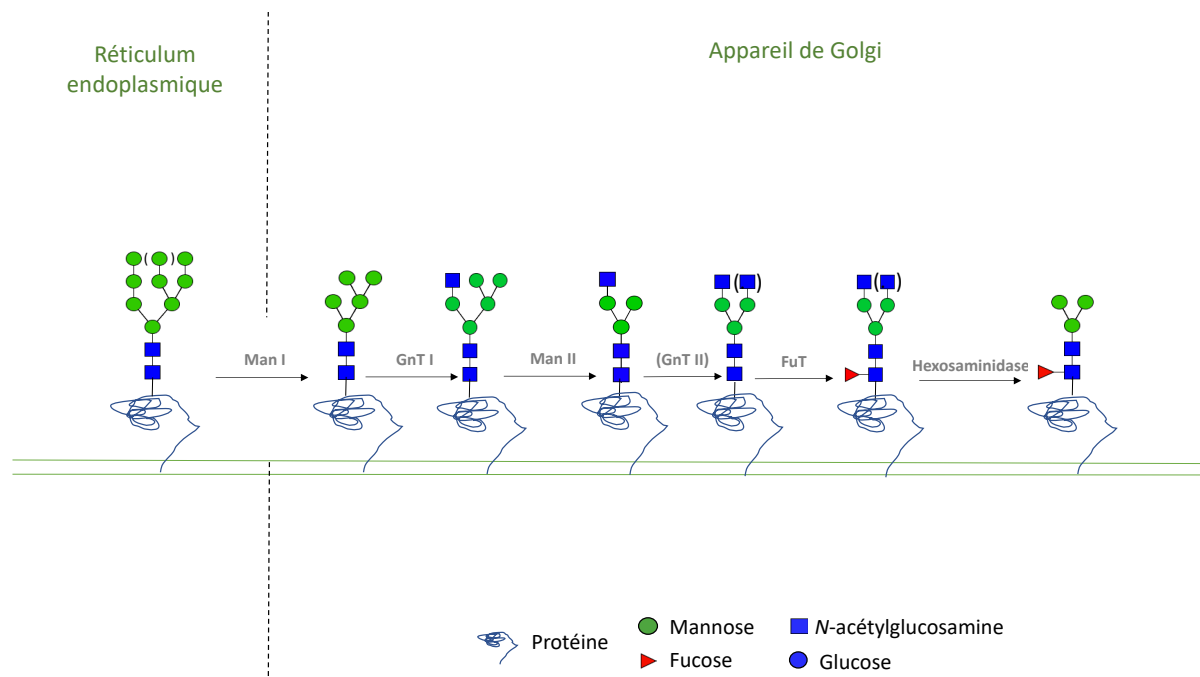


Figure 21 : Voie de la biosynthèse des *N*-glycannes complexes chez *P. tricornutum* (Baïet *et al.*, 2011 ; Mathieu-Rivet *et al.*, 2014 ; Zhang *et al.*, 2019)

Le précurseur oligosaccharidique de type $\text{Man}_{9-8}\text{GlcNAc}_2$, précédemment synthétisé dans le réticulum endoplasmique est transformé en $\text{Man}_5\text{GlcNAc}_2$ par la Man I dans l'appareil de Golgi. Cette structure étant le substrat accepteur de la GnT I, permet la synthèse de *N*-glycannes complexes de type $\text{GlcNAcMan}_5\text{GlcNAc}_2$. Cette structure intermédiaire est ensuite maturée par l'activité de la Man II, l'activité potentielle d'une GnT II et par l'ajout d'un résidu fucose lié en $\alpha(1,3)$ résultant de l'action d'une $\alpha(1,3)$ -fucosyltransférase. L'activité de l'hexosaminidase qui éliminerait les résidus GlcNAc précédemment transférés par la GnT I et la GnT II permettrait la synthèse d'une structure paucimannosidique de type $\text{Man}_3\text{FucGlcNAc}_2$. Les structures *N*-glycanniques sont dessinées en accord avec la nomenclature internationale (Varki *et al.*, 2015).

Man I : α -mannosidase I ; GnT I : *N*-acetylglucosaminyltransférase I ; GnT II : *N*-acetylglucosaminyltransférase II ; Man II : α -mannosidase II ; FuT : $\alpha(1,3)$ -fucosyltransférase

Objectifs de la thèse

Dans un objectif de production d'anticorps monoclonaux recombinants chez *P. tricornutum*, il est nécessaire à terme d'humaniser sa voie de *N*-glycosylation de sorte à obtenir des anticorps recombinants compatibles avec certaines applications thérapeutiques chez l'Homme. Les données disponibles en début de la thèse étaient celles publiées par Baïet *et al.*, 2011, Levy-Ontman *et al.*, 2014, Mathieu-Rivet *et al.*, 2014 et Ovide, 2016 que l'on peut résumer ainsi :

- Des données bio-informatiques prédisant dans le génome de *P. tricornutum* des séquences codant potentiellement des acteurs de la voie de la *N*-glycosylation au niveau du RE (ALG, OST, etc...) et au niveau de l'appareil de Golgi (GnT I, Man II, GnT II, FuT et hexosaminidases).
- Des données biochimiques permettant de prouver la fonctionnalité de la GnT I et une étude initiale de la *N*-glycosylation des protéines endogènes de *P. tricornutum* mettant en évidence des structures oligomannosidiques allant de Man₉GlcNAc₂ à Man₅GlcNAc₂ et des traces d'un *N*-glycane paucimannosidique de type Man₃FucGlcNAc₂.

Compte-tenu de ces informations, une voie potentielle de la *N*-glycosylation des protéines de *P. tricornutum* avait été proposée. Elle est résumée dans la **Figure 21 (Baïet et al., 2011 ; Mathieu-Rivet et al., 2014)**. Cependant, le très faible taux de maturation des structures oligomannosidiques en *N*-glycannes complexes nécessite des investigations complémentaires du processus de *N*-glycosylation chez *P. tricornutum*.

En conséquence, l'objectif de la thèse était, dans un premier temps, de valider ou d'invalider la voie de la *N*-glycosylation de *P. tricornutum* proposée par Baïet et collaborateurs en 2011 par des données d'analyses structurales détaillées des *N*-glycannes des protéines endogènes. Dans un second temps des stratégies d'ingénierie métabolique ont été initiées afin de mener à la synthèse de *N*-glycannes complexes homologues à ceux présents sur les protéines humaines dans cet organisme.

A. Analyse du précurseur oligosaccharidique chez *P. tricornutum*

Les analyses bio-informatiques du génome de *P. tricornutum* réalisées antérieurement à ce travail ont suggéré l'absence de l'ALG10, une enzyme résidente du RE codant une $\alpha(1,2)$ -glucosyltransférase transférant un dernier résidu glucose nécessaire à la dernière étape de la synthèse du précurseur oligosaccharidiques des *N*-glycannes chez l'Homme (Baïet et al., 2011 Levy-Ontman et al., 2014). Ce défaut de transfert du troisième résidu glucose sur le précurseur oligosaccharidique lié au dolichol conduirait en conséquence à la synthèse chez *P. tricornutum* d'un oligosaccharide tronqué de type Glc₂Man₉GlcNAc₂. Dans ce contexte, le premier objectif de ma thèse a été de développer une méthode originale de purification et

d'analyse structurale des précurseurs oligosaccharidiques liés au dolichol. Ce protocole a été validé en confirmant la structure du précurseur oligosaccharide de *C. reinhardtii* (**Vanier et al., 2017**). Puis, ce protocole a été appliqué au précurseur synthétisé par *P. tricornutum* dans le RE afin de démontrer pour la première fois, la structure du précurseur oligosaccharidique des *N*-glycannes chez *P. tricornutum*.

B. Analyse structurale détaillée des *N*-glycannes oligomannosidiques chez *P. tricornutum*

Pendant cette thèse, les analyses par spectrométrie de masse des profils de *N*-glycosylation des protéines de *P. tricornutum* ont révélé la présence de *N*-glycannes allant de $\text{Man}_9\text{GlcNAc}_2$ à $\text{Man}_5\text{GlcNAc}_2$, confirmant les résultats antérieurs (**Baïet et al., 2011**). Cependant, plusieurs isomères, liés au positionnement des liaisons osidiques engagées entre les résidus dans le *N*-glycane, peuvent exister pour chaque type de *N*-glycannes oligomannosidiques. Ceci dépend des spécificités des α -mannosidases impliquées dans leur maturation golgienne. Dès lors, il est essentiel de déterminer la structure des différents isomères présents chez *P. tricornutum*. Cette question se pose de façon encore plus prégnante pour l'oligosaccharide composé de 5 résidus mannose $\text{Man}_5\text{GlcNAc}_2$. En effet, ce *N*-glycane est le substrat de la GnT I qui assure le transfert sur ce $\text{Man}_5\text{GlcNAc}_2$ d'un résidu GlcNAc terminal, ouvrant ainsi la voie vers la synthèse golgienne des *N*-glycannes complexes. Malgré la présence d'une GnT I fonctionnelle dans la diatomée, les analyses biochimiques réalisées sur les glycoprotéines endogènes de *P. tricornutum* n'ont pas permises de mettre en évidence de *N*-glycannes portant des résidus GlcNAc en position terminale. Une des explications de cette absence est que le $\text{Man}_5\text{GlcNAc}_2$ synthétisé puisse être différent de l'isomère retrouvé chez la majorité des eucaryotes (**Figure 22 ; Aebi, 2013**) Cet isomère $\text{Man}_5\text{GlcNAc}_2$ retrouvé chez la majorité des eucaryotes sera appelé « canonique » dans la suite de ce manuscrit. Un isomère $\text{Man}_5\text{GlcNAc}_2$ différent, comme démontré chez *C. reinhardtii*, sera appelée linéaire (**Vanier et al., 2017**). Ainsi, le second objectif de ce travail a été de déterminer les structures détaillées des *N*-glycannes oligomannosidiques synthétisés chez *P. tricornutum* afin de confirmer ou d'infirmer que le $\text{Man}_5\text{GlcNAc}_2$ possède une structure dit « canonique » nécessaire à la catalyse de la GnT I (**Figure 22**).

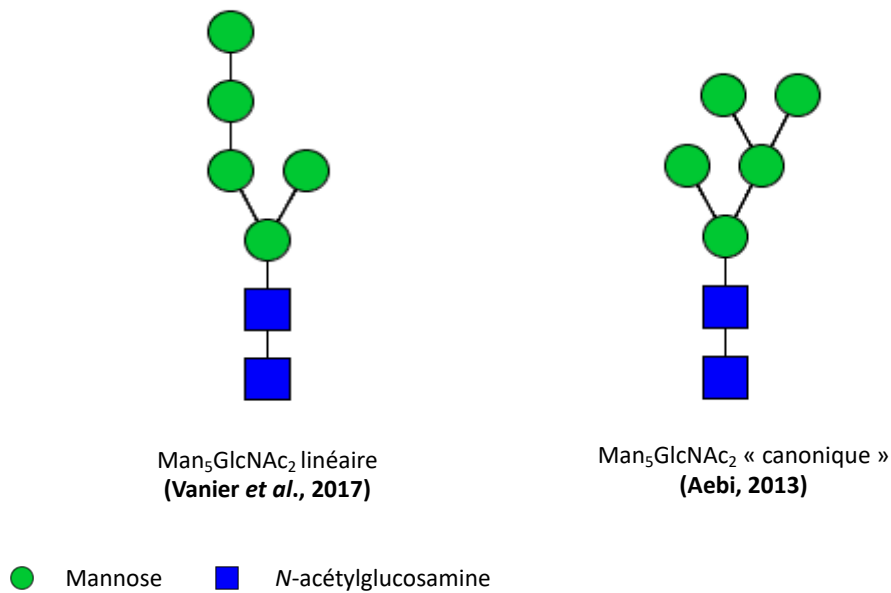


Figure 22 : Isomère du $\text{Man}_5\text{GlcNAc}_2$ dit « canonique » retrouvé chez la majorité des eucaryotes (Aebi, 2013) et linéaire comme retrouvé chez *C. reinhardtii*

L'isomère $\text{Man}_5\text{GlcNAc}_2$ dit « canonique » est retrouvé chez la majorité des eucaryotes (Aebi 2013). Cet isomère est le substrat accepteur de la GnT I, enzyme primordiale à la synthèse de *N*-glycannes complexes. L'isomère $\text{Man}_5\text{GlcNAc}_2$ linéaire retrouvé chez *C. reinhardtii* n'est, quant à lui, pas le substrat accepteur de la GnT I (Vanier et al., 2017)

Man : mannose ; GlcNAc : *N*-acétylglucosamine

C. Tentative d'humanisation de la voie de la *N*-glycosylation chez la diatomée *P. tricornutum*

Compte-tenu du très faible taux de maturation des oligomannosides en *N*-glycannes complexes, le troisième objectif de cette thèse a été de générer des transformants de *P. tricornutum* surexprimant la GnT I de *P. tricornutum*. La surexpression de la GnT I avait pour objectif de palier à la faible activité apparente de l'enzyme endogène. A terme, l'objectif final de cette thèse est la synthèse majoritaire chez *P. tricornutum* de *N*-glycannes portant des résidus *N*-acétylglucosamine en position terminale. L'obtention d'un tel *N*-glycanne sera le point de départ pour la formation de *N*-glycannes complexes analogues aux *N*-glycannes présents sur les anticorps monoclonaux et qui portent des résidus galactose sur les 2 branches de l'oligosaccharide en position terminale.

Chapitre 2

Matériel et méthode

I. Matériels

A. Matériel biologique

1. *Phaeodactylum tricornutum*

L'ensemble des travaux effectués durant cette thèse ont été réalisés avec la souche Pt 1.8.6 de *Phaeodactylum tricornutum* fournie par le laboratoire de Physiologie et de Biotechnologie des Algues de l'IFREMER de Nantes et par le laboratoire LISBP de l'INRA de Toulouse.

2. *Chlamydomonas reinhardtii*

L'ensemble des travaux effectués durant cette thèse ont été réalisés avec la souche CC-5325 de *Chlamydomonas reinhardtii*. Les mutants XTA et XTB utilisés dans ces travaux ont été obtenus à partir de la banque de mutants CLiP (**Site internet :** <https://www.chlamylibrary.org/>) et générés par insertion d'une cassette d'ADN double brin CIB1 permettant d'apporter la résistance à la paromomycine.

3. Souche bactérienne

La souche d'*E. coli* DH5 α (MAX Efficiency™ DH5 α Competent Cells, Invitrogen™) a été utilisée dans les expérimentations de transformation génétique afin de répliquer les plasmides présentés ci-dessous.

B. Vecteur plasmidique

Plusieurs constructions de vecteurs plasmidiques ont été conçues afin d'initier les travaux sur l'ingénierie métabolique de *P. tricornutum*. Les constructions ont été obtenues à partir du vecteur plasmidique Pha-NR utilisé couramment chez *P. tricornutum* (**Figure 23**). Ce vecteur est composé du couple promoteur/terminateur correspondant au gène codant la nitrate réductase (promoteur inductible par le nitrate), d'un gène de résistance à l'ampicilline pour le clonage dans les bactéries *E. coli*, ainsi que d'un gène de résistance à la zéocine afin de sélectionner les clones de *P. tricornutum* transformés. Enfin, les gènes d'intérêt sont

insérés dans ce plasmide au niveau du site de multiclonaage qui est complétement d'une séquence qui donnera à la protéine traduite un épitope « tag » (V5, HA) permettant de l'identifier. Les gènes d'intérêt utilisés pour la transformation génétique de *P. tricornutum* furent respectivement : le gène endogène codant la *N*-acétylgucosaminyltransférase I de la diatomée (Phatr3_J54844, EnsemblProtists) complétement de la séquence du tag V5 : GKPIPPLLGLDST, le gène codant pour la β (1,4)-galactosyltransférase I humaine (P15291, Uniprot) complétement de la séquence du tag HA triplement répété: YPYDVPDYA.

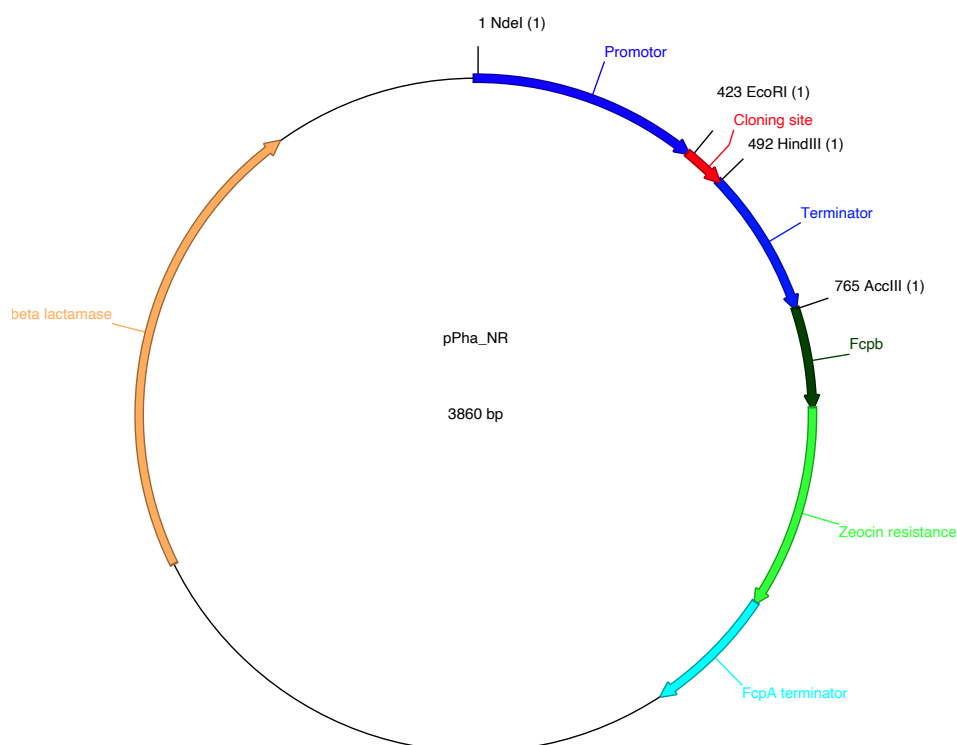


Figure 23 : Carte du vecteur plasmidique Pha-NR utilisé pour les transformations génétiques de *Phaeodactylum tricornutum*

C. Amorces utilisées pour le criblage des clones transformés

Différents jeux d'amorces ont été utilisés pour le criblage des clones ayant inséré les gènes codant soit la GnT I – V5, soit la GalT – HA ou les 2 pour les doubles transformants. Des jeux d'amorces amplifiant deux gènes de référence ont été également utilisés pour l'étape de criblage lors de l'amplification génique afin de valider les conditions expérimentales utilisées.

Les jeux d’amorces, leurs séquences ainsi que leurs températures d’hybridation sont résumés dans le tableau ci-dessous (**Tableau 2**).

Tableau 2 : Nom et séquence des amorces utilisées lors du criblage par amplification génique

AMORCES	SEQUENCES NUCLEOTIDIQUE	CIBLES	TAILLE DE L'AMPLICON	TEMPERATURES D'HYBRIDATION
PHATR3_J54844-FOR4	ATCTCCCGCACCTACCACTT	Gène <i>GnT 1</i> de <i>P. tricornutum</i> complémenté par le peptide V5	532pb	62°C
FUCT46109 NM-V5-C-RV2	GACCGAGGAGAGGGTTAGGG			
GALT-HA-C_FW2	TCACTCACCTACCAGGTGCT	Gène β 4GALT1 humaine complémenté par le peptide HA	163pb	60°C
GALT-HA-C_RV2	CGTCATATGGATAGGACCCTGC			
CDKA1F1	TGGAGCTCCCGATTCTCAAG	Gène de référence CDKA1 de <i>P. tricornutum</i>	221pb	60°C
CDKA1R1	TGATGAGCAAGTTAGGGGGC			
EF1AF1	TCGCGTTGTTCACTCTCTCG	Gène de référence EF1A de <i>P. tricornutum</i>	193pb	60°C
EF1AR1	ACGCTTATCGATACCTCCACAC			

II. Méthodes de culture

A. *Phaeodactylum tricornutum*

1. Sur des boîtes gélosées

Les cellules de *P. tricornutum* ont été cultivées sur des boîtes gélosées contenant 1,3% (p/v) d’agar, et complémentées avec 1mL de milieu nutritif Conway dont la composition, fournie par l’IFREMER est détaillée dans le tableau suivant (**Tableau 3**) ou du milieu nutritif F/2 (**Tableau 3**). La zeocine (ZeocinTM, Invitrogen) (75µg/mL) a été ajoutée afin de sélectionner les clones transformés ou de maintenir la pression de sélection sur les cellules de microalgues. Les boîtes ont ensuite été placées dans une armoire de culture sous une radiation lumineuse de 50 µE.m⁻².s⁻¹ et à 19°C avec une alternance jour/nuit de 16h/8h. Les cellules sont incubées pendant 15 jours ou plus dans ces conditions avant d’être utilisées.

Tableau 3 : Composition des milieux Conway et F/2 utilisés pour cultiver les différentes souches de *P. tricornutum*

	Milieu Conway			Milieu F/2			
	Souche sauvage	Souche transformée		Souche sauvage	Souche transformée		
Solution principale	NaNO ₃	100g	-	NaNO ₃	75g	-	
	NH ₄ Cl	-	80,25g	NH ₄ Cl	-	80g	
	NaH ₂ PO ₄ , H ₂ O		26g	NaH ₂ PO ₄ , H ₂ O		5g	
	H ₃ BO ₃		33,6	Na ₂ SiO ₃ , 9H ₂ O		30g	
	Na ₂ EDTA, 2H ₂ O		45g	Eau distillée		1L	
	FeCl ₃ , 6H ₂ O		1,28g	Trace métaux		1mL	
	MnCl ₂ , 4H ₂ O		0,36g	Solution vitaminique		1mL	
	Eau distillée		1L	Trace métaux	FeCl ₃ , 6H ₂ O		3,15g
	Trace métaux		1mL		Na ₂ EDTA, 2H ₂ O		4,36g
	Solution vitaminique		50mL		CuSO ₄ , 5H ₂ O		0,0098g
Trace métaux	ZnCl ₂ , 7H ₂ O		0,021g		Na ₂ MoO ₄ , 2H ₂ O		0,0063g
	CoCl ₂ , 6H ₂ O		0,02g		ZnSO ₄ , 7H ₂ O		0,022g
	(NH ₄) ₆ Mo ₇ O ₂₄		0,009g	CoCl ₂ , 6H ₂ O		0,01g	
	CuSO ₄ , 5H ₂ O		0,02	MnCl ₂ , 4H ₂ O		0,180g	
	Eau distillée		1mL	Eau distillée		1L	
Solution vitaminique	Cyanocobalamin (B12)		0,1g	Solution vitaminique	Cyanocobalamin (B12)		0,0005g
	Thiamine HCl (B1)		0,005g		Biotin (H)		0,0005g
	Eau distillée		50mL		Thiamine HCl (B1)		0,1g
				Eau distillée		1L	

2. En milieu liquide

Les cellules de *P. tricornutum* ont été cultivées dans le milieu « eau de mer » (EDM) contenant 33,3g/L de sel marin (Instant Ocean®, Truffaut) filtré avec un filtre préalablement autoclavé pendant 1h à 2bars, dont la taille des pores était de 0,22µm. Puis, le filtrat a été complété avec 1mL par litre de culture de milieu Conway (**Tableau 3**). Deux mL de metasilicate de sodium à 40g/L par litre de culture ont également été ajoutés. Les cellules ont été ensuite placées dans une armoire de culture sous agitation à 150rpm (révolutions par minute), sous une irradiation lumineuse de 50µE.m⁻².s⁻¹, à 19°C avec une alternance jour/nuit de 16h/8h. Les temps de culture ont varié de 3 à 10 jours suivant les expériences réalisées. La densité cellulaire à 10 jours, obtenue à partir du compteur cellulaire, atteignait environ 2,5x10⁷ cellules.mL⁻¹.

B. *Chlamydomonas reinhardtii*

Les souches de *C. reinhardtii* ont été mises en culture dans des volumes de 200mL dans un erlenmeyer autoclavé pendant 1h à 120°C sous une pression de 2bars. Les souches ont été cultivées dans un milieu nutritif composé de Tris Acétate Phosphate (TAP) (**Tableau 4**) préalablement filtré et autoclavé pendant 1h à 2bar.

Les cultures ont été mise en culture pendant 4 jours sous une température de 23°C avec une intensité lumineuse de $150\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ en lumière continue sous une agitation de 150rpm. La densité cellulaire à 4 jours atteignait environ 2×10^7 cellules.mL⁻¹.

C. Cultures cellulaires de bactéries

Les souches d'*E. coli* DH5 α (Invitrogen, 18265-017) utilisées lors des étapes de clonage ont été cultivées dans du milieu Lysogeny Broth (LB) (Sigma L-2897) à une concentration de 20g par litre d'eau milliQ et contenant 100 $\mu\text{g}\cdot\text{mL}^{-1}$ d'ampicilline pendant une nuit à 37°C sous une agitation de 150rpm. Les souches d'*E. coli* cultivées sur des boîtes de pétri ont été déposées sur un milieu gélosé contenant 20g de LB par litre d'eau milliQ, 1,2% (p/v) d'agar et contenant 100 $\mu\text{g}\cdot\text{mL}^{-1}$ d'ampicilline. Les milieux liquides ou gélosés ont été préalablement autoclavés pendant 1h à 2bars.

Tableau 4 : Composition du milieu Tris Acétate Phosphate (TAP) (selon les conditions décrites dans la thèse de Pierre-Louis Lucas, 2019)

TAP (1X) 1 L			
Produit		Quantité	
Tris base		2,42 g	
Beijerink		25 mL	
Éléments trace		1 mL	
Acide acétique		1 mL	
Tampon phosphate		0,375 mL	
H ₂ O		Qsp 1 L	

Beijerink (40X) 1 L		Éléments trace(200X) 1 L	
Produit	Quantité	Produit	Quantité
NH ₄ Cl	16 g	Na ₂ EDTA	50 g
CaCl ₂ , 2H ₂ O	2 g	ZnSO ₄ , 7H ₂ O	22 g
MgSO ₄ 7H ₂ O	4 g	H ₃ BO ₃	11,4 g
H ₂ O	Qsp 1 L	MnCl ₂ , 4H ₂ O	5,06 g
Tampon phosphate 100 mL		CoCl ₂ , 6H ₂ O	1,61 g
Produit	Quantité	CuSO ₄ , 5H ₂ O	1,57 g
K ₂ HPO ₄	16 g	(NH ₄) ₆ Mo ₇ O ₂₄ ,4H ₂ O	1,10 g
KH ₂ PO ₄	4 g	FeSO ₄ ,7H ₂ O	4,99 g
H ₂ O	Qsp 100 mL	H ₂ O	Qsp 1 L

III. Méthodes de biologie moléculaire

A. Préparation plasmidique

1. Transformation dans des cellules d' *E. coli*

Les plasmides Pha-NR contenant les gènes d'intérêt codant la GnT I-V5 ou la GalT-HA étaient déjà disponibles au labo au démarrage de mon projet de thèse. L'amplification et la purification de ces plasmides a alors été réalisée comme suit. Une première transformation est réalisée en insérant le plasmide d'intérêt dans des bactéries *E. coli*. Brièvement, 50µL d'une aliquote de cellules DH5α™ sont ajoutés à environ 1 à 10ng de plasmide puis incubés pendant 30min dans de la glace. Ensuite, la préparation contenant les bactéries et les plasmides est placée à 42°C pendant 20 secondes, afin de réaliser un choc thermique permettant aux plasmides de rentrer à l'intérieur des bactéries. Puis, la préparation est de nouveau placée dans de la glace pendant 2 minutes. Ensuite, 950µL de milieu LB préchauffés à 37°C sont ajoutés à la préparation et incubés pendant 1h à la même température sous une agitation à 150rpm. Puis, 100µL de la préparation sont étalés en conditions stériles sur un milieu gélosé LB contenant 100µg/mL d'ampicilline. Le milieu gélosé ainsiensemencé est ensuite mis en culture pendant une nuit à 37°C. Les cellules ayant poussé grâce à l'intégration du plasmide Pha-NR leur conférant une résistance à l'ampicilline forment des colonies isolées.

2. Amplification et purification plasmidique

Une colonie isolée est prélevée stérilement sur la gélose et reprise dans 2mL de milieu LB contenant 100µg/mL d'ampicilline. La pré-culture obtenue est ensuite incubée pendant une demi-journée à 37°C sous une agitation de 150rpm. Ensuite, 1mL de cette pré-culture est prélevé puis versé dans 100mL de milieu LB. La culture est ensuite incubée pendant une nuit à 37°C sous une agitation à 150rpm.

La purification des plasmides est réalisée le lendemain matin en utilisant le kit NucleoBond® Xtra plasmid purification produit par Macherey-Nagel. Tous les tampons cités dans cette partie sont issus du kit. Tout d'abord, la culture de 100mL est centrifugée à 4 000xg pendant 10min. Le surnageant est éliminé et le culot bactérien est repris dans 8mL de tampon

RES + RNase A puis transféré dans un tube de 50mL. Huit mL de tampon LYS sont ajoutés. Ce dernier est mélangé 5 fois de suite par retournement et incubé à température ambiante pendant 5 minutes. Au terme de cette incubation, 8mL du tampon NEU sont ajoutés puis mélangés aussitôt jusqu'à la disparition de la couleur bleue de la solution. Le mélange contenant le culot bactérien lysé est ensuite appliqué sur une colonne NucleoBond®Xtra Column préalablement équilibrée avec 12mL du tampon EQ fourni dans le kit. Deux lavages successifs de la colonne sont ensuite réalisés avec respectivement 5mL de tampon EQU et 8mL de tampon WASH. Les plasmides retenus par la colonne NucleoBond®Xtra Column sont alors élués dans un tube Falcon de 15mL contenant préalablement 5mL du tampon ELU du kit. La précipitation des plasmides est ensuite réalisée. Pour cela, 3,5 mL d'isopropanol sont ajoutés à l'éluât obtenu précédemment, homogénéisés par vortex puis laissés reposer pendant 2min. Le tout est passé sur un filtre NucleoBond® Finalizer, puis lavé avec 2mL d'éthanol 70% (v/v). Le filtre est séché en lui injectant de l'air à 6 reprises grâce à la seringue fournie. Pour finir, l'ADN plasmidique est élué une première fois avec 400µL d'eau DNase/RNase free dans un tube eppendorf de 2mL. La seconde élution du filtre est réalisée avec les 400µL d'éluât obtenus lors de la première élution, dans le but de concentrer l'ADN plasmidique.

3. Dosage de l'ADN plasmidique

L'ADN plasmidique est dosé en utilisant un Thermo Scientific™ NanoDrop™ One^C Microvolume UV-Vis Spectrophotometer. Un µL d'échantillon est chargé sur l'appareil permettant de déterminer la concentration d'ADN de l'échantillon. Le ratio 260/280nm détermine la pureté de l'échantillon. Ce ratio doit être proche de 1,8 dans le cas d'une extraction d'ADN et de 2,0 dans le cas d'une extraction ARN.

4. Digestion enzymatique des plasmides

La digestion enzymatique des plasmides est réalisée afin de vérifier la qualité de la purification réalisée. Les enzymes de restriction BamHI et HindIII ainsi que le tampon E fournis par Promega™ ont été utilisés pour cette vérification. La digestion enzymatique des plasmides a été réalisée dans un tube Eppendorf de 1,5mL à partir des quantités données dans le tableau suivant (**Tableau 5**). Le mélange est incubé pendant une nuit à 37°C.

Tableau 5 : Composition des milieux utilisés lors des digestions enzymatiques effectuées lors du contrôle de la purification plasmidique

	Volume/quantité à ajouter
ADN plasmidique	5µg
Tampon de restriction E (Promega™)	2µL
Enzyme de restriction 1 BamHI	1µL
Enzyme de restriction 2 HindIII	1µL
Qsp 20µL H ₂ O	11,2µL

A la fin de l'incubation, 4µL de tampon de chargement (Thermo Scientific, #R0611) sont ajoutés au mélange de digestion puis ce milieu est chargé sur un gel préparé avec 0,5% (v/v) de tampon Tris Acetate EDTA (TAE) contenant 1% (p/v) d'agarose et 1% (p/v) de SafeView™. Le profil du gel d'agarose est ensuite obtenu sur l'appareil de révélation UV (Ebox, Vilber).

B. Transformation génétique de *Phaeodactylum tricornutum* par biolistique

1. Culture de la diatomée et étalement sur boîtes gélosées

Une culture de *P. tricornutum* a été réalisée dans un milieu à base d'eau de mer (EDM) contenant 40g/L de sel marin (Tropico Salt, SIGMA) récemment autoclavé à 1 bar pendant environ 1 heure puis complété de 1mL/L d'un milieu nutritif F/2 préalablement préparé (**Tableau 3**). Les microalgues ont été cultivées à 23°C sous une lumière continue jusqu'à l'obtention d'une densité cellulaire d'environ 2 000 000 de cellules par mL. Les cellules ont été comptées au microscope optique droit avec une cellule de Nageotte. Environ 10⁸ cellules ont été reprises dans 150µL de milieu EDM puis étalées au centre d'une boîte gélosée contenant 20g/L de sel marin (Tropico Salt, SIGMA), 10% (p/v) d'agar et complétée avant solidification, avec 1mL/L de milieu F/2 préalablement préparé.

2. Préparation de la biolistique et transformations des cellules de *P. tricornutum*

Le milieu d'incubation nécessaire à la biolistique est réalisée avant la transformation des microalgues à 4°C. Ainsi, 1,5 ou 3µg d'ADN plasmidique, 50µL de CaCl₂ 2,5M et 20µL de spermidine 0,1M produite par SIGMA® (référence S2626) sont ajoutés successivement et homogénéisés au vortex avec 50µL de billes d'or (Biorad Gold macrocarriers 0,5µm, référence : 16522167) dans un tube Eppendorf de 1,5mL permettant l'obtention d'une solution appelée « mix plasmidique ». Une centrifugation à 13 000xg est ensuite réalisée pendant 2 minutes et le surnageant est éliminé. Deux lavages successifs du culot sont réalisés avec 250µL d'éthanol à 100% (v/v). Entre chaque lavage, une centrifugation à 13 000xg est effectuée et le surnageant est éliminé. Au final, le culot obtenu contenant le complexe bille d'or/plasmide est repris dans 60µL d'éthanol à 100% (v/v). La transformation des diatomées est réalisée avec un appareil à biolistique PDS-1000/He™ System produit par Biorad® (**Figure 24**). Pour cela, des grilles, des disques de rupture psi 1200, macrocarriers et des supports Holder sont nettoyés dans de l'éthanol à 100% (v/v) puis laissés sécher sous une hotte à flux laminaire (Figure 23). Le « mix plasmidique » précédemment préparé (20µL) est déposé sur un macroporteur et laissé sécher (Figure 23). Par la suite, les différents éléments sont assemblés dans la chambre principale de l'appareil à biolistique. La transformation est réalisée sous vide sur des cellules précédemment étalées sur des boîtes gélosées dont la composition est décrite plus haut. Après la transformation, les boîtes gélosées sont mises dans le noir à 19°C enveloppées dans du papier aluminium pendant 48 heures.

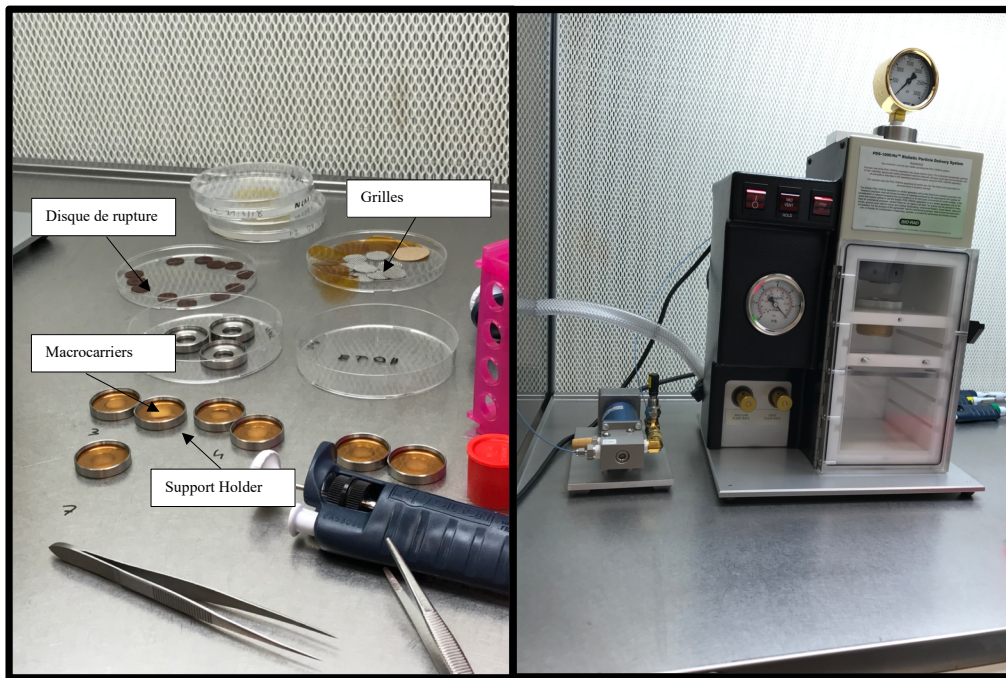


Figure 24 : Photos représentant le matériel nécessaire à une transformation génétique chez *P. tricornutum* par biolistique

A gauche, les différents éléments nécessaires à la réalisation d'une transformation par biolistique. A droite, l'appareil à biolistique PDS-1000/He™ System (Biorad).

3. Étape de sélection des clones après la transformation

Les boîtes gélosées issues de la transformation sont récupérées. Les cellules contenues sont reprises dans 100µL de milieu à base d'eau de mer à 100% (v/v) puis étalées sur une nouvelle boîte gélosée contenant 1,3% (p/v) d'agar et 20g/L de sel marin (Instant Ocean®, Truffaut) ainsi que 75µg/mL de zéocine afin de sélectionner les clones ayant inséré le plasmide Pha-NR contenant une résistance à la zéocine. Les boîtes gélosées sont incubées à 19°C sous une irradiation de 50 µE.m⁻².s⁻¹ pendant 3 semaines jusqu'à l'apparition de colonies isolées. Les clones isolés sont ensuite de nouveau repiqués sur des boîtes gélosées contenant 1,3% (v/v) d'agar et 33,3g/L de sel marin, contenant 75µg/mL de zeocine, 1mL/L de milieu Conway et 2mL/L d'une solution de metasilicate de sodium à 40g/L.

C. Criblage génique des clones isolés

1. Extraction de l'ADN des clones isolés

Les clones isolés sont repris dans 2mL de milieu EDM 100% (33,3g/L de sel marin) avec 1mL/L de milieu Conway et 2mL/L d'une solution de metasilicate de sodium à 40g/L. Le tout est laissé sous une agitation à 150rpm à 19°C et sous une irradiation de $50 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ pendant 3 jours. La culture est ensuite centrifugée à 4 500xg pendant 10min et le surnageant est éliminé. Puis, le culot cellulaire est repris dans 200 μL d'eau milliQ. Une nouvelle centrifugation est réalisée à 4 500g pendant 10min et le surnageant est de nouveau éliminé. Le culot est finalement repris dans 30 μL d'eau milliQ et chauffé à 95°C pendant 10min afin de lyser les cellules. Le broyat cellulaire est centrifugé à 4 500g afin de sédimenter les débris cellulaires. Le surnageant, contenant l'ADN, est stocké dans un tube Eppendorf de 1,5mL et congelé à -20°C.

2. Amplification génique par PCR (Polymerase Chain Reaction)

Un mix PCR est préparé suivant le tableau ci-dessous (**tableau 6**). L'enzyme polymérase utilisée est la GoTAq® Polymérase (M300) produite par Promega®. Chaque tube est placé dans un thermocycleur Sensoquest labcycler gradient (référence : 049963) et soumis à différentes températures selon les températures d'hybridation optimales des amorces utilisées (**tableau 2**).

3. Électrophorèse des ADN

Dix μL d'extraits issus des PCR sont analysés par une électrophorèse sur gel d'agarose préparé avec 0,5% (v/v) de tampon Tris Acetate EDTA (TAE) contenant de 1% à 2% (p/v) d'agarose et 1% (p/v) de SafeView™. Le profil électrophorétique du gel d'agarose est ensuite révélé sur l'appareil de révélation UV (Ebox, Vilber).

Tableau 6 : Composition du mix PCR utilisé pour l'amplification du gène d'intérêt

Composition	Concentration stock	Concentration finale	Volume pour 20µL
Eau	-	-	12,5µL
Tampon GoTAq	5X	1X	4µL
dNTP mix	2,5mM	0,25mM	1,6µL
Amorce Forward	100mM	10mM	0,4µL
Amorce reverse	100mM	10mM	0,4µL
Polymérase	-	-	0,1µL
Matrice ADN	-	-	1µL
TOTAL			20µL

D. Criblage des transcrits des clones isolés

1. Extraction de l'ARN des clones isolés

Les cellules (l'équivalent de 5mL) sont récoltées par une centrifugation à 4 500xg pendant 10min à 19°C en utilisant une centrifugeuse Beckman® Allegra X15 R. Le culot cellulaire est repris dans 1mL de Nucléozol. Le mélange est ensuite transféré dans un tube Lysing Matrix E produit par MP biomedicalTM puis les cellules sont broyées dans un broyeur à billes MP BiomedicalsTM Instrument FastPrep-24TM 5G, 4 fois à 6,5m.s⁻¹. Le tube est ensuite laissé reposer à température ambiante pendant 5min puis centrifugé à 12 000xg pendant 5min. Le surnageant est transféré dans un tube de 2mL RNase free. Un volume de 400µL d'H₂O RNase free est ajouté au surnageant. Le tout est mélangé par retournement puis incubé à température ambiante pendant 15min. Le mélange est ensuite centrifugé à 16 000xg pendant 15min à 4°C. Au terme de cette centrifugation, le surnageant est récupéré puis transféré dans un nouveau tube de 2mL RNase free. Un volume de 900µL d'éthanol 100% (Biology molecular Grade) est ajouté au surnageant précédemment récupéré puis homogénéisé par pipetage. 700µL du mélange est ensuite transféré sur la colonne bleue issue du kit NucleospinTM RNAplus (Macherey-Nagel, 15300205). Un tube de collection fourni dans le kit est placé sous

la colonne. Une fois déposée, la colonne est centrifugée pendant 30s à 8 000xg. Le tube de collection est éliminé et la colonne bleue est placée sur un nouveau tube de collection. Un volume de 200µL de la solution WB1 est ajouté à la colonne puis centrifugé 30s à 11 000xg. Deux lavages sont ensuite réalisés avec la solution WB2, une première fois avec 600µL avec une centrifugation de 11 000xg pendant 30s. Le second lavage avec 250µL avec une centrifugation de 2min de 11 000xg.

2. Élué des ARN de la colonne

Un volume de 30µL d'H₂O RNase free est ajouté au centre de la membrane de la colonne bleue puis laissé reposer pendant 1min à température ambiante, un tube de collection est placé sous la colonne bleue contenant les ARNs. Le tout est centrifugé pendant 1min à 11 000xg. Le volume de 30µL récupéré contenant les ARNs est de nouveau déposé sur la colonne puis centrifugé pendant 1min à 11 000xg. Cette étape permet d'éluer deux fois les ARNs présents avec le même volume (30µL) de façon à obtenir une concentration plus importante en ARN.

3. Élimination d'ADN résiduels par une réaction de DNase

La réaction est réalisée à partir du kit turbo DNase-free™ (ThermoFisher, AM1907). Un mélange contenant 3µL de tampon 10X, 1µL de turbo DNase free et 26µL de la solution précédemment éluée contenant les ARNs est incubé pendant 30min à 37°C. Au terme de ces 30min, 3µL de DNase inactivation est ajouté. Le tout est bien homogénéisé puis incubé 5min à température ambiante pendant 5min. Une centrifugation à 10 000xg pendant 1,5min est ensuite réalisée. Le surnageant obtenu est récupéré dans un nouveau tube de 2mL RNase free.

4. Synthèse d'ADNc à partir d'ARN et amplification d'ADNc par PCR

Une réaction de reverse transcriptase (RT) est réalisée à partir de 2µg d'ARN, dosé dans les conditions identiques présentées dans le chapitre (III. A. 3.). Pour cela, le protocole issu du kit High-capacity cDNA Reverse Transcription (Applied biosystem™, 4368814) est appliqué. Cette réaction permet de convertir de l'ARN extrait en ADN complémentaire (ADNc).

Cet ADNc est ensuite amplifié puis déposé sur gel d'ADN dans les mêmes conditions décrites dans le chapitre (III. C. 2. et 3.).

IV. Méthodes de biochimie

A. Extraction des oligosaccharides-PP-Dolichol (LLO)

1. Extraction microsomale

Les cellules sont récoltées par une centrifugation à 4 500xg pendant 10min à 19°C en utilisant une centrifugeuse Beckman® Allegra X15 R. Le culot cellulaire est repris dans 1 mL de tampon Tris-HCl 0,1M pH 7,2 contenant une tablette d'inhibiteurs de protéases SIGMAFAST™ Protease Inhibitor Cocktail Tablets, EDTA-Free produite par Sigma. Le mélange est ensuite transféré dans un tube Lysing Matrix E produit par MP biomedicals™ puis les cellules sont broyées dans un broyeur à billes MP Biomedicals™ Instrument FastPrep-24™ 5G, 3 fois pendant 30 secondes à 6,5m.s⁻¹. Le broyat est ensuite centrifugé à 4 500xg pendant 10min. Le surnageant est gardé et le culot contenant les débris cellulaires est éliminé. Les étapes de broyage, de centrifugation sont répétées pendant 3 fois. Les surnageants sont regroupés ensemble puis centrifugés à 20 000xg pendant 30 minutes à 4°C dans une ultracentrifugeuse Beckman® Optima XPN-80. Cette première étape permet de sédimenter majoritairement la fraction plastidiale des cellules de la diatomée. Ensuite, le surnageant obtenu est récupéré puis de nouveau centrifugé à 100 000xg pendant 1 heure à 4°C. Le culot obtenu correspond à la fraction microsomale qui est reprise dans le tampon adapté à la suite de l'expérimentation.

2. Extraction des LLO

L'extraction des LLO a été effectuée selon le protocole de **Garénaux *et al.*, 2008**. Les culots microsomaux précédemment obtenus sont repris dans des tubes en verre dans un mélange contenant 400µL d'eau (LC-MS grade, Fisher Chemical), 800µL de méthanol (LC-MS grade, Fisher Chemical) ainsi que 1,2mL de chloroforme (LC-MS grade, Fisher Chemical). Les tubes sont homogénéisés au vortex puis centrifugés à 4 000xg pendant 2min. Deux phases sont visibles, séparées par une interphase. La phase inférieure colorée est éliminée. Cette étape permet d'éliminer les pigments.

Trois mL de mélange A sont ensuite ajoutés dans les tubes en verre. Ce mélange A est préparé à partir de chloroforme/méthanol/MgCl₂ 4mM dans de l'eau (LC-MS grade, Fisher Chemical) (86/14/1) (v/v/v). Après une homogénéisation et une centrifugation dans les mêmes conditions que précédemment, la phase inférieure est de nouveau éliminée. Cette étape est répétée pendant 3 fois. Trois mL de mélange B contenant du méthanol/MgCl₂ 100mM dans de l'eau/chloroforme (16/15/1) (v/v/v) ont été ajoutés dans les tubes en verre. Après une homogénéisation et une centrifugation dans les mêmes conditions que précédemment décrites, la phase supérieure, contenant les protéines est cette fois-ci éliminée. Cette étape est répétée pendant 3 fois. Après 5 minutes de séchage des culots sous une rampe à air, les LLO ont été extraits par l'ajout de 3mL du mélange C contenant du méthanol/chloroforme/eau (10/10/3) (v/v/v). Après une centrifugation à 4 000xg pendant 2min, le surnageant est gardé et placé dans un nouveau tube en verre. Cette étape est répétée pendant 3 fois. Les surnageants récupérés ont été séchés sous une rampe à air pendant environ 2 heures.

3. Hydrolyse acide

Un mL d'acide trifluoroacétique (TFA) (LC-MS grade, Thermo Scientific) à 0,1M est ajouté dans les tubes en verre contenant les LLO. Ces tubes sont ensuite incubés pendant 1 heure à 80°C afin de libérer l'oligosaccharide du LLO. Les échantillons sont ensuite lyophilisés et repris dans 500µL d'eau (LC-MS grade, Thermo Scientific).

4. Purification des oligosaccharides sur colonne carbograph

Les oligosaccharides ont été purifiés par une élution sur des colonnes carbograph (Hypersep hypercarb, Thermo Fisher) de 3mL. Les colonnes sont au préalable conditionnées sous une pression de 5 mmHg avec une enceinte manifold (Phenomenex SPE vacuum manifold). Trois mL de NaOH 1M (LC-MS grade, Fisher Chemical) ont tout d'abord été déposés sur les colonnes. Six mL d'H₂O (LC-MS grade, Fisher Chemical) sont ensuite ajoutés sur les colonnes. Trois mL de CH₃COOH 30% (v/v) suivi de 3mL d'eau sont ajoutés. Les colonnes sont ensuite conditionnées dans 3mL d'acétonitrile (LC-MS grade, Fisher Chemical) à 50% suivis de 6mL d'acétonitrile à 5% (v/v) (LC-MS grade, Fisher Chemical). La colonne est ensuite rincée avec 6mL d'eau (LC-MS grade, Fisher Chemical). La totalité des échantillons, repris dans 500µL

d'eau, est ajoutée à la colonne puis rincée avec 3mL d'eau (LC-MS grade, Fisher Chemical) puis 3mL d'acétonitrile à 5% (v/v) (LC-MS grade, Fisher Chemical). Les échantillons sont élués à la pression atmosphérique dans des tubes en verre avec 4x500µL d'acétonitrile à 50% (v/v) (LC-MS grade, Fisher Chemical) contenant 0,1% (v/v) de TFA. Les échantillons sont ensuite lyophilisés.

5. Perméthylation des oligosaccharides

Environ 2g d'hydroxyde de sodium (NaOH pellet, Fisher Chemical) broyés sont repris dans 3mL de DMSO anhydre (SIGMA-Aldrich). Ce mélange (1,2mL) est ensuite ajouté aux oligosaccharides purifiés, suivi de 1,2mL d'iodométhane (SIGMA-Aldrich). Les tubes sont homogénéisés au vortex puis incubés sous une agitation rotative à 15rpm pendant 2 heures. Un mL d'eau est ajouté afin de stopper la réaction. Un mL de chloroforme suivi de 5mL d'eau sont ajoutés permettant d'obtenir un mélange biphasique. Ce mélange est homogénéisé au vortex puis laissé reposer quelques secondes de façon à laisser apparaître clairement deux phases. La phase supérieure contenant de l'eau et du NaOH est éliminée. Cette étape est recommencée plusieurs fois jusqu'à ce que le pH de la phase aqueuse soit autour de 7.

Les échantillons contenus dans la phase contenant le chloroforme sont séchés sous une rampe à air puis repris dans 200µL de méthanol à 80% (v/v) (LC-MS grade, Fisher Chemical).

6. Purification des oligosaccharides perméthylés sur la colonne C18

Les oligosaccharides préalablement perméthylés sont purifiés sur une colonne C18 de 3mL (Hypersep C18, Thermo Fisher). Les colonnes sont d'abord conditionnées dans 5mL de méthanol à 100% (v/v) (LC-MS grade, Fisher Chemical), suivis successivement de 5mL d'eau (LC-MS grade, Fisher Chemical), de 5mL d'acétonitrile (LC-MS grade, Fisher Chemical) et enfin à nouveau de 5mL d'eau (LC-MS grade, Fisher Chemical). Les oligosaccharides perméthylés lyophilisés sont ensuite repris dans 200µL de méthanol à 80% (v/v) puis déposés sur la colonne C18. Celle-ci est alors rincée avec 5mL d'eau (LC-MS grade, Fisher Chemical). Les oligosaccharides perméthylés, retenus sur la phase stationnaire, sont élués dans des tubes Eppendorf de 2mL différents avec successivement 4x0,5mL d'une solution à 15% (v/v) d'acétonitrile, 35% d'acétonitrile (v/v), 50% d'acétonitrile (v/v) et 75% d'acétonitrile (v/v) (LC-

MS grade, Fisher). Les fractions éluées à 50% et à 75% (v/v) d'acétonitrile contenant les oligosaccharides perméthylés sont lyophilisées puis analysées par spectrométrie de masse, respectivement.

B. Criblage protéique des clones isolés

1. Préparation rapide de la fraction microsomale

Ce protocole est adapté de l'article **Abas and Luschnig, 2010**. Brièvement, les cellules des diatomées sont cultivées pendant 7 jours sous les conditions décrites en section (II. A. 2) en utilisant un milieu Conway contenant une source azotée de type NH_4Cl (**Tableau 3**). Au 6^{ème} jour, 980 μM d'une solution NaNO_3 ont été ajoutés permettant l'expression du gène PtGnT I-V5. Au 7^{ème} jour, les cellules sont ensuite récoltées par une centrifugation à 4 500xg pendant 10min à 19°C en utilisant une centrifugeuse Beckman® Allegra X15 R. Le culot cellulaire est repris dans 150 μL de tampon EB contenant 100mM de Tris-HCl pH 7.5, 5% glycérol (v/v), 25% de saccharose (p/v), 10mM EDTA, 5mM KCl et une capsule d'inhibiteurs de protéases SIGMAFAST™ Protease Inhibitor Cocktail Tablets, EDTA-Free produite par SIGMA. La solution contenant les cellules est placée dans un tube de type Lyzing Matrix D produit par MP biomedicals™. Puis, les cellules sont broyées dans un broyeur à billes MP Biomedicals™ Instrument FastPrep-24™ 5G, pendant 3 fois pendant 1 minute à 6,5m/s. Le contenu du tube est ensuite récupéré puis incubé avec 20 μL d'une solution de polyvinylpolypyrrolidone (PVPP) (1mg de PVPP pour 20 μL de tampon EB). Ce mélange est homogénéisé par vortexage puis placé dans de la glace pendant 5min. Le tout est ensuite centrifugé pendant 3min à 700xg à 4°C afin de sédimenter le PVPP. Cent μL du surnageant obtenu sont mélangés avec 100 μL d'eau milliQ puis centrifugés pendant 1 heure 30 à 21 000xg à 4°C. Pour finir, le surnageant obtenu est éliminé. Le culot obtenu correspond à une fraction enrichie en microsome appelée par la suite « microsomale-like fraction ». Les échantillons sont ensuite analysés par Western Blot suivant le protocole décrit ci-dessous (du IV. B. 3. Jusqu'au IV. B. 11.).

2. Extraction des protéines totales

Les cellules sont récoltées par une centrifugation à 4 500xg pendant 5min à 19°C en utilisant une centrifugeuse Beckman® Allegra X15 R. Le culot cellulaire est repris dans 1 mL de

tampon Tris-HCl 0,1M pH 7,2 contenant une tablette d'inhibiteurs de protéases SIGMAFAST™ Protease Inhibitor Cocktail Tablets, EDTA-Free produite par SIGMA. Le mélange est ensuite transféré dans un tube Lysing Matrix E produit par MP biomedical™ puis les cellules sont broyées dans un broyeur à billes MP Biomedicals™ Instrument FastPrep-24™ 5G, 3 fois pendant 30s à 6,5m/s. Le broyat est ensuite centrifugé à 4 500xg pendant 5min. Le surnageant est gardé et le culot éliminé. Les étapes de broyage, de centrifugation et de conservation du surnageant sont répétées pendant 3 fois. Les trois surnageants issus des 3 centrifugations successives sont regroupés.

