



**HAL**  
open science

**SweetBioharm : Sweet engineering of the green  
microalgae *Chlamydomonas reinhardtii* for  
biopharmaceuticals production**

Savignien Leprovost

► **To cite this version:**

Savignien Leprovost. SweetBioharm : Sweet engineering of the green microalgae *Chlamydomonas reinhardtii* for biopharmaceuticals production. Biochemistry, Molecular Biology. Normandie Université, 2024. English. NNT : 2024NORMR079 . tel-04903810

**HAL Id: tel-04903810**

**<https://theses.hal.science/tel-04903810v1>**

Submitted on 21 Jan 2025

**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**

**SweetBioPharm: Sweet engineering of the green microalgae  
Chlamydomonas reinhardtii for biopharmaceuticals production**

Présentée et soutenue par  
**SAVIGNIEN LEPROVOST**

**Thèse soutenue le 16/12/2024**  
devant le jury composé de :

MME MURIEL BARDOR	Professeur des Universités - Université de Rouen Normandie	Directeur de thèse
M. MICHAEL HIPPLER	Professeur des Universités - Université de Munster (ALLEMAGNE)	Co-directeur de thèse
M. CHRISTOPHE BIOT	Professeur des Universités - Université de Lille	Membre du jury
M. STEPHANE LEMAIRE	Directeur de Recherche - Sorbonne Université	Membre du jury
MME ELODIE RIVET	Maître de Conférences - Université de Rouen Normandie	Membre du jury
M. RICHARD STRASSER	Professeur Associé - BOKU University	Rapporteur du jury

Thèse dirigée par **MURIEL BARDOR** (GLYCOBIOLOGIE ET MATRICE EXTRACELLULAIRE VEGETALE) et **MICHAEL HIPPLER** (Université de Munster (ALLEMAGNE))



# THÈSE

Pour obtenir le diplôme de doctorat

Spécialité **ASPECTS MOLECULAIRES ET CELLULAIRES DE LA BIOLOGIE**

Préparée au sein de l'**Université de Rouen Normandie**

**SweetBioPharm**  
Sweet engineering of the green microalgae *Chlamydomonas reinhardtii* for  
biopharmaceutical production

Présentée et soutenue par  
**SAVIGNIEN LEPROVOST**

Thèse soutenue le 16 décembre 2024 devant le jury composé de:		
Mr. Richard STRASSER	Maître de conférences HDR, Institut für Pflanzenbiotechnologie und Zellbiologie (IPBT), Pflanzenglykobiologie, BOKU University (Autriche)	Rapporteur
Mme. Antje VON SCHAEWEN	Professeur, Institut für Biologie und Biotechnologie der Pflanzen (IBBP), Molekulare Physiologie der pflanzen, Universität Münster (Allemagne)	Rapporteur
Mr. Stéphane LEMAIRE	Directeur de recherche CNRS, Biologie Computationnelle et Quantitative UMR 7238 CNRS, Synthetic and Systems Biology of Microalgae, Sorbonne Université (France)	Examineur
Mr. Christophe BIOT	Professeur, Unité de Glycobiologie Structurale et Fonctionnelle UMR8576, Chemical GlycoBiology, Université de Lille (France)	Examineur
Mme. Elodie RIVET	Maître de conférences HDR, Laboratoire de Glycobiologie et Matrice Extracellulaire végétale UR 4358, Université de Rouen Normandie (France)	Co-encadrant
Mme. Muriel BARDOR	Professeur, Laboratoire de Glycobiologie et Matrice Extracellulaire végétale UR 4358, Université de Rouen Normandie (France)	Directeur de thèse
Mr. Michael HIPPLER	Professeur, Institut für Biologie und Biotechnologie der Pflanzen (IBBP), Plant Biochemistry and Biotechnologie, Universität Münster (Allemagne)	Directeur de thèse

Thèse dirigée par **Muriel BARDOR** (GLYCOBIOLOGIE ET MATRICE EXTRACELLULAIRE) et **Michael HIPPLER** (INSTITUT FÜR BIOLOGIE UND BIOTECHNOLOGIE DER PFLANZEN)



Plant Biochemistry and  
Biotechnology  
Institut für Biologie und  
Biotechnologie der  
Pflanzen - IBBP



# SweetBioPharm

Sweet engineering of the green  
microalga *Chlamydomonas reinhardtii*  
for biopharmaceutical production.



To my late beloved Grandpa.





« But you know happiness can be found even in the darkest of times,  
when one only remembers to turn on the light. »

Albus Dumbledore

Harry Potter and the prisoner of Azkaban - J.K. Rowling



# Acknowledgements



First of all, I would like to thank Richard STRASSER and Antje VON SCHAEWEN, associate professor at the BOKU University and professor at the University of Münster respectively, for doing me the honour of being my thesis reviewers.

I am also grateful to Stéphane LEMAIRE, research director at the Sorbonne University, for following me during the last four years as a member of my PhD committee and agreeing to be part of my thesis jury as examiner and president.

I would like to thank Christophe BIOT, professor at the University of Lille, for being part of my defence jury. You've followed me from undergraduate level right up to the present day. Your presence was as obvious as it was dear to my heart.

I would like to extend my warmest thank to my thesis management team for the trust they have placed in me and their help throughout this thesis project.

To Elodie RIVET, Lecturer at the University of Rouen Normandie, for providing me with both direct and remote supervision when I was in Germany. It was a great pleasure to work with you throughout the four years of my thesis.

To Muriel BARDOR, Professor at the University of Rouen Normandie, for your sound advice and your attentiveness, which have been precious to me.

To Michael HIPPLER, Professor at the University of Münster, for teaching me to fight for my ideas and become independent in my research work. Meeting you meant a lot to me.

All of this was also made possible by the financial support of the Normandy region, the University of Münster and the EUR XL-CHEM, to whom I would like to express my infinite gratitude.

I would like to thank Alain MANRIQUE and Quentin LEMERCIER for taking part in the meetings of the thesis committee as members of the Scientific and Pedagogical Committee and as doctoral student representatives, respectively.

Many thanks to Carole PLASSON who was a key player in my thesis through her involvement and efficiency.

Muito obrigado to my friend, André VIDAL-MEIRELES. You taught me a lot about molecular biology, while at the same time being a darkly humorous sidekick.

I would like to express my gratitude to Kyle LAUERSEN from KAUST for welcoming me to his friendly team in the Kingdom of Saudi Arabia and for allowing me to broaden my field of expertise through this enriching experience

Some of my experiments would not have been possible without the help of Anja HEMSCHEMEIER from the University of Bochum. Many thanks to you.

Thank you to the colleagues at the Pissaro and Primacén platforms for your invaluable help and advice over the past 4 years.

I would like to thank all the members of the Glyco-MEV laboratory. These four years in your presence have been full of joy and good humour. Thank you for all the moments of conviviality and for all the discussions.

Thank you to all my dear comrades among the 'youngsters' of the Glyco-MEV, this thesis would have been much less fun without your company.

I would also like to say 'vielen Dank' to all my fellows from the AG HIPPLER in Münster. Your warm welcome and unfailing good mood meant a great deal to me. Thank you also for the many FAB (Feierabendbier) around the pond or at Pinkus.

Felix and Martin, thank you for these little Frenchy moments. You've shown a surprising mastery of this language. Felix, it must be the French consonance of your name.

Afifa, I'll never forget your kindness when I first arrived in Germany. I'll also remember the Pakistani notion of the term 'spicy'.

Yuval, Tzlil, Gali and Mika, you are the family I felt at home with. Thank you for your warm welcome and the moments of conviviality spent together.

I'm full of admiration and gratitude for the people who once shared my office in France or Germany. You've put up with my endless discussions, my complaints and my jokes. Juliette B, Marie-Laure, Charlotte, Younas, Ophélie, Jaber, all my thoughts are with you as I write these lines. I'd like to extend these thanks to the people with whom I didn't share an office, but who also provided psychological support and, depending on the day, acted as an audience. Vincent, Eulalie, Thomas, Pierre, Juliette C, Nesrine, Margaux and many others, you have my endless gratitude.

« Le J, c'est le S. » There's nothing more to say for you to recognise yourself. Your meeting was one of the best.

I'd like to make a special mention of Alexia (even if you can't tell a beer from a wine), Mélanie and Lara, with whom I started my thesis. Since the beginning of our theses, you have been faultless companions who have allowed themselves to be drawn into my space-time rifts of discussion.

I'd also like to thank all the post-docs, trainees and alumni of my two laboratories, as well as my undergraduate students. You can be sure that I will have very fond memories of you all.

This fourth year of thesis would not have been possible without the kind and competent help of Lauriane VAREA. You played a key role in my success.

It is with great emotion that I thank my family, in-laws and the friends who have become family from the bottom of my heart. You have been my support during this long and eventful challenge. Each of you was there in your own way and carried me through to the defence.

Finally, I can't find the words to thank my future wife, Morgan, who has supported me through thick and thin, day and night, while at the same time brilliantly completing her own thesis. You've always believed in me and my ability to complete this project. Writing these words marks the end of a chapter and the beginning of new adventures together.





# Table of content



Illustrations & Tables .....	1
Abbreviations .....	6
Presentation of the candidate .....	11
Introduction .....	16
I.    Biopharmaceutical production.....	18
A.    General information.....	18
II.   Importance of PTMs .....	20
III.  Current production systems and drawbacks.....	22
A.    Mammalian cells .....	22
B.    Non mammalian .....	25
IV. <i>N</i> -glycosylation.....	29
A.    Within the Endoplasmic Reticulum.....	29
B.    Within the Golgi apparatus .....	35
V.   Engineering the <i>N</i> -glycosylation pathway.....	40
A.    Mammals .....	41
B.    Yeast.....	42
C.    Plants.....	45
VI.  Microalgae as alternative .....	49
A.    General information.....	49
B. <i>Chlamydomonas reinhardtii</i> .....	52
C.    Recombinant protein in <i>C. reinhardtii</i> .....	54
D. <i>N</i> -glycosylation.....	57
E.    Glycoengineering in <i>C. reinhardtii</i> .....	61
VII.  Objectives of the thesis.....	64
A.    Characterisation of a therapeutic glycoprotein reporter: human erythropoietin .....	64
B.    Elimination of immunogenic epitopes by genome editing .....	65
C.    Targeting the Golgi apparatus .....	66
Materials & Methods.....	67
I.    Biological Material .....	69
A.    Bacterial strains.....	69
B. <i>Chlamydomonas reinhardtii</i> strains.....	69
II.   Culture conditions .....	71
A. <i>E. coli</i> .....	71
B. <i>C. reinhardtii</i> .....	71

III.	Biochemical analysis .....	73
A.	Whole cell proteins extraction .....	73
B.	Culture medium proteins isolation .....	74
C.	SDS-PAGE .....	74
D.	Western Blot.....	76
E.	<i>N</i> -glycans isolation. ....	78
F.	Filter Aided Sample Preparation (FASP) for glycoproteomic analysis. ....	80
IV.	Mass spectrometry .....	80
A.	<i>N</i> -glycans profile analysis by MALDI-TOF-MS .....	80
B.	LC-ESI MS/MS Analysis .....	81
V.	Molecular biology .....	82
A.	Plasmids .....	82
B.	Genomic DNA extraction.....	84
C.	PCR and nucleic acids electrophoresis.....	85
D.	<i>C. reinhardtii</i> transformation .....	85
E.	CRISPR/Cas9 gene editing protocol .....	86
VI.	Microscopy and fluorescence screening.....	89
A.	Fluorescence screening.....	89
B.	Confocal microscopy.....	89
VII.	<i>In silico</i> analysis.....	90
A.	Transmembrane domain prediction.....	90
B.	Multiple sequence alignment .....	90
C.	Prediction of the three-dimensional structure of proteins .....	90
Chapter 1: Fine-tuning the <i>N</i> -glycosylation of recombinant human erythropoietin using <i>Chlamydomonas reinhardtii</i> mutants.....		91
I.	Preamble.....	93
Chapter 2: Elimination of immunogenic epitopes by genome editing .....		111
I.	Context .....	113
II.	Results.....	116
A.	CRISPR/Cas9 genome editing.....	116
B.	Biochemical analysis .....	121
C.	Mass spectrometry analysis of <i>N</i> -glycans composition .....	123
D.	Acetate requirement .....	126
III.	Conclusion and discussion .....	127
A.	<i>N</i> -glycosylation.....	127
B.	Acetate requirement .....	129

Chapter 3: Targeting the Golgi apparatus.....	130
I. Context .....	132
II. Results.....	134
A. Choice of the CTS domains .....	134
B. Plasmid construction.....	136
C. Expression in strains with compartment markers .....	140
III. Discussion and Conclusion.....	143
A. Choice of the CTS domain .....	143
B. Plasmid construction.....	144
C. Expression in strains with compartment markers .....	145
Conclusion & Perspectives .....	149
I. Conclusion.....	151
A. Chapter I .....	152
B. Chapter II .....	153
C. Chapter III.....	154
II. Perspectives.....	155
General discussion .....	157
I. Complementation of <i>C. reinhardtii</i> N-glycosylation pathway. ....	159
II. To humanise or not to humanise.....	161
III. <i>C. reinhardtii</i> production yield .....	162
IV. What about O-glycosylation? .....	164
Résumé étendu en français .....	169
I. Introduction .....	171
II. Chapitre I .....	174
III. Chapitre II .....	175
IV. Chapitre III.....	176
V. Conclusion et perspectives.....	177
VI. Discussion .....	179
References .....	180



# Illustrations & Tables



Figure 1 : Evolution of the biopharmaceutical product approvals (Walsh and Walsh, 2022). .....	19
Figure 2 : Statistical representation of experimental PTMs of the most abundant proteins modified from (Khoury <i>et al.</i> , 2011). .....	20
Figure 3 : Comparison of therapeutic protein production in mammalian and non mammalian cells (Walsh and Walsh, 2022). .....	22
Figure 4 : Evolution of mAbs approvals and sales (Walsh and Walsh, 2022). .....	23
Figure 5 : : Diagram showing the part of different mammalian cell models used in the manufacture of commercial-scale biopharmaceuticals from 1987 to February 2021 (Al-Majmaie <i>et al.</i> , 2021). .....	24
Figure 6 : Detailed structure of the Glc <sub>3</sub> Man <sub>9</sub> GlcNAc <sub>2</sub> complete oligosaccharide precursor modified from (Breitling and Aebi, 2013). .....	30
Figure 7 : Summary diagram of the reticular stages of <i>N</i> -glycosylation (Stanley <i>et al.</i> 2022). .....	32
Figure 8 : Scheme of Quality control of <i>N</i> -glycosylated proteins in the endoplasmic reticulum from (Suzuki <i>et al.</i> , 2022). .....	34
Figure 9 : Scheme of a type II transmembrane protein adapted from (Loos and Steinkellner, 2014). .....	35
Figure 10 : Structure of the <i>N</i> -glycans of fungal (a), plant (b) and animal (c) proteins, modified from (Wang <i>et al.</i> , 2017). .....	39
Figure 11 : Humanisation of the <i>N</i> -glycosylation pathway in yeast extract from Lucas, 2019 adapted from De Pourcq <i>et al.</i> , 2010. ....	44
Figure 12 : Example of humanisation of the <i>N</i> -glycosylation pathway in plants, extract from P.L Lucas's thesis (Lucas, 2019) modified from (Chen, 2016). ....	46
Figure 13 : Microalgae phylogenetic tree from (Keeling, 2013). ....	49
Figure 14 : Diagram showing the range of value-added molecules derived from microalgae from (Parameswari. and Lakshmi, 2022). ....	50
Figure 15 : <i>C. reinhardtii</i> cell structure (Sasso <i>et al.</i> , 2018). ....	53
Figure 16 : Proposed representation of the <i>C. reinhardtii</i> <i>N</i> -glycosylation pathway. ....	58
Figure 17 : Scheme comparing the <i>N</i> -glycan biosynthesis in land plants and mammals (left) and proposed <i>N</i> -glycosylation pathway in <i>C. reinhardtii</i> (right) modified from (Vanier <i>et al.</i> , 2017)...	60
Figure 18 : Proposed scheme of xylosylation processes in the green microalga <i>C. reinhardtii</i> (Lucas, 2019). ....	62
Figure 19 : Comparison of the amino acid sequences of different plant and <i>C. reinhardtii</i> xylosyltransferases and their phylogenetic relationships modified from (Lucas <i>et al.</i> , 2020). ....	63
Figure 20 : Schematic diagram of the stacking in the cassette for western blot transfer of proteins. ....	76
Figure 21 : Schematic representation of the main <i>N</i> -glycan structures identified in glycoengineered mutants compared with a wild-type strain modified from (Oltmanns <i>et al.</i> , 2020). ....	114
Figure 22 : Scheme of a ribonucleoproteine (RNP) complex acting on genomic DNA. ....	116
Figure 23 : PCR Screening of the CRISPR/Cas9 mutants obtained in the IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> strain (A. and B.) and the UVM4 strain (C., D. and E.). ....	117
Figure 24 : Spot culture of A. IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> triple mutant, B. IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre08.g361250 and C. IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre13.g588750 on TAP agar medium in normal light at 25°C. ....	118
Figure 25 : Sequence alignment of amino acid sequences .....	119
Figure 26 : 3D modelisation predictions of the truncated proteins compared to the corresponding native predicted 3D structure. ....	120
Figure 27 : Immuno blot analysis of proteins from the culture medium of CC125, IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> triple mutant (TM) and the two mutants obtained via genome editing	



IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre08.g361250 (TMΔCre08.g361250) and IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre13.g588750 (TMΔCre13.g588750) .	121
Figure 28 : N-glycan composition of IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> (TM) strain and the IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre08.g361250 (TMΔCre08.g361250) and IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre13.g588750 mutants (TMΔCre13.g588750) .	123
Figure 29 : MALDI-TOF spectra of glycans isolated from whole cell proteins in the IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> strain and the IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre08.g361250 and IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre13.g588750 mutants .	125
Figure 30 : Acetate requirement test comparing spot cultures of A. IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> and B. IM <sub>XTAx</sub> IM <sub>XTBx</sub> IM <sub>FucT</sub> ΔCre08.g361250 on TAP and TP solid media .	126
Figure 31 : Transmembrane domain prediction using deepTHMHMM 1.0 (Hallgren <i>et al.</i> , 2022) .	135
Figure 32 : Localisation of the catalytic domain of the human β(1,4)-galactosyltransferase in fusion with different CTS domains within plant cells Golgi cisternae modified from (Loos and Steinkellner, 2014) .	135
Figure 33 : Lambda scan of CSI FC2D02 .	136
Figure 34 : Schematic representation of the constructs used for expressing proteins consisting in CTS domains fused to a mCherry (λ <sub>Ex</sub> =587nm, λ <sub>Em</sub> =610nm) fluorescent protein .	137
Figure 35 : Amplification scheme of the sequence corresponding to the predicted Cre10.g439500 and Cre16.g661700 CTS domains on gDNA .	138
Figure 36 : Amplification scheme on gDNA of sequence corresponding to the N-terminal part of Cre16.g678997 (XTB) containing the predicted transmembrane helix .	138
Figure 37 : Schematic representation of the plasmid containing sequence corresponding to the N-terminal part of XTB containing the predicted transmembrane helix .	139
Figure 38 : Confocal imaging of A. FC1E02, B. FC2D02 and C. FC1H12 strains from the CSI (Mackinder <i>et al.</i> , 2017) .	140
Figure 39 : Example of 3D confocal imaging of the endogenous XTB N-terminal region in fusion with the mClover fluorescent protein in UVM4 (Scale bar : 1μm) .	142
Figure 40 : Transmembrane domain prediction of Cre10.g439500, Cre16.g661700, Cre09.g391282 (XTA) and Cre16.g678997 (XTB) using TMHMM-2.0 (Krogh <i>et al.</i> , 2001) .	148
Figure 41 : MALDI-TOF spectrum of N-glycans isolated from whole cell proteins in the UVM4 strain with m/z ranging from 950 to 1330 .	167
Figure 42 : MALDI-TOF spectrum of N-glycans isolated from whole cell proteins in the UVM4 strain with m/z ranging from 1340 to 1720 .	168

Table 1 : Recombinant proteins produced by plants which have obtained marketing authorisation as biopharmaceuticals from (Eidenberger <i>et al.</i> , 2023). .....	27
Table 2 : Summary of algae-made biopharmaceutical products with proven biological activity and potential therapeutic use (Dehghani <i>et al.</i> , 2022). .....	51
Table 3 : Recombinant proteins expressed in <i>C. reinhardtii</i> until 2020 modified from (Barolo <i>et al.</i> , 2020).....	56
Table 4 : Concentration of selection antibiotic for <i>E. coli</i> .....	71
Table 5 : Detailed composition of the Tris-Acetate-Phosphate (TAP) medium.....	72
Table 6 : Detailed composition of the Beijerinck’s salts solution used in TAP medium. ....	72
Table 7 : Detailed composition of the phosphate buffer used in TAP medium. ....	72
Table 8 : Detailed composition of traces elements solution used in TAP medium (Hutner <i>et al.</i> , 1950). ....	72
Table 9 : Concentration of selection antibiotic for <i>C. reinhardtii</i> . ....	73
Table 10 : Composition of the protein extraction buffer. ....	73
Table 11 : Composition of acrylamide stacking and separating gels.....	75
Table 12 : Composition of the Loading sample buffer. ....	75
Table 13 : Composition of the Tris-Glycin SDS-PAGE electrophoresis migration buffer. ....	75
Table 14 : Composition of the Tris-Buffered Saline buffer. ....	77
Table 15 : Dilutions of primary antibodies for western blot analysis .....	77
Table 16 : : Primers sequences for CTS domain amplification.....	82
Table 17 : Primers sequences for HiFi assembly .....	84
Table 18 : Sequences of CRISPR/Cas9 targets and primers on each putative XT.....	87
Table 19 : 30 bp FlagV3 primers sequences .....	88
Table 20 : Excitation and emission wavelength used for confocal microscopy analysis. ....	90
Table 21 : Example of a PCR screening of some of the colonies transformed with the pLM006 based plasmids in the CSI strains. ....	141



# Abbreviations



**(NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>, 4H<sub>2</sub>O**: Ammonium molybdate tetrahydrate

**2-AB**: 2-Aminobenzamide

**A. *saitoi*** : *Aspergillus saitoi*

**A. *thaliana***: *Arabidopsis thaliana*

**ACN**: Acetonitrile

**ACTH**: Adenocorticotropic hormone

**ALG**: Asparagine Linked Glycosylation

**APS**: Ammonium persulfate

**Asn**: Asparagine

**AtGnT**: *A. thaliana* N-Acetylglucosaminyltransferase

**AtMNSI**: *A. thaliana* mannosidase I

**ATP**: Adenosyltriphosphate

**AtXylT**: *A. thaliana* xylosyltransferase

**BCA**: BiCinchoninic acid

**BHK**: Baby Hamster Kidney Fibroblast Cells

**bp**: base pair

**C. *reinhardtii*** / **Cr**: *Chlamydomonas reinhardtii*

**C. *vulgaris***: *Chlorella vulgaris*

**CaCl<sub>2</sub>, 2H<sub>2</sub>O**: Calcium chloride dihydrate

**Cas9**: CRISPR associated protein 9

**CDG**: Congenital Disorder of Glycosylation

**cDNA**: Complementary DNA

**CDS**: Coding Sequence

**CHO**: Chinese Hamster Ovary cells

**CLiP**: Chlamydomonas Library Project

**CMAH**: CMP-Neu5Ac hydroxylase

**CMP**: Cytidine monophosphate

**CMP-Sia**: Cytidine monophosphate-sialic acid

**CNX**: Calnexin

**CoCl<sub>2</sub>, 6H<sub>2</sub>O**: Cobalt (II) chloride hexahydrate

**CRISPR**: Clustered Regularly Interspaced Short Palindromic Repeats

**crRNA**: CRISPR RNA

**CRT**: Calreticulin

**CSI**: Chlamydomonas Spatial Interactome

**CTS**: Cytoplasmic, Transmembrane and Stem region

**CuSO<sub>4</sub>, 5H<sub>2</sub>O**: Copper (II) sulfate pentahydrate

**Cys**: Cysteine

**D. *salina***: *Dunaliella salina*

**DHB**: 2,5-dihydroxybenzoic acid

**dHex**: deoxyhexose

**DMSO**: Diméthylsulfoxyde

**DNA**: Desoxyribonucleic Acid

**DTT**: Dithiothréitol

**E. *coli***: *Escherichia coli*

**EDTA**: Ethylenediaminetetraacetic acid

**ELISA**: Enzyme-linked immunosorbent assay

**EPO**: Erythropoietin

**ER**: Endoplasmic reticulum

**ERAD**: ER-associated degradation

**Farnesyl-PP**: Farnesyl pyrophosphate

**FASP**: Filter Aided Sample Preparation

**FDA**: Food and Drug Administration

**FeSO<sub>4</sub>, 7H<sub>2</sub>O**: Ferrous sulfate heptahydrate

**FMDC2A**: Foot-and-mouth disease virus 2A peptide

**Fuc**: Fucose

**FuT/FucT**: Fucosyltransferase

**GAA**: Acid α-glucosidase

**Gal**: Galactose

**GalNAc**: *N*-acetylgalactosamine

**GalT** : Galactosyltransferase

**gDNA**: Genomic DNA

**GFP**: Green Fluorescent Protein

**Glc**: Glucose

**GlcNAc**: *N*-acetylglucosamine

**GlcNAc-1-P**: *N*-acetylglucosamine-1-phosphate

**GMO**: Genetically Modified Organisms

**GnT**: *N*-acetylglucosaminyltransferase

**GRAS**: Generally Recognised As Safe

**gRNA**: Guide-RNA

**GSI/II**: Glucosidase I/II

**GT**: glycosyltransferase

***H. polymorpha***: *Hansenula polymorpha*

**H<sub>2</sub>O**: Water

**H<sub>3</sub>BO<sub>3</sub>**: Boric acid

**HCl**: Chlorhydric acid

**HEK293**: Human embryonic kidney cells

**HeLa**: Henrietta Lacks cells

**hEPO** : Human EPO

**Hex**: Hexose

**HexNAc**: *N*-acetylhexose

**HEXO**: β-*N*-acetylhexosaminidase

**HPLC**: High Performance Liquid Chromatography

**HRP**: Horse Radish Peroxidase

**IgG**: Immunoglobulin G

**IM**: Insertional mutant

**K<sub>2</sub>HPO<sub>4</sub>**: Potassium phosphate dibasic anhydrous

**KH<sub>2</sub>PO<sub>4</sub>**: Potassium dihydrogen phosphate

**LLO**: Lipid Linked Oligosaccharide

**LSB**: Loading Sample Buffer

**mAbs**: Monoclonal antibodies

**MALDI-TOF**: Matrix Assisted Laser Desorption Ionisation – Time Of Flight

**Man**: Mannose

**MeHex**: Methylhexose

**MeOH**: Methanol

**MgSO<sub>4</sub>, 7H<sub>2</sub>O**: Magnesium (II) sulfate heptahydrate

**MnCl<sub>2</sub>, 4H<sub>2</sub>O**: Manganese (II) chloride tetrahydrate

**MNS**: α-mannosidase

**M-Pol**: Mannan polymerase

**MTA**: Material Transfer Agreement

**Na<sub>2</sub>EDTA**: Ethylenedinitrilotetraacetic acid disodium salt dihydrate

**NaBH<sub>3</sub>CN**: Sodium cyanoborohydride

**NaCl**: Sodium Chloride

**NaOH**: Sodium Hydroxyde

**Neu5Ac**: *N*-acetylneuraminic acid

**Neu5Gc**: *N*-glycolylneuraminic acid

**NH<sub>4</sub>Cl**: Ammonium Chloride

**NS0/Sp2/0**: Non secreting murine myeloma cells

**OST**: Oligosaccharyltransferase

***P. pastoris***: *Pischia pastoris*

***P. tricornutum***: *Phaeodactylum tricornutum*

**PCR**: Polymerase Chain Reaction

**PDI**: Protein Disulfide Isomerase

**P-Dol**: Dolichyl-phosphate

**Pent**: Pentose

**PER C6**: Primary human embryonic retinal cells

**PNGase**: Peptide:N-glycosidase

**PP-Dol**: Dolichyl-pyrophosphate

**PTM**: Post translation modification

**RFP**: Red fluorescent protein

**rhEPO**: Recombinant human EPO

**RNA**: Ribonucleic Acid

**RNP**: RiboNucleoProteins

**RnST**: *R. norvegicus* sialyltransferase

**Rpm**: Rotation per minute

**RT**: Room temperature

***S. cerevisiae***: *Saccharomyces cerevisiae*

***S. pombe***: *Scizosacchromyces pombe*

**SDS**: Sodium dodecyl sulfate

**SDS-PAGE**: sodium dodecyl sulfate–polyacrylamide gel electrophoresis

**Ser**: Serine

**Sia**: Sialic Acid

**SialT**: Sialic acid transferase

**siRNA**: Small interfering RNA

***T. reesei***: *Trichoderma reesei*

**TAE**: Tris-Acetate-EDTA buffer

**TALEN**: Transcription activator-like effector nucleases

**TAP**: Tris Acetate Phosphate

**TBS**: Tris Buffered Saline

**TBST**: Tris Buffered Saline + Tween

**TCA**: Trichloroacetic acid

**TFA**: Trifluoroacetic acid

**Thr**: Threonine

**TMD**: Transmembrane domain

**TNF $\alpha$** : Tumour necrosis factor  $\alpha$

**TP**: Tris-Phosphate

**tracrRNA** trans-activating CRISPR RNA

**UDP**: Uridine diphosphate

**UDP-GlcNAc**: Uridine diphosphate *N*-acetylglucosamine

**UGGT1**: UDP-glucose: glycoprotein glucosyltransferase 1

**UV**: Ultraviolet radiation

**WT**: Wild type

**XTA**: Xylosyltransferase A

**XTB**: Xylosyltransferase B

**Xyl**: Xylose

***Y. lipolytica***: *Yarrowia lipolytica*

**YFP**: Yellow fluorescent protein

**ZFNs**: Zinc-finger nucleases

**ZnSO<sub>4</sub>, 7H<sub>2</sub>O**: Zinc sulfate heptahydrate







# Presentation of the candidate







# Savignien LEPROVOST

 11/07/1995  
 25 rue Montgolfier 59700 Marcq en Baroeul  
 +33 (0)640784985  
 savignien.leprovost@gmail.com

## PROFILE

In 2020, I set myself the challenge of completing the Franco-German SweetBioPharm project within the GlycoMEV and IBBP laboratories. I have been confronted through this ambitious project with the interdisciplinarity needed to move forward in the innovative field of metabolic engineering of microalgae for the production of biopharmaceuticals. At the end of these 4 years, I was part of tomorrow's bioproduction through my published work.

I would like to take on a new challenge and put my expertise, skills and experience to good use in research and development specialist roles.

## WORK EXPERIENCE

2023-2024	<b>Temporary Teaching and Research Associate</b> University of Rouen Normandy Tutorials and practical courses in metabolic biochemistry and bio-organic chemistry for biology undergraduates.
13 to 27 May 2023	<b>King Abdullah University Science and Technology, Kingdom of Saudi Arabia.</b> Participation in a project aimed at expressing a glycosylation enzyme in the extremophilic microalga <i>C. merolae</i> , a collaboration between the laboratories of Pr Michael HIPPLER, WWU Münster and Dr Kyle LAUERSEN. Transformation of microalgae using a method optimised by the partner.
2022-2023	<b>Temporary teacher at WWU Münster University (D)</b> Practical teaching in English - Biochemistry, 1st year biology degree. Separation and purification of haemoglobin and cytochrome C by ion exchange chromatography Measurement of the redox properties of cytochrome C.  - Molecular physiology, biochemistry and biotechnology of microalgae, 3rd year biology degree. CRISPR/Cas9 genomic editing in the microalga <i>Chlamydomonas reinhardtii</i> Colony PCR screening.
2020-2022	<b>Temporary teacher at the University of Rouen Normandie</b> Practical teaching Metabolic biochemistry, 3rd year biology degree. Cell fractionation from rat livers and isolation of the various organelles by centrifugation/ultracentrifugation. Protein assay by colorimetry Nucleic acid assay by spectrophotometry Measurement of the redox properties of cytochrome C.
2020-2024	<b>PhD in student – Project manager</b> University of Rouen Normandy / WWU Münster (D) SweetBioPharm project: This project aims to explore and humanize the N-glycosylation pathway in the microalga <i>C. reinhardtii</i> for the production of biopharmaceutical glycoproteins. EUR XL-CHEM (ANR-18-EURE-0020 XL-CHEM).
2020 (6 mois)	<b>Trainee Master 2 DMPK/ADME</b> Bioprojet BIOTECH Development and validation of a method for quantifying a drug candidate using UPLC-MS/MS.

2019 (2 mois)	<b>Trainee</b> GRITA – Mass spectrometry platform EA7365 Simultaneous extraction and quantification by HPLC-MS/MS of bile acids and their C4 precursor
2018 (1 mois)	<b>Trainee</b> UGSF UMR8576 Heterologous production of a potential <i>M. tuberculosis</i> lectin and purification by IMAC chromatography.
2017 (1 mois)	<b>Trainee</b> UGSF UMR8576 Introduction to methodologies for the study of glycoconjugates by MALDI-TOF mass spectrometry.

## EDUCATION AND QUALIFICATIONS

2020 - 2024	<b>PhD in biology, biotechnology</b> University of Rouen Normandy / WWU Münster SweetBioPharm: Sweet engineering of the green microalgae <i>Chlamydomonas reinhardtii</i> for biopharmaceuticals production. EUR XL-CHEM (ANR-18-EURE-0020 XL-CHEM).
2018-2020	<b>Master degree in Chemistry and Life Sciences</b> <b>Bioanalytical Chemistry speciality</b> University of Lille
2015-2018	<b>Bachelor's degree in Life Sciences</b> <b>Biochemistry speciality</b> University of Lille

## SKILLS

<ul style="list-style-type: none"> <li>• Project management</li> <li>• Thematic mobility</li> <li>• Teamwork</li> <li>• Scientific communication</li> <li>• Isolation, purification and maintenance of microalgal and bacterial strains</li> <li>• Cultivation of microorganisms</li> <li>• Genetic transformation of microorganisms</li> <li>• Genome editing using CRISPR/Cas9</li> <li>• Nucleic acid extraction and amplification by PCR, qPCR, RT-PCR</li> <li>• Design and optimisation of plasmid vectors for recombinant protein expression</li> <li>• Extraction and purification of biomolecules by chromatography</li> <li>• Extraction of glycans</li> <li>• Glycoproteomic and glycomic analysis by mass spectrometry</li> <li>• Liquid chromatography coupled to mass spectrometry</li> <li>• Confocal microscopy</li> </ul>
--

## LANGUAGES

<ul style="list-style-type: none"> <li>• French – Native (C2 level)</li> <li>• English – Fluent (C1 level)</li> <li>• Espagnol – Fluent (C1 level)</li> <li>• German – Beginner (A1.1 level)</li> </ul>
---

## PUBLICATION

2024	<b>Leprovost, S., Plasson, C., Balieu, J., Walet-Balieu, M.-L., Lerouge, P., Bardor, M. and Mathieu-Rivet, E. (2024), Fine-tuning the N-glycosylation of recombinant human erythropoietin using <i>Chlamydomonas reinhardtii</i> mutants.</b> Plant Biotechnol. J. <a href="https://doi.org/10.1111/pbi.14424">https://doi.org/10.1111/pbi.14424</a>
------	--

2023	<p><b>International meeting 1 :</b>  <b>The 20th International Conference on the Cell and Molecular Biology of Chlamydomonas. Princeton University, NJ, USA</b>  Poster presentation:  <u>Towards the use of <i>Chlamydomonas reinhardtii</i> as a biopharmaceutical expression system for non-xylosylated <i>N</i>-glycosylated proteins.</u>  Savignien LEPROVOST, André VIDAL MEIRELES, Muriel BARDOR, Elodie MATHIEU-RIVET and Michael HIPPLER</p> <p><b>International meeting 2 :</b>  <b>The 21st European Carbohydrate Symposium. Paris, France</b>  Présentation d'un poster:  <u>Towards the use of <i>Chlamydomonas reinhardtii</i> as a biopharmaceutical expression system for non-xylosylated <i>N</i>-glycosylated proteins.</u>  Savignien LEPROVOST, André VIDAL MEIRELES, Muriel BARDOR, Elodie MATHIEU-RIVET and Michael HIPPLER</p>
2022	<p><b>National meeting :</b>  <b>28<sup>èmes</sup> Journées du Groupe Français des Glycosciences. Branville, France</b>  Poster presentation:  <u>Towards understanding the targeting of glycosyltransferases in the green microalga, <i>Chlamydomonas reinhardtii</i>.</u>  Savignien LEPROVOST, Damien SCHAPMAN, Carole PLASSON, Ludovic GALAS, Christophe DUBESSY, Lara HOEPFNER, Michael HIPPLER, Muriel BARDOR and Elodie MATHIEU-RIVET</p>
2021	<p><b>International meeting :</b>  <b>The 19<sup>th</sup> International conference on the Cell and Molecular Biology of Chlamydomonas. Iles des Embiez, France</b>  Poster presentation:  <u>Xylosylation of protein <i>N</i>-linked glycans in <i>Chlamydomonas reinhardtii</i> is heterogeneous and mediated by a multigene xylosyltransferase family</u>  Savignien LEPROVOST, Pierre-Louis LUCAS, CHAN Philippe, Anne OLTMANNNS, Corinne LOUTELIER-BOURHIS, Carole PLASSON, Carlos AFONSO, Patrice LEROUGE, Narimane MATI-BAOUCHE, Elodie MATHIEU-RIVET, Michael HIPPLER, Muriel BARDOR</p>

## REFERENCES

- Pr. Muriel BARDOR  
GlycoMEV UR4358, URN  
muriel.bardor@univ-rouen.fr  
+33 (0)2 35 14 67 51
- Pr. Dr. Michael HIPPLER  
IBBP, WWU Münster  
mhippler@uni-muenster.de  
+49 251 83-24790
- Dr Elodie MATHIEU-RIVET  
GlycoMEV UR4358, URN  
elodie.rivet@univ-rouen.fr



# Introduction



This chapter will be devoted to a review of the state of the art to provide a better understanding of the SWEETBIOPHARM project (Sweet Engineering of the green microalga *Chlamydomonas reinhardtii* for biopharmaceutical production). Firstly, biopharmaceutical production will be described, in particular its figures and trends, as well as the various production models and their limitations. Next, the importance of *N*-glycosylation as a post-translational modification (PTM) will be presented in general terms, before introducing its engineering depending on the production models. Finally, microalgae will be presented as alternative production models for biomedicines, particularly *Chlamydomonas reinhardtii*.