3. Migration sur gel d'électrophorèse

Les protéines sont déposées sur un gel de type NuPAGE™ 4-12% Bis-Tris Protein Gels acheté chez ThermoFisher® ou sur un gel préparé au laboratoire dont la composition est décrite dans le tableau suivant (**Tableau 7**). Le gel de séparation des gels préparés au laboratoire est coulé dans une cassette vide (ThermoFisher®) puis laissé polymériser en ajoutant une couche d'isopropanol sur le dessus afin de permettre une polymérisation homogène. Par la suite, le gel de concentration est préparé puis placé également dans la même cassette en ajoutant un peigne avec un nombre adapté de puits (**Tableau 7**).

Tableau 7 : Composition d'un gel Bis-Tris 4-12%

	Gel de séparation (mL)	Gel de concentration (mL)
H ₂ O	2,15	1,45
40% Acrylamide / Bis-Acrylamide (29 :1) (p/p) (Fisher)	1,5	0,34
1,5 M Tris pH 8,8	1,25	-
1,5 M Tris pH 6,8	-	0,17
10% SDS (p/v)	0,05	0,02
10% APS (p/v)	0,05	0,02
TEMED	0,002	0,002

Lors du chargement des protéines sur le gel électrophorétique, 2,5µL de protéines de calibration de masses moléculaires connues (PageRuler™ Prestained Protein Ladder, Fisher Scientific) sont ajoutés dans un des puits du gel afin de calibrer la migration du gel

d'électrophorèse et de pouvoir estimer la masse moléculaire des protéines ayant migré (**Figure 25**). La migration électrophorétique des protéines est alors effectuée durant environ 1 heure sous une tension constante de 180V. A terme, les protéines sont soit colorées au bleu de Coomassie, soit transférées sur une membrane de nitrocellulose.

Gel type	Tris-Glycine						Tris-Acetate*		Bis-Tris*						
Gel concentration	4-20%	8-16%	10-20%	8%	10%	12%	15%	3-8%	7%	4-12%		10%		12%	
Running buffer	Tris-Glycine						Tris-Acetate		MOPS	MES	MOPS	MES	MOPS	MES	
Apparent Molecular Weights, kDa															
% length of gel	10														
	20	180	180	180	180	180	180	130	130	150	140	140	140	140	140
	30	130	130	100	130	100	70	100	100	120	115	115	115	115	115
	40	100	100	70	100	70	55	70	70	80	80	80	80	80	80
	50	70	70	55	70	55	40	55	55	65	65	65	65	65	65
	60	55	55	40	55	40	35	40	40	50	50	50	50	50	50
	70	40	40	35	40	35	25	35	35	40	40	40	40	40	40
	80	35	35	25	35	25	15	25	25	30	30	30	30	30	30
	90	25	25	15	25	15	10	15	15	20	20	20	20	20	20
	100	15	15	10	15	10	10	10	10	15	15	15	15	15	15

Figure 25 : Masses moléculaires des protéines de calibration utilisées pour l'électrophorèse SDS-PAGE (PageRuler™ Prestained Protein Ladder, Fisher Scientific)

4. Coloration des protéines au bleu de Coomassie

Les protéines présentes sur le gel d'électrophorèse sont colorées au bleu de Coomassie R250 (45% d'éthanol (v/v), 10% d'acide acétique (v/v) contenant 0,003% (p/v) de Coomassie Brilliant Blue R250 Thermo Scientific) pendant 1 heure à température ambiante sous une agitation de 50rpm. Le gel est ensuite décoloré une nuit sous une agitation de 50rpm et à la température ambiante dans une solution de fixation/décoloration contenant 20% d'éthanol (v/v) et 10% d'acide acétique (v/v).

5. Coloration des protéines au nitrate d'argent

Les protéines présentes sur le gel d'électrophorèse ont été colorées au nitrate d'argent qui est une technique de révélation plus sensible que la coloration au bleu de Coomassie R250. Les protéines présentes dans le gel d'électrophorèse sont tout d'abord fixées pendant 2h avec

30% d'éthanol (v/v) (LC-MS grade, Fisher Chemical) et 10% d'acide acétique (v/v) sous une agitation de 15rpm. Le gel est ensuite lavé 2 fois pendant 10min avec une solution d'éthanol à 10% (v/v) puis rincé 3 fois pendant 10min dans de l'eau déionisée.

Le gel d'électrophorèse est ensuite baigné pendant 1min dans une solution contenant 0,02% de thiosulfate de sodium (p/v) puis rincé rapidement à l'eau déionisée avant d'être coloré avec une solution contenant 0,1% de nitrate d'argent (p/v) (Acros Organics) pendant 30 minutes. Finalement, le gel est rincé rapidement avec de l'eau déionisée puis incubé dans une solution contenant 1,2% de carbonate de sodium (p/v), 0,00008% de thiosulfate de sodium (p/v) (Fisher Chemical) ainsi que 0,0001% de formaldéhyde à 37% (v/v) (SIGMA-Aldrich) jusqu'à l'apparition des bandes protéiques colorées. La coloration est stoppée avec une solution contenant 1% d'acide acétique (v/v).

6. Transfert électrophorétique sur une membrane de nitrocellulose et révélation en western blot

Les protéines ont été transférées sur une membrane de nitrocellulose après leur migration sur un gel d'électrophorèse de type SDS-PAGE (**Figure 26**). Un « sandwich » comprenant 2 couches de papier Whatman, de part et d'autre de la membrane de nitrocellulose posée sur le gel d'électrophorèse contenant les protéines, est constitué. Chacun des éléments du montage est préalablement incubé pendant 10min dans la solution Pierce™ 1-Step Transfert Buffer (ThermoFisher). Ensuite, le transfert protéique est réalisé avec un courant constant de 2,0 A pendant 15min.

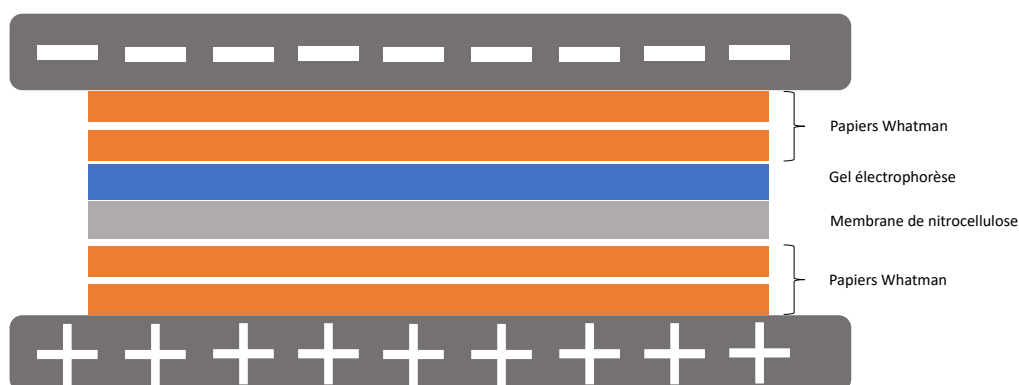


Figure 26 : Schéma représentant le montage nécessaire pour réaliser un transfert des protéines du gel d'électrophorèse sur une membrane de nitrocellulose

7. Vérification de la qualité du transfert par coloration au rouge Ponceau

Au terme du transfert, les protéines présentes sur la membrane de nitrocellulose sont colorées pendant 1min avec une solution contenant 1% de rouge ponceau (p/v) (Serra, C.I.27195) et 1% de TCA (p/v). La membrane est ensuite rincée à l'eau distillée. La membrane est ensuite saturée dans du TBST1X pendant une nuit. Le TBST1X est préparé avec 0,24% de Tris Base (m/v) (Fisher Scientific), 2,92% de NaCl (Fisher Scientific) (m/v) et 0,1% Tween20 (p/v) (Acros Organics) dans de l'eau milliQ.

8. Incubation de la membrane avec l'anticorps primaire

La membrane de nitrocellulose est incubée à la température ambiante dans une pochette plastique scellée avec l'anticorps primaire préalablement dilué dans du TBST1X. Les dilutions des différents anticorps primaires utilisés sont résumées dans le tableau ci-dessous (**Tableau 8**). L'incubation est réalisée pendant 3 heures afin de permettre aux anticorps de se lier aux antigènes. Quatre lavages avec 20mL de TBST1X sont réalisés sous vide à l'aide du Snap I.D® 2.0 de Millipore.

Tableau 8 : Récapitulatif des anticorps primaires utilisés durant ce projet de thèse

Anticorps primaire	Hôte	Dilution
Anti-V5 (PA1-993, invitrogen)	Lapin	1/2000 ^{eme}
Anti-HA (H6908, SIGMA Aldrich)	Lapin	1/2000 ^{eme}
Anti β 1,2 Xylose (AS07 267, Agrisera)	Lapin	1/5000 ^{eme}
Anti α 1,3 Fucose (AS07 268, Agrisera)	Lapin	1/15000 ^{eme}

9. Incubation de la membrane avec l'anticorps secondaire

La membrane est ensuite incubée avec un anticorps secondaire couplé à la peroxidase de raifort (HRP pour Horseradish peroxidase) dilué au 1/30 000^{eme} dans du TBST1X.

L'incubation est réalisée à la température ambiante sous une agitation de 50rpm pendant 1 heure.

La membrane est ensuite lavée pendant 3 fois avec 20mL de TBST1X puis pendant 1 fois avec 20mL de TBS1X sous vide à l'aide du Snap I.D® 2.0 de Millipore afin d'éliminer le tween pouvant gêner la réaction de chimioluminescence. Le TBS1X est préparé avec 0,24% de Tris Base (p/v) (Fisher Scientific) et 2,92% de NaCl (Fisher Scientific) (p/v) dans de l'eau milliQ

10. Immunodétection révélée par chimioluminescence avec le réactif PICO Plus

La membrane de nitrocellulose est incubée dans un réactif de chimioluminescence préparé à partir du kit PICO-plus (SuperSignal™ West Pico PLUS Chemiluminescent Substrate, Thermofisher Scientific) pour permettre la révélation de l'immunodétection. Le réactif est préparé en mélangeant 2mL de la solution de « peroxyde » avec 2mL de solution « luminol ». Ce mélange est ensuite réparti uniformément sur l'ensemble de la membrane de nitrocellulose puis laissé incuber pendant 5min.

11. Révélation par Imager Fusion-FX

La membrane de nitrocellulose est maintenant révélée sur l'Imager Fusion-FX (Vilber) et le logiciel associé Evolution CaptEdge. Brièvement, la membrane est égouttée puis placée sur un support noir. Ce support est inséré dans l'Imager. Le temps d'intégration sélectionné, variant de 10s à 2min, dépend de la quantité de matériel chargée et de la dilution de l'anticorps primaire utilisé. Une fois terminée, la photographie obtenue est sauvegardée et annotée.

C. Eastern Blot et galactosylation *in vitro*

1. Galactosylation *in vitro* des glycoprotéines de *P. tricornutum*

50µg de protéines totales de *P. tricornutum* sont incubés dans un tube Eppendorf de 2mL dans 200µL de tampon glycine 100mM contenant 4mM (p/v) de MnCl₂, 0,89mM (v/v) d'UDP-galactose et 2mU de la β(1,4)-galactosyltransférase bovine (G5507, Sigma). Ce mélange est incubé pendant 24h à 37°C. Le lendemain, le mélange est évaporé au Speedvac.

2. Migration sur gel électrophorétique et transfert sur une membrane de nitrocellulose

Cinquante μg de protéines totales sont déposés et séparés sur un gel d'électrophorèse SDS-PAGE puis transférés sur une membrane de nitrocellulose en suivant le protocole décrit dans la section (IV. B. 2, 3, 6 et 7.)

3. Incubation avec la lectine RCA120 marquée à la FITC

Après la saturation dans du TBST 1X, la membrane de nitrocellulose est incubée dans du TBST1X contenant $5\mu\text{g}$ par mL de lectine RCA120 couplée à la FITC (Fluorescein isothiocyanate) (Vector labs) pendant 2 heures à la température ambiante et à l'obscurité sous une agitation de 50rpm. La membrane est ensuite lavée 3 fois avec 20mL de TBST1X puis 1 fois avec 20mL de TBS1X (afin d'éliminer le tween pouvant gêner l'émission de fluorescence)

4. Révélation de la fluorescence sur Typhoon

La fluorescence émise à l'endroit où s'est fixée la RCA120 couplée à la FITC sur la membrane de nitrocellulose est révélée sur un appareil Typhoon™ FLA9000 (GE Healthcare) en utilisant une longueur d'onde d'excitation de 495nm et une longueur d'onde d'émission de 521nm.

D. Extraction glycanique du $\text{Man}_5\text{GlcNAc}_2$

1. Dénaturation des protéines et digestion à la PNGase F

Les protéines sont extraites dans les mêmes conditions que décrites dans la section (IV. B. 2.). L'extrait protéique total est ensuite centrifugé pendant 30min à 15 000xg de façon à éliminer les débris cellulaires. Le surnageant obtenu est conservé puis dialysé pendant 24h contre de l'eau milliRo en utilisant des boudins de dialyse possédant un seuil de coupure de 6000-8000 Da (Fisher Chemical).

Après la dialyse, dix milligrammes de protéines sont ensuite lyophilisés dans un tube Falcon de 15mL puis repris dans 1mL de tampon Tris-HCl 0,1M pH 7,5 (LC-MS grade, Fisher Chemical) contenant 0,1% de SDS (p/v) (SDS micropellets Fisher Chemical) (p/v). Les tubes contenant les protéines sont chauffés à 100°C pendant 10min de façon à dénaturer les protéines. Les tubes sont ensuite refroidis à la température ambiante. Un mL de tampon Tris-HCl 0,1M pH 7,5 contenant 0,5% d'IGEPAL® (v/v) (MP bio, CA-630) est ajouté de façon à inhiber l'action du SDS puis les échantillons sont homogénéisés au vortex.

Quinze unités de PNGase F (Roche) sont ajoutées dans les tubes Falcon contenant les protéines dénaturées pour effectuer leur déglycosylation, puis laissés pendant 24 heures à 37°C sous une agitation de 100rpm.

2. Précipitation à l'éthanol des protéines déglycosylées

Au terme des 24h de déglycosylation, 8mL d'éthanol à 100% (v/v) (LC-MS grade, Fisher Chemical) sont ajoutés aux 2mL échantillon, soit l'équivalent de 4 volumes d'éthanol pour un volume d'échantillon. Le mélange est incubé pendant 24h à -20°C afin de permettre la précipitation des protéines et de récupérer dans la phase éthanolique les *N*-glycannes libres. Les tubes sont ensuite centrifugés à 10 000xg pendant 15min. Le surnageant contenant les *N*-glycannes libres est conservé puis séché sous une rampe à air.

3. Purification des oligosaccharides sur une colonne carbograph

Le protocole utilisé pour la purification sur colonne carbograph des *N*-glycannes après la digestion à la PNGase F est le même que celui détaillé dans la section (IV. A. 4.)

4. Marquage au 2-AB

Cette étape de dérivation des *N*-glycannes au 2-AminoBenzamide (2-AB) permet d'ajouter un groupement fluorophore, le 2-AB, sur les *N*-glycannes de façon à améliorer la sensibilité de l'analyse par spectrométrie de masse. Pour réaliser cette dérivation, dix milligrammes de 2-AB sont repris dans 200µL d'une solution DMSO anhydre/acide acétique (SIGMA) (70/30, v/v). Douze mg de cyanoborohydrure de sodium (NaBH₃CN) sont ensuite ajoutés et solubilisés dans ces 200µL. Dix µL du mélange obtenu sont ajoutés aux échantillons

préalablement lyophilisés puis incubés à 60°C pendant 2h. Au terme de cette incubation, les échantillons sont refroidis puis repris dans 500µL d'acétonitrile 96% (v/v) (LC-MS grade, Fisher Chemical).

5. Dessalage des glycanes marqués au 2-AB

Cette étape cruciale permet d'éliminer l'excédent de molécules de 2-AB qui n'ont pas été utilisées. Une colonne D1 (Ludger) de 2mL est au préalable conditionnée sous une pression de 5 mmHg par une enceinte manifold (Phenomenex SPE vacuum manifold) avec 2mL d'acétonitrile à 96% (v/v) (LC-MS grade, Fisher Chemical) et 1mL d'acétonitrile à 100% (v/v) (LC-MS grade, Fisher Chemical). Les échantillons, repris dans 500µL d'acétonitrile à 96% (v/v), sont ajoutés à la colonne et incubés pendant 15 minutes. Une étape de lavage est ensuite réalisée par l'ajout de 1mL d'acétonitrile à 100% (v/v), suivi de 5mL d'acétonitrile 96% (v/v). Les *N*-glycanes marqués au 2-AB, retenus dans la phase stationnaire de la colonne, sont élués dans un tube Eppendorf de 2mL par 4 ajouts successifs de 500µL d'acétonitrile à 20% (v/v). Cette étape d'éluion se fait sous la pression atmosphérique.

6. Perméthylation et purification sur colonne C18

Le protocole utilisé ici est le même que celui présenté dans les sections (IV. A. 5. et 6.) Après purification, les *N*-glycanes perméthylés sont ensuite analysés par spectrométrie de masse MALDI-TOF, ESI-MSⁿ et IMS-MS. Ces méthodes seront détaillées dans la section (V).

E. Extraction glycanique par la PNGase A

1. Dénaturation des protéines et digestion protéique par la trypsine et la chymotrypsine

Les protéines sont extraites dans les mêmes conditions que la section (IV. B. 2.). L'extrait protéique total est ensuite centrifugé pendant 30min à 15 000xg. Le surnageant obtenu est gardé puis dialysé pendant 24h contre de l'eau milliRo en utilisant des boudins de dialyse de seuil de coupure de 6000-8000Da (Fisher Chemical). Dix milligrammes de protéines sont ensuite lyophilisés dans un tube Falcon de 15mL. Le lyophilisat obtenu est repris dans

2mL de tampon de bicarbonate d'ammonium 100mM (LC-MS grade, Fisher Chemical) pH 8,3 puis 600µg de trypsine bovine (SIGMA) sont ajoutés. Les échantillons sont incubés pendant 24h à 37°C sous une agitation de 150rpm. La chymotrypsine (600µg) est ensuite ajoutée. Les tubes contenant les échantillons sont laissés pendant 4h à 37°C sous une agitation de 150rpm. La trypsine et la chymotrypsine sont ensuite inactivées en chauffant les tubes à 100°C pendant 10min. L'échantillon est ensuite lyophilisé.

2. Déglycosylation des glycoprotéines par la PNGase A

Les échantillons sont repris dans 500µL de tampon acétate de sodium 100mM pH 5,5 avant l'ajout d'1mU de PNGase A (Roche). Cette enzyme est capable de rompre la liaison *N*-osidique située entre le *N*-glycane et la protéine, même si le *N*-glycane possède un résidu $\alpha(1,3)$ -fucose lié sur le premier résidu GlcNAc du chitobiose. Le mélange est incubé pendant 36h à 37°C sous une agitation de 150rpm.

3. Purification des *N*-glycannes par les colonnes C18 et carbograph

La purification des *N*-glycannes précédemment libérés par la PNGase A est obtenue grâce à une colonne C18 (Hypersep C18, Thermo Fisher). Brièvement, la colonne est conditionnée dans 5mL d'éthanol (LC-MS grade, Fisher Chemical) puis 5mL d'eau (LC-MS grade, Fisher Chemical). Les échantillons sont passés entièrement dans la colonne puis 5mL d'eau sont ajoutés. Le tout est récupéré dans un tube Falcon puis lyophilisé. Les *N*-glycannes libres ainsi récupérés sont ensuite de nouveaux purifiés sur une colonne carbograph selon le protocole de la section (IV. A. 4.).

4. Perméthylation et purification des *N*-glycannes perméthylés par la colonne C18

Les *N*-glycannes sont ensuite perméthylés puis purifiés par une colonne C18 selon les protocoles détaillés dans la section (IV. A. 5. et 6.) Les *N*-glycannes perméthylés sont ensuite analysés par MALDI-TOF et LC-ESI-MS (voir leur protocole dans la section (V)).

F. Purification et analyse biochimique de la GnT I

1. Extraction microsomale de la GnT I

Les cellules de diatomées sont cultivées pendant 10 jours sous les conditions décrites en section (II. A. 2) en utilisant un milieu Conway contenant une source azotée de type NH_4Cl (**Tableau 3**). Au 6^{ème} jour, une première induction est réalisée. Les cellules sont alors récoltées par une centrifugation à 4 500xg pendant 10min à 19°C en utilisant une centrifugeuse Beckman® Allegra X15 R. Les cellules récoltées sont ensuite reprises dans un milieu de culture dans les conditions décrites en section (II. A. 2), contenant du NaNO_3 . Au 8^{ème} jour de culture, une seconde induction est réalisée avec 980µM d'une solution NaNO_3 ajoutée au milieu de culture. Au 10^{ème} jour, les cellules sont récoltées par une centrifugation à 4 500xg pendant 10min à 19°C en utilisant une centrifugeuse Beckman® Allegra X15 R. Le culot cellulaire est repris dans 1mL de tampon Tris-HCl 0,1M pH 7,2 contenant une tablette d'inhibiteurs de protéases SIGMAFAST™ Protease Inhibitor Cocktail Tablets, EDTA-Free produite par SIGMA. Le mélange est ensuite transféré dans un tube Lysing Matrix E produit par MP biomedicals™ puis les cellules sont broyées dans un broyeur à billes MP Biomedicals™ Instrument FastPrep-24™ 5G, 3 fois pendant 30s à 6,5m/s. Le broyat est ensuite centrifugé à 4 500xg pendant 10min. Le surnageant est gardé et le culot contenant les débris cellulaires est éliminé. Les étapes de broyage, de centrifugation sont répétées pendant 3 fois. Les surnageants sont regroupés ensemble puis centrifugés à 10 000rpm pendant 1h à 4°C dans une ultracentrifugeuse Beckman® Optima XPN-80. Cette première étape permet de sédimenter majoritairement la fraction plastidiale des cellules de la diatomée. Ensuite, le surnageant obtenu est récupéré puis de nouveau centrifugé à 100 000xg pendant 1h30 à 4°C. Le culot obtenu correspond à la fraction microsomale qui est reprise dans le tampon adapté à la suite de l'expérimentation.

2. Immunoprécipitation de la GnT I

L'immunoprécipitation de la GnT I a été réalisée à partir d'un extrait de microsomes solubilisés de *P. tricornutum* suivant le protocole (IV. A. 1.). Cette immunoprécipitation a été réalisée avec le kit Pierce® Crosslink Magnetic IP/Co-IP Kit, 40rxn. Brièvement, 5µg d'anticorps polyclonaux anti-V5 de lapin (V5 Epitope Tag Antibody, ThermoFisher®) ont été préalablement incubés avec 25µL de la suspension de billes magnétiques recouvertes de protéines A et de

protéines G d'*E. coli*. Les IgG adsorbées sur les billes ont été ensuite couplées de façon covalente sur les protéines A et G par un tampon de couplage fourni avec le kit. Puis, l'extrait de microsomes solubilisés, contenant la GnT I dotée de l'épitope V5, a été incubé pendant une nuit à 4°C sous une agitation rotative de 15rpm en présence des billes magnétiques porteuses des IgGs fixées dirigées contre l'épitope V5. Les complexes immuns formés pendant l'incubation et retenus par les billes, ont été attirés avec elles par le portoir magnétique (Magnetic Bands®, ThermoScientific). Ceci permet de séparer la protéine d'intérêt étiquetée en une étape de l'extrait liquide contenant les protéines microsomales contaminantes. Par la suite, la GnT I a été éluée des complexes immuns grâce à 40µL d'un tampon d'éluion à base de glycine pH 2 provenant du kit Pierce® Crosslink Magnetic IP/Co-IP Kit, 40rxn puis l'éluât a été neutralisé à pH 7 avec 4µL de tampon de neutralisation de pH 8,5 fourni également par le kit.

3. Vérification de l'immunoprécipitation par Western Blot

Dans le but de vérifier que l'immunoprécipitation de la GnT I a été couronnée de succès, un Western Blot a été réalisé en utilisant un anticorps primaire spécifiquement dirigé contre l'étiquette V5 présent à l'extrémité C-terminale de la GnT I suivant les conditions de la section (IV. B. 3./6-11).

4. Digestion trypsique et chymotrypsique de la GnT I

La protéine immunoprécipitée obtenue est déposée dans sa totalité sur un gel d'électrophorèse puis colorée suivant les conditions décrites dans la section IV. C. 2 et 3. La bande colorée obtenue correspondant à la GnT I migrant à environ 56kDa est découpée à l'aide d'un scalpel stérile puis découpée en morceaux d'environ 1mm³. Les bandes découpées sont ensuite incubées à 56°C pendant 1h dans une solution 100mM de bicarbonate d'ammonium. Les résidus de cystéine sont ensuite bloqués à l'obscurité pendant 20min à température ambiante dans une solution contenant 55mM d'iodoacetamide (SIGMA). Les protéines présentes dans le gel sont ensuite digérées par la trypsine à un ratio de 3% (Promega, Charbonnières-les-Bains, France) pendant une nuit à température ambiante. Après digestion toute une nuit, 1% d'acide formique (v/v) est ajouté puis le tout est agité au termomixer à 25°C pendant 20min. Le surnageant sont transférés dans de nouveaux tubes. Les peptides obtenus sont ensuite séchés au SpeedVac puis resuspendus dans 10 à 15µL

d'acétonitrile 3% (v/v) /acide formique (v/v) 0,1% (v/v) pour l'analyse par spectrométrie de masse (V. E. et F.)

V. Méthodes de spectrométrie de masse

A. Analyse par MALDI-TOF

1. Analyse du précurseur oligosaccharidique et du *N*-glycane Man₅GlcNAc₂

Les analyses par MALDI-TOF (Matrix Assisted Laser Desorption Ionisation – Time Of Flight) ont été effectuées avec un spectromètre de masse Autoflex III (Bruker Daltonics (Bremen, Germany)). L'appareil est équipé d'un laser grenat d'yttrium-aluminium dopé au néodyme (Nd : YAG laser, longueur d'onde 355-nm). L'acquisition des spectres a été effectuée en mode réflectron positif sur une gamme de masse variant de 500 m/z à 4 000 m/z avec une fréquence laser de 100 Hz, et une moyenne minimale de 3 000 spectres individuels. L'accélération de voltage de l'instrument a été fixée à 19 kV. La calibration du MALDI-TOF a été effectuée en utilisant les ions monoisotopiques monochargés de l'angiotensine II (5 pmol.μL⁻¹), la substance P (5 pmol.μL⁻¹), l'adencorticotropique hormone (ACTH) 18-39 (5 pmol.μL⁻¹) et l'insuline B (10 pmol μL⁻¹). L'acide 2,5- dihydroxybenzoïque (DHB) a été utilisé comme matrice. Un volume de 1 μL de chaque échantillon est mélangé dans 1μL de matrice DHB (10 mg.mL⁻¹ dans de l'ACN/eau (70/30: v/v) contenant 0,1% de TFA (v/v) puis, 0,7μL du mélange échantillon/matrice est déposé et séché sur la plaque MALDI (Bruker Daltonics, Bremen, Germany). Les logiciels FlexControl 3.3 et FlexAnalysis 3.3 ont été utilisés afin d'analyser les spectres obtenus. Les masses théoriques des *N*-glycannes ont été obtenues par le logiciel GlycoWorkBench 2.0.

2. Analyse des *N*-glycannes issus de la fraction protéique totale

Les conditions expérimentales sont identiques à celles présentées dans la section (V. A. 1.) exceptée la concentration de DHB qui est cette fois-ci de 20mg.mL⁻¹.

B. Analyse par ESI-MSⁿ

Les précurseurs oligosaccharidiques perméthylés ainsi que le Man₅GlcNAc₂ perméthylé ont été analysés par ESI-MSⁿ (Electrospray ionization mass spectrometry, n = 1 à 4) avec un spectromètre de masse Bruker HCT Ultra ETD II comportant un piège à ions de type quadripolaire (QIT). Les voltages du capillaire et du « end plate » ont été respectivement fixés à -3,5 kV et -3,0 kV en mode d'ionisation positive. Les voltages du « skimmer » et de la sortie du capillaire ont été fixés respectivement à 40 V et 200 V. Le gaz vecteur de nébulisation utilisé est du N₂. La pression en hélium dans le piège à ions était de $1,2 \text{ \AA} \sim 10^{-5} \text{ mbar}$. L'acquisition des données a été effectuée sur la gamme de masse m/z 100 à m/z 2 200, avec une vitesse de scan de m/z 26 000 par seconde. Le nombre d'ions entrant dans le piège a été automatiquement ajusté en contrôlant le temps d'accumulation en utilisant le mode de contrôle des charges ioniques (ICC) (target 100 000) avec un temps d'accumulation maximum de 50 ms. La valeur du LMCO (Low mass Cut-Off) a été de m/z 100 pour le Man₅GlcNAc₂ dérivé ou de m/z 140 pour le précurseur oligosaccharidique. Le nombre moyen de spectres a été de 6 et la moyenne mobile de 2. Les expériences d'ESI-MSⁿ ont été induites par des collisions CID utilisant l'hélium comme gaz de collision et conduites par dissociation. L'ion précurseur et les ions intermédiaires ont été isolés avec une fenêtre de m/z 1,5 en utilisant une fréquence de résonance avec une amplitude de 0,7 à 1,0 V_{p-p}. La solution contenant l'échantillon a été introduite dans la source avec un débit de $180 \mu\text{L}\cdot\text{h}^{-1}$ à l'aide d'une seringue (Cole-Palmer, Vernon Hills, IL, USA). Les logiciels Esquire control 6.2 et Data Analysis 4.0 (Bruker Daltonics, Bremen, Germany) ont été utilisés pour l'analyse des spectres de masse obtenus. Les masses théoriques des ions parents et fragments correspondant aux N-glycannes ont été calculés en utilisant le logiciel GlycoWorkBench 2.0.

C. Analyse par ESI-IMS

Les expériences d'ESI-IMS (Electrospray Ionization - Ion Mobility Spectrometry) ont été réalisées avec le spectromètre de masse Waters SYNAPT G2 hybrid quadrupole/HDMS équipé d'une source ionisation ESI LockSprayTM et des logiciels massLynx 4.1 et DriftScope 2.2 (Waters, Manchester, UK). Le SYNAPT HDMS a été calibré en utilisant des groupes d'ions

généralisés par le formate de sodium ($2\text{mg}\cdot\text{mL}^{-1}$) et opéré en mode sensible (résolution de 20 000 FWHM). Les paramètres ESI en mode positif sont les suivants : voltage capillaire de 3.1kV ; voltage du cône de 90V ; voltage du cône d'extraction de 5V ; la température de la source de 110°C ; la température de désolvatation de 250°C ; le flux du gaz de désolvatation (N_2) à $700\text{L}\cdot\text{h}^{-1}$ et le flux de gaz du cône (N_2) à $20\text{L}\cdot\text{h}^{-1}$. Les données ont été acquises en utilisant une gamme de masse allant de m/z 50 à 2000 avec 1 seconde de durée de scan et 0,02 seconde de délai d'interscan. Les solutions d'échantillons ont été infusées dans la source à un débit de $300\mu\text{L}\cdot\text{h}^{-1}$ à l'aide d'une seringue (Cole-Palmer, Vernon Hills, Illinois, USA). Les conditions IMS ont été optimisées de la manière suivante : le débit du gaz (N_2) à $90\text{mL}\cdot\text{min}^{-1}$; la pression IMS à 2.95mbar ; l'amplitude à 40V et la vitesse de l'onde à $700\text{m}\cdot\text{s}^{-1}$. Le spectre de mobilité ionique a été affiné en utilisant le logiciel Origin 9.0 (OriginLab). Les masses théoriques des ions parents et fragments correspondant aux *N*-glycannes ont été calculés en utilisant le logiciel GlycoWorkBench 2.0.

D. Analyse des glycannes perméthylés par nanoLC-ESI-MS

Les glycannes perméthylés sont resuspendus dans une solution contenant 30% d'acétonitrile (v/v) /0,1% acide formique (v/v) avant d'être analysés avec un système nano-LC1200 couplé à un spectromètre de masse Q-TOF 6545 XT AdvanceBio et équipé d'une source d'ionisation Nanospray et d'une interface HPLC-chip cube (Agilent Technologies). Un gradient linéaire de 19min (30 à 90% d'acétonitrile (v/v) dans 0,1% d'acide formique (v/v)), à un flux de $350\text{nL}\cdot\text{min}^{-1}$. Brièvement, les glycannes perméthylés ont été enrichis et dessalés sur une colonne 40nL RP-C18 et séparés sur une colonne C18 ou PGC (150mm de long x $75\mu\text{m}$ de diamètre ; Agilent Technologies) de type Zorbax (taille des pores de 30nm, taille des particules de $5\mu\text{m}$). Un scan complet autoMS1 de m/z 290 à m/z 3200 et autoMS2 de m/z 59 à m/z 3200 a été enregistré. Dans chaque cycle, un maximum de 5 précurseurs classés selon leur état de charge (les ions monochargés étant préférentiels et le reste exclu) a été isolé et fragmenté dans la chambre de collision avec une énergie de collision (CE) fixée (CE 25, 30, 35 et 45 eV). L'exclusion de ces précurseurs a été activée dans les 0,2min après l'obtention d'un spectre, et le seuil de sélection des précurseurs a été fixé à 1000 coups (seuil relatif de 0.001%).

La vitesse de balayage variait en fonction de l'abondance des précurseurs (15 000

comptages/spectre) et avec une rigueur de pureté de 100% et un seuil de pureté de 30%. Les spectres ont été analysés en utilisant le logiciel Qualitative Analysis B.06.00 (Agilent Technologies). Les masses théoriques des ions parents et fragments correspondant aux *N*-glycannes ont été calculés en utilisant le logiciel GlycoWorkBench 2.0.

E. Analyse de la GnT I par QTOF XT

Les échantillons repris dans 3% d'acétonitrile (v/v) et 0,1% d'acide formique (v/v) obtenus dans la section (IV. F. 4.) ont été analysés avec un nano-LC1200 couplé à un spectromètre de masse QTOF 6545 XT AdvanceBio équipé d'une source Nanospray et d'une interface HPLC-chip cube (Agilent Technologies, Les Ulis, France). Brièvement, les peptides ont été enrichis et dessalés sur une colonne 40nL RP-C18 et séparés sur une colonne C18 (150mm de long x 75µm de diamètre) de type Zorbax (taille des pores de 30nm, taille des particules de 5µm). Un gradient linéaire de 26min (3%-80% acétonitrile (v/v) dans 0,1% d'acide formique (v/v)) avec un flux de 350nL/min a été utilisé, et l'éluât a été analysé avec un spectromètre de masse de type QTOF. Un scan complet autoMS1 de m/z 290 à m/z 2000 et autoMS2 de m/z 59 à m/z 3200 a été enregistré. Dans chaque cycle, un maximum de 5 précurseurs classés selon leur état de charge (les ions dichargés préférés et les ions monochargés exclus) a été isolé et fragmenté dans la chambre de collision. L'énergie de collision a été automatiquement ajustée selon le ratio m/z . L'exclusion de ces précurseurs a été activée dans les 0,2 min après l'obtention d'un spectre, et le seuil de sélection des précurseurs a été fixé à 1000 coups. Les spectres ont été analysés en utilisant le logiciel Qualitative Analysis B.06.00 (Agilent Technologies, USA).

F. Analyse de la GnT I par Qexactive Plus

Les échantillons repris dans 3% d'acétonitrile (v/v) et 0,1% d'acide formique (v/v) obtenus dans la section (IV. F. 4.) ont été analysés sur le Q Exactive Plus (Thermo Scientific, France) équipé d'une source nanoESI. Les peptides ont été chargés sur une colonne d'enrichissement (C18 Pepmap100 precolumn (300µm i.d. x5mm, 5µm, 100 Å, Thermo Scientific) et séparés sur une colonne EASY-spray (15cm, 3m, 0.075mm, 100 Å, Thermo Scientific), avec un flux de 300µL.min⁻¹. La phase mobile est composée d'une solution

contenant 0,1% d'acide formique (v/v) dans de l'eau (Tampon A) et une solution composée d'ACN/H₂O (80/20 (v/v)) et 0,1% d'acide formique (v/v) (Tampon B). La durée du gradient d'élution était de 30min : 0-16min, 2-50% avec tampon B ; 16-17min, 50-90% avec tampon B ; 17-20min, 90% avec tampon B ; 20,5-27min, 2% avec tampon B. La température de la colonne a été fixée à 40°C. Les paramètres d'acquisition du spectre de masse sont les suivants : le temps d'injection maximum de 100ms, le voltage capillaire de 1.6KV, la température capillaire de 275°C, le scan complet MS de m/z @400-1800 avec une résolution de 70 000 en MS et 17 500 en MS/MS. Les 10 ions les plus intenses ont été sélectionnés puis fragmentés avec de l'azote comme gaz de collision. L'énergie de collision normalisée a été fixée à 27 eV. Les spectres de masse ont été analysés avec le logiciel MZmine 2.53.

Chapitre 3

Résultats

A. Analyse des précurseurs oligosaccharidiques liés au dolichol par spectrométrie de masse chez la diatomée *P. tricornutum*

Ce premier chapitre de résultats présenté sous la forme d'un article scientifique relate à la fois l'optimisation d'un protocole d'extraction des précurseurs oligosaccharidiques à partir de cellules de microalgues et l'analyse de ces précurseurs par des approches multiples de spectrométrie de masse. La méthodologie développée a été tout d'abord validée sur la microalgue verte, *C. reinhardtii* puis utilisée pour étudier la structure détaillée du précurseur oligosaccharidique de la diatomée, *P. tricornutum*. La méthodologie optimisée peut maintenant s'appliquer à n'importe quelle autre microalgue. Ces travaux ont fait l'objet d'une publication parue en Décembre 2018 dans le journal « Plant Methods ».

Chez l'Homme, la voie de la *N*-glycosylation des protéines débute dans le RE par la synthèse d'un précurseur oligosaccharidique de type $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$, précurseur universel à tous les eucaryotes (**Aebi, 2013**). Des études bio-informatiques effectuées chez les microalgues *C. reinhardtii* et *P. tricornutum* ont montré l'absence des gènes codant certaines enzymes ALG participant à la synthèse de ce précurseur oligosaccharidique (**Baïet et al., 2011 ; Mathieu-Rivet et al., 2013 ; Levy-Ontman et al., 2014 ; Mathieu-Rivet et al., 2014**). L'absence de ces ALG (3 chez *C. reinhardtii* et 1 chez *P. tricornutum*) aurait pour conséquence la synthèse d'un précurseur oligosaccharidique différent de celui couramment identifié chez les eucaryotes. Dans ce contexte, une analyse structurale des précurseurs localisés au niveau du RE s'avère donc nécessaire.

Dans ce chapitre de résultats, nous avons développé à la fois un protocole original de purification des précurseurs et une analyse structurale par spectrométrie de masse multidimensionnelle. Les masses moléculaires de ces précurseurs ont ensuite été déterminées par MALDI-TOF et leurs structures à partir des spectres de fragmentation obtenus en ESI-MSⁿ (n=4). Le protocole a été validé dans un premier temps sur la microalgue verte *C. reinhardtii*. Les oligosaccharides précurseurs extraits d'une souche sauvage et de mutants déficients en activités xylosyltransférases golgiennes ont été analysés. Les résultats obtenus ont permis de mettre en évidence une structure de type $\text{Glc}_3\text{Man}_5\text{GlcNAc}_2$ chez *C. reinhardtii* en accord avec les résultats obtenus précédemment (**Vanier et al., 2017**). Comme attendu, aucune différence

structurale n'a été observée dans les précurseurs oligosaccharidiques issus du sauvage et des mutants déficients en activités xylosyltransférases. Le précurseur oligosaccharidique de *P. tricornutum* a été extrait et étudié pour la première fois dans le cadre de ce travail. Les analyses par spectrométrie de masse ont permis de mettre en évidence la synthèse d'un précurseur oligosaccharidique tronqué d'un hexose de type $\text{Glc}_2\text{Man}_9\text{GlcNAc}_2$ chez *P. tricornutum*. Ces résultats corroborent les prédictions bio-informatiques publiées précédemment (**Baïet et al., 2011 ; Mathieu-Rivet et al., 2014 ; Levy-Ontman et al., 2014**) et nous permettent de proposer une voie de la synthèse du précurseur oligosaccharidique chez *P. tricornutum* (**Figure 27**).

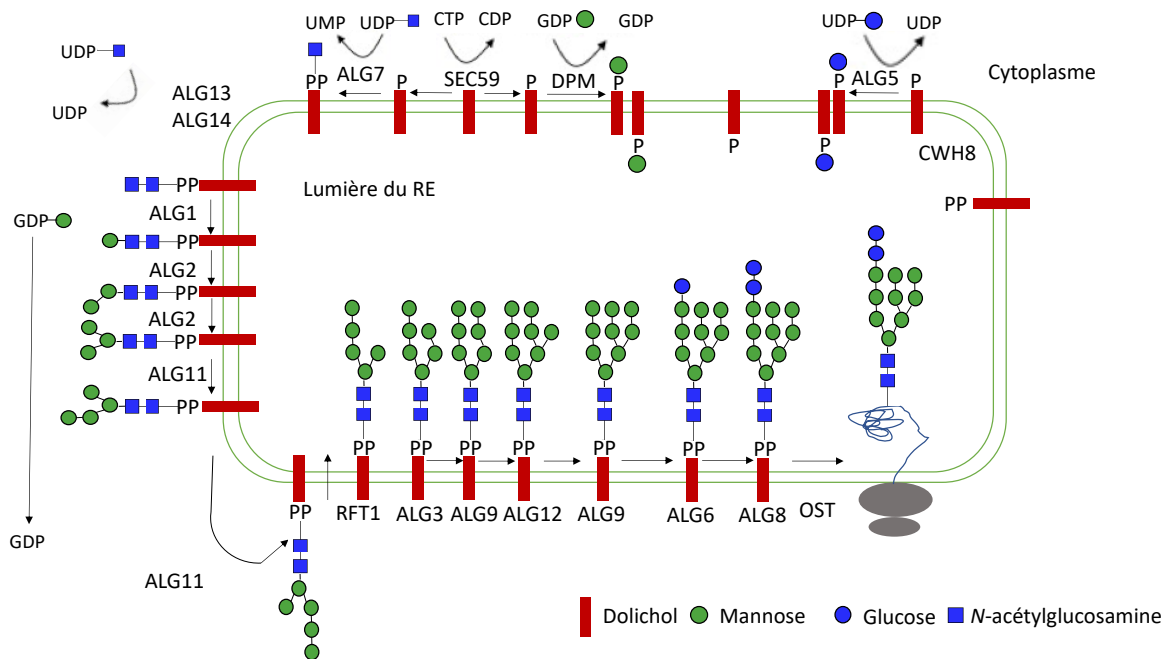


Figure 27 : Voie de la biosynthèse du précurseur glycanique des *N*-glycannes des protéines chez *P. tricornutum*

Ce schéma résume les étapes de synthèse du précurseur se déroulant dans le réticulum endoplasmique de la diatomée *P. tricornutum*. Le changement d'orientation entre les phase cytosolique et luminale serait opéré par une enzyme flippase RFT1 putative. Les étapes qui consistent en l'ajout de différents monosaccharides sont catalysées par une succession d'enzymes appelées ALG. Contrairement à la majorité des eucaryotes, ces étapes conduiraient à la formation d'un précurseur oligosaccharidique lié au dolichol tronqué d'un hexose composé de 3 résidus *N*-acétylglucosamine, 9 résidus mannose et 2 résidus glucose. Ce précurseur oligosaccharidique est ensuite transféré sur une protéine en cours de synthèse par l'OST. Les structures *N*-glycanniques sont dessinés en accord avec la nomenclature internationale (Varki *et al.*, 2015).

UDP : uridine diphosphate ; UMP uridine monophosphate ; GDP : guanosine diphosphate ; ALG : asparagine linked glycosylation ; DPM : dolichyl-phosphate mannosyltransférase I ; OST : oligosaccharyltransférase ; CTP : cytosine triphosphate ; CDP : cytosine diphosphate ; PP : pyrophosphate ; RE : réticulum endoplasmique

METHODOLOGY

Open Access



User-friendly extraction and multistage tandem mass spectrometry based analysis of lipid-linked oligosaccharides in microalgae

Pierre-Louis Lucas^{1†}, Rodolphe Dumontier^{1†}, Corinne Loutelier-Bourhis², Alain Mareck¹, Carlos Afonso², Patrice Lerouge¹, Narimane Mati-Baouche¹ and Muriel Bardor^{1,3*}

Abstract

Background: Protein *N*-glycosylation is initiated within the endoplasmic reticulum through the synthesis of a lipid-linked oligosaccharides (LLO) precursor. This precursor is then transferred *en bloc* on neo-synthesized proteins through the action of the oligosaccharyltransferase giving birth to glycoproteins. The *N*-linked glycans bore by the glycoproteins are then processed into oligomannosides prior to the exit of the glycoproteins from the endoplasmic reticulum and its entrance into the Golgi apparatus. In this compartment, the *N*-linked glycans are further matured in complex type *N*-glycans. This process has been well studied in a lot of eukaryotes including higher plants. In contrast, little information regarding the LLO precursor and synthesis of *N*-linked glycans is available in microalgae.

Methods: In this report, a user-friendly extraction method combining microsomal enrichment and solvent extractions followed by purification steps is described. This strategy is aiming to extract LLO precursor from microalgae. Then, the oligosaccharide moiety released from the extracted LLO were analyzed by multistage tandem mass spectrometry in two models of microalgae namely the green microalgae, *Chlamydomonas reinhardtii* and the diatom, *Phaeodactylum tricornutum*.

Results: The validity of the developed method was confirmed by the analysis of the oligosaccharide structures released from the LLO of two xylosyltransferase mutants of *C. reinhardtii* confirming that this green microalga synthesizes a linear Glc₃Man₂GlcNAc₂ identical to the one of the wild-type cells. In contrast, the analysis of the oligosaccharide released from the LLO of the diatom *P. tricornutum* demonstrated for the first time a Glc₂Man₉GlcNAc₂ structure.

Conclusion: The method described in this article allows the fast, non-radioactive and reliable multistage tandem mass spectrometry characterization of oligosaccharides released from LLO of microalgae including the ones belonging to the Phaeodactylaceae and Chlorophyceae classes, respectively. The method is fully adaptable for extracting and characterizing the LLO oligosaccharide moiety from microalgae belonging to other phyla.

Keywords: Diatom, *Chlamydomonas reinhardtii*, *Phaeodactylum tricornutum*, Lipid-linked oligosaccharides, Microalgae, Multistage tandem mass spectrometry

*Correspondence: muriel.bardor@univ-rouen.fr

[†]Pierre-Louis Lucas and Rodolphe Dumontier contributed equally to this work

¹ UNIROUEN, Laboratoire Glyco-MEV EA4358, Normandie Univ, 76000 Rouen, France

Full list of author information is available at the end of the article



© The Author(s) 2018. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Background

N-glycosylation is a protein post-translational modification playing many roles in biological process [1]. For example, interactions between host and pathogens or between cells involve specific recognition between oligosaccharides and lectins [1]. Moreover, *N*-glycosylation ensures the conformation, the stability and the biological function of glycoproteins [2]. In plants, *N*-glycosylation has been widely studied, mainly in *Arabidopsis thaliana* and *Nicotiana benthamiana* [3–7]. The *N*-glycan synthesis starts in the endoplasmic reticulum (ER) with a two steps processing involving the action of a set of specific enzymes named Asparagine-Linked Glycosylation (ALG) [8]. The first step, beginning in the cytoplasmic side of the ER, results in the assembly of an oligosaccharide on a lipid carrier called dolichol pyrophosphate (PP-Dol) by successive additions of two *N*-acetylglucosamine and five mannose residues respectively through the action of ALG7, 13, 14, 1, 2 and 11. The lipid-linked oligosaccharide precursor (LLO) is then flipped towards the luminal side of the ER where the ALG3, 9 and 12 add four additional mannose residues, then respective actions of the ALG6, 8 and 10 elongate the LLO structure with three terminal glucose residues [9]. Such assembly of the LLO leads to the Glc₃Man₅GlcNAc₂-PP-Dol precursor in plants which is identical to the structures identified in other eukaryotes [8, 9]. The second step involves the transfer of the oligosaccharide part of the LLO from the PP-Dol to an asparagine residue of a neosynthesized protein by the oligosaccharyltransferase complex (OST) [10, 11]. This is directly followed by the quality control of the protein folding involving, the removal of the three terminal glucose residues as a final step, leading to the formation of oligomannosidic *N*-linked glycans [12, 13]. This nascent folded glycoprotein is then transferred to the Golgi apparatus where *N*-linked glycans will undergo specific maturation by glycosyltransferases and glycosidases which are different among species and lead to different complex-type *N*-glycans. In plants, the *N*-glycosylation maturation process generates complex-type *N*-glycans GlcNAc₂Man₃GlcNAc₂ bearing alpha(1,3)-core fucose and beta(1,2)-core xylose and eventually Lewis a epitopes [7, 14, 15].

So far, little attention has been paid to protein *N*-glycosylation in microalgae. Indeed, all the published work focused on the analysis of the *N*-glycan structures of four main microalgae: *Phaeodactylum tricornutum*, *Botryococcus braunii*, *Chlamydomonas reinhardtii* and *Porphyridium sp.* [16–20]. Major differences in complex *N*-glycan structures of these microalgae were recently reviewed [21]. For example, complex *N*-glycans containing a xylose moiety have been identified in *C. reinhardtii*, *Porphyridium sp.* and *Botryococcus braunii* but not in

P. tricornutum. Currently, no functional study characterizing the enzymes involved in the LLO synthesis in microalgae is available in the literature. Preliminary bioinformatic genome analyses indicated that some ALG could be missing. Indeed, neither ALG3, ALG9 and ALG12 in *C. reinhardtii* nor ALG10 in both *P. tricornutum* and *C. reinhardtii* have been identified so far [16, 18, 22]. Because those genes are crucial for the complete synthesis of the Glc₃Man₅GlcNAc₂ oligosaccharide, the lack of these steps could severely impact the synthesis of the final LLO precursor [9]. To the best of our knowledge, only one publication reported so far the structure of the LLO precursor from *C. reinhardtii* wild-type [23], demonstrating that *C. reinhardtii* accumulates a predominant linear truncated Glc₃Man₅GlcNAc₂ LLO in the ER. Such a truncated structure probably results from the absence of the ALG genes as already described in protozoan parasites. Indeed, parasite LLO extracted from *Cryptosporidium parvum* and *Toxoplasma gondii* are characterized by the synthesis of a truncated single arm precursor with five mannose residues and two or three glucose residues [24, 25]. These studies tend to demonstrate that the LLO synthesis is not so well preserved in eukaryotes and deserves to be studied in other species like microalgae. Moreover, information regarding the functional property of LLO in microalgae is currently missing.

In this article, an user-friendly extraction method and multistage tandem mass spectrometry were combined to characterize oligosaccharides released from the LLO of microalgae. This method is elaborated from previous reported works in plants and parasites and has been validated on two xylosyltransferase mutants of the microalga *C. reinhardtii* which were compared to the wild-type cells. The oligosaccharide moiety released from the *C. reinhardtii* mutant LLO was demonstrated to correspond to a linear Glc₃Man₅GlcNAc₂ structure confirming previous results [23]. Then, the protocol has been applied to the Pt1.8.6 strain of the diatom *P. tricornutum* in order to determine the structure of its LLO-released oligosaccharide. Such results demonstrate the adaptability of the method described herein for further studies of LLO extracted from other microalgae.

Results and discussion

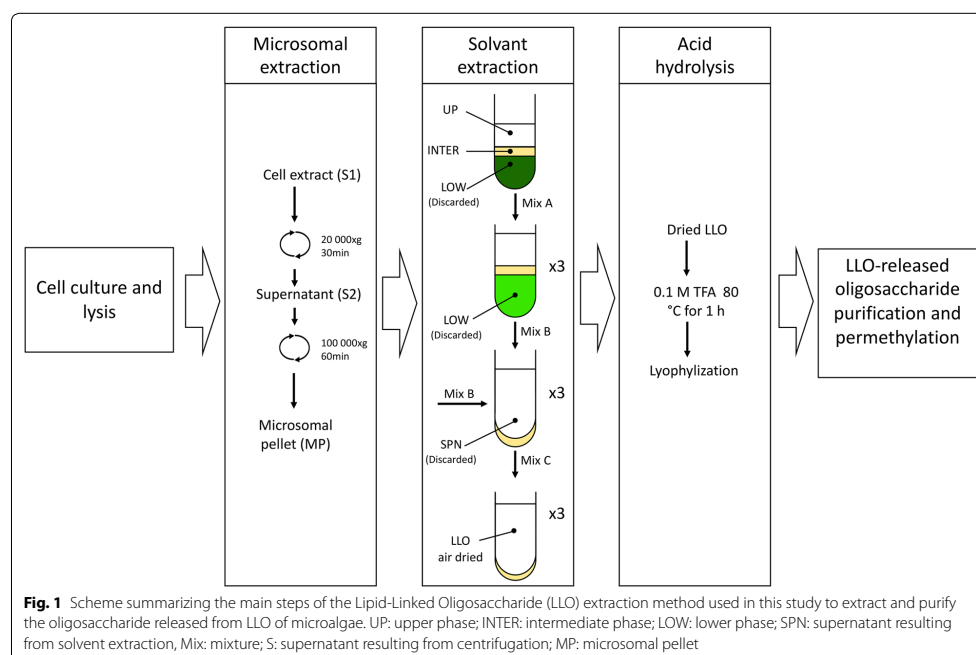
Optimization of a protocol for the LLO precursor extraction and characterization in microalgae

Nowadays, most of the protocols reported in the literature for LLO extractions and analyses are based on radioactivity labelling. Cells are incubated with radioactive ³H- or ¹⁴C-labelled metabolic precursors which are incorporated into LLO through the cell metabolism [26–28]. This strategy allows detection and analysis of low quantity of LLO but requires specific health and

safety practices and cannot be performed in all laboratories due to the hazard management. Moreover, as incubation times with radioactive precursors are often very short, the incorporated radioactivity may not reflect the complete set of oligosaccharides produced by the organisms and the labelled material represents minute amount which is not compatible with a detailed structural analysis of the radioactive compounds.

Here, we developed a simple protocol for LLO extraction without any radioactive labelling. In this purpose, the LLO were optimally extracted by the combination of two methods previously reported for *T. gondii* and *A. thaliana*, respectively [24, 29]. The first step of the protocol consisted in cell homogenization and a centrifugation step at 10 000xg. The cell homogenization was optimized for *C. reinhardtii* and *P. tricornutum* respectively as their cell wall composition is completely different. The diatom cells containing a more rigid cell wall require more grinding cycles for efficient cell lysis. Therefore, in future application, we would recommend for other microalgae species to apply either the specie-specific cell lysis protocol if available or the one used herein for *P. tricornutum*. It is to note that our extraction procedure does not include lipase inhibitors. However, the PMSF used during

the cell lysis step might help decreasing lipase activity as previously described [30, 31]. Therefore in future application, if low LLO-released oligosaccharides are recovered using the protocol described here, it may be worth adding lipase inhibitors before cell breakage to make sure that lipase activities are not degrading the linkage between the dolichol phosphate and the oligosaccharide. Then, the microsomal fraction containing the LLO anchored in the ER were recovered from the other cell components by successive centrifugation steps at 20,000xg followed by an ultracentrifugation step at 100,000xg. The resulting pellet corresponding to the enriched microsomal fraction was then collected (MP in Fig. 1). The second step corresponds to the LLO purification using successive solvent extraction procedures. Finally, two different methods have been tested in order to hydrolyze the pyrophosphate bond between the lipid and the oligosaccharide of the LLO. The first one consisted in resuspending the LLO in 0.01 M HCl in tetrahydrofuran and incubated the sample at 50 °C for 2 h [24]. The second method involved the hydrolysis of the LLO using 0.1 M of trifluoroacetic acid (TFA) at 80 °C for 1 h [29]. In our hands, only the mild acidic hydrolysis using 0.1 M of TFA gave sufficient amount of LLO-released oligosaccharide for further

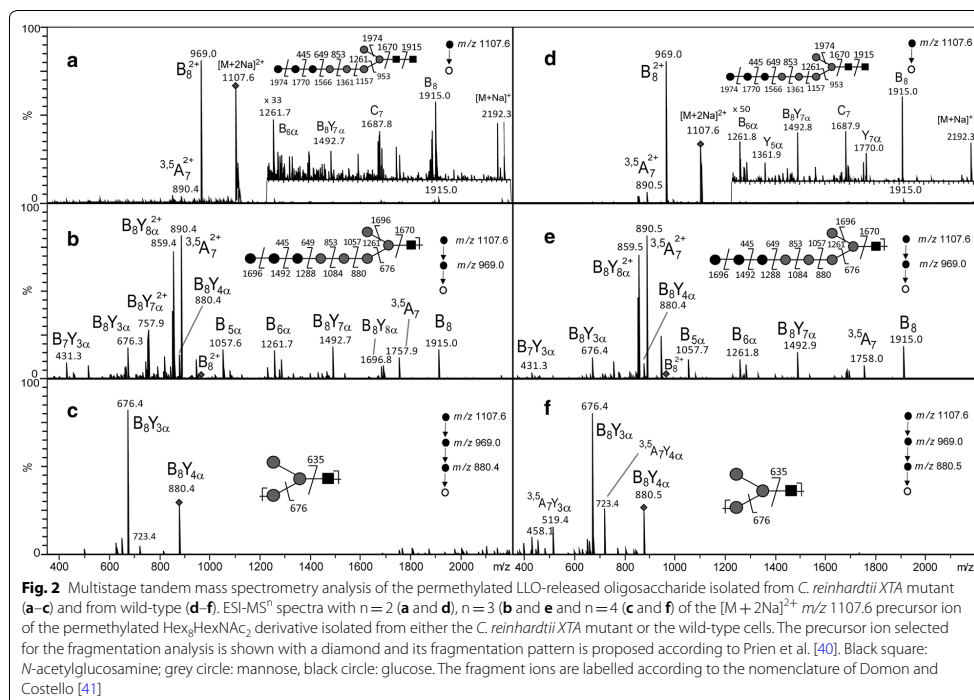


multistage tandem mass spectrometry analysis. Figure 1 summarizes the overall user-friendly protocol of LLO extraction in microalgae. After purification of the LLO-released oligosaccharide using a Carboglyph column, the oligosaccharide was permethylated as previously described [32]. Permethylated samples presents two main analytical advantages. Firstly, permethylated oligosaccharides are less polar than their native forms and so are more efficiently ionized in MALDI and ESI [33, 34]. Secondly, this greatly facilitates the discrimination between glycan isomers when permethylated oligosaccharides are fragmented by collision-induced dissociation multistage tandem mass spectrometry [35–37].

Identification of oligosaccharide isolated from the LLO of the green microalgae *Chlamydomonas reinhardtii*

The optimised protocol was first applied to the characterization of the LLO-released oligosaccharide isolated from the LLO of two *C. reinhardtii* mutants which are impaired in two different genes encoding putative xylosyltransferases (*XTA* and *XTB*) involved in the xylosylation of complex-type *N*-glycans in *C. reinhardtii* [18, 38, 39]. In

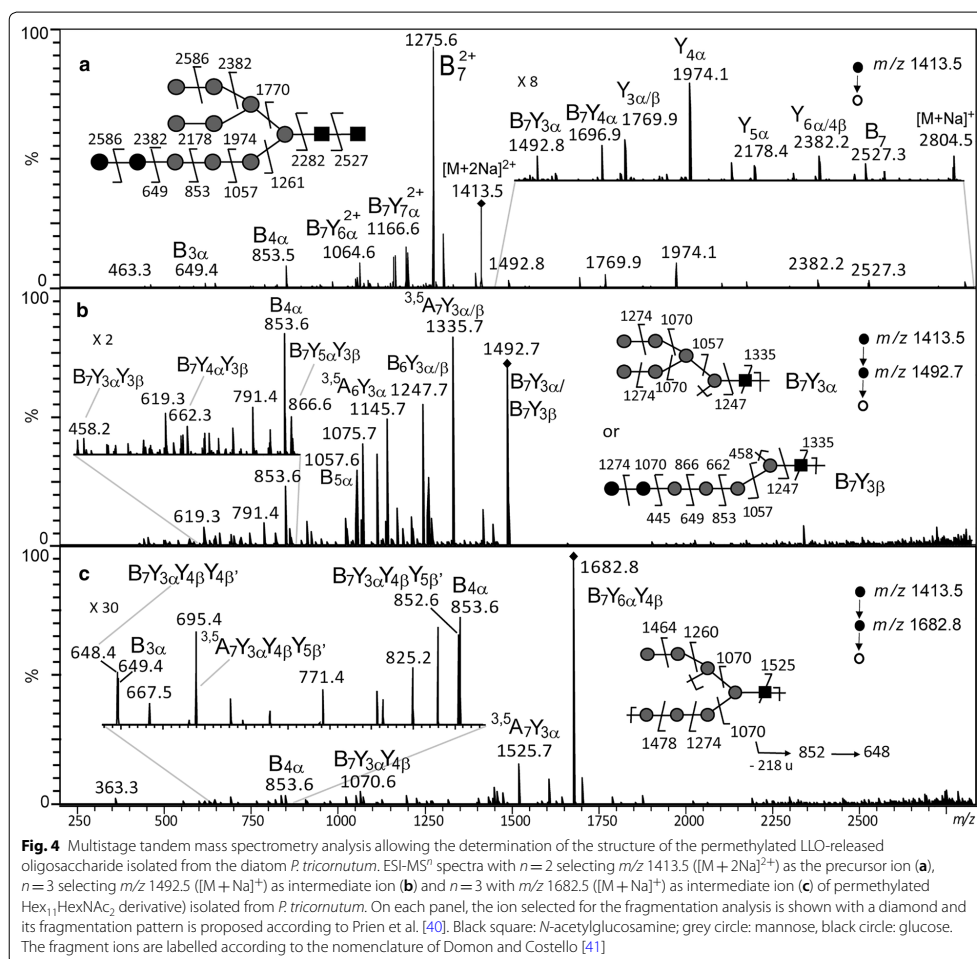
parallel, wild-type *C. reinhardtii* cells were submitted to the same LLO extraction protocol. After extraction, mild hydrolysis and permethylation, LLO-released oligosaccharides of *XTA*, *XTB* mutants and wild-type have been analyzed using MALDI-TOF-MS. The resulting mass spectra for all three samples show a predominant ion at m/z 2192 which was assigned to be the sodium adduct of a Hex₈HexNAc₂ (data not shown). To gain insight into the structural information about this LLO-released oligosaccharide, multistage tandem mass spectrometry (ESI-MS^{*n*}) experiments have been performed, as this method has been proven to be powerful to determine the chemical structures of polysaccharide fragments [33, 34, 37] and glycans [23, 40]. Here, the [M+2Na]²⁺ ion (m/z 1107.6) of Hex₈HexNAc₂ has been selected for fragmentation (Fig. 2). The ESI-MS^{*n*} ($n=2, 3$ and 4) spectra show mainly B- and Y-ion series which are fragment ions usually observed in MS/MS analyses of oligosaccharides [23, 37, 41]. They arise from cleavages of the glycosidic bond with loss of either the reducing terminal portion for B-ions or the non-reducing terminal portion for Y-ions. Few BY- ions resulting from two glycosidic cleavages



as well as some A-ions arising from cross ring cleavages were also observed. The fragment ions are labelled according to the nomenclature introduced by Domon and Costello in 1988 [41]. For clarity in the figures, we did not detail all the fragment ions observed on the MSⁿ spectra, but we choose to focus on those corresponding to diagnostic transitions which allow determining the oligosaccharide structure and even discriminating potential isomers. Thus, the specific fragmentation pathway m/z 1107.6 ($[M+2Na]^{2+}$) \rightarrow m/z 969.0 (B_8^{2+}) \rightarrow m/z 880.4 (B_8Y_{4a}) \rightarrow m/z 676.4 (B_8Y_{3a}) which was observed in both mutants (Fig. 2a–c for *C. reinhardtii* XTA mutant and Additional file 1 for *C. reinhardtii* XTB mutant) as well as in wild-type strains (Fig. 2d–f) is consistent with the oligosaccharide structure previously described for the LLO of the *C. reinhardtii* wild-type cells [23]. Figure 3 shows the fragmentation pattern of the $[M+2Na]^{2+}$ ion (m/z 1107.6) with assignments and detailed structures of the diagnostic fragment ions. These results confirmed the presence of a linear LLO structure in *C. reinhardtii* XTA and XTB mutants as compared to the wild-type cells. Indeed, as in wild-type, *C. reinhardtii* XTA and XTB mutants accumulate a linear LLO-released oligosaccharide containing 8 hexose and 2 *N*-acetylglucosamine residues which largely differs from the $Glc_3Man_9GlcNAc_2$ synthesized in most of the eukaryotes [9, 11]. The bioinformatic analysis of *C. reinhardtii* genome failed in predicting orthologous genes for ALG3, 9 and 10 responsible for the completion of the LLO synthesis into $Glc_3Man_9GlcNAc_2$ -PP-Dolichol [18, 22, 38]. Based on these predictions, *C. reinhardtii* LLO-released oligosaccharide structure should be a linear $Glc_2Man_5GlcNAc_2$. However, the MS analysis suggested the accumulation in *C. reinhardtii* of a $Hex_8HexNAc_2$ LLO-released oligosaccharide, thus exhibiting one additional hexose residue than the LLO structure predicted from genomic data. Therefore, a non-yet identified glycosyltransferase should act on the LLO precursor to transfer this additional hexose. As the specific ion transitions m/z 1107.6 \rightarrow m/z 969.0 \rightarrow m/z 880.4 \rightarrow m/z 676.4 evidenced by ESI-MSⁿ analyses is consistent with the linear glycan structure depicted in Figs. 2 and 3, we proposed that LLO from *C. reinhardtii* mutants exhibit a trihexyl extension which could probably be a triglucosyl one. $Glc_3Man_5GlcNAc_2$ LLO-released oligosaccharide may result from the action of an additional glycosyltransferase adding a third glucose to the LLO precursor. This hypothesis is further supported by the prediction in the *C. reinhardtii* genome of a putative alpha (1, 2) glucosidase I (GSI) [18, 22, 38]. This enzyme would be responsible for the cleavage of the terminal glucose of the triglucosyl extension on the glycan precursor after its transfer onto the nascent glycoprotein [42, 43].

Identification of the LLO-released oligosaccharide from *Phaeodactylum tricornutum*

To investigate whether the developed method would be suitable for other microalgae species, we investigated the oligosaccharide structure released from the LLO of the diatom *P. tricornutum* which has never been reported so far in the literature. LLO was isolated using the same protocol and its oligosaccharide moiety was cleaved from the dolichol lipid anchor and permethylated prior to MS analysis. MALDI-TOF mass spectrum exhibited two ions at m/z 2804 and m/z 2820 (data not shown). They respectively correspond to sodium and potassium adducts of a $Hex_{11}HexNAc_2$ structure, differing by one hexose residue from the $Glc_3Man_9GlcNAc_2$ LLO precursor identified in mammals and land plants [9, 11, 44]. In addition, ESI-MSⁿ ($n=1$ to 4) analyses have been performed by selecting the $[M+2Na]^{2+}$ of $Hex_{11}HexNAc_2$ structure as a precursor ion (m/z 1413.5). ESI-MSⁿ ($n=3$ and 4) experiments (Fig. 4 and Additional files 2, 3 and 4) were performed and were consistent with the tri-antennae structure proposed in Figs. 4a, 5, 6 and Additional file 2. For example, the ESI-MS³ spectrum selecting m/z 1492.7 as the intermediate fragment ion (Figs. 4b, 5) showed the presence of the discriminant fragment ions m/z 458.2 ($B_7Y_{3a}Y_{3\beta}$) m/z 662.3 ($B_7Y_{4a}Y_{3\beta}$) and m/z 866.6 ($B_7Y_{5a}Y_{3\beta}$) of the antennae a in addition to m/z 1070.6 ($B_7Y_{6a}Y_{3\beta}$) and m/z 1274.7 ($B_7Y_{7a}Y_{3\beta}$) which are not discriminant, because the two latter can arise from the fragmentation of either the antenna a or the di-antenna branch bc. It is worth noting that m/z 1492.7 first fragments into m/z 1274.7 by elimination of 218 u (terminal hexose residue). Then, successive losses of 204 u (hexose residue that was linked with two other residues) explain the ion series m/z 1070.6, m/z 866.6, m/z 662.3 and m/z 458.2. This fragmentation pattern is characteristic of the antennae a (Fig. 5). The ESI-MS³ spectrum selecting m/z 1682.5 as the intermediate fragment ion (Figs. 4c, 6) revealed the presence of discriminant fragment ions at m/z 852.6 and m/z 648.4 corresponding respectively to the $Hex_3HexNAc$ and $Hex_2HexNAc$ structures arising from the fragmentation of a branched LLO. It is worth noting that m/z 1682.5 can also be assigned to the structure depicted in the Additional file 2 that gives rise to the fragment ions m/z 649.3 (B_{3a}) and m/z 853.6 (B_{4a}) in Fig. 4c. The presence of a tri-antenna LLO-released oligosaccharide structure was also evidenced by the ESI-MS⁴ spectrum showing the fragmentation pattern " m/z 1413.5 \rightarrow m/z 1057.6 \rightarrow m/z 839.4 \rightarrow product ions (Additional file 3). The fragment ions m/z 417.2 and m/z 621.3 can only result from fragmentation of a branched ion m/z 839.4 ($B_{3\beta}Y_{5\beta}$ from antenna bc) while m/z 431.2 can only arise from fragmentation of a linear m/z 839.4 ($B_{5a}Y_{7a}$) characteristic of antenna a. Moreover, the presence of tri-antenna was also confirmed by complementary ESI-MSⁿ experiments



(Additional file 4). From these data, two structures, Glc₃Man₈GlcNAc₂ or Glc₂Man₉GlcNAc₂, can be proposed. The genomic analysis favors the Glc₂Man₉GlcNAc₂ for the structure of the oligosaccharide released from the LLO. Indeed, bioinformatics analysis of *P. tricornutum* genome reveals the presence of all the enzymes involved in the synthesis of the precursor, except ALG10 and the GSI which are respectively involved in the addition and removal of the third glucose residue from the LLO precursor [16, 45]. Moreover, previous analysis of the *N*-glycans attached to the endogenous proteins of *P. tricornutum* showed a

high abundance of oligomannosidic *N*-glycans especially the one composed of 9 mannose residues [16]. Thus, all together, these results demonstrate that the structure of the oligosaccharide released from the LLO of *P. tricornutum* is a Glc₂Man₉GlcNAc₂ structure.

Conclusions

In this communication, we described a simple and user-friendly protocol to extract and analyze the oligosaccharide structure of the LLO from microalgae. This method allows to collect enough material to perform MS detailed

analyses, thus representing a powerful approach allowing identification of the structure of the LLO-released oligosaccharide. Moreover, the contribution of multistage tandem mass spectrometry to elucidate chemical structure of complex polymers is once more comforted. The advantage of this technique over radio-labelling methods is the access to the structural identification in close details of either known structures or those to elucidate. The validity of the method has been confirmed by comparison of wild-type and mutant data from *C. reinhardtii* and the elucidation of the LLO-released oligosaccharide structure of *P. tricornutum*. These results illustrate the adaptability of the optimized method as it allows successful identification of the LLO-released oligosaccharide of the diatom, *P. tricornutum* as well as the one from *C. reinhardtii* which belongs to the chlorophyceae lineage. Now, it could be conceived to use this user-friendly method to extract and analyse the LLO-released oligosaccharide structures in other microalgae, allowing understanding of the evolution of the LLO.

Methods

The water used throughout the protocol was deionized water. LC–MS grade water was purchased from Thermo-Scientific and was used for the purification of the LLO-released oligosaccharide. All the extraction steps up to the mild acidic hydrolysis were performed at 4 °C.

Chlamydomonas reinhardtii cell culture and lysis

The *C. reinhardtii* mutant strains *XTA* (LMJ, RY0402.087519) and *XTB* (LMJ, RY0402.118417) were ordered from the CLIP library (<https://www.chlamylibrary.org>). Those mutants were generated by random insertion of a paromomycin resistance cassette (CIB1) [46]. Cassette occupancy site in those two mutants had been checked as recommended in [46].

One liter of *C. reinhardtii* cell culture was grown in TAP medium (Tris Acetate Phosphate) at 23 °C, agitated at 150 rpm until reaching an optical density (OD) at 680 nm of 1. A centrifugation at 4500 × *g* using a Beckman Coulter Allegra® X-15R centrifuge with SX4750A rotor for 10 min was then performed in order to pellet the cells. The cells were washed twice with 20 mM phosphate buffer pH 7.4. Then, the cell pellets were resuspended in 6 mL of 20 mM phosphate buffer pH 7.4 containing 1 mM ethylene diamine tetraacetic acid, 0.5 mM dithiothreitol and 0.5 mM phenylmethylsulfonyl fluoride (PMSF). This solution was prepared extemporaneously. Cells were grinded using glass beads in the Lysing Matrix D tubes (MPbiomedicals) and a homogenizer (Fast Prep-24 5G™ High Speed Homogenizer, MPbiomedicals). Five cycles of 30 s at 6.5 m s⁻¹ with chilling for 2 min on ice between each cycle were performed. Samples were

then centrifuged at 10,000 × *g* for 10 min at 4 °C and the supernatant was transferred in ultracentrifugation tubes (Beckman Coulter). Each remaining pellet was resuspended in 1 mL of the same extraction buffer and then, 5 new cycles of 30 s at 6.5 m s⁻¹ with chilling for 2 min on ice between cycles were achieved. The resulting supernatants containing the cell lysate (S1) were pooled with the previous ones.

Phaeodactylum tricornutum cell culture and lysis

One liter of Pt 1.8.6 strain (CCAP 1055/1) was grown in sea water (33 g L⁻¹ of Instant Ocean Sel Marin, Tru-faut) enriched with 0.1% (v/v) of Conway medium and 0.2% (v/v) of 40 g L⁻¹ sodium silicate [16] until reaching 1.10⁷ cells mL⁻¹. The culture medium was then centrifuged at 4500 × *g* using Beckman Coulter Allegra® X-15R (with SX4750A rotor) for 10 min to pellet the cells. The resulting cell pellets were resuspended as described for *C. reinhardtii* cells. The diatom cells were grinded using a homogenizer (Fast Prep-24 5G™ High Speed Homogenizer, MPbiomedicals) and the Lysing Matrix E tubes (MPbiomedicals) performing 6 cycles of 30 s at 6.5 m s⁻¹ with chilling for 2 min on ice between cycles. Samples were then treated as per *C. reinhardtii*. The resulting supernatants contained the diatom cell lysate were named S1.

Microsomal fraction preparation

S1 was centrifuged at 20,000 × *g* at 4 °C in a Beckman Coulter ultracentrifuge Optima™ XL-100 K for 30 min using a type 70.1Ti rotor. The resulting pellet was discarded and the supernatant (S2) was centrifuged at 100,000 × *g* in the same conditions for 1 h. The new supernatant (S3) was discarded and the resulting pellet containing the microsomal fraction (MP) was conserved for further LLO purification steps.

Solvent extraction

The microsomal fraction was resuspended with 400 µL of water (LC–MS Grade, Fisher Chemical) in a glass tube, 800 µL of methanol (LC–MS Grade, Fisher Chemical) and 1.2 mL of chloroform (ACS reagent grade, Acros Organics) were added (1/2/3, v/v/v). The sample was then homogenized by vortex and centrifuged at 4000 × *g* at 4 °C for 2 min in the Beckman Coulter Allegra® X-15R centrifuge using the SX4750A rotor. After centrifugation, the sample exhibited a soluble upper phase (UP), an insoluble lower phase (LOW) separated by an interphase (INTER) (Fig. 1). The lower phase (LOW) containing pigments, was gently discarded using a glass pipette.

Three mL of the mixture A (Mix A in Fig. 1) containing chloroform/methanol and 4 mM MgCl₂ in ultrapure water (86/14/1, v/v/v) was then added to the remaining

upper phase. After mixing and centrifugation at $4000\times g$ for 2 min at 4 °C, the lower phase was discarded. This step was repeated 3 times to remove all pigments. The remaining upper phase (UP, Fig. 1) was then washed four times with 3 mL of the mixture B (Mix B in Fig. 1) containing methanol/4 mM $MgCl_2$ in water/chloroform (16/15/1, v/v/v). After homogenizing by vortex and centrifuging ($4000\times g$ at 4 °C for 2 min), the supernatant was discarded. This step was performed 3 times to remove soluble proteins. Finally, the pellet partially air dried, was re-suspended three times with 3 mL of the mixture C (Mix C in Fig. 1) containing chloroform/methanol/water (10/10/3, v/v/v). After homogenization by vortex and centrifugation ($4000\times g$ at 4 °C for 2 min), the supernatant containing the LLO was transferred in a new glass tube and dried with air stream.

Mild acid treatment

In order to remove the dolichol lipid moiety from the oligosaccharide, dried LLO was hydrolyzed by a mild acid treatment using 0.1 M of TFA (LC-MS grade, Thermo Scientific) at 80 °C for 1 h. In these conditions, the glycosidic bond between the dolichol phosphate and the oligosaccharide was efficiently hydrolyzed. Released oligosaccharides, later called LLO-released oligosaccharide, were then lyophilized.

LLO oligosaccharides purification

LLO-released oligosaccharides were resuspended in 500 μL of water (LC-MS grade, Fisher Chemical) and then purified using carbograph columns (Hypersep™ Hypercarb™ SPE Cartridges, 200 mg, 3 mL, ThermoFisher Scientific). First, the columns were conditioned by successive washing with 3 mL of 1 M sodium hydroxide (Laboratory reagent grade, Fisher scientific), 6 mL of water (LC-MS grade, Fisher Chemical), 3 mL of acetic acid 30% (v/v) Acros Organics) and finally 3 mL of water (LC-MS grade, Fisher Chemical). These steps may be carried out under vacuum (5 mmHg) using a Manifold. Then, the columns were successively washed with 3 mL of 50% acetonitrile in water (v/v) (LC-MS grade, Fisher Chemical), 6 mL of 5% acetonitrile in water (v/v) and finally with 6 mL of water (LC-MS grade, Fisher Chemical). The LLO-released oligosaccharide samples were resuspended in 500 μL of water (LC-MS grade, Fisher Chemical) and were loaded into the column. After washing with 3 mL of water and 3 mL of 5% acetonitrile in water (v/v), LLO-released oligosaccharides were eluted by 4×0.5 mL of 50% acetonitrile in 0.1 M of TFA (v/v). The eluted fractions were recovered and were lyophilized in glass tubes prior to permethylation.

Permethylation

A suspension of sodium hydroxide in dimethylsulfoxide was prepared by grinding about 0.90 g of dried sodium hydroxide (4 tablets) in 3 mL of dimethylsulfoxide (Sigma-Aldrich). 0.6 mL of this preparation was added to the lyophilized sample. Then, 0.5 mL of iodomethane (ACS reagent grade, Acros Organics) was added and the sample was incubated under agitation for 1 h. One mL of water was added to quench the reaction. One mL of chloroform and 3 mL of water (LC-MS grade, Thermo Scientific) were then added to the samples and mixed. When two phases appeared clearly, the upper aqueous phase was discarded, and 3 mL of water were added. This step was repeated until pH was neutralized. The chloroform phase containing permethylated glycans was then dried.

Permethylated LLO oligosaccharides purification

After permethylation, the LLO-released oligosaccharide was purified using C18 columns (Hypersep™ C18 Cartridges, 200 mg, 3 mL, Thermo Fisher). Columns were conditioned by successive washing with 5 mL of methanol, 5 mL of water (LC-MS grade, Fisher Chemical), 5 mL of acetonitrile and finally with 5 mL of water (LC-MS grade, Fisher Chemical). Permethylated sample was resuspended in 200 μL of 80% methanol (v/v) and were loaded on the column. Successive elutions with 2 mL of 15% (v/v), 35% (v/v) and 75% acetonitrile (v/v) were performed. Permethylated oligosaccharides were recovered in the 75% acetonitrile fraction. The resulting samples were then dried in a ThermoFisher SPD111 V SpeedVac®. After drying, the samples were ready to be analyzed using the Matrix Assisted Laser Desorption Ionization—Time Of Flight Mass Spectrometry and ElectroSpray Ionization Multistage tandem Mass Spectrometry (ESI-MSⁿ).

Matrix assisted laser desorption ionization—mass spectrometry (MALDI-TOF-MS)

MALDI-TOF-MS experiments were performed using an Autoflex III time-of-flight mass spectrometer (Bruker Daltonics (Bremen, Germany) equipped with a Nd: YAG laser (355-nm wavelength), FlexControl 3.3 and FlexAnalysis 3.3 software package. Spectra were acquired in the positive-ion reflector mode using a mass range m/z 500–4000, 100 Hz laser frequency, 10 ns pulsed ion extraction delay and 50% attenuation laser. At least 3 000 individual spectra were averaged. The instrument accelerating voltage was 19 kV. External calibration of MALDI-TOF mass spectra was carried out using singly charged monoisotopic peaks of angiotensin II (5 pmol μL^{-1}), substance P (5 pmol μL^{-1}), ACTH 18–39 fragment (5 pmol μL^{-1}) and insulin B (10 pmol μL^{-1}) in the same experimental (matrix conditions, sample deposit,...) and instrumental

conditions (accelerating voltage, laser frequency, ion extraction delay, attenuation laser,...) than those used for the LLO-released oligosaccharide sample analyses. 2,5-dihydroxybenzoic acid (DHB) was used as matrix: 1 μL of sample was mixed within 1 μL of matrix solution (10 mg mL^{-1} in 0.1% aqueous TFA/acetonitrile (70/30: v/v)), then 0.7 μL of the sample/matrix mixture was spotted and dried onto the MALDI polished stainless-steel target plate (Bruker Daltonics, Bremen, Germany).

Electrospray ionization multistage tandem mass spectrometry (ESI-MSⁿ)

ESI-MSⁿ ($n=1$ to 4) experiments were performed using a Bruker HCT Ultra ETD II quadrupole ion trap (QIT) mass spectrometer equipped with an electrospray source, the Esquire control 6.2 and Data Analysis 4.0 softwares (Bruker Daltonics, Bremen, Germany). The ESI parameters were as followed: capillary and end plate voltages respectively set at -4.0 kV and -3.5 kV in positive ion mode, skimmer and capillary exit voltages set at 40 V and 120 V, respectively, nebulizer gas (N_2), pressure, drying gas (N_2) flow rate and drying gas temperature were 10 psi, 7.0 L min^{-1} and 300 °C, respectively. The data were acquired using a mass range of either m/z 200–2200 range (for $\text{Hex}_8\text{HexNAc}_2$) or m/z 200–2850 range (for $\text{Hex}_{11}\text{HexNAc}_2$) and using a scan speed of m/z 8100 per second. The number of ions entering the trap cell was automatically adjusted by controlling the accumulation time using the ion charge control (ICC) mode (target 100,000) with a maximum accumulation time of 50 ms. The injection low-mass cut-off (LMCO) value was m/z 90 (for $\text{Hex}_8\text{HexNAc}_2$) or m/z 140 for $\text{Hex}_{11}\text{HexNAc}_2$ derivatives. The values of spectra average and rolling average were 6 and 2, respectively. ESI-MSⁿ experiments were carried out by collision-induced dissociation (CID) using helium as the collision gas, isolation width of 1 or 1.5 m/z unit for the precursor ions and for the intermediate ions using a resonant excitation frequency with an amplitude from 0.8 to 1.0 Vp-p. Sample solutions (resuspension in 50 μL of $\text{CH}_3\text{CN}/\text{H}_2\text{O}$) were infused into the source at a flow rate of 180 $\mu\text{L h}^{-1}$ by means of a syringe pump (Cole-Palmer, Vernon Hills, IL, USA). External calibration was performed using the 'tuning mix' from Agilent Technologies (Santa Rosa, CA, USA). All mass spectra were manually interpreted, based on literature [23, 40, 41]. The Glycoworkbench software v2.1 was used to support and confirm the assignment of the fragment ions. The ESI-MSⁿ spectra showed diagnostic fragment ions resulting from the characteristic fragmentation pathways of a specific chemical structure and consequently helped in identifying the oligosaccharide structure according to Prien et al. [40]. The extraction of LLO and MS analyses have been performed at least twice independently.

Additional files

Additional file 1. Multistage tandem mass spectrometry analyses of the structure of the LLO-released oligosaccharide isolated from *C. reinhardtii* XTb mutant by ESI-MSⁿ. ESI-MSⁿ spectra with $n=2$ (panel A), $n=3$ (panel B) and $n=3$ (panel C) of permethylated $\text{Hex}_8\text{HexNAc}_2$ derivative (m/z 1107.6 corresponding to $[\text{M}+2\text{Na}]^{2+}$ precursor ion) isolated from *C. reinhardtii* XTb mutant. On each panel, the ion selected for the fragmentation analysis is shown with a diamond and its fragmentation pattern is proposed according to Prien et al. [40]. Black square: *N*-acetylglucosamine; grey circle: mannose, black circle: glucose. The fragment ions are labelled according to the nomenclature of Domon and Costello [41].

Additional file 2. Scheme representing a fragmentation pathway complementary to the one depicted in Fig. 5. Cleavages of the glycosidic bond and cross ring cleavages are represented by dotted lines. The fragment ions are labelled according to the nomenclature of Domon and Costello [41].

Additional file 3. Multistage tandem mass spectrometry analyses of the structure of the LLO-released oligosaccharide isolated from *P. tricornutum*. ESI-MSⁿ spectra with $n=3$ (panel A) and $n=4$ (panel B) of permethylated $\text{Hex}_{11}\text{HexNAc}_2$ derivative (m/z 1413.5 corresponding to $[\text{M}+2\text{Na}]^{2+}$ precursor ion) isolated from *P. tricornutum*. On each panel, the ion selected for the fragmentation analysis is shown with a diamond and its fragmentation pattern is proposed according to Prien et al. [40]. Black square: *N*-acetylglucosamine; grey circle: mannose, black circle: glucose. The fragment ions are labelled according to the nomenclature of Domon and Costello [41]. Note that loss of 204 u correspond to the elimination of a 1,2-; 1,3-; 1,4- or 1,6-linked hexose residue while loss of 218 u correspond to the elimination of a terminal hexose residue. Therefore, the successive losses of two 218 u (m/z 1057.5 \rightarrow m/z 839.4 \rightarrow m/z 621.3) indicate a di-antenna ion m/z 1057.5 while the successive losses of 204 u (m/z 1057.5 \rightarrow m/z 853.4 \rightarrow m/z 649.3) indicate a linear ion m/z 1057.5.

Additional file 4. Multistage tandem mass spectrometry analyses of the structure of the LLO-released oligosaccharide isolated from *P. tricornutum*. ESI-MSⁿ spectra with $n=3$ (panels A, C and E) and $n=4$ (panel B and D) of permethylated $\text{Hex}_{11}\text{HexNAc}_2$ derivative (m/z 1413.5 corresponding to $[\text{M}+2\text{Na}]^{2+}$ precursor ion) isolated from *P. tricornutum*. On each panel, the ion selected for the fragmentation analysis is shown with a diamond and its fragmentation pattern is proposed according to Prien et al. [40]. Black square: *N*-acetylglucosamine; grey circle: mannose, black circle: glucose. The fragment ions are labelled according to the nomenclature of Domon and Costello [41]. Note that loss of 190 u correspond to the elimination of a hexose residue which was linked with three other residues in the oligosaccharide (for example, a 1,3,6-linked hexose residue). Therefore, the successive losses of two 218 u (m/z 1247.6 \rightarrow m/z 1029.5 \rightarrow m/z 811.5) indicate a di-antenna ion m/z 1247.6 while the loss of 190 u followed by successive losses of 204 u (m/z 1247.6 \rightarrow m/z 1057.5 \rightarrow m/z 853.5 \rightarrow m/z 649.4 \rightarrow m/z 445.3) indicate a linear ion m/z 1247.6.

Abbreviations

ALG: asparagine-linked glycosylation; ESI: electrospray ionization; Glc: glucose; GlcNAc: *N*-acetylglucosamine; GS: alpha glucosidase I; Hex: hexose; LLO: lipid-linked oligosaccharide; Man: mannose; MALDI-TOF: matrix assisted laser desorption ionization—time of flight; MS: mass spectrometry; PMSF: phenylmethylsulfonyl fluoride; PP-Dol: dolichol pyrophosphate; XTA: xylosyltransferase mutant A; XTb: xylosyltransferase mutant B.

Authors' contributions

PLL, MB designed the method; PLL, RD, CLB performed the experiments; PLL, RD, CLB, CA, AM, NMB, PL, MB analyzed the results and wrote the paper. All authors read and approved the final manuscript.

Author details

¹ UNIROUEN, Laboratoire Glyco-MEV EA4358, Normandie Univ, 76000 Rouen, France. ² UNIROUEN, INSA Rouen, CNRS, COBRA, Normandie Univ, 76000 Rouen, France. ³ Institut Universitaire de France (IUF), 75000 Paris, France.

Acknowledgements

Authors would like to thank Dr Elodie Mathieu-Rivet for her help in screening the *C. reinhardtii* mutant clones.

Competing interests

The authors declare to have no competing of interests.

Availability of data and materials

Data and material would be made available upon request.

Consent for publication

All authors have consented for publication.

Ethics approval and consent to participate

Non applicable.

Funding

The authors thank the European Regional Development Fund (ERDF, N°31708), the Region Normandie (Crunch Network, N°20 – 13), the Labex SynOrg (ANR-11-LABX-0029), the Institut Universitaire de France (I.U.F.), the University of Rouen Normandie for financial support.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 30 March 2018 Accepted: 23 November 2018

Published online: 05 December 2018

References

- Varki A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*. 1993;3:97–130.
- Lingg N, Zhang P, Song Z, Bardor M. The sweet tooth of biopharmaceuticals: importance of recombinant protein glycosylation analysis. *Biotechnol J*. 2012;7:1462–72.
- Strasser R, Altmann F, Mach L, Glössl J, Steinkellner H. Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking β 1,2-linked xylose and core α 1,3-linked fucose. *FEBS Lett*. 2004;561:132–6.
- Strasser R, Stadlmann J, Schähls M, Stiegler G, Quendler H, Mach L, et al. Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol J*. 2008;6:392–402.
- Strasser R. Biological significance of complex N-glycans in plants and their impact on plant physiology. *Front Plant Sci* [Internet]. 2014 [cited 2018 Feb 20];5. <https://www.frontiersin.org/articles/10.3389/fpls.2014.00363/full>.
- Strasser R. Plant protein glycosylation. *Glycobiology*. 2016;26:926–39.
- Schoberer J, Strasser R. Plant glyco-biotechnology. *Semin Cell Dev Biol* [Internet]. 2017 [cited 2018 Feb 16]. <http://www.sciencedirect.com/science/article/pii/S1084952116303603>.
- Schoberer J, Shin Y-J, Vavra U, Veit C, Strasser R. Analysis of protein glycosylation in the ER. *Plant endoplasmic reticulum* [Internet]. Humana Press, New York, NY; 2018 [cited 2018 Feb 16]. p. 205–22. https://link.springer.com/protocol/10.1007/978-1-4939-7389-7_16.
- Aebi M. N-linked protein glycosylation in the ER. *Biochim Biophys Acta BBA Mol Cell Res*. 2013;1833:2430–7.
- Mohorko E, Glockshuber R, Aebi M. Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. *J Inher Metab Dis*. 2011;34:869–78.
- Breitling J, Aebi M. N-Linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol*. 2013;5:a013359–a013359.
- Määttänen P, Gehring K, Bergeron JJM, Thomas DY. Protein quality control in the ER: the recognition of misfolded proteins. *Semin Cell Dev Biol*. 2010;21:500–11.
- Ruggiano A, Foresti O, Carvalho P. ER-associated degradation: protein quality control and beyond. *J Cell Biol*. 2014;204:869–79.
- Fitchette-Lainé AC, Gomord V, Cabanes M, Michalski JC, Saint Macary M, Foucher B, et al. N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J Cell Mol Biol*. 1997;12:1411–7.
- Fitchette A-C, Cabanes-Macheteau M, Marvin L, Martin B, Satiat-Jeunemaitre B, Gomord V, et al. Biosynthesis and immunolocalization of Lewis a-Containing N-Glycans in the plant cell. *Plant Physiol*. 1999;121:333–44.
- Baïet B, Burel C, Saint-Jean B, Louvet R, Menu-Bouaouiche L, Kiefer-Meyer M-C, et al. N-Glycans of *Phaeodactylum tricoratum* diatom and functional characterization of its N-acetylglucosaminyltransferase I. *Enzyme. J Biol Chem*. 2011;286:6152–64.
- Levy-Ontman O, Arad SM, Harvey DJ, Parsons TB, Fairbanks A, Tekoah Y. Unique N-glycan moieties of the 66-kDa cell wall glycoprotein from the red microalga *Porphyridium* sp. *J Biol Chem*. 2011;286:21340–52.
- Mathieu-Rivet E, Scholz M, Arias C, Dardelle F, Schulze S, Le Mauff F, et al. Exploring the N-glycosylation pathway in *Chlamydomonas reinhardtii* unravels novel complex structures. *Mol Cell Proteomics MCP*. 2013;12:3160–83.
- Levy-Ontman O, Fisher M, Shotland Y, Weinstein Y, Tekoah Y, Arad SM. Genes involved in the endoplasmic reticulum N-glycosylation pathway of the Red *Microalga porphyridium* sp.: a bioinformatic study. *Int J Mol Sci*. 2014;15:2305–26.
- Schulze S, Urzica E, Reijnders MJMF, van de Geest H, Warris S, Bakker LV, et al. Identification of methylated GnTI-dependent N-glycans in *Botryococcus braunii*. *New Phytol*. 2017;215:1361–9.
- Dumontier R, Mareck A, Mati-Baouche N, Bardor PL and M. Toward future engineering of the N-glycosylation pathways in microalgae for optimizing the production of biopharmaceuticals. *Microalgal Biotechnol* [Internet]. 2018 [cited 2018 Sep 3]. <https://www.intechopen.com/books/microalgal-biotechnology/toward-future-engineering-of-the-n-glycosylation-pathways-in-microalgae-for-optimizing-the-productio>.
- Mathieu-Rivet E, Kiefer-Meyer M-C, Vanier G, Ovide C, Burel C, Lerouge P, et al. Protein N-glycosylation in eukaryotic microalgae and its impact on the production of nuclear expressed biopharmaceuticals. *Front Plant Sci*. 2014;5:359.
- Vanier G, Lucas P-L, Loutelier-Bourhis C, Vanier J, Plasson C, Walet-Balieu M-L, et al. Heterologous expression of the N-acetylglucosaminyltransferase I dictates a reinvestigation of the N-glycosylation pathway in *Chlamydomonas reinhardtii*. *Sci Rep*. 2017;7:10156.
- Garénaux E, Shams-Eldin H, Chirat F, Bieker U, Schmidt J, Michalski J-C, et al. The dual origin of toxoplasma gondii N-glycans. *Biochemistry*. 2008;47:12270–6.
- Haserick JR, Leon DR, Samuelson J, Costello CE. Asparagine-linked glycans of *Cryptosporidium parvum* contain a single long arm, are barely processed in the endoplasmic reticulum (ER) or golgi, and show a strong bias for sites with threonine. *Mol Cell Proteomics*. 2017;16:S42–53.
- Li E, Tabas I, Kornfeld S. The synthesis of complex-type oligosaccharides. I. Structure of the lipid-linked oligosaccharide precursor of the complex-type oligosaccharides of the vesicular stomatitis virus G protein. *J Biol Chem*. 1978;253:7762–70.
- Chapman A, Li E, Kornfeld S. The biosynthesis of the major lipid-linked oligosaccharide of Chinese hamster ovary cells occurs by the ordered addition of mannose residues. *J Biol Chem*. 1979;254:10243–9.
- Villers C, Cacan R, Mir AM, Labiau O, Verbert A. Release of oligomannoside-type glycans as a marker of the degradation of newly synthesized glycoproteins. *Biochem J*. 1994;298:135–42.
- Pabst M, Grass J, Toegel S, Liebming E, Strasser R, Altmann F. Isomeric analysis of oligomannosidic N-glycans and their dolichol-linked precursors. *Glycobiology*. 2012;22:389–99.
- Salleh AB, Rahman NZRA, Basri M. *New lipases and proteases*. New York: Nova Publishers; 2006.
- Ranjbar M, Zibae A, Sendi JJ. Purification and characterization of a digestive lipase in the midgut of *Ectomyeloid ceratoniae* Zeller (Lepidoptera: Pyralidae). *Front Life Sci*. 2015;8:64–70.
- Ciucanu I, Kerek F. A simple and rapid method for the permethylation of carbohydrates. *Carbohydr Res*. 1984;131:209–17.
- Kang P, Mechref Y, Novotny MV. High-throughput solid-phase permethylation of glycans prior to mass spectrometry. *Rapid Commun Mass Spectrometry RCM*. 2008;22:721–34.
- Mechref Y, Novotny MV. Structural investigations of glycoconjugates at high sensitivity. *Chem Rev*. 2002;102:321–69.

35. Matamoros Fernández LE, Obel N, Scheller HV, Roepstorff P. Characterization of plant oligosaccharides by matrix-assisted laser desorption/ionization and electrospray mass spectrometry. *J Mass Spectrom.* 2003;38:427–37.
36. Matamoros LF, Obel N, Scheller HV, Roepstorff P. Differentiation of isomeric oligosaccharide structures by ESI tandem MS and GC-MS. *Carbohydr Res.* 2004;339:655–64.
37. Plancot B, Vanier G, Maire F, Bardor M, Lerouge P, Farrant JM, et al. Structural characterization of arabinoxylans from two African plant species *Eragrostis nindensis* and *Eragrostis tef* using various mass spectrometric methods. *Rapid Commun Mass Spectrom.* 2014;28:908–16.
38. Mathieu-Rivet E, Lerouge P, Bardor M. *Chlamydomonas reinhardtii*: protein glycosylation and production of biopharmaceuticals. *Chlamydomonas Biotechnol Biomed* [Internet]. Springer, Cham; 2017 [cited 2018 Feb 28]. p. 45–72. https://link.springer.com/chapter/10.1007/978-3-319-66360-9_3.
39. Schulze S, Oltmanns A, Machnik N, Liu G, Xu N, Jarmatz N, et al. N-Glycoproteomic characterization of mannosidase and xylotransferase Mutant Strains of *Chlamydomonas*. *Plant Physiol.* 2018.
40. Prien JM, Ashline DJ, Lapadula AJ, Zhang H, Reinhold VN. The high mannose glycans from bovine ribonuclease B isomer characterization by ion trap MS. *J Am Soc Mass Spectrom.* 2009;20:539–56.
41. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj J.* 1988;5:397–409.
42. Burda P, Aebi M. The ALG10 locus of *Saccharomyces cerevisiae* encodes the α -1,2 glucosyltransferase of the endoplasmic reticulum: the terminal glucose of the lipid-linked oligosaccharide is required for efficient N-linked glycosylation. *Glycobiology.* 1998;8:455–62.
43. Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem.* 2004;73:1019–49.
44. Burda P, Aebi M. The dolichol pathway of N-linked glycosylation. *Biochim Biophys Acta BBA.* 1999;1426:239–57.
45. Gügi B, Le Costaouec T, Burel C, Lerouge P, Helbert W, Bardor M. Diatom-specific oligosaccharide and polysaccharide structures help to unravel biosynthetic capabilities in diatoms. *Mar Drugs.* 2015;13:5993–6018.
46. Zhang R, Patena W, Armbruster U, Gang SS, Blum SR, Jonikas MC. High-throughput genotyping of green algal mutants reveals random distribution of mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. *Plant Cell.* 2014;26:1398–409.

Ready to submit your research? Choose BMC and benefit from:

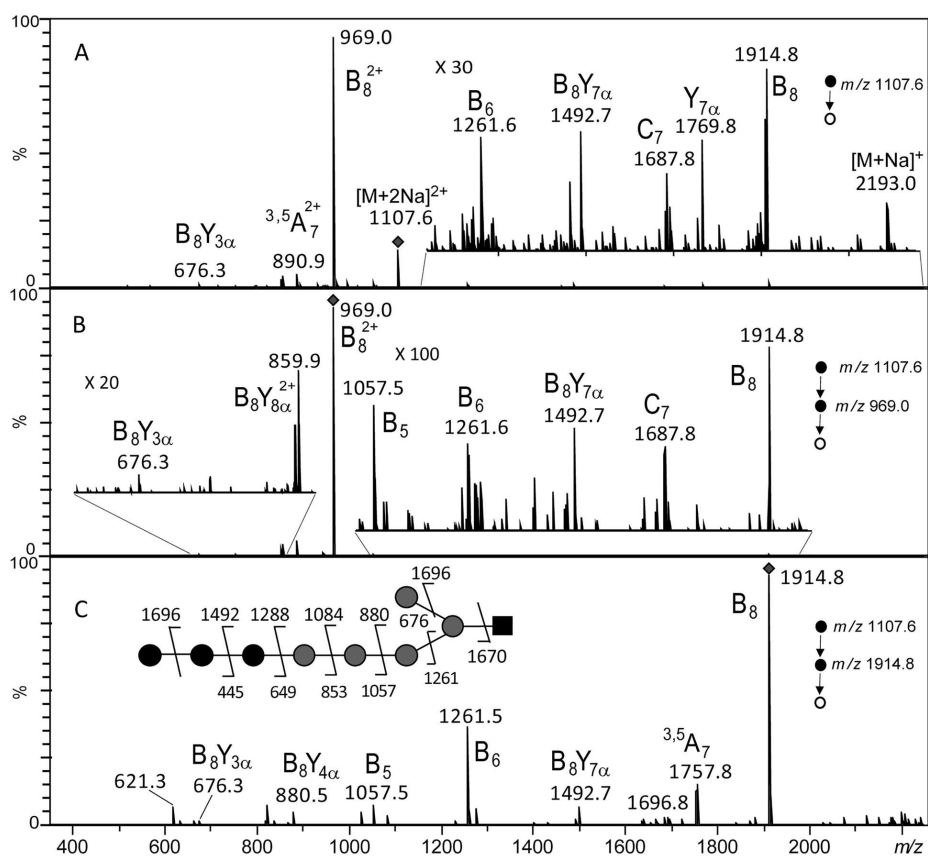
- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

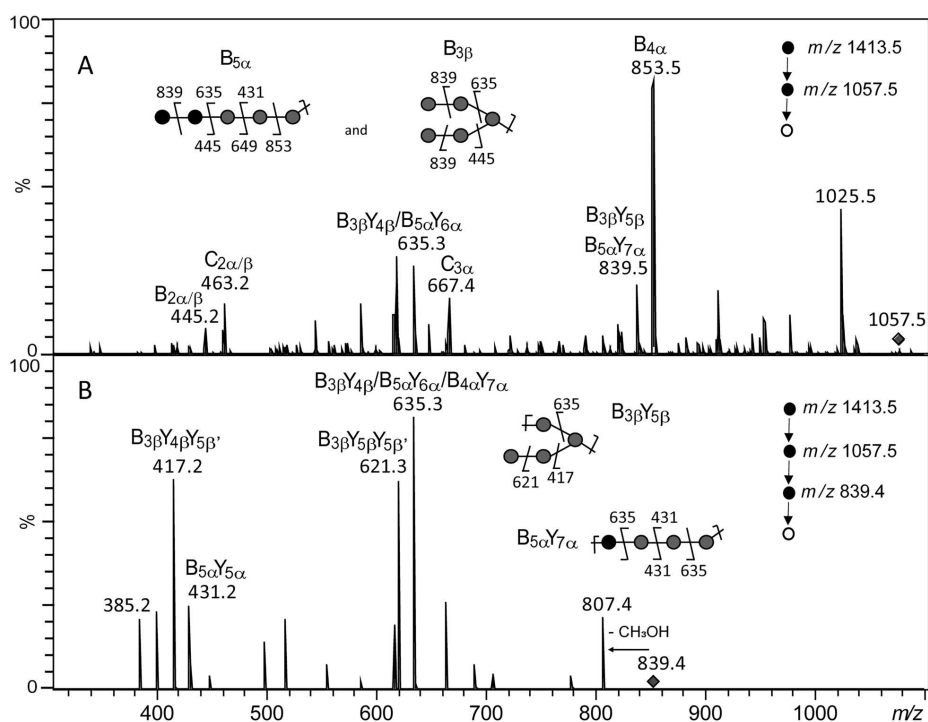
Learn more biomedcentral.com/submissions



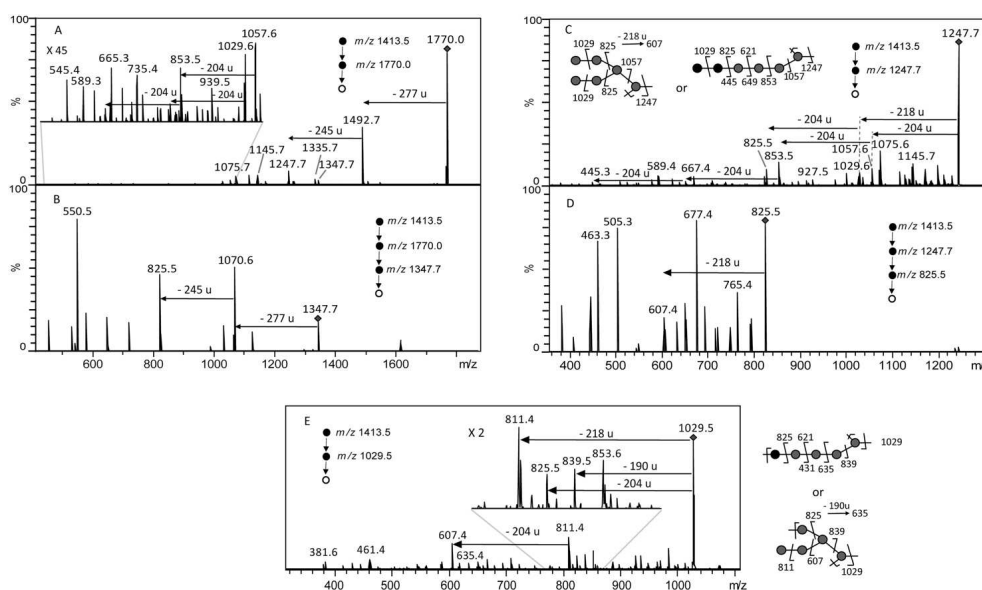
Figure supplémentaire du chapitre I



Additional file 1. Multistage tandem mass spectrometry analyses of the structure of the LLO-released oligosaccharide isolated from *C. reinhardtii* XTb mutant by ESI-MSn. ESI-MSn spectra with $n=2$ (panel A), $n=3$ (panel B) and $n=3$ (panel C) of permethylated Hex₈HexNAC₂ derivative (m/z 1107.6 corresponding to $[M + 2Na]^{2+}$ precursor ion) isolated from *C. reinhardtii* XTb mutant. On each panel, the ion selected for the fragmentation analysis is shown with a diamond and its fragmentation pattern is proposed according to Prien et al. [40]. Black square: *N*-acetylglucosamine; grey circle: mannose, black circle: glucose. The fragment ions are labelled according to the nomenclature of Domon and Costello [41].



Additional file 3. Multistage tandem mass spectrometry analyses of the structure of the LLO-released oligosaccharide isolated from *P. tricornutum*. ESI-MSⁿ spectra with $n = 3$ (panel A) and $n = 4$ (panel B) of permethylated Hex₁₁HexNAc₂ derivative (m/z 1413.5 corresponding to $[M + 2Na]^{2+}$ precursor ion) isolated from *P. tricornutum*. On each panel, the ion selected for the fragmentation analysis is shown with a diamond and its fragmentation pattern is proposed according to Prien et al. [40]. Black square: *N*-acetylglucosamine; grey circle: mannose, black circle: glucose. The fragment ions are labelled according to the nomenclature of Domon and Costello [41]. Note that loss of 204 u correspond to the elimination of a 1,2-; 1,3-; 1,4- or 1,6-linked hexose residue while loss of 218 u correspond to the elimination of a terminal hexose residue. Therefore, the successive losses of two 218 u (m/z 1057.5 \rightarrow m/z 839.4 \rightarrow m/z 621.3) indicate a di-antenna ion m/z 1057.5 while the successive losses of 204 u (m/z 1057.5 \rightarrow m/z 853.4 \rightarrow m/z 649.3) indicate a linear ion m/z 1057.5.



Additional file 4. Multistage tandem mass spectrometry analyses of the structure of the LLO-released oligosaccharide isolated from *P. tricornutum*. ESI-MSⁿ spectra with $n = 3$ (panels A, C and E) and $n = 4$ (panel B and D) of permethylated Hex₁₁HexNAC₂ derivative (m/z 1413.5 corresponding to $[M + 2Na]^{2+}$ precursor ion) isolated from *P. tricornutum*. On each panel, the ion selected for the fragmentation analysis is shown with a diamond and its fragmentation pattern is proposed according to Prien et al. [40]. Black square: N-acetylglucosamine; grey circle: mannose, black circle: glucose. The fragment ions are labelled according to the nomenclature of Domon and Costello [41]. Note that loss of 190 u correspond to the elimination of a hexose residue which was linked with three other residues in the oligosaccharide (for example, a 1,3,6-linked hexose residue). Therefore, the successive losses of two 218 u (m/z 1247.6 \rightarrow m/z 1029.5 \rightarrow m/z 811.5) indicate a di-antenna ion m/z 1247.6 while the loss of 190 u followed by successive losses of 204 u (m/z 1247.6 \rightarrow m/z 1057.5 \rightarrow m/z 853.5 \rightarrow m/z 649.4 \rightarrow m/z 445.3) indicate a linear ion m/z 1247.6.