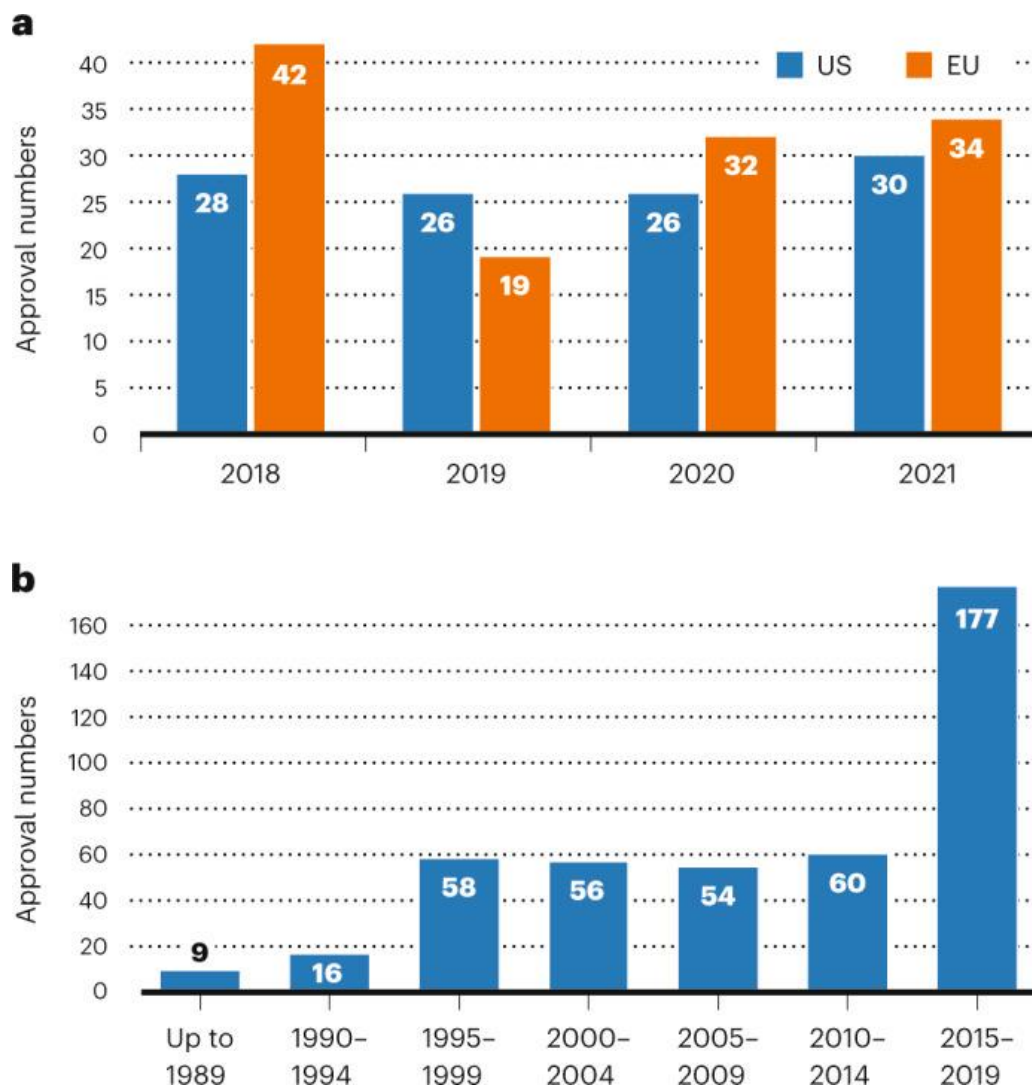
## I. Biopharmaceutical production

### A. General information

In the past few years, biopharmaceuticals took a starring role triggered by the worldwide pandemic of COVID-19 and largely influenced by the world's ageing population, increasing of chronic diseases, and rising incidences of cancers, all of driving the growing worldwide demand for biopharmaceuticals. The sustainability and growth of this market is also ensured by a desire to circumvent the undesirable effects caused by certain treatments based on the use of small synthetic molecules, as well as by sometimes cumbersome surgical solutions. However, the cost of these therapies remains very high due to the constraints associated with their production. Production costs are reflected in the affordability of these drugs, which can cost up to €1600 for 80mg of Humira, also known as adalimumab, a therapeutic monoclonal antibody used in the treatment of Crohn's disease (Arrêté du 23 juillet 2021 modifiant la liste des spécialités pharmaceutiques remboursables aux assurés sociaux). By 2024, the value of this market is predicted to reach \$389.0 billion (Kesik-Brodacka, 2018; O'Flaherty *et al.*, 2020; Walsh and Walsh, 2022).

The term biopharmaceuticals, often associated with biomedicines or biologics, is defined by health authorities as a product isolated from natural sources or containing active substances from a biological source such as cells or living organisms produced by bioengineering or other cutting-edge technologies (Chapitre 1er : Dispositions générales. (Articles L5121-1 à L5121-21) - Légifrance; Les médicaments biosimilaires dans l'UE - Guide d'information destiné aux professionnels de la santé; Research, 2019).

Those medicinal substances purified into active principal ingredients are large molecules, representing above hundred times the size of conventional pharmaceutical molecules, and can vary in their composition and structural complexity (Parkins and Lashmar, 2000). In this manuscript, the term "biomedicines" will refer to all therapeutic products based on recombinant DNA or hybridomas (Walsh, 2002).



**Figure 1 : Evolution of the biopharmaceutical product approvals (Walsh and Walsh, 2022).**  
 a, product approval numbers per year (by product trade name) in Europe compared to USA. b, Number of product approvals in one or both regions between 1989 and 2019.

The use of recombinant DNA has considerably increased the availability of recombinant proteins. Among biomedicines, recombinant therapeutic proteins produced by recombinant DNA technology in different expression systems have been the most widely produced molecules and the ones that have seen the greatest growth on the pharmaceuticals market (Walsh and Walsh, 2022).

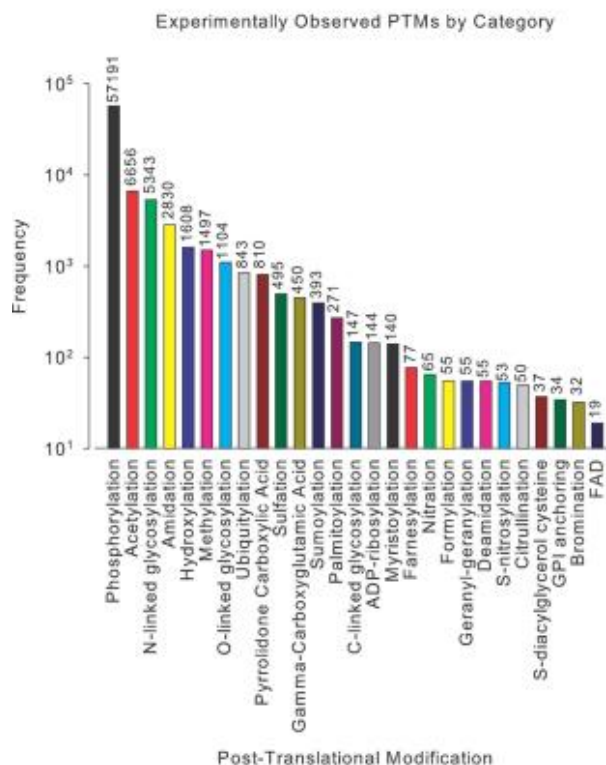


Since the development of recombinant DNA technologies in the 1970s, the approval of recombinant proteins derived from this technology has grown from 9 in 1989 to 177 between 2015 and 2019 in Europe and/or United States of America. More recently, the approval of 197 biopharmaceutical products, including 96 monoclonal antibodies (mAbs), has been observed from January 2018 to June 2022 in both regions (Figure 1) (Walsh and Walsh, 2022).

## II. Importance of PTMs

PTMs involve the covalent addition of groups and/or chemical molecules to the side chain of an amino acid residue. These reversible or irreversible modifications are mediated by enzymes and will determine and/or influence the intrinsic properties of the proteins concerned, such as solubility, cell regulation, subcellular localisation, lifespan, immunogenicity or activity (Ramazi and Zahiri, 2021).

Among the 400 or so PTMs known to date, the most common are phosphorylation, acetylation and glycosylation (Figure 2) (Khoury *et al.*, 2011; Ramazi and Zahiri, 2021).



**Figure 2 : Statistical representation of experimental PTMs of the most abundant proteins modified from (Khoury *et al.*, 2011).**

This histogram represents the prediction frequencies of experimental PTMs present on proteins from the Swiss-Prot database.

Glycosylation is a ubiquitous modification that involves a specific machinery that depends on a wide variety of enzymes catalysing the covalent addition or removal of carbohydrate residues called glycans (Moremen *et al.*, 2012). More precisely, glycosylation is a conserved co-translational and post-translational modification of proteins destined to be secreted or integrated into the plasma membrane in eukaryotic organisms. A better understanding of this modification was made possible by the emergence of glycobiology forty years ago (Olden *et al.*, 1985; Rademacher *et al.*, 1988).

Although aglycosylated proteins exist, this reversible addition of glycans occurs on around 50% of proteins in humans, demonstrating its great importance, particularly in the processes of immune response and cell signal modulation (An *et al.*, 2009). People suffering from disorders in this machinery face severe problems which affect different parts of the body depending on the type of Congenital Disorder of Glycosylation (CDG). As an example, CDG-Ia is the most common one that corresponds to a deficiency in phosphomannomutase 2 (PPM2) catalysing the isomerisation of mannose 6-phosphate to mannose 1-phosphate, a precursor that will be later involved in the *N*-glycosylation processes. This CDG leads to neurological deficiencies, muscular hypotonia and in 20% of the cases, death during the childhood. (Lam and Krasnewich, 1993; Grunewald *et al.*, 2002).

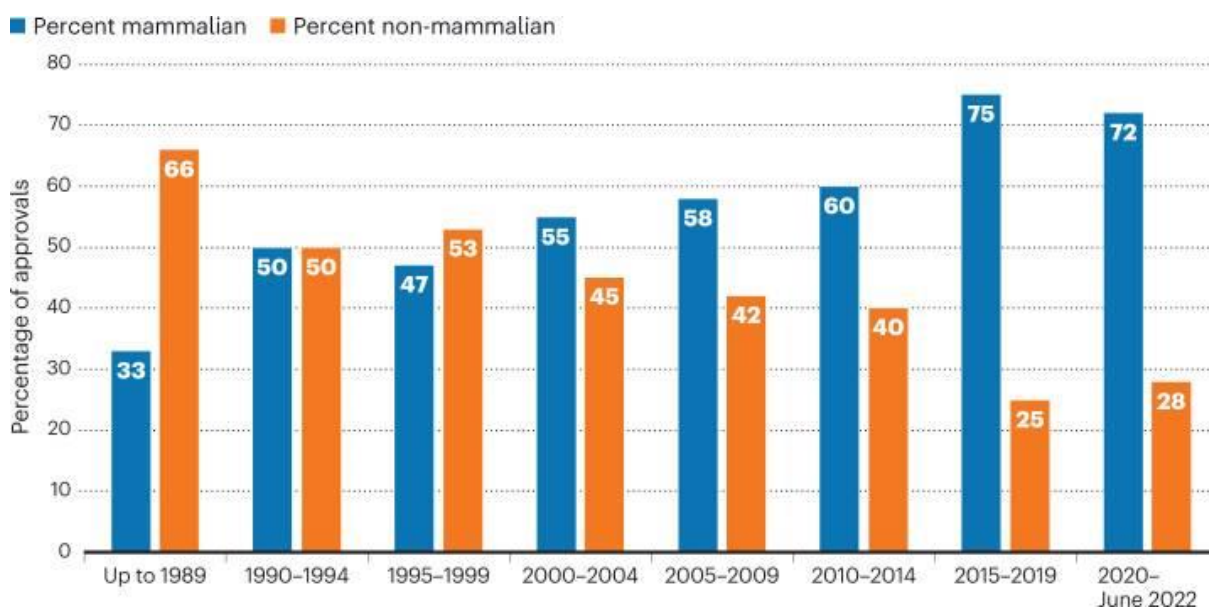
As a consequence, glycosylation is of major importance in biopharmaceutical production, since 70% of the therapeutic proteins produced require glycosylation (Sethuraman and Stadheim, 2006). There are several forms of protein glycosylation, of which *N*- and *O*-glycosylation are the main processes. The distinction is made by the type of linkage involved. In the case of *O*-glycosylation, the linkage is between *N*-acetylgalactosamine (GalNAc) and the oxygen (O) of the hydroxyl group of a serine or threonine in the protein, whereas in the case of *N*-glycosylation, the linkage is between a *N*-acetylglucosamine (GlcNAc) and the nitrogen (N) of an asparagine residue. This manuscript focuses solely on *N*-glycosylation, that will be detailed in part IV of this introduction.

The choice of the production system to meet the PTMs requirements of biopharmaceutical products is crucial as their biological and therapeutical properties depend on these PTMs (Walsh, 2010).

### III. Current production systems and drawbacks

The different expression systems currently used to produce biopharmaceuticals, as well as the alternative production systems being studied to meet the growing demand for these biomedicines and offer less costly production processes, are described in the following section. The main drawbacks of these models will also be explained.

Systems for the heterologous production of recombinant proteins of pharmaceutical interest can be classified into two categories: non-mammalian models and mammalian models. In the 1990s, the two categories were used equally, but since the 2000s the use of non-mammalian models has fallen sharply, with a drop of 25% between the periods from 1995 to 1999 and from 2020 to June 2022 (Figure 3).



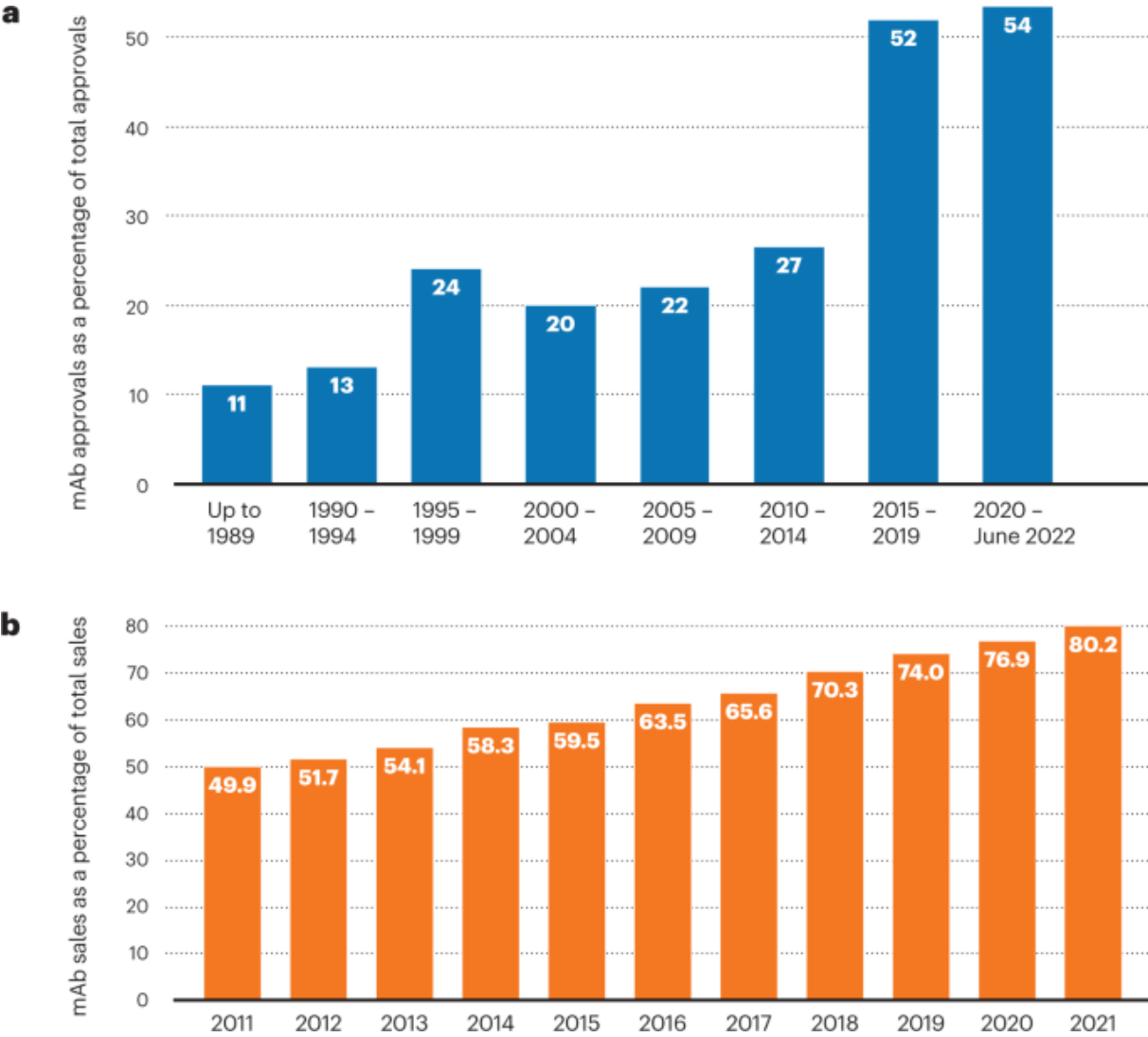
**Figure 3 : Comparison of therapeutic protein production in mammalian and non mammalian cells (Walsh and Walsh, 2022).**

Diagram showing the percentage of therapeutic proteins marketed between 1989 and June 2022 according to the cell type in which they were produced, mammalian in blue or non-mammalian in red.

#### A. Mammalian cells

Since twenty years, the main production system for therapeutical glycoproteins has been and remains mammalian cells producing more than 70% of the molecules authorised on the market between 2015 and June 2022 (Figure 4).

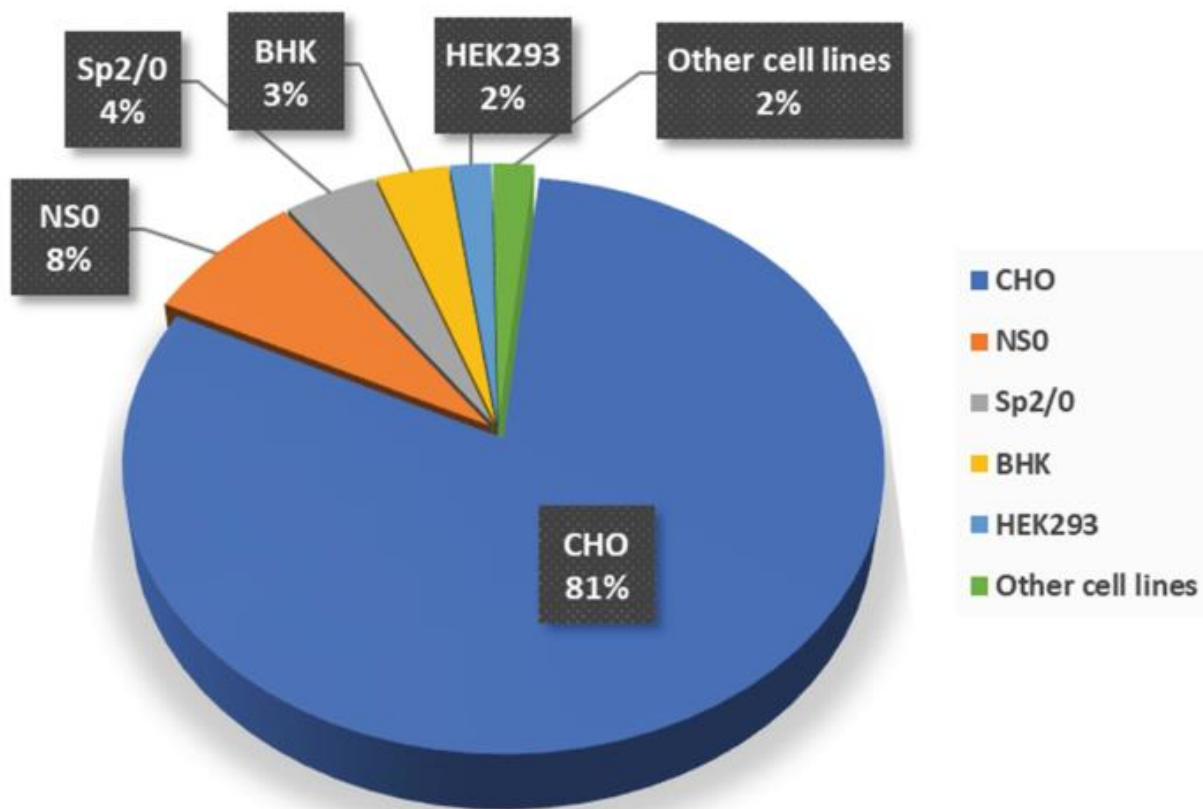
This trend is linked to the growing demand for glycosylated mAbs, which are currently the most widely produced molecules, accounting for more than half of all marketing authorisations between 2020 and 2022 and more than 80% of biopharmaceutical protein sales in 2021 (Figure 4) (Walsh and Walsh, 2022).



**Figure 4 : Evolution of mAbs approvals and sales (Walsh and Walsh, 2022).**  
 a, Trend in newly approved mAbs, expressed as a percentage of the total number of biopharmaceuticals approved for the first time, between 1989 and June 2022. b, Evolution of the global mAbs market over ten years, expressed as a percentage of total global sales of protein-based biopharmaceutical products. Financial data from LaMerie Business Intelligence.

Indeed, mammalian cells were chosen for the production of recombinant therapeutic proteins because of the similarity of their PTMs to human PTMs, as well as their ability to secrete the proteins of interest into the culture medium at concentrations up to 5 g/L (Al-Majmaie *et al.*, 2021).

Mammalian cell lines include non-secreting murine myeloma cells such as NS0 and Sp2/0, Baby Hamster Kidney Fibroblast Cells (*BHK*) established in 1961 by I. A. Macpherson and M. G. P. Stoker (Stoker and Macpherson, 1964), Human cell lines derived from embryonic kidney cells HEK293 as well as other cell lines (Figure 5). These include PER C6 immortalized primary human embryonic retinal cells for example. (Butler, 2005; O’Flaherty *et al.*, 2020; Al-Majmaie *et al.*, 2021; Walsh and Walsh, 2022).



**Figure 5 : : Diagram showing the part of different mammalian cell models used in the manufacture of commercial-scale biopharmaceuticals from 1987 to February 2021 (Al-Majmaie *et al.*, 2021).**

CHO: Chinese Hamster Ovary cells, NS0: Non Ig-secreting murine cells, Sp2/0: Mouse myeloma cell line, BHK: Baby Hamster Kydney cells, HEK293: Human Embryonic Kydney cells.

Among the mammalian cell culture systems used so far, Chinese Hamster Ovary (CHO) cells represent the strongest production system that can reach up to 8 g/L at production scale. CHO cells were used to produce 81% of biopharmaceutical produced in mammal cells between 1987 and 2021 (Figure 5) and this percentage rose to 89% in 2022.

Although animal cells are the main production models because of their resemblance to human cells, they also carry the same risks, particularly those of viral contamination by pathogens (Merten, 2002). Similarly, they have complex and important nutritional needs that require meticulous maintenance. All these disadvantages have a major impact on the cost of producing biopharmaceuticals in these cells, because of the need to use purification systems to avoid contamination. That include viral inactivation steps as well as biotesting. Such requirements increase the production coasts.

Finally, glycosylation can be also a limiting factor, since differences can be observed between animal cells and human cells, such as the presence of the  $\alpha(1,3)$ -galactose epitope, or *N*-glycolylneuraminic acid (Neu5Gc) (Bardor *et al.*, 2005; Ghaderi *et al.*, 2010) which can induce an immune response in humans (van Beers and Bardor, 2012).

## B. Non mammalian

### 1) *Escherichia coli* (*E. coli*)

Since the emergence of biopharmaceutical industry in the 80's, the gram-negative bacteria *E. coli* has been a pioneering expression system for recombinant proteins and became the most popular cell factory, especially the K12 strain for which the National Institutes of Health provided guideline for safety. Indeed, since the successful production of human insulin in 1979 (Chance and Frank, 1993), this prokaryotic organism is still used nowadays with 36 biopharmaceuticals approved between 2018 and 2022 (Walsh and Walsh, 2022). This choice is justified by its well-characterized genetics, the fast growth rate, cost-effectiveness and scalability. This bacteria is an efficient expression system for therapeutic proteins but its lack in sophisticated PTMs including the formation of disulfide bonds and the ability to perform an essential process for the pharmacokinetics and pharmacodynamics of therapeutic proteins, the *N*-glycosylation, are its major limitations (Kamionka, 2011; Huang *et al.*, 2012; Rosano and Ceccarelli, 2014; Selas Castiñeiras *et al.*, 2018).

In contrast, non-mammal eukaryotic organisms, capable of making these PTMs, guarantee better protein stability and activity, thus ensuring the efficacy and safety in medical applications.

## 2) Yeasts

Yeasts have emerged as an interesting platform for the production of biomedicines, notably thanks to their ease of cultivation in bioreactors in inexpensive media, combined with the availability of genetic tools. The development of these tools was made possible by the sequencing of the *Saccharomyces cerevisiae* (*S. cerevisiae*) genome in 1996 (Goffeau *et al.*, 1996) and the *Pischia pastoris* (*P. pastoris*) genome in 2009 (De Schutter *et al.*, 2009). In addition, yeast cultures can reach high cell densities in a short time on an industrial scale, which justifies their use for the production of recombinant proteins such as insulin and analogues since 1987 (Chance and Frank, 1993; Nielsen, 2013; De Wachter *et al.*, 2021; Kulagina *et al.*, 2021).

These eukaryotic unicellular organisms benefit from the GRAS status awarded by the FDA, in particular the baker's yeast *S. cerevisiae*, but also *P. pastoris*, *Yarrowia lipolytica* (*Y. lipolytica*), *Hansenula polymorpha* (*H. polymorpha*) and *Scizosacchomyces pombe* (*S. pombe*) (Madhavan *et al.*, 2021).

Unlike bacteria, they possess the post-translational machinery that enables glycosylation of secreted heterologous proteins of pharmaceutical interest. Therefore, during the period from 2018 to June 2022, nine health products were approved, four of which were produced by *S. cerevisiae* and five by *P. pastoris* (Walsh and Walsh, 2022) some of them being glycosylated such as Eptinezumab, a therapeutic humanised mAb for preventive treatment of migraine (Pederson *et al.*, 2021).

Nevertheless, yeast glycan structures are not identical to human structures. Indeed, it is well known that yeast synthesizes hypermannosylated glycans that can lead to changes in the properties of biopharmaceuticals and, consequently, to undesirable effects such as immunogenic reactions (De Pourcq *et al.*, 2012; Dubey *et al.*, 2023). Moreover, they also face challenges related to protein secretion capacity, inappropriate protein modifications and protease activity. Attempts to overcome the protease activity lead to more fragile strains particularly in *P. pastoris* (Berlec and Strukelj, 2013).

### 3) Plants

Since the expression of human serum albumin correctly processed in potato and tobacco plants in 1990 (Sijmons *et al.*, 1990), plants have emerged as an alternative biopharmaceutical production system to the usual ones. The absence of pathogen cross-contamination between plants and humans is an advantage. What's more, as eukaryotic organisms they are capable of glycosylation (Fischer *et al.*, 2004).

**Table 1 : Recombinant proteins produced by plants which have obtained marketing authorisation as biopharmaceuticals from (Eidenberger *et al.*, 2023).**

Product	Expression host	Application	Company	References
Avidin and $\beta$ -glucuronidase	Transgenic <i>Zea mays</i>	Research reagents	Merck/subsidiary MilliporeSigma	(Hood <i>et al.</i> , 1999)
Allergens, antibodies, cytokines, enzymes, viral-like proteins, viral proteins	Transient <i>Nicotiana benthamiana</i>	Research reagents	Leaf Expression Systems	<a href="https://www.leafexpressionsystems.com/">https://www.leafexpressionsystems.com/</a>
Lectins, growth factors, cytokines	Transient <i>N. benthamiana</i>	Research reagents	iBio, Inc.	<a href="https://ibioinc.com/">https://ibioinc.com/</a>
Growth factors	Transgenic <i>Hordium vulgare</i>	Cosmetics	ORF Genetics	(Magnusdottir <i>et al.</i> , 2013)
Glucocerebrosidase	Stable <i>Daucus carota</i> cells	Enzyme replacement	Protalix Biotherapeutics	(Zimran <i>et al.</i> , 2018)
Pegunigalsidas e alfa	Stable <i>D. carota</i>	Enzyme replacement	Protalix Biotherapeutics	(van der Veen <i>et al.</i> , 2020)
ZMapp	Transient <i>N. benthamiana</i>	Therapeutic antibody	Mapp Biopharmaceutical, Inc.	(Qiu <i>et al.</i> , 2014)
Influenza virus virus-like particles	Transient <i>N. benthamiana</i>	Human vaccine	Medicago, Inc.	(B. J. Ward <i>et al.</i> , 2021)
SARS-CoV-2 virus-like particles	Transient <i>N. benthamiana</i>	Human vaccine	Medicago, Inc.	(Hager <i>et al.</i> , 2022)



In 2012, the first therapeutic molecule produced in plant cells, glucocerebrosidase, produced in carrot cells under the name « Taliglucerase alfa », was approved and commercialised (Table 1) (Haddley, 2012). Thanks to a sound knowledge of molecular biology, it is now possible to design DNA constructs that can be used to transiently produce biomedicines or to obtain stable lines. Strategies can also lead to the storage of molecules of interest in the seeds or other parts of the plant, allowing to production yields to reach the order of g/kg of leaves and production costs well below those generated by production in mammalian cells. (Lobato Gómez *et al.*, 2021; Lee *et al.*, 2023; Eidenberger *et al.*, 2023)

Other biopharmaceutical products, such as vaccines against SARS-COV2, are currently being under pre-clinical or clinical studies with a view to being brought to market for human or veterinarian treatment purposes (Su *et al.*, 2023). The first Canadian vaccine against SARS-COV2 developed by Medicago and approved by Canadian Health can be cited as an example (Hager *et al.*, 2022)

Nonetheless, numerous factors limit the use of plants as an alternative production model. Although authorised on the American continent, in Asia and in Africa, the cultivation of GMOs outdoors for commercial purposes has been banned in Europe since 2008, with the exception of MON810 maize.

Finally, the demonstration of the ability of recombinant plant-produced proteins with core xylose and  $\alpha$ 1-3 fucose to trigger immune responses in rodents and other animals, such as goats, suggests that such responses could also occur in humans (Bardor *et al.*, 2003; Faye *et al.*, 1993; Kurosaka *et al.*, 1991). On one hand, this could be an advantage for vaccination as it can contribute to boost the immune system but, on another hand, for antibodies for example, that can lead to allergy or decrease of the half-life of the antibodies. In the last case, *N*-glycosylation features would therefore hinder their use in human therapy.

## IV. *N*-glycosylation

Protein *N*-glycosylation is the most common co- and post-translational modification in all eukaryotic species, occurring on above 70% of the proteins. This operation takes place in the endoplasmic reticulum (ER) and in the Golgi apparatus, which divide the *N*-glycosylation pathway into two main stages.

### A. Within the Endoplasmic Reticulum

#### 1) The Lipid Linked Oligosaccharide (LLO) synthesis.

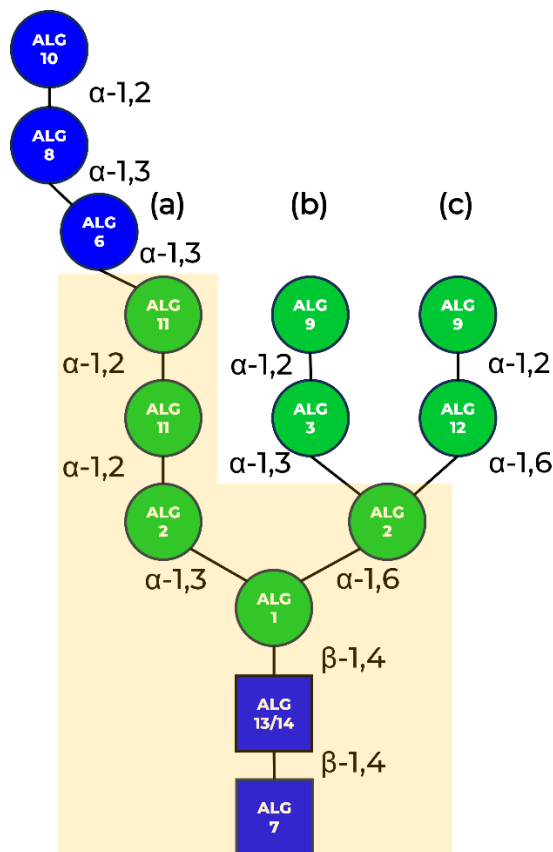
The ER is the starting point for the cytosolic synthesis of the polyprenol carrier dolichyl-pyrophosphate (PP-Dol) composed of 11 to 23 isoprenic units depending on the species (Gryz *et al.*, 2019; Löw *et al.*, 1991; Swiezewska and Danikiewicz, 2005). The lipidic membrane anchor is a derivative of the mevalonate pathway that leads to the farnesyl pyrophosphate (Farnesyl-PP) through enzymatic reactions. This latest molecule is a key element in the biosynthesis of both the cholesterol and non-sterol isoprenoids including dolichol that will undergo a phosphorylation catalysed by the dolichol kinase to lead to the activated Dolichyl-phosphate (P-Dol) (Cantagrel and Lefeber, 2011; Chojnacki and Dallner, 1988; Jones *et al.*, 2009).

Once the anchor has been synthesised, an enzymatic ballet will enable the formation of the LLO, from which the oligosaccharidic moiety is destined to be transferred to the protein finally. Thanks to the presence of uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc), that is synthesised in the cytosol, and the help of the asparagine linked glycosylation (ALG) 7 enzyme, a *N*-acetylglucosamine-1-phosphate (GlcNAc-1-P) is first transferred on the P-Dol. Another GlcNAc residue is added with a  $\beta(1,4)$  linkage to PP-dol-GlcNAc by two enzymes, respectively named ALG 13 and 14 (Figure 6).

Three mannose residues are then added by ALG 1 and 2 (Aebi, 2013). The first mannose is added onto the terminal GlcNAc with a  $\beta(1,4)$  linkage, the two others mannose residues are linked respectively in  $\alpha(1,3)$  and  $\alpha(1,6)$  on the previously added mannose residue opening the door to the formation of a bi-antennated form of the precursor. ALG 11 allows the addition of two further  $\alpha(1,2)$ -mannose residues to create the antenna (a) before the Man<sub>5</sub>GlcNAc<sub>2</sub> motif that is then transferred from the cytoplasm to the lumen of the ER by a flippase, RFT1 (Helenius *et al.*, 2002) and the

synthesis of a glycan precursor continues on the luminal face of the ER membrane (Figure 6).

The addition of four new mannose residues by the intervention of new ALG (3, 9 and 12) forms the triantennal structure  $\text{Man}_9\text{GlcNAc}_2$ . The (b) and (c) antennas are then created by the subsequent addition of two  $\alpha(1,2)$ -mannose residues on each lately added mannoses. Finally, the serial addition of respectively two  $\alpha(1,3)$ -glucoses and one  $\alpha(1,2)$ -glucose by ALG 6, 8 and 10 on branch (a) leads to the synthesis of the  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$  complete oligosaccharide precursor (Figure 6) (Aebi, 2013; Breitling and Aebi, 2013).



**Figure 6 : Detailed structure of the  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$  complete oligosaccharide precursor modified from (Breitling and Aebi, 2013).**

The different enzymes responsible for adding the monosaccharides, as well as the linkages, involved in arms a, b and c are described. GlcNAc residues are represented by blue squares, mannoses, and glucoses by green and blue circles respectively. The orange shadow highlights the part of the structure that is synthesised during the cytosolic phase. The rest is added during the luminal phase of the ER synthesis.

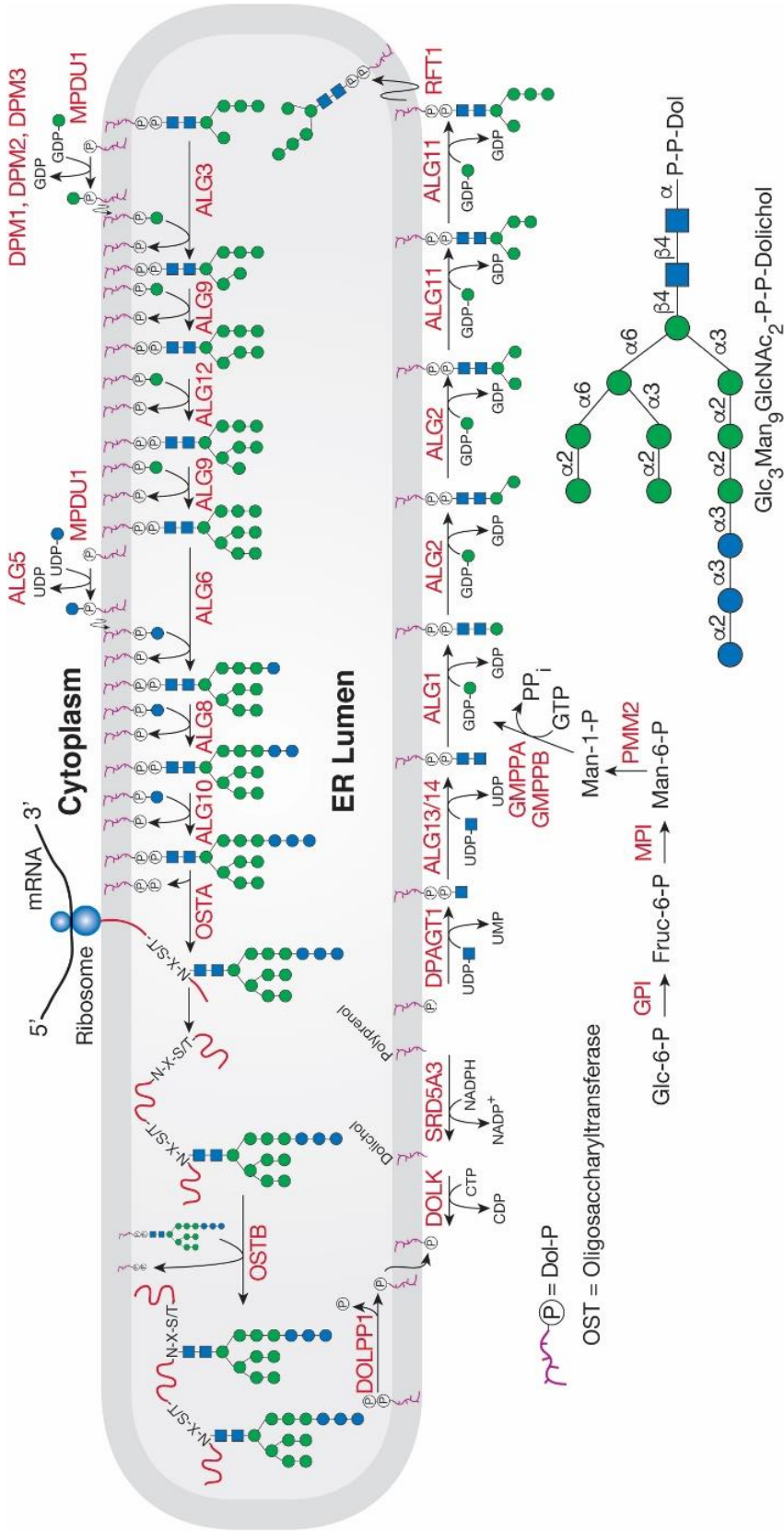
## 2) Oligosaccharyltransferase

The Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> oligosaccharide precursor is then transferred onto a newly synthesised protein or that is being synthesised, by the hetero-oligomeric oligosaccharyltransferase (OST), onto the amide group of an asparagine side chain depending on specific consensus sequences of the Asn-X-Ser/Thr/Cys type, where X can be any amino acid except proline (Kornfeld and Kornfeld, 1985). Indeed, the proline bends the polypeptide chain, thus creating a steric hindrance to the addition of glycan units.

ER membrane located OST enzyme complex is made up of 7 to 8 transmembrane subunits (Mohorko *et al.*, 2011). These sub-units have mainly been characterised in mammals (ribophorin I, DAD1, N33/IAP, OST4, STT3A/STT3B, OST48, ribophorin II) and yeast (OST1, OST2, OST3/OST6p, OST4p, OST5p, STT3p, Wbp1p, Swp1p) (Kelleher and Gilmore, 2006). Among those subunits, one is highly conserved in eukaryotes and correspond to the catalytic subunit called STT3. It exists in two forms in mammals (STT3A and STT3B) and leads to the formation of two isoforms of the OST complex (Figure 7). The isoform containing STT3A will preferentially transfer the glycan precursor to a protein emerging from the translocon through a reaction mediated by DC2 and KCP2 subunits (Shrimal *et al.*, 2017) whereas the isoform containing STT3B will carry out this transfer post-translationally, independently of the translocon (Ruiz-Canada *et al.*, 2009).

To carry out the transfer of the oligosaccharide from the LLO to the nascent glycoprotein, the OST first recognises the glycan motif Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> of the LLO for which it has a high affinity (Figure 7) (Karaoglu *et al.*, 2001). It seems that the efficiency of LLO recognition by OST is linked to the presence of terminal glucose at the non-reducing end but also to the presence of the 2 GlcNAc moieties at the reducing end. Moreover, it has been shown *in vitro* that the absence of the 3 glucose residues reduces glycosylation efficiency by 75%. Paradoxically, it has also been shown that LLOs in which the oligosaccharide units corresponds to GlcNAc<sub>2</sub>Man<sub>5</sub> or GlcNAc<sub>2</sub> reduce glycosylation by only 50% compared to tetradecasaccharide (Ramírez *et al.*, 2022). Therefore, substrates are ranging from Dol-PP-GlcNAc<sub>2</sub> to Dol-PP-GlcNAc<sub>2</sub>Man<sub>9</sub>Glc<sub>3</sub> depending on the species (Kelleher *et al.*, 2007; Schwarz and Aebi, 2011).

Once the oligosaccharide moiety has been transferred from the LLO to the target protein, the newly glycosylated protein undergoes a quality control.



**Figure 7 : Summary diagram of the reticular stages of N-glycosylation (Stanley et al. 2022).**

The synthesis of the lipid anchor carrying the LLO and the transfer of the latter to the protein being synthesised by the OST complex are detailed right up to quality control. The enzymes involved are described, as well as the structure of the LLO and the links between the monosaccharides). In mammalian cells, the OSTA complex is associated with the translocon in the ER membrane and preferentially glycosylates nascent polypeptides in mammalian cells, there are two OST complexes. The OSTA complex, associated with the translocon in the ER membrane, preferentially glycosylates proteins undergoing synthesis that cross the translocon. The OSTB complex modifies proteins found in the lumen of the ER. Enzyme names and abbreviations have been taken from the Human Genome Nomenclature Committee (HGNC).

### 3) Quality control and ERAD system

Quality control is a key process that takes place in the ER. It is now well known that blocking the first step of *N*-glycosylation with tunicamycin, an antibiotic analogue of UDP-GlcNAc leads, among other consequences, to an accumulation of misfolded proteins in the ER, allowing the conclusion that *N*-glycosylation plays an important role in their 3D structure (Helenius and Aebi, 2004; Rudd and Dwek, 1997).

Glucose residues attached with  $\alpha(1,2)$  and  $\alpha(1,3)$  linkage are removed respectively by glucosidase I, a ER type I transmembrane enzyme whose catalytic site is oriented towards the lumen (Shailubhai *et al.*, 1991) and the soluble heterodimeric glucosidase II which contains an HDEL retention signal at its C-terminal end (Trombetta *et al.*, 1996). The monoglucosylated intermediate  $\text{Glc}_1\text{Man}_9\text{GlcNAc}_2$  can therefore be recognised by the lectins calnexin (CNX) and calreticulin (CRT). Both are chaperone proteins which are involved in the quality control cycle (CNX/CRT cycle) by maintaining proteins that are not fully folded in the reticulum (Vassilakos *et al.*, 1998).

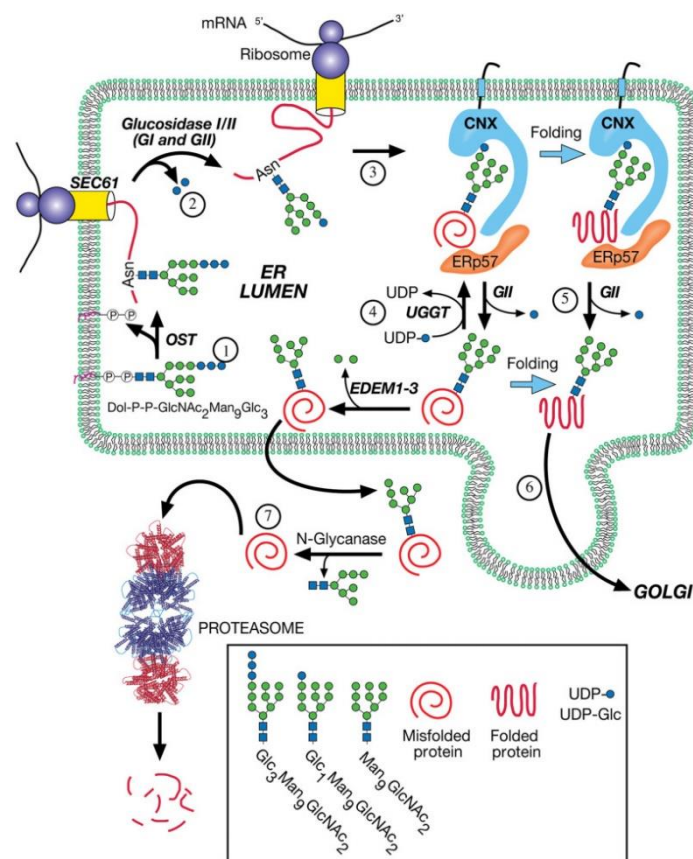
The glycoprotein then becomes a substrate for various enzymes from the Protein Disulfide Isomerase (PDI) family such as the oxidoreductase ERp57 and the peptidyl-prolyl isomerase CypB, which are recruited and help to ensure the correct conformation of the protein by promoting the formation of disulphide bridges and thus preventing the aggregation of misfolded nascent proteins (Molinari and Helenius, 1999).

Properly folded glycoproteins leave the CNX/CRT cycle, the last glucose residue is removed by GSII, preventing any further interaction with the cycle. Then, proteins are transported to the Golgi apparatus via COPII-coated vesicles (Barlowe *et al.*, 1994).

Proteins that remain incorrectly folded will be recognised by a UDP-glucose: glycoprotein glucosyltransferase 1 (UGGT1), which will add an  $\alpha(1,3)$ -glucose on the antenna (a) (Figure 8) that will promote a new interaction with the CNX/CRT cycle to increase the chances that these proteins become correctly folding (Ellgaard and Helenius, 2003; Ruddock and Molinari, 2006; Sousa *et al.*, 1992).

Proteins which, after several cycles, have still not acquired the correct conformation are retained to undergo the ER-associated degradation (ERAD). Firstly, the antenna (b) of the glycan is trimmed by the membrane-bound class I  $\alpha(1,2)$ -mannosidase I, which leads to an 8-mannose residue  $\text{Man}_8\text{GlcNAc}_2$  (Herscovics, 2001).

The loss of one mannose will reduce affinity with UGGT1 and another mannose residue will be removed on the antenna (c). The  $\text{Man}_7\text{GlcNAc}_2$  residue thus formed can be recognised by lectins such as OS9 and XTP3-B in mammals or their yeast counterpart Yos9 promoting the degradation of the protein by the ERAD machinery. Misfolded proteins are then translocated to the cytosol and ubiquitin-tagged before being recognised by the ATPase p97/ VCP (Cdc48 in yeast) complex that extract the protein out of the ER membrane. The protein is then deglycosylated by *N*-glycanase if necessary and degraded into peptides by the proteasome (Hebert and Molinari, 2007; Helenius and Aebi, 2004; Kostova and Wolf, 2003).



**Figure 8 : Scheme of Quality control of *N*-glycosylated proteins in the endoplasmic reticulum from (Suzuki *et al.*, 2022).**

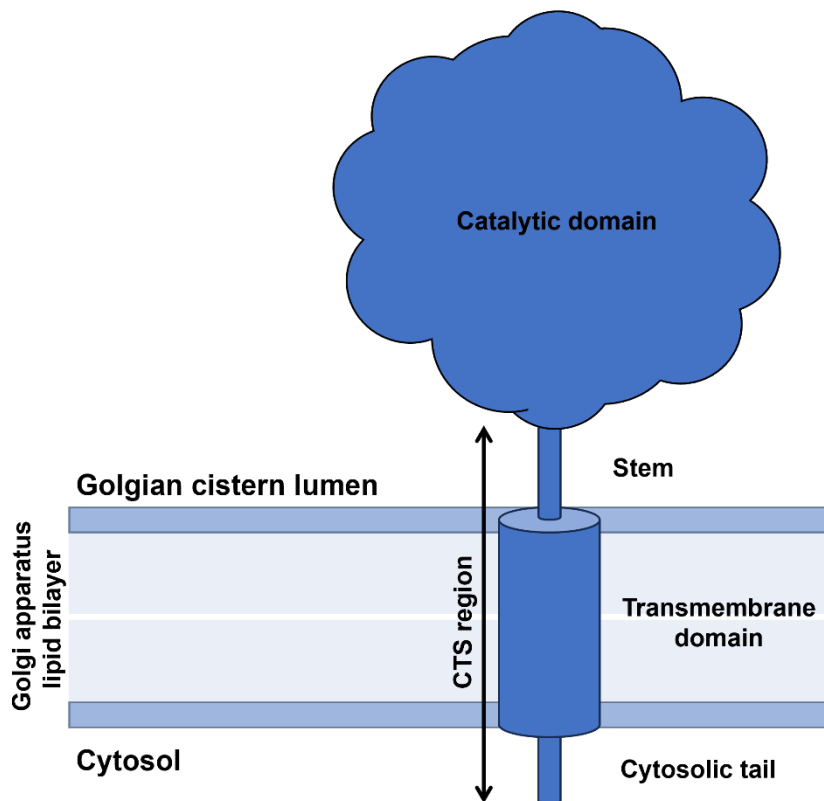
After receiving the LLO from the OST complex, the protein will be recognised by the calnexin/calreticulin complex, which has an affinity for monoglucosylated glycans.

## B. Within the Golgi apparatus

The process continues in the Golgi apparatus through maturation stages that lead to the synthesis of complex *N*-glycans. These different stages involve two main types of enzymatic players, glycosyltransferases, and glycosidases. Many of them are type II transmembrane proteins that enable monosaccharides to be added to or removed from an oligosaccharide, respectively (Loos and Castilho, 2015).

### 1) Type II transmembrane proteins

Type II transmembrane proteins share an *N*-terminal organisation involving a cytoplasmic tail, a single transmembrane domain and a stem region (CTS region). The C-terminal substrate specific catalytic domain is oriented towards the lumen of the Golgian cisternae (Figure 9) (Paulson and Colley, 1989).



**Figure 9 : Scheme of a type II transmembrane protein adapted from (Loos and Steinkellner, 2014).** A *N*-glycosyltransferase is here represented with the C-terminal catalytic domain directed to the Golgi lumen and the *N*-terminal CTS region integrated in the Golgi cistern membrane. The *N*-terminal cytosolic tail is most of the time composed of basic amino acids whereas the transmembrane domain (TMD) is mostly composed of hydrophobic amino acids. The length of the TMD can vary depending on the composition of the lipid bilayer. The stem is rather a disordered region linked to the catalytic domain.



The complex addressing mechanisms of golgian glycoenzymes have been the subject of much research since the 1980s and is still not completely clear (Colley, 1997; Welch and Munro, 2019).

However, it has been demonstrated that sub-localisation within the *cis*, *medial* and *trans* cisternae of the Golgi apparatus is dependent on the CTS domain but also on the cell type as shown by numerous studies such as those by Catherine Rabouille and collaborators in HeLa cells (Rabouille *et al.*, 1995) or Jennifer Schoberer and Richard Strasser in plant cells (Schoberer and Strasser, 2011). Indeed, the composition in amino acids of the CTS domain and its length has been shown to be correlated to the thickness and the lipid composition of the organelle membrane (Schoberer *et al.*, 2019; Sharpe *et al.*, 2010).

The Golgi stages differ between mammals, plants and yeast due to an enzymatic diversity specific to each lineage. Indeed, each species has its own order of assembly with an overlapping distribution of the enzymatic repertoire across the Golgi apparatus enabling the *N*-glycan maturation machinery to function like a sequential assembly line and thus leading to a diversity of complex mature *N*-glycans (Wang *et al.*, 2017).

## 2) Mammal *N*-glycans

In mammals, the Golgi stages begin with the action of  $\alpha$ -mannosidase I (MNSI), which removes 3 to 4 mannose residues from the glycan to obtain  $\text{Man}_5\text{GlcNAc}_2$ . This is the substrate for *N*-acetylglucosaminyltransferase I (GnTI), which uses UDP-GlcNAc as a donor to add a  $\beta(1,2)$ -GlcNAc residue onto the  $\alpha(1,3)$ -mannose thus completing the maturation of the (a) branch of the glycan core. The GnTI is a key enzyme located in the *medial-Golgi* paving the way for the synthesis of so called hybrid and complex glycans in most eukaryotes (von Schaewen *et al.*, 1993).

Two mannose residues with  $\alpha(1,3)$  and  $\alpha(1,6)$  linkage are also respectively removed from branches (b) and (c) by the  $\alpha$ -mannosidase II (MNSII). Finally, the *N*-acetylglucosaminyltransferase II (GnTII) adds a  $\beta$ 1-2 GlcNAc residue, thus completing the synthesis of the biantennated  $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$  glycan core common to most eukaryotic species. There are also tri- and tetra-antennated motifs which structures are the result of the addition of GlcNAc residues to the  $\alpha(1,3)$  and  $\alpha(1,6)$ -mannoses by two different GnT (GnTIV and GnT V). The GnTIII transfers a bisecting GlcNAc to the

core mannose. It is worth noting that the bisecting GlcNAc has been reported to inhibit the terminal modifications of *N*-glycans (Nakano *et al.*, 2019).

The complex glycans are then matured by the addition of  $\beta(1,4)$ -galactose moieties to the two terminal GlcNAc, catalysed by a  $\beta(1,4)$ -galactosyltransferase (GalT), resulting in structure ranging from  $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$  to  $\text{Gal}_4\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$ . This galactosylated glycan is the substrate of two sialyltransferases,  $\alpha(2,6)$ -sialyltransferase ( $\alpha(2,6)$ -SialT) and  $\alpha(2,3)$ -sialyltransferase ( $\alpha(2,3)$ -SialT), which use cytosine monophosphate-sialic acid (CMP-Sia) as donor and galactose residues as acceptors (Figure 10c). It is worth noting that the SialTs are divided into four families depending on the receptor and the glycosidic linkage through the vertebrate lineage. In mammals, a subdivision of these families has been done according to the sequence similarities resulting in 20 sub-families thus explaining some species variability regarding the level or the type ( $\alpha(2,6)$ - or  $\alpha(2,3)$ - linkage) of sialylation (Harduin-Lepers, 2023). As an example, Human has both  $\alpha(2,3)$ -SialT and  $\alpha(2,6)$ -SialT activities with at least 18 enzymes ensuring the sialylation whereas CHO cells are lacking the  $\alpha(2,6)$ -SialT activity (Lee *et al.*, 1989; Harduin-Lepers, 2023).

Finally, an  $\alpha(1,6)$ -fucosyltransferase (FuT8) completes the maturation process by transferring an  $\alpha(1,6)$ -fucose residue to the proximal GlcNAc of the chitobiose moiety. However, this enzyme is capable of processing less complex glycans and thus of intervening earlier in the golgian maturation chain (Fisher *et al.*, 2019).

### 3) Yeast *N*-glycans

Almost all the studies presented in this section were carried out on *S. cerevisiae*. The other yeast species whose *N*-glycosylation has been studied have shown a similar function, with a few species-specific particularities such as the nature of the osidic bond or the presence of moieties that are surprising in yeast, such as  $\alpha(1,3)$ -galactose residues in *S. pombe* (Gemmill and Trimble, 1999).

Yeasts essentially synthesise oligomannosidic glycans, which can be divided into two categories: inner core glycans and hyperglycosylated glycans mostly harboured on cell wall glycoproteins. Proteins entering the Golgi from the ER bear  $\text{Man}_8\text{GlcNAc}_2$  type *N*-glycans in *S. cerevisiae* that are not the substrate of  $\alpha$ -mannosidases but rather of mannosyltransferases or mannosylphosphate

transferases. The first Golgian step is therefore the addition of a first mannose by an  $\alpha(1,6)$ -mannosyltransferase called Och1 (Nakayama *et al.*, 1992). The resulting glycan is the common structure between the two types of glycans found in yeast.

Core-type structures receive  $\alpha(1,2)$ -linked mannose on  $\alpha(1,6)$ -mannose added by Och1p *via* an enzyme that has not been characterised to date, followed by mannosylphosphate grafting by Mnn4p and Mnn6p. Finally,  $\alpha(1,3)$ -linked terminal mannoses are added by Mnn1p (De Pourcq *et al.*, 2010).

Hypermannosylated glycans (Figure 10a) are composed mainly of mannan chains, which are added *via* two enzyme complexes called mannan polymerase I and II (M-Pol I and M-pol II) and reaching up to 200 mannose residues (Stolz and Munro, 2002).

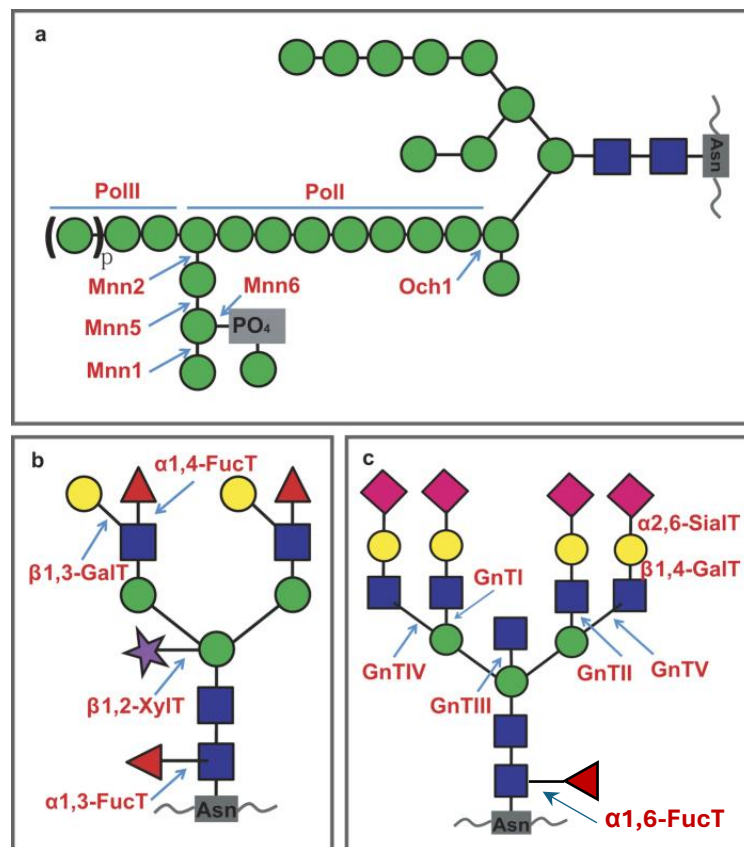
The *cis-Golgi* located M-Pol I complex is made up of two mannosidases, Mnn9p and Van1p, which are responsible for adding 10 to 15  $\alpha(1,6)$ -mannose units. The M-Pol II complex is a heteropentamere composed of subunits, called Mnn9p, Anp1p, Mnn10p, Mnn11p and Hoc1p, that will extend the backbone mannan chains with more  $\alpha(1,6)$ -linked mannose residues (Jungmann *et al.*, 1999; Jungmann and Munro, 1998).

Finally, the backbone is branched by other mannosyltransferases acting successively, such as Mnn2p, Mnn5p and Mnn1p, the first two being responsible for adding  $\alpha(1,2)$ -man before the last adds  $\alpha 1$ -3man (Rayner and Munro, 1998; Wiggins and Munro, 1998). Mannosylphosphates can also be grafted by two enzymes called Mnn4p et Mnn6p (Odani *et al.*, 1997; Wang *et al.*, 1997)

#### 4) Plant *N*-glycans

In plants, the Golgi stages of *N*-glycosylation are initiated by the removal of  $\alpha(1,2)$ -mannoses by MNSI, which provides GnTI with its substrate. As explained above, GnTI initiates the maturation steps that lead to hybrid and complex glycans (von Schaewen *et al.*, 1993). The action of GnTI precedes that of alpha mannosidase II (ManII) (Strasser *et al.*, 2006), then GnTII grafts a second GlcNAc to form the glycan GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>. These steps are identical to those previously detailed for mammalian cells (Strasser, 2014).

However, plants also have maturation stages that are absent in mammals. Plant glycans are xylosylated with a  $\beta(1,2)$  linkage. This step is carried out by a  $\beta(1,2)$ -xylosyltransferase that use UDP-xylose as donor to act on glycans bearing one or two terminal GlcNAc (Bencúr *et al.*, 2005; Kajiura *et al.*, 2012). This enzyme, unique in *A. thaliana*, can therefore accept as a substrate  $\text{GlcNAcMan}_5\text{GlcNAc}_2$ ,  $\text{GlcNAcMan}_3\text{GlcNAc}_2$  and  $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$  and even compete with the redundant fucosyltransferases FUT11 and FUT12 (Strasser, 2016). The two fucosyltransferases are responsible for the addition of core  $\alpha(1,3)$ -fucose (Strasser *et al.*, 2004; Wilson *et al.*, 2001). Action of these enzymes leads to the formation of  $\text{GlcNAc}_2\text{Man}_3\text{XylFucGlcNAc}_2$  complex glycans or  $\text{GlcNAc}_1\text{Man}_5\text{XylFucGlcNAc}_2$  hybrid glycans.



**Figure 10 : Structure of the *N*-glycans of fungal (a), plant (b) and animal (c) proteins, modified from (Wang *et al.*, 2017).**

*N*-glycans are attached to asparagine residues in the consensus sequence AsN-X-Ser/Thr/Cys. Abbreviations in red are enzymes for steps involved in lineage-specific modifications of *N*-glycans. a) All enzymes are mannosyltransferases; b) GalT: Galactosyltransferase, FucT: Fucosyltransferase, XylIT: Xylosyltransferase; c) GnTI: *N*-acetylglucosaminyltransferase I, GnTII: *N*-acetylglucosaminyltransferase II, SialTs: Sialyltransferases, GnTIII: *N*-acetylglucosaminyltransferase III, GnTIV: *N*-acetylglucosaminyltransferase IV, GnTV: *N*-acetylglucosaminyltransferase V, GalT: Galactosyltransferase.

Complex glycans can be further processed by a galactosyltransferase (GalT) that graft a  $\beta(1,3)$ -galactose on the terminal GlcNAc (Strasser, Bondili, Vavra, *et al.*, 2007) and a fucosyltransferase (FUT13) responsible for the addition of a  $\alpha(1,4)$ -fucose residue onto the  $\beta(1,3)$ -galactose previously added (Figure 10b) (Léonard *et al.*, 2002). This results in the presence of Lewis<sup>a</sup> in plants (Fitchette-Lainé *et al.*, 1997; Lerouge *et al.*, 1998; Melo *et al.*, 1997).

Finally, a last type of glycan is found in plants, particularly in vacuoles, resulting from the removal of terminal GlcNAc on complex-type glycans that have previously been xylosylated and fucosylated on the core or not. So-called paucimannosidic glycan structures ( $\text{Man}_3\text{GlcNAc}_2$  and  $\text{Man}_3\text{XylFucGlcNAc}_2$ ) are then generated *via* the action of the HEXO enzyme family composed of 3  $\beta$ -*N*-acetylhexosaminidases (HEXO1, HEXO2 and HEXO3). In both *A. thaliana* and *N. benthamiana*, the vacuolar HEXO 1 and the apoplastic HEXO3 are reputed to be involved in the formation of paucimannosidic *N*-glycans through their action on complex *N*-glycans whereas the membrane-bound HEXO2 is rather known to act on chitotriose structures and has been shown recently to be active on apoplastic glycoproteins with a specificity for terminal GalNAc (Strasser, Bondili, Schoberer, *et al.*, 2007; Gutternigg *et al.*, 2007; Liebminger *et al.*, 2011; Castilho *et al.*, 2014; Shin *et al.*, 2017; Alvisi *et al.*, 2021).

## V. Engineering the *N*-glycosylation pathway

Despite the ability of the aforementioned heterologous eukaryotic models to introduce the *N*-glycans that are essential for protein conformation, species-specific maturation represents a barrier to the use of these recombinant molecules in a therapeutic context. A species-specific glycoepitope could be recognised as 'non-self' by the human organism and could become immunogenic.

Glycoengineering has made it possible to avoid this pitfall by making minor or major modifications depending on the extent of the differences in *N*-glycosylation between production systems and humans. Glycoengineering consists of remodelling or optimizing by adding or inactivating enzymes involved in the *N*-glycosylation pathway. This work has been made easier by the emergence of genome editing tools over the last twenty years. The transcription activator-like effector nucleases (TALENs) (Christian *et al.*, 2010) and zinc-finger nucleases (ZFNs) (Bibikova *et al.*, 2002) were

the first user friendly tools before CRISPR/Cas9 (Doudna and Charpentier, 2014; Steentoft *et al.*, 2014) became democratised. The resulting humanisation of the *N*-glycosylation pathways can therefore improve drug metabolism and pharmacokinetics.

## A. Mammals

Mammalian cells are, by definition, the most similar to humans in terms of *N*-glycosylation. Glycoengineering in these cells has therefore been aimed more at improving pharmacokinetic parameters, particularly in CHO and HEK models, for the production of particular types of biomedicine depending on their purpose and in the light of human metabolism (Narimatsu *et al.*, 2021).

The first signs of divergence between mammals in general and humans were observed during immunogenic reactions following xenotransplantation. An in-depth study of the glycan profiles revealed epitopes responsible for these reactions, such as galactose linked in  $\alpha(1,3)$ , called  $\alpha$ -galactose epitope (Cooper *et al.*, 1993; Oriol *et al.*, 1993). Up to 1% of antibodies specifically directed against these epitopes were found in human serum (Sandrin *et al.*, 1993). This is because humans and certain primates differ from other mammals in the absence of  $\alpha(1,3)$ -galactosyltransferase, an enzyme allowing the transfer of an  $\alpha(1,3)$ -galactose onto the  $\beta(1,4)$ -linked galactose of the *N*-glycan (Galili *et al.*, 1988). These residues were also found in one of the biopharmaceutical production models, CHO cells (Bosques *et al.*, 2010).

In response to this discovery, a successful inactivation of the  $\alpha(1,3)$ -galactosyltransferase have been performed in mice (Tearle *et al.*, 1996), enabling complete elimination of the immunogenic galactose  $\alpha(1,3)$ -galactose epitope and thus opening up new possibilities for xenografts and the production of biotherapeutics such as antibodies (Basnet *et al.*, 2010).

*N*-glycolylneuraminic acid (Neu5Gc) have also been identified as an immunogenic residue for humans. This sialic acid is a hydroxylated derivative of *N*-acetylneuraminic acid (Neu5Ac) that has been processed by an enzyme eliminated in humans during evolution (Peri *et al.*, 2018). However, it has been demonstrated that human cells can incorporate Neu5Gc from heterologous source such as the culture media (Bardor *et al.*, 2005). Moreover, although deprived of the ability to synthesise this immunogenic sialic acid, human tissues have been detected with Neu5Gc due to

this incorporation mechanism and thus this phenomena can occur as well during the production process of biologics if no attention are paid to the culture media (Tangvoranuntakul *et al.*, 2003). As this epitope was found on recombinant proteins expressed in murine cells (Ghaderi *et al.*, 2010), it was necessary to use glycoengineering to inactivate or at least reduce the enzymatic activity of CMP-Neu5Ac hydroxylase (CMAH). This was achieved in CHO cells using antisense ribonucleic acid (RNA) directed against the gene encoding this enzyme or genome editing tools, reducing the level of terminal Neu5Gc from 4% to 1% (Chai *et al.*, 2020; Chenu *et al.*, 2003). Finally, the use of these glycoengineered strains in modified culture media containing, in particular, human Neu5Ac could significantly reduce the presence of Neu5Gc on the proteins produced (Bardor *et al.*, 2005; Borys *et al.*, 2010; Ghaderi *et al.*, 2010).

In addition to the nature of the sialic acids, the linkage of these terminal residues on the *N*-glycans differs in humans compared with the mammalian expression systems used for biopharmaceutical production. In humans, sialic acids are predominantly bound with a  $\alpha(2,6)$ -linkage, with a minority bound with a  $\alpha(2,3)$ -linkage. The latter residue is predominantly found on glycans *N*-linked to proteins produced in animal cells (Lee *et al.*, 1989). To overcome this problem, the gene encoding  $\alpha(2,6)$ -sialyltransferase was expressed in the CHO cell genome, thereby modifying the glycan profile (Bragonzi *et al.*, 2000; Fukuta *et al.*, 2000; Houeix and Cairns, 2019; Lin *et al.*, 2015).