B. Analyse structurale des *N*-glycannes oligomannosidiques synthétisés chez la diatomée *P. tricornutum*

Ce second chapitre de résultats, présenté sous la forme d'un article scientifique en préparation, décrit l'analyse structurale détaillée des isomères des *N*-glycannes oligomannosidiques synthétisés chez *P. tricornutum*. Une analyse structurale plus poussée de la structure du $\text{Man}_5\text{GlcNAc}_2$ a été réalisée afin de déterminer si celui-ci possède une structure stéréochimique identique à celle de l'isomère $\text{Man}_5\text{GlcNAc}_2$ dit « canonique » qui est le substrat accepteur de l'enzyme golgienne GnT I chez les mammifères. Cet article sera soumis dans le journal « Analytical Chemistry ».

Chez la majorité des eucaryotes, les voies de la *N*-glycosylation aboutissent à la synthèse dans le RE d'un *N*-glycane de type oligomannosidique de structure $\text{Man}_9\text{GlcNAc}_2$ identique chez les mammifères, les insectes et les plantes. Ce *N*-glycane est ensuite transformée par l'élimination de résidus mannose par la Man I formant successivement des structures $\text{Man}_8\text{GlcNAc}_2$, $\text{Man}_7\text{GlcNAc}_2$ et $\text{Man}_6\text{GlcNAc}_2$ existant chacune sous la forme de plusieurs isomères. La dernière étape réalisée par la Man I aboutit normalement à la formation d'un *N*-glycane dit « canonique » de type $\text{Man}_5\text{GlcNAc}_2$ qui est le précurseur de la synthèse des *N*-glycannes complexes chez la majorité des eucaryotes.

Dans ce chapitre de résultats, nous avons d'abord voulu identifier les différents isomères des *N*-glycannes oligomannosidiques synthétisés par la diatomée. Ces analyses structurales ont été réalisées en combinant les approches de spectrométrie de masse, ESI-MSⁿ et IMS-MS, et en comparant la structure des *N*-glycannes oligomannosidiques de *P. tricornutum* avec les structures des *N*-glycannes oligomannosidiques de la ribonucléase B bovine. Cette approche a permis de déterminer la structure d'un isomère unique pour les *N*-glycannes oligomannosidiques, $\text{Man}_9\text{GlcNAc}_2$ et $\text{Man}_5\text{GlcNAc}_2$, structures identiques à celles retrouvées chez les mammifères (**Fu et al., 1994**). Ces analyses ont également déterminé que les *N*-glycannes de type $\text{Man}_8\text{GlcNAc}_2$, $\text{Man}_7\text{GlcNAc}_2$ et $\text{Man}_6\text{GlcNAc}_2$ peuvent exister sous la forme de plusieurs isomères au sein de la diatomée.

L'analyse structurale détaillée réalisée sur le $\text{Man}_5\text{GlcNAc}_2$ a permis de montrer que, contrairement à d'autres microalgues qui synthétisent un Man 5 linéaire (**Vanier *et al.*, 2017 ; Lucas *et al.*, 2020**), *P. tricornutum* synthétise un Man 5 « canonique » substrat accepteur de l'enzyme golgienne GnT I. La formation de cet isomère « canonique » est primordiale comme base moléculaire pour la synthèse des *N*-glycannes complexes chez l'Homme mais aussi chez la diatomée dans l'optique de l'humanisation de ses voies de la *N*-glycosylation des protéines.

Title: Towards deciphering structural features of the oligomannoside isomers of *N*-glycans in the diatom, *Phaeodactylum tricornutum* using a combination of cutting-edge mass spectrometry techniques

Rodolphe Dumontier^{1,2,‡}, Corinne Loutelier-Bourhis^{3,‡}, Marie-Laure Walet-Balieu^{1,2}, Carole Burel^{1,2}, Alain Mareck^{1,2}, Carlos Afonso³, Patrice Lerouge^{1,2,‡,*} and Muriel Bardor^{1,2,4,5,‡,*}

¹ Normandie Univ, UNIROUEN, Laboratoire Glycobiologie et Matrice Extracellulaire végétale (Glyco-MEV) EA4358, 76000 Rouen, France ;

² Normandie Univ, UNIROUEN, SFR NORVEGE, 76000 Rouen, France ;

³ Normandie Univ, UNIROUEN, INSA Rouen, CNRS, COBRA (UMR6014), 76000 Rouen, France ;

⁴ Institut Universitaire de France (IUF), 75000 Paris, France.

⁵ Univ Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F- 59000 Lille, France ;

* Corresponding authors: P.L. e-mail: patrice.lerouge@univ-rouen.fr and M.B. e-mail: muriel.bardor@univ-rouen.fr

‡ Equal contribution of these authors

Key words: IMS-MS, glycans, microalgae, oligomannosides, ESI-MSⁿ

Running Title: Isomers of Oligomannosides in *Phaeodactylum tricornutum*

Abstract

Proteins from microalgae carry *N*-linked glycans that result from the transfer and processing of a lipid-linked oligosaccharide precursor onto the protein in the endoplasmic reticulum (ER) and then processing of the protein *N*-glycan in the Golgi apparatus as reported for other eukaryotic cells. This maturation processing gives rise to a wide diversity of *N*-glycan structures including oligomannosides derived from the trimming of the precursor. Microalgae are currently considered as emerging expression

systems for the production of recombinant proteins useful for therapeutic applications in human. In this context, evaluation of microalgae-made biopharmaceuticals, which are often glycoprotein-based pharmaceuticals products, will require efficient analytical methodologies dedicated to the profiling of the recombinant protein *N*-glycosylation. Moreover, to ensure a good batch-to-batch reproducibility of the biopharmaceuticals produced in microalgae, a good understanding of the regulation of their *N*-glycosylation pathway is necessary. Herein, we described the profiling of the oligomannosides isolated from the diatom *Phaeodactylum tricornutum* that has already been used successfully to produce functional recombinant monoclonal antibodies. In this work, the combination of Ion Mobility Spectrometry–Mass Spectrometry (IMS-MS) and Electrospray Ionization-multistage tandem mass spectrometry (ESI-MSⁿ) allows us to decipher the structural features of the isomers of oligomannosides synthesized by the diatom.

Introduction

N-linked glycosylation is the most common eukaryotic post-translational modification of secreted proteins¹. It results from the attachment of an oligosaccharide onto the asparagine residues belonging to the consensus sequence Asn-X-Ser/Thr/Cys with X being any amino acid except proline^{2,3,4,5}. This process starts in the endoplasmic reticulum (ER) with the biosynthesis of a lipid-linked oligosaccharide (LLO) precursor composed of a Glc₃Man₉GlcNAc₂ oligosaccharide linked to a membrane-anchor dolichol pyrophosphate (Dol-PP). This LLO is then, enzymatically transferred onto the asparagine residue of the *N*-glycosylation site of the nascent proteins through the action of the oligosaccharyltransferase (OST)⁶. Trimming of glucose residues together with the interactions between the glycoprotein and ER-resident chaperones ensure the folding and quality control of the glycoprotein⁷. The glycoproteins then enter the Golgi apparatus where their *N*-glycans are processed through the action of the α -mannosidase I into a unique canonical Man₅GlcNAc₂ (Man-5). Subsequently, *N*-acetylglucosaminyltransferase I (GnT I), an α -mannosidase II and a *N*-acetylglucosaminyltransferase II (GnT II) give rise to the GlcNAc₂Man₃GlcNAc₂ core, which is common to mammals, insects and land plants^{8,9,10}. This core undergoes further processing into organism specific complex-type *N*-glycans that are involved in multiple essential physiological and biological functions¹¹.

Nowadays, protein *N*-glycosylation in microalgae is gaining an increasing interest. Several publications have reviewed the use of microalgae for the production of biopharmaceuticals^{12,13,14,15,16}. In this marine biotech context, the understanding of the *N*-glycosylation pathway in microalgae becomes crucial as the production of recombinant proteins for therapeutic applications in human will require the engineering of the microalgal endogenous *N*-glycosylation pathway for the production of biopharmaceuticals exhibiting human-compatible *N*-glycans^{17,18,19}. Recent scientific articles describing the structure of *N*-glycans from different microalgae species have revealed that these unicellular organisms synthesize a wide diversity of *N*-glycan structures, far exceeding those of plant and insect cells^{8,9}. Indeed, in addition to oligomannosides ranging from Man-5 to Man₉GlcNAc₂ (Man-9) that result from the trimming of the ER LLO precursor, microalgae synthesize mature oligosaccharides consisting of a Man₂GlcNAc₂ (Man-2) to Man-5 core substituted with various pentoses (Xylp, Arap, Araf), hexoses (Galp, Galf, Fucp, GlcNAcp) and *O*-methyl groups^{20,21,22,23,24,25,26,27,28}. Moreover, differences in the ER processing of LLO have been predicted in the different microalgae species²². It was also demonstrated that the green microalgae *Chlamydomonas reinhardtii* and the diatom *Phaeodactylum tricornutum* synthesize respectively truncated Glc₃Man₅GlcNAc₂ and Glc₂Man₉GlcNAc₂ LLO oligosaccharides instead of the Glc₃Man₉GlcNAc₂ found in mammals, insects, yeast and plants²⁹. With regards to oligomannosides, microalgae synthesize various isomers that rely on the structure of the LLO precursor from which they derived and the specificity of α -mannosidases involved in their trimming. As an example, the green microalga *C. reinhardtii* proteins carry a linear Man-5 that possess a Man α 1-2Man α 1-2Man trimannosyl sequence α (1,3)-linked to the β -Man of the core instead of the canonical Man-5 isomer observed in plants, insects and mammals^{24,30}. A mixture of Man-5 isomers have also been identified in *Chlorella vulgaris*²⁶.

Among microalgae species, the diatom *P. tricornutum* appears as one of the most promising expression systems investigated for the production of biopharmaceuticals such as monoclonal antibodies. Indeed, functional monoclonal antibodies directed against viruses have been efficiently produced in *P. tricornutum*^{31,32,33,34}. Consistently with the *N*-glycan profile identified on total *P. tricornutum* protein extracts²⁰, the diatom-made monoclonal antibodies carry *N*-linked oligomannosides ranging from Man-5 to Man-9³⁵. The presence of such *N*-glycans on

the algae-made monoclonal antibodies might alter their effector functions, thus leading to a possible decrease of its efficacy in human immunotherapy. Furthermore, the presence of oligomannosides on the algae-made antibodies might increase its *in vivo* clearance due to the binding to the macrophage mannose receptor in the liver³⁵. Therefore, as far as production of biopharmaceuticals exhibiting human-compatible *N*-glycans is concerned in this diatom, the engineering of its endogenous *N*-glycosylation pathway would require that *P. tricornutum* is able to process oligomannosides into a canonical Man-5 oligomannoside¹⁸. Indeed, this canonical Man-5 isomer is substrate for the GnT I, the first Golgi transferase that initiate the synthesis of the GlcNAc₂Man₃GlcNAc₂ core in mammals¹⁰. *P. tricornutum* oligomannosides may exist as multiple positional isomers depending on the glycosidic linkages between constitutive α -Man residues. As a consequence, analytical strategies are required for the detailed profiling of the oligomannoside populations either for the understanding of their ER and Golgi biosynthesis in *P. tricornutum* and for the evaluation of appropriate strategies for their glyco-engineering into human-like *N*-glycan structures. Discrimination between those oligomannoside positional isomers is highly challenging because they are closely related structures and are usually available only in very low amounts. In this context, mass spectrometry appears as a suitable sensitive methodology which allows the structural characterization of oligosaccharide mixtures using tandem mass spectrometry (MS/MS). For instance, multistage tandem mass spectrometry (ESI-MSⁿ) analyses of *N*-glycans from the ribonuclease B was demonstrated to be able to structurally differentiate the three isomers of Man-7 and Man-8³⁶. Since a few years, the coupling of ion mobility spectrometry to mass spectrometry (IMS-MS) also permitted the characterization of numerous biological compounds as well as isomers differentiation. Moreover, this technique allows the experimental determination of collision cross sections (CCS) of ions in a specific buffer gas. These CCS are specific chemical descriptors that can be used to increase the confidence in the structural identification of complex molecules³⁷. Herein, we described the structural analysis of the oligomannosides ranging from Man-5 to Man-9 synthesized by *P. tricornutum* by combination of Ion Mobility Spectrometry–Mass Spectrometry (IMS-MS) and Electrospray Ionization-multistage tandem mass spectrometry (ESI-MSⁿ).

Results and discussion

The diatom P. tricornutum synthesizes a canonical Man₅GlcNAc₂

N-glycans were released from *P. tricornutum* proteins using PNGase F and then labelled with 2-aminobenzamide (2AB). Then, the sample was divided in two parts, one being permethylated³⁸. First, the 2AB derivatives of the diatom Man-5 oligomannoside was analyzed by IMS-MS, an analytical method which allows the separation of oligosaccharide isomers^{24,30,39,40,41,42}. The ion mobility of the [M + Na]⁺ adduct of 2AB Man-5 was compared to the ion mobilities of 2AB derivatives of the canonical Man-5 isolated from the bovine ribonuclease B and of the linear 2AB Man-5 prepared from *C. reinhardtii* total endogenous proteins (Figure 1 A). The canonical Man-5 isolated from the bovine ribonuclease B and the linear Man-5 structures from *C. reinhardtii* are displayed in Figure 1C^{30,43}. As shown in Figure 1A, the canonical 2AB Man-5 from bovine ribonuclease B and the Man-5 derivative from *P. tricornutum* proteins exhibited a unique ion mobility with a drift time t_D of about 11.55 +/- 0.03 ms. This suggested that they share the same structure. In contrast, linear 2AB Man-5 from *C. reinhardtii* exhibited a t_D of 10.25 +/- 0.03 ms. Similar results were observed when the analysis by IMS-MS was performed on the [M + 2Na]²⁺ ion of permethylated 2AB Man-5 derivative from *P. tricornutum* as compared to the canonical Man-5 isolated from the bovine ribonuclease B and the linear Man-5 structures from *C. reinhardtii* (Figure 1 B). Indeed, as observed for non-permethylated 2AB Man-5 derivatives, permethylated 2AB Man-5 from bovine ribonuclease B and *P. tricornutum* proteins exhibited similar ion mobilities differing from the one recorded for the linear *C. reinhardtii* permethylated 2AB Man-5 (Figure 1B).

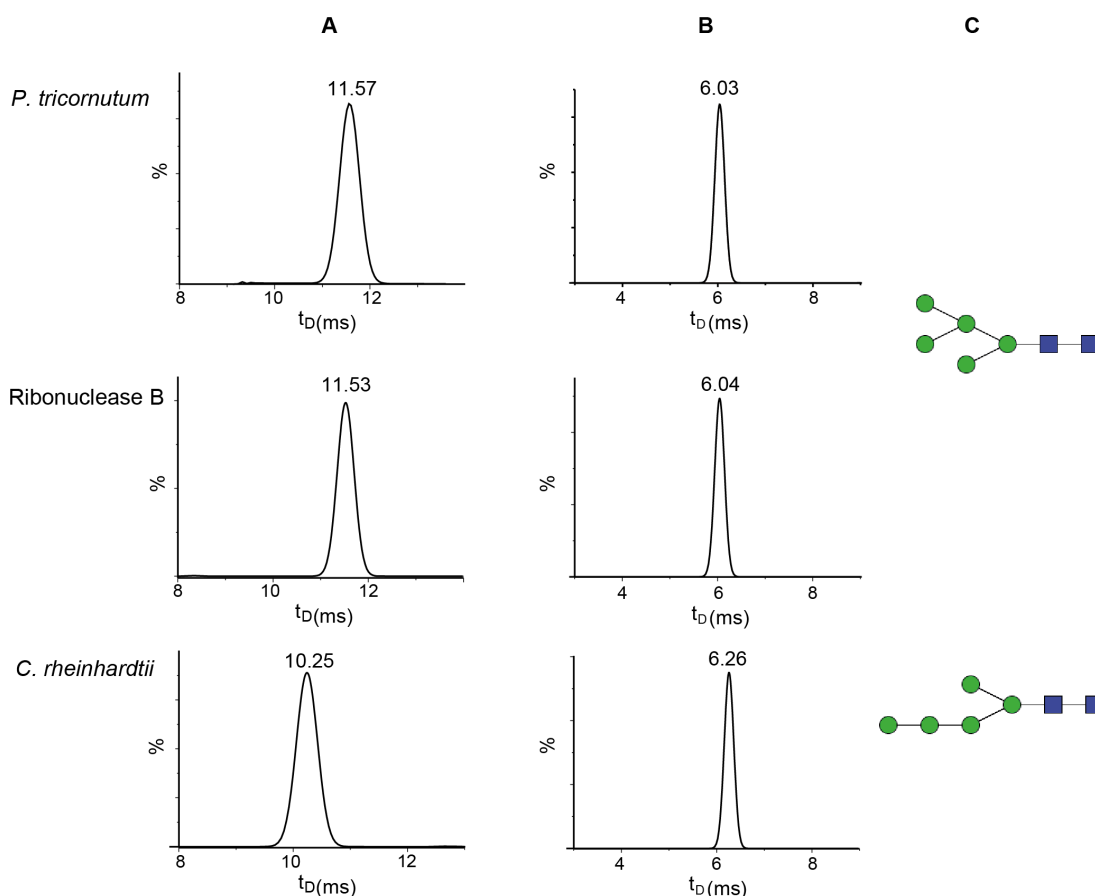


Figure 1. Drift time (t_D) recorded by Ion Mobility Spectrometry–Mass Spectrometry (IMS-MS) of non-permethylated (A) and permethylated (B) 2AB Man-5 derivatives isolated from *P. tricornutum*, bovine ribonuclease B and *C. reinhardtii* proteins. $[M+Na]^+$ ions were analyzed for non-permethylated 2AB Man-5 derivatives whereas t_D for permethylated 2AB Man-5 derivatives were recorded for $[M+2Na]^{2+}$ ions. (C) Structures of the canonical (middle) Man-5 N-linked to bovine ribonuclease B and the linear Man-5 as described in *C. reinhardtii* (bottom), respectively.

Experimental collision cross section (CCS, Ω) values of non-permethylated and permethylated 2AB derivatives were determined by IMS-MS using dextran oligomers as references⁴⁴. CCS are useful structural descriptors which provide information about the shape and size of ions in a buffer gas. In metabolomics or lipidomics studies, CCS are always used to increase the confidence in compound identification^{45,46,47}. Moreover, it has been reported that isomers of oligosaccharides can be identified based on their CCS^{37,48,49,50}. Figure 2 showed that permethylated 2AB Man-5 derivatives present higher CCS values than non-permethylated 2AB Man-5 due to the increase of the glycan size by methylation of hydroxyl groups. Data presented in Figure

2 also indicated that canonical permethylated 2AB-Man-5 are more compact than linear permethylated 2AB-Man-5 whereas the opposite behavior is observed for non-permethylated samples.

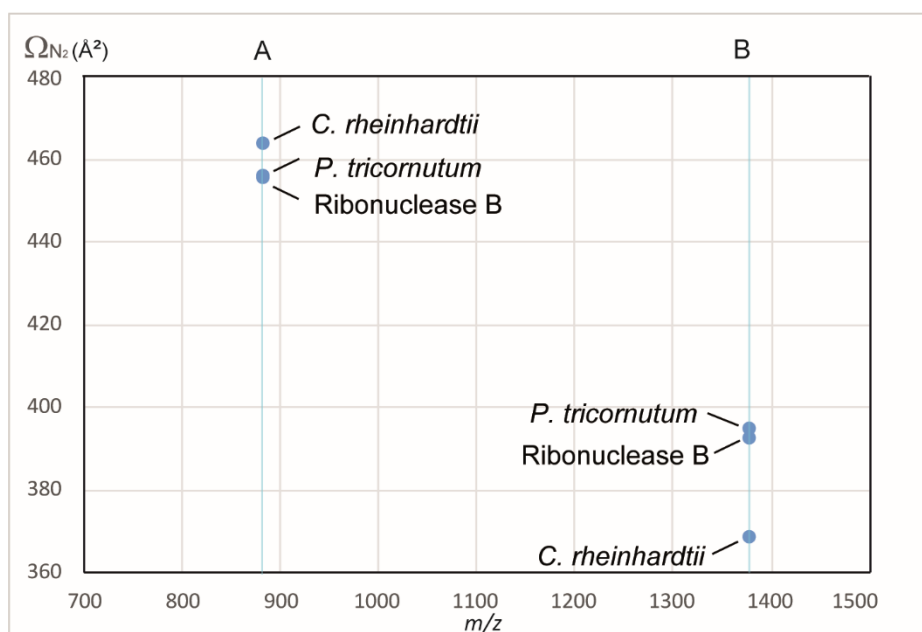


Figure 2. Collision cross sections (CCS, Ω) determined in nitrogen by ion mobility spectrometry-mass spectrometry (IMS-MS) of permethylated (A; $[M + 2Na]^{2+}$ adduct) and non-permethylated (B; $[M + Na]^+$ adduct) 2AB Man-5 isolated from bovine ribonuclease B, *P. tricornutum* and *C. reinhardtii* proteins.

Data recorded by IMS-MS suggested that *P. tricornutum* proteins carry a canonical Man-5 isomer. For confirmation, the structure of *P. tricornutum* Man-5 isomer was further investigated by ESI-MSⁿ with $n = 2$, $n = 3$ and $n = 4$. Fragmentation pattern of the doubly charged $[M + 2Na]^{2+}$ ion of permethylated 2AB-Man-5 (m/z 882) from *P. tricornutum* revealed a fragmentation pattern m/z 882 \rightarrow m/z 1302 \rightarrow m/z 866 \rightarrow m/z 648 that is specific for the canonical isomer of permethylated 2AB Man-5 (Figure 3). Similar pattern of fragmentation was observed for $[M + 2Na]^{2+}$ precursor ion of permethylated 2AB Man-5 obtained from bovine ribonuclease B³⁰ (Supplemental Figure 1B) and differed from the ESI-MSⁿ fragmentation pattern of the linear sequence permethylated 2AB Man-5 obtained from *C. reinhardtii* proteins³⁰. Taken together, IMS-MS data recorded for both non-permethylated and permethylated 2AB Man-5

Figure 3. Fragmentation pattern for m/z 882.6 precursor ion ($[M+2Na]^{2+}$) and multistage tandem mass spectra ($n= 2, 3$ and 4) of permethylated 2AB Man-5 derivative isolated from *P. tricornutum* proteins.

Structural analysis of the *P. tricornutum* oligomannoside population

The structural analysis of *P. tricornutum* oligomannosides was then extended to Man-6 to Man-9. Analysis of the whole diatom population was investigated by IMS-MS taking advantage of data recorded for Man-5 to Man-9 oligomannosides isolated from bovine ribonuclease B, which structures have been already reported^{36,43,51}. The experimental CCS determined by IMS-MS using nitrogen as buffer gas showed that values from *P. tricornutum* oligomannosides were very close to those from bovine ribonuclease B oligomannosides suggesting that both oligomannoside populations are similar (Table 1). Similar observation and conclusion can be drawn from experimental CCS recorded in helium as buffer gas (Supplemental Table 1).

Table 1: Experimental CCS ($^{TW}CCS_{N_2 \rightarrow N_2}$) determined for permethylated 2AB derivatives of Man-5 to Man-9 oligomannosides isolated from *P. tricornutum* proteins and from the bovine ribonuclease B⁵².

Permethylated 2AB derivatives	$[M+2Na]^{2+}$ (m/z)	Ω_{N_2} (Å ²)	
		Ribonuclease B	<i>P. tricornutum</i>
Man-5	882.5	456.0	455.6
Man-6	984.6	494.0	492.8
Man-7	1086.7	528.9	527.8
Man-8	1188.7	561.1	560.0
Man-9	1290.8	595.1	591.3

The presence of positional isomers for each oligomannoside was investigated by measuring the ion mobility peak width (full width at half maximum, FWHM) that has been recently reported as a new isomeric descriptor^{53,54}. In fact, the presence of isomers or conformers can be detected based on the peak broadening recorded in the ion mobility spectrum. FWHM for permethylated 2AB Man-5 to Man-9 ion mobility

peaks were extracted from each ion mobility peaks according to the method previously described for other classes of compounds^{53,54}. It has been reported that bovine ribonuclease B, used as a reference in this study, is *N*-glycosylated by unique isomers of Man-5 and Man-9 respectively, whereas different positional isomers of Man-6, Man-7 and Man-8 have been structurally characterized by proton NMR⁴³ and mass spectrometry coupled to Liquid Chromatography^{36,51}. The plots of FWHM ($\Delta t_{D1/2}$) values in function of drift time t_D for permethylated 2AB derivatives of bovine ribonuclease B oligomannosides present higher values of FWHM than expected (considering the sole diffusion phenomenon), especially for the Man-6, Man-7 and Man-8 (Figure 4A). Thus, in contrast to Man-5 and Man-9 that both exist as unique isomer (smaller $\Delta t_{D1/2}$), the higher $\Delta t_{D1/2}$ observed for Man-7 and Man-8 is likely due to the presence of three positional isomers for these oligomannosides^{36,43,51}. As shown in Figure 4, data recorded for bovine ribonuclease B and *P. tricornutum* oligomannosides are similar. It can thus be concluded that Man-5 and Man-9 oligomannoside derivatives synthesized in *P. tricornutum* exist as single isomers while this diatom synthesizes different positional isomers of Man-6, Man-7 and Man-8.

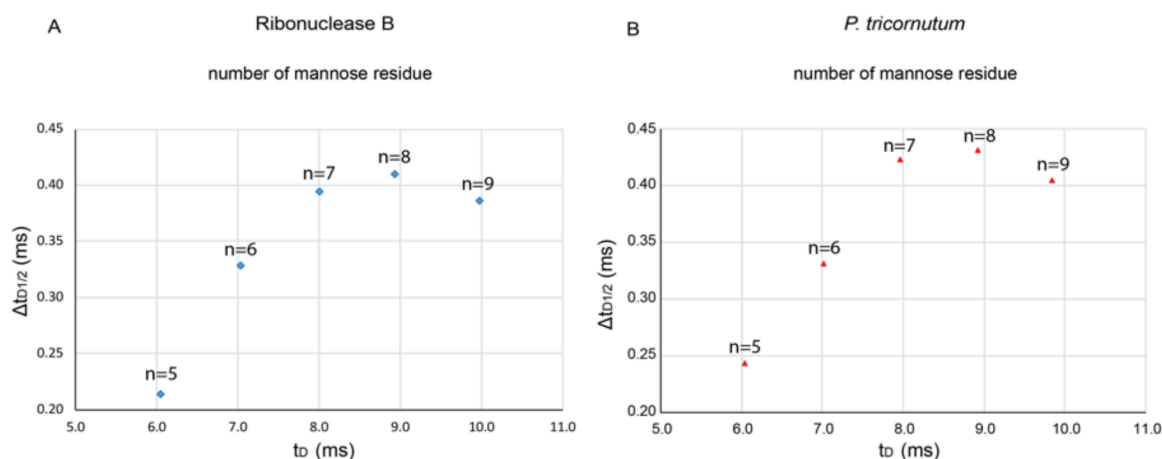


Figure 4. Full width at half maximum (FWHM, $\Delta t_{D1/2}$) according to the drift time (t_D) for permethylated 2AB Man-5 to Man-9 isolated from (A) bovine ribonuclease B and (B) *P. tricornutum*. n: number of mannose residues.

The observed weak peak broadening for Man-6 is likely due to the presence of minor isomers together with a major one (Figure 4B). The structure of the main Man-6 isomer was investigated by ESI-MSⁿ and compared to the Man-6 of the bovine ribonuclease B (Table 1). ESI-MSⁿ patterns of the doubly charged $[M + 2Na]^{2+}$ ion of

the permethylated 2AB-Man-6 (m/z 984.6) from *P. tricornutum* proteins and bovine ribonuclease B revealed similar fragmentation patterns m/z 984.6 \rightarrow m/z ,1506.9 \rightarrow m/z 880.5 (or \rightarrow m/z 866.5) corresponding to the fragmentation of the Man-6 isomer reported in Figure 5 (Supplemental Figure 2).

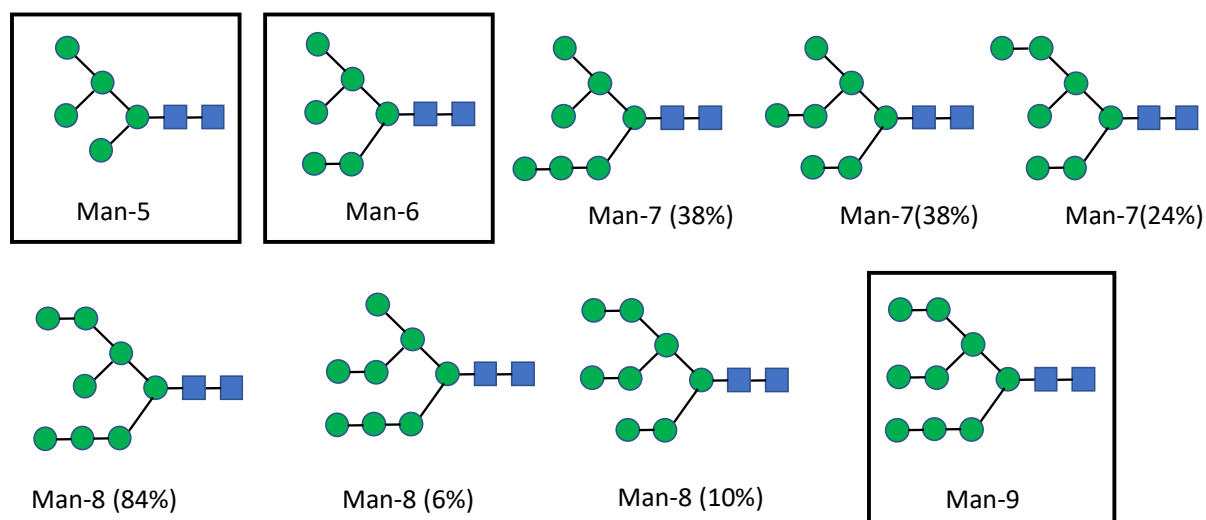


Figure 5. Structures of main oligomannosides N-linked to bovine ribonuclease B according to ⁴³. Same structures were determined for Man-5, Man-6 and Man-9 (in black boxes) isolated from *P. tricornutum* proteins.

Conclusion

Structures of oligomannosides from *P. tricornutum* proteins were investigated by IMS-MS and ESI-MSⁿ using glycans N-linked to bovine ribonuclease B as references. Ion mobility spectrometry - mass spectrometry has previously been reported as a very efficient methodology for the analysis of glycan structural isomers allowing the separation and identification of closely related structures^{39,40,41,42,48}. The structural analysis of Man-5 from *P. tricornutum* proteins clearly demonstrated that this diatom synthesizes a unique canonical Man₅GlcNAc₂ isomer as observed in mammals (Figure 5). This result is consistent with the expression in *P. tricornutum* of a GnT I that is supposed to transfer a terminal GlcNAc residue specifically on this canonical Man-5 oligomannoside. This suggests that this diatom is able to initiate the processing of oligomannosides into complex type N-glycans as in other eukaryotes. Structures for Man-6 to Man-9 oligomannosides N-linked to *P. tricornutum* proteins could also be

proposed based on the determined CCS values and FWHM measurement of the signals recorded in IMS-MS. This study demonstrated for the first time that oligomannosides Man-5, as well as Man-9, *N*-linked to *P. tricornutum* proteins are single isomers. On the other hand, Man-6, Man-7 and Man-8 appear to present different isomers as observed for the oligomannosides isolated from bovine ribonuclease B. Electrospray Ionization-multistage tandem mass spectrometry of Man-6 major isomer from *P. tricornutum* suggested a *N*-glycan structure that is identical to the Man-6 structure reported for bovine ribonuclease B.

Experimental section

Materials

Bovine ribonuclease B was from Sigma (R7884). *Flavobacterium meningosepticum* PNGase F produced in *E. coli* was from Roche (NGLY-R0). All buffer, solvent and reagents were LC-MS grade.

N-glycans preparation.

Total proteins from *P. tricornutum* were extracted as previously described elsewhere⁵⁵. After extraction, the protein extract was spin down during 30 min at 15,000g. The supernatant was dialyzed 24h against water using a molecular cut-off of 6 to 8,000 Da. After dialysis, above 10 mg of proteins from *P. tricornutum* were freeze dried prior to deglycosylation using the Peptide *N*-glycosidase F (PNGase F) as previously reported²⁰. One mg of bovine ribonuclease B was digested and then treated in parallel to the *P. tricornutum* samples. Deglycosylated proteins were then precipitated with 4V of Ethanol at -20°C during 24h. Samples were then spin down at 20,000 g during 15 minutes. The supernatant containing the ethanol soluble compounds such as the released *N*-glycans were dried. The *N*-glycan fractions from *C. reinhardtii* were prepared as previously reported in the literature^{24,30}.

N-glycans derivatization.

The free *N*-glycans were then labelled with 2-aminobenzamide (2AB) according to the literature⁵⁶. Excess of reagent was removed using a cartridge D1 from Ludger according to the manufacturer's instructions. Permethylation of samples was performed as described in the literature³⁸. Permethylated 2AB-labelled *N*-glycans were finally clean-up as reported in ⁵⁷ prior to IMS-MS and ESI-MSⁿ analyses.

IMS-MS

The ESI-IM-MS experiments were performed using a Waters SYNAPT G2 hybrid quad-rupole/HDMS instrument equipped with an ESI LockSpray™ source, the MassLynx 4.1 and the Drift Scope 2.2 softwares (Waters, Manchester, UK). The SYNAPT HDMS system was calibrated using sodium formate cluster ions (2 mg.mL⁻¹) and operated in 'V' resolution mode (resolution 20,000 FWHM). The ESI parameters were in positive ion mode: capillary voltage, 3.1 kV; sample cone voltage, 70 V; source temperature, 90 °C; desolvation temperature, 300 °C; desolvation gas flow (N₂), 700 L.h⁻¹. The data were acquired using a 50–2000 *m/z* range with 1 s scan time and 0.02 s interscan delay. Sample solutions were infused into the source at a flow rate of 400 μL.h⁻¹ with a syringe pump (Cole-Palmer, Vernon Hills, Illinois, USA). The IMS conditions were: gas flow (N₂), 90 mL.min⁻¹; IMS cell pressure, 2.98 mbar; IMS traveling wave height voltage, 40 V and T-wave velocity, 550 m.s⁻¹ in the case of 2AB glycans and 700 ms⁻¹ for permethylated 2AB derivatives. Helium gas flow was 180 mL.min⁻¹. IMS wave delay was 450 μs. Data were recorded from three independent preparations of *N*-glycans.

As previously mentioned, IM-MS not only afforded two dimensional ion separation^{44,48}, but provided information about the three-dimensional structures of ions through their collision cross-section (CCS, Ω)^{46,47,48,60}. Indeed, gas phase separation of ions occurs when ions migrate under the influence of an electric field in an ion mobility (IM) device containing a buffer gas (He, N₂, ...). Depending on their charge, shape and size (related to Ω) the ion migrate at different velocities and pass through the IM cell at different drift times (*t_d*). The ion velocity (*v*) is related to the ion mobility (*K*) by equation (1):

$$K = \frac{v}{E} = \frac{L}{t_d E} \quad (1)$$

where *E* is the electric field and *L* the length of the IM cell.

The ion mobility *K* is also related to Ω in a definite buffer gas according to equation (2)⁶⁰:

$$K = \frac{3ze}{16N} \times \sqrt{\left(\frac{1}{m} + \frac{1}{M}\right)} \times \sqrt{\frac{2\pi}{k_b T}} \times \frac{T}{273.15} \times \frac{760}{P} \times \frac{1}{\Omega} \quad (2)$$

where ze is the ion charge, m and M are respectively the mass of the ion and of the buffer gas, k_b is the Boltzmann constant, N , P and T are the number density, the pressure and the temperature of the buffer gas.

Thus, in uniform electric field (drift tube ion mobility, DTIMS), the CCS can be determined directly from the drift time t_d , according to the Mason-Schamp equation (3)⁶¹:

$$\Omega = \frac{3ze}{16N} \times \sqrt{\left(\frac{1}{m} + \frac{1}{M}\right)} \times \sqrt{\frac{2\pi}{k_b T}} \times \frac{T}{273.15} \times \frac{760}{P} \times \frac{t_d E}{L} \quad (3)$$

In TWIMS cells, the electric field is not constant and uniform. The equation (3) is not applicable and the relation between t_d and Ω is then expressed by the equation (4)⁵⁸:

$$\Omega = \frac{3ze}{16N} \times \sqrt{\left(\frac{1}{m} + \frac{1}{M}\right)} \times \sqrt{\frac{2\pi}{k_b T}} \times \frac{760}{P} \times \frac{T}{273.15} \times A t_d^B \quad (4)$$

where A and B are the correction factors related to the electric field E and to the compensation of the non-linearity of E , respectively.

This relation can be simplified to equation (5):

$$\Omega' = A' t_d^B \quad (5)$$

where

$$A' = \frac{3}{16N} \times \sqrt{\frac{2\pi}{kT}} \times \frac{760}{P} \times \frac{T}{273.15} \times A$$

and

$$\Omega' = \frac{\Omega}{ze \sqrt{\left(\frac{1}{m} + \frac{1}{M}\right)}} = \frac{\Omega}{ze \sqrt{1/\mu}}$$

To obtain CCS of an analyte ion, the coefficients A' and B need to be determined *via* the calibration of the TWIMS cell using reference compounds of known CCS, previously determined using DTIMS devices. These reference ions must exhibit the same charge state as analyte ion. Various calibrating substances are available which CCS have been reported either in helium or in nitrogen^{48,62}. Note that experimental collision cross section determined with a TWIMS cell can be labelled as ^{TW}CCS, and more precisely as ^{TW}CCS_{N₂→He} when estimated in helium and as ^{TW}CCS_{N₂→N₂} when estimated in nitrogen as recommended⁵⁹.

Correlating measured drift time (t_d) of calibrant ions with corresponding reduced collision cross section Ω' according to Smith *et al.* method permit to plot a calibration curve $\text{Ln } \Omega' = \text{Ln } A' + B \text{ Ln } t_d$.

In our case, two sets of TWIMS cell calibration were performed, the first to determine ^{TW}CCS of non-permethylated 2AB labelled glycans for which $[M+Na]^+$ adducts of

dextran were used, the second to determine ^{TW}CCS of permethylated 2AB derivatives for which $[M+2Na]^{2+}$ ions of dextran were used.

The drift times of both sample and reference ions were accurately determined by fitting the extracted ion mobility spectra with a Gaussian regression (OriginPro 2016 b9.3.226 software, OriginLab).

ESI-MSⁿ analysis

Permethylated *N*-glycans were analyzed by ESI-MSⁿ ($n = 2$ to 3) using a Bruker HCT Ultra ETD II quadrupole ion trap (QIT) mass spectrometer equipped with the Esquire control 6.2 and Data Analysis 4.0 softwares (Bruker Daltonics, Bremen, Germany). The ESI parameters, in positive ion mode, were as followed: capillary voltage set at -3.5 kV, end plate offset at -500 V, skimmer and capillary exit voltages set at 40 V and 200 V, respectively, nebulizer gas (N_2), pressure, drying gas (N_2) flow rate and drying gas temperature were 10 psi, 7.0 L min^{-1} and 300 °C, respectively. Helium pressure in the ion trap was 1.3×10^{-5} mbar. The data were acquired using a 200 – 2200 m/z range, using a scan speed of 8100 m/z per second. The number of ions entering the trap cell was automatically adjusted by controlling the accumulation time using the ion charge control (ICC) mode (target $100,000$) with a maximum accumulation time of 50 ms. The injection low-mass cut-off (LMCO) value was m/z 120 . The values of spectra averages and rolling average were 6 and 2 . ESI-MSⁿ experiments were carried out by collision-induced dissociation (CID) using helium as the collision gas, isolation width of 1 m/z unit for the precursor ions and 2 m/z unit for the intermediate ions using a resonant excitation frequency with an amplitude from 0.8 to 0.9 Vp–p. Sample were resuspended in 100 μ L of CH_3OH/H_2O , $1/1$, v/v and then diluted twice or 100 -fold in the same solvent system for *P. tricornutum* and Ribonuclease B, respectively. The final solutions were infused into the source at a flow rate of 180 μ L. h^{-1} by means of a syringe pump (Cole-Palmer, Vernon Hills, IL, USA). External calibration was performed using the ‘tuning mix’ from Agilent Technologies (Santa Rosa, CA, USA). All mass spectra were manually interpreted, based on literature^{48,36,62}. The Glycoworkbench software v2.1 was used to support and confirm the assignment of the fragment ions.

AUTHOR INFORMATION

Corresponding Author

*Correspondence should be addressed to Patrice Lerouge (patrice.lerouge@univ-rouen.fr) and Muriel Bardor (muriel.bardor@univ-rouen.fr)

Author Contributions

R.D., C.L.-B., M.-L.W.-B., C. B. performed the experimental work; R.D., M.-L.W.-B., C.L.-B., M.B., P.L. analysed the data; The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

‡These authors contributed equally.

ACKNOWLEDGMENT

The authors are grateful for generous and financial support from the IUF (Institut Universitaire de France, 2014-2019); the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement 774078 (Pharma-Factory); the European Regional Development Fund (ERDF, N°31708), the Region Normandie (Crunch Network, N°20-13; SFR NORVEGE), the Labex SynOrg (ANR-11-LABX-0029) and the University of Rouen Normandie.

REFERENCES

- (1) Khoury, G. A.; Baliban, R. C.; Floudas, C. A. Proteome-Wide Post-Translational Modification Statistics: Frequency Analysis and Curation of the Swiss-Prot Database. *Sci Rep* **2011**, *1* (1), 1–5. <https://doi.org/10.1038/srep00090>.
- (2) Burda, P.; Aebi, M. The Dolichol Pathway of N-Linked Glycosylation. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1999**, *1426* (2), 239–257. [https://doi.org/10.1016/S0304-4165\(98\)00127-5](https://doi.org/10.1016/S0304-4165(98)00127-5).
- (3) Gil, G.-C.; Velander, W. H.; Cott, K. E. V. N-Glycosylation Microheterogeneity and Site Occupancy of an Asn-X-Cys Sequon in Plasma-Derived and Recombinant Protein C. *PROTEOMICS* **2009**, *9* (9), 2555–2567. <https://doi.org/10.1002/pmic.200800775>.
- (4) Zielinska, D. F.; Gnad, F.; Wiśniewski, J. R.; Mann, M. Precision Mapping of an In Vivo N-Glycoproteome Reveals Rigid Topological and Sequence Constraints. *Cell* **2010**, *141* (5), 897–907. <https://doi.org/10.1016/j.cell.2010.04.012>.
- (5) Matsui, T.; Takita, E.; Sato, T.; Kinjo, S.; Aizawa, M.; Sugiura, Y.; Hamabata, T.; Sawada, K.; Kato, K. N-Glycosylation at Noncanonical Asn-X-Cys Sequences in Plant Cells. *Glycobiology* **2011**, *21* (8), 994–999. <https://doi.org/10.1093/glycob/cwq198>.
- (6) Mohorko, E.; Glockshuber, R.; Aebi, M. Oligosaccharyltransferase: The Central Enzyme of N-Linked Protein Glycosylation. *J Inherit Metab Dis* **2011**, *34* (4), 869–878. <https://doi.org/10.1007/s10545-011-9337-1>.

- (7) Määttänen, P.; Gehring, K.; Bergeron, J. J. M.; Thomas, D. Y. Protein Quality Control in the ER: The Recognition of Misfolded Proteins. *Seminars in Cell & Developmental Biology* **2010**, *21* (5), 500–511. <https://doi.org/10.1016/j.semcdb.2010.03.006>.
- (8) Lerouge, P.; Cabanes-Macheteau, M.; Rayon, C.; Fischette-Lainé, A.-C.; Gomord, V.; Faye, L. N-Glycoprotein Biosynthesis in Plants: Recent Developments and Future Trends. *Plant Mol Biol* **1998**, *38* (1), 31–48. <https://doi.org/10.1023/A:1006012005654>.
- (9) Shi, X.; Jarvis, D. L. Protein N-Glycosylation in the Baculovirus-Insect Cell System. *Curr Drug Targets* **2007**, *8* (10), 1116–1125.
- (10) Stanley, P.; Taniguchi, N.; Aebi, M. N-Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2015.
- (11) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49. <https://doi.org/10.1093/glycob/cww086>.
- (12) Barrera, D. J.; Mayfield, S. P. High-Value Recombinant Protein Production in Microalgae. In *Handbook of Microalgal Culture*; John Wiley & Sons, Ltd, 2013; pp 532–544. <https://doi.org/10.1002/9781118567166.ch27>.
- (13) Rosales-Mendoza, S. Perspectives for the Algae-Made Biopharmaceuticals Field. In *Algae-Based Biopharmaceuticals*; Rosales-Mendoza, S., Ed.; Springer International Publishing: Cham, 2016; pp 143–163. https://doi.org/10.1007/978-3-319-32232-2_9.
- (14) Hempel, F.; Maier, U. G. Microalgae as Solar-Powered Protein Factories. In *Advanced Technologies for Protein Complex Production and Characterization*; Vega, M. C., Ed.; Advances in Experimental Medicine and Biology; Springer International Publishing: Cham, 2016; pp 241–262. https://doi.org/10.1007/978-3-319-27216-0_16.
- (15) Mathieu-Rivet, E.; Lerouge, P.; Bardor, M. Chlamydomonas Reinhardtii: Protein Glycosylation and Production of Biopharmaceuticals. In *Chlamydomonas: Biotechnology and Biomedicine*; Hippler, M., Ed.; Microbiology Monographs; Springer International Publishing: Cham, 2017; pp 45–72. https://doi.org/10.1007/978-3-319-66360-9_3.
- (16) Rosales-Mendoza, S.; Solís-Andrade, K. I.; Márquez-Escobar, V. A.; González-Ortega, O.; Bañuelos-Hernandez, B. Current Advances in the Algae-Made Biopharmaceuticals Field. *Expert Opinion on Biological Therapy* **2020**, *0* (0), 1–16. <https://doi.org/10.1080/14712598.2020.1739643>.
- (17) Mathieu-Rivet, E.; Kiefer-Meyer, M.-C.; Vanier, G.; Ovide, C.; Burel, C.; Lerouge, P.; Bardor, M. Protein N-Glycosylation in Eukaryotic Microalgae and Its Impact on the Production of Nuclear Expressed Biopharmaceuticals. *Front. Plant Sci.* **2014**, *5*. <https://doi.org/10.3389/fpls.2014.00359>.
- (18) Dumontier, R.; Mareck, A.; Mati-Baouche, N.; Lerouge, P.; Bardor, M. Toward Future Engineering of the N-Glycosylation Pathways in Microalgae for Optimizing the Production of Biopharmaceuticals. *Microalgal Biotechnology* **2018**. <https://doi.org/10.5772/intechopen.73401>.
- (19) Barolo, L.; Abbriano, R. M.; Commault, A. S.; George, J.; Kahlke, T.; Fabris, M.; Padula, M. P.; Lopez, A.; Ralph, P. J.; Pernice, M. Perspectives for Glyco-Engineering of Recombinant Biopharmaceuticals from Microalgae. *Cells* **2020**, *9* (3), 633. <https://doi.org/10.3390/cells9030633>.