## B. Yeast

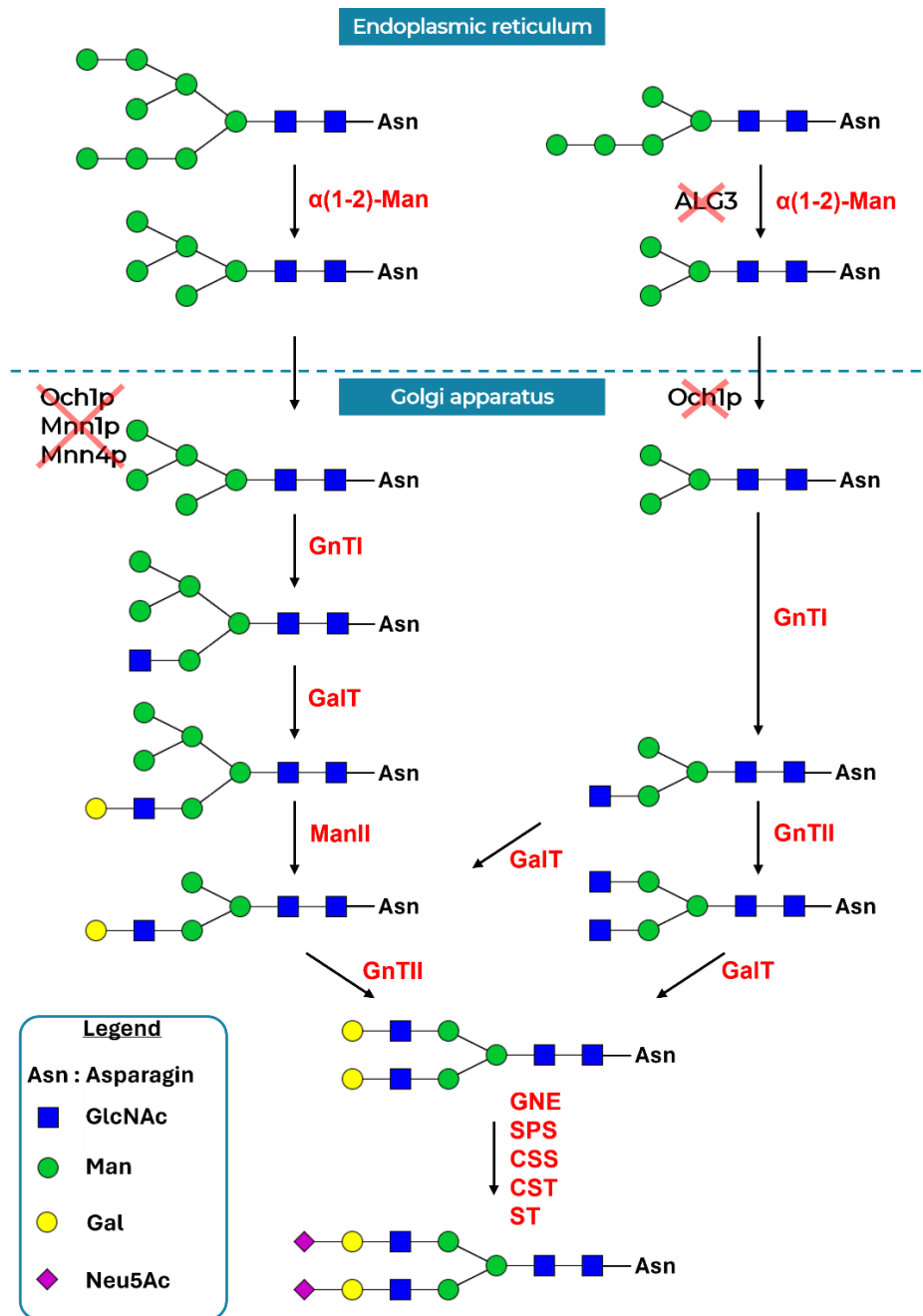
The main challenge in yeast *N*-glycosylation engineering is to move from high-mannose and hypermannosylated *N*-glycans to Man<sub>5</sub>GlcNAc<sub>2</sub> glycans, which is the substrate for GnTI, essential for the formation of complex human glycans. Moreover, it has been shown that recombinant proteins bearing yeast high-mannose *N*-glycans are easily cleared once administered to humans (Goetze *et al.*, 2011). Immunogenicity of yeast recombinant proteins is also increased by the presence of  $\beta(1,2)$ -mannose residues (Han *et al.*, 1997) and terminal  $\alpha(1-3)$  mannose residues (Raschke *et al.*, 1973) found respectively in *P. pastoris* (also known as *Komagataella phaffii*) and *S. cerevisiae*.

This section presents the main changes made to the most promising models: *S. cerevisiae* and *P. pastoris*. As explained above, *N*-glycosylation takes place in the ER and the Golgi apparatus, so two main strategies have been developed for humanising *N*-glycosylation pathways.

Historically, the first method was to modify this pathway by eliminating Golgi-resident glycosyltransferases responsible for hypermannosylation. Hence, the Och1  $\alpha(1,6)$ -mannosyltransferase described earlier has been inactivated in both *S. cerevisiae* and *P. pastoris* (Krainer *et al.*, 2013; Nakayama *et al.*, 1992) leading to Man<sub>8</sub>GlcNAc<sub>2</sub> *N*-glycans. Then, the inactivation of Mnn1 and Mnn4 genes responsible respectively for the grafting of  $\alpha(1,3)$ -mannose and phosphomannose residues in the *S. cerevisiae*  $\Delta$ Och1 mutants allowed the production of homogeneously mannosylated *N*-glycans with only  $\alpha(1,2)$ -linkage (Nakanishi-Shindo *et al.*, 1993). This step was not necessary in *P. pastoris* as it does not have these immunogenic  $\alpha(1,3)$ -mannose residues (Choi *et al.*, 2003).

To achieve the goal of producing Man<sub>5</sub>GlcNAc<sub>2</sub> glycans, heterologous  $\alpha(1,2)$ -mannosidases from other fungal species such as *Aspergillus saitoi* (*A. saitoi*) and *Trichoderma reesei* (*T. reesei*) have been respectively expressed in the *S. cerevisiae* Och1p x Mnn1 x Mnn4 triple mutant (Chiba *et al.*, 1998) and also in the *P. pastoris* Och1 mutant (Vervecken *et al.*, 2004). These enzymes have been HDEL-tagged or fused to other signals to retain them in the early stages of the *N*-glycosylation in the ER. In parallel, a strategy aiming at directly modifying the glycan precursor in the ER have been developed by inactivating in each of the models the genes encoding the alg3 protein, a Dol-P-Man: Man<sub>5</sub>GlcNAc<sub>2</sub>-PP-Dol mannosyltransferase responsible for forming the 'b' antenna of the precursor (Nakanishi-Shindo *et al.*, 1993). This mutation alone, or coupled with another mutation such as the one affecting Och1 in *P. pastoris* (Davidson *et al.*, 2004), provided a substrate for the  $\alpha(1,2)$ -mannosidase from *A. saitoi* mentioned earlier, enabling the formation of a 3 mannose precursor rather than a 5 mannose precursor.





**Figure 11 : Humanisation of the *N*-glycosylation pathway in yeast extract from Lucas, 2019 adapted from De Pourcq *et al.*, 2010.**

ALG : Asparagine linked glycosylation,  $\alpha(1,2)$ -mannosidase :  $\alpha(1,2)\text{-man}$ , *N*-acetylglucosaminyltransferase I and II : GnT I and II, mannosidase II : Man II, galactosyltransferase : GalT,  $\alpha(1,6)$ -mannosyltransferase : Och1p, mannosylphosphates-transferases : Mnn4p,  $\alpha(1,3)$ -mannosyltransferase : Mnn1p, GNE : UDP-Nacetylglucosamine-2-epimerase/*N*-acetylmannosamine kinase, SPS : *N*-acetylneuraminic-9-phosphate synthase, CSS : CMP-acide sialique synthase, CST : CMP-acide sialique transporter, ST : sialyltransferases.

Each of these strategies allowed to establish the conditions required for subsequent humanisation by expressing a GnTI and a GnTII, resulting in the structure GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> (Choi *et al.*, 2003; Hamilton *et al.*, 2003). The synthesis of this basic framework for all the complex glycans also required the expression of sugar nucleotide transporter to carry the elementary building blocks such as UDP-GlcNAc to the dedicated compartment, as well as the expression of a ManII to eliminate the  $\alpha(1,3)$ - and  $\alpha(1,6)$ -mannose residues, depending on the chosen strategy (Hamilton *et al.*, 2003).

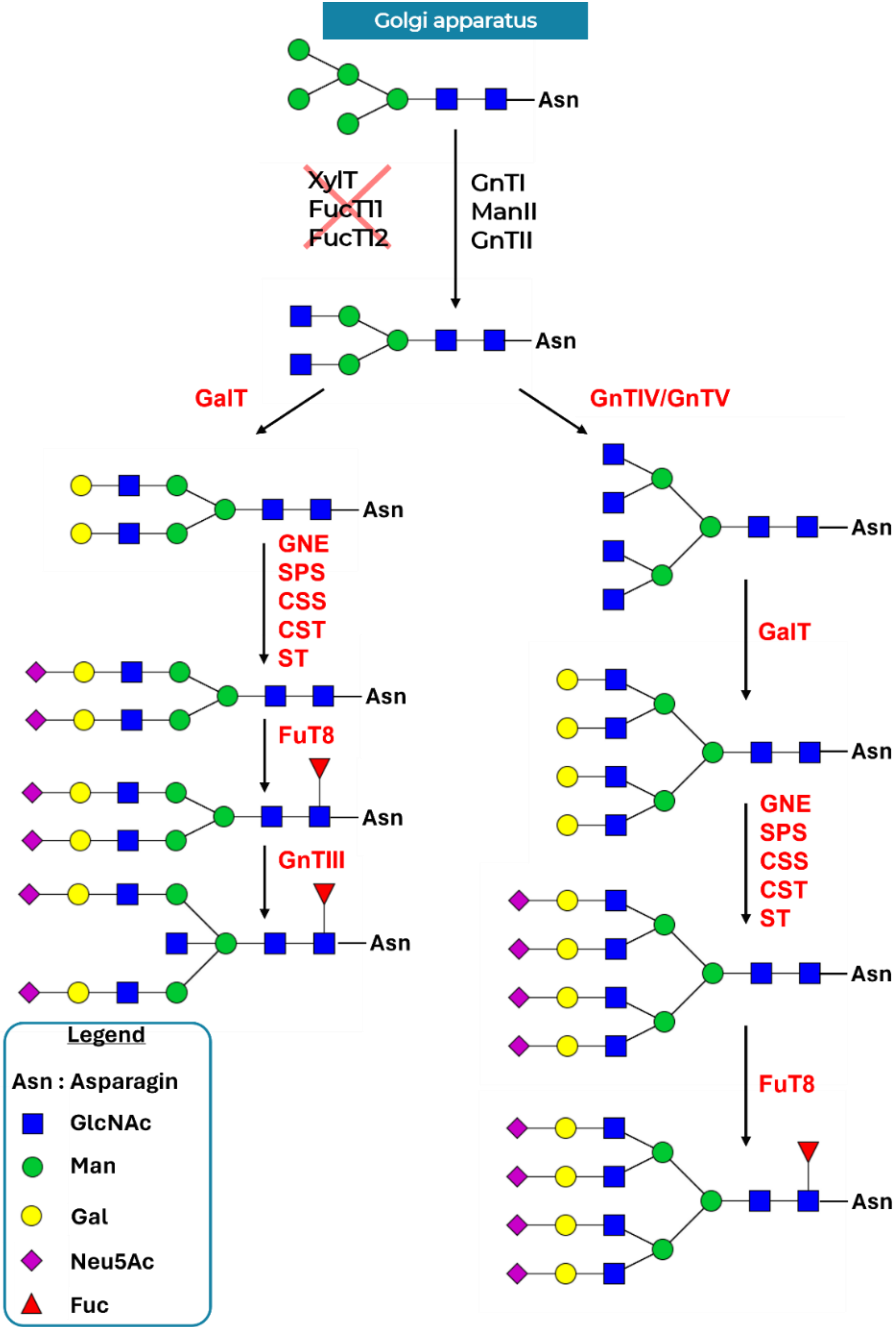
Engineering allowing galactosylation have been performed such as the expression of a human  $\beta(1,4)$ -GalT enabling  $\beta(1,4)$ -galactoses to be added to the terminal position in *P. pastoris* and *S. cerevisiae* (Bobrowicz *et al.*, 2004; Piirainen *et al.*, 2022). This gave yeast the enzymatic capacity to form the Gal<sub>2</sub>GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> motif.

The full humanisation of *N*-glycosylation pathways in yeast has been made possible by the expression of a  $\alpha(2,6)$ -sialT, which adds sialic acids in the terminal position. It was therefore necessary first to go through a feat of metabolic engineering involving the creation of a process to transform GlcNAc into activated sialic acid (CMP-sialic acid), its transport to the correct golgian sub-compartment and its transfer onto a terminal  $\beta(1,4)$ -galactoses (Hamilton *et al.*, 2006; Hamilton and Gerngross, 2007). All these steps have been made possible by the heterologous expression of 6 genes encoding enzymes including a epimerase, a kinase, a phosphatase and 2 synthases involved in the conversion of UDP-GlcNAc into CMP-sialic acid, a CMP-sialic acid transporter and a sialyltransferase (Hamilton *et al.*, 2003, 2006; Hamilton and Gerngross, 2007).

### C. Plants

Plants present the advantage of having similar Golgi stages identical to the human ones until the formation of the GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> type *N*-glycan is obtained in the Golgi apparatus. However, they also contain  $\beta(1,2)$ -xylose and core  $\alpha(1,3)$ -fucose residues, which are known to be potentially immunogenic for humans (Bardor *et al.*, 2003). The first efforts have therefore been made to remove these epitopes in several plant species including tobacco and *Arabidopsis thaliana* (*A. thaliana*) before

trying to humanise them by knock-out and/or knock-in strategies (Please refer to Eidenberger *et al.*, 2023 and Strasser, 2023 for recent review).



**Figure 12 : Example of humanisation of the N-glycosylation pathway in plants, extract from P.L Lucas's thesis (Lucas, 2019) modified from (Chen, 2016).**

The various epitopes found on the glycan structures were obtained by N-glycosylation engineering consisting of the inactivation and/or heterologous expression of glycoenzymes. Endogenous enzymes are written in black and crossed when inactivated. Heterologously expressed enzymes are written in red. Asn: asparagine, Man II: mannosidase II, GnT I to V: N-acetylglucosaminyltransferase I to V, XylIT: b(1,2)-xylosyltransferase, FucT11/12: a(1,3)-fucosyltransferase, FuT8: a(1,6)-fucosyltransferase, GalT : b(1,4)-galactosyltransferase, GNE: UDP-Nacetylglucosamine- 2-epimerase/N-acetylmannosamine kinase, SPS: N-acetylneuraminic-9-phosphate synthase, CSS: CMP-sialic acid synthase, CST: CMP-sialic acid transporter, ST: sialyltransferases.

Glycosylation mutants were therefore used to generate double mutants in which the genes encoding the  $\beta(1,2)$ -xylosyltransferase and  $\alpha(1,3)$ -fucosyltransferase were inactivated (Strasser *et al.*, 2004). Other methods, such as the use of siRNA (Strasser *et al.*, 2008) or genome editing tools (Göritzer *et al.*, 2022; Jansing *et al.*, 2019; Mercx *et al.*, 2017), have also been employed to achieve the same goal. In this way, therapeutic recombinant proteins expressed in plants lacking these immunogenic epitopes were created (Cox *et al.*, 2006; Schähs *et al.*, 2007; Strasser *et al.*, 2009a).

The expression of GnT III, GnT IV and GnT V has also been necessary to complete the humanisation of plant *N*-glycosylation. Heterologous GnT III from human or rat have been shown to be capable of adding the bisecting GlcNAc to plant *N*-glycans (Frey *et al.*, 2009; Rouwendal *et al.*, 2007). The expression of genes encoding the GnT III, GnT IV, GnT V and FuT8 enzymes in mutants deficient in plant immunogenic epitopes was then used to obtain tri- and tetra-antennated glycans in *Nicotiana benthamiana* (*N. benthamiana*) (Castilho *et al.*, 2011).

Plants have been shown to be capable of expressing a human  $\beta(1,4)$ -galactosyltransferase with a modification of the *N*-glycosylation profile (Bakker *et al.*, 2001, 2006; Palacpac *et al.*, 1999). The expression of a such an enzyme in the trans-Golgi of double mutant  $\Delta$ XT/ $\Delta$ FuT plants has taken the scientific community a step closer to humanising *N*-glycosylation pathways and allowed once more the production of humanised plant-made biopharmaceuticals that show a lower immunogenicity (Bohlender *et al.*, 2020; Strasser *et al.*, 2009a).

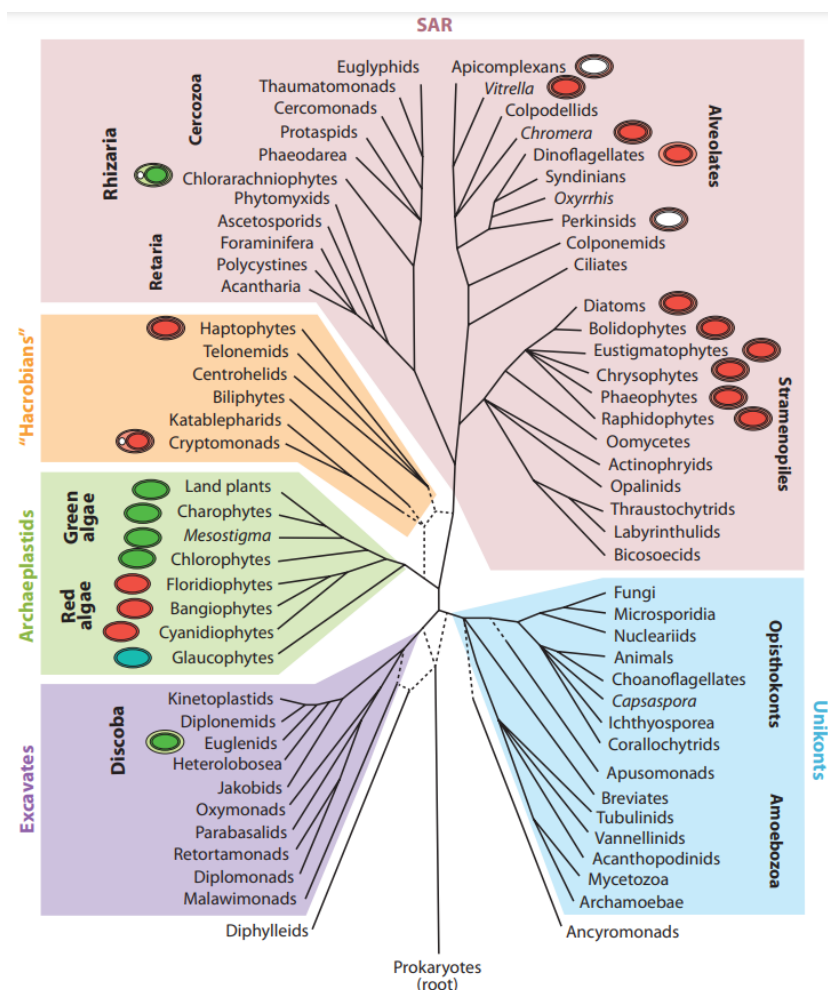
Finally, the lifespan of therapeutic molecules has been then improved by adding a step to the engineered plant *N*-glycan construction chain. Indeed, plants are unable to graft terminal  $\alpha(2,3)$  and  $\alpha(2,6)$ -sialic acid such as *N*-acetylneuraminic acid (Neu5Ac), a residue present in humans (Angata and Varki, 2002). It was therefore necessary to design a pathway for synthesising CMP-sialic acid from UDP-GlcNAc and the machinery required to deliver it to the Golgi compartment to which the heterologous sialyltransferases (SialTs) were addressed (Bohlender *et al.*, 2020; Castilho *et al.*, 2008, 2010; Kallolimath *et al.*, 2016; Paccalet *et al.*, 2007). In this way, biopharmaceuticals such as EPO could be synthesised by glycoengineered plants in a biologically active form, as demonstrated by Pr Herta Steinkellner's team and

collaborators (Castilho *et al.*, 2013; Jez *et al.*, 2013; Loos *et al.*, 2014; Schneider *et al.*, 2014).

## VI. Microalgae as alternative

### A. General information

Over 300 000 species of microalgae have been classified based primarily on differences in pigments, nature of the reserves and cellular structures present in these organisms. These photosynthetic unicellular micro-organisms include prokaryotic cells such as cyanobacteria, also known as blue-green algae. These photosynthetic unicellular micro-organisms include prokaryotic cells such as cyanobacteria, also known as blue-green algae. They also include eukaryotic microalgae divided into green microalgae, brown microalgae and red microalgae, whose species are classified in the following groups: Euglenophytes, Bacillariophytes, Chlorophytes, Chrysophytes, Dinophytes, Charophytes, Glaucophytes, Rhodophytes, Cryptophytes, Phaeophytes and Haptophytes (Figure 13) (Goshtasbi *et al.*, 2023; Keeling, 2013).



**Figure 13 : Microalgae phylogenetic tree from (Keeling, 2013).**

Hacrobia and SAR (stramenopile-alveolate-rhizarian) groups combined would be corresponding to Chromalveolates. Oval coloured shapes and their structures are representative of the organism's structure and its affiliation with red or green algae. Whites indicate non-photosynthetic organisms.

Their phylogenetic diversity reflects their adaptation capacity to various environments, so they can be found in freshwater as well as in extreme environments such as acidic thermal springs.

For several years now, microalgae have been attracting interest from industry as alternative systems. These organisms are already being used in a number of fields because of their advantages. Indeed, they can be easily grown under controlled conditions, such as in photobioreactors, which allow GMOs to be grown without dispersing them into the environment and without gene pollution. Their high and rapid growth rate in inexpensive environments leading to high yields is a further advantage for industries producing value-added molecules such as biofuels with the lipids accumulated in these microalgae (Parameswari. and Lakshmi, 2022; Wang *et al.*, 2024). They are also used in the design of food supplements thanks to their capacity to produce antioxidants such as  $\alpha$ -carotene,  $\beta$ -carotene, astaxanthin, fucoxanthin, lutein and zeaxanthin (Figure 14).

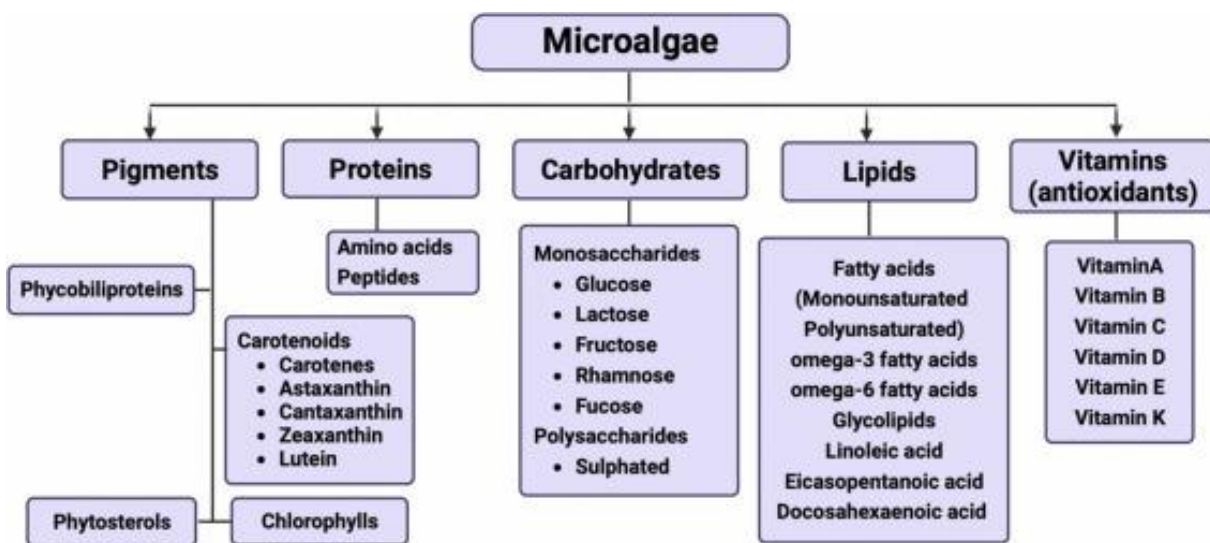


Figure 14 : Diagram showing the range of value-added molecules derived from microalgae from (Parameswari. and Lakshmi, 2022).

With the introduction in the 1980s of genetic transformation tools for microalgae (biolistic, electroporation, glass beads, virus mediated delivery) (Gutiérrez and Lauersen, 2021; Specht *et al.*, 2010) the production of recombinant protein, such as biomedicines, is therefore also conceivable.

In addition to the ease of cultivation mentioned above and the tools available, microalgae offer other advantages, particularly for protein production.

Unlike conventional systems such as mammalian production systems, algae offer a significant reduction in production costs. This is due, in particular, to the elimination of purification steps linked to the risk of contamination by viruses or prions that occurs in mammalian cells culture systems. One example is the contamination of blood-derived products by the human immunodeficiency virus in the 80s and 90s, which had disastrous consequences. Since microalgae cannot be infected by this type of human pathogen, they have a clear advantage. What's more, eukaryotic microalgae have retained throughout evolution the cellular machinery needed to fold complex human proteins with co- and post-translational modifications. Finally, the FDA's granting of GRAS status to numerous species of green algae once again justifies the elimination of purification stages, thereby helping to reduce production costs.

**Table 2 : Summary of algae-made biopharmaceutical products with proven biological activity and potential therapeutic use (Dehghani *et al.*, 2022).**

Recombinant protein	Expression Host	Targeted disease	References
Human interleukin 2	<i>C. reinhardtii</i> , <i>C. vulgaris</i> and <i>D. salina</i>	Cancer	(Dehghani <i>et al.</i> , 2020)
Human anti-hepatitis B surface antigen antibody (CL4mAb)	<i>P. tricornutum</i>	Hepatitis B	(Hempel <i>et al.</i> , 2011; Hempel and Maier, 2012)
Hepatitis B surface antigen	<i>P. tricornutum</i>	Hepatitis B	(Hempel <i>et al.</i> , 2011)
RBD of SARS-CoV-2	<i>P. tricornutum</i>	COVID-19	(Slattery <i>et al.</i> , 2022)
Human growth hormone	<i>C. reinhardtii</i>	Turner syndrioem	(Wannathong <i>et al.</i> , 2016)
Hepatitis B surface antigen	<i>D. salina</i>	Hepatitis B	(Geng <i>et al.</i> , 2003)
Human interferon $\alpha$ 2a	<i>C. reinhardtii</i>	Cancer	(Commault <i>et al.</i> , 2020)
Spike glycoprotein of SARS-CoV-2	<i>C. reinhardtii</i>	COVID-19	(Kiefer <i>et al.</i> , 2022)
RBD of Spike SARS-CoV-2	<i>C. reinhardtii</i>	COVID-19	(Berndt <i>et al.</i> , 2021)
BCB; a multi epitope protein	<i>Schyzochitrium sp.</i>	Breast cancer	(Hernández-Ramírez <i>et al.</i> , 2020)
Human VEGF-165	<i>C. reinhardtii</i>	Wound healing	(Jarquín-Cordero <i>et al.</i> , 2020)



In this respect, microalgae offer new hopes of finding more environmentally friendly production systems at lower cost. As a result, the following species *C. reinhardtii*, *Dunaliella salina* (*D. salina*), *Phaeodactylum tricornutum* (*P. tricornutum*), *Chorella vulgaris* (*C. vulgaris*), among others, have been used to produce recombinant proteins with proven biological activity for therapeutic purposes (Table 2). Among the mentioned models, this thesis is focusing on the green microalga *C. reinhardtii*.

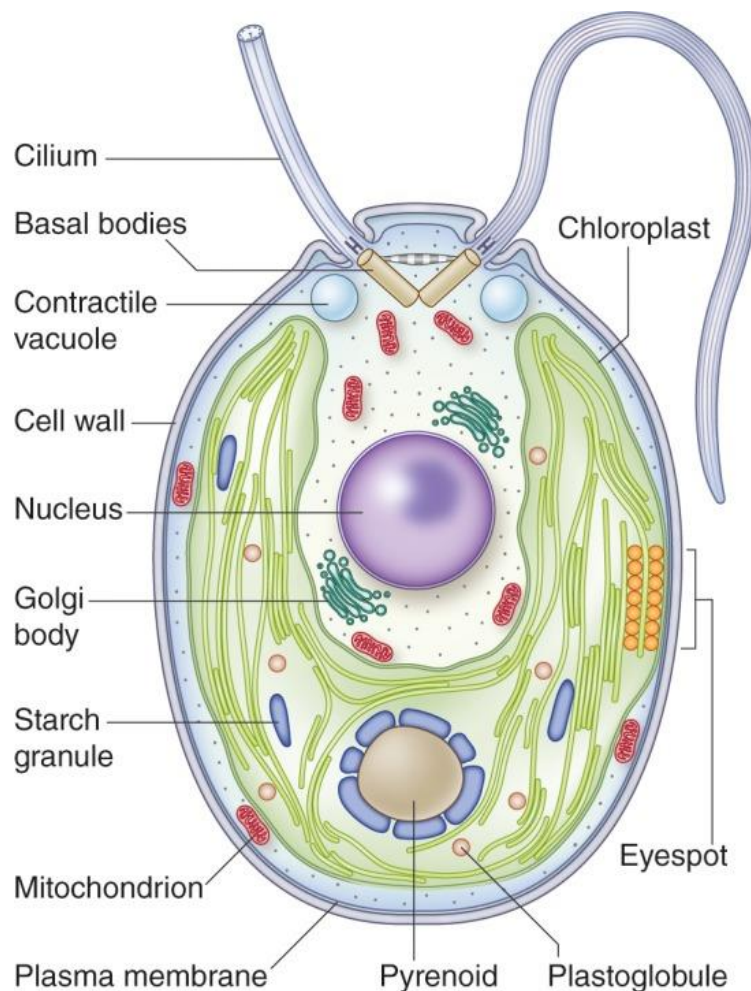
## B. *Chlamydomonas reinhardtii*

*C. reinhardtii* is a single-celled green microalga isolated for the first time in 1945 by Gilbert Morgan Smith (Flowers *et al.*, 2015). The cells measure around 10µm and are bounded by a cell wall. *C. reinhardtii* is a eukaryotic organism and therefore has a nucleus, mitochondria, an endoplasmic reticulum and at least one Golgi apparatus. This photosynthetic organism belongs to the *Chlorophyta* phyla and consequently has a single chloroplast representing 40% of the cell volume, a pyrenoid and a eyespot or stigma involved in phototaxis (Figure 15) (Takouridis *et al.*, 2015)

*C. reinhardtii* is mainly photosynthetic when light conditions are sufficient to ensure the production of ATP and the organic carbon compounds required for growth and reproduction. However, in the absence of light or when nitrogen or organic carbon sources are deprived, this microalga can absorb organic nutrients, such as acetate, from its surrounding source. This ability to simultaneously use autotrophic photosynthesis and heterotrophic organic nutrient uptake is known as mixotrophy (Blifernez-Klassen *et al.*, 2012; McATEER *et al.*, 1985)

The two motile cilia on the apical pole of the cell (Figure 15) ensure mobility, adhesion and are also involved in reproduction. While *C. reinhardtii* is haploid most of the time, with asexual reproduction by fission giving it a short generation time of around eight hours, two vegetative cells of opposite sexual polarity, positive and negative mating type (mt+ and mt-), can differentiate into haploid gametes triggered by environmental signals (for example, a nitrogen deprivation). The gametes fusion, mediated by the cilia, will result in diploid zygotes and thus initiate so-called sexual reproduction. The zygospores remain dormant until more favourable conditions allow them to germinate in the presence of light, giving rise to four haploid cells (Harris, 2001).

The great adaptability and versatility of *C. reinhardtii* has made it a widely used model for biological studies since the 1960s leading to a better understanding of fundamental biology, photosynthesis and energy production, the evolution of organelles and diversity within the plant kingdom. In addition, it has been awarded GRAS (Generally Recognized As Safe) status by the Food and Drug Administration in 2018. It therefore presents no danger to human health. These factors, together with the ease with which these organisms can be grown in photobioreactors, which means that GMO crops can be totally controlled, reinforced its use also as a biotechnological model. Industries have taken advantage of the benefits of cultivation, notably by using *C. reinhardtii* to produce biohydrogen (Torzillo *et al.*, 2015) or to bioremediate soil polluted by heavy metals (Cadoret *et al.*, 2008).



**Figure 15 : *C. reinhardtii* cell structure (Sasso *et al.*, 2018).**

The cell contains a nucleus, a single chloroplast occupying 40% of the cell volume, as well as mitochondria, a Golgi apparatus and reserve organs such as a pyrenoid and starch grains. The cell also has an organ called the stigma, which uses light intensity to control the mobility of the cell via its flagella. The subcellular structures are contained within a cell wall and a plasma membrane.

The complete sequencing of its chloroplast, mitochondrial and nuclear genomes (Craig *et al.*, 2022) and the appearance of genetic transformations in microalgae since the 1980s that are now well developed in *C. reinhardtii* (Gao *et al.*, 2015; Sizova *et al.*, 2013; Zhang *et al.*, 2020) led to the development of mutant libraries (Li *et al.*, 2019; Zhang *et al.*, 2014) such as the CLiP Library. More than 65,000 insertional mutants are available for the scientific community since 2015. These mutants can be crossed by sexual reproduction (Jiang and Stern, 2009) facilitating, over the past decade, the studies aiming at deciphering metabolic pathways in particular, the *N*-glycosylation pathway.

### C. Recombinant protein in *C. reinhardtii*

The knowledge gained from studying the different genomes of *C. reinhardtii* and the mutants generated, as well as the different mechanisms in this alga, has also made it an intriguing and interesting model for the expression of recombinant proteins.

The first attempt at heterologous expression of a green fluorescent protein (GFP) from *Aequorea Victoria* in the *C. reinhardtii* genome was made in 1999 by Fuhrmann and colleagues. The low production rate obtained was attributed to a number of hypothetical factors, including silencing due to an epigenetic effect or a transcriptional defect (Fuhrmann *et al.*, 1999). Although GFP is not of pharmaceutical interest, its expression marks the start of a series of other attempts to express a recombinant protein in the *C. reinhardtii* study model.

Despite efforts to identify promoters capable of over-expressing heterologous genes (Fischer and Rochaix, 2001), the development of synthetic hybrid promoters (Schroda *et al.*, 2000) and the optimisation of the codon bias of the transgene derived from the *C. reinhardtii* nuclear genome (Fuhrmann *et al.*, 1999, 2004; Shao and Bock, 2008), which have enabled the expression of nuclear genes to be optimised, the expression levels achieved remain unsatisfactory for industrial production. Given the obstacles encountered in production by targeting nuclear DNA, most proteins have been expressed by targeting chloroplast DNA. The large volume of the chloroplast is a real advantage, as it allows the storage of recombinant proteins representing 2 to 20% of total soluble proteins, giving a relatively satisfactory yield (Rasala and Mayfield, 2011). In addition, chloroplast DNA is a circular DNA that allows autonomous duplication in this organelle.

The first production of a monoclonal antibody directed against glycoprotein D of the herpes virus in this organism has been carried out in the early 2000s (Mayfield *et al.*, 2003). A second monoclonal antibody directed against the PA 83 protein, involved in anthrax (Tran *et al.*, 2009), was also expressed in this organism. Both proteins have been expressed via a chloroplast codon optimised plasmid (Table 3).

Other molecules have been produced *via* a transformation of the chloroplast genome. Dauvillée and collaborators produced the antigens for a malaria vaccine in 2010. In 2016, Rasala and collaborators produced hEPO. Nevertheless, despite these encouraging results, protein expression *via* chloroplast DNA does not allow PTMs such as glycosylation on the recombinant proteins produced. As mentioned previously, the glycan part of a glycoprotein is essential to its functionality. Thus, glycoproteins can only be produced *via* nuclear DNA expression.

However, it is well known that nuclear expression levels in *C. reinhardtii* are low and are accompanied by silencing phenomena that restrict the production of recombinant proteins and therefore result in low production yields. Strains with stable nuclear expression have emerged since, such as UVM4 and UVM11 (Neupert *et al.*, 2009), and are now widely used for high yield recombinant protein expression. Both carry mutation, generated by UV rays exposure, in a gene encoding Sir2-like NAD(+) dependent protein deacetylase that may allow the DNA that newly enters the cells to be expressed with no silencing and a better expression level (Neupert *et al.*, 2020). Moreover, the last decade has brought many new molecular biology tools for a better nuclear expression of recombinant proteins.