- (20) Baïet, B.; Burel, C.; Saint-Jean, B.; Louvet, R.; Menu-Bouaouiche, L.; Kiefer-Meyer, M.-C.; Mathieu-Rivet, E.; Lefebvre, T.; Castel, H.; Carlier, A.; Cadoret, J.-P.; Lerouge, P.; Bardor, M. N-Glycans of *Phaeodactylum Tricornutum* Diatom and Functional Characterization of Its N-Acetylglucosaminyltransferase I Enzyme. *J. Biol. Chem.* **2011**, *286* (8), 6152–6164. <https://doi.org/10.1074/jbc.M110.175711>.
- (21) Mathieu-Rivet, E.; Scholz, M.; Arias, C.; Dardelle, F.; Schulze, S.; Mauff, F. L.; Teo, G.; Hochmal, A. K.; Blanco-Rivero, A.; Loutelier-Bourhis, C.; Kiefer-Meyer, M.-C.; Fufezan, C.; Burel, C.; Lerouge, P.; Martinez, F.; Bardor, M.; Hippler, M. Exploring the N-Glycosylation Pathway in *Chlamydomonas Reinhardtii* Unravels Novel Complex Structures. *Molecular & Cellular Proteomics* **2013**, *12* (11), 3160–3183. <https://doi.org/10.1074/mcp.M113.028191>.
- (22) Levy-Ontman, O.; Fisher, M.; Shotland, Y.; Weinstein, Y.; Tekoah, Y.; Arad, S. M. Genes Involved in the Endoplasmic Reticulum N-Glycosylation Pathway of the Red Microalga *Porphyridium Sp.*: A Bioinformatic Study. *International Journal of Molecular Sciences* **2014**, *15* (2), 2305–2326. <https://doi.org/10.3390/ijms15022305>.
- (23) Schulze, S.; Oltmanns, A.; Machnik, N.; Liu, G.; Xu, N.; Jarmatz, N.; Scholz, M.; Sugimoto, K.; Fufezan, C.; Huang, K.; Hippler, M. N-Glycoproteomic Characterization of Mannosidase and Xylosyltransferase Mutant Strains of *Chlamydomonas Reinhardtii*. *Plant Physiology* **2018**, *176* (3), 1952–1964. <https://doi.org/10.1104/pp.17.01450>.
- (24) Lucas, P.-L.; Mathieu-Rivet, E.; Song, P. C. T.; Oltmanns, A.; Loutelier-Bourhis, C.; Plasson, C.; Afonso, C.; Hippler, M.; Lerouge, P.; Mati-Baouche, N.; Bardor, M. Multiple Xylosyltransferases Heterogeneously Xylosylate Protein N-Linked Glycans in *Chlamydomonas Reinhardtii*. *The Plant Journal* **2020**, *102* (2), 230–245. <https://doi.org/10.1111/tpj.14620>.
- (25) Oltmanns, A.; Hoepfner, L.; Scholz, M.; Zinzius, K.; Schulze, S.; Hippler, M. Novel Insights Into N-Glycan Fucosylation and Core Xylosylation in *C. Reinhardtii*. *Front Plant Sci* **2019**, *10*, 1686. <https://doi.org/10.3389/fpls.2019.01686>.
- (26) Mócsai, R.; Figl, R.; Troschl, C.; Strasser, R.; Svehla, E.; Windwarder, M.; Thader, A.; Altmann, F. N-Glycans of the Microalga *Chlorella Vulgaris* Are of the Oligomannosidic Type but Highly Methylated. *Sci Rep* **2019**, *9* (1), 1–8. <https://doi.org/10.1038/s41598-018-36884-1>.
- (27) Mócsai, R.; Figl, R.; Sützl, L.; Fluch, S.; Altmann, F. A First View on the Unsuspected Intra-genus Diversity of N-Glycans in *Chlorella* Microalgae. *The Plant Journal* *n/a* (n/a). <https://doi.org/10.1111/tpj.14718>.
- (28) Mócsai, R.; Blaukopf, M.; Svehla, E.; Kosma, P.; Altmann, F. The N-Glycans of *Chlorella Sorokiniana* and a Related Strain Contain Arabinose but Have Strikingly Different Structures. *Glycobiology*. <https://doi.org/10.1093/glycob/cwaa012>.
- (29) Lucas, P.-L.; Dumontier, R.; Loutelier-Bourhis, C.; Mareck, A.; Afonso, C.; Lerouge, P.; Mati-Baouche, N.; Bardor, M. User-Friendly Extraction and Multistage Tandem Mass Spectrometry Based Analysis of Lipid-Linked Oligosaccharides in Microalgae. *Plant Methods* **2018**, *14* (1), 107. <https://doi.org/10.1186/s13007-018-0374-8>.
- (30) Vanier, G.; Lucas, P.-L.; Loutelier-Bourhis, C.; Vanier, J.; Plasson, C.; Walet-Balieu, M.-L.; Tchi-Song, P. C.; Remy-Jouet, I.; Richard, V.; Bernard, S.; Driouich, A.; Afonso, C.; Lerouge, P.; Mathieu-Rivet, E.; Bardor, M. Heterologous

- Expression of the N -Acetylglucosaminyltransferase I Dictates a Reinvestigation of the N -Glycosylation Pathway in *Chlamydomonas Reinhardtii*. *Sci Rep* **2017**, *7* (1), 1–12. <https://doi.org/10.1038/s41598-017-10698-z>.
- (31) Hempel, F.; Lau, J.; Klingl, A.; Maier, U. G. Algae as Protein Factories: Expression of a Human Antibody and the Respective Antigen in the Diatom *Phaeodactylum Tricornutum*. *PLoS One* **2011**, *6* (12). <https://doi.org/10.1371/journal.pone.0028424>.
- (32) Hempel, F.; Maier, U. G. An Engineered Diatom Acting like a Plasma Cell Secreting Human IgG Antibodies with High Efficiency. *Microbial Cell Factories* **2012**, *11* (1), 126. <https://doi.org/10.1186/1475-2859-11-126>.
- (33) Hempel, F.; Maurer, M.; Brockmann, B.; Mayer, C.; Biedenkopf, N.; Kelterbaum, A.; Becker, S.; Maier, U. G. From Hybridomas to a Robust Microalgal-Based Production Platform: Molecular Design of a Diatom Secreting Monoclonal Antibodies Directed against the Marburg Virus Nucleoprotein. *Microb Cell Fact* **2017**, *16* (1), 131. <https://doi.org/10.1186/s12934-017-0745-2>.
- (34) Vanier, G.; Stelter, S.; Vanier, J.; Hempel, F.; Maier, U. G.; Lerouge, P.; Ma, J.; Bardor, M. Alga-Made Anti-Hepatitis B Antibody Binds to Human Fcγ Receptors. *Biotechnology Journal* **2018**, *13* (4), 1700496. <https://doi.org/10.1002/biot.201700496>.
- (35) Vanier, G.; Hempel, F.; Chan, P.; Rodamer, M.; Vaudry, D.; Maier, U. G.; Lerouge, P.; Bardor, M. Biochemical Characterization of Human Anti-Hepatitis B Monoclonal Antibody Produced in the Microalgae *Phaeodactylum Tricornutum*. *PLoS One* **2015**, *10* (10). <https://doi.org/10.1371/journal.pone.0139282>.
- (36) Prien, J. M.; Ashline, D. J.; Lapadula, A. J.; Zhang, H.; Reinhold, V. N. The High Mannose Glycans from Bovine Ribonuclease B Isomer Characterization by Ion Trap MS. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (4), 539–556. <https://doi.org/10.1021/jasms.8b03421>.
- (37) Struwe, W. B.; Pagel, K.; Benesch, J. L. P.; Harvey, D. J.; Campbell, M. P. GlycoMob: An Ion Mobility-Mass Spectrometry Collision Cross Section Database for Glycomics. *Glycoconj J* **2016**, *33* (3), 399–404. <https://doi.org/10.1007/s10719-015-9613-7>.
- (38) Ciucanu, I.; Kerek, F. A Simple and Rapid Method for the Permethylolation of Carbohydrates. *Carbohydrate Research* **1984**, *131* (2), 209–217. [https://doi.org/10.1016/0008-6215\(84\)85242-8](https://doi.org/10.1016/0008-6215(84)85242-8).
- (39) Zhu, M.; Bendiak, B.; Clowers, B.; Hill, H. H. Ion Mobility-Mass Spectrometry Analysis of Isomeric Carbohydrate Precursor Ions. *Anal Bioanal Chem* **2009**, *394* (7), 1853–1867. <https://doi.org/10.1007/s00216-009-2865-y>.
- (40) Plasencia, M. D.; Isailovic, D.; Merenbloom, S. I.; Mechref, Y.; Clemmer, D. E. Resolving and Assigning N-Linked Glycan Structural Isomers from Ovalbumin by IMS-MS. *J Am Soc Mass Spectrom* **2008**, *19* (11), 1706–1715. <https://doi.org/10.1016/j.jasms.2008.07.020>.
- (41) Zhu, F.; Lee, S.; Valentine, S. J.; Reilly, J. P.; Clemmer, D. E. Mannose7 Glycan Isomer Characterization by IMS-MS/MS Analysis. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (12), 2158–2166. <https://doi.org/10.1021/jasms.8b04202>.
- (42) Harvey, D. J.; Scarff, C. A.; Edgeworth, M.; Struwe, W. B.; Pagel, K.; Thalassinou, K.; Crispin, M.; Scrivens, J. Travelling-Wave Ion Mobility and Negative Ion Fragmentation of High-Mannose N-Glycans. *J Mass Spectrom* **2016**, *51* (3), 219–235. <https://doi.org/10.1002/jms.3738>.
- (43) Fu, D.; Chen, L.; O'Neill, R. A. A Detailed Structural Characterization of Ribonuclease B Oligosaccharides by ¹H NMR Spectroscopy and Mass

- Spectrometry. *Carbohydr. Res.* **1994**, *261* (2), 173–186. [https://doi.org/10.1016/0008-6215\(94\)84015-6](https://doi.org/10.1016/0008-6215(94)84015-6).
- (44) Hoffmann, W.; Hofmann, J.; Pagel, K. Energy-Resolved Ion Mobility-Mass Spectrometry—A Concept to Improve the Separation of Isomeric Carbohydrates. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (3), 471–479. <https://doi.org/10.1021/jasms.8b04709>.
- (45) Levy, A. J.; Oranzi, N. R.; Ahmadireskety, A.; Kemperman, R. H. J.; Wei, M. S.; Yost, R. A. Recent Progress in Metabolomics Using Ion Mobility-Mass Spectrometry. *TrAC Trends in Analytical Chemistry* **2019**, *116*, 274–281. <https://doi.org/10.1016/j.trac.2019.05.001>.
- (46) Paglia, G.; Angel, P.; Williams, J. P.; Richardson, K.; Olivos, H. J.; Thompson, J. W.; Menikarachchi, L.; Lai, S.; Walsh, C.; Moseley, A.; Plumb, R. S.; Grant, D. F.; Palsson, B. O.; Langridge, J.; Geromanos, S.; Astarita, G. Ion Mobility-Derived Collision Cross Section As an Additional Measure for Lipid Fingerprinting and Identification. *Anal. Chem.* **2015**, *87* (2), 1137–1144. <https://doi.org/10.1021/ac503715v>.
- (47) Paglia, G.; Astarita, G. Metabolomics and Lipidomics Using Traveling-Wave Ion Mobility Mass Spectrometry. *Nat Protoc* **2017**, *12* (4), 797–813. <https://doi.org/10.1038/nprot.2017.013>.
- (48) Pagel, K.; Harvey, D. J. Ion Mobility–Mass Spectrometry of Complex Carbohydrates: Collision Cross Sections of Sodiated N-Linked Glycans. *Anal. Chem.* **2013**, *85* (10), 5138–5145. <https://doi.org/10.1021/ac400403d>.
- (49) Barroso, A.; Giménez, E.; Konijnenberg, A.; Sancho, J.; Sanz-Nebot, V.; Sobott, F. Evaluation of Ion Mobility for the Separation of Glycoconjugate Isomers Due to Different Types of Sialic Acid Linkage, at the Intact Glycoprotein, Glycopeptide and Glycan Level. *Journal of Proteomics* **2018**, *173*, 22–31. <https://doi.org/10.1016/j.jprot.2017.11.020>.
- (50) Manz, C.; Pagel, K. Glycan Analysis by Ion Mobility-Mass Spectrometry and Gas-Phase Spectroscopy. *Current Opinion in Chemical Biology* **2018**, *42*, 16–24. <https://doi.org/10.1016/j.cbpa.2017.10.021>.
- (51) Costello, C. E.; Contado-Miller, J. M.; Cipollo, J. F. A Glycomics Platform for the Analysis of Permethylated Oligosaccharide Alditols. *J Am Soc Mass Spectrom* **2007**, *18* (10), 1799–1812. <https://doi.org/10.1016/j.jasms.2007.07.016>.
- (52) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; Pauw, E. D.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T. Recommendations for Reporting Ion Mobility Mass Spectrometry Measurements. *Mass Spectrometry Reviews* **2019**, *38* (3), 291–320. <https://doi.org/10.1002/mas.21585>.
- (53) Farenc, M.; Paupy, B.; Marceau, S.; Riches, E.; Afonso, C.; Giusti, P. Effective Ion Mobility Peak Width as a New Isomeric Descriptor for the Untargeted Analysis of Complex Mixtures Using Ion Mobility-Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (11), 2476–2482. <https://doi.org/10.1021/jasms.8b05409>.
- (54) Mendes Siqueira, A. L.; Beaumesnil, M.; Hubert-Roux, M.; Loutelier-Bourhis, C.; Afonso, C.; Bai, Y.; Courtiade, M.; Racaud, A. Atmospheric Solid Analysis Probe Coupled to Ion Mobility Spectrometry-Mass Spectrometry, a Fast and Simple

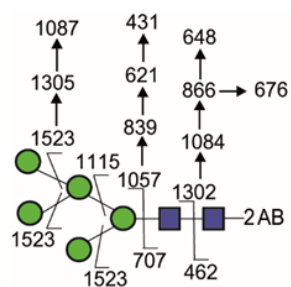
- Method for Polyalphaolefin Characterization. *J. Am. Soc. Mass Spectrom.* **2018**, 29 (8), 1678–1687. <https://doi.org/10.1021/jasms.8b05887>.
- (55) Zhang, P.; Burel, C.; Plasson, C.; Kiefer-Meyer, M.-C.; Ovide, C.; Gügi, B.; Wan, C.; Teo, G.; Mak, A.; Song, Z.; Driouich, A.; Lerouge, P.; Bardor, M. Characterization of a GDP-Fucose Transporter and a Fucosyltransferase Involved in the Fucosylation of Glycoproteins in the Diatom *Phaeodactylum Tricornutum*. *Front. Plant Sci.* **2019**, 10. <https://doi.org/10.3389/fpls.2019.00610>.
- (56) Séveno, M.; Cabrera, G.; Triguero, A.; Burel, C.; Leprince, J.; Rihouey, C.; Vézina, L.-P.; D'Aoust, M.-A.; Rudd, P. M.; Royle, L.; Dwek, R. A.; Harvey, D. J.; Lerouge, P.; Cremata, J. A.; Bardor, M. Plant N-Glycan Profiling of Minute Amounts of Material. *Analytical Biochemistry* **2008**, 379 (1), 66–72. <https://doi.org/10.1016/j.ab.2008.04.034>.
- (57) North, S. J.; Huang, H.-H.; Sundaram, S.; Jang-Lee, J.; Etienne, A. T.; Trollope, A.; Chalabi, S.; Dell, A.; Stanley, P.; Haslam, S. M. Glycomics Profiling of Chinese Hamster Ovary Cell Glycosylation Mutants Reveals N-Glycans of a Novel Size and Complexity. *J. Biol. Chem.* **2010**, 285 (8), 5759–5775. <https://doi.org/10.1074/jbc.M109.068353>.
- (58) Smith, D. P.; Knapman, T. W.; Campuzano, I.; Malham, R. W.; Berryman, J. T.; Radford, S. E.; Ashcroft, A. E. Deciphering Drift Time Measurements from Travelling Wave Ion Mobility Spectrometry-Mass Spectrometry Studies. *Eur J Mass Spectrom (Chichester)* **2009**, 15 (2), 113–130. <https://doi.org/10.1255/ejms.947>.
- (59) May, J. C.; Morris, C. B.; McLean, J. A. Ion Mobility Collision Cross Section Compendium. *Anal. Chem.* **2017**, 89 (2), 1032–1044. <https://doi.org/10.1021/acs.analchem.6b04905>.
- (60) Paglia, G.; Williams, J. P.; Menikarachchi, L.; Thompson, J. W.; Tyldesley-Worster, R.; Halldórsson, S.; Rolfsson, O.; Moseley, A.; Grant, D.; Langridge, J.; Palsson, B. O.; Astarita, G. Ion Mobility Derived Collision Cross Sections to Support Metabolomics Applications. *Anal. Chem.* **2014**, 86 (8), 3985–3993. <https://doi.org/10.1021/ac500405x>.
- (61) Revercomb, H. E.; Mason, E. A. Theory of Plasma Chromatography/Gaseous Electrophoresis. Review. *Anal. Chem.* **1975**, 47 (7), 970–983. <https://doi.org/10.1021/ac60357a043>.
- (62) Bush, M. F.; Campuzano, I. D. G.; Robinson, C. V. Ion Mobility Mass Spectrometry of Peptide Ions: Effects of Drift Gas and Calibration Strategies. *Anal. Chem.* **2012**, 84 (16), 7124–7130. <https://doi.org/10.1021/ac3014498>.
- (63) Domon, B.; Costello, C. E. A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconjugate J* **1988**, 5 (4), 397–409. <https://doi.org/10.1007/BF01049915>.

Figure supplémentaire du chapitre II

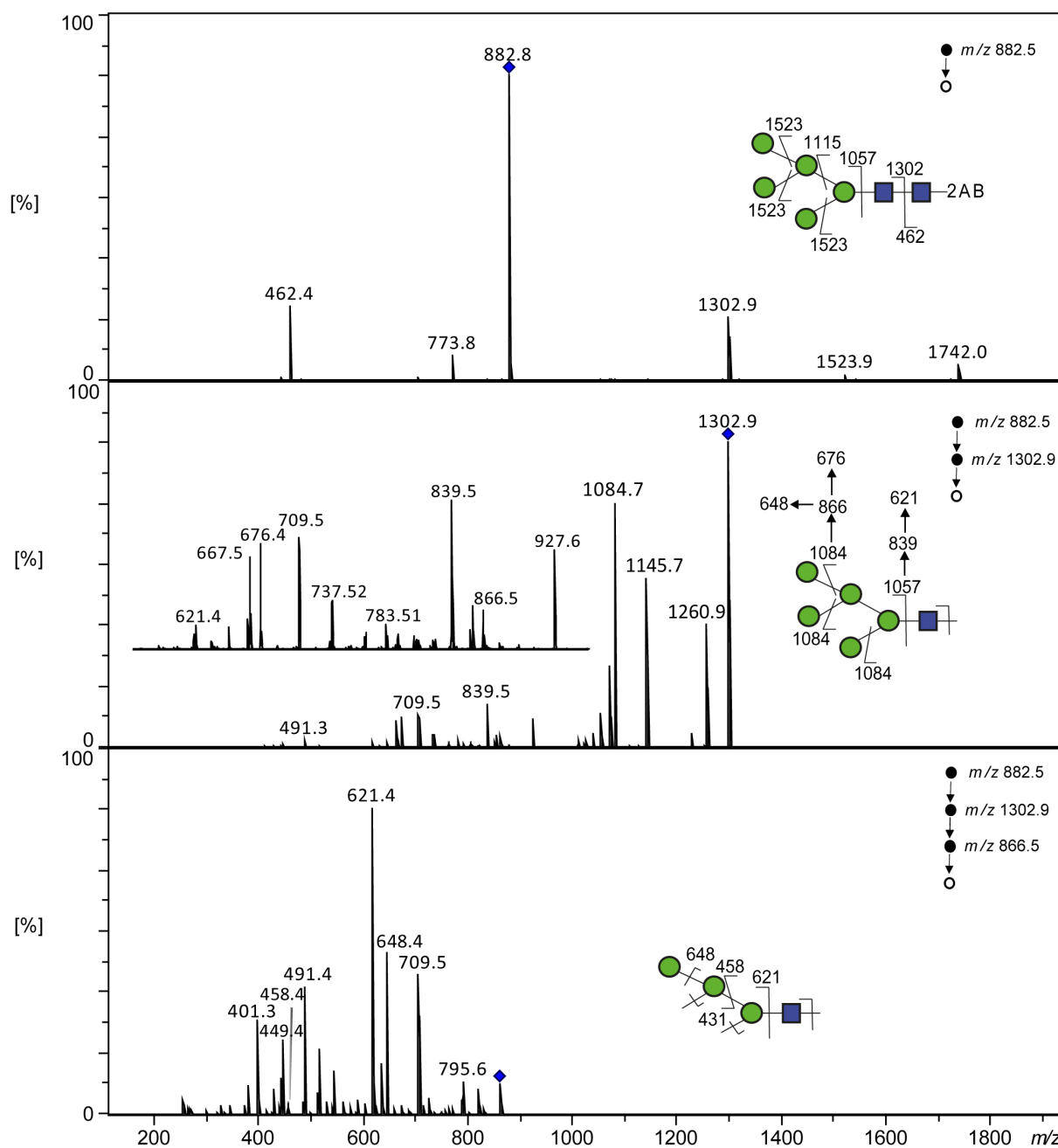
Supplemental Table 1: Experimental CCS (${}^{TW}CCS_{N_2 \rightarrow He}$) determined for permethylated 2AB Man-5 to Man-9 oligomannoside from *P. tricornutum* proteins and Ribonuclease B.

Permethylated 2AB derivatives	$[M+2Na]^{2+}$ (m/z)	$\Omega_{He} (\text{\AA}^2)$	
		Ribonuclease B	<i>P. tricornutum</i>
Man-5	882.5	343.8	344.3
Man-6	984.6	391.7	390.3
Man-7	1086.7	435.3	435.6
Man-8	1188.7	479.0	478.9
Man-9	1290.8	523.5	522.3

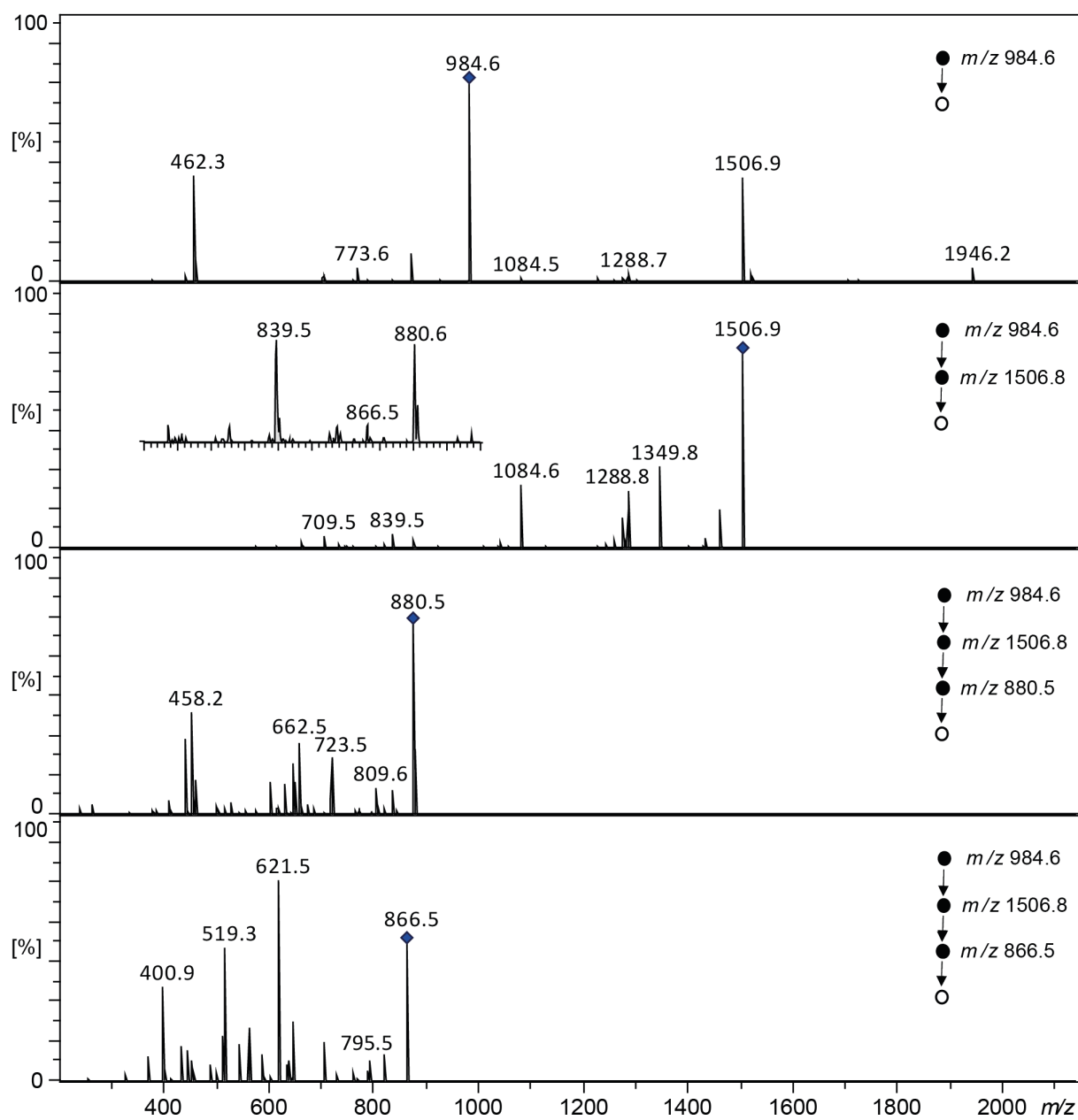
A

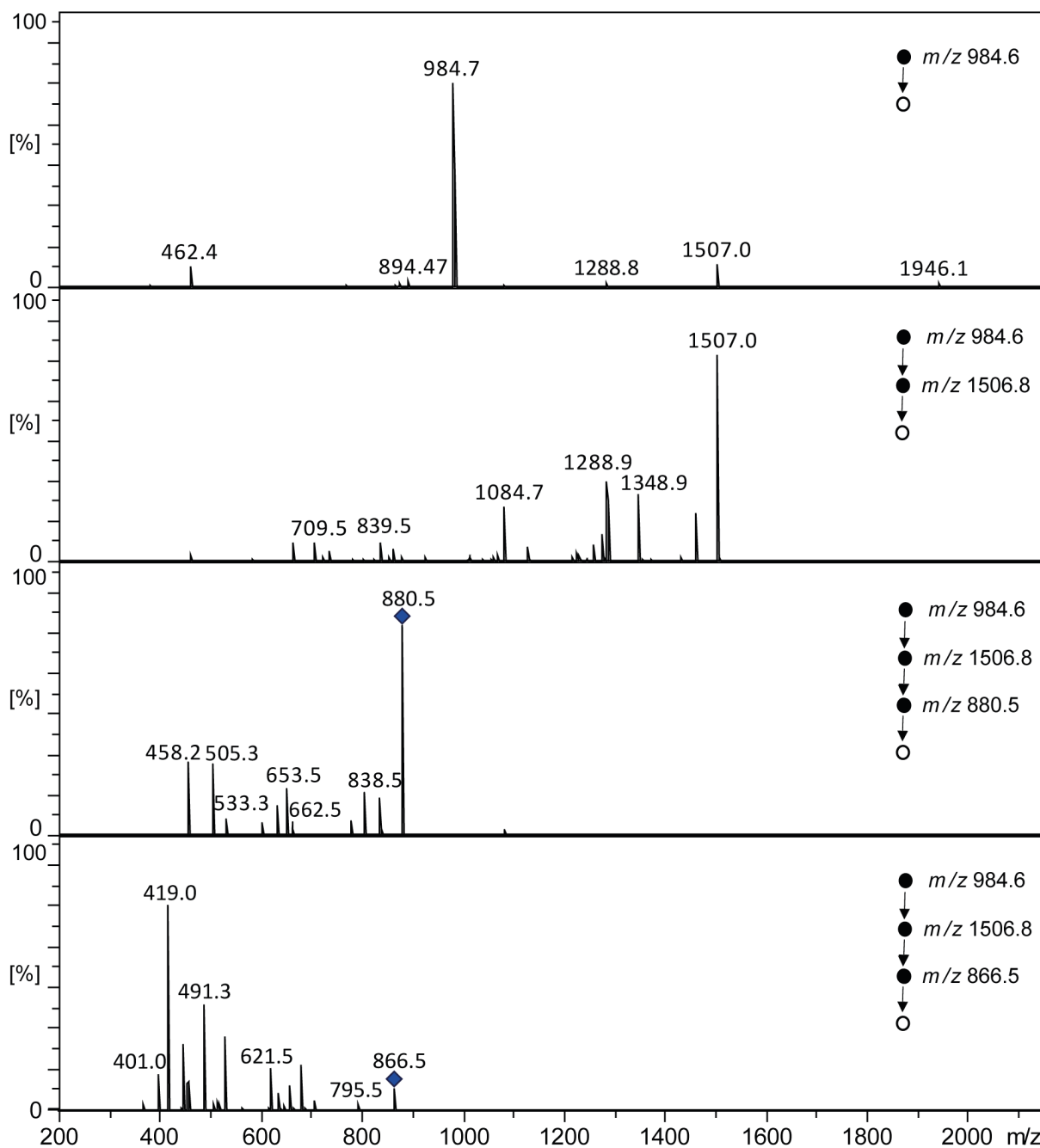


B



Supplemental Figure 1. (A) Scheme representing the fragmentation patterns for m/z 882.6 precursor ion ($[M+2Na]^{2+}$) of canonical 2AB-Man-5. (B) ESI-MSⁿ spectra ($n=2, 3$ and 4) of permethylated 2AB-Man-5 derivatives prepared from bovine ribonuclease B.

A

B

Supplemental Figure 2. (A) ESI-MSⁿ ($n= 2, 3$ and 4) of m/z 984.6 precursor ion ($[M+2Na]^{2+}$) of permethylated 2AB-Man-6 derivative isolated from (A) bovine ribonuclease B and (B) *P. tricornutum* proteins.

C. Tentative d'humanisation de la voie de la *N*-glycosylation chez la diatomée *P. tricornutum*

Comme rappelé dans l'Introduction, des anticorps monoclonaux fonctionnels dirigés contre des virus ont été produits avec succès chez la diatomée *P. tricornutum* (**Hempel et al., 2011 ; Hempel and Maier, 2012 ; Vanier et al., 2015 ; Hempel et al., 2017**). Ces travaux ont également montré que ces anticorps monoclonaux produits par *P. tricornutum* portent des structures oligomannosidiques de types Man₉GlcNAc₂ à Man₅GlcNAc, à l'instar des *N*-glycannes majoritairement identifiés sur les protéines de cette microalgue (**Baïet et al., 2011 ; Vanier et al., 2015**). Si *P. tricornutum* est donc apte à glycosyler des protéines recombinantes destinées à la thérapie humaine, les *N*-glycannes portés par les anticorps recombinants produits dans cette microalgue ne sont pas adaptés pour certaines applications thérapeutiques. En effet, comme évoqué dans l'introduction bibliographique, les IgG humaines portent majoritairement des *N*-glycannes bi-antennés nécessaires à l'acquisition de certaines de ses fonctions biologiques (voir le chapitre II. C. de l'Introduction).

L'absence de maturation des *N*-glycannes oligomannosidiques décorant les glycoprotéines de la diatomée *P. tricornutum* pose question. En effet, comme indiqué dans le chapitre VI. B. de l'Introduction, plusieurs travaux publiés montrent que *P. tricornutum* possède une machinerie enzymatique capable de maturer dans l'appareil de Golgi les *N*-glycannes oligomannosidiques des protéines en *N*-glycannes complexes. Ainsi, **Baïet et al., 2011** ont identifié par homologies de séquences plusieurs gènes de glycosyltransférases putatives impliquées dans les processus de *N*-glycosylation des protéines dans le génome de *P. tricornutum*. Les fonctionnalités de certains acteurs protéiques de la voie de la *N*-glycosylation chez *P. tricornutum* ont été démontrées. Par exemple, l'activité de la GnT I, enzyme responsable de l'initiation de la formation des *N*-glycannes complexes, a été démontrée par complémentation fonctionnelle chez le mutant CHO *lec1* déficient en cette activité (**Baïet et al., 2011**). De plus, des travaux récents ont permis de caractériser la fonctionnalité d'un transporteur golgien de GDP-fucose et celle d'une $\alpha(1,3)$ -fucosyltransférase ($\alpha(1,3)$ FuT) chez *P. tricornutum* (**Zhang et al., 2019**). Il est intéressant de noter que les transférases GnT I et $\alpha(1,3)$ FuT sont localisées dans les mêmes saccules médial-trans de l'appareil de Golgi de la microalgue (**Zhang et al., 2019**). De plus, les analyses bio-

informatiques réalisées par **Baïet *et al.*, 2011** ont montré la présence d'un gène codant une $\beta(1,2)$ -xylosyltransférase putative qui pourrait être responsable du transfert des résidus $\beta(1,2)$ -xylose sur le core du *N*-glycane. Toutes ces données expérimentales et bio-informatiques indiquent que la diatomée possède potentiellement toute la machinerie enzymatique capable de synthétiser des *N*-glycannes complexes portant des résidus *N*-acétylglucosamine en position terminale et arborant des résidus fucose liés en $\alpha(1,3)$ et des résidus xyloses liés en $\beta(1,2)$. Parallèlement, les analyses par spectrométrie de masse des *N*-glycannes libérés des protéines de la diatomée montrent qu'elles portent majoritairement des *N*-glycannes de type oligomannosidique et seulement des traces de *N*-glycannes complexes constitués d'un core $\text{Man}_3\text{GlcNAc}_2$ substitué par un résidu fucose lié en $\alpha(1,3)$ (**Figure 28A ; Baïet *et al.* 2011 ; Zhang *et al.*, 2019**).

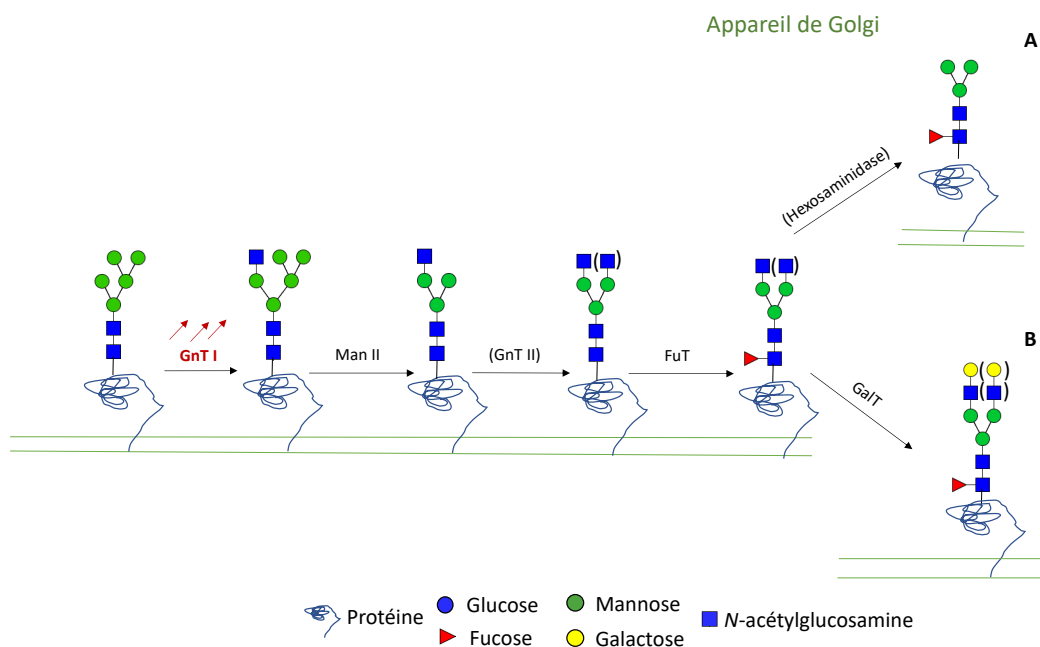


Figure 28 : Stratégie d'humanisation des voies de la N-glycosylation chez *P. tricornutum*

(A) Voie de la *N*-glycosylation chez la diatomée conduisant à la synthèse de *N*-glycannes complexes constitués d'un core $\text{Man}_3\text{GlcNAc}_2$ substitué par un résidu fucose lié en $\alpha(1,3)$.

(B) Stratégie choisie pour initier l'humanisation de la voie de synthèse des *N*-glycannes chez *P. tricornutum* et consistant en la surexpression de la GnT I puis le transfert de résidus galactose en position terminale par expression de la $\beta(1,4)$ -galactosyltransférase humaine dans l'optique de favoriser l'ajout de résidus galactose sur les *N*-glycannes possédant des résidus GlcNAc terminaux. La GnT II et les hexosaminidases indiquées dans ce schéma sont putatives. En conséquence, leurs implications dans la synthèse des *N*-glycannes complexes de *P. tricornutum* ne sont pas expérimentalement validées à ce jour.

GnT I et II : *N*-acétylglucosaminyltransférases I et II ; Man II : α -mannosidase II ; FuT : $\alpha(1,3)$ -fucosyltransférase.

Le très faible taux de maturation des *N*-glycannes complexes et l'absence de résidus GlcNAc terminaux sur ces structures peuvent avoir plusieurs origines possibles. Les résultats rapportés dans le chapitre II de ce manuscrit montrent que la structure de type $\text{Man}_5\text{GlcNAc}_2$ qui s'accumule sur les protéines chez *P. tricornutum* est une structure branchée qui est l'isomère substrat de la GnT I. En conséquence, la faible maturation de ce $\text{Man}_5\text{GlcNAc}_2$ en *N*-glycannes complexes reflète soit i) la faible activité de cette transférase chez *P. tricornutum*, soit ii) la présence d'activités hexosaminidases qui pourraient rapidement éliminer le résidu *N*-acétylglucosamine transféré par la GnT I (**Figure 28A**). En effet, deux gènes candidats codant des hexosaminidases sont prédites dans le génome de *P. tricornutum* (**Baïet et al., 2011**). Ces hexosaminidases présentent de fortes homologues de séquence avec les hexosaminidases impliquées dans l'élimination des résidus *N*-acétylglucosamine terminaux chez les plantes (voir le chapitre IV. C. de l'Introduction). La caractérisation de ces hexosaminidases putatives chez *P. tricornutum* est actuellement en cours au laboratoire dans le cadre d'un autre projet de recherche.

Dans ce contexte, la stratégie choisie pour initier l'humanisation des *N*-glycannes glycosylant les protéines synthétisées chez la diatomée *P. tricornutum* est décrite dans la **Figure 28B**. Elle se décline en deux volets : le premier volet consiste à produire des diatomées transformées surexprimant la GnT I endogène (PtGnT I). Cette surexpression a pour objectif d'augmenter l'activité GnT I dans la cellule et de contrebalancer une éventuelle activité hexosaminidase pour ainsi accumuler des *N*-glycannes complexes portant des résidus *N*-acétylglucosamine en position terminale sur les protéines de cette diatomée.

Le second volet de cette stratégie consiste en l'insertion d'un gène codant la $\beta(1,4)$ -galactosyltransférase humaine dans le génome de *P. tricornutum*. Cette enzyme est capable de transférer des résidus galactose sur des *N*-glycannes complexes portant des résidus *N*-acétylglucosamine en position terminale (**Figure 28B**). En conséquence, la transformation par la $\beta(1,4)$ -galactosyltransférase doit permettre à la diatomée de la protéger de l'élimination des résidus *N*-acétylglucosamine des *N*-glycannes par les hexosaminidases endogènes (si les résidus galactose sont ajoutés suffisamment rapidement). Cette complémentation doit, d'autre part, permettre d'obtenir des *N*-glycannes portant des résidus galactose en position terminale analogues à celles observées sur les *N*-glycannes des anticorps monoclonaux produits dans des cellules de mammifères (**Figure 28B**). Les résultats concernant ce second

volet étant préliminaires au moment du rendu de ce manuscrit seront présentés dans le chapitre « conclusions et perspectives ».

1. Surexpression de la GnT I de *P. tricornutum* (PtGnT I)

Le gène codant la PtGnT I a été introduit dans le génome de la diatomée sous le contrôle d'un promoteur inductible par le nitrate, permettant ainsi d'induire sa surexpression par l'ajout de nitrate dans le milieu de culture. Le gène codant la PtGnT I a été fusionné avec une séquence codant le peptide antigénique V5 pour faciliter la détection de la glycosyltransférase surexprimée chez *P. tricornutum* (PtGnT I-V5 ; **Figure 29A**). Le plasmide Pha-NR, utilisé lors de la transformation, comprend également un gène de sélection codant une résistance à un antibiotique, la zéocine. La transformation génétique de *P. tricornutum* a été réalisée par biolistique. Les cellules transformées ont été étalées sur des boîtes gélosées contenant de la zéocine comme agent de sélection. Au bout de 15 jours, les 121 colonies obtenues ont été de nouveaux étalées sur des boîtes gélosées. La vérification de l'insertion du transgène a été réalisée par PCR et a permis d'identifier 60 transformants positifs ayant intégré le transgène (**Figure 29B**). L'expression du gène codant la PtGnT I-V5 a ensuite été analysée sur les 60 clones positifs. Pour cela, une extraction microsomale à partir d'environ 5 millions de cellules a été réalisée. La totalité de l'extrait a été chargée sur un gel d'électrophorèse, transférée sur une membrane de nitrocellulose avant de réaliser un Western blot au moyen d'un anticorps spécifiquement dirigé contre le peptide V5. Les extraits protéiques de dix-neuf transformants présentaient une protéine immunodétectée à la taille attendue de 56 kDa (**Figure 29C**), ce qui démontrait une expression effective du transgène PtGnT I-V5 chez ces clones. La quantité de cellules utilisées pour le Western blot, étant à peu près identique, nous pouvons donc comparer approximativement l'expression de la PtGnT I-V5 dans les différents clones. Les clones 2 ; 65 ; 69 ; 105 ; 107 et 109 ont été sélectionnés pour la suite des analyses présentant les plus forts taux de PtGnT I-V5 exprimée (**Figure 29C**). Le clone GX5 qui surexprime PtGnT I-V5 et qui est issu d'une précédente étude (**Ovide et al., 2018 ; Zhang et al., 2019**), a également été inclus pour la suite de ce travail.

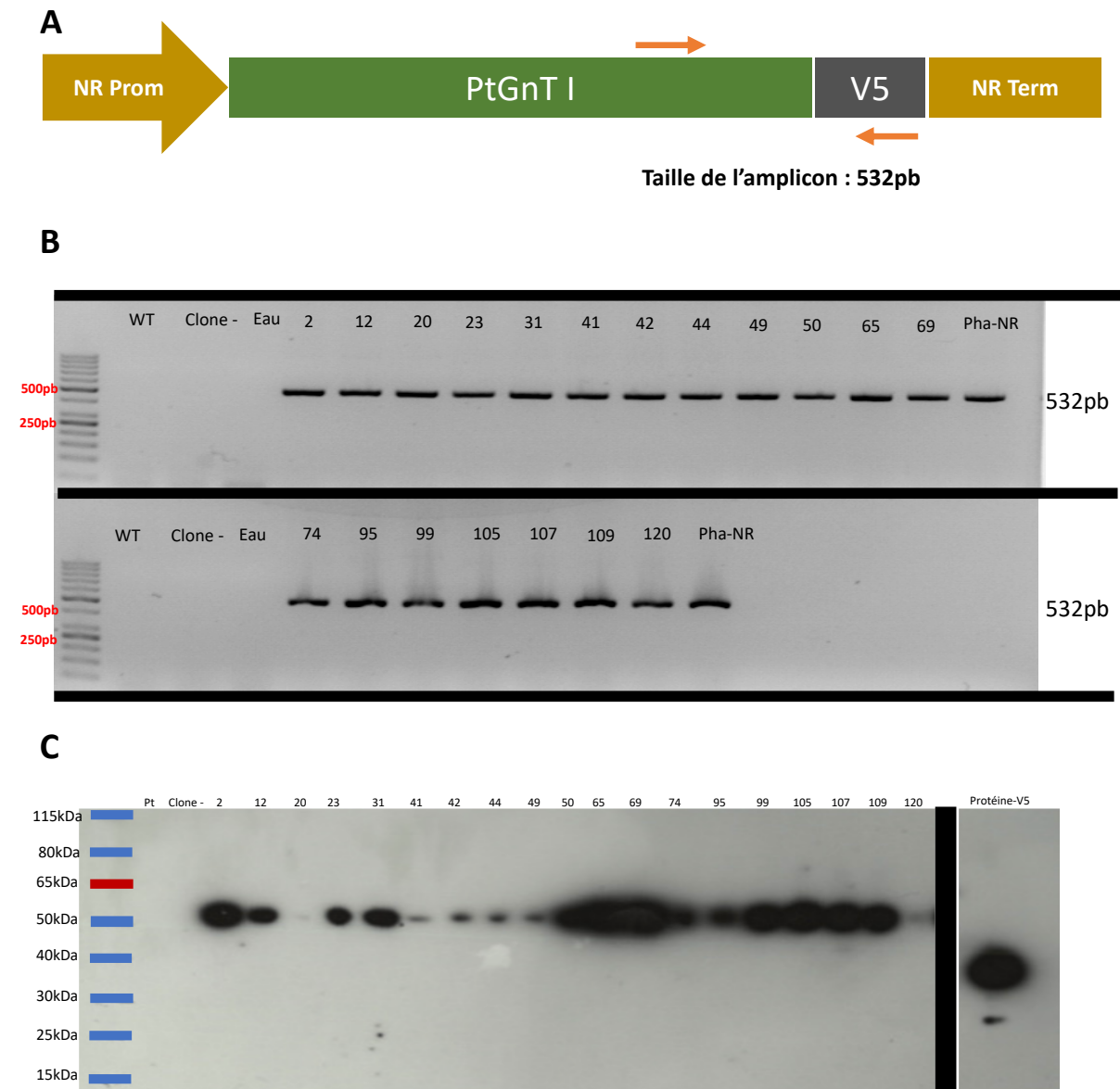


Figure 29 : Criblages génique et protéique des clones ayant inséré le gène codant la PtGnT I-V5

(A) Positionnement des couples d'amorces (flèches de couleur orange) utilisés pour l'amplification par PCR de la séquence PtGnT I-V5. Prom et Term : promoteur et terminateur du gène codant la nitrate réductase (NR).

(B) Produits de PCR obtenus à partir des transformants surexprimant la PtGnT I-V5. Le témoin positif est le plasmide Pha-NR utilisé lors de la transformation par biolistique. Les témoins négatifs sont les produits de PCR de la souche sauvage (WT), de l'eau et d'un clone correspondant à une transformation par le vecteur pPha-NR mais dépourvu des gènes d'intérêt (Clone -).

(C) Immunodétection par des anticorps spécifiquement dirigés contre le peptide V5 des protéines microsomales isolées à partir des transformants de *P. tricornutum* surexprimant la PtGnT I-V5. Le témoin positif est une protéine commerciale contenant le peptide V5. Il est à noter que cette figure a été construite à partir d'un montage d'un blot qui a été découpé pour ne laisser apparaître que les informations intéressantes dans le cadre de ce projet.

En conclusion, le criblage a permis d'obtenir un rendement de 49,6% en transformants ayant inséré le gène codant la PtGnT I-V5 et de 15,7% en transformants exprimant la transférase PtGnT I-V5 (**Tableau 9**).

Tableau 9 : Bilan du criblage des transformants PtGnTI-V5

Nombre de colonies isolées	Nombre de colonies ayant intégré le gène de la <i>PtGnT I-V5</i>	Transformants possédant le gène <i>PtGnT I-V5</i> (%)	Nombre de clones exprimant la protéine PtGnT I-V5	Rendement final de transformation (%)
121	60	49,58	19	15,70

2. Séquençage protéique de la PtGnT I- V5 par chromatographie liquide couplée à de la spectrométrie de masse

Afin de confirmer que la protéine détectée par Western blot correspond effectivement à PtGnT I-V5, une analyse protéomique a été réalisée. Pour cela, un extrait des protéines microsomales du clone GX5 a été immunoprécipité au moyen de l'anticorps dirigé contre le peptide V5. L'analyse par chromatographie liquide couplée à la spectrométrie de masse des peptides obtenus à la suite de la digestion par la trypsine de la protéine immunoprécipitée a permis d'obtenir un taux de recouvrement de 41,2% de la séquence correspondant à la protéine PtGnT I-V5. Ce résultat permet de confirmer que la protéine détectée sur le gel est bien la transférase PtGnT I-V5 (**Figure 30**).

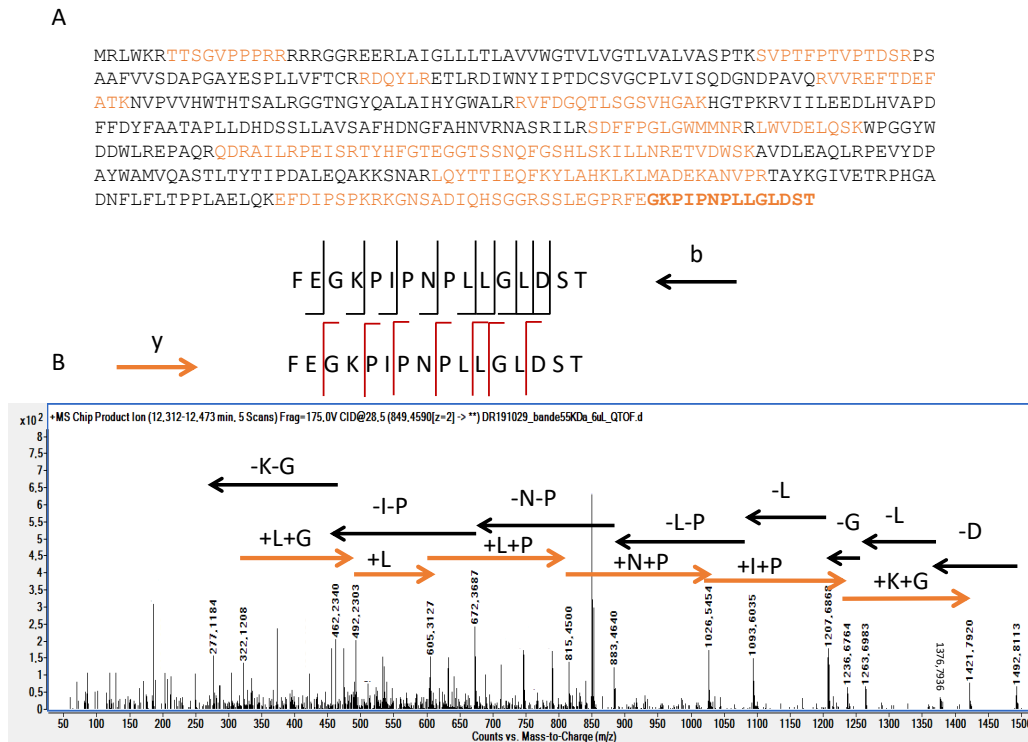


Figure 30 : Séquençage peptidique de la protéine PtGnT I-V5 immunoprécipitée à partir de la fraction microsomale du clone GX5

(A) Séquence protéique de la PtGnTI-V5 surexprimée chez *P. tricornutum*. Les peptides générés par traitement à la trypsine de la protéine immunoprécipitée (clone GX5) et identifiés par chromatographie liquide couplée à de la spectrométrie de masse sont indiqués en orange. Le peptide V5 situé du côté C-terminal de la protéine est indiqué en gras. Le taux de recouvrement des peptides avec la séquence protéique obtenu lors de cette analyse est de 41,2%.

(B) Spectre de fragmentation MS² du peptide FEGKPIPNLLGLDST comprenant le peptide V5 situé du côté C-terminal de la PtGnT I-V5. b et y correspondent aux fragments peptidiques observés.

3. Analyses des transformants sélectionnés

a. Analyse des transcrits des clones sélectionnés

Une analyse plus poussée des clones sélectionnés lors de l'étape de criblage précédente a été réalisée. Chaque clone a été cultivé pendant 10 jours, avec ou sans induction par le nitrate. Une première induction par le nitrate a été réalisée à J6 de culture, suivie d'une seconde induction à J8 pour induire l'expression de PtGnT I-V5 dans les cellules. Une extraction des ARN des clones sélectionnés, induits ou non induits par le nitrate, suivie d'une RT-PCR a permis l'amplification d'un fragment d'ADNc. Cette amplification a été réalisée en

utilisant des amorces spécifiques du gène *PtGnT I-V5* positionnées sur les mêmes sites que lors du criblage précédent (**Figure 29A**). L'analyse des transcrits montrent une amplification exclusive d'ADNc chez les clones après l'induction au nitrate, confirmant ainsi que le gène est exclusivement transcrit lors d'un apport de nitrate dans le milieu de culture (**Figure 31**). Les données d'expression montrent que les clones 2, 65 et 107 expriment le gène codant la *PtGnT-V5* de façon plus importante que les autres clones. Lors de l'amplification des ADNc, des fragments amplifiés non expliqués à une taille d'environ 300pb ont pu être observés chez les clones 2, 69 et 109.

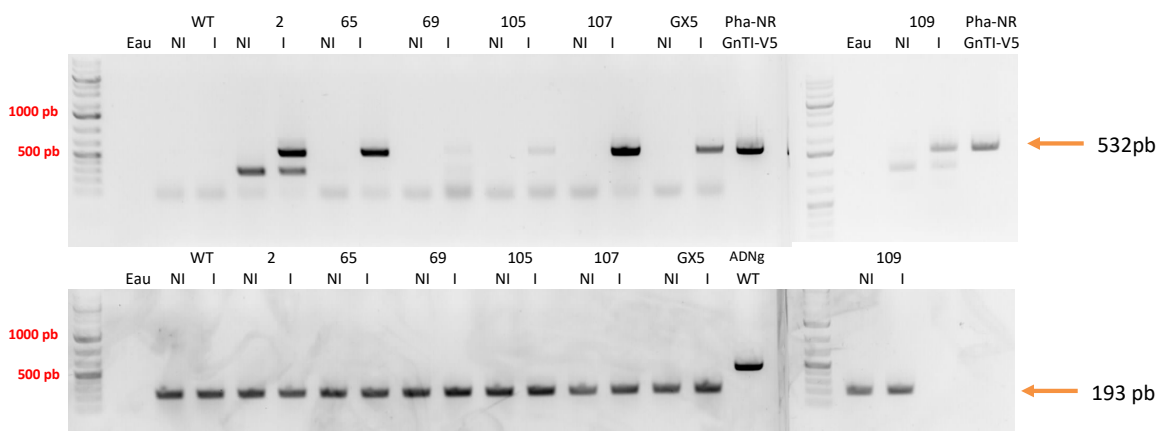


Figure 31 : Analyses par RT-PCR validant la dépendance de la transcription du gène codant la *PtGnT I-V5* de l'induction des cellules par le nitrate.

L'amplification d'un fragment d'ADNc a été réalisée sur des préparations issues de cultures non-induites (NI) et induites (I) de chaque clone. Les fragments d'ADNc ont été obtenus après une extraction d'ARN totaux des clones suivi d'une réaction de RT-PCR permettant de convertir les ARNm purifiés en ADNc. EF1A, codant un facteur d'élongation, est utilisé comme un gène de référence (193 pb). Le témoin négatif de réaction PCR correspond aux produits de PCR de l'eau pure et le témoin négatif de transformation génétique correspond aux produits de PCR de la souche sauvage *Pt.1.8.6* (WT). Le témoin positif est le vecteur pPha-NR utilisé pour la transformation génétique possédant le gène codant la *PtGnT I-V5*.

b. Analyse par Western blot des clones sélectionnés

Une seconde analyse a consisté à vérifier que l'expression de la protéine *PtGnT I-V5* est également dépendante de l'induction par le nitrate. Une extraction microsomale suivie d'une analyse par Western blot réalisée à partir d'extraits microsomaux a permis de conforter les résultats obtenus lors de l'analyse des transcrits et ainsi de montrer que l'expression de la protéine *GnT I-V5* (56 kDa) est en effet dépendante de l'induction par le nitrate (**Figure 32**).

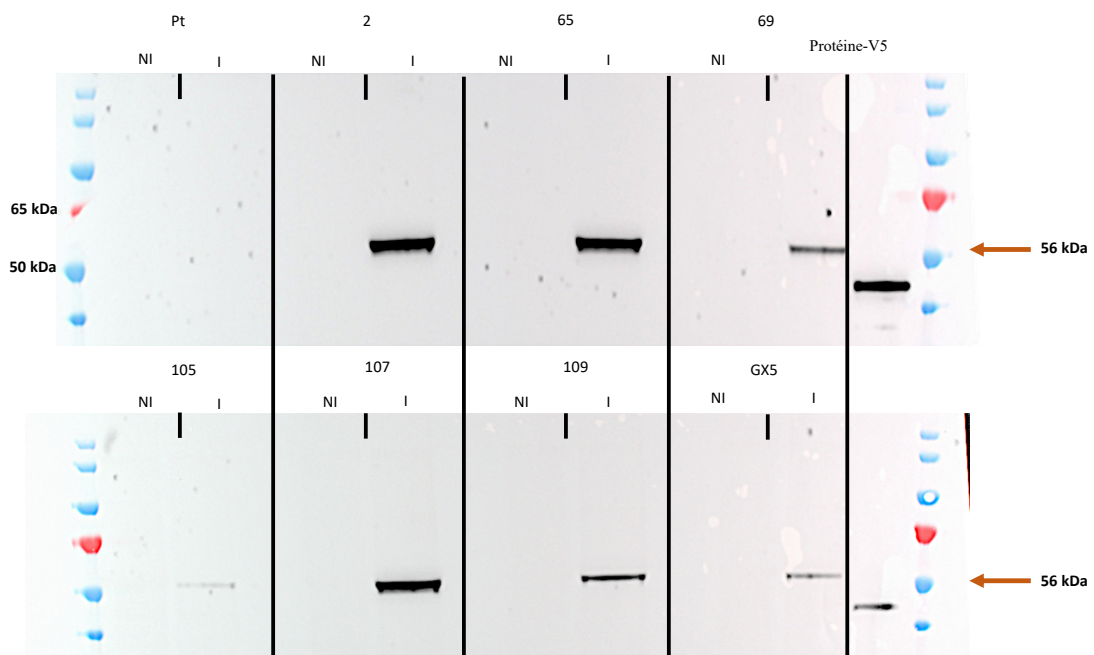


Figure 32 : Analyse par Western blot montrant que l'expression de la protéine PtGnT I-V5 est dépendante de l'induction des cellules de *P. tricornutum* par le nitrate.

Immunodétection par les anticorps anti-V5 des protéines microsomales issues des cultures non-induites (NI) et induites (I) de chaque clone. Le témoin négatif d'induction et de l'immunodétection est les protéines microsomales de la souche sauvage Pt 1.8.6 (Pt). Le témoin positif d'immunodétection est une protéine commerciale contenant le peptide antigénique V5. Les marqueurs de taille sont commercialisés par ThermoScientific.

4. Analyse de la N-glycosylation des protéines des clones sélectionnés

La surexpression de la transférase PtGnTI-V5 a été confirmée chez les clones sélectionnés comme le montre le paragraphe précédent. Ce paragraphe se focalise sur les conséquences de cette surexpression sur une éventuelle modification de la maturation des *N*-glycannes portés par les protéines de cette diatomée. Ainsi, des analyses de la *N*-glycosylation des protéines totales synthétisées par les différents clones de *P. tricornutum* surexprimant la transférase PtGnT I-V5 ont donc été réalisées par Western blot au moyen d'un anticorps dirigé contre l'épitope core $\alpha(1,3)$ -fucose et d'un anticorps dirigé contre l'épitope core $\beta(1,2)$ -xylose (Ebert *et al.*, 2015 ; Jansing *et al.*, 2019 ; Zhang *et al.*, 2019 ; Lucas *et al.*, 2020). L'objectif de cette étude est d'évaluer si l'induction de la surexpression de la PtGn T I-V5 par le nitrate conduit à l'accumulation de *N*-glycannes complexes portant les épitopes core $\alpha(1,3)$ -fucose et/ou core $\beta(1,2)$ -xylose sur les protéines de la diatomée.

Dans un second temps, des expériences de transfert enzymatique *in vitro* de galactose ont été réalisées sur les extraits protéiques totaux issus des différents clones. Le principe de ce test est d'incuber les glycoprotéines des transformants avec une $\beta(1,4)$ -galactosyltransférase bovine commerciale qui est capable de transférer des résidus galactose sur des *N*-glycannes possédant des résidus *N*-acétylglucosamine terminaux (**Bardor *et al.*, 2003b ; Bosco *et al.*, 2008 ; Thomann *et al.*, 2015 ; Figure 33**). Le résultat de cette réaction de transfert de galactose *in vitro* est ensuite soumis à une affinodétection en utilisant la lectine RCA120 (*Ricinus communis*) qui reconnaît le motif lactosamine (Gal β 1-4GlcNAc) éventuellement produit par la réaction enzymatique (**Bardor *et al.*, 2003b ; Itakura *et al.*, 2007 ; Baïet *et al.*, 2011 ; Figure 33**). La reconnaissance de la lactosamine par la lectine RCA120 est donc le reflet de la présence de *N*-glycannes portant des résidus *N*-acétylglucosamine en position terminale sur les glycoprotéines de *P. tricornutum*.

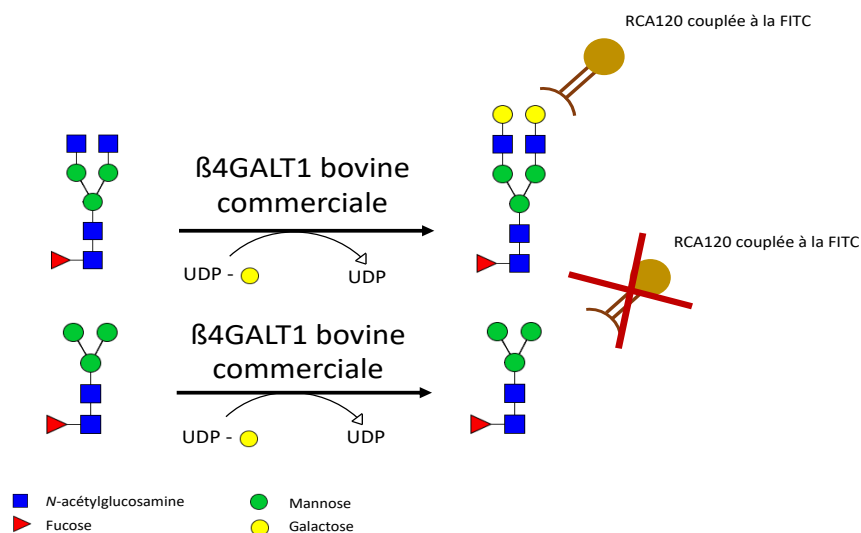


Figure 33 : Principe de l'affinodétection des épitopes glucidiques obtenus après la galactosylation enzymatique *in vitro* des *N*-glycannes. Cette affinodétection des épitopes glucidiques a été effectuée sur une empreinte protéique obtenue après un transfert sur une membrane de nitrocellulose.