In addition to optimising the transgene by adapting the use of codons specific to *C. reinhardtii*, it has been shown that the RBCS2 intronic sequence should be inserted sequentially along the transgene to fragment it and thus mimic the length of the exons found in *C. reinhardtii* to improve recruitment of the transcriptional machinery (Baier *et al.*, 2018). This optimisation stage has been made easier thanks to the development of an online tool, Intronserter (Jaeger *et al.*, 2019). Finally, as for the other expression system, it has been demonstrated that the combination Promoter/Terminator plays a key role in the expression level of a transgene (Einhaus *et al.*, 2021; Geisler *et al.*, 2021; Kumar *et al.*, 2013; López-Paz *et al.*, 2017).

The rapid building of such optimised plasmids is now enabled by the MoClo Toolkit, a flexible tool inspired from golden-gate technology that is constantly being improved thanks to input from the overall *Chlamydomonas* community (Crozet *et al.*, 2018).

**Table 3 : Recombinant proteins expressed in *C. reinhardtii* until 2020 modified from (Barolo *et al.*, 2020).**

Organelle	Protein	Source
Chloroplast	E7 of HPV-16	(Demurtas <i>et al.</i> , 2013)
	D2-CTB	(Dreesen <i>et al.</i> , 2010)
	$\alpha$ -galactosidase	(Georgianna <i>et al.</i> , 2013)
	Phytase	(Georgianna <i>et al.</i> , 2013)
	Xylanase	(Georgianna <i>et al.</i> , 2013)
	Pfs25	(Gregory <i>et al.</i> , 2012)
	Pfs28	(Gregory <i>et al.</i> , 2012)
	Pfs25-CTB	(Gregory <i>et al.</i> , 2013)
	E2	(He <i>et al.</i> , 2007)
	Pfs48/45	(Jones <i>et al.</i> , 2013)
	M-SAA	(Manuell <i>et al.</i> , 2007)
	Anti-HSV glycoprotein D Isc	(Mayfield <i>et al.</i> , 2003)
	12FN3	(Rasala <i>et al.</i> , 2010)
	Erythropoietin	(Rasala <i>et al.</i> , 2010)
	HMGB1	(Rasala <i>et al.</i> , 2010)
	Interferon $\beta$	(Rasala <i>et al.</i> , 2010)
	Proinsulin	(Rasala <i>et al.</i> , 2010)
	SAA-10FN3	(Rasala <i>et al.</i> , 2010)
	VEGF	(Rasala <i>et al.</i> , 2010)
	Allophycocyanin	(Su <i>et al.</i> , 2005)
	VP1-CTB	(Sun <i>et al.</i> , 2003)
	V28	(Surzycki <i>et al.</i> , 2009)
	Anti-PA 83 anthrax IgG1	(Tran <i>et al.</i> , 2009)
	Anti-CD22-gelonin sc	(Tran, Henry, <i>et al.</i> , 2013)
	Anti-CD22-ETA sc	(Tran, Van, <i>et al.</i> , 2013)
	GAD65	(Wang <i>et al.</i> , 2008)
TRAIL	(Yang <i>et al.</i> , 2006)	
Phytase (AppA)	(Yoon <i>et al.</i> , 2011)	
Metallothionein-2	(Zhang <i>et al.</i> , 2006)	
Nucleus	Human Epidermal Growth Factor	(Baier, Kros, <i>et al.</i> , 2018)
	VEGF-165	(Chávez <i>et al.</i> , 2016)
	GBSS-AMA1	(Dauvillée <i>et al.</i> , 2010)
	GBSS-MSP1	(Dauvillée <i>et al.</i> , 2010)
	Erythropoietin	(Eichler-Stahlberg <i>et al.</i> , 2009)
	Sep-15	(Hou <i>et al.</i> , 2013)
	<i>Lolium Perenne</i> IBP	(Lauersen, Vanderveer, <i>et al.</i> , 2013)
	$\beta$ -1,4-endoxylanase	(Rasala <i>et al.</i> , 2012)

In 2013, it has been shown that the addition of the signal peptide of the native extracellular *C. reinhardtii* carbonic anhydrase allows the efficient secretion of recombinant protein in the culture medium (Lauersen, Berger, *et al.*, 2013).

In 2018, the list of secretory signal peptides have been extended by Molino and collaborators with the aim of helping to use this organism with greater secretion yield for industrial purposes (Molino *et al.*, 2018).

However, the emphasis in this introduction has been on the fact that the availability of tools and the ability of an organism to produce and secrete recombinant proteins are not the only prerequisites for being considered as an alternative bioproduction system for therapeutic purposes. Indeed, such ambitions first require an in-depth understanding of its *N*-glycosylation pathways.

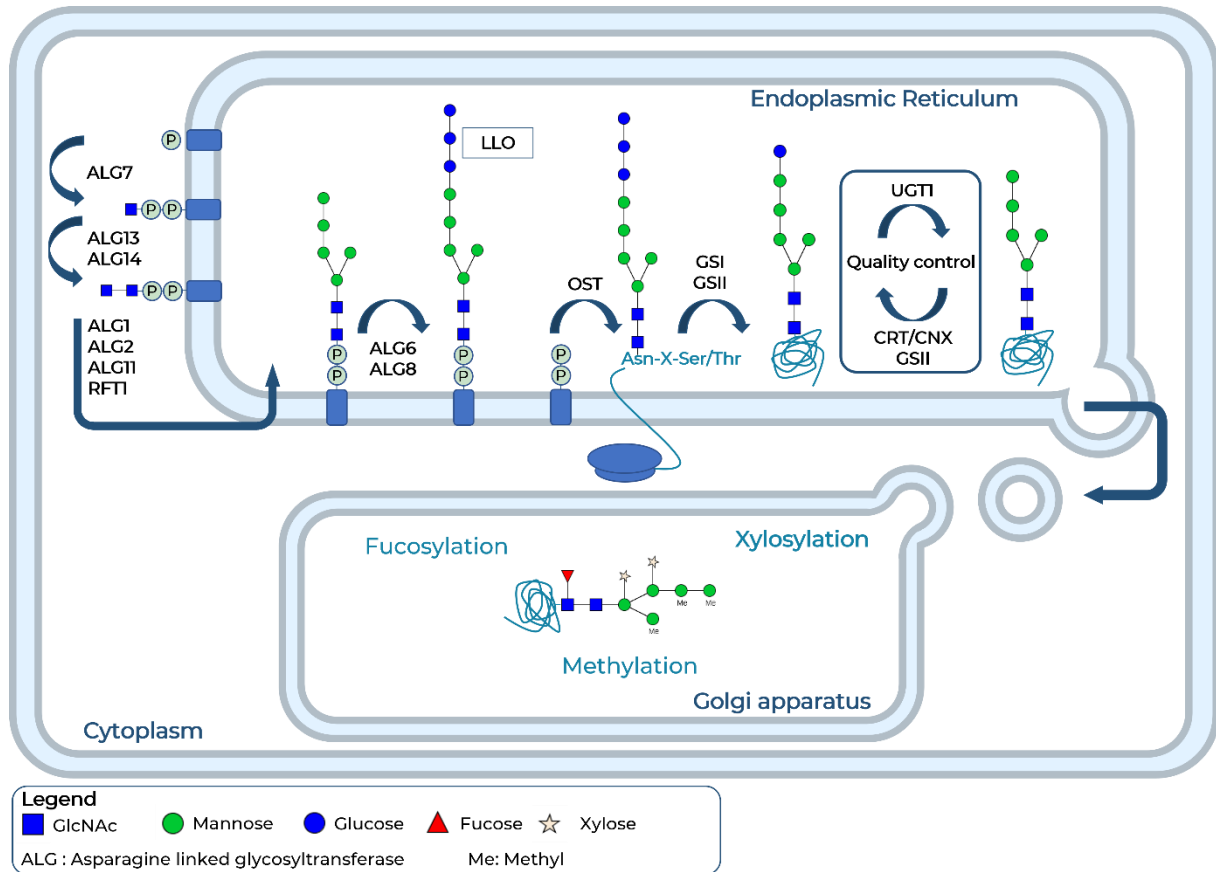
#### D. *N*-glycosylation

The study and characterisation of *N*-glycosylation in *C. reinhardtii* began in 2013 and the following paragraphs summarise the studies carried out since then (Mathieu-Rivet *et al.*, 2013; Vanier *et al.*, 2017; Lucas *et al.*, 2018; Schulze *et al.*, 2018; Lucas *et al.*, 2020; Oltmanns *et al.*, 2020).

LLO synthesis begins on the cytosolic side of the ER membrane. It is initiated by the action of a phosphate transferase (CrALG7) on the dolichol pyrophosphate anchor, leading to the formation of GlcNAc1-PP-Dol. The addition of another GlcNAc, mediated by the CrALG13/CrALG14 complex, leads to the formation of the GlcNAc2-PP-Dol intermediate. This is then mannosylated five times thanks to the respective activities of  $\beta$ (1,4)-mannosyltransferase (CrALG1), then  $\alpha$ (1,3)- and  $\alpha$ (1,2)-mannosyltransferases (CrALG2 and CrALG11). The cytosolic steps thus lead to the formation of the Man<sub>5</sub>GlcNAc<sub>2</sub>-PP-Dol motif characterised in 2018 (Lucas *et al.*, 2018). This oligosaccharide precursor is then flip-flopped inside the ER across the reticulum membrane by CrRFT flippase. The transferred precursor is then directly subject to glucosylation steps ensured by two  $\alpha$ (1,3)-glucosyltransferases (ALG6 and ALG8), which alone transfer three glucose residues, since bioinformatic analyses of the nuclear genome failed to identify a gene encoding the ALG10 enzyme responsible for adding an  $\alpha$ (1,2)-glucose. (Gomord *et al.*, 2010; Levy-Ontman *et al.*, 2014; Mathieu-Rivet *et al.*, 2013).

Therefore, the glycan precursor accumulated in the *C. reinhardtii* ER is a LLO with a predominant truncated linear structure of the Glc<sub>3</sub>Man<sub>5</sub>GlcNAc<sub>2</sub> type distinct from that found in other eukaryotes (Figure 16) (Lucas *et al.*, 2018).

This structure is consistent with *in silico* studies analysis demonstrating that no gene encoding the luminal mannosyltransferases ALG3, ALG9, and ALG12, that allow the addition of additional mannoses, has been found in the genome (Mathieu-Rivet *et al.*, 2013). Similarly, It has been observed in other organisms that the loss of ALG glycosyltransferase assemblies is at the origin of the diversity of dolichol-related precursors (Samuelson *et al.*, 2005).



**Figure 16 : Proposed representation of the *C. reinhardtii* N-glycosylation pathway.**

*C. reinhardtii* synthesises a pentamannosylated linear glycan precursor of the  $\text{Glc}_3\text{Man}_5\text{GlcNAc}_2$  type on a dolicholpyrophosphate (PP-Dol) membrane anchor (Lucas *et al.*, 2018). A transfer of the oligosaccharide group from the lipid-linked oligosaccharide (LLO) precursor to the protein under synthesis is carried out by the predicted oligosaccharyltransferase (OST) complex (Vanier *et al.*, 2017). This is followed by deglycosylation to form a  $\text{Man}_5\text{GlcNAc}_2$  structure (Vanier *et al.*, 2017). After transfer of the newly synthesised protein to the Golgi apparatus, the N-glycans are matured to become partially methylated, fucosylated and xylosylated.

The oligosaccharide part of the LLO synthesised in this way is transferred *en bloc* by the oligosaccharyltransferase (OST) complex to the protein being synthesised, onto the consensus sequence previously described (Figure 16).

Deglucosylation by  $\alpha$ -glucosidases (GSI, GSIIa and GSIIb) allows the nascent glycoprotein to enter the quality control cycle, before being sent to the Golgi apparatus.

In addition to the predominantly oligomannosidic *N*-glycan structures, structures with two to five mannose exhibiting  $\alpha(1,2/6)$  terminal residues, glycans with one or two xylose residues, as well as methylations; are synthesised in *C. reinhardtii* (Lucas *et al.*, 2020; Mathieu-Rivet *et al.*, 2020; Oltmanns *et al.*, 2020; Schulze *et al.*, 2020). Notably the presence of  $\beta(1,2)$  xylose groups grafted onto the glycan body, as observed in plants by Lerouge *et al.* in 1998, has been characterized. A second xylose residue and methyl groups on mannoses were also identified (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020).

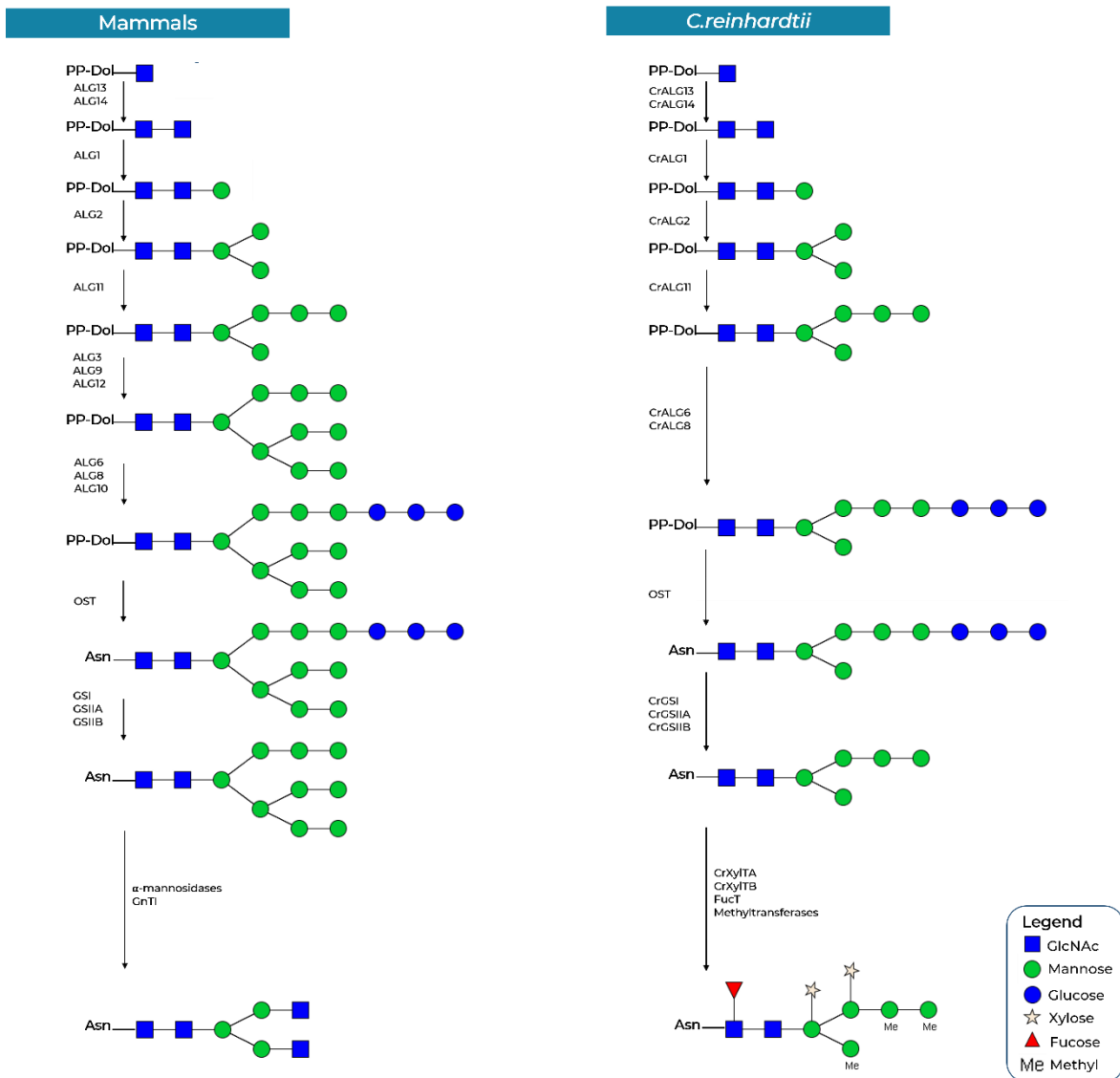
Recently, two genes named *XTA* and *XTB*, encoding xylosyltransferases, were characterised in *C. reinhardtii* (Lucas *et al.*, 2020; Schulze *et al.*, 2018). This is consistent with previous results obtained in 2018, by Pr. Michaël Hippler's team at the University of Münster in Germany who showed that the inactivation of the first *XTA* led to a significant reduction in xylosylation of the glycan's core of secreted proteins (Schulze *et al.*, 2018). The xylosylation pathway is detailed in Figure 18.

This work also showed that more or less methylated glycans bearing  $\beta(1,2)$ -xylose residues were greatly reduced by the absence of mannosidase I activity in *C. reinhardtii*. This suggests that xylosylation and methylation are dependent on the activity of this mannosidase I (Schulze *et al.*, 2018). Fucose residues linked in  $\alpha(1,3)$  to *N*-glycans were highlighted by Oltmanns *et al.*, 2020 (Oltmanns *et al.*, 2020).

It is noteworthy that *C. reinhardtii* differs from plants and mammals by the absence of GnTI activity, the key enzyme in the Golgi apparatus that initiates the formation of diversified complex *N*-glycans (Mathieu-Rivet *et al.*, 2014, 2013). In mammals and plants, GnTI transfers an *N*-acetylglucosamine residue to the Man<sub>5</sub>GlcNAc<sub>2</sub> intermediate motif, leading to the synthesis of GlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub>-Asn (Lerouge *et al.*, 1998; Strasser, 2016) and paving the way to the formation of complex structures harbouring terminal GlcNAc residues (Figure 17).

To sum up, this microalga synthesizes non canonical oligomannosidic *N*-glycans ranging from Man<sub>5</sub>GlcNAc<sub>2</sub> to Man<sub>3</sub>GlcNAc<sub>2</sub> that are bearing specific residues such as  $\beta(1,2)$ ,  $\beta(1,4)$ -xylose and  $\alpha(1,3)$ -fucose that are absent in mammalian glycoproteins (Figure 17) (Lucas *et al.*, 2020; Oltmanns *et al.*, 2019; Schulze *et al.*, 2018).





**Figure 17 : Scheme comparing the *N*-glycan biosynthesis in land plants and mammals (left) and proposed *N*-glycosylation pathway in *C. reinhardtii* (right) modified from (Vanier *et al.*, 2017).**

The synthesis of a pentamannosylated linear Glc<sub>3</sub>Man<sub>5</sub>GlcNAc<sub>2</sub> precursor *C. reinhardtii* is due to the lack of the ALG3, 9, 10 and 12 enzymes. Plants and mammals synthesize a Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> precursors onto a membrane-anchor dolichol pyrophosphate (PP-Dol) in the endoplasmic reticulum. The Golgi stages are independent from the GnTI enzyme in *C. reinhardtii* and can't lead to mammal type complex glycans. However, *N*-glycans are partially methylated, fucosylated and xylosylated (Oltmanns *et al.*, 2020).

## E. Glycoengineering in *C. reinhardtii*

Glycoengineering projects in *C. reinhardtii* are supported by the fact that, among the microalgae studied so far as emerging alternative expression systems for the biopharmaceutical production, it benefits from numerous molecular and biochemical tools for the engineering of metabolic pathways (Schroda and Remacle, 2022)

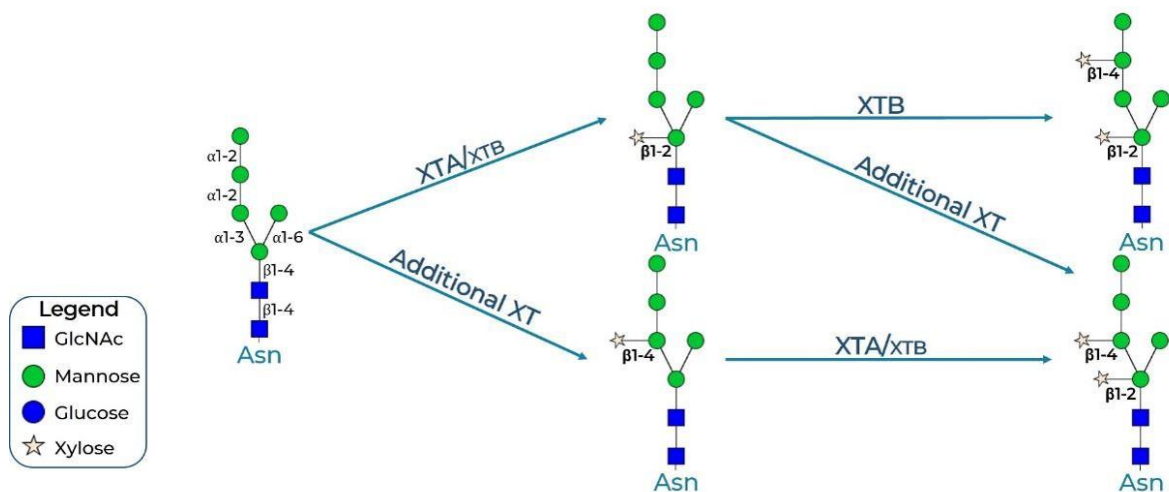
The presence of two xylose groups, including a  $\beta(1,2)$ -linked group on a mannose attached to the core chitobiose, was observed, as were methyl groups in *C. reinhardtii*. As the immunogenic potential in humans of the  $\beta(1,2)$ -xylose of plant proteins is the subject of controversy, the removal of this epitope is strongly recommended in the context of the production of injectable recombinant proteins. Fucosylation has also recently been demonstrated on *C. reinhardtii* *N*-glycans (Oltmanns *et al.*, 2020). Furthermore, as previously explained, the *N*-glycosylation pathway in this microalga is independent of GnTI, which does not allow the synthesis of human-type *N*-glycans. These differences in the enzymatic machinery of this microalga compared with mammals justify a glycoengineering attempt.

In 2017, Vanier and colleagues set out to humanise the *N*-glycosylation pathway in *C. reinhardtii*. In this study, the heterologous expression of GnTI known to be active in *A. thaliana* and *P. tricornutum* was carried out in order to complement the GnTI-independent Golgi processing in a common cw92 *C. reinhardtii* strain. GnTI from *P. tricornutum* has been previously shown to restore biosynthesis of complex *N*-glycan in Chinese hamster Lec1 ovary cells deficient in endogenous GnTI activity (Baïet *et al.*, 2011). Mass spectrometry analysis revealed that no additional GlcNAc residues have been transferred on the *N*-glycans of endogenous proteins in *C. reinhardtii* lines expressing heterologous PtGnTI or AtGnTI despite an expression level allowing the immunoblot detection. Moreover, the accumulation of starch granules reflecting stress linked to GnTI expression have been observed. The *N*-glycans therefore remained unchanged even though GnTI was expressed (Vanier *et al.*, 2017). This phenotype was attributed to the presence of Man5 structures that is not a suitable substrate for the heterologous GnTI enzymes expressed.

The work carried out in 2020 showed that the XTA protein is mainly responsible for transferring a  $\beta(1,2)$ -xylose residue to the *N*-glycan core. On the other hand, the XTB protein is responsible for the transfer of a  $\beta(1,4)$ -xylose onto a mannose of the

linear branch of the *N*-glycan of synthesised proteins; and to a lesser extent, it is responsible for the addition of  $\beta(1,2)$ -xylose onto the core of the *N*-glycan (Figure 18) (Lucas *et al.*, 2020).

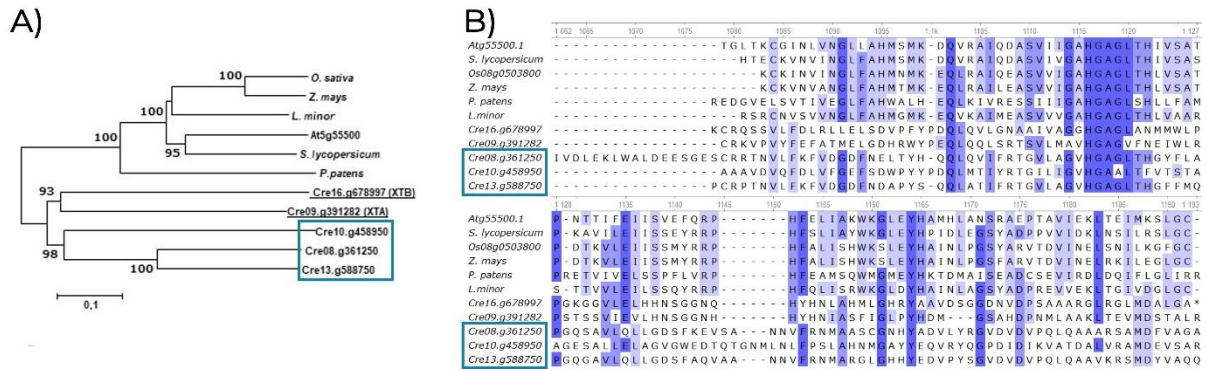
This work has highlighted the presence of great heterogeneity in xylosylated glycan structures in *C. reinhardtii*  $IM_{XTA}$ ,  $IM_{XTB}$  single mutants and  $IM_{XTA}IM_{XTB}$  double mutant. Indeed some residual xylose residues have been observed in the aforementioned double mutant leading to the hypothesis that another xylosyltransferase may also be involved in the maturation of *N*-glycans in *C. reinhardtii* (Lucas *et al.*, 2020).



**Figure 18 : Proposed scheme of xylosylation processes in the green microalga *C. reinhardtii* (Lucas, 2019).**

The figure shows the positions of the xylose residues on the mannoses of the glycan core by means of a star. The addition of these residues is mediated by different xylosyltransferases governed by a multigene family.

Three new candidate genes (Cre10.g458950, Cre13.g588750 and Cre08.g361250) were identified in the *C. reinhardtii* genome for this additional xylosyltransferase activity (Figure 19). It would therefore appear that xylosylation in *C. reinhardtii* is governed by a multigene family, which will provide guidance for future glycoengineering work.



**Figure 19 : Comparison of the amino acid sequences of different plant and *C. reinhardtii* xylosyltransferases and their phylogenetic relationships modified from (Lucas *et al.*, 2020).** The amino acid sequences of putative xylosyltransferases from *C. reinhardtii* were aligned to sequences corresponding to xylosyltransferases from *Arabidopsis thaliana* (At5g55500), *Solanum lycopersicum* (NP\_001311390.1), *Oryza sativa* (Os08g0503800), *Zea mays* (NP\_001105845.1), *Physcomitrella patens* (CAD22108.1) and *Lemna minor* (ABG89269.1). The protein sequences alignment was performed with ClustalW and then the tree (A) was deduced using the Neighbor-Joining method (Saitou and Nei, 1987). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985). (B) Sequences encoding the three new xylosyltransferases candidate genes (Cre10.g458950, Cre13.g588750 and Cre08.g361250) share a characteristic GT61 motif with other XTs in their C-terminal part. Conserved amino acid residues are shaded dark blue (identities) and light blue (similarities). Dashed lines represent gaps inserted for optimal alignment of the sequences.

## VII. Objectives of the thesis

In this context, the aim of this project was to glycoengineer the green microalga *Chlamydomonas reinhardtii* for biopharmaceutical production. Indeed, as previously demonstrated, the *N*-glycosylation of this species is very different from that of humans and is therefore incompatible. The thesis work focused on the creation of a strain lacking the immunogenic epitopes and the identification of signals capable of addressing glycoenzymes in the Golgi apparatus for a future tailor-made engineering. The modifications were monitored using a therapeutic glycoprotein reporter.

The SWEETBIOPHARM project is a French-German project financially supported by the Normandy region, France and the University of Münster, Germany. This thesis is also part of the XL-CHEM university research school in coordination with the Integrative Biology, Health and Environment doctoral school (ED NBISE No. 497). The work carried out by XL-CHEM has received support from the French government, managed by the Agence Nationale de la Recherche (ANR-18-EURE-0020), as part of the “Investissements d'Avenir” program, and from the Normandy Region as part of the RIN LABEL XL-CHEM.

### A. Characterisation of a therapeutic glycoprotein reporter: human erythropoietin

Among biopharmaceutical glycoproteins, the human erythropoietin (hEPO), a regulatory hormone involved in erythropoiesis and which can be used to treat anaemia, has been a textbook case since the 1980s demonstrating that *N*-glycosylation is essential and that the choice of production model is therefore just as important (Dubé *et al.*, 1988).

In 2009, Eichler Stahlberg and collaborators transformed the nuclear genome of the *cw15arg<sup>-</sup>* *Chlamydomonas reinhardtii*'s genetic background to produce the hEPO in the culture medium using ARS2 secretion signal peptide with an observed molecular weight of 33kDa which is close to the 34kDa in human (Eichler-Stahlberg *et al.*, 2009). Those results allow the hypothesis that a glycosylated EPO was produced for the first time in the green microalga. Indeed, this well studied hormone contains three *N*-glycosylation sites (Asn24, Asn38, Asn83) and one *O*-glycosylation site (Ser126) justifying the choice of using it as a reporter.

However, despite the use of transgenes containing multiple introns, expression yields were too low to collect enough protein for further and detailed structural analysis of the recombinant hEPO produced in *C. reinhardtii* (Eichler-Stahlberg *et al.*, 2009).

In agreement, the first objective of this project was to produce hEPO in *Chlamydomonas reinhardtii* strains with a sufficient yield to purify and characterise its *N*-glycosylation by both glycoproteomic and glycomic analysis. To achieve this, plasmids containing the codon optimized sequence of the hEPO were designed in the light of the latest advances in molecular biology in *C. reinhardtii*.

The publication arising from these results will constitute the first chapter of this manuscript and is, to the best of our knowledge, the first to report on the characterisation of the *N*-glycosylation of a therapeutic molecule expressed in *C. reinhardtii*.

## B. Elimination of immunogenic epitopes by genome editing

As 10 years of studies on *N*-glycosylation in *C. reinhardtii* have shown, complex glycans carry residues that are known to be immunogenic in humans, such as xyloses and fucoses (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020; Schulze *et al.*, 2018). The study of strains carrying mutations in the genes encoding the enzyme responsible for the graft of such residues and their crosses generated fucosylation-free progeny whose xylosylation activity was reduced but not eliminated, indicating possible redundancy and the potential presence of a multigenic family (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020).

The second objective was therefore to identify the genes putatively involved in xylosylation in the *C. reinhardtii* genome and to inactivate them by genome editing using CRISPR-Cas9 tools in order to create a strain producing proteins lacking immunogenic epitopes. The implementation of this approach will be detailed in Chapter II of this manuscript.

### C. Targeting the Golgi apparatus

As exposed previously in this introduction, *C. reinhardtii* synthesises non canonical oligomannosidic *N*-glycans ranging from Man<sub>5</sub>GlcNAc<sub>2</sub> to Man<sub>3</sub>GlcNAc<sub>2</sub> (Vanier *et al.*, 2017). Moreover, *N*-glycans do not contain terminal *N*-acetylglucosamine residues which is consistent with bioinformatic analyses that revealed a lack of the gene encoding GnTI, a Golgi enzyme that is necessary for the subsequent construction of mammalian biantennary *N*-glycans. These important differences raise the question of the engineering of the *N*-glycosylation pathway to make it suitable for the expression of recombinant glycoproteins compatible with human therapy.

Although it is now well known that the addressing of glycoenzymes in the different Golgi cisternae depends on the Cytoplasm, Transmembrane and Stem (CTS) regions (Schoberer and Strasser, 2011), the targeting mechanisms of enzymes in the Golgi apparatus remain unexplored in *C. reinhardtii*. The aim of this work, summarised in Chapter III of the results section of this manuscript, was therefore to understand the targeting of glycosyltransferases (GTs) in *C. reinhardtii* by mapping the distribution of CTS domains from either endogenous or heterologous type II glycoenzymes within the Golgi apparatus thanks to their fusion with a fluorescent protein.



# Materials & Methods





This chapter is dedicated to the materials and methods that have been used during the SweetBioPharm project. To avoid repetition, it has been decided not to go into further detail on the materials and methods related to the first chapter of the results, which is presented in the form of an article.

## I. Biological Material

### A. Bacterial strains

The *E. coli* DH5a strain (MAX Efficiency™ DH5a Competent Cells, Invitrogen™) has been used to replicate the plasmids.

### B. *Chlamydomonas reinhardtii* strains

#### 1) Wild type

- CC-5325 (CC-4533) cw15 mt-

This *Chlamydomonas reinhardtii* haploid strain is the result of a cross between strains 4A- (Dent *et al.*, 2005) and D66+ (Schnell and Lefebvre, 1993). It has been selected for its ability to withstand cryopreservation, its ability to be transformed by electroporation, its ease of crossing, its ability to be grown in the dark or in the light without forming aggregates in liquid culture, its ability to swim normally and finally its ability not to adhere to glass. It should also be noted that this genetic background has been used to construct the CLiP insertion mutant library (*Chlamydomonas* Library Project ([www.chlamylibrary.org](http://www.chlamylibrary.org))) and is identical to the genetic background CC-4533 cw15.

- UVM4

UVM4 strain has been kindly provided by Pr Ralph BOCK from the Max-Planck-Institut für Molekulare Pflanzenphysiologie, Potsdam under an MTA agreement. This strain carries mutations in a gene encoding for Sir2-like NAD(+) dependent protein deacetylase that may allow a better insertion of foreign DNA that newly enters the cells with no silencing and a better expression level (Neupert *et al*, 2020).

## 2) Mutants

- $IM_{XTA}IM_{XTB}$

The double mutant  $IM_{XTA}IM_{XTB}$  used in this study was generated by Lucas *et al.*, in 2020. Briefly, this double mutant results from crossing the insertional mutant  $IM_{XTB}$  obtained by insertion of the CIB1 cassette into Exon 1 of the Cre16.g678997 gene encoding the xylosyltransferase B, with  $IM_{XTA}$ , obtained by insertion of the CIB1 cassette into Exon 1 of the Cre09.g391282 gene encoding the xylosyltransferase A.

- $IM_{XTA}IM_{XTB}IM_{FucT}$

The triple mutant  $IM_{XTA}IM_{XTB}IM_{FucT}$  has been obtained by crossing an insertional mutant  $IM_{XTA}$  from the library described in (Cheng *et al.*, 2017), with the double mutant  $IM_{XTB}IM_{FucT}$  that was generated from strains LMJ.RY0402.049160 (Cre18.g749697, XP\_001695259.2) and LMJ.RY0402.118417 (Cre16.g678997, XP\_042916231.1) from the CLiP library as described in (Oltmanns *et al.*, 2020).

## 3) Chlamydomonas Spatial Interactome (CSI) strains

The CSI strains have been generated by expressing mVenus-tagged genes in the cMJ030 (CC-4533) genetic background (Mackinder *et al.*, 2017).

- FC2D02

This strain expresses Cre16.g661700-Venus-3xFLAG. Cre16.g661700 is a gene encoding a remorin type protein reported to be golgian (Mackinder *et al.*, 2017; Vilarrasa-Blasi *et al.*, 2021).

- FC1H12

This strain expresses Cre03.g171350-Venus-3xFLAG. Cre03.g171350 is the gene encoding a Sec61 protein that is an ER marker (Lang *et al.*, 2017).

- FC1E02

This strain expresses Cre10.g439500-Venus-3xFLAG. Cre10.g439500 is a gene encoding an unknown protein predicted to be golgian (Mackinder *et al.*, 2017).

## II. Culture conditions

### A. *E. coli*

The bacterial strains used have been grown in Lysogeny Broth medium (20 g of LB (LB Broth, Duchefa Biochimie) in 1 L of milliQ water) overnight at 37°C with continuous shaking at 150 rpm.

Bacterial strains have been kept either on plate containing Lysogeny Broth Agar medium (20 g LB broth with agar (Sigma L-2897) in 1 L of milliQ water) or at -80°C in glycerol stock (16% glycerol).

Where necessary, antibiotic supplementation was carried out at the concentrations described in the table below.

**Table 4 : Concentration of selection antibiotic for *E. coli***

Antibiotic	Concentration
ampicilin	10 µg/mL
spectinomycin	50 µg/mL

All experiments have been done under sterile conditions.

### B. *C. reinhardtii*

*C. reinhardtii* strains have been grown photoheterotrophically in Erlenmeyer flasks containing Tris-Acetate-Phosphate (TAP) medium (Gorman and Levine, 1965). The medium composition of which is described in Tables 5 to 8, contains acetate, a directly assimilable carbon source, enabling cells to grow more rapidly under continuous illumination.

Cultures have been carried out in continuous light (60 to 80 µE), at a constant temperature (25°C) on an orbital shaker (130 rpm). The durations and volumes have been adapted according to the needs of the different experiments and will be precise for each type of experiment described below.

For acetate requirement tests, spot cultures have been carried out on both TAP and Tris-Phosphate (TP) solid media in continuous light (60 to 80 µE), at a constant temperature (25°C).

All experiments have been done under sterile conditions.

**Table 5 : Detailed composition of the Tris-Acetate-Phosphate (TAP) medium**

TAP medium for 1 L (pH 7,0)	
Chemicals	Quantities
1M Tris base	2,42 g
Phosphate buffer (Table 7)	0,375 mL
Beijerinck's salts solution (Table 6)	25 mL
Traces elements (Table 8)	1 mL
Acetic acid	1 mL

TP medium is based on the same recipe as the TAP medium (Table 5), except for acetate, which is absent from the TP medium.

**Table 6 : Detailed composition of the Beijerinck's salts solution used in TAP medium.**

Beijerinck's salts solution (1 L of 40x)	
Chemicals	Quantities
NH <sub>4</sub> Cl	16 g
CaCl <sub>2</sub> , 2H <sub>2</sub> O	2 g
MgSO <sub>4</sub> , 7H <sub>2</sub> O	4 g

**Table 7 : Detailed composition of the phosphate buffer used in TAP medium.**

Phosphate buffer (100 mL)	
Chemicals	Quantities
K <sub>2</sub> HPO <sub>4</sub>	16 g
KH <sub>2</sub> PO <sub>4</sub>	4 g

**Table 8 : Detailed composition of traces elements solution used in TAP medium (Hutner *et al.*, 1950).**

Traces elements (1 L of 200x)	
Chemicals	Quantities
Na <sub>2</sub> EDTA	50 g
ZnSO <sub>4</sub> , 7H <sub>2</sub> O	22 g
H <sub>3</sub> BO <sub>3</sub>	11,4 g
MnCl <sub>2</sub> , 4H <sub>2</sub> O	5,06 g
CoCl <sub>2</sub> , 6H <sub>2</sub> O	1,61 g
CuSO <sub>4</sub> , 5H <sub>2</sub> O	1,57 g
(NH <sub>4</sub> ) <sub>6</sub> Mo <sub>7</sub> O <sub>24</sub> , 4H <sub>2</sub> O	1,10 g
FeSO <sub>4</sub> , 7H <sub>2</sub> O	4,99 g

All volumes have been adjusted with milliQ water.

Where necessary, antibiotic supplementation was carried out at the concentrations described in the table below (Table 9).

**Table 9 : Concentration of selection antibiotic for *C. reinhardtii*.**

Antibiotic	Concentration
Hygromycin	15 to 25 µg/mL
Paromomycin	15 µg/mL
Zeocin	15 µg/mL
Nourseothricin	15 µg/mL

### III. Biochemical analysis

#### A. Whole cell proteins extraction

Cell pellets centrifuged at 4500 g (RC-5C PLUS, Sorvall) for 15 min have been resuspended in 1 mL protein extraction buffer (Table 10).

**Table 10 : Composition of the protein extraction buffer.**

Protein extraction buffer	
Chemicals	Concentration
Tris Base	35 mM
Tris-HCl	26,25 mM
EDTA	125 mM
Lithium dodecyl sulfate	0,5%
Glycerol	2,5%

Cells have been then lysed in 2 mL tubes containing ceramic beads (Lysing Matrix D MP Biomedicals™) using a bead beating grinder (FastPrep-24™ 5G, MP Biomedicals™). Cell lysis has been performed at 6.5 m.s<sup>-1</sup> for 30 s with a 2 min pause in ice between each cycle. This has been repeated 5 times.

The mixtures have been then centrifuged for 10 min at 10.000 g to collect the protein-containing supernatant. The addition of 1mL of lysis buffer to the cell debris pellet enabled a second extraction to be performed under the same conditions as before. The supernatants were pooled and dialysed against milliRo water for 48 hours at 4°C (Dialysis Tubing 3.5 kDa, Fisherbrand).

Protein were quantified using a BCA protein assay (Pierce™ BCA Protein Assay Kits, ThermoScientific) with Bovine Serumalbumin (BSA) standards ranging from 25 µg.mL<sup>-1</sup> to 2 mg.mL<sup>-1</sup>.

## B. Culture medium proteins isolation

This method has been used to generate the results presented in the second chapter of results during my stay at the University of Münster, Germany.

Cells have been cultivated in a volume of 100 mL TAP medium in a sterile 300 mL Erlenmeyer. Cell density has been adjusted to reach a final concentration of 30 µg/ml chlorophyll after 4 days in normal light (60 to 80 µE) on an orbital shaker at 130 rpm.

After 4 days of culture, supernatants have been collected by centrifugating the culture at 4.500 g for 5 minutes, then transferred to Beckmann centrifugation jars and kept on ice from this step.

Cell debris have been pelleted at 48.000 g for two hours at 4°C.

Cleaned supernatants have been then transferred to Schott bottles and kept on ice while step by step concentrated to a volume of about 1,5 mL in Amicon® Ultra-15 Centrifugal Filter Unit (30 kDa MWCO) (Merck, Millipore®) 5 min to 20 min, 4500 g at 4°C.

After concentration, the samples have been stored at -80°C until further processing.

## C. SDS-PAGE

Proteins have been denaturated in Loading sample buffer (LSB) (Table 12) at 100°C for 10 min prior loading them on a Tris-glycine electrophoresis gel, with a 7% acrylamide stacking gel and a 10 to 12% separating gel (Table 11). A molecular weight marker (Thermo Scientific™ PageRuler™ Prestained Protein Ladder, 10 to 180 kDa) has been loaded on the gel as well. The proteins are separated in Tris-Glycine migration buffer (Table 13) at 120 V. The gel has been rinsed briefly with MilliRo water before a Coomassie blue staining or transferring the proteins onto a nitrocellulose membrane for Western Blot analysis.

**Table 11 : Composition of acrylamide stacking and separating gels.**

<b>Stacking gel</b>	<b>7%</b>	
<b>Chemicals</b>	<b>Quantities for 5 mL</b>	
H <sub>2</sub> O mQ	3,65 mL	
Acrylamide 40%	850 µL	
Tris 1,5 M pH6,8	417 µL	
SDS 10%	50 µL	
APS 10%	50 µL	
TEMED	4 µL	
<b>Separating gel</b>	<b>10%</b>	<b>12%</b>
<b>Chemicals</b>	<b>Quantities for 5 mL</b>	
H <sub>2</sub> O mQ	2,4 mL	2,15 mL
Acrylamide 40%	1,25 mL	1,5 mL
Tris 1,5 M pH8,8	1,25 mL	1,25 mL
SDS 10%	50 µL	50 µL
APS 10%	50 µL	50 µL
TEMED	2 µL	2 µL

**Table 12 : Composition of the Loading sample buffer.**

<b>Loading sample buffer 5X</b>	
<b>Chemicals</b>	<b>Quantities for 30 mL</b>
Tris HCl 1 M pH6,8	9,375 mL
SDS	3 g
β-mercaptoethanol	3,75 mL
Glycerol	15 mL
Bromophenol blue	A spatula tip

**Table 13 : Composition of the Tris-Glycin SDS-PAGE electrophoresis migration buffer.**

<b>Tris-glycin SDS 10X buffer pH 8,5/9</b>		
<b>Chemicals</b>	<b>Quantities for 5 L</b>	<b>Final concentration 10X</b>
Tris base	151,425 g	250 mM
Glycine	720,0672 g	1,92 M
SDS	50 g	1%

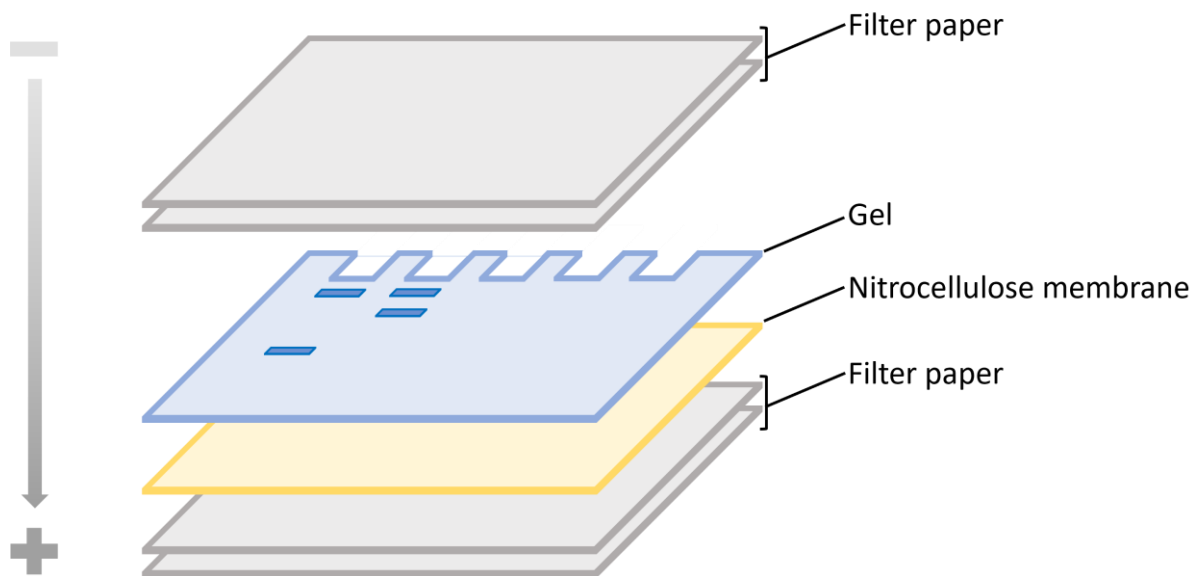


## D. Western Blot

Proteins separated by SDS-PAGE have been transferred to a nitrocellulose membrane (0,2  $\mu$ , Amersham™ Protran®). The transfer has been carried out using a tank blotting cell (Mini Trans-Blot® Cell and Criterion™ Blotter, Biorad) or using the Pierce™ Power Blotter system (Thermo Scientific™) and semi-dry blotting method according to the manufacturer's instructions. Two Whatman papers (Invitrogen) soaked in transfer buffer (1X Tris-glycin SDS with a final concentration of 20% ethanol) have been placed one on top of the other on the positive terminal of the cassette, followed by the nitrocellulose membrane, the gel followed by two new Whatman papers soaked in buffer (Figure 20). The transfer has been then performed at 65 V for 2 hours.

### 1) Ponceau Red

After the transfer, the nitrocellulose membrane has been then soaked for 5 min in a Ponceau Red solution (Ponceau Red (Serra, C.I.27195) 1% (w/v) in 3% trichloroacetic acid), then rinsed with water to check that protein transfer has been carried out correctly.



**Figure 20 : Schematic diagram of the stacking in the cassette for western blot transfer of proteins.**

## 2) Nitrocellulose membrane saturation

The nitrocellulose has been briefly incubated with Tris Buffered Saline (TBS) (Table 14) buffer (10 mM Tris HCl, 500 mM NaCl pH7.5) containing 0.1% Tween (TWEEN 20, Sigma Aldrich) (TBST) to remove all the Ponceau red, then the membrane has been saturated overnight in TBST with 3% milk (w/v) (Skimmed milk powder, Régilait) on an orbital shaker to saturate the nitrocellulose membrane and prevent any future aspecific recognition.

**Table 14 : Composition of the Tris-Buffered Saline buffer.**

TBS (Tris-Buffered Saline) 10X pH 7,5	
Chemicals	Quantities for 5 L
Tris base	121,14 g
NaCl	1461 g

## 3) Primary antibody incubation

The membrane has been incubated with agitation on a carousel in sealed plastic bags with the primary antibody diluted in 3% milk TBST according to the dilution described in Table 15. After 3 hours of incubation, the membranes have been rinsed 4 times with TBST for 10 min under agitation to remove all primary antibodies that do not bind to the targeted proteins and potential milk lumps.

**Table 15 : Dilutions of primary antibodies for western blot analysis**

Antibody	Dilution
Anti-HRP (Oltmanns <i>et al.</i> , 2020)	1/666

## 4) Secondary antibody incubation

The membrane has been incubated with the horseradish peroxidase (HRP)-coupled secondary antibody (Goat anti-Rabbit or Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP from Thermo Fisher Scientific, Invitrogen A16041, and 65-6120) in TBST at 1:30.000 dilution under agitation for 1 hour. To remove excess of secondary antibodies, the membrane has been washed 4 times for 10 min with TBST, then rinsed 3 times with TBS for 10 min to remove excess tween that could interfere with revelation based on a chemiluminescence reaction.

## 5) Chemiluminescence reaction

Revelation of antibody-bound proteins has been performed using the PICO-plus kit (SuperSignal™ West Pico PLUS Chemiluminescent Substrate, ThermoFisher Scientific). The membrane has been incubated with a mixture of 2mL each of the two reagents (Peroxide solution and luminol solution) from the PICO-plus kit. The reagent has been then removed from the nitrocellulose membrane and the chemiluminescent signal has been collected using a Vilber FX imager.

## E. *N*-glycans isolation.

### 1) *N*-glycans cleavage using *N*-glycosidase F (PNGase F).

PNGase F is an enzyme that cleaves all types of protein-bound *N*-glycans with the exception of *N*-glycans with an  $\alpha(1,3)$ -fucose residue bound to the proximal *N*-acetylglucosamine (GlcNAc) (Maley *et al.*, 1989).

An amount ranging from 1 to 2 mg of total protein, extracted as described in in part III. A. of this material and methods chapter, was resuspended in 500  $\mu$ L of 100 mM Tris-HCl buffer pH8 with 0.1% SDS (w/v), then, underwent denaturation by heating in a water bath at 100°C for 10 min. Once the samples had cooled to room temperature, a volume of 500  $\mu$ L of 100 mM Tris-HCl buffer pH 8 containing 0.5% Nonidip (IGEPAL) (v/v) was added to neutralise excess SDS. Next, 10 U of peptide *N*-glycosidase F (PNGase F) (100 units in 0.1 mL, Roche) were added per tube, before incubation in a thermoblock at 37°C for 24 h with continuous shaking at 850 rpm.

### 2) Ethanol precipitation of proteins

The deglycosylated samples were taken up in 4 times their volume with 100% HPLC grade ethanol and left at -20°C for 72 h to precipitate the proteins. After centrifugation for 15 min at 10.000 g (Allegra X-30R centrifuge, Beckman Coulter), the ethanolic supernatant containing the free *N*-glycans was recovered, then dried for 3 h under an air ramp before the purification step.

### 3) *N*-glycans purification

PNGase F-released *N*-glycans have been purified using a carbograph (Hypersep Hypercarb, Thermo Scientific) column on a vacuum ramp following the protocol described below.

Column conditioning:

- Step 1: 1 mL MeOH
- Step 2: 1 mL NaOH 1 M
- Step 3: 2 mL H<sub>2</sub>O
- Step 4: 2 mL acetic acid 30%
- Step 5: 2 mL H<sub>2</sub>O
- Step 6: 1 mL acetonitrile 50%, TFA 0,1%
- Step 7: 2 mL acetonitrile 5%, TFA 0,1%

The sample obtained in the previous step is resuspended in 500 µL H<sub>2</sub>O, then this volume is placed on the column. The flowthrough is collected.

The column has been washed with 2 mL H<sub>2</sub>O, followed by 1 mL acetonitrile 5%, TFA 0.1%. The *N*-glycans have been then eluted with 1 mL acetonitrile 50%, TFA 0.1%. Acetonitrile has been evaporated under air and the sample is then freeze-dried.

#### 4) 2-Aminobenzamide (2-AB) labelling

A quantity of 5 mg of 2-AB has been weighed and solubilised in an Eppendorf tube with 100 µL of a mixture of DMSO/glacial acetic acid (70/30, (v/v)). This mixture has been then transferred to another tube containing 6 mg of sodium cyanoborohydride (NaBH<sub>3</sub>CN). The whole mixture has been homogenised by pipetting up and down and formed the 2-AB mix. A volume of 10 µL of the 2-AB mix has been added to the lyophilise *N*-glycans, then the tubes have been vortexed and centrifuged before being incubated in a water bath at 60°C for 2 h. Following this incubation step, the samples have been cooled to room temperature and then centrifuged at 10.000g for 10 sec to ensure that the glycans have been at the bottom of the tube.

The 2-AB-labelled *N*-glycans have been then purified using Ludger D1 columns. Briefly, columns have been prepared by washing them successively with 1 mL of water, 5 mL of acetic acid, 3 mL of acetonitrile. Finally, an extra acetonitrile washing with a volume of 1mL have been performed.

The samples have been then loaded on the cartridge by spreading them on the entire membrane surface and have been left for 15min to allow the glycans to adsorb. Cartridges have been washed with 1mL acetonitrile followed by 5 x 1mL of a 96% acetonitrile solution. Glycans have been eluted with 3 x 0,5mL of water. Samples have been freeze-dried until further use.

## F. Filter Aided Sample Preparation (FASP) for glycoproteomic analysis.

Filter Aided Sample Preparation (FASP) has been performed loading 60 to 100 µg protein onto Amicon ultra centrifugal filters (0.5 mL, 30 kDa MWCO, Millipore). Samples have been concentrated to 30 µL by centrifugation at RT and 14k g.

After disulfide bonds have been reduced using 200 µL of 100 mM DTT in buffer A (100 mM Tris pH=7,6 supplied with 8 M urea) for 30 min at RT, samples have been centrifuged at 14000 g for 15 min (if not indicated otherwise, following centrifugation steps resemble these settings).

After two washing steps with buffer A, cysteine residues have been derivatized with carbamidomethyl residues (100 µL of 50 mM iodoacetamide in buffer A; incubation in the dark for 20 min at RT). Three additional washing steps with 100 µL buffer A have been followed by five washing steps with 200 µL of 50 mM ammonium bicarbonate (ABC) buffer (centrifugation at 14k g for 10 min).

Proteins have been digested using trypsin (sequencing-grade modified, Promega, Madison, WI, USA; trypsin to protein ratio of 1/50 to 1/100 in 50 mM ABC buffer) for 16 h at 37°C.

Peptides have been eluted by centrifugation and washing two times with 50 µL H<sub>2</sub>O and afterwards have been dried in a vacuum centrifuge and stored at -80 °C until the measurement.

## IV. Mass spectrometry

### A. N-glycans profile analysis by MALDI-TOF-MS

MALDI-TOF (Matrix Assisted Laser Desorption Ionisation - Time Of Flight) were carried out using an Autoflex III mass spectrometer (Bruker Daltonics (Bremen, Germany)). The instrument is equipped with a neodymium-doped yttrium-aluminium garnet laser (Nd: YAG laser, wavelength 355 nm). The spectra were acquired in positive reflectron mode over a mass range from 500 m/z to 4.000 m/z at a laser frequency of laser frequency of 100 Hz, and a minimum average of 3.000 individual spectra.

The instrument's voltage acceleration was set at 19 kV. Calibration of the MALDI-TOF was performed using monocharged monoisotopic ions of angiotensin II (5 pmol.µL<sup>-1</sup>), substance P (5 pmol.µL<sup>-1</sup>), adrenocorticotrophic hormone (ACTH) 18-39 (5 pmol.µL<sup>-1</sup>) and insulin B (10 pmol µL<sup>-1</sup>). 2,5-dihydroxybenzoic acid (DHB) was used as a matrix. A volume of 1 µL of each sample was mixed in 1 µL of matrix DHB matrix (20 mg.mL<sup>-1</sup> in ACN/water (70/30: v/v) containing 0.1% TFA (v/v)) and then 0.7 µL of the sample/matrix mixture was deposited and dried on the MALDI plate (Bruker Daltonics, Bremen, Germany). FlexControl 3.3 and FlexAnalysis 3.3 were used to analyse the spectra obtained. analyse the spectra obtained.

The theoretical masses of the *N*-glycans were obtained using GlycoWorkBench 2.0 software (Ceroni *et al.*, 2008).

## B. LC-ESI MS/MS Analysis

### 1) Culture medium protein analysis of the putative XT mutants

This protocol is adapted from the one described in (Oltmanns *et al.*, 2020).

Tryptic peptides were reconstituted in 2% (v/v) acetonitrile/0.1% (v/v) formic acid in ultrapure water and separated with an Ultimate 3000 RSLC nano System (Thermo Scientific). Subsequently, the sample was loaded on a trap column (C18 PepMap 100, 300 µm x 5 mm, 5 mm particle size, 100 Å pore size; Thermo Scientific) and desalted for 5 min using 0.05% (v/v) TFA/2% (v/v) acetonitrile in ultrapure water with a flow rate of 10 µL.min<sup>-1</sup>. Following, peptides were separated on a separation column (Acclaim PepMap100 C18, 75 mm i.D., 2 mm particle size, 100 Å pore size; Thermo Scientific) with a length of 50 cm. Glycoproteomic data obtained with this method have been analysed using an in-house developed Python tool SugarPy (Schulze *et al.*, 2020).

Ursgal: Version 0.6.9

Search engines: XTandem\_alanine, msgfplus\_v2019\_07\_03, msfragger\_3\_0

Validation engine: percolator\_3\_4\_0

Sugarpy: Version 1.0.0

## V. Molecular biology

### A. Plasmids

#### 1) CTS domains

CTS domains in endogenous genes (Cre10.g439500, Cre16.g661700 and Cre16.g678997) have been identified *in silico* and heterologous CTS domains from plant or rat glycoenzymes (AtGnTI, AtXylT, AtMNSI, RnST) used in other glycoengineering projects have been selected.

Sequences coding for endogenous CTS domains have been amplified from gDNA (Table 16 and 17) whereas sequences coding for CTS domains from plant and rat glycoenzymes have been optimized and synthesized to be inserted into backbone plasmids allowing the expression of C-terminal mCherry-tagged CTS domains.

**Table 16 : : Primers sequences for CTS domain amplification**

Endogenous	Forward 5'→3' Sequence	Reverse 5'→3' Sequence
Cre10.g439500	ATGGCTAAACCAACGCACGC	CAGGTTGGGGATAGGCGAG
		CGAAGACGTGTTGGGGAATG
Cre16.g661700	ATGTCTCCAAGGATCGTGGT	CTTGTAGGCTATGACCTTGC
		CTTGCCGGTGGGGCTGACC

A first series of 8 plasmids based on the pBR9 plasmid (Rasala *et al.*, 2013) has been cloned. For this purpose, primers have been designed to amplify CTS sequences by PCR (Q5® High-Fidelity DNA Polymerase, New England Biolabs® Inc) with a 5' Apal and a 3' XhoI restriction sites prior fusing them to the 3' part of a zeocin resistance cassette (BLE) that contains two copies of the first RBCS2 intron *via* the sequence encoding a foot-and-mouth disease virus 2A peptide (FMDV2A).

A volume equivalent to 50 ng of backbone vector was used in a ligation mix consisting of 1 µL of 10x Buffer and 0.3 µL of T4 DNA ligase (New England Biolabs® Inc.). The volume was adjusted to 10 µL with the insert solution. The mix was then incubated overnight at 4°C before being used for *E. coli* transformation.

The expression of such fusion proteins is driven *via* the combination of the pAR (RBCS2/HSP70A) tandem promoter with RBCS2 3'UTR.

A second series of 8 plasmids has been produced based on pLM006 from Luke Mackinder/Martin Jonikas, University of York, UK/Princeton, April 2017 using Gibson assembly. Expression of fusion proteins in this plasmid is driven by the combination of pSAD promoter with RBCS2 3'UTR. The plasmid contains a second transcriptional unit allowing the resistance to hygromycin. CTS domains have been amplified by PCR (Q5® High-Fidelity DNA Polymerase, New England Biolabs® Inc) with primers adding 15-20 bp overlapping ends prior assembling them with the backbone plasmid (Gibson Assembly® Cloning Kit, New England Biolabs® Inc). Finally, the amplified sequences have been cloned in pLM006 using HiFi assembly (NEBuilder® HiFi DNA Assembly Master Mix, New England Biolabs® Inc).

In both series an ampicillin resistance cassette has been used to select *E. coli* colonies transformed to amplify the plasmid.

Endogenous putative CTS domain sequence from Cre16.g678997 (*XTB*) gene has been amplified by PCR from CC5325 *C. reinhardtii* gDNA extraction using a high-fidelity DNA polymerase (Q5® High-Fidelity DNA Polymerase, New England Biolabs® Inc). Primers have been predicted using the Snapgene Gibson assembly tool and ordered to Metabion International. The amplicon has been then purified (NucleoSpin Gel and PCR Clean-up, Macherey-Nagel, Düren, Germany) and inserted into a backbone plasmid containing the gene encoding a 3' terminus 6HN-tagged fluorescent protein according to the protocol provided with the Gibson assembly kit. A backbone vector of ours, whose sequence has been optimised based on the strategy described in Baier et al, 2018, with the tool described in Jaeger et al, 2019 (Baier, Wichmann, et al., 2018; Jaeger et al., 2019), has been used. The 5' part of the putative *XTB* CTS domain has been fused to the 3' part of a nourseothricin resistance cassette (NAT1) via the sequence encoding FMDV2A.

Three copies of the first RBCS2 intron as well as one copy of the second RBCS2 intron were added. The expression of this optimized sequence was then driven via the combination of the pAR (RBCS2/HSP70A) tandem promoter with RBCS2 3'UTR. A short linker has been inserted between the sequences encoding the N-terminal transmembrane helix of *XTB* and the fluorescent protein.

Finally, a spectinomycin resistance cassette has been used to select *E. coli* colonies transformed to amplify the plasmid.



For plasmids cloned using HiFi assembly, a volume equivalent to 50 ng of backbone vector was used with 2-fold excess of each insert. Then, 10 µL of HiFi assembly master mix has been added prior adjusting the volume to 20 µL with deionised water.

Samples have been incubated in a thermocycler at 50°C for 15 minutes before being stored on ice or at –20°C for subsequent transformation.

After cloning, plasmids integrity has been verified by Sanger sequencing.

**Table 17 : Primers sequences for HiFi assembly**

	<b>5'→3' Sequence</b>
XTB-CTS domain Forward	TGGAGAGCAACCCGGGCCCGTGCAGCAAATTACACCATCGG
XTB-CTS domain Reverse	TTGCTCACGCGGCCCTCGATCCTAGAGCCTAGGTACATGACCATG
Backbone plasmid Forward	ATCGAGGGCCGCGTG
Backbone plasmid Reverse	GGGGCCCGGGTTGC

## 2) CRISPR/Cas9

The pPH360 plasmid containing the codon optimized gene coding the *Streptomyces rimosus* aminoglycoside phosphotransferase aphVII selection marker (hygromycin B resistance) driven by the HSP70A/RBCS2 tandem promoter and RBCS2 3' UTR, has been used to select the transformed colonies (Sizova *et al.*, 2021). The plasmid has been ordered on the online Chlamydomonas resource centre.

### B. Genomic DNA extraction

gDNA extraction from *C. reinhardtii* has been performed on colonies cultured on TAP agar medium for 4 days. A small quantity of cells (volume equivalent to 1 µL) has been collected with an Eppendorf pipette tip and re-solubilised in PCR tubes with 20 µL of 10 mM EDTA in milliQ water. Samples have been heated at 100°C for 10 min. After homogenisation by vortexing, the tubes have been centrifuged for 1 minute at 10.000g. This gDNA has been then used as template in the PCR reactions.

### C. PCR and nucleic acids electrophoresis

PCR, or polymerase chain reaction, is a molecular biology technique used to amplify a portion of DNA between two defined DNA segments called primers using an enzyme called polymerase. In this work, different primer pairs have been used to amplify genomic DNA (gDNA) and complementary DNA (cDNA) from several strains of *C. reinhardtii*.

Cycles and PCR mix have been adapted according to the manufacturer's recommendations depending on the kit.

Few  $\mu\text{L}$  of PCR reaction mix were analysed by agarose gel electrophoresis prepared with 0.5% (v/v) Tris Acetate EDTA (TAE) buffer (Invitrogen™ Ultrapure™ TAE Buffer) containing 1% to 3% (w/v) agarose depending on the size of the amplicon to analyse and 1% (w/v) SafeView™ or Midori green advance (Nippon genetics). The electrophoretic profile of the agarose gel is then revealed on a UV imager (E-BOX CX5 EDGE Vilber-Lourmat) using the default settings.

### D. *C. reinhardtii* transformation

From a 48-hours pre-culture, 400 mL of culture has been brought to a cell density of around  $10^6$  cells/mL. A volume corresponding to  $2 \times 10^8$  cells has been centrifuged at 1.000 g for 5 min at room temperature. The cell pellet has been resuspended in 1 mL TAP 40 mM sucrose and incubated for 5 min at  $16^\circ\text{C}$ .

The transformation mix has been prepared directly in the 2 mm electroporation cuvette by successive addition of 125  $\mu\text{L}$  of cell suspension and a volume equivalent to 1  $\mu\text{g}$  of linearised plasmid, homogenised by gentle agitation.

Linearisation was carried out in a 50  $\mu\text{L}$  reaction volume containing 5  $\mu\text{L}$  of 10X NEB Buffer and a volume equivalent to at least 1  $\mu\text{g}$  of DNA. The volume was adjusted with sterile deionised water before being incubated at  $37^\circ\text{C}$  for a minimum of one hour. The linearised plasmid was then purified on an agarose gel.

The electroporator is set to 400 V and 25  $\mu\text{F}$  and the electric shock have been applied.

The transformed cells have been then transferred to a 50 mL flask containing 10 mL of TAP 40 mM sucrose and placed in the dark at  $25^\circ\text{C}$  with 130 rpm shaking to recover overnight.

The following day, the cells have been centrifuged at 1000 g for 5 minutes at room temperature. The supernatant has been discarded, and the cell pellet has been resuspended in the last drop of medium before being plated on 2% TAP agar supplemented with the selection antibiotic (Table 16). The plates have been then placed in the dark overnight before being transferred in low light (25  $\mu$ E) for 2 days and then placed in normal light until colonies appeared.

#### E. CRISPR/Cas9 gene editing protocol

In order to find the best candidate sequences for the target gene, first the coding sequence (CDS region) has been copied from Phytozome (<https://phytozome-next.jgi.doe.gov/>, search in *Chlamydomonas reinhardtii* v5.6 database) and pasted in the webtools CRISPor (<http://crispor.tefor.net/>) and CRISPRdirect (<https://crispr.dbcls.jp/>). In these webtools, the PAM sequence requirement has been set to “NGG” specific to *Streptococcus pyogenes* Cas9 (S.p. Cas9) and the genome/specificity check has been selected for *Chlamydomonas reinhardtii*.

Once the servers found the possible targets, in CRISPRdirect, the list has been narrowed down to the “highly specific” targets only, and then these sequences have been searched into the list of results from CRISPor, to confirm that the second webtool also determined them to have a high specificity score, and to make sure no “Inefficient” errors have been suggested for that particular target. Based on these analyses, two targets have been selected, ideally in the first 1/3 of the CDS of the gene of interest.

Once targets (Table 18) have been identified, the CRISPR RNA (crRNA), trans-activating CRISPR RNA (tracrRNA), S.p. Cas9 enzyme (IDT-DNA Alt-R® S.p. Cas9 Nuclease V3), and reaction buffer have been ordered from Integrated DNA Technologies (IDT) (<https://eu.idtdna.com/pages>). The IDT-DNA duplex buffer, provided by IDT, is a commercial buffer suitable for CRISPR experiment using IDT materials.

Generation of mutants has been done based on (Kelterborn *et al.*, 2022). Briefly, cells of different genetic backgrounds have been grown in TAP medium under moderate light conditions (about 60-80  $\mu$ E illumination, 16h:8h “light:dark” regime) from TAP 2% agar plates for 3 days, re-diluted with fresh TAP medium and grown for 3 extra days, before the cell density has been adjusted to about  $0.5 \times 10^6$  cells/mL.

**Table 18 : Sequences of CRISPR/Cas9 targets and primers on each putative XT.**

Gene	Target 5'→3'	Forward primer 5'→3'	Reverse primer 5'→3'
Cre08.g361 250	CTCCGTGTTTCGTGTCA CGGG	TAACTTCCGACCCCAAT CGC	GTACTGCAACCCCACAT AGC
	TGCTGCGGAAATCGAG GCTG	CTATGTGGGGTTGCAGT ACCT	GTGGCTTTGGCTTTGAC TACG
Cre10.g458 950	AGCGCTTAGATCCTAT GTGC	CCAGCGAAGCTTTTGTG TCC	GGCTGCTTGTTTTCTGA GGC
	ATACTTCTGGTCCGTG ATAA	ATGATGCAATGCGGTTC GTG	CGGTACTTAGCCAACG GTGA
Cre13.g588 750	AAGATGGACTTTCGGG AAAA	GCTGAGGCGCAGAGTT CTAT	AAGACTGATGCTCACCT CGC
	GCTGCGGCTCAAAAAA CGGA	CCTTGCTGCAAATCACA CCC	CTGCTCCGGATGGAGT CAAG

On the morning of the transformation, the cell density has been checked, and  $1.0 \times 10^8$  cells have been harvested via centrifugation (2.000 g, 5 min, RT). After carefully removing the supernatant, cells have been re-suspended in 1 mL of MAX Efficiency™ Transformation Reagent for Algae (Thermo Fisher Scientific) supplemented with 75 µL of 0,22 µm filter-sterilized 20% sucrose and transferred into a 1.5 mL conical tube for incubation at 40 °C with 450 rpm shaking for 30 minutes. After the heat-shock, cells have been allowed to recover at RT for 30 minutes.

In parallel, RiboNucleoProteins (RNPs) and FLAG<sub>v3</sub> solutions have been prepared. First, guide-RNAs (gRNAs) have been assembled by mixing 1 µL tracrRNA with 1 µL crRNA (the volume has been completed to 10 µL with 8 µL of IDT-DNA duplex buffer) and incubation at 95 °C for 2 minutes, before ramping down to 20 °C at a rate of -0,1 °C/s. Then, RNPs have been assembled by mixing 9 µL of gRNA with 1.5 µL of IDT-DNA Alt-R® S.p. Cas9 Nuclease V3 and 3 µL of NEB 10X 3.1 buffer (New England Biolabs®) (volume has been completed to 30 µL by adding 16.5 µL of double-distilled, autoclaved water) and incubation at 37 °C for 15 min.

The RNPs have been kept on ice until future used. For the FLAG<sub>v3</sub> preparation, a sequence that contains several stop codons sense and anti-sense, 15 µL of the pre-

ordered synthetic forward and reverse sequences (Table 19) have been mixed with 120 µL of IDT-DNA duplex buffer and incubated at 95 °C for 2 min, before ramping down to 20 °C at a rate of -0,1 °C/s.

**Table 19 : 30 bp FlagV3 primers sequences**

Oligo name	Sequence (5' – 3')
FLAG <sub>v3</sub> Fw	T*T*A*GCTAAGCCTCCCCAAAGCCTGGCCAGGGTC*T*A*G
FLAG <sub>v3</sub> Rv	C*T*A*GACCCTGGCCAGGCTTTGGGGAGGCTTAGC*T*A*A

An electroporation master mix has been prepared by mixing 120 µL of cells with 3 µL of a 1 µg/µL antibiotic selection plasmid, 6 µL of FLAG<sub>v3</sub>, and 15 µL of RNP. The 2-mm gap electroporation cuvettes have been pre-washed with TAP medium to make sure cells do not stick to it before 40 µL of the electroporation master mix have been transferred into them and used for each electroporation replicate (3 replicates in total). In the NEPA21 Super Electroporator (NEPAGENE), the following parameters have been set:

- Pore forming pulse: Voltage (300 V – cell wall containing strains; 200 V – cell wall deficient strains), Number of pulses (1 for cell wall containing strains; 2 for cell wall deficient strains), Length (8 ms), Interval (50 ms), Decay rate (40%), Polarity (+)
- Transfer pulse: Voltage (20 V), No. (5), Length (50 ms), Interval (50 ms), Decay rate (40%), Polarity (+/-).