Cette analyse permet l'évaluation de la présence de *N*-glycannes complexes portant des résidus *N*-acétylglucosamine terminaux sur les protéines des transformants de *P. tricornutum* exprimant la PtGnT I-V5. Le principe de la galactosylation enzymatique *in vitro* consiste à transférer des résidus galactose par une incubation avec la $\beta(1,4)$ -galactosyltransférase bovine et le substrat donneur UDP-galactose sur des résidus *N*-acétylglucosamine terminaux présents potentiellement sur les *N*-glycannes. Le motif lactosamine ainsi obtenu est ensuite affinodéTECTÉ par la RCA120 couplée à un fluorochrome (FITC).

a. Immunodétection des protéines des clones sélectionnés par des anticorps dirigés contre les épitopes core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose

Des extraits protéiques des clones sélectionnés ont été séparés sur un gel d'électrophorèse puis transférés sur une membrane de nitrocellulose. Les résultats obtenus après l'immunodétection par les anticorps dirigés contre l'épitope core $\alpha(1,3)$ -fucose ou l'épitope core $\beta(1,2)$ -xylose montrent une augmentation très faible de la teneur en glycoprotéines fucosylées et xylosylées après l'induction de la surexpression de la PtGnT I-V5 par le nitrate pour les clones 65, 105 et 107 de *P. tricornutum* (**Figure 34**). Pour les autres clones, aucune augmentation significative du signal n'est observée en comparant l'expression des clones induits et non induits (données non présentées). Ces résultats suggèrent donc que la voie de la *N*-glycosylation des protéines de la diatomée semble être faiblement impactée par la surexpression de la PtGnT I-V5 chez les clones 65, 105 et 107. L'accumulation de *N*-glycannes complexes portant à la fois des résidus $\alpha(1,3)$ -fucose et $\beta(1,2)$ -xylose semble plus significative pour le clone 105 (**Figure 34**). Il est à noter également que les clones 65 et 107 semblent accumuler fortement des résidus $\beta(1,2)$ -xylose et plus faiblement des résidus $\alpha(1,3)$ -fucose après la surexpression de la PtGnT I-V5. Les clones 2 et GX5 ont également été conservés pour la suite de l'étude.

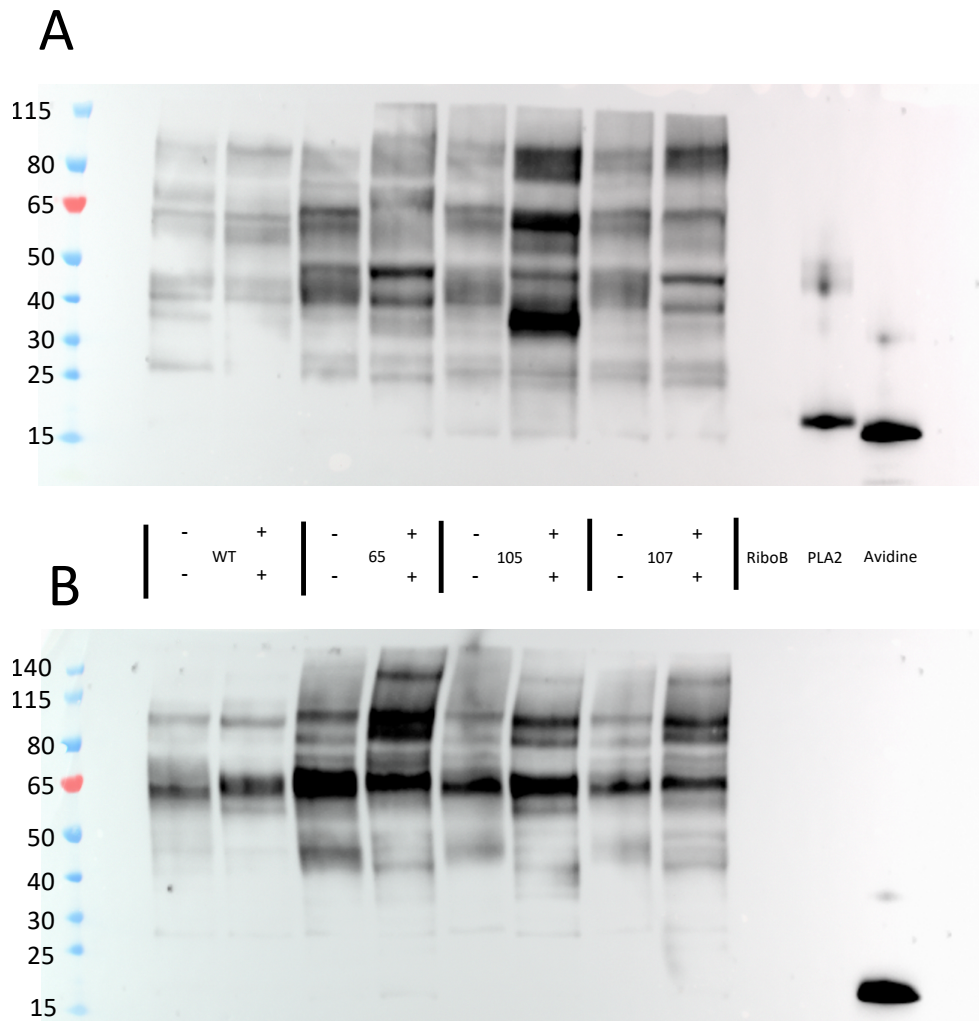


Figure 34 : Immunodétection des épitopes core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose portés par les *N*-glycannes des protéines totales de *P. tricornutum* isolées des clones 65, 105 et 107.

(A) Immunodétection par l'anticorps anti $\alpha(1,3)$ -fucose des épitopes core $\alpha(1,3)$ -fucose des *N*-glycannes associés aux protéines totales des clones sélectionnés. Le témoin négatif utilisé ici est de la ribonucléase B (RiboB), protéine portant des *N*-glycannes oligomannosidiques (Joao and Dwek, 1993). Les témoins positifs sont de la phospholipase A2 du venin d'abeille (PLA2) portant un épitope core $\alpha(1,3)$ -fucose sur ses *N*-glycannes (Lai and Her, 2002) et de l'avidine recombinante produite dans le maïs, protéine glycosylée portant des épitopes core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose (Bardor et al., 2003a). Les extraits protéiques totaux de *P. tricornutum* ont été isolés à partir de clones non induits (–) ou induits (+) par le nitrate.

(B) Immunodétection par l'anticorps anti $\beta(1,2)$ -xylose des épitopes core $\beta(1,2)$ -xylose des *N*-glycannes associés aux protéines totales des clones sélectionnés. Les témoins négatifs utilisés ici sont de la ribonucléase B (Joao and Dwek, 1993) et de la PLA2 de venin d'abeille (Lai and Her, 2002). Le témoin positif est de l'avidine recombinante produite dans le maïs. Les extraits protéiques totaux de *P. tricornutum* ont été isolés à partir de clones non induits (–) ou (+) induits par le nitrate.

b. Tests de galactosylation *in vitro* réalisés sur les protéines totales issues des clones sélectionnés

Pour chaque clone sélectionné (clones 2, 65, 69, 105, 107, 109) et du témoin positif GX5, nous avons réalisé un test de galactosylation enzymatique *in vitro* sur les protéines totales issues des diatomées induites (+) ou non induites (-) par le nitrate. Pour chaque échantillon, un témoin négatif (-) réalisé sans l'ajout de la $\beta(1,4)$ -galactosyltransférase bovine a été effectué en parallèle. Les extraits protéiques totaux ont ensuite été séparés sur un gel d'électrophorèse, transférés sur une membrane de nitrocellulose puis soumis à une affindétection par la RCA120 couplée à un fluorochrome. La présence de *N*-glycannes portant des résidus *N*-acétylglucosamine en position terminale chez les glycoprotéines des transformants sera validée par l'absence de bandes fluorescentes dans les conditions dites témoins (-/- ; -/+ ; +/-) (**Figure 35A**) et la présence de bandes fluorescentes après l'induction par le nitrate et l'ajout de $\beta(1,4)$ -galactosyltransférase bovine lors du test condition noté +/+.

Les résultats obtenus à la suite du test de galactosylation *in vitro* n'ont pas permis de mettre en évidence de reconnaissance par la RCA 120 pour 5 des 6 clones testés ici (données non montrées), notamment pour les clones 2 et GX5 (**Figure 35A**). Ces clones n'accumulent donc pas ou trop peu de *N*-glycannes possédant des résidus *N*-acétylglucosamine en position terminale. Le clone 105 est le seul clone qui montre plusieurs bandes fluorescentes correspondant à des glycoprotéines migrant entre 50 et 65kDa (**Figure 35A**). Ce résultat suggère que ce clone est capable d'accumuler en faible quantité des *N*-glycannes complexes portant des résidus *N*-acétylglucosamine en position terminale.

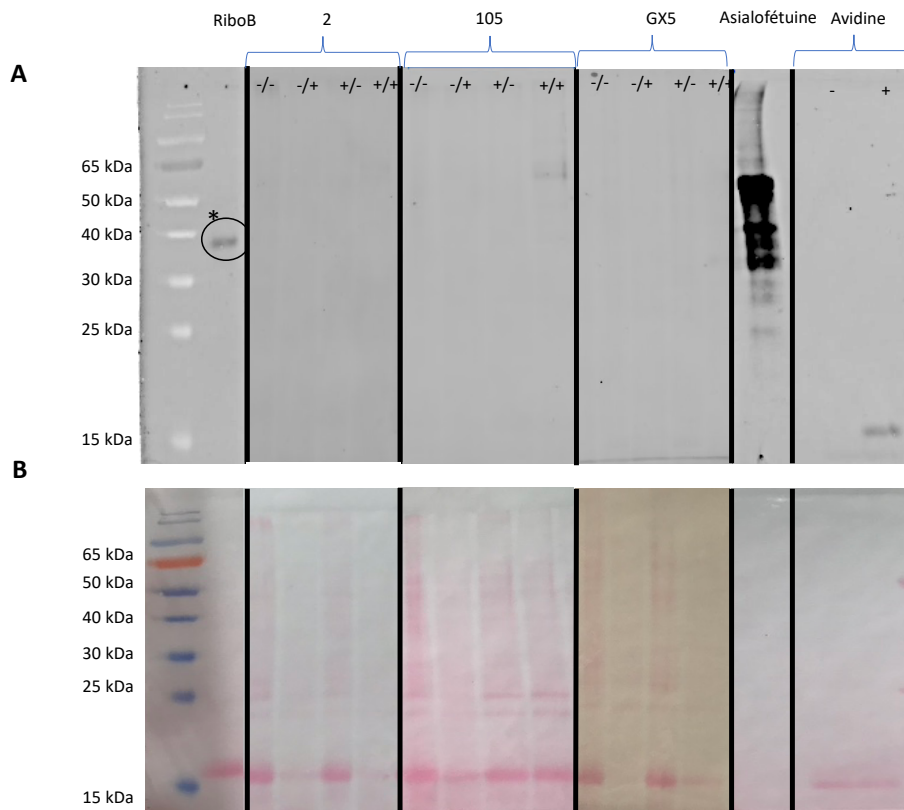


Figure 35 : Tests de galactosylation *in vitro* réalisés sur les protéines totales extraites à partir du clone 105 de *P. tricornutum* surexprimant la PtGnT I-V5

Affinodétection par la lectine RCA120 couplée à la FITC des protéines totales du clone 2, 105 et GX5 (A). -/- : protéines obtenues sans une induction par le nitrate et sans une galactosylation *in vitro* ; -/+ : sans une induction par le nitrate et avec une galactosylation *in vitro* ; +/- : avec une induction par le nitrate et sans une galactosylation *in vitro* ; +/+ : avec une induction par le nitrate et avec une galactosylation *in vitro*. La ribonucléase B (RiboB), protéine portant des *N*-glycannes oligomannosidiques (Joao and Dwek, 1993) et d'une taille d'environ 16 kDa, est utilisée comme un témoin négatif. L'asialofétuine, portant des résidus $\beta(1,4)$ -galactose en position terminale de ces *N*-glycannes, est le témoin positif de la révélation à la RCA120 (Yet et al., 1988). Il est à noter qu'une quantité trop faible d'asialofétuine (0,5 μ g) a été déposée pour être visible après la coloration au rouge ponceau (B). L'avidine recombinante produite dans le maïs est utilisée comme un témoin positif de la galactosylation *in vitro* puisqu'elle porte des *N*-glycannes ayant des résidus GlcNAc terminaux. La bande réactive entourée et signalée par une étoile est une protéine contaminante galactosylée présente dans la ribonucléase B. Il est à noter que cette figure a été construite à partir de trois Eastern blot différents révélés en même temps.

c. Analyse par spectrométrie de masse des *N*-glycannes des clones 2 et GX5

En complément des résultats présentés ci-dessus, nous avons entrepris l'analyse des profils de *N*-glycosylation des protéines des clones 2, 105 et GX5. L'analyse des *N*-glycannes du clone 105 est en cours à la date de transmission de ce manuscrit. Pour les clones 2 et GX5, l'immunodétection des épitopes core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose (données non présentées) et le test de la galactosylation enzymatique *in vitro* (**Figure 35A et 35B**) suggèrent que les profils des glycoprotéines de ces deux clones ne sont que très peu affectés par l'expression de la PtGnT I-V5. Les glycoprotéines ont donc été extraites puis traitées par la PNGase A capable de cliver la liaison osidique entre la *N*-acétylglucosamine proximale du *N*-glycanne et l'asparagine présente sur le site consensus de glycosylation de la protéine. Ces *N*-glycannes ont ensuite été perméthylés pour en améliorer la détection et l'analyse par spectrométrie de masse. Les oligosaccharides perméthylés ont ensuite été analysés par chromatographie liquide couplée à un spectromètre de masse électrospray (LC-ESI MS). Les résultats obtenus montrent que la population en *N*-glycannes du sauvage est identique à celle retrouvée chez les transformants surexprimant la PtGnT I-V5 (**Figure 36**). Les structures majoritaires sont des *N*-glycannes oligomannosidiques allant de $\text{Man}_9\text{GlcNAc}_2$ à $\text{Man}_5\text{GlcNAc}_2$ et un glycanne complexe de structure $\text{Man}_3\text{FucGlcNAc}_2$ comme décrit par **Baïet et al., 2011**. Les proportions relatives des *N*-glycannes oligomannosidiques chez les transformants sont similaires comparées à celles de la souche sauvage avec respectivement environ 50%, 19%, 17%, 9% et 5% pour le $\text{Man}_9\text{GlcNAc}_2$, $\text{Man}_8\text{GlcNAc}_2$, $\text{Man}_7\text{GlcNAc}_2$, $\text{Man}_6\text{GlcNAc}_2$ et $\text{Man}_5\text{GlcNAc}_2$ (**Figure 36**). Des structures complexes très minoritaires sont également observées dans ce spectre. Cependant, ces oligosaccharides sont exempts de résidu *N*-acétylglucosamine en position terminale et leurs intensités ne diffèrent pas de façon notable entre la souche sauvage et les transformants (**Figure 36**). Leurs structures n'ont pas pu être déterminées avec certitude compte tenu de leurs faibles intensités. Ces résultats suggèrent donc que la surexpression de la PtGnT I-V5 ne modifie pas de façon détectable les profils de glycosylation des glycoprotéines des clones 2 et GX5. Ces mêmes résultats ont été observés suite à l'analyse des *N*-glycannes du clone 105.

5. Conclusion

La transformation génétique de *P. tricornutum* par l'insertion d'un gène codant la PtGnT I-V5 a permis d'isoler 19 transformants surexprimant cette protéine. L'analyse protéomique de la PtGnT I-V5 issue du clone GX5 a permis de confirmer l'expression de cette protéine. Cependant, les analyses par Western blot et les tests de galactosylation *in vitro* n'ont pas mis en évidence une accumulation de *N*-glycannes complexes chez les transformants. Les résultats les plus notables ont été obtenus pour les clones 65, 105 et 107 mais suggèrent cependant que si l'expression de PtGnT I-V5 a impacté la voie de *N*-glycosylation des protéines de *P. tricornutum*, cet impact est probablement faible.

Les analyses par spectrométrie de masse des profils des *N*-glycannes des clones 2, 105 et GX5 n'ont pas révélé de modifications de la voie de *N*-glycosylation des protéines de *P. tricornutum*.

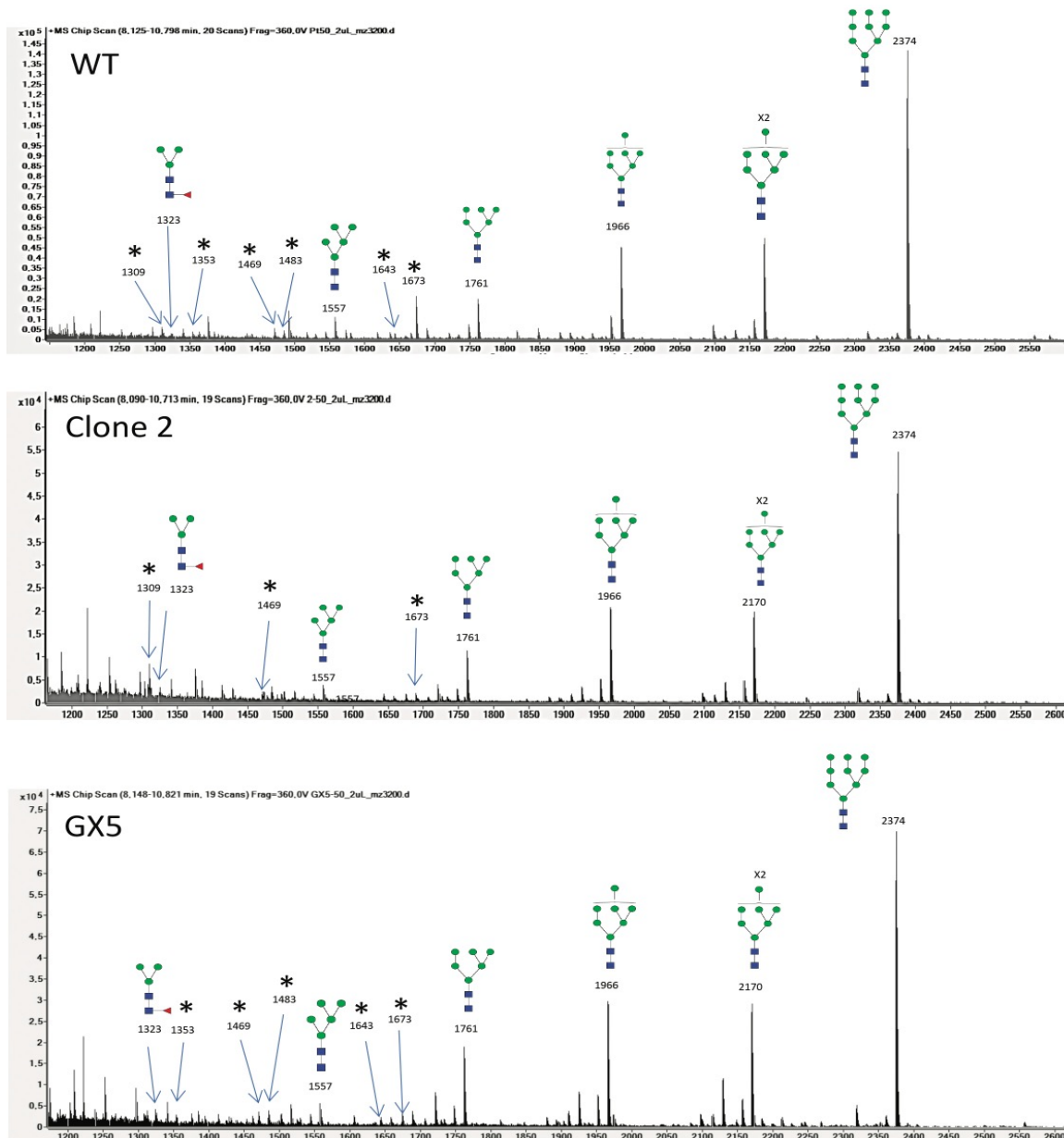


Figure 36 : Analyse des *N*-glycannes de *P. tricornutum* obtenue par chromatographie liquide couplée à la spectrométrie de masse de la souche sauvage (WT) et des transformants surexprimant PtGnT I-V5 (clones 2 et GX5)

Spectre de masse des *N*-glycannes extraits à partir des protéines totales de la souche sauvage (WT) comparé aux spectres de masse des *N*-glycannes extraits à partir des protéines totales des deux transformants surexprimant PtGnT I-V5. L'intensité des espèces ioniques est représentée en ordonnée et le rapport m/z en abscisse. Les spectres des transformants 2 et GX5 sont similaires à celui du sauvage, signifiant que la surexpression de PtGnT I-V5 n'affecte pas à un niveau détectable la formation des *N*-glycannes complexes portant des résidus *N*-acétylglucosamine terminaux chez les clones 2 et GX5. Les structures majoritaires retrouvées sont des *N*-glycannes oligomannosidiques allant de $\text{Man}_9\text{GlcNAc}_2$ à $\text{Man}_5\text{GlcNAc}_2$ et le glycanne complexe $\text{Man}_3\text{FucGlcNAc}_2$. D'autres structures complexes minoritaires sont également détectées (*).

GlcNAc : *N*-acétylglucosamine ; Man : mannose ; WT : wild-type

Chapitre 4

Conclusion-discussion-perspectives

Conclusion

Les anticorps monoclonaux sont des molécules biologiques utilisées dans le cadre de la thérapie humaine. La majorité des anticorps monoclonaux commercialisés à ce jour sont produits dans des cellules de mammifères (Walsh, 2018). Les contraintes importantes liées à la production de ces anticorps monoclonaux dans des cellules de mammifères ont poussé la communauté scientifique à identifier de nouvelles cellules hôtes pour ce type de production. Des publications récentes ont montré que *P. tricornutum*, une microalgue unicellulaire de la famille des diatomées, est aujourd'hui capable de produire des anticorps monoclonaux fonctionnels et glycosylés (Hempel et al., 2011 ; Hempel and Maier, 2012 ; Vanier et al., 2015 ; Hempel et al., 2017 ; Vanier et al., 2018). Cependant, la *N*-glycosylation réalisée par cette diatomée ne permettrait pas à l'heure actuelle d'utiliser ces anticorps monoclonaux en thérapie humaine pour certaines applications thérapeutiques. L'objectif de ce projet de thèse était donc de compléter les connaissances sur le processus de la *N*-glycosylation des protéines chez *P. tricornutum* et d'initier des stratégies d'ingénierie métabolique pour permettre l'accumulation chez cette diatomée de *N*-glycannes complexes portant des résidus GlcNAc terminaux. La synthèse par *P. tricornutum* de tels *N*-glycannes est l'étape nécessaire pour entreprendre par la suite l'humanisation des structures glycaniques. Les données collectées principalement par analyses bioinformatiques et les résultats biochimiques préliminaires avaient conduit la communauté scientifique à proposer une voie de la *N*-glycosylation des protéines de *P. tricornutum* résumée dans la **Figure 21** (Baïet et al., 2011 ; Mathieu-Rivet et al., 2014). L'objectif principal de la thèse était d'apporter des données biochimiques solides sur la structure des *N*-glycannes portés par les protéines de la diatomée. Ces données d'analyse structurale ont permis par la suite, d'envisager des stratégies d'ingénierie pertinentes de façon à initier l'humanisation des voies de *N*-glycosylation de *P. tricornutum*.

A. Le précurseur oligosaccharidique est tronqué d'un résidu glucose chez *P. tricornutum*

Ce projet de thèse s'est tout d'abord concentré sur l'étude biochimique du précurseur oligosaccharidique synthétisé dans le RE. Ces analyses réalisées par spectrométrie de masse

ont permis de mettre en évidence la structure d'un précurseur oligosaccharidique tronqué d'un glucose en comparaison de la structure du précurseur oligosaccharidique retrouvée chez la majorité des eucaryotes. Ces données structurales confirment donc chez *P. tricornutum* l'absence de l'activité ALG10, enzyme permettant le transfert du troisième résidu glucose sur le précurseur oligosaccharidique. Elles sont également cohérentes avec l'absence de l'activité α -glucosidase I éliminant ce même résidu glucose lors du processus de contrôle qualité (**Figure 7**). Ces résultats confirment les prédictions bioinformatiques de Baïet et ses collaborateurs en 2011.

B. Analyse structurale détaillée des *N*-glycannes oligomannosidiques chez *P. tricornutum*

Le second objectif de la thèse était de déterminer les structures détaillées des *N*-glycannes oligomannosidiques depuis le $\text{Man}_9\text{GlcNAc}_2$ synthétisé dans le RE jusqu'au $\text{Man}_5\text{GlcNAc}_2$ synthétisé dans l'appareil de Golgi, structure qui est le substrat de la GnT I, transférase initiant la synthèse des *N*-glycannes complexes (**Aebi, 2013**). Ces étapes de maturation sont orchestrées par la Man I. Les différentes analyses réalisées en utilisant plusieurs approches par spectrométrie de masse ont permis de conclure que *P. tricornutum* possède un isomère de $\text{Man}_9\text{GlcNAc}_2$ et un isomère de $\text{Man}_5\text{GlcNAc}_2$ qui sont uniques, ainsi qu'un mélange d'isomères de $\text{Man}_8\text{GlcNAc}_2$, $\text{Man}_7\text{GlcNAc}_2$ et $\text{Man}_6\text{GlcNAc}_2$. Il peut donc être conclu que la voie de synthèse des *N*-glycannes oligomannosidiques chez *P. tricornutum* allant de $\text{Man}_9\text{GlcNAc}_2$ à $\text{Man}_5\text{GlcNAc}_2$ est identique à celle décrite chez les mammifères, les insectes et les plantes. Ces analyses structurales permettent également de confirmer que le $\text{Man}_5\text{GlcNAc}_2$ produit est l'isomère canonique, substrat de la GnT I. A ce stade, les résultats des deux premiers chapitres démontrent donc que les étapes de la voie de la *N*-glycosylation des protéines de *P. tricornutum* sont jusqu'à la formation du $\text{Man}_5\text{GlcNAc}_2$ remarquablement analogues à celles décrites chez les mammifères (**Figure 37**).

conséquence, le dernier objectif de la thèse a consisté à surexprimer la GnT I endogène par l'insertion dans le génome de la diatomée du gène codant cette transférase portant une étiquette V5, sous le contrôle d'un promoteur inductible au nitrate. La transformation génétique a permis d'obtenir 19 clones surexprimant la transférase. Cependant, les différentes analyses réalisées montrent que cette surexpression n'induit aucune modification importante des profils de la *N*-glycosylation chez les transformants.

Discussion perspective

A. Le cycle de contrôle qualité de la diatomée

L'absence de transfert du troisième résidu glucose sur le précurseur oligosaccharidique a été décrit comme induisant un impact sur la *N*-glycosylation des glycoprotéines chez *A. thaliana*. En effet, l'analyse glycanique par Western blot des protéines totales chez un mutant *alg10* a montré une diminution importante des protéines portant les épitopes core $\alpha(1,3)$ -fucose et core $\beta(1,2)$ -xylose, indiquant une diminution de la maturation des *N*-glycannes complexes présents sur les glycoprotéines chez ce mutant (**Farid et al., 2011**). Au niveau physiologique, cette mutation provoque une diminution de la taille des feuilles de plantules d'*A. thaliana* âgées de 6 semaines. De plus, une sensibilité accrue aux stress abiotiques, notamment au stress salin, a pu être observée. En effet, une diminution de la croissance racinaire a pu être observée après avoir placé les plantules dans un milieu Murashige et Skoog (MS) de culture riche en NaCl (**Farid et al., 2011**). Cet effet est d'autant plus important que la concentration en sel augmente (**Farid et al., 2011**). Chez l'Homme, l'absence de l'ALG10 ne provoque pas de pathologie connue (**Thiel, 2003**). Cette étape n'a pas été conservée au cours de l'évolution chez *P. tricornutum*. Des analyses bio-informatiques sur plusieurs autres microalgues ont permis de mettre en évidence que, comme pour *P. tricornutum*, le gène *ALG10* et celui codant l' α -glucosidase I étaient également absents dans le génome de la majorité des diatomées (**Levy-Ontman et al., 2014**). Ces absences suggèrent que les diatomées ont, au cours de l'évolution, perdu ces gènes ayant pour conséquence la synthèse d'un précurseur oligosaccharidique tronqué d'un glucose (**Aebi, 2013 ; Levy-Ontman et al., 2014**).

La synthèse d'un tel précurseur tronqué amène également à se poser la question de la structure de l'OST de la diatomée, complexe protéique transférant ce précurseur oligosaccharidique sur une protéine en cours de synthèse. Comme expliqué lors de l'introduction de ce manuscrit, l'OST humaine reconnaît un oligosaccharide précurseur de type $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$. Après une analyse bio-informatique basée sur les homologues de séquences, les gènes codant les sous-unités impliquées dans l'organisation de l'OST humaine, semblent absents du génome de la diatomée, excepté celui de la sous-unité catalytique STT3. Un alignement réalisé entre la séquence du gène de la sous-unité STT3a humaine et le génome de la diatomée, met en évidence deux gènes putatifs, *Phatr3_J55197* et *Phatr3_J55198*, codant une sous-unité STT3. Les protéines traduites théoriquement à partir de ces gènes possèdent respectivement 31,6 et 31,2% d'identité avec la protéine STT3A humaine, d'après l'alignement réalisé sur les protéines entières. La même approche d'alignement entre la protéine STT3B humaine et le protéome de la diatomée, apporte un résultat similaire avec une identité de 31,4%. Tous les domaines protéiques essentiels à l'activité enzymatique des sous-unités STT3A et B humaines sont conservés sur les séquences protéiques *Phatr3_J55197* et *Phatr3_J55198* de la diatomée, excepté le motif DK qui tout comme le motif WWDYG est impliqué dans la formation de la poche catalytique. Parmi les motifs conservés, on retrouve les deux motifs DxD impliqués dans l'interaction avec l'ion divalent ainsi que le motif WWDYG. Le motif SVSE, impliqué dans l'activité catalytique de l'OST, est présent dans les séquences protéiques de *Phatr3_J55197* et *Phatr3_J55198*, mais possède une mutation qui substitue le troisième acide aminé du motif qui est une sérine (S) en une alanine (A).

Ces informations suggèrent que l'OST pourrait être organisée différemment chez la diatomée que chez l'Homme. Une organisation différente a déjà pu être observée pour l'OST de *Trypanosoma brucei* capable de synthétiser un précurseur oligosaccharidique tronqué de type $\text{Man}_5\text{GlcNAc}_2$ ou $\text{Man}_9\text{GlcNAc}_2$. En effet, les analyses bio-informatiques et biochimiques ont montré que cette espèce possède exclusivement les 3 gènes codant la sous-unité catalytique SST3. Ces trois gènes sont appelés TbSTT3a, b et c (**Kellegher and Gilmore, 2006 ; Izquierdo et al., 2009**). Tout comme chez *P. tricornutum*, les autres sous-unités, impliquées dans l'organisation de l'OST chez les mammifères, sont absentes du génome du parasite (**Kellegher and Gilmore, 2006**). Des travaux ont pu montrer que TbSST3a reconnaissait spécifiquement un précurseur oligosaccharidique de type $\text{Man}_5\text{GlcNAc}_2$ (**Izquierdo et al., 2009**). La sous-unité TbSST3b privilégie quant à elle, le transfert d'un précurseur

oligosaccharidique de type $\text{Man}_9\text{GlcNAc}_2$ (**Izquierdo et al., 2009**). Il est souhaitable à l'avenir de caractériser les deux gènes putatifs de *P. tricornutum* codant les sous-unités STT3 et de réaliser une analyse structurale de l'OST de *P. tricornutum*. Ces données permettraient de comprendre comment l'OST de la diatomée reconnaît et procède au transfert du précurseur oligosaccharidique sur une protéine en cours de synthèse.

B. Hypothèses expliquant pourquoi l'activité GnT I ne conduit pas à l'accumulation de *N*-glycannes complexes chez la diatomée ?

Au terme de ce projet de thèse, la principale question qui se pose est : pourquoi l'action de la GnT I, endogène ou surexprimée, ne conduit-elle pas à l'accumulation de *N*-glycannes complexes chez *P. tricornutum* ? Initialement, l'hypothèse émise était que le $\text{Man}_5\text{GlcNAc}_2$ n'était pas substrat accepteur de la GnT I. En effet, des données biochimiques obtenues récemment chez deux microalgues vertes démontrent que ces organismes peuvent synthétiser un glycanne de type $\text{Man}_5\text{GlcNAc}_2$ linéaire et très différent de celui synthétisé par la majorité des eucaryotes (**Vanier et al., 2017 ; Mócsai et al., 2019 ; Lucas et al., 2020**). Les données biochimiques obtenues lors de cette thèse ont validé que cette diatomée synthétise un isomère de $\text{Man}_5\text{GlcNAc}_2$ dit « canonique ». Cet isomère est substrat accepteur de la GnT I, éliminant donc l'hypothèse que la diatomée synthétiserait un $\text{Man}_5\text{GlcNAc}_2$ différent, substrat non reconnu par l'enzyme. La GnT I est par ailleurs fonctionnelle. Il a été démontré qu'elle était capable de compléter un mutant CHO déficient en activité GnT I endogène (**Baïet et al., 2011**). Une première hypothèse expliquant l'absence de GlcNAc en position terminale des *N*-glycannes pourrait être la non-disponibilité du substrat donneur à savoir le nucléotide sucre de type UDP-GlcNAc dans l'appareil de Golgi. *P. tricornutum* possède la machinerie enzymatique pour la synthèse de l'UDP-GlcNAc dans le cytosol (**Gügi et al., 2015**). Comme preuve, la synthèse cytosolique du précurseur oligosaccharidique au début de la voie de la *N*-glycosylation commence par le transfert dans le cytosol par les ALG 7, 13 et 14 de deux résidus GlcNAc à partir d'UDP-GlcNAc sur le dolichol, formant ainsi le chitobiose du *N*-glycanne. Cependant, il est aujourd'hui impossible de savoir si ce nucléotide-sucre est transféré et donc disponible dans l'appareil de Golgi. En effet, peu de données bibliographiques sur les transporteurs nucléotidiques sont disponibles chez la diatomée.

Récemment, des analyses bio-informatiques réalisées sur le génome de la diatomée ont permis de mettre en évidence 3 gènes putatifs codant des transporteurs du GDP-fucose (**Zhang et al., 2019**). Parmi ces 3 gènes, *Phatr3J_43174* semble coder une protéine apparentée aux transporteurs de GDP-fucose humains et de *C. elegans* (**Zhang et al., 2019**). Les deux autres gènes semblent coder des transporteurs plus apparentés à ceux retrouvés chez *A. thaliana* (**Zhang et al., 2019**). De plus, comme expliqué dans l'introduction de ce manuscrit de thèse, la protéine codée par le gène *Phatr3J_43174* a pu être caractérisée biochimiquement par complémentation dans des cellules CHO dépourvues du transporteur du GDP-fucose (**Zhang et al., 2019**). Cette protéine, localisée dans les membranes golgiennes, a permis de restaurer la fucosylation des *N*-glycannes dans l'appareil de Golgi chez le mutant CHO (**Zhang et al., 2019**). Ces données confirment donc que le mécanisme de transfert de ces nucléotides-sucres dans l'appareil de Golgi de la diatomée est semblable à celui retrouvé chez les eucaryotes. En effet, les nucléotides-sucres sont incapables de passer la membrane golgienne sans la présence d'un transporteur spécifique (**Caffaro and Hirschberg, 2006 ; Liu et al., 2010**).

Aucune information prouvant l'existence d'un transporteur golgien spécifique d'UDP-GlcNAc n'est disponible chez *P. tricornutum*. L'absence d'un tel transporteur pourrait empêcher le transfert d'UDP-GlcNAc depuis le cytosol vers l'appareil de Golgi et expliquer l'absence de résidus GlcNAc en position terminale sur les *N*-glycannes complexes de la diatomée. Dans les perspectives de ce projet de thèse, le transport effectif de l'UDP-GlcNAc du cytosol vers la lumière du Golgi réalisé chez la diatomée devient donc primordial pour permettre le transfert golgien des résidus GlcNAc par la GnT I sur le Man₅GlcNAc₂ « canonique ». Des analyses bio-informatiques du génome de la diatomée permettent de mettre en évidence 5 candidats codant des transporteurs potentiels d'UDP-GlcNAc, avec plus

de 30% d'homologies, alignement réalisé sur la protéine entière du transporteur humain **(Figure 38)**.

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
✓	predicted protein [Phaeodactylum tricornutum CCAP 1055/1]	281	281	90%	5e-94	49.36%	XP_002179931.1
✓	predicted protein [Phaeodactylum tricornutum CCAP 1055/1]	188	188	97%	8e-58	37.69%	XP_002183716.1
✓	predicted protein [Phaeodactylum tricornutum CCAP 1055/1]	138	138	92%	4e-38	34.94%	XP_002183461.1
✓	predicted protein [Phaeodactylum tricornutum CCAP 1055/1]	115	115	97%	1e-29	32.10%	XP_002180456.1
✓	predicted protein [Phaeodactylum tricornutum CCAP 1055/1]	75.5	75.5	60%	6e-16	30.29%	XP_002186260.1

Figure 38 : Liste des transporteurs putatifs de l'UDP-GlcNAc obtenue par la recherche dans le génome de *P. tricornutum* de séquences homologues du gène de l'isoforme 1 du transporteur golgien de l'UDP-GlcNAc humain (Uniprot : Q9Y2D2)

Le logiciel utilisé pour l'alignement des séquences est BlastP (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Ces données bio-informatiques ne sont pas une preuve définitive de l'existence d'un transport efficace d'UDP-GlcNAc entre le cytosol et l'appareil de Golgi de la diatomée. Cependant, la présence chez *P. tricornutum* de l'UDP-GlcNAc cytosolique, d'une GnT I fonctionnelle et de la synthèse de son substrat accepteur nous permet de supposer qu'un (ou des) transporteur(s) soi(en)t localisé(s) dans les membranes golgiennes et importe(nt) l'UDP-GlcNAc nécessaire à la machinerie de la voie de la *N*-glycosylation des protéines, sans pouvoir exclure que ce transport soit limitant. Afin de vérifier cette hypothèse, il pourrait être, à l'avenir, utile de caractériser les gènes identifiés codant un transporteur potentiel d'UDP-GlcNAc. Pour réaliser cette caractérisation, il serait nécessaire de compléter avec les gènes putatifs de *P. tricornutum* une lignée cellulaire déficiente en activité de transporteur d'UDP-GlcNAc. La génération d'une souche CHO déficient en transport d'UDP-GlcNAc dans l'appareil de Golgi a déjà été réalisée par l'approche RNAi en ciblant spécifiquement le gène *SLC35A3*. Cette approche permet d'inhiber l'expression du gène *SLC35A3* codant pour un transporteur de l'UDP-GlcNAc chez CHO (**Maszczak-Seneczko et al., 2013**). Les résultats obtenus par analyse glycanique de ces cellules CHO modifiées ont montré une légère baisse de la proportion de *N*-glycannes complexes portant du GlcNAc en position terminale, reflétant la baisse du transport d'UDP-GlcNAc dans l'appareil de Golgi de la souche modifiée (**Maszczak-Seneczko et al., 2013**). Ces données suggéreraient que d'autres gènes sont impliqués dans le transport de l'UDP-GlcNAc dans l'appareil de Golgi des cellules CHO (**Maszczak-Seneczko et al., 2013**). La génération d'un mutant dépourvu de toute activité de transport d'UDP-GlcNAc dans

l'appareil de Golgi pourrait être réalisée dans des cellules CHO en utilisant l'approche CRISPR-cas9 (**Sergeeva et al., 2019**). La complémentation dans ces cellules dépourvues de transporteur de l'UDP-GlcNAc par l'expression des gènes de *P. tricornutum*, permettrait de déterminer lesquels de ces gènes codent un transporteur de l'UDP-GlcNAc et sa localisation subcellulaire, en particulier de déterminer si le transporteur est localisé ou non dans les membranes golgiennes.

Une seconde hypothèse peut être proposée pour expliquer l'absence apparente d'action de la GnT I sur les *N*-glycannes des protéines synthétisées par la diatomée. L'action de la GnT I et/ou des enzymes golgiennes nécessaires à la maturation des *N*-glycannes complexes pourrait être dépendante des conditions physiologiques. Par exemple, un pH de 6,8 au lieu de 7,4 du milieu de culture des cellules de mammifères peut réguler l'activité de la $\beta(1,4)$ GalT en diminuant la proportion de *N*-glycannes portant des résidus galactose en position terminale de l'anticorps monoclonal R24 produits (**Müthing et al., 2003 ; Hossler et al., 2009**). Également, Une régulation des GnT III et V impliquée dans la synthèse des *N*-glycannes polyantennés a notamment été décrite chez les mammifères dans certaines pathologies (**Ko et al., 1999 ; Taniguchi et al., 1999**). Chez *P. tricornutum*, lors de stress ou de conditions de cultures particulières, les cellules sont capables de changer de morphotype en évoluant des morphotypes fusiforme et triradié vers le morphotype ovale (**De Martino et al., 2007 ; Ovide et al., 2018 ; Song et al., 2020**), et/ou d'adapter leur métabolisme aux conditions environnementales (**Matthijs et al., 2016 ; Remmers et al., 2018 ; Zhang et al., 2020**). Les cultures Pt 1.8.6 utilisées dans ce travail de thèse sont composées majoritairement de cellules sous le morphotype fusiforme (**De Martino et al., 2007 ; Ovide et al., 2018**). Les analyses transcriptomiques menées sur les différents morphotypes de *P. tricornutum* n'ont pas montré de différences flagrantes dans les niveaux de transcription des gènes codant les glycosyltransférases golgiennes, et en premier lieu la GnT I (**Ovide et al., 2018**). Seuls les gènes *Phatr3_J49563* et *Phatr3_J45073* codant les hexosaminidases présentent un niveau de transcription plus faible chez le morphotype ovale comparé au morphotype fusiforme. Des conclusions analogues émanent des analyses des transcriptomes de la diatomée soumise à des carences en azote (**Matthijs et al., 2016 ; Remmers et al., 2018**) ou à une forte pression en CO₂ (**Zhang et al., 2020**). Dans ces conditions, l'expression des gènes codant les acteurs golgiens de la voie de maturation des *N*-glycannes n'est que peu affectée. Des analyses des profils de glycosylation des protéines synthétisées dans ces différentes conditions permettrait

cependant de mieux évaluer l'impact des conditions environnementales sur la maturation des *N*-glycannes.

Parmi les autres hypothèses, on peut supposer que la maturation des *N*-glycannes est restreinte à des populations particulières de glycoprotéines, notamment les protéines sécrétées dans le milieu de culture. Il serait intéressant à l'avenir d'analyser la glycosylation des protéines sécrétées dans le milieu de culture ou d'une protéine spécifique utilisée comme rapportrice et exprimée de manière sécrétée ou non. Pour cela, la transformation de *P. tricornutum* par le gène codant l'avidine permettrait de disposer d'une glycoprotéine facile à purifier par affinité et dont le profil de glycosylation permettrait d'évaluer l'impact des conditions environnementales et/ou de son lieu de stockage en fonction de son adressage dans la diatomée.

En complément, la production d'un anticorps monoclonal comme l'anticorps anti-Hépatite B (**Vanier et al., 2015**) par les clones surexprimant la GnT I permettrait d'évaluer l'impact de cette surexpression sur les *N*-glycannes de cet anticorps monoclonal qui est glycosylé avec des *N*-glycannes de type Man₉GlcNAc₂ à Man₅GlcNAc₂ quand il est produit dans des cellules de *P. tricornutum* sauvage (**Vanier et al., 2015**).

C. Ingénierie corrective des étapes golgiennes de maturation des *N*-glycannes de *P. tricornutum* en vue de leur humanisation

Les données biochimiques obtenues lors de cette thèse, accompagnées de nouvelles données bibliographiques sur les glycosyltransférases golgiennes, nous permettent de dresser un état des lieux plus rigoureux des étapes golgiennes de la *N*-glycosylation des protéines de *P. tricornutum* et ainsi de mieux définir en perspective les stratégies d'inactivation ou de complémentation de gènes pour permettre à plus longue échéance la synthèse d'anticorps portant une glycosylation « humanisée » chez cette microalgue. Les paragraphes suivants réactualisent les connaissances concernant les enzymes golgiennes de la maturation des *N*-glycannes de *P. tricornutum* et les stratégies d'ingénierie nécessaires qui en découlent.

1. Les fucosyltransférases

Les FuT transfèrent des résidus fucose sur les *N*-glycannes en cours de maturation dans l'appareil de Golgi. Au sein du génome de la diatomée, 3 gènes codant des FuT putatives ont été identifiés : *Phatr3_J46109*, *Phatr3_J46110* et *Phatr3_J54599*. Ces FuTs sont regroupées au sein de la famille CAZy GT10 des glycosyltransférases ([Site internet : Cazy GT10](#) ; **Wilson et al., 2001** ; **Both et al., 2011** ; **Zhang et al., 2019**). Des analyses bio-informatiques ont montré que ces 3 gènes ont conservé les motifs d'acides aminés essentiels à l'activité enzymatique des fucosyltransférases (**Zhang et al., 2019**). De plus, toute la machinerie enzymatique pour la synthèse de GDP-fucose est disponible chez la diatomée (**Gügi et al., 2015**). La souche *CHO-gmt5* est un mutant dépourvu du transporteur GDP-fucose, empêchant la fucosylation des protéines (**Haryadi et al., 2013**). La complémentation du gène *Phatr3_J43174* de *P. tricornutum* dans ce mutant *CHO-gmt5* a permis de restaurer la fucosylation des *N*-glycannes portés par les protéines. Ces travaux ont permis de mettre en évidence que le gène *Phatr3_J43174* code un transporteur du GDP-fucose (**Zhang et al., 2019**). De plus, les analyses par microscopie confocale, ont montré que ce transporteur du GDP-fucose était adressé dans les membranes golgiennes de la souche *CHO-gmt5*, suggérant donc que cette protéine est vraisemblablement adressée dans les membranes golgiennes également chez la diatomée (**Zhang et al., 2019**). Les prédictions bio-informatiques des fucosyltransférases putatives *Phatr3_J46109* et *Phatr3_J46110* indiquent que ces gènes codent respectivement des protéines de 798 acides aminés et de 707 acides aminés. Ces deux gènes codent des protéines étant de 25 à 30% plus grandes que les fucosyltransférases caractérisées chez d'autres espèces eucaryotes (**Ovide, 2016**). *Phatr3_J46109* ne posséderait pas de domaine transmembranaire et *Phatr3_J46110* posséderait un domaine transmembranaire sur la partie *N*-terminale de la protéine (**Ovide, 2016**). Les analyses par microscopie confocale effectuées à la suite de la fusion de ces deux protéines avec une protéine fluorescente GFP, montrent qu'elles semblent être adressées dans l'appareil de Golgi, prouvant la présence d'un domaine transmembranaire nécessaire pour leur bon adressage golgien (**Ovide, 2016**). En accord, les analyses des séquences protéiques montrent la présence d'un domaine GT10, semblable à celui retrouvé dans la séquence protéique des fucosyltransférases issues de plantes (**Ovide, 2016**). Aucune donnée biochimique démontrant l'activité de transfert du fucose de ces deux protéines n'est disponible actuellement.

Le troisième gène, *Phatr3_J54599* code une protéine d'une taille de 481 acides aminés, taille similaire aux fucosyltransférases retrouvées chez les plantes. Il possède en plus du domaine, GT10, un domaine GT10_N présent sur les fucosyltransférases des invertébrés (Ovide, 2016). Ce gène a été récemment caractérisé comme exprimant une activité $\alpha(1,3)$ -fucosyltransférase fonctionnelle et localisée dans l'appareil de Golgi (Zhang *et al.*, 2019). Tous ces résultats sont cohérents avec la présence de *N*-glycannes paucimannosidiques de type Man₃FucGlcNAc₂ détectés par Baïet *et al.*, 2011 sur les glycoprotéines endogènes chez *P. tricornutum*.

2. L' α -mannosidase II et la *N*-acétylglucosaminyltransférase II

La recherche des motifs caractéristiques des domaines catalytiques des enzymes Man II et GnT II au sein des protéines de *P. tricornutum* par BLAST P montre qu'ils sont présents au sein d'une protéine unique de type II constituée de 1453 acides aminés (Phatr3_52248). Des protéines présentant une organisation analogue sont prédites chez d'autres diatomées comme *Fragilariopsis cylindrus* CCMP1102 : OEU08391.1, 1543 acides aminés et *Pseudonitzschia multistriata* : VEU35400.1, 1367 acides aminés (Figure 39).



Figure 39 : Structure schématique de la protéine de fusion Man II/GnT II.

Ce schéma, représentant la protéine de fusion Man II/GnT II, est composé d'une séquence homologue de la Man II du côté *N*-terminal et d'une séquence homologue de la GnT II du côté *C*-terminal.

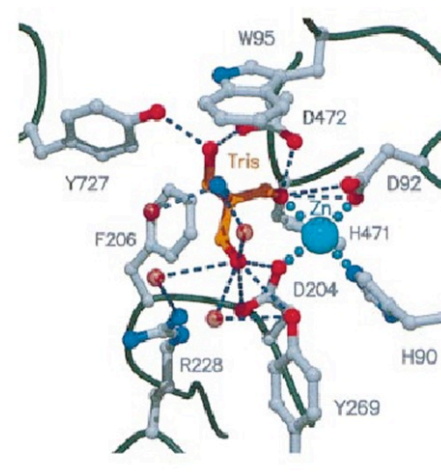
TMD : domaine transmembranaire ; Man II : α -mannosidase II ; GnT II : *N*-acétylglucosaminyltransférase II

a. Le domaine Man II

Le domaine en *N*-terminal (entre les acides aminés 200 et 900) de la séquence Phatr3_52248 de *P. tricornutum* présente une forte homologie (49% d'identité) avec la Man II de *Drosophila melanogaster* dont la structure tridimensionnelle a été déterminée par diffraction aux rayons X (van den Elsen *et al.*, 2001). Les Man II sont regroupées dans la famille CAZy GH38 des glycosylhydrolases et ont la particularité d'hydrolyser successivement les deux

résidus mannose dans le même site catalytique ([Site internet : Cazy GH38](#) ; [Shah et al., 2008](#)). Le site catalytique de la Man II de *D. melanogaster* est caractérisé par la présence d'un ion Zn^{2+} pris en charge dans une structure pyramidale impliquant les résidus H₉₀, D₉₂, D₂₀₄, D₄₇₂ and H₄₇₁ (**Figure 40A**). L'acide aspartique D₂₀₄ est directement impliqué dans l'hydrolyse des résidus mannose (**Figure 40A**). Tous ces acides aminés essentiels sont retrouvés au sein des motifs conservés dans la séquence de *P. tricornutum* (**Figure 40B**). De plus, les analyses biochimiques des *N*-glycannes des protéines de la diatomée réalisées montrent la synthèse de structures *N*-glycanniques de type Man₃FucGlcNAc₂ prouvant indirectement que cette Man II est active ([Baïet et al., 2011](#)).

A



B

Phaeo 52248	196	IDMRQLVRTLSFDNPDGGVWVKQGWNVNTARDDFTNAQP	LQVFVPHSH CD PGW IKTFFDY
CAA10755.1	42	VQMLELYDRMSFKDIDGGVWVKQGWNIKYDPLKYNAAHKK	LKVFVPHSH ₉₀ ND ₉₂ PGW ₉₅ IQTFFDY
Phaeo 52248	255	FQSQTRQILTTVVQALQRDQRRKFIWAEISYFEWVWYREQDDTRKVVQTLDDNKQLQFVT	
CAA10755.1	104	YQHDTKHILSNALRHLDNPEMKFIWAEISYFARFYHDLGENKKLQMKSIK NGQLEFVT	
Phaeo 52248	315	GGWVQPDEANSELYAMEIQLQEGHDWIRQTFGDAHV PKHG	WSID PF GYSPTMAYLLQKYG
CAA10755.1	164	GGWVMPDEANSHWRNVLLQLTEGQTWIKQFMNVT— PTAS	WAID ₂₀₄ PF ₂₀₆ GHSPMPYILQKSG
Phaeo 52248	375	FKAMLIQRVHYAVKKELAQRRHLEFYWRQTWEDATTAGTHDI	FTHVMPF FSYDVPHTCGP
CAA10755.1	221	FKNMLIQRTHYSVKKELAQRRHVEFLWRQIWDN-- - KGDTA	FTHMMPF ₂₆₉ YSYDIPHTCGP
Phaeo 52248	435	DPSVCCQDFRQRT----TCPWLKAPQTITSKNVAERAMLLLDQYRKKAALYRS NVVLAP	
CAA10755.1	278	DPKVCQDFRKRMGSGFLSCPWKVPPRTISDQNVAAERSDLLVDQWKKKAEYRNTVVLLP	
Phaeo 52248	490	LGDDFRYLTAQEAEAYTNYQAIFDYVNNANLP-NVKMQFGT LSD YF DAV----- VGSFDT	
CAA10755.1	338	LGDDFRFKQNTTEWDVQRVNYERLFEHINSQAHFNVQAQFGTLQEYF DAVHQERAGQAEF	
Phaeo 52248	544	PILQGSFFTYSDIDQD YWSGYF TSRVSDKALGRWLERVLYSATQMGA-----SKQD	
CAA10755.1	398	PTLSGDDFFTYADRSNDYWSGYTSPYHKRMDRVLMMHYVRAAEMLSAWHSWDGMARIEER	
Phaeo 52248	595	LQAPRRALSFLQHH D GVTGTAKTH	VHEDYARQMMDAIHTTDEWMLRAIQ
CAA10755.1	457	LEQARRELSFLQH H ₄₇₁ D ₄₇₂ GITGTAKTHV	VVDYEQR MQ EALKACQ MVMQQSVYR

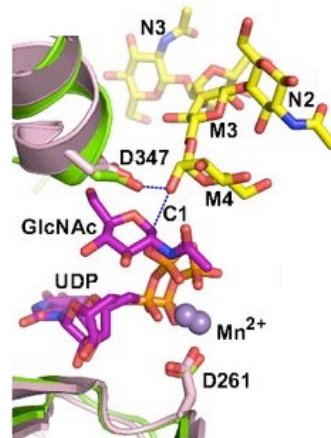
Figure 40 : Analyse comparée des séquences protéiques correspondant aux sites catalytiques de la Man II de drosophile et de *P. tricornutum*

(A) Structure du site catalytique de la Man II de *D. melanogaster* déterminée par diffraction aux rayons X d'après **van den Elsen *et al.*, 2001**. (B) Alignements entre la séquence N-terminale de la protéine Phatr3_J52248 de *P. tricornutum* (Phaeo 52248) et la séquence de la Man II de *D. melanogaster* (CAA10755.1). Les acides aminés impliqués dans le site catalytique sont encadrés en noir et les acides aminés essentiels du site catalytique de la Man II sont surlignés en jaune.

b. Le domaine GnT II

Les GnT II sont regroupées dans la famille CAZy GT16 des glycosyltransférases ([Site internet : Cazy GT16](#)). Sur la base des alignements des séquences protéiques, la présence dans la séquence de Phatr3_52248 d'un motif conservé avec des GnT II eucaryotes avait été reportée par **Mathieu-Rivet *et al.*, 2014**. La publication récente de la structure tridimensionnelle de la GnT II humaine permet de ré-évaluer cette hypothèse avec plus de certitude. Les 350 derniers amino acides en position C-terminale de la séquence de Phatr3_52248 présentent une identité de 28 % avec la GnT II humaine (**Kadirvelraj *et al.*, 2018**). La comparaison des deux séquences montre que le motif E₂₅₉ED₂₆₁ est conservé (**Figure 41A**). Ce motif, ainsi que l'histidine H₃₇₄, permettent la fixation d'un ion Mn²⁺ impliqué dans l'interaction avec les résidus phosphate de l'UDP-GlcNAc. Au sein du site catalytique de la GnT II humaine, l'acide aspartique 347 (D₃₄₇) est positionné de façon à pouvoir déprotoner la fonction hydroxyle nucléophile en C6 de l' α (1,6)-mannose du core comme attendu pour un mécanisme catalytique d'inversion de configuration (formation d'une liaison β entre le GlcNAc et le mannose) (**Figure 41A ; Kadirvelraj *et al.*, 2018**). Le résidu β -Man du core N-glycannique est maintenu par des liaisons hydrogène formées avec les résidus Y344 and N345 du motif D₃₄₂DYNWDW₃₄₈. De même, le résidu GlcNAc est stabilisé par des liaisons hydrogène avec les chaînes latérales des résidus W346 et E259 (**Figure 41B ; Kadirvelraj *et al.*, 2018**). Ces motifs responsables de la reconnaissance des substrats oligosaccharidiques et du mécanisme enzymatique sont également conservés dans la séquence candidate de la diatomée (**Figure 41A**).