Immediately after electroporation, the cuvettes have been washed with 0.5 mL of fresh TAP medium, and the cells have been transferred into 24-well plates for overnight incubation under low light intensity (about 20 µE illumination, 16h:8h “light:dark” regime), after which they have been plated in TAP 2% agar plates supplemented with selection antibiotic (Table 9).

After 7-10 days, colonies start to appear and are transferred into new TAP 2% agar plates, before being randomly selected for colony PCR, using the Phire Plant Direct PCR Master Mix (Thermo Fisher Scientific) according to the manufacturer’s instructions and primers spanning the genomic region of interest, where CRISPR is supposed to act.

After PCR, run samples on a 3% TAE-agarose gel.

## VI. Microscopy and fluorescence screening

### A. Fluorescence screening

The positive clones expressing a fluorescent protein have been screened based on the fluorescence level. Individual colonies grown on solid TAP medium supplemented with selection antibiotic (Table 16) were picked and resuspended in a small volume of TAP (160  $\mu$ L to 1mL) without antibiotic into a well plate.

After 5 days on an orbital shaker at 25°C and normal light conditions, cultures have been transferred in a black well plate with clear bottom (Corning® Costar®) and the fluorescence has been analysed using a plate reader with fluorescence detector such as the FlexStation® 3 Multi-Mode Microplate Reader (Molecular Devices LLC, San Jose, CA) according to the fluorescent protein's intrinsic properties provided by the webtool FPbase (Lambert, 2019).

Fluorescence was read in Endpoint+ mode from the bottom. A cutoff was chosen according to the emission and excitation wavelengths of the fluorescent proteins. The number of readings per well was set to 6 and the photomultiplier sensitivity was set to medium.

Fluorescence has been normalized to OD<sub>720</sub> to take account for differences in cell density.

### B. Confocal microscopy

Confocal microscopy analyses have been realised in collaboration with the PRIMACEN microscopy platform.

Colonies that have been screened by fluorescence as described above, have been cultivated in the same conditions to analyse them by confocal microscopy.

Both Leica SP5 X and Leica SP8 VIS/MP equipped with hybrid detectors have been used for observations of *C. reinhardtii* cells immobilised on Polysine™ Adhesion Microscope slides (Epredia™).

Excitation and emission wavelength used for confocal microscopy analyses are listed in Table 20.

**Table 20 : Excitation and emission wavelength used for confocal microscopy analysis.**

Fluorescent protein	$\lambda$ Excitation (nm)	$\lambda$ Emission (nm)
Autofluorescence	514	685/40
mVenus	514	543/22
mCherry	561	600/20
mClover	514	543/22

## VII. *In silico* analysis

### A. Transmembrane domain prediction

CTS domains from various proteins have been predicted using the TMHMM online prediction tool (TMHMM 2.0) (Krogh *et al.*, 2001) and DEEPTMHMM1.0 (Hallgren *et al.*, 2022)

### B. Multiple sequence alignment

Multiple sequence alignments have been performed using the Clustal Omega online tool on the *EMBL's* European Bioinformatics Institute (Madeira *et al.*, 2022).

### C. Prediction of the three-dimensional structure of proteins

All 3D protein structures have been predicted using the Phyre<sup>2</sup> webtool in intensive mode (Kelley *et al.*, 2015).

Chapter 1: Fine-tuning the *N*-glycosylation of recombinant human erythropoietin using *Chlamydomonas reinhardtii* mutants





## I. Preamble

This chapter presents the first characterisation of the *N*-glycosylation of recombinant human erythropoietin (rhEPO), a glycoprotein of therapeutic interest, expressed in different glycomutants of the green microalga *Chlamydomonas reinhardtii*. This original work has been published in the Plant Biotechnology Journal in June 2024.

As indicated in the introduction to this manuscript, the vast majority of biopharmaceutical health products are glycoproteins. *N*-glycosylation is well known to date to influence the pharmacokinetic parameters and functionalities of such molecules. Current production organisms, such as CHO cells, have therefore been chosen for their ability to carry out these post-translational modifications as closely as possible to those found in humans. However, the costs involved in using such production systems are particularly high, not only because of the cultivation requirements, but also because of the need to use extensive purification methods due to the sensitivity of these cell cultures to contamination by pathogens such as viruses.

In view of the growing demand for biomedicines, the scientific community is striving to find alternative production models capable of *N*-glycosylation, compatible with current cellular production lines and offering lower production costs. Microalgae have therefore emerged as a potential alternative production system for biopharmaceuticals. Until now, it has been possible to express rhEPO, notably in *C. reinhardtii* in 2009 by Alke Eichler-Stahlberg and collaborators. However, despite hypotheses concerning the capacity of this organism to glycosylate recombinant human proteins, to date, no study has been able to characterise the glycosylation profile of such proteins.

We therefore set out to express the rhEPO in *C. reinhardtii* glycoengineered  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  and  $IM_{XTAx}IM_{XTB}$  mutants in which xylosylation and fucosylation capacities have been impacted. Results obtained in these strains have been compared to those obtained in the cc5325 wild-type strain used as a reference to evaluate the efficiency of knock-out strategies developed for the deletion of *N*-glycan immunogenic epitopes in *C. reinhardtii*. After purification, biochemical analyses including a glycoproteomic study, showed that rhEPO harbours mature *N*-glycans, in which  $\beta(1,2)$ -xylose and  $\alpha(1,3)$ -fucose residues can be found depending on the strain, on the 3 *N*-

glycosites. Such *N*-glycans are consistent with those found on endogenous glycoproteins in the different strains used in this work.

Besides, these analyses have brought new insight by highlighting the presence of *O*-methyl fucose residues linked to the proximal GlcNAc of the chitobiose core of *N*-glycans in the two strains with a conserved fucosylation capacity.

The mature glycans that are present in the different variants of rhEPO reflect the *N*-glycosylation profile specific to each mutant, proving that this microalga is capable of glycosylating recombinant proteins for therapeutic purposes, even when the *N*-glycosylation pathway has been modified.

# Fine-tuning the *N*-glycosylation of recombinant human erythropoietin using *Chlamydomonas reinhardtii* mutants

S. Leprovost<sup>1,2,†</sup>, C. Plasson<sup>1,†</sup>, J. Balieu<sup>1</sup>, M-L. Walet-Balieu<sup>3</sup>, P. Lerouge<sup>1</sup>, M. Bardor<sup>1</sup> and E. Mathieu-Rivet<sup>1,\*</sup> 

<sup>1</sup>Université de Rouen Normandie, Normandie Univ, GlycoMEV UR 4358, SFR Normandie Végétal FED 4277, Innovation Chimie Carnot, IRIB, GDR CNRS Chemobiologie, Rouen, France

<sup>2</sup>Institute for Plant Biology and Biotechnology (IBBP), University of Münster, Münster, Germany

<sup>3</sup>Infrastructure de Recherche HeRaLeS, Plate-forme protéomique PISSARO, Université de Rouen Normandie, Rouen, France

Received 6 March 2024;

revised 6 June 2024;

accepted 18 June 2024.

\*Correspondence (Tel 00 (33)2 35 14 67 24;

email [elodie.rivet@univ-rouen.fr](mailto:elodie.rivet@univ-rouen.fr))

<sup>†</sup>Equal contribution of the co-authors.

## Summary

Microalgae are considered as attractive expression systems for the production of biologics. As photosynthetic unicellular organisms, they do not require costly and complex media for growing and are able to secrete proteins and perform protein glycosylation. Some biologics have been successfully produced in the green microalgae *Chlamydomonas reinhardtii*. However, post-translational modifications like glycosylation of these *Chlamydomonas*-made biologics have poorly been investigated so far. Therefore, in this study, we report on the first structural investigation of glycans linked to human erythropoietin (hEPO) expressed in a wild-type *C. reinhardtii* strain and mutants impaired in key Golgi glycosyltransferases. The glycoproteomic analysis of recombinant hEPO (rhEPO) expressed in the wild-type strain demonstrated that the three *N*-glycosylation sites are 100% glycosylated with mature *N*-glycans containing four to five mannose residues and carrying core xylose, core fucose and *O*-methyl groups. Moreover, expression in *C. reinhardtii* insertional mutants defective in *xylosyltransferases A and B* and *fucosyltransferase* resulted in drastic decreases of core xylosylation and core fucosylation of glycans *N*-linked to the rhEPOs, thus demonstrating that this strategy offers perspectives for humanizing the *N*-glycosylation of the *Chlamydomonas*-made biologics.

**Keywords:** *Chlamydomonas reinhardtii*, biologics, erythropoietin, glycosylation, glycoengineering.

## Introduction

The biopharmaceutical market size was about US \$343 billion in 2021 (Walsh and Walsh, 2022). In this market, sales of recombinant proteins like antibodies or hormones represent US \$271 billion with an increase of 44% between 2017 and 2021. Currently, recombinant proteins are mainly produced in Chinese HAMSTER OVARY (CHO) cells but this type of production is expensive and provide variable outcomes. As a consequence, there is a growing interest in the development of new, cheaper and effective expression systems. In this context, microalgae represent an attractive alternative. As photosynthetic unicellular organisms, they do not require costly and complex media for growing and most of them are classified by the USA Food and Drug Administration as 'Generally Recognized As Safe'. Moreover, as other eukaryotes, they are able to secrete proteins in the culture medium and perform protein *N*-glycosylation. A few studies have reported the expression of glycosylated therapeutic proteins in microalgae, highlighting the potential of these organisms as innovative cell biofactories. For example, a fully assembled recombinant IgG antibody directed against hepatitis B surface antigen has been produced in the diatom *Phaeodactylum tricoratum* (Hempel *et al.*, 2011) and has been demonstrated to be of good quality, glycosylated and able to efficiently bind human Fcγ receptors (Vanier *et al.*, 2015, 2017).

With regards to green microalgae, about 10 proteins, including human erythropoietin (hEPO), growth factors, interleukin 2 and

the receptor binding domain of the viral protein Spike, have been produced in *Chlamydomonas reinhardtii* (Berndt *et al.*, 2021; Deghani *et al.*, 2020; Eichler-Stahlberg *et al.*, 2009; Jarquín-Cordero *et al.*, 2020; Kiefer *et al.*, 2022). However, post-translational modifications including glycosylation of these therapeutic proteins expressed in *C. reinhardtii* are poorly documented (Smyth *et al.*, 2021). This is a major concern as the *N*-glycan biosynthesis pathways differ between *C. reinhardtii* and mammals and give rise to structurally different final glycan structures. For example, some of the biologics produced in CHO cells can exhibit sialylated polyantennary *N*-glycans that may be required for their *in vivo* half-time and bioactivity. In contrast, mature glycans *N*-linked to *C. reinhardtii* endogenous proteins exhibit α(1,3)-fucose (core-Fuc) linked to the proximal *N*-acetylglucosamine (GlcNAc) of the core chitobiose of non-canonical Man<sub>4</sub>GlcNAc<sub>2</sub> (M4) and Man<sub>5</sub>GlcNAc<sub>2</sub> (M5), in which the arrangement of mannose residues differs from the canonical structures described in plants or mammals (Table S1). Moreover, two xylose (Xyl) residues are linked to M4 and M5, one of them being β(1,2) linked to the core mannose (core-Xyl) and the other one on the trimannosyl linear branch. It should be noted that another pentose, arabinose, was recently identified on protein *N*-glycans from another microalga species belonging to the Chlorophyta (Mócsai *et al.*, 2020), and as a consequence, we cannot definitively rule out that arabinose residues are not present in *Chlamydomonas* *N*-glycans. In addition, mannose residues are partially *O*-methylated (Mathieu-Rivet *et al.*, 2013;

Please cite this article as: S.Leprovost, C.Plasson, J.Balieu, M.L.Walet-Balieu, P.Lerouge, M.Bardor and E.Mathieu-Rivet (2024) Fine-tuning the *N*-glycosylation of recombinant human erythropoietin using *Chlamydomonas reinhardtii* mutants. *Plant Biotechnol. J.*, <https://doi.org/10.1111/pbi.14424>.

Vanier *et al.*, 2017). Core-Xyl and core-Fuc have been shown to potentially induce immune reactions in mammals (Bardor *et al.*, 2003). Thus, the production of biologics in *C. reinhardtii* raises the question of the engineering of the *N*-glycosylation of Chlamydomonas-made biologics in order to remove putative immunogenic glycoepitopes and to make them suitable for future human therapy. Recently, the genes encoding some of the transferases responsible for the transfer of core-Xyl and core-Fuc epitopes, *xylosyltransferases A and B* (Cre09.g391282, *XTA* and Cre16.g678997, *XTB*) and *fucosyltransferase* (Cre18.g749697, *FUT*), have been identified. Moreover, the investigation of the insertional double-mutant  $IM_{XTA} \times IM_{XTB}$  and triple-mutant  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  demonstrated that the inactivation of these genes resulted in drastic decreases in core xylosylation and core fucosylation of glycans *N*-linked to their endogenous proteins and as a consequence, reduces the putative immunogenicity of *C. reinhardtii* proteins (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020; Schulze *et al.*, 2018).

Herein, we report on the structural investigation of a therapeutic human glycoprotein hormone, hEPO, recombinantly expressed in a wild-type *C. reinhardtii* strain and in  $IM_{XTA} \times IM_{XTB}$  and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  mutants for which the *N*-glycosylation pathway has been impaired. hEPO was selected as a model biopharmaceutical molecule because it exhibits three *N*-glycosylation sites on N24, N38 and N83, together with one *O*-glycosylation site on S126. Moreover, glycans account for 40% of the protein molecular weight and are essential for its stability and *in vivo* biological activity (Wasley *et al.*, 1991). The main objective of this study is to investigate whether knock-out strategies developed for the deletion of *N*-glycan immunogenic epitopes in *C. reinhardtii* are suitable for the glycoengineering of a therapeutic protein since this aspect is crucial for further exploitation of microalgae for the production of human-compatible biologics.

## Results

### rhEPO expressed in various genetic backgrounds is *N*-glycosylated

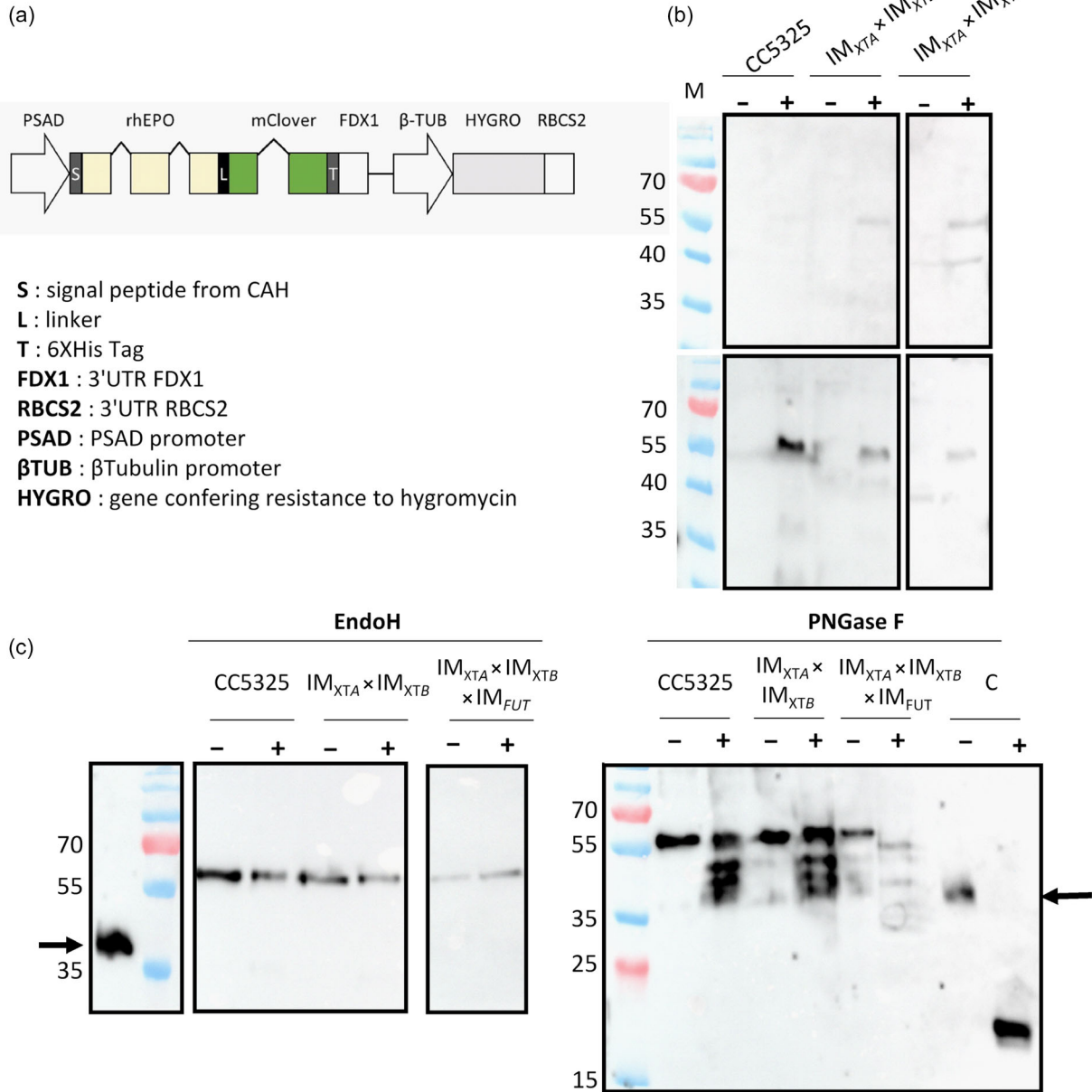
The sequence encoding hEPO was synthesized in fusion with the fluorescent protein mClover at the 3'-end of *hEPO* to help the selection of the best transformants exhibiting highest expression levels of the fusion protein rhEPO::mClover (Berndt *et al.*, 2021; Einhaus *et al.*, 2021, 2022; Freudenberg *et al.*, 2021, 2022; Lauersen *et al.*, 2016, 2018; Perozeni *et al.*, 2020; Wichmann *et al.*, 2018). The sequence coding for the signal peptide of the carbonic anhydrase CAH1 was added upstream *rhEPO::mClover* sequence to allow the secretion of the protein in the culture medium (Molino *et al.*, 2018). Finally, two copies of *RBCS2* intron 1 as well as one copy of *RBCS2* intron 2 were added inside the *rhEPO::mClover* sequence, as this strategy led to significant improvements of the expression levels of nuclear transgenes (Baier *et al.*, 2018, 2020). The expression of this optimized *hEPO* sequence was then driven *via* the combination of the PSAD promoter with FDX1 terminator (Figure 1a). CC5325 wild-type (WT) *C. reinhardtii* strain and mutants impaired for *XTA* and *XTB* activities (double-mutant  $IM_{XTA} \times IM_{XTB}$ ) and for both xylosyltransferases and *FUT* activities (triple-mutant  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$ ) were transformed with this construct. Positive clones expressing the highest levels of rhEPO were then selected according to the fluorescence level of mClover for further analysis by Western blot (Figure 1b). In all transformants, a signal at the

molecular weight of ~65 kDa was immunodetected using an anti-hEPO antibody in proteins isolated from the culture medium, whereas only a very weak band was immunodetected in the total protein extracts isolated from the cell pellets (Figure 1b). This suggests that the major part of expressed rhEPOs is secreted in the culture medium. The amount of secreted rhEPO in the culture medium was then quantified by an ELISA assay. rhEPO was found to be expressed at levels ranging from 0.02 in the triple mutant to 4.46 ng/mL in the reference strain. Despite the efforts made for optimizing the construction allowing the expression of rhEPO, these expression yields remained low compared to data published by Eichler-Stahlberg *et al.* (2009). This is likely due to the genetic background of the strain used for this study in which epigenetic mechanisms probably limit the expression of nuclear transgenes (Neupert *et al.*, 2020).

Recombinant human EPO fused to mClover was purified from the culture media by affinity chromatography. The electrophoretic mobility of purified rhEPOs was then investigated after treatment with deglycosylating enzymes to get insights into their *N*-glycosylation profiles. It is to note that the mClover does not contain any *N*-glycosylation site. The analysis was performed on rhEPOs expressed in WT as well as in the double-mutant  $IM_{XTA} \times IM_{XTB}$  and the triple-mutant  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  (Figure 1c). Whereas Endo H is able to only release oligomannosidic *N*-glycans ranging from  $Man_5GlcNAc_2$  to  $Man_9GlcNAc_2$ , PNGase F is able to cleave all glycans *N*-linked to Asn of glycosylation sites, except those having a core-Fuc (Tretter *et al.*, 1991). As depicted in Figure 1c, the electrophoretic mobilities of rhEPOs were not affected after deglycosylation with Endo H, suggesting that rhEPOs do not carry oligomannosidic *N*-glycans. In contrast, after PNGase F treatment, additional signals at lower molecular weights than rhEPOs were immunodetected for rhEPOs expressed in WT strain,  $IM_{XTA} \times IM_{XTB}$  and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  mutants (Figure 1c). These results suggested that rhEPOs expressed in *C. reinhardtii* are *N*-glycosylated with mature *N*-glycans rather than oligomannoside structures. Moreover, the occurrence of additional electrophoretic bands after PNGase F treatment demonstrated that the three glycosylation sites are occupied in rhEPOs with either core-Fuc containing *N*-glycans (resistant to PNGase F) and mature *N*-glycans devoid of this specific glycoepitope (sensitive to PNGase F), each electrophoretic bands of lower mobility corresponding to rhEPOs lacking core-Fuc on one to three *N*-glycosylation sites of rhEPO.

### Expression of rhEPOs in glycosylation mutants results in the secretion of glyco-engineered rhEPO lacking immunogenic core-Xyl and core-Fuc epitopes

The peptide sequence and the *N*-glycosylation profile of rhEPOs expressed either in WT or in the double  $IM_{XTA} \times IM_{XTB}$  and triple  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  mutants were then determined through a glycoproteomic approach (Balieu *et al.*, 2022). The purified proteins were separated by SDS-PAGE and digested by trypsin and Glu-C. The resulting peptides and glycopeptides were analysed by nano-liquid chromatography coupled to electrospray mass spectrometry (LC-ESI MS/MS). The overall protein sequence coverages were determined to be about 60% (Figure S1). Peptides containing N24, N38 or N83 were not detected in any sample, thus suggesting that they are fully *N*-glycosylated. In contrast, the peptide E117-R131 devoid of any *O*-glycan modification was detected in the LC-ESI MS/MS analyses of all samples, suggesting that the *O*-glycosylation site at S126 is not

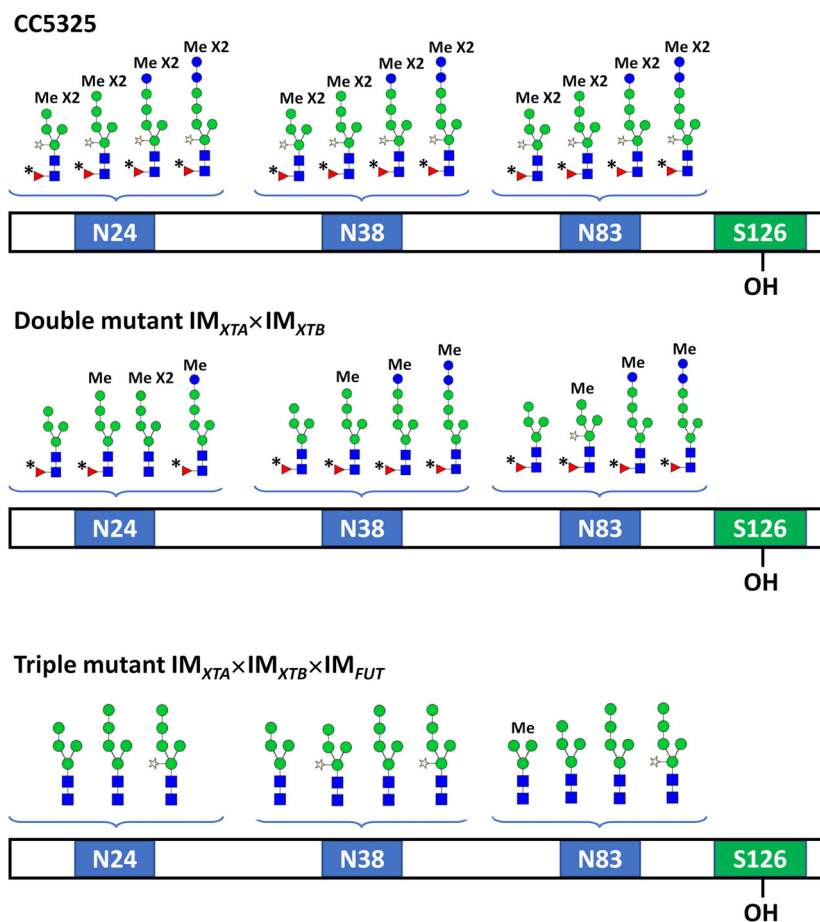


**Figure 1** rhEPO is secreted in the culture medium in *C. reinhardtii* transformants and is N-glycosylated. (a) Construct used for the expression rhEPO in *C. reinhardtii*. (b) Western blot analyses using an anti-hEPO antibody performed on total proteins from the cell pellets (upper panel) and TCA precipitated proteins from the culture medium (bottom panel) of CC5325,  $IM_{XTA} \times IM_{XTB}$  and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  mutants. (–) Proteins from non-transformed strains used as negative controls. (c) rhEPOs were purified from the culture media, submitted to a deglycosylation with Endo H or PNGase F and then analysed by Western blot using an anti-hEPO antibody. (–) and (+): Samples non-treated or treated by Endo H or PNGase F. C: positive control, corresponding to commercial rhEPO. The black arrows point out on intact commercial rhEPO.

occupied in rhEPOs produced in *C. reinhardtii* (Figure S1). In order to determine the N-glycan structures and their distribution on the three N-glycosylation sites, the mixtures of peptides and glycopeptides released by the endoprotease digestions were submitted to a targeted LC-ESI MS/MS analysis (Balieu *et al.*, 2022). To this end, peptides giving MS/MS spectra exhibiting N-glycan diagnostic fragment ions at m/z 204 (GlcNAc) and 366 (Man-GlcNAc) were selected as being glycopeptides. Whatever the strain in which rhEPO was expressed, the three

glycosylation sites of rhEPOs were found to be occupied in rhEPOs by a mixture of mature glycans with relative abundances depending on the protein N-glycosylation site (Figure 2 and Table S2).

Sequences of N-linked glycans of rhEPO glycopeptides from WT were determined on the basis of the MS/MS fragmentation patterns of glycopeptides as illustrated for  $[M + 3H]^{3+}$  ions assigned to HCSSLNENITVPDTK (Figure 3a). Major ions were assigned to M4- and M5-based structures depicted in Table S1.



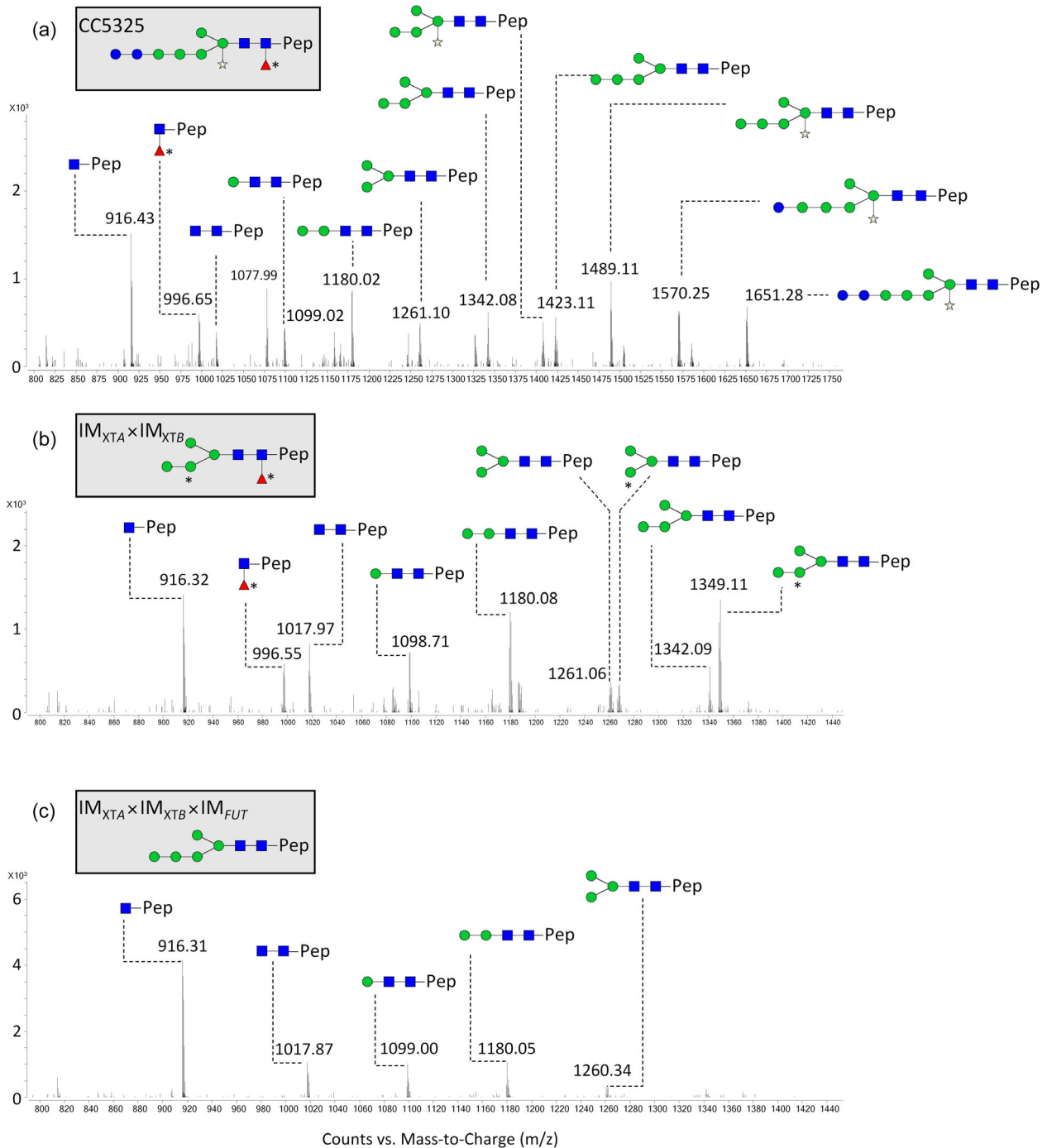
**Figure 2** Distribution of *N*-glycans on the three *N*-glycosylation sites of rhEPOs produced in *C. reinhardtii* CC5325,  $IM_{XTA} \times IM_{XTB}$  and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  mutants on rhEPOs. For structures containing one xylose residue, this xylose residue is represented as  $\beta(1,2)$  linked to the core mannose but it may also be linked to the trimannosyl linear branch (Lucas *et al.*, 2020). Other *N*-glycans were identified either lacking Fuc on the proximal GlcNAc or with additional *O*-methylations. Blue square: GlcNAc; green circle: Man; blue circle: Glc; yellow star: xylose, red triangle:  $\alpha(1,3)$ -fucose, red triangle with a star: *O*-methyl-fucose, Me: methyl group on Man.

Based on *m/z* values of molecular ions and fragmentation patterns, these structures were found to be M4 and M5 substituted by Xyl and Fuc residues, and *O*-methyl groups in accordance with previous reports on the *N*-glycosylation of endogenous proteins in *C. reinhardtii* (Lucas *et al.*, 2020; Mathieu-Rivet *et al.*, 2013; Oltmanns *et al.*, 2020; Schulze *et al.*, 2018) (Figure 2 and Table S2). Although we cannot definitively conclude on the location of the Xyl residues on *N*-glycans, we can assume that at least one of them is attached to the core mannose via a  $\beta(1,2)$  linkage as the purified rhEPO is recognized by the anti- $\beta(1,2)$ -Xyl antibodies that specifically recognize the core Xyl (Figure S2; Faye *et al.*, 1993). Surprisingly, MS/MS fragmentation of rhEPO glycopeptides from WT also revealed the presence of one *O*-methyl fucose residue linked to the proximal GlcNAc. Fragment ions assigned to the peptide attached to a GlcNAc are usually observed in the fragmentation patterns of *N*-glycopeptides together with ions differing by 146 Da for *N*-glycans having a Fuc residue on the proximal GlcNAc. However, we observed ions differing by 160 Da from peptide carrying a single GlcNAc residue which suggests a substitution of the proximal GlcNAc by an *O*-methylated Fuc instead of a Fuc (Figure 3a). Core fucosylation was previously reported in *C. reinhardtii* (Mathieu-Rivet *et al.*, 2013; Oltmanns

*et al.*, 2020) but not *O*-methyl fucosylation. This suggests that *O*-methylation of proteins *N*-glycans in *C. reinhardtii* is not restricted to mannose residues but also occurs on the core-Fuc.

Compared to structures identified on rhEPO expressed in WT, rhEPO from the double  $IM_{XTA} \times IM_{XTB}$  mutant impaired for XTA and XTB activities displayed structures containing also four to five mannose residues as well as *O*-methyl-Fuc, but lacking most Xyl substitutions (Figure 3b). Likewise, mainly M4 and M5 oligomannosides were found on the three *N*-glycosylation sites of rhEPO expressed in the  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  triple mutant (Figure 2 and Table S2). A few xylosylated structures were detected in this triple mutant which is consistent with the results of the Western blot analysis using an anti- $\beta(1,2)$ -Xyl antibody that has been performed on purified rhEPO (Figure S2). It is worth noting that fragment assigned to an *O*-methyl Fuc residue linked to the proximal GlcNAc was not observed in MS/MS spectra of glycopeptides from the  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  triple mutant that is defective in core *N*-glycan fucosylation (Figure 3c).

Glycoproteomic analyses of rhEPO indicated some oligosaccharides attached to the three glycosylation sites are *N*-glycans containing 6 or 7 hexoses (Figure 2). These are M5 *N*-glycans with two additional terminal Glc residues that are transiently transferred on M5 in the ER to ensure the quality control of the



**Figure 3** MS/MS spectra of  $[M + 3H/3]^{3+}$  ion at  $m/z$  1154.144 (a), 952.750 (b) and 948.738 (c) that were assigned to rEPO peptide HCSLNENITVPDTK (Pep) *N*-linked to G2M5XF\* in CC5325, M4F\*\* in  $IM_{XTA} \times IM_{XTB}$  and M5 in  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$ . Main doubly charged ions were assigned to the glycopeptide fragments. Blue square: GlcNAc; green circle: Man; blue circle: Glc, yellow star:  $\beta(1,2)$ -xylose and red triangle:  $\alpha(1,3)$ -fucose. \*O-methylation.

glycoproteins (GM5 and G2M5; Table S1) (Lucas *et al.*, 2018). These minor mono- and diglycosylated *N*-glycans, together with their xylosylated and *O*-methylated derivatives, have been previously reported in *C. reinhardtii* *N*-glycan profiles (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020; Schulze *et al.*, 2018). However, to check whether these glucosylations do not result from the expression of rEPO, a glycan analysis of endogenous proteins from WT and mutant lines was carried out. As shown in

Figure S3, GM5 and G2M5, as well as GM5X and G2M5X, were observed in MALDI TOF MS glycan profiles obtained from total proteins from the different preparations.

## Discussion

The development of alternative systems for the expression of biologics combining reduced production cost and reduced



safety risks like adventitious contamination from virus represents a major challenge in the near future (Brady and Love, 2021). During the last 20 years, *C. reinhardtii* has been extensively studied as a microalga model for the production of recombinant proteins. However, despite recent progress allowing increasing the expression level of secreted recombinant proteins (Schroda and Remacle, 2022), the glycosylation of a glycoprotein recombinantly expressed in this model has never been reported so far.