A



B

Phaeo 52248	1194	LIHPHSCYE	PNEFP GDDP	-----	-TLNEGFAG----	DSYGNPRSAWITCC	KHHF
Homo sapiens	175	VFFPFSIQLY	PNEFP GSDP	RDCPRDLPKNAALK	LG CINA EY PDSFGHYREAKFSQT	KHHW	
Phaeo 52248	1240	TWMLHTVFR	DFTDPAVDTF	LFLE ED	YIVAP	TIYSAVIAGLNVMEDMDKEIPGGFFGLG	
Homo sapiens	235	WWKLHFVWERV	KILRDYAGLI	LFLE₂₅₉ED₂₆₁	HYLAF	DFYHVFKKMWKLKQ---QECPECDV-LS	
Phaeo 52248	1298	MDPSMANTAF	FEPYKKA---	TWYVEAFKSGP	M T MNRDMFKKLQ	HAKEY	CTFDDY N WD WS
Homo sapiens	291	LGTYSASRSFY	GMA DKVDVKT	WKST EHNMG	-LALTRNAYQ	KLIE CTDTF	CTYDDY₃₄₄N₃₄₅WD₃₄₇WT
Phaeo 52248	1355	IVHLQSKLLPR	---TLLMPSKVLAK	H	IGVKE GMH	TNKSFGKDFSDLFPNYAPRREQLTTH	
Homo sapiens	350	LQYL-TVSCLPK	FWKVLVPQIPR	F	H₃₇₄AG-DCGMH	HKKTCRPS-----TQSAQIESLLNN	
Phaeo 52248	1413	FAEYRFTGNTAAALH	---TEFNP---	GY	GGWGHPKDHEHC	1453	
Homo sapiens	402	NKQYMFPELTISEK	FTVVAISPPRKN		GGWGDIRDH ELC	440	

Figure 41 : Analyse comparée des séquences protéiques correspondant aux sites catalytiques de la GnT II humaine et de *P. tricornutum*

(A) Structure du site catalytique de la GnT II humaine déterminée par diffraction aux rayons X d'après [Kadirvelraj et al., 2018](#). (B) Alignements entre la séquence C-terminale de la protéine Phatr3_J52248 de *P. tricornutum* (Phaeo 52248) et la GnT II humaine (*Homo sapiens*). Les acides aminés impliqués dans le site catalytique sont encadrés en noir et les acides aminés essentiels du site catalytique de la GnT II sont surlignés en jaune.

3. Les hexosaminidases

Les hexosaminidases sont capables d'éliminer des résidus β -N-acétylglucosamine terminaux des N-glycannes dans le système de sécrétion ou sur le lieu de stockage de la glycoprotéine. Elles ont été décrites chez les plantes ([Lerouge et al., 1998](#) ; [Strasser et al., 2007](#)) et les cellules d'insectes ([Altmann et al., 1995](#) ; [Geisler et al., 2008](#)). Ces enzymes sont capables, dans ces organismes, de cliver les résidus GlcNAc présents en position terminale des

N-glycannes, donnant ainsi naissance à des *N*-glycannes paucimannosidiques. Elles appartiennent à la famille CAZy GH20 des glycosylhydrolases ([Site internet : Cazy GH20](#)). Deux séquences de gènes de *P. tricornutum* présentent de fortes homologues avec les gènes des β -hexosaminidases d'*A. thaliana* (*At3G55260*, *At1G65590* et *At1G05590*). Les séquences des protéines associées, Phatr3_J45073 et Phatr3_J49563, ont des longueurs respectives de 1085 et 973 acides aminés.

En collaboration avec le Dr Richard Strasser (Université of BOKU, Autriche), le laboratoire a initié une caractérisation des hexosaminidases de *P. tricornutum*. Les données préliminaires montrent que la protéine Phatr3_J45073 est détectée dans l'appareil de Golgi à la suite de son expression transitoire chez *A. thaliana* et dans des feuilles de tabac. De plus, Phatr3_J45073 est capable de compléter un mutant d'*A. thaliana* déficient en activités hexosaminidases (**Liebmingier et al., 2011**). Ces travaux suggèrent, dans l'état des données actuelles, qu'au moins une des deux protéines candidates est une β -hexosaminidase impliquée dans la dégradation des *N*-glycannes complexes. Ces travaux préliminaires devront être confirmés, notamment en évaluant l'impact de l'inactivation des deux protéines candidates sur la structure des *N*-glycannes synthétisés chez *P. tricornutum*.

L'action de ces hexosaminidases pourrait être également réduite considérablement par l'insertion d'un gène codant la $\beta(1,4)$ -galactosyltransférase humaine ($\beta(1,4)$ GalT) dans le génome de *P. tricornutum*. En effet, cette enzyme est capable de transférer des résidus galactose sur des *N*-glycannes complexes portant des résidus GlcNAc en position terminale ce qui pourrait bloquer l'action des hexosaminidases. Des analyses bio-informatiques et biochimiques ont montré que *P. tricornutum* synthétise de l'UDP-galactose, le nucléotide-sucre utilisé par les GalT (**Gügi et al., 2015**). Rappelons que les expressions hétérologues du transporteur du GDP-fucose (**Zhang et al., 2019**) et de la GnT I (**Baïet et al., 2011**) de *P. tricornutum* dans des mutants CHO ont permis de restaurer les activités déficientes de ces protéines dans les cellules de mammifères. Ceci suggère que les mécanismes d'adressage et de rétention des protéines golgiennes sont conservés entre les mammifères et *P. tricornutum*. Ces résultats laissent présager qu'après la transformation par le gène codant la $\beta(1,4)$ GalT humaine, la transférase recombinante sera correctement localisée et active dans l'appareil de Golgi de la microalgue. En conséquence, la transformation par le gène humain codant la $\beta(1,4)$ GalT doit permettre à la diatomée de transférer des résidus galactose sur les résidus *N*-acétylglucosamine localisés en position terminale des *N*-glycannes de ses protéines.

L'expression de la $\beta(1,4)$ GalT humaine dans l'appareil de Golgi de *P. tricornutum* doit non seulement empêcher, si les résidus galactose sont ajoutés suffisamment rapidement, l'élimination des résidus *N*-acétylglucosamine des *N*-glycannes par les hexosaminidases endogènes mais également permettre l'obtention de *N*-glycannes complexes portant des résidus galactose en position terminale, analogues à celles observées chez les anticorps monoclonaux humains.

Cette stratégie de blocage de l'élimination des GlcNAc en position terminale a été initiée par la complémentation, dans le génome de la diatomée, du gène codant la $\beta(1,4)$ GalT humaine. Brièvement, le gène codant la $\beta(1,4)$ GalT a été introduit dans le génome de la diatomée sous le contrôle d'un promoteur inductible par le nitrate et a été fusionné à une séquence codant le peptide antigénique HA répétée 3 fois pour faciliter sa détection. Parmi les 122 colonies obtenues, 72 transformants ont inséré le transgène. Ces 72 clones ont été ensuite criblés par Western blot au moyen d'un anticorps dirigé contre le peptide HA (**Figure 42**). Trois transformants ont montré une bande protéique reconnue par l'anticorps, soit un rendement de 2,45% parmi les clones transformés.

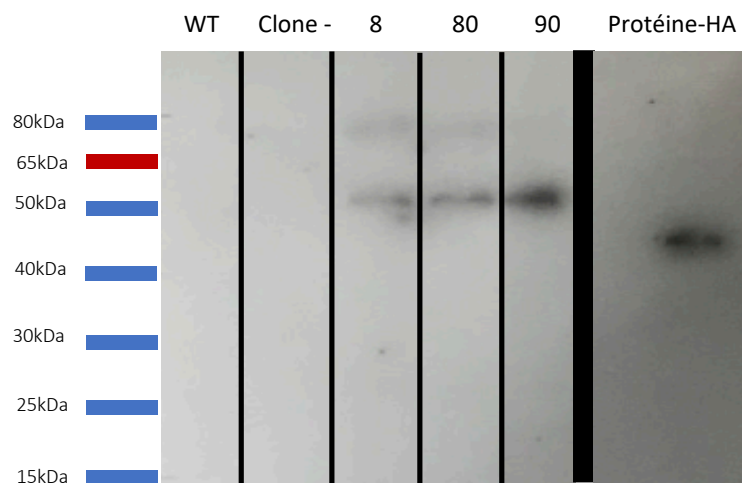


Figure 42 : Criblage des transformants de *P. tricornutum* exprimant le gène humain codant la $\beta(1,4)$ GalT-HA

Immunodétection par des anticorps dirigés contre le peptide HA des protéines microsomales isolées des transformants exprimant la $\beta(1,4)$ GalT-HA. Le témoin positif est une protéine commerciale contenant le peptide HA. Les témoins négatifs utilisés sont les protéines microsomales de la souche sauvage Pt 1.8.6 (WT) et d'un clone transformé par le vecteur pPha-NR dépourvu des gènes d'intérêt mais ayant intégré le gène de la résistance à la zéocine (Clone -).

Le gène codant la $\beta(1,4)$ GalT-HA humaine et le gène codant la GnT I de *P. tricornutum* ont également été introduits simultanément dans le génome de la diatomée sous le contrôle d'un promoteur inductible par le nitrate, permettant ainsi d'induire leur expression par l'ajout de nitrate dans le milieu de culture. La transformation génétique de *P. tricornutum* par biolistique a permis d'obtenir 182 colonies dont 52 transformants ayant intégré les deux gènes. La co-expression de la $\beta(1,4)$ GalT-HA et de la PtGnT I-V5 a ensuite été recherchée chez ces 52 clones par Western blot au moyen d'un anticorps dirigé contre le peptide V5 puis d'un anticorps dirigé contre le peptide HA (**Figure 43**). Au terme de ce criblage, seulement 3 doubles transformants montrent une expression simultanée de la $\beta(1,4)$ GalT-HA et de la PtGnT I-V5, soit un rendement faible de 1,64%.

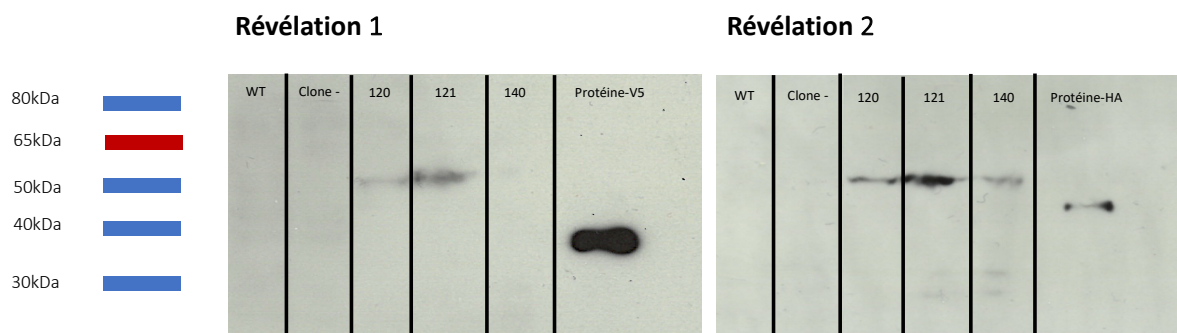


Figure 43 : Criblage protéique des doubles transformants ayant inséré simultanément les gènes codant la $\beta(1,4)$ GalT-HA et la PtGnT I-V5

Immunodétection par des anticorps dirigés contre le peptide V5 (Révélation 1) et contre le peptide HA (Révélation 2) des protéines microsomales isolées à partir des doubles transformants de *P. tricornutum* exprimant les $\beta(1,4)$ GalT-HA et PtGnT I-V5. A la fin de la première révélation, les complexes anticorps primaire/secondaire présents sur le blot ont été éliminés avant de procéder à la seconde révélation. Le témoin positif est une protéine commerciale contenant les peptides HA et V5. Les témoins négatifs utilisés sont les protéines microsomales de la souche sauvage Pt 1.8.6 (WT) et d'un clone transformé avec le vecteur pPha-NR dépourvu des gènes d'intérêt mais ayant inséré le gène de la résistance à la zéocine (Clone -).

Il serait donc nécessaire de continuer l'analyse biochimique de ces simples et doubles transformants afin de vérifier si ces clones permettent d'accumuler des *N*-glycannes complexes portant des résidus galactose en position terminale. Si une telle accumulation est confirmée, une avancée majeure concernant l'humanisation des voies de *N*-glycosylation aura été atteinte. L'humanisation des voies de *N*-glycosylation de la diatomée pourra donc être

prolongée à terme par l'insertion dans son génome des gènes impliqués dans les voies de la biosynthèse des acides sialiques, dans leur transport dans l'appareil de Golgi et dans leur transfert sur les *N*-glycannes. Parallèlement, l'expression d'un anticorps chez les transformants permettrait de vérifier à chaque étape de l'ingénierie métabolique de la voie de *N*-glycosylation, si l'anticorps recombinant porte lui aussi des *N*-glycannes humanisés permettant d'envisager une utilisation à des fins thérapeutiques.

Chapitre 5

Bibliographie

- Abas, L., Luschnig, C., 2010. Maximum yields of microsomal-type membranes from small amounts of plant material without requiring ultracentrifugation. *Anal Biochem* 401, 217–227. <https://doi.org/10.1016/j.ab.2010.02.030>
- Abdullahi, A.S., Underwood, G.J.C., Gretz, M.R., 2006. Extracellular Matrix Assembly in Diatoms (bacillariophyceae). V. Environmental Effects on Polysaccharide Synthesis in the Model Diatom, *Phaeodactylum Tricornutum*. *Journal of Phycology* 42, 363–378. <https://doi.org/10.1111/j.1529-8817.2006.00193.x>
- Aebi, M., 2013. N-linked protein glycosylation in the ER. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1833, 2430–2437. <https://doi.org/10.1016/j.bbamcr.2013.04.001>
- Aeby, A., Prigogine, C., Vilain, C., Malfilatre, G., Jaeken, J., Lederer, D., Bogaert, P.V., 2016. RFT1-congenital disorder of glycosylation (CDG) syndrome: a cause of early-onset severe epilepsy. *Epileptic Disorders* 18, 92–96. <https://doi.org/10.1684/epd.2016.0802>
- Alaimo, C., Catrein, I., Morf, L., Marolda, C.L., Callewaert, N., Valvano, M.A., Feldman, M.F., Aebi, M., 2006. Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *The EMBO Journal* 25, 967–976. <https://doi.org/10.1038/sj.emboj.7601024>
- Altmann, F., 2007. The Role of Protein Glycosylation in Allergy. *IAA* 142, 99–115. <https://doi.org/10.1159/000096114>
- Altmann, F., Schwihla, H., Staudacher, E., Glössl, J., März, L., 1995. Insect cells contain an unusual, membrane-bound beta-N-acetylglucosaminidase probably involved in the processing of protein N-glycans. *J. Biol. Chem.* 270, 17344–17349. <https://doi.org/10.1074/jbc.270.29.17344>
- Anthony, R.M., Nimmerjahn, F., Ashline, D.J., Reinhold, V.N., Paulson, J.C., Ravetch, J.V., 2008. Recapitulation of IVIG Anti-Inflammatory Activity with a Recombinant IgG Fc. *Science* 320, 373–376. <https://doi.org/10.1126/science.1154315>
- Anthony, R.M., Wermeling, F., Ravetch, J.V., 2012. Novel roles for the IgG Fc glycan. *Ann. N. Y. Acad. Sci.* 1253, 170–180. <https://doi.org/10.1111/j.1749-6632.2011.06305.x>
- Apweiler, R., Hermjakob, H., Sharon, N., 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* 1473, 4–8.
- Arnold, J.N., Wormald, M.R., Sim, R.B., Rudd, P.M., Dwek, R.A., 2007. The Impact of Glycosylation on the Biological Function and Structure of Human Immunoglobulins. *Annual Review of Immunology* 25, 21–50. <https://doi.org/10.1146/annurev.immunol.25.022106.141702>
- Aviezer, D., Brill-Almon, E., Shaaltiel, Y., Hashmueli, S., Bartfeld, D., Mizrachi, S., Liberman, Y., Freeman, A., Zimran, A., Galun, E., 2009. A Plant-Derived Recombinant Human Glucocerebrosidase Enzyme—A Preclinical and Phase I Investigation. *PLoS ONE* 4, e4792. <https://doi.org/10.1371/journal.pone.0004792>
- Baïet, B., Burel, C., Saint-Jean, B., Louvet, R., Menu-Bouaouiche, L., Kiefer-Meyer, M.-C., Mathieu-Rivet, E., Lefebvre, T., Castel, H., Carlier, A., Cadoret, J.-P., Lerouge, P., Bardor, M., 2011. N-Glycans of *Phaeodactylum tricornutum* Diatom and Functional Characterization of Its N-Acetylglucosaminyltransferase I Enzyme. *J Biol Chem* 286, 6152–6164. <https://doi.org/10.1074/jbc.M110.175711>

- Baker, K.N., Rendall, M.H., Hills, A.E., Hoare, M., Freedman, R.B., James, D.C., 2001. Metabolic control of recombinant protein N-glycan processing in NS0 and CHO cells. *Biotechnol. Bioeng.* 73, 188–202. <https://doi.org/10.1002/bit.1051>
- Bakker, H., Bardor, M., Molthoff, J.W., Gomord, V., Elbers, I., Stevens, L.H., Jordi, W., Lommen, A., Faye, L., Lerouge, P., Bosch, D., 2001. Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. U.S.A.* 98, 2899–2904. <https://doi.org/10.1073/pnas.031419998>
- Bakker, H., Rouwendal, G.J.A., Karnoup, A.S., Florack, D.E.A., Stoop, G.M., Helsper, J.P.F.G., Ree, R. van, Die, I. van, Bosch, D., 2006. An antibody produced in tobacco expressing a hybrid β -1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *PNAS* 103, 7577–7582. <https://doi.org/10.1073/pnas.0600879103>
- Bardor, M., Faveeuw, C., Fitchette, A.-C., Gilbert, D., Galas, L., Trottein, F., Faye, L., Lerouge, P., 2003a. Immunoreactivity in mammals of two typical plant glyco-epitopes, core α (1,3)-fucose and core xylose. *Glycobiology* 13, 427–434. <https://doi.org/10.1093/glycob/cwg024>
- Bardor, M., Loutelier-Bourhis, C., Paccalet, T., Cosette, P., Fitchette, A.-C., Vézina, L.-P., Trépanier, S., Dargis, M., Lemieux, R., Lange, C., Faye, L., Lerouge, P., 2003b. Monoclonal C5-1 antibody produced in transgenic alfalfa plants exhibits a N-glycosylation that is homogenous and suitable for glyco-engineering into human-compatible structures. *Plant Biotechnology Journal* 1, 451–462. <https://doi.org/10.1046/j.1467-7652.2003.00041.x>
- Bardor, M., Nguyen, D.H., Diaz, S., Varki, A., 2005. Mechanism of uptake and incorporation of the non-human sialic acid N-glycolylneuraminic acid into human cells. *J. Biol. Chem.* 280, 4228–4237. <https://doi.org/10.1074/jbc.M412040200>
- Barrera, D.J., Mayfield, S.P., 2013. High-value Recombinant Protein Production in Microalgae, in: Emeritus, A.R.P.D., Ph.D, Q.H. (Eds.), *Handbook of Microalgal Culture*. John Wiley & Sons, Ltd, pp. 532–544. <https://doi.org/10.1002/9781118567166.ch27>
- Barrera, D.J., Rosenberg, J.N., Chiu, J.G., Chang, Y.-N., Debatis, M., Ngoi, S.-M., Chang, J.T., Shoemaker, C.B., Oyler, G.A., Mayfield, S.P., 2015. Algal chloroplast produced camelid VHH antitoxins are capable of neutralizing botulinum neurotoxin. *Plant Biotechnology Journal* 13, 117–124. <https://doi.org/10.1111/pbi.12244>
- Basnet, N.B., Ide, K., Tahara, H., Tanaka, Y., Ohdan, H., 2010. Deficiency of N-glycolylneuraminic acid and Gal α 1-3Gal β 1-4GlcNAc epitopes in xenogeneic cells attenuates cytotoxicity of human natural antibodies. *Xenotransplantation* 17, 440–448. <https://doi.org/10.1111/j.1399-3089.2010.00610.x>
- Benyair, R., Ogen-Shtern, N., Mazkereth, N., Shai, B., Ehrlich, M., Lederkremer, G.Z., 2015. Mammalian ER mannosidase I resides in quality control vesicles, where it encounters its glycoprotein substrates. *Molecular Biology of the Cell* 26, 172–184. <https://doi.org/10.1091/mbc.E14-06-1152>
- Bernardo, A.D., Macor, P., Guarnotta, C., Franco, G., Florena, A.M., Tedesco, F., Tripodo, C., 2010. Humoral immunotherapy of multiple myeloma: perspectives and perplexities. *Expert Opinion on Biological Therapy* 10, 863–873. <https://doi.org/10.1517/14712591003774063>
- Bobrowicz, P., Davidson, R.C., Li, H., Potgieter, T.I., Nett, J.H., Hamilton, S.R., Stadheim, T.A., Miele, R.G., Bobrowicz, B., Mitchell, T., Rausch, S., Renfer, E., Wildt, S., 2004. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia*

- pastoris: production of complex humanized glycoproteins with terminal galactose. *Glycobiology* 14, 757–766. <https://doi.org/10.1093/glycob/cwh104>
- Bork, K., Horstkorte, R., Weidemann, W., 2009. Increasing the sialylation of therapeutic glycoproteins: The potential of the sialic acid biosynthetic pathway. *J. Pharm. Sci.* 98, 3499–3508. <https://doi.org/10.1002/jps.21684>
- Bosco, M., Gall, S.L., Rihouey, C., Couve-Bonnaire, S., Bardor, M., Lerouge, P., Pannecoucke, X., 2008. 6-Azido d-galactose transfer to N-acetyl-d-glucosamine derivative using commercially available β -1,4-galactosyltransferase. *Tetrahedron Letters* 49, 2294–2297. <https://doi.org/10.1016/j.tetlet.2008.02.018>
- Both, P., Sobczak, L., Breton, C., Hann, S., Nöbauer, K., Paschinger, K., Kozmon, S., Mucha, J., Wilson, I.B.H., 2011. Distantly related plant and nematode core α 1,3-fucosyltransferases display similar trends in structure–function relationships. *Glycobiology* 21, 1401–1415. <https://doi.org/10.1093/glycob/cwr056>
- Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otilar, R.P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J.A., Brownlee, C., Cadoret, J.-P., Chiovitti, A., Choi, C.J., Coesel, S., De Martino, A., Detter, J.C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman, M.J.J., Jenkins, B.D., Jiroutova, K., Jorgensen, R.E., Joubert, Y., Kaplan, A., Kröger, N., Kroth, P.G., La Roche, J., Lindquist, E., Lommer, M., Martin–Jézéquel, V., Lopez, P.J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L.K., Montsant, A., Secq, M.-P.O., Napoli, C., Obornik, M., Parker, M.S., Petit, J.-L., Porcel, B.M., Poulsen, N., Robison, M., Rychlewski, L., Rynearson, T.A., Schmutz, J., Shapiro, H., Siaut, M., Stanley, M., Sussman, M.R., Taylor, A.R., Vardi, A., von Dassow, P., Vyverman, W., Willis, A., Wyrwicz, L.S., Rokhsar, D.S., Weissenbach, J., Armbrust, E.V., Green, B.R., Van de Peer, Y., Grigoriev, I.V., 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456, 239–244. <https://doi.org/10.1038/nature07410>
- Boyd, P.N., Lines, A.C., Patel, A.K., 1995. The effect of the removal of sialic acid, galactose and total carbohydrate on the functional activity of Campath-1H. *Mol. Immunol.* 32, 1311–1318. [https://doi.org/10.1016/0161-5890\(95\)00118-2](https://doi.org/10.1016/0161-5890(95)00118-2)
- Bragonzi, A., Distefano, G., Buckberry, L.D., Acerbis, G., Foglieni, C., Lamotte, D., Campi, G., Marc, A., Soria, M.R., Jenkins, N., Monaco, L., 2000. A new Chinese hamster ovary cell line expressing alpha2,6-sialyltransferase used as universal host for the production of human-like sialylated recombinant glycoproteins. *Biochim. Biophys. Acta* 1474, 273–282. [https://doi.org/10.1016/S0304-4165\(00\)00023-4](https://doi.org/10.1016/S0304-4165(00)00023-4)
- Breitling, J., Aebi, M., 2013. N-Linked Protein Glycosylation in the Endoplasmic Reticulum. *Cold Spring Harbor Perspectives in Biology* 5, a013359–a013359. <https://doi.org/10.1101/cshperspect.a013359>
- Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., Imberty, A., 2006. Structures and mechanisms of glycosyltransferases. *Glycobiology* 16, 29R–37R. <https://doi.org/10.1093/glycob/cwj016>
- Brockhausen, I., 2006. Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. *EMBO Reports* 7, 599. <https://doi.org/10.1038/sj.embor.7400705>
- Brockhausen, I., 1999. Pathways of O-glycan biosynthesis in cancer cells. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1473, 67–95. [https://doi.org/10.1016/S0304-4165\(99\)00170-1](https://doi.org/10.1016/S0304-4165(99)00170-1)

- Budzianowski, J., 2015. Tobacco against Ebola virus disease. *Prz. Lek.* 72, 567–571.
- Burki, F., Roger, A.J., Brown, M.W., Simpson, A.G.B., 2020. The New Tree of Eukaryotes. *Trends in Ecology & Evolution* 35, 43–55. <https://doi.org/10.1016/j.tree.2019.08.008>
- Cabanes-Macheteau, M., Fitchette-Lainé, A.-C., Loutelier-Bourhis, C., Lange, C., Vine, N.D., K. C. Ma, J., Lerouge, P., Faye, L., 1999. N-Glycosylation of a mouse IgG expressed in transgenic tobacco plants. *Glycobiology* 9, 365–372. <https://doi.org/10.1093/glycob/9.4.365>
- Caffaro, C.E., Hirschberg, C.B., 2006. Nucleotide Sugar Transporters of the Golgi Apparatus: From Basic Science to Diseases. *Acc. Chem. Res.* 39, 805–812. <https://doi.org/10.1021/ar0400239>
- Castilho, A., Pabst, M., Leonard, R., Veit, C., Altmann, F., Mach, L., Glössl, J., Strasser, R., Steinkellner, H., 2008. Construction of a Functional CMP-Sialic Acid Biosynthesis Pathway in Arabidopsis. *Plant Physiol* 147, 331–339. <https://doi.org/10.1104/pp.108.117572>
- Castilho, A., Strasser, R., Stadlmann, J., Grass, J., Jez, J., Gatteringer, P., Kunert, R., Quendler, H., Pabst, M., Leonard, R., Altmann, F., Steinkellner, H., 2010. In planta protein sialylation through overexpression of the respective mammalian pathway. *J. Biol. Chem.* 285, 15923–15930. <https://doi.org/10.1074/jbc.M109.088401>
- Cavazzana-Calvo, M., Debiais, D., 2011. Le biomédicament, pour une définition commune, in: *Que sais-je ? Presses Universitaires de France, Paris*, pp. 9–23.
- CAZy - CE13 [WWW Document], n.d. URL <http://www.cazy.org/CE13.html> (accessed 5.6.20).
- CAZy - GH20 [WWW Document], n.d. URL <http://www.cazy.org/GH20.html> (accessed 5.15.20).
- CAZy - GH38 [WWW Document], n.d. URL <http://www.cazy.org/GH38.html> (accessed 5.15.20).
- CAZy - GT10 [WWW Document], n.d. URL <http://www.cazy.org/GT10.html> (accessed 5.15.20).
- CAZy - GT16 [WWW Document], n.d. URL <http://www.cazy.org/GT16.html> (accessed 5.15.20).
- Chacón-Lee, T.L., González-Mariño, G.E., 2010. Microalgae for “Healthy” Foods—Possibilities and Challenges. *Comprehensive Reviews in Food Science and Food Safety* 9, 655–675. <https://doi.org/10.1111/j.1541-4337.2010.00132.x>
- Chargelegue, D., Drake, P.M.W., Obregon, P., Ma, J.K.-C., 2005. Production of Secretory IgA in Transgenic Plants, in: *Molecular Farming*. John Wiley & Sons, Ltd, pp. 159–169. <https://doi.org/10.1002/3527603638.ch10>
- Chenu, S., Grégoire, A., Malykh, Y., Visvikis, A., Monaco, L., Shaw, L., Schauer, R., Marc, A., Goergen, J.-L., 2003. Reduction of CMP-N-acetylneuraminic acid hydroxylase activity in engineered Chinese hamster ovary cells using an antisense-RNA strategy. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1622, 133–144. [https://doi.org/10.1016/S0304-4165\(03\)00137-5](https://doi.org/10.1016/S0304-4165(03)00137-5)
- Chiba, Y., Suzuki, M., Yoshida, S., Yoshida, A., Ikenaga, H., Takeuchi, M., Jigami, Y., Ichishima, E., 1998. Production of Human Compatible High Mannose-type (Man5GlcNAc2) Sugar Chains in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 273, 26298–26304. <https://doi.org/10.1074/jbc.273.41.26298>
- Chiovitti, A., Harper, R.E., Willis, A., Bacic, A., Mulvaney, P., Wetherbee, R., 2005. Variations in the Substituted 3-Linked Mannans Closely Associated with the Silicified Walls of Diatoms¹. *Journal of Phycology* 41, 1154–1161. <https://doi.org/10.1111/j.1529-8817.2005.00140.x>
- Chlamy Library [WWW Document], n.d. URL <https://www.chlamylibrary.org/> (accessed 5.14.20).

- Choi, B.-K., Bobrowicz, P., Davidson, R.C., Hamilton, S.R., Kung, D.H., Li, H., Miele, R.G., Nett, J.H., Wildt, S., Gerngross, T.U., 2003. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. *Proceedings of the National Academy of Sciences* 100, 5022–5027. <https://doi.org/10.1073/pnas.0931263100>
- Chung, C.H., Mirakhur, B., Chan, E., Le, Q.-T., Berlin, J., Morse, M., Murphy, B.A., Satinover, S.M., Hosen, J., Mauro, D., Slebos, R.J., Zhou, Q., Gold, D., Hatley, T., Hicklin, D.J., Platts-Mills, T.A.E., 2008. Cetuximab-induced anaphylaxis and IgE specific for galactose- α -1,3-galactose. *N. Engl. J. Med.* 358, 1109–1117. <https://doi.org/10.1056/NEJMoa074943>
- Clarke, S.J., Smith, J.T., Gebbie, C., Sweeney, C., Olszewski, N., 2009. A phase I, pharmacokinetic (PK), and preliminary efficacy assessment of ALD518, a humanized anti-IL-6 antibody, in patients with advanced cancer. *JCO* 27, 3025–3025. https://doi.org/10.1200/jco.2009.27.15_suppl.3025
- Cobb, B.A., 2020. The history of IgG glycosylation and where we are now. *Glycobiology* 30, 202–213. <https://doi.org/10.1093/glycob/cwz065>
- Code de la santé publique - Article L5111-1, n.d. , Code de la santé publique.
- Cox, K.M., Sterling, J.D., Regan, J.T., Gasdaska, J.R., Frantz, K.K., Peele, C.G., Black, A., Passmore, D., Moldovan-Loomis, C., Srinivasan, M., Cuisson, S., Cardarelli, P.M., Dickey, L.F., 2006. Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nature Biotechnology* 24, 1591. <https://doi.org/10.1038/nbt1260>
- Cruz de Carvalho, M.H., Sun, H.-X., Bowler, C., Chua, N.-H., 2016. Noncoding and coding transcriptome responses of a marine diatom to phosphate fluctuations. *New Phytol.* 210, 497–510. <https://doi.org/10.1111/nph.13787>
- Daboussi, F., Leduc, S., Maréchal, A., Dubois, G., Guyot, V., Perez-Michaut, C., Amato, A., Falciatore, A., Juillerat, A., Beurdeley, M., Voytas, D.F., Cavarec, L., Duchateau, P., 2014. Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nat Commun* 5, 3831. <https://doi.org/10.1038/ncomms4831>
- De Martino, A., Meichenin, A., Shi, J., Pan, K., Bowler, C., 2007. Genetic and phenotypic characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) accessions1. *Journal of Phycology* 43, 992–1009. <https://doi.org/10.1111/j.1529-8817.2007.00384.x>
- De Pourcq, K., De Schutter, K., Callewaert, N., 2010. Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Appl. Microbiol. Biotechnol.* 87, 1617–1631. <https://doi.org/10.1007/s00253-010-2721-1>
- Delmas, F., Séveno, M., Northey, J.G.B., Hernould, M., Lerouge, P., McCourt, P., Chevalier, C., 2008. The synthesis of the rhamnogalacturonan II component 3-deoxy-D-manno-2-octulosonic acid (Kdo) is required for pollen tube growth and elongation. *J Exp Bot* 59, 2639–2647. <https://doi.org/10.1093/jxb/ern118>
- Deshmukh, S., Bala, K., Kumar, R., 2019. Selection of microalgae species based on their lipid content, fatty acid profile and apparent fuel properties for biodiesel production. *Environ Sci Pollut Res* 26, 24462–24473. <https://doi.org/10.1007/s11356-019-05692-z>
- Dhar, C., Sasmal, A., Varki, A., 2019. From “Serum Sickness” to “Xenosialitis”: Past, Present, and Future Significance of the Non-human Sialic Acid Neu5Gc. *Front. Immunol.* 10. <https://doi.org/10.3389/fimmu.2019.00807>

- Ebert, B., Rautengarten, C., Guo, X., Xiong, G., Stonebloom, S., Smith-Moritz, A.M., Herter, T., Chan, L.J.G., Adams, P.D., Petzold, C.J., Pauly, M., Willats, W.G.T., Heazlewood, J.L., Scheller, H.V., 2015. Identification and Characterization of a Golgi-Localized UDP-Xylose Transporter Family from *Arabidopsis*. *The Plant Cell* 27, 1218–1227. <https://doi.org/10.1105/tpc.114.133827>
- Ecker, D.M., Jones, S.D., Levine, H.L., 2015. The therapeutic monoclonal antibody market. *mAbs* 7, 9–14. <https://doi.org/10.4161/19420862.2015.989042>
- Farid, A., Pabst, M., Schoberer, J., Altmann, F., Glössl, J., Strasser, R., 2011. *Arabidopsis thaliana* alpha1,2-glucosyltransferase (ALG10) is required for efficient N-glycosylation and leaf growth. *The Plant Journal* 68, 314–325. <https://doi.org/10.1111/j.1365-313X.2011.04688.x>
- Feige, M.J., Nath, S., Catharino, S.R., Weinfurtner, D., Steinbacher, S., Buchner, J., 2009. Structure of the murine unglycosylated IgG1 Fc fragment. *J. Mol. Biol.* 391, 599–608. <https://doi.org/10.1016/j.jmb.2009.06.048>
- Ferris, S.P., Kodali, V.K., Kaufman, R.J., 2014. Glycoprotein folding and quality-control mechanisms in protein-folding diseases. *Disease Models & Mechanisms* 7, 331–341. <https://doi.org/10.1242/dmm.014589>
- Fitchette-Lainé, A.-C., Gomord, V., Cabanes, M., Michalski, J.-C., Macary, M.S., Foucher, B., Cavelier, B., Hawes, C., Lerouge, P., Faye, L., 1997. N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *The Plant Journal* 12, 1411–1417. <https://doi.org/10.1046/j.1365-313x.1997.12061411.x>
- Fu, D., Chen, L., O'Neill, R.A., 1994. A detailed structural characterization of ribonuclease B oligosaccharides by ¹H NMR spectroscopy and mass spectrometry. *Carbohydrate Research* 261, 173–186. [https://doi.org/10.1016/0008-6215\(94\)84015-6](https://doi.org/10.1016/0008-6215(94)84015-6)
- Galli, C., Bernasconi, R., Soldà, T., Calanca, V., Molinari, M., 2011. Malectin Participates in a Backup Glycoprotein Quality Control Pathway in the Mammalian ER. *PLoS ONE* 6, e16304. <https://doi.org/10.1371/journal.pone.0016304>
- Gao, H., Wright, D.A., Li, T., Wang, Y., Horken, K., Weeks, D.P., Yang, B., Spalding, M.H., 2014. TALE activation of endogenous genes in *Chlamydomonas reinhardtii*. *Algal Research* 5, 52–60. <https://doi.org/10.1016/j.algal.2014.05.003>
- Gao, X.-D., Nishikawa, A., Dean, N., 2004. Physical interactions between the Alg1, Alg2, and Alg11 mannosyltransferases of the endoplasmic reticulum. *Glycobiology* 14, 559–570. <https://doi.org/10.1093/glycob/cwh072>
- Garénaux, E., Shams-Eldin, H., Chirat, F., Bieker, U., Schmidt, J., Michalski, J.-C., Cacan, R., Guérardel, Y., Schwarz, R.T., 2008. The Dual Origin of *Toxoplasma gondii* N-Glycans. *Biochemistry* 47, 12270–12276. <https://doi.org/10.1021/bi801090a>
- Geisler, C., Aumiller, J.J., Jarvis, D.L., 2008. A fused lobes Gene Encodes the Processing β -N-Acetylglucosaminidase in Sf9 Cells. *J. Biol. Chem.* 283, 11330–11339. <https://doi.org/10.1074/jbc.M710279200>
- Gemmill, T.R., Trimble, R.B., 1999. Overview of N- and O-linked oligosaccharide structures found in various yeast species. *Biochim. Biophys. Acta* 1426, 227–237. [https://doi.org/10.1016/S0304-4165\(98\)00126-3](https://doi.org/10.1016/S0304-4165(98)00126-3)
- Gerken, ThomasA., 2012. O-Glycoprotein Biosynthesis: Site Localization by Edman Degradation and Site Prediction Based on Random Peptide Substrates, in: McGuckin, M.A., Thornton, D.J.

- (Eds.), *Mucins, Methods in Molecular Biology*. Humana Press, pp. 81–108. https://doi.org/10.1007/978-1-61779-513-8_5
- Ghaderi, D., Taylor, R.E., Padler-Karavani, V., Diaz, S., Varki, A., 2010. Implications of the presence of N-glycolylneuraminic acid in recombinant therapeutic glycoproteins. *Nat. Biotechnol.* 28, 863–867. <https://doi.org/10.1038/nbt.1651>
- Gomord, V., Faye, L., 2004. Posttranslational modification of therapeutic proteins in plants. *Current Opinion in Plant Biology* 7, 171–181. <https://doi.org/10.1016/j.pbi.2004.01.015>
- Gordon, R.D., Sivarajah, P., Satkunarajah, M., Ma, D., Tarling, C.A., Vizitiu, D., Withers, S.G., Rini, J.M., 2006. X-ray Crystal Structures of Rabbit N-acetylglucosaminyltransferase I (GnT I) in Complex with Donor Substrate Analogues. *Journal of Molecular Biology* 360, 67–79. <https://doi.org/10.1016/j.jmb.2006.04.058>
- Grabowski, G.A., Golembo, M., Shaaltiel, Y., 2014. Taliglucerase alfa: An enzyme replacement therapy using plant cell expression technology. *Molecular Genetics and Metabolism* 112, 1–8. <https://doi.org/10.1016/j.ymgme.2014.02.011>
- Grilo, A.L., Mantalaris, A., 2018. The Increasingly Human and Profitable Monoclonal Antibody Market. *Trends in Biotechnology*. <https://doi.org/10.1016/j.tibtech.2018.05.014>
- Gudelj, I., Lauc, G., Pezer, M., 2018. Immunoglobulin G glycosylation in aging and diseases. *Cellular Immunology, Special issue: The role of glycans in immunology* 333, 65–79. <https://doi.org/10.1016/j.cellimm.2018.07.009>
- Gügi, B., Le Costaouec, T., Burel, C., Lerouge, P., Helbert, W., Bardor, M., 2015. Diatom-Specific Oligosaccharide and Polysaccharide Structures Help to Unravel Biosynthetic Capabilities in Diatoms. *Mar Drugs* 13, 5993–6018. <https://doi.org/10.3390/md13095993>
- Gustafsson, C., Govindarajan, S., Minshull, J., 2004. Codon bias and heterologous protein expression. *Trends in Biotechnology* 22, 346–353. <https://doi.org/10.1016/j.tibtech.2004.04.006>
- Halim, R., Danquah, M.K., Webley, P.A., 2012. Extraction of oil from microalgae for biodiesel production: A review. *Biotechnology Advances* 30, 709–732. <https://doi.org/10.1016/j.biotechadv.2012.01.001>
- Hamieh, A., Cartier, D., Abid, H., Calas, A., Burel, C., Bucharles, C., Jehan, C., Grumolato, L., Landry, M., Lerouge, P., Anouar, Y., Lihmann, I., 2017. Selenoprotein T is a novel OST subunit that regulates UPR signaling and hormone secretion. *EMBO reports* 18, 1935–1946. <https://doi.org/10.15252/embr.201643504>
- Hamilton, S.R., Bobrowicz, P., Bobrowicz, B., Davidson, R.C., Li, H., Mitchell, T., Nett, J.H., Rausch, S., Stadheim, T.A., Wischnewski, H., Wildt, S., Gerngross, T.U., 2003. Production of Complex Human Glycoproteins in Yeast. *Science* 301, 1244–1246. <https://doi.org/10.1126/science.1088166>
- Hamilton, S.R., Davidson, R.C., Sethuraman, N., Nett, J.H., Jiang, Y., Rios, S., Bobrowicz, P., Stadheim, T.A., Li, H., Choi, B.-K., Hopkins, D., Wischnewski, H., Roser, J., Mitchell, T., Strawbridge, R.R., Hoopes, J., Wildt, S., Gerngross, T.U., 2006. Humanization of Yeast to Produce Complex Terminally Sialylated Glycoproteins. *Science* 313, 1441–1443. <https://doi.org/10.1126/science.1130256>

- Hamilton, S.R., Gerngross, T.U., 2007. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Current Opinion in Biotechnology, Expression technologies / Tissue and cell engineering* 18, 387–392. <https://doi.org/10.1016/j.copbio.2007.09.001>
- Hammond, C., Braakman, I., Helenius, A., 1994. Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc Natl Acad Sci U S A* 91, 913–917. <https://doi.org/10.1073/pnas.91.3.913>
- Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M.A., Samyn-Petit, B., Julien, S., Delannoy, P., 2001. The human sialyltransferase family. *Biochimie* 83, 727–737. [https://doi.org/10.1016/S0300-9084\(01\)01301-3](https://doi.org/10.1016/S0300-9084(01)01301-3)
- Haryadi, R., Zhang, P., Chan, K.F., Song, Z., 2013. CHO-gmt5, a novel CHO glycosylation mutant for producing afucosylated and asialylated recombinant antibodies. *Bioengineered* 4, 90–94. <https://doi.org/10.4161/bioe.22262>
- Hayes, J.M., Frostell, A., Cosgrave, E.F.J., Struwe, W.B., Potter, O., Davey, G.P., Karlsson, R., Anneren, C., Rudd, P.M., 2014. Fc gamma receptor glycosylation modulates the binding of IgG glycoforms: a requirement for stable antibody interactions. *J. Proteome Res.* 13, 5471–5485. <https://doi.org/10.1021/pr500414q>
- Hayes, J.M., Wormald, M.R., Rudd, P.M., Davey, G.P., 2016. Fc gamma receptors: glycobiology and therapeutic prospects. *J Inflamm Res* 9, 209–219. <https://doi.org/10.2147/JIR.S121233>
- He, L., Han, X., Yu, Z., 2014. A Rare *Phaeodactylum tricornutum* Cruciform Morphotype: Culture Conditions, Transformation and Unique Fatty Acid Characteristics. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0093922>
- Helenius, A., Aebi, M., 2004. Roles of N-Linked Glycans in the Endoplasmic Reticulum. *Annual Review of Biochemistry* 73, 1019–1049. <https://doi.org/10.1146/annurev.biochem.73.011303.073752>
- Hempel, F., Lau, J., Klingl, A., Maier, U.G., 2011. Algae as protein factories: expression of a human antibody and the respective antigen in the diatom *Phaeodactylum tricornutum*. *PLoS ONE* 6, e28424. <https://doi.org/10.1371/journal.pone.0028424>
- Hempel, F., Maier, U.G., 2016. Microalgae as Solar-Powered Protein Factories, in: *Advanced Technologies for Protein Complex Production and Characterization, Advances in Experimental Medicine and Biology*. Springer, Cham, pp. 241–262. https://doi.org/10.1007/978-3-319-27216-0_16
- Hempel, F., Maier, U.G., 2012. An engineered diatom acting like a plasma cell secreting human IgG antibodies with high efficiency. *Microbial Cell Factories* 11, 126. <https://doi.org/10.1186/1475-2859-11-126>
- Hempel, F., Maurer, M., Brockmann, B., Mayer, C., Biedenkopf, N., Kelterbaum, A., Becker, S., Maier, U.G., 2017. From hybridomas to a robust microalgal-based production platform: molecular design of a diatom secreting monoclonal antibodies directed against the Marburg virus nucleoprotein. *Microbial Cell Factories* 16, 131. <https://doi.org/10.1186/s12934-017-0745-2>
- Hiatt, A., Cafferkey, R., Bowdish, K., 1989. Production of antibodies in transgenic plants. *Nature* 342, 76–78. <https://doi.org/10.1038/342076a0>

- Higel, F., Seidl, A., Sörgel, F., Friess, W., 2016. N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. *European Journal of Pharmaceutics and Biopharmaceutics* 100, 94–100. <https://doi.org/10.1016/j.ejpb.2016.01.005>
- Hilger, C., Fischer, J., Wölbing, F., Biedermann, T., 2019. Role and Mechanism of Galactose-Alpha-1,3-Galactose in the Elicitation of Delayed Anaphylactic Reactions to Red Meat. *Curr Allergy Asthma Rep* 19. <https://doi.org/10.1007/s11882-019-0835-9>
- Hoagland, K.D., Rosowski, J.R., Gretz, M.R., Roemer, S.C., 1993. Diatom Extracellular Polymeric Substances: Function, Fine Structure, Chemistry, and Physiology. *Journal of Phycology* 29, 537–566. <https://doi.org/10.1111/j.0022-3646.1993.00537.x>
- Hodoniczky, J., Zheng, Y.Z., James, D.C., 2005. Control of Recombinant Monoclonal Antibody Effector Functions by Fc N-Glycan Remodeling in Vitro. *Biotechnology Progress* 21, 1644–1652. <https://doi.org/10.1021/bp050228w>
- Holland, M., Yagi, H., Takahashi, N., Kato, K., Savage, C.O.S., Goodall, D.M., Jefferis, R., 2006. Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis. *Biochimica et Biophysica Acta (BBA) - General Subjects, Glycoproteomics* 1760, 669–677. <https://doi.org/10.1016/j.bbagen.2005.11.021>
- Hossler, P., Khattak, S.F., Li, Z.J., 2009. Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* 19, 936–949. <https://doi.org/10.1093/glycob/cwp079>
- Houeix, B., Cairns, M.T., 2019. Engineering of CHO cells for the production of vertebrate recombinant sialyltransferases. *PeerJ* 7, e5788. <https://doi.org/10.7717/peerj.5788>
- Hu, J., Liu, J., Yang, D., Lu, M., Yin, J., 2014. Physiological roles of asialoglycoprotein receptors (ASGPRs) variants and recent advances in hepatic-targeted delivery of therapeutic molecules via ASGPRs. *Protein Pept. Lett.* 21, 1025–1030. <https://doi.org/10.2174/0929866521666140626102429>
- Hubbard, S.C., Robbins, P.W., 1979. Synthesis and processing of protein-linked oligosaccharides in vivo. *J. Biol. Chem.* 254, 4568–4576.
- Irie, A., Suzuki, A., 1998. CMP-N-Acetylneuraminic Acid Hydroxylase Is Exclusively Inactive in Humans. *Biochemical and Biophysical Research Communications* 248, 330–333. <https://doi.org/10.1006/bbrc.1998.8946>
- Itakura, Y., Nakamura-Tsuruta, S., Kominami, J., Sharon, N., Kasai, K., Hirabayashi, J., 2007. Systematic Comparison of Oligosaccharide Specificity of Ricinus communis Agglutinin I and Erythrina Lectins: a Search by Frontal Affinity Chromatography. *J Biochem* 142, 459–469. <https://doi.org/10.1093/jb/mvm153>
- Izquierdo, L., Schulz, B.L., Rodrigues, J.A., Güther, M.L.S., Procter, J.B., Barton, G.J., Aebi, M., Ferguson, M.A.J., 2009. Distinct donor and acceptor specificities of Trypanosoma brucei oligosaccharyltransferases. *EMBO J* 28, 2650–2661. <https://doi.org/10.1038/emboj.2009.203>
- Jacobs, P.P., Callewaert, N., 2009. N-glycosylation Engineering of Biopharmaceutical Expression Systems. *Current Molecular Medicine* 9, 774–800. <https://doi.org/10.2174/156652409789105552>

- Jansing, J., Sack, M., Augustine, S.M., Fischer, R., Bortesi, L., 2019. CRISPR/Cas9-mediated knockout of six glycosyltransferase genes in *Nicotiana benthamiana* for the production of recombinant proteins lacking β -1,2-xylose and core α -1,3-fucose. *Plant Biotechnology Journal* 17, 350–361. <https://doi.org/10.1111/pbi.12981>
- Joao, H.C., Dwek, R.A., 1993. Effects of glycosylation on protein structure and dynamics in ribonuclease B and some of its individual glycoforms. *European Journal of Biochemistry* 218, 239–244. <https://doi.org/10.1111/j.1432-1033.1993.tb18370.x>
- Jones, A.J.S., Papac, D.I., Chin, E.H., Keck, R., Baughman, S.A., Lin, Y.S., Kneer, J., Battersby, J.E., 2007. Selective clearance of glycoforms of a complex glycoprotein pharmaceutical caused by terminal N-acetylglucosamine is similar in humans and cynomolgus monkeys. *Glycobiology* 17, 529–540. <https://doi.org/10.1093/glycob/cwm017>
- Kadirvelraj, R., Yang, J.-Y., Sanders, J.H., Liu, L., Ramiah, A., Prabhakar, P.K., Boons, G.-J., Wood, Z.A., Moremen, K.W., 2018. Human N-acetylglucosaminyltransferase II substrate recognition uses a modular architecture that includes a convergent exosite. *PNAS* 115, 4637–4642. <https://doi.org/10.1073/pnas.1716988115>
- Kaneko, Y., Nimmerjahn, F., Ravetch, J.V., 2006. Anti-Inflammatory Activity of Immunoglobulin G Resulting from Fc Sialylation. *Science* 313, 670–673. <https://doi.org/10.1126/science.1129594>
- Kannicht, C., Fuchs, B., 2008. Post-Translational Modifications of Proteins, in: *Molecular Biomethods Handbook*. Humana Press, pp. 427–449. https://doi.org/10.1007/978-1-60327-375-6_28
- Karas, B.J., Diner, R.E., Lefebvre, S.C., McQuaid, J., Phillips, A.P.R., Noddings, C.M., Brunson, J.K., Valas, R.E., Deerinck, T.J., Jablanovic, J., Gillard, J.T.F., Beerli, K., Ellisman, M.H., Glass, J.I., Iii, C.A.H., Smith, H.O., Venter, J.C., Allen, A.E., Dupont, C.L., Weyman, P.D., 2015. Designer diatom episomes delivered by bacterial conjugation. *Nat Commun* 6, 1–10. <https://doi.org/10.1038/ncomms7925>
- Kelleher, D.J., Gilmore, R., 2006. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* 16, 47R-62R. <https://doi.org/10.1093/glycob/cwj066>
- Khoury, G.A., Baliban, R.C., Floudas, C.A., 2011. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific Reports* 1. <https://doi.org/10.1038/srep00090>
- Kim, Y.K., Kim, K.R., Kang, D.G., Jang, S.Y., Kim, Y.H., Cha, H.J., 2009. Suppression of beta-N-acetylglucosaminidase in the N-glycosylation pathway for complex glycoprotein formation in *Drosophila* S2 cells. *Glycobiology* 19, 301–308. <https://doi.org/10.1093/glycob/cwn138>
- Ko, J.H., Miyoshi, E., Noda, K., Ekuni, A., Kang, R., Ikeda, Y., Taniguchi, N., 1999. Regulation of the GnT-V Promoter by Transcription Factor Ets-1 in Various Cancer Cell Lines. *J. Biol. Chem.* 274, 22941–22948. <https://doi.org/10.1074/jbc.274.33.22941>
- Ko, K., Tekoah, Y., Rudd, P.M., Harvey, D.J., Dwek, R.A., Spitsin, S., Hanlon, C.A., Rupprecht, C., Dietzschold, B., Golovkin, M., Koprowski, H., 2003. Function and glycosylation of plant-derived antiviral monoclonal antibody. *PNAS* 100, 8013–8018. <https://doi.org/10.1073/pnas.0832472100>
- Konno, Y., Kobayashi, Y., Takahashi, K., Takahashi, E., Sakae, S., Wakitani, M., Yamano, K., Suzawa, T., Yano, K., Ohta, T., Koike, M., Wakamatsu, K., Hosoi, S., 2012. Fucose content of monoclonal

antibodies can be controlled by culture medium osmolality for high antibody-dependent cellular cytotoxicity. *Cytotechnology* 64, 249–265. <https://doi.org/10.1007/s10616-011-9377-2>

- Krapp, S., Mimura, Y., Jefferis, R., Huber, R., Sondermann, P., 2003. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J. Mol. Biol.* 325, 979–989. [https://doi.org/10.1016/s0022-2836\(02\)01250-0](https://doi.org/10.1016/s0022-2836(02)01250-0)
- Lai, C.-C., Her, G.-R., 2002. Analysis of N-glycosylation of phospholipase A2 from venom of individual bees by microbore high-performance liquid chromatography–electrospray mass spectrometry using an ion trap mass spectrometer. *Journal of Chromatography B* 766, 243–250. [https://doi.org/10.1016/S0378-4347\(01\)00479-0](https://doi.org/10.1016/S0378-4347(01)00479-0)
- Lairson, L.L., Henrissat, B., Davies, G.J., Withers, S.G., 2008. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annual Review of Biochemistry* 77, 521–555. <https://doi.org/10.1146/annurev.biochem.76.061005.092322>
- Landry, N., Ward, B.J., Trépanier, S., Montomoli, E., Dargis, M., Lapini, G., Vézina, L.-P., 2010. Preclinical and clinical development of plant-made virus-like particle vaccine against avian H5N1 influenza. *PLoS ONE* 5, e15559. <https://doi.org/10.1371/journal.pone.0015559>
- Larsen, R.D., Rivera-Marrero, C.A., Ernst, L.K., Cummings, R.D., Lowe, J.B., 1990. Frameshift and nonsense mutations in a human genomic sequence homologous to a murine UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-galactosyltransferase cDNA. *J. Biol. Chem.* 265, 7055–7061.
- Le Costaouëc, T., Unamunzaga, C., Mantecon, L., Helbert, W., 2017. New structural insights into the cell-wall polysaccharide of the diatom *Phaeodactylum tricornutum*. *Algal Research* 26, 172–179. <https://doi.org/10.1016/j.algal.2017.07.021>
- Leach, M.R., Cohen-Doyle, M.F., Thomas, D.Y., Williams, D.B., 2002. Localization of the Lectin, ERp57 Binding, and Polypeptide Binding Sites of Calnexin and Calreticulin. *J. Biol. Chem.* 277, 29686–29697. <https://doi.org/10.1074/jbc.M202405200>
- Lee, E.U., Roth, J., Paulson, J.C., 1989. Alteration of terminal glycosylation sequences on N-linked oligosaccharides of Chinese hamster ovary cells by expression of beta-galactoside alpha 2,6-sialyltransferase. *J. Biol. Chem.* 264, 13848–13855.
- Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.-C., Gomord, V., Faye, L., 1998. N-Glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol Biol* 38, 31–48. <https://doi.org/10.1023/A:1006012005654>
- Levy-Ontman, O., Fisher, M., Shotland, Y., Weinstein, Y., Tekoah, Y., Arad, S.M., 2014. Genes Involved in the Endoplasmic Reticulum N-Glycosylation Pathway of the Red Microalga *Porphyridium* sp.: A Bioinformatic Study. *Int J Mol Sci* 15, 2305–2326. <https://doi.org/10.3390/ijms15022305>
- Li, H., Sethuraman, N., Stadheim, T.A., Zha, D., Prinz, B., Ballew, N., Bobrowicz, P., Choi, B.-K., Cook, W.J., Cukan, M., Houston-Cummings, N.R., Davidson, R., Gong, B., Hamilton, S.R., Hoopes, J.P., Jiang, Y., Kim, N., Mansfield, R., Nett, J.H., Rios, S., Strawbridge, R., Wildt, S., Gerngross, T.U., 2006. Optimization of humanized IgGs in glycoengineered *Pichia pastoris*. *Nature Biotechnology* 24, nbt1178. <https://doi.org/10.1038/nbt1178>
- Li, X., Zhang, R., Patena, W., Gang, S.S., Blum, S.R., Ivanova, N., Yue, R., Robertson, J.M., Lefebvre, P.A., Fitz-Gibbon, S.T., Grossman, A.R., Jonikas, M.C., 2016. An Indexed, Mapped Mutant

- Library Enables Reverse Genetics Studies of Biological Processes in *Chlamydomonas reinhardtii*. *Plant Cell* 28, 367–387. <https://doi.org/10.1105/tpc.15.00465>
- Liebminger, E., Veit, C., Pabst, M., Batoux, M., Zipfel, C., Altmann, F., Mach, L., Strasser, R., 2011. β -N-Acetylhexosaminidases HEXO1 and HEXO3 Are Responsible for the Formation of Paucimannosidic N-Glycans in *Arabidopsis thaliana*. *J. Biol. Chem.* 286, 10793–10802. <https://doi.org/10.1074/jbc.M110.178020>
- Lingg, N., Zhang, P., Song, Z., Bardor, M., 2012. The sweet tooth of biopharmaceuticals: Importance of recombinant protein glycosylation analysis. *Biotechnology Journal* 7, 1462–1472. <https://doi.org/10.1002/biot.201200078>
- Liu, H., Guo, L., Zhang, J., Zhou, Y., Zhou, J., Yao, J., Wu, H., Yao, S., Chen, B., Chai, Y., Qi, J., Gao, G.F., Tan, S., Feng, H., Yan, J., 2019. Glycosylation-independent binding of monoclonal antibody toripalimab to FG loop of PD-1 for tumor immune checkpoint therapy. *MAbs* 11, 681–690. <https://doi.org/10.1080/19420862.2019.1596513>
- Liu, L., 2015. Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins. *J Pharm Sci* 104, 1866–1884. <https://doi.org/10.1002/jps.24444>
- Liu, L., Stadheim, A., Hamuro, L., Pittman, T., Wang, W., Zha, D., Hochman, J., Prueksaritanont, T., 2011. Pharmacokinetics of IgG1 monoclonal antibodies produced in humanized *Pichia pastoris* with specific glycoforms: a comparative study with CHO produced materials. *Biologicals* 39, 205–210. <https://doi.org/10.1016/j.biologicals.2011.06.002>
- Liu, L., Xu, Y.-X., Hirschberg, C.B., 2010. The role of nucleotide sugar transporters in development of eukaryotes. *Seminars in Cell & Developmental Biology, Mitochondrial Morphology: Shaping the Cell Function* 21, 600–608. <https://doi.org/10.1016/j.semcd.2010.02.002>
- Lizak, C., Gerber, S., Numao, S., Aebi, M., Locher, K.P., 2011. X-ray structure of a bacterial oligosaccharyltransferase. *Nature* 474, 350–355. <https://doi.org/10.1038/nature10151>
- Lucas, P.-L., 2019. Etude et ingénierie de la N-glycosylation des protéines chez la microalgue verte *Chlamydomonas reinhardtii*. Université de Rouen Normandie.
- Lucas, P.-L., Mathieu-Rivet, E., Song, P.C.T., Oltmanns, A., Loutelier-Bourhis, C., Plasson, C., Afonso, C., Hippler, M., Lerouge, P., Mati-Baouche, N., Bardor, M., 2020. Multiple xylosyltransferases heterogeneously xylosylate protein N-linked glycans in *Chlamydomonas reinhardtii*. *The Plant Journal* n/a. <https://doi.org/10.1111/tpj.14620>
- Luo, C., Chen, S., Xu, N., Wang, C., Sai, W. bo, Zhao, W., Li, Y. chun, Hu, X. jing, Tian, H., Gao, X. dong, Yao, W. bing, 2017. Glycoengineering of pertuzumab and its impact on the pharmacokinetic/pharmacodynamic properties. *Sci Rep* 7. <https://doi.org/10.1038/srep46347>
- Ma, J.K.-C., Hikmat, B.Y., Wycoff, K., Vine, N.D., Chargelegue, D., Yu, L., Hein, M.B., Lehner, T., 1998. Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nat Med* 4, 601–606. <https://doi.org/10.1038/nm0598-601>
- Ma, J.K.-C., Lehner, T., Stabila, P., Fux, C.I., Hiatt, A., 1994. Assembly of monoclonal antibodies with IgG1 and IgA heavy chain domains in transgenic tobacco plants. *European Journal of Immunology* 24, 131–138. <https://doi.org/10.1002/eji.1830240120>

- Maeda, E., Kita, S., Kinoshita, M., Urakami, K., Hayakawa, T., Kakehi, K., 2012. Analysis of Nonhuman N-Glycans as the Minor Constituents in Recombinant Monoclonal Antibody Pharmaceuticals. *Anal. Chem.* 84, 2373–2379. <https://doi.org/10.1021/ac300234a>
- Malphettes, L., Freyvert, Y., Chang, J., Liu, P.-Q., Chan, E., Miller, J.C., Zhou, Z., Nguyen, T., Tsai, C., Snowden, A.W., Collingwood, T.N., Gregory, P.D., Cost, G.J., 2010. Highly efficient deletion of FUT8 in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. *Biotechnology and Bioengineering* 106, 774–783. <https://doi.org/10.1002/bit.22751>
- Maszczak-Seneczko, D., Sosicka, P., Olczak, T., Jakimowicz, P., Majkowski, M., Olczak, M., 2013. UDP-N-acetylglucosamine Transporter (SLC35A3) Regulates Biosynthesis of Highly Branched N-Glycans and Keratan Sulfate. *J. Biol. Chem.* 288, 21850–21860. <https://doi.org/10.1074/jbc.M113.460543>
- Mata, T.M., Martins, A.A., Caetano, Nidia.S., 2010. Microalgae for biodiesel production and other applications: A review. *Renewable and Sustainable Energy Reviews* 14, 217–232. <https://doi.org/10.1016/j.rser.2009.07.020>
- Mathieu-Rivet, E., Kiefer-Meyer, M.-C., Vanier, G., Ovide, C., Burel, C., Lerouge, P., Bardor, M., 2014. Protein N-glycosylation in eukaryotic microalgae and its impact on the production of nuclear expressed biopharmaceuticals. *Front Plant Sci* 5, 359. <https://doi.org/10.3389/fpls.2014.00359>
- Mathieu-Rivet, E., Scholz, M., Arias, C., Dardelle, F., Schulze, S., Le Mauff, F., Teo, G., Hochmal, A.K., Blanco-Rivero, A., Loutelier-Bourhis, C., Kiefer-Meyer, M.-C., Fufezan, C., Burel, C., Lerouge, P., Martinez, F., Bardor, M., Hippler, M., 2013. Exploring the N-glycosylation pathway in *Chlamydomonas reinhardtii* unravels novel complex structures. *Mol. Cell Proteomics* 12, 3160–3183. <https://doi.org/10.1074/mcp.M113.028191>
- Matthijs, M., Fabris, M., Broos, S., Vyverman, W., Goossens, A., 2016. Profiling of the Early Nitrogen Stress Response in the Diatom *Phaeodactylum tricornutum* Reveals a Novel Family of RING-Domain Transcription Factors. *Plant Physiology* 170, 489–498. <https://doi.org/10.1104/pp.15.01300>
- Mayfield, S.P., Franklin, S.E., Lerner, R.A., 2003. Expression and assembly of a fully active antibody in algae. *Proc Natl Acad Sci U S A* 100, 438–442. <https://doi.org/10.1073/pnas.0237108100>
- Mengle-Gaw, L., McCoy-Haman, M.F., Tiemeier, D.C., 1991. Genomic structure and expression of human β -1,4-galactosyltransferase. *Biochemical and Biophysical Research Communications* 176, 1269–1276. [https://doi.org/10.1016/0006-291X\(91\)90423-5](https://doi.org/10.1016/0006-291X(91)90423-5)
- Merchant, S.S., Kropat, J., Liu, B., Shaw, J., Warakanont, J., 2012. TAG, You're it! *Chlamydomonas* as a reference organism for understanding algal triacylglycerol accumulation. *Current Opinion in Biotechnology, Energy biotechnology • Environmental biotechnology* 23, 352–363. <https://doi.org/10.1016/j.copbio.2011.12.001>
- Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Maréchal-Drouard, L., Marshall, W.F., Qu, L.-H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V.V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.-L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K.,

- Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A.V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J.V., Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S.V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.-P., Riaño-Pachón, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.-J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y.W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W.C.A., Otilar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I.V., Rokhsar, D.S., Grossman, A.R., 2007. The *Chlamydomonas* Genome Reveals the Evolution of Key Animal and Plant Functions. *Science* 318, 245–250. <https://doi.org/10.1126/science.1143609>
- Mercx, S., Smargiasso, N., Chaumont, F., De Pauw, E., Boutry, M., Navarre, C., 2017. Inactivation of the $\beta(1,2)$ -xylosyltransferase and the $\alpha(1,3)$ -fucosyltransferase genes in *Nicotiana tabacum* BY-2 Cells by a Multiplex CRISPR/Cas9 Strategy Results in Glycoproteins without Plant-Specific Glycans. *Front. Plant Sci.* 8. <https://doi.org/10.3389/fpls.2017.00403>
- Mercx, S., Tollet, J., Magy, B., Navarre, C., Boutry, M., 2016. Gene Inactivation by CRISPR-Cas9 in *Nicotiana tabacum* BY-2 Suspension Cells. *Front. Plant Sci.* 7. <https://doi.org/10.3389/fpls.2016.00040>
- Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J.W., Marth, J.D., 1994. Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development. *EMBO J.* 13, 2056–2065.
- Miller, R.K., Mace, K., Polliotti, B., DeRita, R., Hall, W., Treacy, G., 2003. Marginal Transfer of ReoPro™ (Abciximab) Compared with Immunoglobulin G (F105), Inulin and Water in the Perfused Human Placenta In Vitro. *Placenta* 24, 727–738. [https://doi.org/10.1016/S0143-4004\(03\)00101-2](https://doi.org/10.1016/S0143-4004(03)00101-2)
- Mócsai, R., Blaukopf, M., Svehla, E., Kosma, P., Altmann, F., 2020b. The N-glycans of *Chlorella sorokiniana* and a related strain contain arabinose but have strikingly different structures. *Glycobiology*. <https://doi.org/10.1093/glycob/cwaa012>
- Mócsai, R., Figl, R., Sützl, L., Fluch, S., Altmann, F., 2020a. A first view on the unsuspected intra-genus diversity of N-glycans in *Chlorella* microalgae. *The Plant Journal* n/a. <https://doi.org/10.1111/tpj.14718>
- Mócsai, R., Figl, R., Troschl, C., Strasser, R., Svehla, E., Windwarder, M., Thader, A., Altmann, F., 2019. N-glycans of the microalga *Chlorella vulgaris* are of the oligomannosidic type but highly methylated. *Sci Rep* 9, 1–8. <https://doi.org/10.1038/s41598-018-36884-1>
- Mohorko, E., Glockshuber, R., Aebi, M., 2011. Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. *J Inherit Metab Dis* 34, 869–878. <https://doi.org/10.1007/s10545-011-9337-1>
- Molinari, M., Calanca, V., Galli, C., Lucca, P., Paganetti, P., 2003. Role of EDEM in the Release of Misfolded Glycoproteins from the Calnexin Cycle. *Science* 299, 1397–1400. <https://doi.org/10.1126/science.1079474>

- Münster-Kühnel, A.K., Tiralongo, J., Krapp, S., Weinhold, B., Ritz-Sedlacek, V., Jacob, U., Gerardy-Schahn, R., 2004. Structure and function of vertebrate CMP-sialic acid synthetases. *Glycobiology* 14, 43R-51R. <https://doi.org/10.1093/glycob/cwh113>
- Müthing, J., Kemminer, S.E., Conradt, H.S., Sagi, D., Nimtz, M., Kärst, U., Peter-Katalinić, J., 2003. Effects of buffering conditions and culture pH on production rates and glycosylation of clinical phase I anti-melanoma mouse IgG3 monoclonal antibody R24. *Biotechnol. Bioeng.* 83, 321–334. <https://doi.org/10.1002/bit.10673>
- Nakanishi, K., Narimatsu, S., Ichikawa, S., Tobisawa, Y., Kurohane, K., Niwa, Y., Kobayashi, H., Imai, Y., 2013. Production of Hybrid-IgG/IgA Plantibodies with Neutralizing Activity against Shiga Toxin 1. *PLOS ONE* 8, e80712. <https://doi.org/10.1371/journal.pone.0080712>
- Nakanishi-Shindo, Y., Nakayama, K., Tanaka, A., Toda, Y., Jigami, Y., 1993. Structure of the N-linked oligosaccharides that show the complete loss of alpha-1,6-polymannose outer chain from och1, och1 mnn1, and och1 mnn1 alg3 mutants of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 268, 26338–26345.
- Navarre, C., Smargiasso, N., Duvivier, L., Nader, J., Far, J., De Pauw, E., Boutry, M., 2017. N-Glycosylation of an IgG antibody secreted by *Nicotiana tabacum* BY-2 cells can be modulated through co-expression of human β -1,4-galactosyltransferase. *Transgenic Res* 26, 375–384. <https://doi.org/10.1007/s11248-017-0013-6>
- Nymark, M., Sharma, A.K., Sparstad, T., Bones, A.M., Winge, P., 2016. A CRISPR/Cas9 system adapted for gene editing in marine algae. *Scientific Reports* 6, 24951. <https://doi.org/10.1038/srep24951>
- Olivier, S., Mehtali, M., 2009. Les systèmes alternatifs de production d'anticorps monoclonaux thérapeutiques - Avantages espérés pour le futur. *Med Sci (Paris)* 25, 1163–1168. <https://doi.org/10.1051/medsci/200925121163>
- Oltmanns, A., Hoepfner, L., Scholz, M., Zinzius, K., Schulze, S., Hippler, M., 2020. Novel Insights Into N-Glycan Fucosylation and Core Xylosylation in *C. reinhardtii*. *Front. Plant Sci.* 10. <https://doi.org/10.3389/fpls.2019.01686>
- O'Neil, B.H., Allen, R., Spigel, D.R., Stinchcombe, T.E., Moore, D.T., Berlin, J.D., Goldberg, R.M., 2007. High Incidence of Cetuximab-Related Infusion Reactions in Tennessee and North Carolina and the Association With Atopic History. *JCO* 25, 3644–3648. <https://doi.org/10.1200/JCO.2007.11.7812>
- O'Neill, E.C., Kuhadomlarp, S., Rejzek, M., Fangel, J.U., Alagesan, K., Kolarich, D., Willats, W.G.T., Field, R.A., 2017. Exploring the Glycans of *Euglena gracilis*. *Biology* 6, 45. <https://doi.org/10.3390/biology6040045>
- Otto, R., Santagostino, A., Schrader, U., 2014. Rapid growth in biopharma: Challenges and opportunities | McKinsey & Company [WWW Document]. URL <http://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/rapid-growth-in-biopharma> (accessed 9.22.17).
- Ovide, C., 2016. Etude de la morphogénèse et de la voie de N-glycosylation chez la diatomée *P. Tricornutum*. Université de Rouen.
- Ovide, C., Kiefer-Meyer, M.-C., Bérard, C., Vergne, N., Lecroq, T., Plasson, C., Burel, C., Bernard, S., Driouich, A., Lerouge, P., Tournier, I., Dauchel, H., Bardor, M., 2018. Comparative in depth RNA

sequencing of *P. tricornutum*'s morphotypes reveals specific features of the oval morphotype. *Scientific Reports* 8, 14340. <https://doi.org/10.1038/s41598-018-32519-7>

- Paccalet, T., Bardor, M., Rihouey, C., Delmas, F., Chevalier, C., D'Aoust, M.-A., Faye, L., Vézina, L., Gomord, V., Lerouge, P., 2007. Engineering of a sialic acid synthesis pathway in transgenic plants by expression of bacterial Neu5Ac-synthesizing enzymes. *Plant Biotechnology Journal* 5, 16–25. <https://doi.org/10.1111/j.1467-7652.2006.00211.x>
- Palacpac, N.Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T., Seki, T., 1999. Stable expression of human β 1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *PNAS* 96, 4692–4697. <https://doi.org/10.1073/pnas.96.8.4692>
- Pérez, S., Rodríguez-Carvajal, M.A., Doco, T., 2003. A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function. *Biochimie, Free and conjugated sugars. Biochemistry and molecular biology. Dedicated to Andre Verbert and Jean Agneray* 85, 109–121. [https://doi.org/10.1016/S0300-9084\(03\)00053-1](https://doi.org/10.1016/S0300-9084(03)00053-1)
- Petrucelli, S., Otegui, M.S., Lareu, F., Dinh, O.T., Fitchette, A.-C., Circosta, A., Rumbo, M., Bardor, M., Carcamo, R., Gomord, V., Beachy, R.N., 2006. A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. *Plant Biotechnology Journal* 4, 511–527. <https://doi.org/10.1111/j.1467-7652.2006.00200.x>
- Popp, O., Moser, S., Zielonka, J., Rüger, P., Hansen, S., Plöttner, O., 2018. Development of a pre-glycoengineered CHO-K1 host cell line for the expression of antibodies with enhanced Fc mediated effector function. *MAbs* 10, 290–303. <https://doi.org/10.1080/19420862.2017.1405203>
- Pujar, N., Low, D., O'Leary, R., 2017. Antibody Purification: Drivers of Change, in: *Process Scale Purification of Antibodies*. John Wiley & Sons, Ltd, pp. 699–716. <https://doi.org/10.1002/9781119126942.ch32>
- Qian, J., Liu, T., Yang, L., Daus, A., Crowley, R., Zhou, Q., 2007. Structural characterization of N-linked oligosaccharides on monoclonal antibody cetuximab by the combination of orthogonal matrix-assisted laser desorption/ionization hybrid quadrupole–quadrupole time-of-flight tandem mass spectrometry and sequential enzymatic digestion. *Analytical Biochemistry* 364, 8–18. <https://doi.org/10.1016/j.ab.2007.01.023>
- Radakovits, R., Eduafo, P.M., Posewitz, M.C., 2011. Genetic engineering of fatty acid chain length in *Phaeodactylum tricornutum*. *Metab. Eng.* 13, 89–95. <https://doi.org/10.1016/j.ymben.2010.10.003>
- Rader, R.A., 2008. (Re)defining biopharmaceutical. *Nat Biotech* 26, 743–751. <https://doi.org/10.1038/nbt0708-743>
- Raju, T.S., 2008. Terminal sugars of Fc glycans influence antibody effector functions of IgGs. *Current Opinion in Immunology, Host–Pathogen Interactions / Antibody Therapeutics* 20, 471–478. <https://doi.org/10.1016/j.coi.2008.06.007>
- Raju, T.S., Briggs, J.B., Chamow, S.M., Winkler, M.E., Jones, A.J.S., 2001. Glycoengineering of Therapeutic Glycoproteins: In Vitro Galactosylation and Sialylation of Glycoproteins with Terminal N-Acetylglucosamine and Galactose Residues. *Biochemistry* 40, 8868–8876. <https://doi.org/10.1021/bi010475i>

- Ramírez, A.S., Kowal, J., Locher, K.P., 2019. Cryo–electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B. *Science* 366, 1372–1375. <https://doi.org/10.1126/science.aaz3505>
- Rastogi, A., Maheswari, U., Dorrell, R.G., Vieira, F.R.J., Maumus, F., Kustka, A., McCarthy, J., Allen, A.E., Kersey, P., Bowler, C., Tirichine, L., 2018. Integrative analysis of large scale transcriptome data draws a comprehensive landscape of *Phaeodactylum tricornutum* genome and evolutionary origin of diatoms. *Sci Rep* 8. <https://doi.org/10.1038/s41598-018-23106-x>
- Remmers, I.M., D’Adamo, S., Martens, D.E., de Vos, R.C.H., Mumm, R., America, A.H.P., Cordewener, J.H.G., Bakker, L.V., Peters, S.A., Wijffels, R.H., Lamers, P.P., 2018. Orchestration of transcriptome, proteome and metabolome in the diatom *Phaeodactylum tricornutum* during nitrogen limitation. *Algal Research* 35, 33–49. <https://doi.org/10.1016/j.algal.2018.08.012>
- Römisch, K., 2005. Endoplasmic Reticulum–Associated Degradation. *Annual Review of Cell and Developmental Biology* 21, 435–456. <https://doi.org/10.1146/annurev.cellbio.21.012704.133250>
- Rosales-Mendoza, S., Solís-Andrade, K.I., Márquez-Escobar, V.A., González-Ortega, O., Bañuelos-Hernandez, B., 2020. Current advances in the algae-made biopharmaceuticals field. *Expert Opinion on Biological Therapy* 0, 1–16. <https://doi.org/10.1080/14712598.2020.1739643>
- Roth, J., Zuber, C., Park, S., Jang, I., Lee, Y., Kysela, K.G., Le Fourn, V., Santimaria, R., Guhl, B., Cho, J.W., 2010. Protein N-glycosylation, protein folding, and protein quality control. *Mol. Cells* 30, 497–506. <https://doi.org/10.1007/s10059-010-0159-z>
- Rudd, P.M., Elliott, T., Cresswell, P., Wilson, I.A., Dwek, R.A., 2001. Glycosylation and the Immune System. *Science* 291, 2370–2376. <https://doi.org/10.1126/science.291.5512.2370>
- Ruggiano, A., Foresti, O., Carvalho, P., 2014. ER-associated degradation: Protein quality control and beyond. *J Cell Biol* 204, 869–879. <https://doi.org/10.1083/jcb.201312042>
- Ruiz-Canada, C., Kelleher, D.J., Gilmore, R., 2009. Cotranslational and Posttranslational N-Glycosylation of Polypeptides by Distinct Mammalian OST Isoforms. *Cell* 136, 272–283. <https://doi.org/10.1016/j.cell.2008.11.047>
- Rup, B., Alon, S., Amit-Cohen, B.-C., Brill Almon, E., Chertkoff, R., Tekoah, Y., Rudd, P.M., 2017. Immunogenicity of glycans on biotherapeutic drugs produced in plant expression systems-The taliglucerase alfa story. *PLoS ONE* 12, e0186211. <https://doi.org/10.1371/journal.pone.0186211>
- Rush, J.S., 2016. Role of Flippases in Protein Glycosylation in the Endoplasmic Reticulum. *Lipid Insights* 8, 45–53. <https://doi.org/10.4137/LPI.S31784>
- Rush, J.S., Waechter, C.J., 1995. Transmembrane movement of a water-soluble analogue of mannosylphosphoryldolichol is mediated by an endoplasmic reticulum protein. *The Journal of Cell Biology* 130, 529–536. <https://doi.org/10.1083/jcb.130.3.529>
- Samraj, A.N., Pearce, O.M.T., Läubli, H., Crittenden, A.N., Bergfeld, A.K., Banda, K., Gregg, C.J., Bingman, A.E., Secrest, P., Diaz, S.L., Varki, N.M., Varki, A., 2015. A red meat-derived glycan promotes inflammation and cancer progression. *Proc. Natl. Acad. Sci. U.S.A.* 112, 542–547. <https://doi.org/10.1073/pnas.1417508112>

- Scaife, M.A., Nguyen, G.T.D.T., Rico, J., Lambert, D., Helliwell, K.E., Smith, A.G., 2015. Establishing *Chlamydomonas reinhardtii* as an industrial biotechnology host. *The Plant Journal* 82, 532–546. <https://doi.org/10.1111/tpj.12781>
- Scallon, B.J., Tam, S.H., McCarthy, S.G., Cai, A.N., Raju, T.S., 2007. Higher levels of sialylated Fc glycans in immunoglobulin G molecules can adversely impact functionality. *Molecular Immunology* 44, 1524–1534. <https://doi.org/10.1016/j.molimm.2006.09.005>
- Schallus, T., Jaeckh, C., Feher, K., Palma, A.S., Liu, Y., Simpson, J.C., Mackeen, M., Stier, G., Gibson, T.J., Feizi, T., Pieler, T., Muhle-Goll, C., 2008. Malectin: A Novel Carbohydrate-binding Protein of the Endoplasmic Reticulum and a Candidate Player in the Early Steps of Protein N-Glycosylation. *Molecular Biology of the Cell* 19, 3404–3414. <https://doi.org/10.1091/mbc.E08-04-0354>
- Scheen A. Nomenclature internationale des différents types d'anticorps monoclonaux. *Revue Médicale de Liège*. 2009 64(5-6).
- Schillberg, S., Raven, N., Spiegel, H., Rasche, S., Buntru, M., 2019. Critical Analysis of the Commercial Potential of Plants for the Production of Recombinant Proteins. *Front Plant Sci* 10. <https://doi.org/10.3389/fpls.2019.00720>
- Schneider, J., Castilho, A., Pabst, M., Altmann, F., Gruber, C., Strasser, R., Gattinger, P., Seifert, G.J., Steinkellner, H., 2015. Characterization of plants expressing the human β 1,4-galactosyltransferase gene. *Plant Physiology and Biochemistry* 92, 39–47. <https://doi.org/10.1016/j.plaphy.2015.04.010>
- Schroeder, H.W., Cavacini, L., 2010. Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology, 2010 Primer on Allergic and Immunologic Diseases* 125, S41–S52. <https://doi.org/10.1016/j.jaci.2009.09.046>
- Schulze, S., Oltmanns, A., Machnik, N., Liu, G., Xu, N., Jarmatz, N., Scholz, M., Sugimoto, K., Fufezan, C., Huang, K., Hippler, M., 2018. N-Glycoproteomic characterization of mannosidase and xylotransferase Mutant Strains of *Chlamydomonas*. *Plant Physiology* pp.01450.2017. <https://doi.org/10.1104/pp.17.01450>
- Sergeeva, D., Camacho-Zaragoza, J.M., Lee, J.S., Kildegaard, H.F., 2019. CRISPR/Cas9 as a Genome Editing Tool for Targeted Gene Integration in CHO Cells. *Methods Mol. Biol.* 1961, 213–232. https://doi.org/10.1007/978-1-4939-9170-9_13
- Serif, M., Dubois, G., Finoux, A.-L., Teste, M.-A., Jallet, D., Daboussi, F., 2018. One-step generation of multiple gene knock-outs in the diatom *Phaeodactylum tricorutum* by DNA-free genome editing. *Nature Communications* 9, 3924. <https://doi.org/10.1038/s41467-018-06378-9>
- Seveno, M., Bardor, M., Paccalet, T., Gomord, V., Lerouge, P., Faye, L., 2004. Glycoprotein sialylation in plants? *Nature biotechnology*. <https://doi.org/10.1038/nbt1104-1351>
- Séveno, M., Séveno-Carpentier, E., Voxeur, A., Menu-Bouaouiche, L., Rihouey, C., Delmas, F., Chevalier, C., Driouich, A., Lerouge, P., 2010. Characterization of a putative 3-deoxy-d-manno-2-octulosonic acid (Kdo) transferase gene from *Arabidopsis thaliana*. *Glycobiology* 20, 617–628. <https://doi.org/10.1093/glycob/cwq011>
- Shah, N., Kuntz, D.A., Rose, D.R., 2008. Golgi alpha-mannosidase II cleaves two sugars sequentially in the same catalytic site. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9570–9575. <https://doi.org/10.1073/pnas.0802206105>

- Shin, Y.-J., Castilho, A., Dicker, M., Sádio, F., Vavra, U., Grünwald-Gruber, C., Kwon, T.-H., Altmann, F., Steinkellner, H., Strasser, R., 2017. Reduced paucimannosidic N-glycan formation by suppression of a specific β -hexosaminidase from *Nicotiana benthamiana*. *Plant Biotechnology Journal* 15, 197–206. <https://doi.org/10.1111/pbi.12602>
- Shinkawa, T., Nakamura, K., Yamane, N., Shoji-Hosaka, E., Kanda, Y., Sakurada, M., Uchida, K., Anazawa, H., Satoh, M., Yamasaki, M., Hanai, N., Shitara, K., 2003. The Absence of Fucose but Not the Presence of Galactose or Bisecting N-Acetylglucosamine of Human IgG1 Complex-type Oligosaccharides Shows the Critical Role of Enhancing Antibody-dependent Cellular Cytotoxicity. *J. Biol. Chem.* 278, 3466–3473. <https://doi.org/10.1074/jbc.M210665200>
- Silberstein, S., Gilmore, R., 1996. Biochemistry, molecular biology, and genetics of the oligosaccharyltransferase. *FASEB J* 10, 849–858. <https://doi.org/10.1096/fasebj.10.8.8666161>
- Sizova, I., Greiner, A., Awasthi, M., Kateriya, S., Hegemann, P., 2013. Nuclear gene targeting in *Chlamydomonas* using engineered zinc-finger nucleases. *Plant J* 73, 873–882. <https://doi.org/10.1111/tpj.12066>
- Snider, M.D., Rogers, O.C., 1984. Transmembrane movement of oligosaccharide-lipids during glycoprotein synthesis. *Cell* 36, 753–761. [https://doi.org/10.1016/0092-8674\(84\)90355-6](https://doi.org/10.1016/0092-8674(84)90355-6)
- Song, Z., Lye, G.J., Parker, B.M., 2020. Morphological and biochemical changes in *Phaeodactylum tricornutum* triggered by culture media: Implications for industrial exploitation. *Algal Research* 47, 101822. <https://doi.org/10.1016/j.algal.2020.101822>
- Sourrouille, C., Marquet-Blouin, E., D'Aoust, M.-A., Kiefer-Meyer, M.-C., Seveno, M., Pagny-Salehabadi, S., Bardor, M., Durambur, G., Lerouge, P., Vezina, L., Gomord, V., 2008. Down-regulated expression of plant-specific glycoepitopes in alfalfa. *Plant Biotechnology Journal* 6, 702–721. <https://doi.org/10.1111/j.1467-7652.2008.00353.x>
- Spearman, M., Dionne, B., Butler, M., 2011. The Role of Glycosylation in Therapeutic Antibodies, in: Al-Rubeai, M. (Ed.), *Antibody Expression and Production, Cell Engineering*. Springer Netherlands, Dordrecht, pp. 251–292. https://doi.org/10.1007/978-94-007-1257-7_12
- Sriraman, R., Bardor, M., Sack, M., Vaquero, C., Faye, L., Fischer, R., Finnern, R., Lerouge, P., 2004. Recombinant anti-hCG antibodies retained in the endoplasmic reticulum of transformed plants lack core-xylose and core- α (1,3)-fucose residues. *Plant Biotechnology Journal* 2, 279–287. <https://doi.org/10.1111/j.1467-7652.2004.00078.x>
- Stadlmann, J., Pabst, M., Altmann, F., 2010. Analytical and Functional Aspects of Antibody Sialylation. *J Clin Immunol* 30, 15–19. <https://doi.org/10.1007/s10875-010-9409-2>
- Strasser, R., Altmann, F., Mach, L., Glössl, J., Steinkellner, H., 2004. Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking β 1,2-linked xylose and core α 1,3-linked fucose. *FEBS Letters* 561, 132–136. [https://doi.org/10.1016/S0014-5793\(04\)00150-4](https://doi.org/10.1016/S0014-5793(04)00150-4)
- Strasser, R., Bondili, J.S., Schoberer, J., Svoboda, B., Liebminger, E., Glössl, J., Altmann, F., Steinkellner, H., Mach, L., 2007. Enzymatic Properties and Subcellular Localization of *Arabidopsis* β -N-Acetylhexosaminidases. *Plant Physiol* 145, 5–16. <https://doi.org/10.1104/pp.107.101162>
- Strasser, R., Castilho, A., Stadlmann, J., Kunert, R., Quendler, H., Gattinger, P., Jez, J., Rademacher, T., Altmann, F., Mach, L., Steinkellner, H., 2009. Improved Virus Neutralization by Plant-

- produced Anti-HIV Antibodies with a Homogeneous β 1,4-Galactosylated N-Glycan Profile. *J. Biol. Chem.* 284, 20479–20485. <https://doi.org/10.1074/jbc.M109.014126>
- Strasser, R., Stadlmann, J., Schähns, M., Stiegler, G., Quendler, H., Mach, L., Glössl, J., Weterings, K., Pabst, M., Steinkellner, H., 2008. Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnology Journal* 6, 392–402. <https://doi.org/10.1111/j.1467-7652.2008.00330.x>
- Stukenberg, D., Zauner, S., Dell'Aquila, G., Maier, U.G., 2018. Optimizing CRISPR/Cas9 for the Diatom *Phaeodactylum tricornutum*. *Front Plant Sci* 9, 740. <https://doi.org/10.3389/fpls.2018.00740>
- Subedi, G.P., Barb, A.W., 2016. The immunoglobulin G1 N-glycan composition affects binding to each low affinity Fc γ receptor. *mAbs* 8, 1512–1524. <https://doi.org/10.1080/19420862.2016.1218586>
- Takashima, S., Seino, J., Nakano, T., Fujiyama, K., Tsujimoto, M., Ishida, N., Hashimoto, Y., 2009. Analysis of CMP-sialic acid transporter-like proteins in plants. *Phytochemistry* 70, 1973–1981. <https://doi.org/10.1016/j.phytochem.2009.09.017>
- Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., Muchmore, E., 2003. Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *PNAS* 100, 12045–12050. <https://doi.org/10.1073/pnas.2131556100>
- Taniguchi, N., Miyoshi, E., Ko, J.H., Ikeda, Y., Ihara, Y., 1999. Implication of N-acetylglucosaminyltransferases III and V in cancer: gene regulation and signaling mechanism. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1455, 287–300. [https://doi.org/10.1016/S0925-4439\(99\)00066-6](https://doi.org/10.1016/S0925-4439(99)00066-6)
- Tepedelen, B.E., Kirmizibayrak, P.B., 2019. Endoplasmic Reticulum-Associated Degradation (ERAD). *Endoplasmic Reticulum*. <https://doi.org/10.5772/intechopen.82043>
- Tesson, B., Gaillard, C., Martin-Jézéquel, V., 2009. Insights into the polymorphism of the diatom *Phaeodactylum tricornutum* Bohlin. *Botanica Marina* 52, 104–116. <https://doi.org/10.1515/BOT.2009.012>
- Thiel, C., 2003. Congenital Disorder of Glycosylation (CDG) - Ih. Identifizierung eines Defekts in der Dol-P-Man:Man1-2GlcNAc2-PP-Dol Mannosyltransferase.
- Thobhani, S., Yuen, C.-T., Bailey, M.J.A., Jones, C., 2009. Identification and quantification of N-linked oligosaccharides released from glycoproteins: An inter-laboratory study. *Glycobiology* 19, 201–211. <https://doi.org/10.1093/glycob/cwn099>
- Thomann, M., Schlothauer, T., Dashivets, T., Malik, S., Avenal, C., Bulau, P., Rüger, P., Reusch, D., 2015. In Vitro Glycoengineering of IgG1 and Its Effect on Fc Receptor Binding and ADCC Activity. *PLoS One* 10. <https://doi.org/10.1371/journal.pone.0134949>
- Tokuşoglu, Ö., Ünal, M.K., 2003. Biomass Nutrient Profiles of Three Microalgae: *Spirulina platensis*, *Chlorella vulgaris*, and *Isochrysis galbana*. *Journal of Food Science* 68, 1144–1148. <https://doi.org/10.1111/j.1365-2621.2003.tb09615.x>
- Tran, M., Henry, R.E., Siefker, D., Van, C., Newkirk, G., Kim, J., Bui, J., Mayfield, S.P., 2013. Production of anti-cancer immunotoxins in algae: Ribosome inactivating proteins as fusion partners. *Biotechnol. Bioeng.* 110, 2826–2835. <https://doi.org/10.1002/bit.24966>

- Tran, M., Zhou, B., Pettersson, P.L., Gonzalez, M.J., Mayfield, S.P., 2009. Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. *Biotechnol. Bioeng.* 104, 663–673. <https://doi.org/10.1002/bit.22446>
- Umaña, P., Jean-Mairet, J., Moudry, R., Amstutz, H., Bailey, J.E., 1999. Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat. Biotechnol.* 17, 176–180. <https://doi.org/10.1038/6179>
- Ünligil, U.M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., Rini, J.M., 2000. X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J* 19, 5269–5280. <https://doi.org/10.1093/emboj/19.20.5269>
- van Beers, M.M.C., Bardor, M., 2012. Minimizing immunogenicity of biopharmaceuticals by controlling critical quality attributes of proteins. *Biotechnol J* 7, 1473–1484. <https://doi.org/10.1002/biot.201200065>
- van den Elsen, J.M., Kuntz, D.A., Rose, D.R., 2001. Structure of Golgi alpha-mannosidase II: a target for inhibition of growth and metastasis of cancer cells. *EMBO J.* 20, 3008–3017. <https://doi.org/10.1093/emboj/20.12.3008>
- Vanier, G., Hempel, F., Chan, P., Rodamer, M., Vaudry, D., Maier, U.G., Lerouge, P., Bardor, M., 2015. Biochemical Characterization of Human Anti-Hepatitis B Monoclonal Antibody Produced in the Microalgae *Phaeodactylum tricornutum*. *PLoS ONE* 10, e0139282. <https://doi.org/10.1371/journal.pone.0139282>
- Vanier, G., Lucas, P.-L., Loutelier-Bourhis, C., Vanier, J., Plasson, C., Walet-Balieu, M.-L., Tchi-Song, P.C., Remy-Jouet, I., Richard, V., Bernard, S., Driouich, A., Afonso, C., Lerouge, P., Mathieu-Rivet, E., Bardor, M., 2017. Heterologous expression of the N -acetylglucosaminyltransferase I dictates a reinvestigation of the N -glycosylation pathway in *Chlamydomonas reinhardtii*. *Scientific Reports* 7, 10156. <https://doi.org/10.1038/s41598-017-10698-z>
- Vanier, G., Stelter, S., Vanier, J., Hempel, F., Maier, U.G., Lerouge, P., Ma, J., Bardor, M., 2018. Alga-made anti-Hepatitis B antibody binds to human Fcγ receptors. *Biotechnology Journal*.
- Varki, A., 2001. Loss of N-glycolylneuraminic acid in humans: Mechanisms, consequences, and implications for hominid evolution. *Am. J. Phys. Anthropol.* 116, 54–69. <https://doi.org/10.1002/ajpa.10018>
- 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., Vliegenthart, J.F., Lütteke, T., Perez, S., Bolton, E., Rudd, P., Paulson, J., Kanehisa, M., Toukach, P., Aoki-Kinoshita, K.F., Dell, A., Narimatsu, H., York, W., Taniguchi, N., Kornfeld, S., 2015. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* 25, 1323–1324. <https://doi.org/10.1093/glycob/cwv091>
- Vasilev, N., Smales, C.M., Schillberg, S., Fischer, R., Schiermeyer, A., 2016. Developments in the production of mucosal antibodies in plants. *Biotechnol. Adv.* 34, 77–87. <https://doi.org/10.1016/j.biotechadv.2015.11.002>
- Vézina, L.-P., Faye, L., Lerouge, P., D’Aoust, M.-A., Marquet-Blouin, E., Burel, C., Lavoie, P.-O., Bardor, M., Gomord, V., 2009. Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnology Journal* 7, 442–455. <https://doi.org/10.1111/j.1467-7652.2009.00414.x>

- Vleugels, W., Haeuptle, M.A., Ng, B.G., Michalski, J.-C., Battini, R., Dionisi-Vici, C., Ludman, M.D., Jaeken, J., Foulquier, F., Freeze, H.H., Matthijs, G., Hennet, T., 2009. RFT1 Deficiency in Three Novel CDG Patients. *Hum Mutat* 30. <https://doi.org/10.1002/humu.21085>
- Volkin, D.B., Hershenson, S., Ho, R.J.Y., Uchiyama, S., Winter, G., Carpenter, J.F., 2015. Two Decades of Publishing Excellence in Pharmaceutical Biotechnology. *Journal of Pharmaceutical Sciences* 104, 290–300. <https://doi.org/10.1002/jps.24285>
- Von Horsten, H.H., Ogorek, C., Blanchard, V., Demmler, C., Giese, C., Winkler, K., Kaup, M., Berger, M., Jordan, I., Sandig, V., 2010. Production of non-fucosylated antibodies by co-expression of heterologous GDP-6-deoxy-D-lyxo-4-hexulose reductase. *Glycobiology* 20, 1607–1618. <https://doi.org/10.1093/glycob/cwq109>
- Walsh, G., 2018. Biopharmaceutical benchmarks 2018. *Nature Biotechnology* 36, 1136–1145. <https://doi.org/10.1038/nbt.4305>
- Walsh, G., 2014. Biopharmaceutical benchmarks 2014. *Nature Biotechnology* 32, 992–1000. <https://doi.org/10.1038/nbt.3040>
- Walsh, G., 2002. Biopharmaceuticals and biotechnology medicines: an issue of nomenclature. *European Journal of Pharmaceutical Sciences* 15, 135–138. [https://doi.org/10.1016/S0928-0987\(01\)00222-6](https://doi.org/10.1016/S0928-0987(01)00222-6)
- Walsh, G., Jefferis, R., 2006. Post-translational modifications in the context of therapeutic proteins. *Nature Biotechnology* 24, 1241–1252. <https://doi.org/10.1038/nbt1252>
- Ward, B.J., Landry, N., Trépanier, S., Mercier, G., Dargis, M., Couture, M., D'Aoust, M.-A., Vézina, L.-P., 2014. Human antibody response to N-glycans present on plant-made influenza virus-like particle (VLP) vaccines. *Vaccine* 32, 6098–6106. <https://doi.org/10.1016/j.vaccine.2014.08.079>
- Westerhof, L.B., Wilbers, R.H.P., Raaij, D.R. van, Nguyen, D.-L., Goverse, A., Henquet, M.G.L., Hokke, C.H., Bosch, D., Bakker, J., Schots, A., 2014. Monomeric IgA can be produced in planta as efficient as IgG, yet receives different N-glycans. *Plant Biotechnology Journal* 12, 1333–1342. <https://doi.org/10.1111/pbi.12251>
- Wild, R., Kowal, J., Eyring, J., Ngwa, E.M., Aebi, M., Locher, K.P., 2018. Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science* 359, 545–550. <https://doi.org/10.1126/science.aar5140>
- Wildt, S., Gerngross, T.U., 2005. The humanization of N-glycosylation pathways in yeast. *Nat Rev Micro* 3, 119–128. <https://doi.org/10.1038/nrmicro1087>
- Williams, D.B., 2006. Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. *Journal of Cell Science* 119, 615–623. <https://doi.org/10.1242/jcs.02856>
- Wilson, I.B.H., Rendić, D., Freilinger, A., Dumić, J., Altmann, F., Mucha, J., Müller, S., Hauser, M.-T., 2001. Cloning and expression of cDNAs encoding α 1,3-fucosyltransferase homologues from *Arabidopsis thaliana*1The cDNA sequences referred to in this publication have been deposited with the EMBL database under the numbers AJ404860 (FucTA), AJ404861 (FucTB) and AJ404862 (FucTC).1. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1527, 88–96. [https://doi.org/10.1016/S0304-4165\(01\)00151-9](https://doi.org/10.1016/S0304-4165(01)00151-9)

- Wright, A., Morrison, S.L., 1997. Effect of glycosylation on antibody function: implications for genetic engineering. *Trends in Biotechnology* 15, 26–32. [https://doi.org/10.1016/S0167-7799\(96\)10062-7](https://doi.org/10.1016/S0167-7799(96)10062-7)
- Wycoff, K.L., 2005. Secretory IgA Antibodies from Plants [WWW Document]. <https://doi.org/info:doi/10.2174/1381612054367508>
- Yamaguchi, Y., Barb, A.W., 2020. A synopsis of recent developments defining how N-glycosylation impacts immunoglobulin G structure and function. *Glycobiology* 30, 214–225. <https://doi.org/10.1093/glycob/cwz068>
- Yamaguchi, Y., Nishimura, M., Nagano, M., Yagi, H., Sasakawa, H., Uchida, K., Shitara, K., Kato, K., 2006. Glycoform-dependent conformational alteration of the Fc region of human immunoglobulin G1 as revealed by NMR spectroscopy. *Biochim. Biophys. Acta* 1760, 693–700. <https://doi.org/10.1016/j.bbagen.2005.10.002>
- Yamane-Ohnuiki, N., Kinoshita, S., Inoue-Urakubo, M., Kusunoki, M., Iida, S., Nakano, R., Wakitani, M., Niwa, R., Sakurada, M., Uchida, K., Shitara, K., Satoh, M., 2004. Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. *Biotechnol. Bioeng.* 87, 614–622. <https://doi.org/10.1002/bit.20151>
- Yao, J., Weng, Y., Dickey, A., Wang, K.Y., 2015. Plants as Factories for Human Pharmaceuticals: Applications and Challenges. *Int J Mol Sci* 16, 28549–28565. <https://doi.org/10.3390/ijms161226122>
- Ye, J., Ly, J., Watts, K., Hsu, A., Walker, A., McLaughlin, K., Berdichevsky, M., Prinz, B., Kersey, D.S., d'Anjou, M., Pollard, D., Potgieter, T., 2011. Optimization of a glycoengineered *Pichia pastoris* cultivation process for commercial antibody production. *Biotechnology Progress* 27, 1744–1750. <https://doi.org/10.1002/btpr.695>
- Yet, M.G., Chin, C.C., Wold, F., 1988. The covalent structure of individual N-linked glycopeptides from ovomucoid and asialofetuin. *J. Biol. Chem.* 263, 111–117.
- Yin, Z., Huang, X., 2012. Recent Development in Carbohydrate Based Anti-cancer Vaccines. *J Carbohydr Chem* 31, 143–186. <https://doi.org/10.1080/07328303.2012.659364>
- Yu, C., Gao, K., Zhu, L., Wang, W., Wang, L., Zhang, F., Liu, C., Li, M., Wormald, M.R., Rudd, P.M., Wang, J., 2016. At least two Fc Neu5Gc residues of monoclonal antibodies are required for binding to anti-Neu5Gc antibody. *Sci Rep* 7. <https://doi.org/10.1038/srep20029>
- Yu, M., Brown, D., Reed, C., Chung, S., Lutman, J., Stefanich, E., Wong, A., Stephan, J.-P., Bayer, R., 2012. Production, characterization and pharmacokinetic properties of antibodies with N-linked Mannose-5 glycans. *MAbs* 4, 475–487. <https://doi.org/10.4161/mabs.20737>
- Zaslavskaja, L.A., Lippmeier, J.C., Kroth, P.G., Grossman, A.R., Apt, K.E., 2000. Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *Journal of Phycology* 36, 379–386. <https://doi.org/10.1046/j.1529-8817.2000.99164.x>
- Zeleny, R., Kolarich, D., Strasser, R., Altmann, F., 2006. Sialic acid concentrations in plants are in the range of inadvertent contamination. *Planta* 224, 222–227. <https://doi.org/10.1007/s00425-005-0206-8>

- Zhang, P., Burel, C., Plasson, C., Kiefer-Meyer, M.-C., Ovide, C., Gügi, B., Wan, C., Teo, G., Mak, A., Song, Z., Driouich, A., Lerouge, P., Bardor, M., 2019. Characterization of a GDP-Fucose Transporter and a Fucosyltransferase Involved in the Fucosylation of Glycoproteins in the Diatom *Phaeodactylum tricornutum*. *Front. Plant Sci.* 10. <https://doi.org/10.3389/fpls.2019.00610>
- Zhang, R., Patena, W., Armbruster, U., Gang, S.S., Blum, S.R., Jonikas, M.C., 2014. High-Throughput Genotyping of Green Algal Mutants Reveals Random Distribution of Mutagenic Insertion Sites and Endonucleolytic Cleavage of Transforming DNA. *The Plant Cell* 26, 1398–1409. <https://doi.org/10.1105/tpc.114.124099>
- Zhang, W., Tang, X., Yang, Y., Zhang, X., Zhang, X., 2020. Elevated pCO₂ Level Affects the Extracellular Polymer Metabolism of *Phaeodactylum tricornutum*. *Front. Microbiol.* 11. <https://doi.org/10.3389/fmicb.2020.00339>
- Zhang, X., Lok, S.H.L., Kon, O.L., 1998. Stable expression of human α -2,6-sialyltransferase in Chinese hamster ovary cells: functional consequences for human erythropoietin expression and bioactivity. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1425, 441–452. [https://doi.org/10.1016/S0304-4165\(98\)00095-6](https://doi.org/10.1016/S0304-4165(98)00095-6)
- Zhang, Y., Li, D., Jin, X., Huang, Z., 2014. Fighting Ebola with ZMapp: spotlight on plant-made antibody. *Sci China Life Sci* 57, 987–988. <https://doi.org/10.1007/s11427-014-4746-7>
- Zheng, K., Bantog, C., Bayer, R., 2011. The impact of glycosylation on monoclonal antibody conformation and stability. *mAbs* 3, 568–576. <https://doi.org/10.4161/mabs.3.6.17922>
- Zhou, Q., Shankara, S., Roy, A., Qiu, H., Estes, S., McVie-Wylie, A., Culm-Merdek, K., Park, A., Pan, C., Edmunds, T., 2008. Development of a simple and rapid method for producing non-fucosylated oligomannose containing antibodies with increased effector function. *Biotechnol. Bioeng.* 99, 652–665. <https://doi.org/10.1002/bit.21598>

Résumé / Abstract

Le marché des anticorps monoclonaux représente actuellement la majorité des protéines recombinantes à intérêt thérapeutique. Les difficultés liées à la production de tels anticorps dans des cellules de mammifères, ont poussé la communauté scientifique à envisager des organismes alternatifs pour leur production. Récemment, les microalgues ont montré leur capacité à produire des anticorps monoclonaux fonctionnels. Il est cependant nécessaire de modifier la glycosylation naturelle de ces microalgues afin qu'elles produisent des anticorps pouvant être utilisés pour certaines applications thérapeutiques chez l'Homme. Dans ce contexte, l'étude de la *N*-glycosylation de la diatomée *Phaeodactylum tricornutum* a été réalisée. Cette étude a consisté dans un premier temps à déterminer la structure du précurseur oligosaccharidique des *N*-glycannes des protéines formées dans le réticulum endoplasmique. Les résultats ont permis de montrer la synthèse par la diatomée d'un précurseur tronqué d'un glucose comparé à la majorité des eucaryotes. Dans un second temps, il a été nécessaire de déterminer la structure détaillée des *N*-glycannes oligomannosidiques. Ces résultats ont permis d'établir que les structures Man₉GlcNAc₂ et Man₅GlcNAc₂ synthétisées sont similaires à celles retrouvées chez les mammifères. De plus, le Man₅GlcNAc₂ synthétisé est l'isomère substrat accepteur de la *N*-acétylglucosaminyltransférase I (GnT I) qui catalyse l'étape primordiale initiant la synthèse des glycannes complexes. Cette étude a donc permis d'initier l'humanisation de la voie de la *N*-glycosylation par la sur-expression de la GnT I chez la diatomée afin d'amplifier les proportions en *N*-glycannes complexes portant des résidus *N*-acétylglucosamine en position terminale.

Mots clés : *Phaeodactylum tricornutum*, *N*-glycane, *N*-acétylglucosaminyltransférase I, humanisation, anticorps monoclonaux

Currently, monoclonal antibodies represent the major class of recombinant proteins used for therapeutical applications in humans. The constraints due to the production of these monoclonal antibodies in mammalian cells led scientists to develop alternative expression system for the production of monoclonal antibodies. Recently, microalgae have been shown to be able to produce functional monoclonal antibodies. However, the engineering of the *N*-glycosylation pathway in these microalgae is required in order to optimize the effector functions of the algae-made antibody for some therapeutic applications. In this context, this PhD thesis focused on the study of the *N*-glycosylation of the diatom *Phaeodactylum tricornutum*. Initially, we focused on the structural analysis of the oligosaccharide precursor synthesized in the endoplasmic reticulum. The results showed that *P. tricornutum* synthesizes a truncated precursor as compared to the other eukaryotes. Then, I have investigated the detailed structure of the oligomannosidic *N*-glycans and demonstrated that the Man₉GlcNAc₂ and Man₅GlcNAc₂ synthesized were similar to those found in mammals. In addition, the Man₅GlcNAc₂ synthesized is the acceptor isomer for the *N*-acetylglucosaminyltransferase I (GnT I) which initiates in eukaryotes the synthesis of complex-type *N*-glycans. Humanization of the *N*-glycosylation pathway was finally initiated by over-expression of the GnT I in the diatom in order to increase the amounts of complex *N*-glycans carrying *N*-acetylglucosamine residues in the terminal position.

Key words: *Phaeodactylum tricornutum*, *N*-glycans, *N*-acetylglucosaminyltransferase I, humanization, monoclonal antibodies