In order to investigate whether this green microalga is a suitable expression system for the production of therapeutic glycoproteins, we report in this study on the first structural investigation of glycans linked to a biopharmaceutical glycoprotein hormone expressed in *C. reinhardtii*. hEPO was chosen because it is a small glycoprotein exhibiting both *N*- and *O*-glycosylation sites and it was already produced successfully in *C. reinhardtii* (Eichler-Stahlberg et al., 2009). hEPO was expressed in *C. reinhardtii*, a wild-type strain and mutants impaired for *N*-glycosylation. ESI-MS/MS analyses of rhEPO demonstrated that the three *N*-glycosylation sites are 100% glycosylated. Glycans *N*-linked to rhEPO expressed in WT strain correspond to mature oligosaccharides containing four or five mannose residues carrying core-Xyl and core-Fuc and *O*-methyl groups (Figure 2). Previous glycomic analyses performed in *C. reinhardtii* showed that the *N*-glycan population released from endogenous proteins was composed of a major part of non-canonical M5 arising from the *N*-glycan processing occurring in the endoplasmic reticulum (ER) and M3 to M5 substituted with core-Xyl, core-Fuc and *O*-methyl groups arising from maturation in the Golgi apparatus (Lucas et al., 2018; Mathieu-Rivet et al., 2013; Oltmanns et al., 2020). Thus, the fact that rhEPO harbours mainly mature structures rather than oligomannosides suggests that *N*-glycans are exposed to the surface of the glycoprotein in a such way that they are fully accessible to the Golgi enzymes. This feature has been also reported for hEPO or recombinant hEPO expressed in CHO cells (Yang et al., 2017). This makes hEPO a suitable glycoprotein model for studying glycoengineering strategies based on the inactivation of Golgi-resident glycosidases and glycosyltransferases. Then, the *N*-glycan structures of rhEPO produced in different glycosylation mutants of *C. reinhardtii* have been analysed. As observed on endogenous proteins of *C. reinhardtii*-mutant strains (Lucas et al., 2020; Oltmanns et al., 2020; Schulze et al., 2018), the inactivation of *XTA* and *XTB*, and *FUT* resulted in a drastic decrease in immunogenic core xylosylation and core fucosylation of glycans *N*-linked on rhEPO expressed in these mutants. However, core xylosylation was not fully suppressed (Figure S2) probably because of the presence of additional xylosyltransferases that are predicted in the *C. reinhardtii* genomes, likely contributing to the protein core xylosylation through a compensation phenomenon (Lucas et al., 2020; Oltmanns et al., 2020). Moreover, the inactivation of protein core xylosylation and core fucosylation in mutants gives rise to mainly short oligomannosides that are more suitable for further glycoengineering to human-like *N*-glycans by expression of human glycosyltransferases to build the missing sialylated lactosamine sequences that are required for *in vivo* half-life and activity of rhEPO (Takeuchi et al., 1989; Yuen et al., 2003).

It is worth noting that our analysis allowed the identification of an *O*-methyl Fuc on the proximal GlcNAc residue of *N*-glycans, whereas so far, only core-Fuc was detected in glycoproteomic analysis performed on endogenous proteins (Mathieu-Rivet et al., 2013; Oltmanns et al., 2020). This residue results probably

from the high accessibility of rhEPO *N*-glycans to Golgi *O*-methyltransferases. Unfortunately, although the construction allowing the expression of rhEPO was designed to optimize the yield of secreted rhEPO, this remained too low to perform the analysis of PNGase-released *N*-glycans from purified rhEPO by MALDI-TOF mass spectrometry. However, these *O*-methyl Fuc-containing structures were not observed in *N*-glycan profiles of total endogenous proteins from *C. reinhardtii*, suggesting that they are resistant to the deglycosylation by PNGase A. We presume that the *O*-methylation of Fuc residue linked to the proximal GlcNAc residue prevents the *N*-glycans cleavage by PNGase A. *O*-methyl Fuc has already been previously reported on *N*-glycans of proteins from *C. elegans* (Altmann et al., 2001), or more recently on terminal lactosamine epitopes of *N*-glycans of proteins from the oysters *Crassostrea gigas* and *Ostrea edulis* (Auger et al., 2023).

*N*-linked glycans containing terminal sialic acids were not found in the glycan profile of rhEPOs expressed in *C. reinhardtii*. This is consistent with previous reports (Lucas et al., 2020; Mathieu-Rivet et al., 2013; Oltmanns et al., 2020; Schulze et al., 2018). This aspect could constitute a limit to the use of such microalga model for the expression of therapeutic proteins like hEPO since its most active form harbours sialylated tetraantennary *N*-glycans. Moreover, it is well established that the presence of sialic acid residues protects EPO from its recognition by galactose-binding receptors located on erythrocytes, thus preventing its clearance and prolonging its half-life in serum (Elliott et al., 2014; Wasley et al., 1991). Therefore, knock-in strategies would be necessary in future studies to express heterologous glycosyltransferases and glycoenzymes that are required for the synthesis of sialylated *N*-glycans in *C. reinhardtii* as previously reported in moss and plants (Bohlender et al., 2020; Kallolimath et al., 2016).

hEPO also exhibits one *O*-glycosylation site on S126. hEPO *O*-glycan is a mucine-type oligosaccharide with a GalNAc *O*-linked to the serine residue (Sasaki et al., 1987). In our study, no *O*-glycan was observed on rhEPOs expressed in *C. reinhardtii*. Only few information is available on protein *O*-glycosylation in *C. reinhardtii*, except for hydroxyproline-rich glycoproteins (Mathieu-Rivet et al., 2020). To our knowledge, GalNAc has not been reported in *C. reinhardtii* and, as a consequence, gaining human glycans *O*-linked to a recombinant therapeutic protein produced in this green microalga would require both the implementation of the enzymatic machinery to supply the Golgi apparatus with the nucleotide-sugar UDP-GalNAc, and also the expression of appropriate glycosyltransferases such as a polypeptide: *N*-acetylgalactosaminyltransferase, as well as a core  $\beta(1,3)$ -galactosyltransferase, that are required for the formation of the most common *O*-glycan extension (Castilho et al., 2012).

## Materials and methods

### Strains and growth media

The insertional mutants  $IM_{XTA} \times IM_{XTB}$  (Lucas et al., 2020) and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  (Oltmanns et al., 2020; Schulze et al., 2018) were used for these experiments. The double-mutant  $IM_{XTA \times XTB}$  has been obtained by crossing the insertional strains LMJ.RY0402.087519 (Cre09.g391282, XP\_042921158.1) and LMJ.RY0402.118417 (Cre16.g678997, XP\_042916231.1) provided by the ClIP library (Li et al., 2016; Zhang et al., 2014). The triple-mutant  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  has been obtained by crossing an insertional mutant  $IM_{XTA}$  from the library described in Cheng et al., 2017, with the double-mutant  $IM_{XTB} \times IM_{FUT}$

that was generated from strains LMJ.RY0402.049160 (Cre18.g749697, XP\_001695259.2) and LMJ.RY0402.118417 (Cre16.g678997, XP\_042916231.1) from the CLIP library. The strain CC5325 was chosen as a wild-type strain (WT) since it is the genetic background that was used to generate the mutants in the CLIP library (Li *et al.*, 2016; Zhang *et al.*, 2014). All the strains were grown on Tris-acetate phosphate (TAP) medium (Gorman and Levine, 1965) at a constant illumination of 400  $\mu\text{mol photons/m}^2/\text{s}$ , at 25 °C.

### Constructs and transformation

The 5'-part of the sequence encoding hEPO was fused to the 3' part of the sequence encoding for mClover, and a sequence coding for the SP from the carbonic anhydrase (SP-CA, Molino *et al.*, 2018) was added downstream to this construct, leading the expression of the fusion protein SP-CA::rhEPO::mClover. This sequence was then optimized using Intronsorter (Jaeger *et al.*, 2019) as follows: two copies of the first RBCS2 intron as well as one copy of the second RBCS2 intron were added, and the sequence was synthesized according to the nuclear codon bias in *C. reinhardtii*. The expression of this optimized sequence was then driven *via* the combination of the PSAD promoter with FDX1 terminator (López-Paz *et al.*, 2017). A second transcriptional unit was added to these constructions allowing the resistance to hygromycin.

After *in silico* design, all plasmids were synthesized *de novo* by polyplus transfection SA (<https://www.polyplus-sartorius.com>, Loos, France). Plasmids were delivered as lyophilized DNA, transformed and maintained in *Escherichia coli* DH5 $\alpha$  grown in lysogeny broth medium (LB)-containing ampicillin. All plasmid vectors were prepared by midi prep using the NucleoBond Xtra Midi kit from Macherey-Nagel (Düren, Germany) and then linearized by digest with EcoRV restriction enzyme at 37 °C. For each transformation, a starter culture of *C. reinhardtii* was grown in TAP media to near saturation ( $\sim 5 \times 10^6$  cells/mL) and then diluted back into  $0.5 \times 10^6$  cells/mL in 500 mL of TAP media 24 h before transformation. The next day, cells were collected by centrifugation at 1000 *g* for 5 min at 16 °C. The cells were then resuspended in 4 mL of TAP containing 40 mM sucrose. Depending on the cell pellet size, the final volume is about 5 mL. For one transformation, 250  $\mu\text{L}$  is transferred in a 4 mm gap distance electroporation cuvette (Fisherbrand; FisherScientific), with 1–2  $\mu\text{g}$  of linearized vector. The mixture is then incubated for 5 min at 16 °C before electroporation using a Gene Pulser Xcell Electroporation System (Bio-Rad Laboratories, Hercules, CA, USA) set to 800 V and 25  $\mu\text{F}$ . Immediately after electroporation, the mixture is resuspended in 8 mL TAP 40 mM sucrose. After overnight recovery in 40 mM sucrose in TAP media with low light ( $\sim 5 \mu\text{mol photons/m}^2/\text{s}$ ), cells were pelleted, resuspended in 5 mL of fresh TAP media and plated on solid TAP containing 15  $\mu\text{g/mL}$  zeocin or 15  $\mu\text{g/mL}$  hygromycin B. The positive clones expressing rhEPO::mClover were screened thanks to the fluorescence level of mClover. Individual colonies grown on TAP medium supplemented with selection antibiotic were picked and resuspended in 160  $\mu\text{L}$  of TAP without antibiotic into a 96-well plate. After 5 days, the mClover fluorescence was analysed using the FlexStation<sup>®</sup> 3 Multi-Mode Microplate Reader (Molecular Devices LLC, San Jose, CA, USA) according to the following parameters: excitation: 480 nm; emission: 530 nm. mClover fluorescence was normalized to  $\text{DO}_{720}$  to take into account differences in cell density.

### Protein extraction

Total proteins were extracted from a 5 mL culture grown until near saturation ( $\text{DO}_{720} \approx 1$ ). After centrifugation and removing of the culture medium, the cell pellet was broken in 1 mL of Agrisera buffer (0.138 M Tris-HCl pH 8.5, 0.5 M LDS, 10% (v/v) glycerol, 0.5 mM EDTA) in a microtube containing lysis matrix D (MP Biomedicals<sup>™</sup>) with a Fastprep system (MP Biomedicals<sup>™</sup>) according to the following parameters: 4 cycles (40 s; 6.5 m/s). After centrifugation at 4000 *g* 10 min, 20  $\mu\text{L}$  of supernatant was used for Western blot analysis. Proteins of the culture medium were precipitated with 10% TCA (v/v) at 4 °C overnight. After 3 washes with cold acetone 80% (v/v), the pellet is resuspended in 60  $\mu\text{L}$  Agrisera buffer and 30  $\mu\text{L}$  was used for Western blot analyses.

### Purification of rhEPO-mClover

rhEPO-mClover was purified using the GFP-Trap<sup>™</sup> Magnetic Agarose Kit (ChromoTek<sup>®</sup>), according to the manufacturer's instructions. Briefly, the medium was collected from a 200 mL culture grown until near saturation ( $\text{DO}_{720} \approx 1$ ) and then lyophilized before being re-suspended in 20 mL of phosphate buffer saline pH 6.7 containing a cocktail of protease inhibitor (Roche). After a centrifugation at 48 000 *g* for 2 h, the supernatant was loaded on an Amicon 50 KDa centrifugal filter unit 5000 *g* for 2 h, to get a final volume of 500  $\mu\text{L}$ . The filter was rinsed with 500  $\mu\text{L}$  of 2 $\times$  dilution buffer provided in the kit, in such a way that secreted proteins were concentrated in final volume of 1 mL. Twenty-five microlitres of beads were then incubated in this volume overnight at 4 °C on a carousel rotator. After three washing using the provided buffer, 80  $\mu\text{L}$  of LSB2X was added and the mixture was warmed at 95 °C for 5 min for further glycoproteomic analysis. For further deglycosylation treatment, rhEPO is eluted using 140  $\mu\text{L}$  of acidic elution buffer provided in the kit and then neutralized with 20  $\mu\text{L}$  of neutralization buffer.

### PNGase F and Endo H treatment

For PNGase F deglycosylation, 110  $\mu\text{L}$  of purified rhEPO was first denaturated in Tris-HCl 0.1 M pH8, 0.1% SDS (w/v) during 5 min at 100 °C. Then, deglycosylation is performed in Tris-HCl 0.1 M pH8, NP40 0.5% (v/v) using 1 U of PNGase F (Roche) during 24 h at 37°C. Deglycosylation with 2000 U EndoH (Promega) was carried out on 50  $\mu\text{L}$  of purified rhEPO according to the manufacturer's instructions.

### Western blot analysis

Immunoblotting analyses were performed on total proteins from cells, the culture medium or on purified rhEPO. Proteins were separated on 12% SDS polyacrylamide gels and transferred to nitrocellulose membranes (0.2  $\mu\text{m}$ ; Amersham<sup>™</sup> Protran<sup>®</sup>) using the Pierce<sup>™</sup> Power Blotter system (Thermo Scientific<sup>™</sup>) and semi-dry blotting method according to the manufacturer's instructions. rhEPO-mClover was detected using either anti-eGFP (1/6000e; Invitrogen AB\_962098) or anti-hEPO (1/2000e; Abcam ab226956) as a primary antibody, and an HRP-conjugated secondary antibody (Invitrogen A16041 for anti-eGFP, and 65–6120 for anti-hEPO). For the analysis of the presence of xylose or fucose residue, rhEPO-mClover was detected using either anti-Xyl (1/3000e; Agrisera AS07 267) or anti-Fuc (1/2000e; Agrisera AS07 268) as primary antibodies and a HRP-conjugated secondary antibody (Invitrogen G5 G120).

## ELISA assay

The amount of rhEPO was quantified using the Human EPO ELISA Kit (Invitrogen BMS2035-2) according to the manufacturer's instructions. For each clone expressing rhEPO, three independent cultures were grown for 5 days. After centrifugation (4500 *g*, 10 min) to pellet the cells, ELISA assays were performed on 50  $\mu$ L of the culture medium. The reported value for each clone corresponds to the average of three technical repeats.

## Glycoproteomic analysis

The protein sequence and *N*-glycan profile of purified rhEPOs were determined through the analysis by LC-ESI MS/MS of peptides and glycopeptides released by trypsin and Glu-C digestions as previously described in Balieu *et al.* (2022). Peptide analyses by nano-LC-ESI-MS/MS were performed using a liquid chromatography system coupled to a Q-TOF 6545 XT AdvanceBio mass spectrometer (Agilent Technologies, Les Ulis, France). The liquid chromatography system consisted of a nano-LC 1200 system and an HPLC-chip cube interface (Agilent Technologies). The samples were resuspended in an acetonitrile/formic acid mixture of 0.1% 3/97 v/v. The samples were enriched and desalted using a 40 nL RP-C18 column, then the compounds were separated using an Agilent Zorbax C18 column (length 150 mm, internal diameter 75  $\mu$ m, particle size 5  $\mu$ m and pore size 30 nm). A linear gradient of 3% to 80% acetonitrile in 0.1% formic acid was carried out over 26 min at a flow rate of 350 nL/min. The mass range was 290–2000 *m/z* in MS1 and 59–3200 *m/z* in MS2. During each cycle, a maximum of five precursors sorted by charge state (discharged priority, mono-charged excluded) are isolated and fragmented in the collision cell with N2. The collision energy is automatically adjusted according to the *m/z* value. Dynamic exclusion of these precursors was activated after 1 spectrum in 0.2 min and the absolute threshold was 1000 (relative threshold 0.001%). The data were recorded and processed using Agilent Data Acquisition B.09.00 and Mass Hunter Qualitative B.09.00 software. For the identification of putative *O*-linked glycans on S126, we searched in the LC-ESI MS/MS data for glycopeptide candidates exhibiting C- or N-terminal peptide fragments of the S126-containing tryptic peptide E117-R131. No glycopeptide candidate was detected by this analytical approach.

## *N*-glycan profiling of *C. reinhardtii* lines

For the determination of *N*-glycan profiles of *C. reinhardtii* CC5325, IM<sub>XTA</sub>IM<sub>XTB</sub> and IM<sub>XTA</sub>IM<sub>XTB</sub>IM<sub>FUT</sub> lines, *N*-glycans were released from proteins by PNGase A, coupled to 2-amidobenzamide (2-AB) and analysed by MALDI-TOF mass spectrometry as reported in Balieu *et al.*, 2022.

## Acknowledgements

This work was supported by the University of Rouen Normandy, by grants from the Graduate School of Research XL-Chem (ANR-18-EURE-0020 XL-Chem), by French government through the ANR agency under the ANR PRCE DAGENTA project (ANR-21-CE20-0038-001) and the program « Grand défi Biomédicament: améliorer les rendements et maîtriser les coûts de production: Nouveaux Systèmes d'Expression – 2020 » (PHAEOMABS project – ANR-21-F2II-0005), and the Région Normandie (projet émergent RIN Sweet Trip). We also thank Pr. Michael Hippler, University of Münster, for providing the triple-mutant

IM<sub>XTA</sub> × IM<sub>XTB</sub> × IM<sub>FUT</sub>, and Dr Gaëtan Vanier who initiated this work.

## Conflict of interest

All the authors declare no competing interests.

## Data availability statement

Data sharing is not applicable to this article. These data have never been shared before on any repository database.

## References

- Altmann, F., Fabini, G., Ahorn, H. and Wilson, I.B.H. (2001) Genetic model organisms in the study of N-glycans. *Biochimie* **83**, 703–712.
- Auger, A., Yu, S.-Y., Guu, S.-Y., Quémérer, A., Euller-Nicolas, G., Ando, H., Desdouts, M. *et al.* (2023) Species-specific N-glycomes and methylation patterns of oysters *Crassostrea gigas* and *Ostrea edulis* and their possible consequences for the norovirus–HBGA interaction. *Mar. Drugs* **21**, 342.
- Baier, T., Jacobebbinghaus, N., Einhaus, A., Lauersen, K.J. and Kruse, O. (2020) Introns mediate post-transcriptional enhancement of nuclear gene expression in the green microalga *Chlamydomonas reinhardtii*. *PLoS Genet.* **16**, e1008944.
- Baier, T., Wichmann, J., Kruse, O. and Lauersen, K.J. (2018) Intron-containing algal transgenes mediate efficient recombinant gene expression in the green microalga *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **46**, 6909–6919.
- Balieu, J., Jung, J.-W., Chan, P., Lomonosoff, G.P., Lerouge, P. and Bardor, M. (2022) Investigation of the N-glycosylation of the SARS-CoV-2 S protein contained in VLPs produced in *Nicotiana benthamiana*. *Molecules* **27**, 5119.
- Bardor, M., Faveeuw, C., Fitchette, A.-C., Gilbert, D., Galas, L., Trottein, F., Faye, L. *et al.* (2003) Immunoreactivity in mammals of two typical plant glycoepitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology* **13**, 427–434.
- Berndt, A.J., Smalley, T.N., Ren, B., Simkovsky, R., Badary, A., Sproles, A.E., Fields, F.J. *et al.* (2021) Recombinant production of a functional SARS-CoV-2 spike receptor binding domain in the green algae *Chlamydomonas reinhardtii*. *PLoS One* **16**, e0257089.
- Bohlender, L.L., Parsons, J., Hoernstein, S.N.W., Rempfer, C., Ruiz-Molina, N., Lorenz, T., Rodríguez Jahnke, F. *et al.* (2020) Stable protein sialylation in *Physcomitrella*. *Front. Plant Sci.* **18**, 610032.
- Brady, J.R. and Love, J.C. (2021) Alternative hosts as the missing link for equitable therapeutic protein production. *Nat. Biotechnol.* **39**, 404–407.
- Castilho, A., Neumann, L., Daskalova, S., Mason, H.S., Steinkellner, H., Altmann, F. and Strasser, R. (2012) Engineering of sialylated mucin-type O-glycosylation in plants. *J. Biol. Chem.* **287**, 36518–36526.
- Cheng, X., Liu, G., Ke, W., Zhao, L., Lv, B., Ma, X., Xu, N. *et al.* (2017) Building a multipurpose insertional mutant library for forward and reverse genetics in *Chlamydomonas*. *Plant Methods* **15**, 36.
- Dehghani, J., Adibkia, K., Movafeghi, A., Pourseif, M.M. and Omid, Y. (2020) Designing a new generation of expression toolkits for engineering of green microalgae; robust production of human interleukin-2. *Bioimpacts* **10**, 259–268.
- Eichler-Stahlberg, A., Weisheit, W., Ruecker, O. and Heitzer, M. (2009) Strategies to facilitate transgene expression in *Chlamydomonas reinhardtii*. *Planta* **229**, 873–883.
- Einhaus, A., Baier, T., Rosenstengel, M., Freudenberg, R.A. and Kruse, O. (2021) Rational Promoter Engineering Enables Robust Terpene Production in Microalgae. *ACS Synth. Biol.* **10**, 847–856.
- Einhaus, A., Steube, J., Freudenberg, R., Barczyk, J., Baier, T. and Kruse, O. (2022) Engineering a powerful green cell factory for robust photoautotrophic diterpenoid production. *Metab. Eng.* **82**, 90.
- Elliott, S., Sinclair, A., Collins, H., Rice, L. and Jelkmann, W. (2014) Progress in detecting cell-surface protein receptors: the erythropoietin receptor example. *Ann. Hematol.* **93**, 181–192.

- Faye, L., Gomord, V., Fitchette-Lainé, A.C. and Chrispeels, M.J. (1993) Affinity purification of antibodies specific for Asn-linked glycans containing alpha 1→3 fucose or beta 1→2 xylose. *Anal. Biochem.* **15**, 104–108.
- Freudenberg, R.A., Baier, T., Einhaus, A., Wobbe, L. and Kruse, O. (2021) High cell density cultivation enables efficient and sustainable recombinant polyamine production in the microalga *Chlamydomonas reinhardtii*. *Bioresour. Technol.* **323**, 124542.
- Freudenberg, R.A., Wittemeier, L., Einhaus, A., Baier, T. and Kruse, O. (2022) Advanced pathway engineering for phototrophic putrescine production. *Plant Biotechnol. J.* **20**, 1968–1982.
- Gorman, D.S. and Levine, R.P. (1965) Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. USA* **54**, 1665–1669.
- Hempel, F., Lau, J., Klingl, A. and 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.
- Jaeger, D., Baier, T. and Lauenstein, K.J. (2019) Intronserter, an advanced online tool for design of intron containing transgenes. *Algal Res.* **42**, 101588.
- Jarquín-Cordero, M., Chávez, M.N., Centeno-Cerdas, C., Bohne, A.-V., Hopfner, U., Machens, H.-G., Egaña, J.T. et al. (2020) Towards a biotechnological platform for the production of human pro-angiogenic growth factors in the green alga *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.* **104**, 725–739.
- Kallolimath, S., Castilho, A., Strasser, R., Grünwald-Gruber, C., Altmann, F., Strubl, S., Galuska, C.E. et al. (2016) Engineering of complex protein sialylation in plants. *Proc. Natl. Acad. Sci. USA* **23**, 9498–9503.
- Kiefer, A.M., Niemeyer, J., Probst, A., Erkel, G. and Schroda, M. (2022) Production and secretion of functional SARS-CoV-2 spike protein in *Chlamydomonas reinhardtii*. *Front. Plant Sci.* **13**, 988870.
- Lauenstein, K.J., Baier, T., Wichmann, J., Wördenweber, R., Mussgnug, J.H., Hübner, W., Huser, T. et al. (2016) Efficient phototrophic production of a high-value sesquiterpenoid from the eukaryotic microalga *Chlamydomonas reinhardtii*. *Metab. Eng.* **38**, 331–343.
- Lauenstein, K.J., Wichmann, J., Baier, T., Kampranis, S.C., Pateraki, I., Møller, B.L. and Kruse, O. (2018) Phototrophic production of heterologous diterpenoids and a hydroxy-functionalized derivative from *Chlamydomonas reinhardtii*. *Metab. Eng.* **49**, 116–127.
- Li, X., Zhang, R., Patena, W., Gang, S.S., Blum, S.R., Ivanova, N., Yue, R. et al. (2016) An indexed, mapped mutant library enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. *Plant Cell* **28**, 367–387.
- López-Paz, C., Liu, D., Geng, S. and Umen, J.G. (2017) Identification of *Chlamydomonas reinhardtii* endogenous genic flanking sequences for improved transgene expression. *Plant J.* **92**, 1232–1244.
- Lucas, P.-L., Dumontier, R., Loutelier-Bourhis, C., Mareck, A., Afonso, C., Lerouge, P., Mati-Baouche, N. et al. (2018) User-friendly extraction and multistage tandem mass spectrometry based analysis of lipid-linked oligosaccharides in microalgae. *Plant Methods* **14**, 107.
- Lucas, P.-L., Mathieu-Rivet, E., Song, P.C.T., Oltmanns, A., Loutelier-Bourhis, C., Plasson, C., Afonso, C. et al. (2020) Multiple xylosyltransferases heterogeneously xylosylate protein N-linked glycans in *Chlamydomonas reinhardtii*. *Plant J.* **102**, 230–245.
- Mathieu-Rivet, E., Mati-Baouche, N., Walet-Balieu, M.-L., Lerouge, P. and Bardor, M. (2020) N- and O-glycosylation pathways in the microalgae polyphyletic group. *Front. Plant Sci.* **11**, 609993.
- Mathieu-Rivet, E., Scholz, M., Arias, C., Dardelle, F., Schulze, S., Le Mauff, F., Teo, G. et al. (2013) Exploring the N-glycosylation pathway in *Chlamydomonas reinhardtii* unravels novel complex structures. *Mol. Cell. Proteomics* **12**, 3160–3183.
- Mócsai, R., Blaukopf, M., Svehla, E., Kosma, P. and Altmann, F. (2020) The N-glycans of *Chlorella sorokiniana* and a related strain contain arabinose but have strikingly different structures. *Glycobiology* **16**, 663–676.
- Molino, J.V.D., de Carvalho, J.C.M. and Mayfield, S.P. (2018) Comparison of secretory signal peptides for heterologous protein expression in microalgae: Expanding the secretion portfolio for *Chlamydomonas reinhardtii*. *PLoS One* **13**, e0192433.
- Neupert, J., Gallaher, S.D., Lu, Y., Strenkert, D., Segal, N., Barahimipour, R., Fitz-Gibbon, S.T. et al. (2020) An epigenetic gene silencing pathway selectively acting on transgenic DNA in the green alga *Chlamydomonas reinhardtii*. *Nat. Commun.* **8**, 6269.
- Oltmanns, A., Hoepfner, L., Scholz, M., Zinzus, K., Schulze, S. and Hippler, M. (2020) Novel insights into N-glycan fucosylation and core xylosylation in *Chlamydomonas reinhardtii*. *Front. Plant Sci.* **10**, 1686.
- Perozeni, F., Cazzaniga, S., Baier, T., Zanon, F., Zoccatelli, G., Lauenstein, K.J., Wobbe, L. et al. (2020) Turning a green alga red: engineering astaxanthin biosynthesis by intragenic pseudogene revival in *Chlamydomonas reinhardtii*. *Plant Biotechnol. J.* **18**, 2053–2067.
- Sasaki, H., Bothner, B., Dell, A. and Fukuda, M. (1987) Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA. *J. Biol. Chem.* **5**, 12059–12076.
- Schroda, M. and Remacle, C. (2022) Molecular advancements establishing *Chlamydomonas* as a host for biotechnological exploitation. *Front. Plant Sci.* **13**, 911483.
- Schulze, S., Oltmanns, A., Machnik, N., Liu, G., Xu, N., Jarmatz, N., Scholz, M. et al. (2018) N-glycoproteomic characterization of mannosidase and xylosyltransferase mutant strains of *Chlamydomonas reinhardtii*. *Plant Physiol.* **176**, 1952–1964.
- Smyth, D.J., Ren, B., White, M.P.J., McManus, C., Webster, H., Shek, V., Evans, C. et al. (2021) Oral delivery of a functional algal-expressed TGF- $\beta$  mimic halts colitis in a murine DSS model. *J. Biotechnol.* **10**, 1–12.
- Takeuchi, M., Inoue, N., Strickland, T.W., Kubota, M., Wada, M., Shimizu, R., Hoshi, S. et al. (1989) Relationship between sugar chain structure and biological activity of recombinant human erythropoietin produced in Chinese hamster ovary cells. *Proc. Natl. Acad. Sci. USA* **86**, 7819–7822.
- Tretter, V., Altmann, F. and März, L. (1991) Peptide-N4-(N-acetyl- $\beta$ -glucosaminyl)asparagine amidase F cannot release glycans with fucose attached  $\alpha$ 1 → 3 to the asparagine-linked N-acetylglucosamine residue. *Eur. J. Biochem.* **199**, 647–652.
- Vanier, G., Hempel, F., Chan, P., Rodamer, M., Vaudry, D., Maier, U.G., Lerouge, P. et al. (2015) Biochemical characterization of human anti-hepatitis B monoclonal antibody produced in the microalgae *Phaeodactylum tricornutum*. *PLoS One* **10**, e0139282.
- Vanier, G., Lucas, P.-L., Loutelier-Bourhis, C., Vanier, J., Plasson, C., Walet-Balieu, M.-L., Tchi-Song, P.C. et al. (2017) Heterologous expression of the N-acetylglucosaminyltransferase I dictates a reinvestigation of the N-glycosylation pathway in *Chlamydomonas reinhardtii*. *Sci. Rep.* **7**, 10156.
- Walsh, G. and Walsh, E. (2022) Biopharmaceutical benchmarks 2022. *Nat. Biotechnol.* **40**, 1722–1760.
- Wasley, L.C., Timony, G., Murtha, P., Stoudemire, J., Dorner, A.J., Caro, J., Krieger, M. et al. (1991) The importance of N- and O-linked oligosaccharides for the biosynthesis and in vitro and in vivo biologic activities of erythropoietin. *Blood* **77**, 2624–2632.
- Wichmann, J., Baier, T., Wentnagel, E., Lauenstein, K.J. and Kruse, O. (2018) Tailored carbon partitioning for phototrophic production of (E)- $\alpha$ -bisabolene from the green microalga *Chlamydomonas reinhardtii*. *Metab. Eng.* **45**, 211–222.
- Yang, Q., An, Y., Zhu, S., Zhang, R., Loke, C.M., Cipollo, J.F. and Wang, L.-X. (2017) Glycan remodeling of human erythropoietin (EPO) through combined mammalian cell engineering and chemoenzymatic transglycosylation. *ACS Chem. Biol.* **12**, 1665–1673.
- Yuen, C.-T., Storrington, P.L., Tiplady, R.J., Izquierdo, M., Wait, R., Gee, C.K., Gerson, P. et al. (2003) Relationships between the N-glycan structures and biological activities of recombinant human erythropoietins produced using different culture conditions and purification procedures. *Br. J. Haematol.* **121**, 511–526.
- Zhang, R., Patena, W., Armbruster, U., Gang, S.S., Blum, S.R. and Jonikas, M.C. (2014) High-throughput genotyping of green algal mutants reveals random distribution of mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. *Plant Cell* **26**, 1398–1409.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Figure S1** Sequence coverage determined by LC-ESI MS/MS analysis of the rhEPO clover expressed in the *C. reinhardtii* CC5325 strain. Black arrows point out the glycosylation sites.

**Figure S2** MALDI-TOF MS of 2AB-labelled *N*-glycans released by PNGase A from proteins of *C. reinhardtii* CC5325 (A),  $IM_{XTA} \times IM_{XTB}$  (B) and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  (C). See Table S1 for structures.

**Table S1** Structure, name and *m/z* values of  $[M + Na]^+$  ions of 2-AB-derivatives of *N*-glycans.

**Table S2** Distribution of *N*-glycans on the three glycosylation sites of rhEPOs produced in *C. reinhardtii* CC5325 strain,  $IM_{XTA} \times IM_{XTB}$  and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  mutants.

# Chapter 1

## Supplemental data

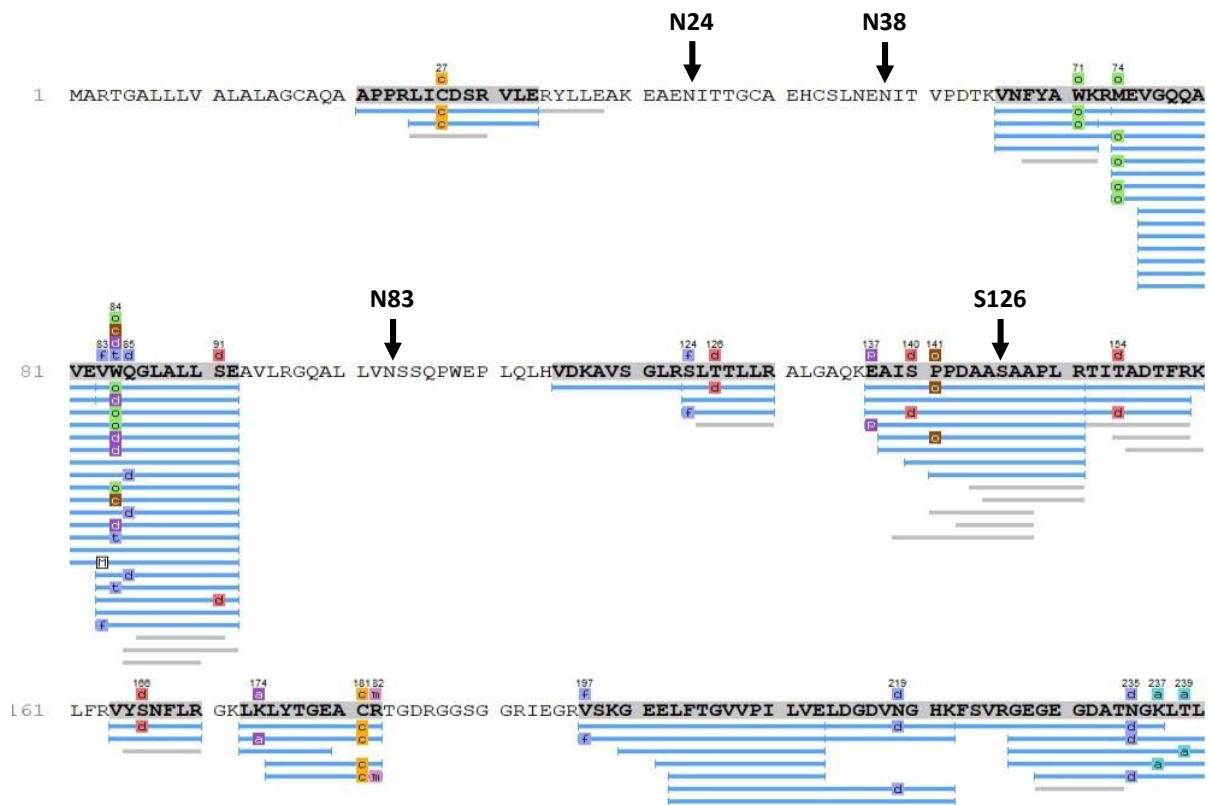
Structure	Name	[M+Na] <sup>+</sup>	Structure	Name	[M+Na] <sup>+</sup>
	M3	1 053.37		M3F	1199.42
	M3X	1 185.41		M3XF	1 331.46
	M4	1 215.41		M4F	1 361.47
	M4X	1 347.47		M4XF	1 493.52
	M5	1 377.53		M5F	1 523.59
	M5X	1 509.52		M5XF	1 655.59
	GM5	1 539.52		GM5F	1 685.58
	GM5X	1 671.57		GM5XF	1 817.59
	G2M5	1 701.57		G2M5F	1 817.63
	G2M5X	1 833.62		G2M5XF	1 979.68

**Table S1.** Structure, name and  $m/z$  values of  $[M+Na]^+$  ions of 2-AB-derivatives of *N*-glycans. Blue square: GlcNAc; green circle: Man; blue circle: Glc, yellow star:  $\beta(1,2)$ -xylose and red triangle:  $\alpha(1,3)$ -fucose. For *O*-methylated derivatives, add 14.01 Da *per* methyl group.

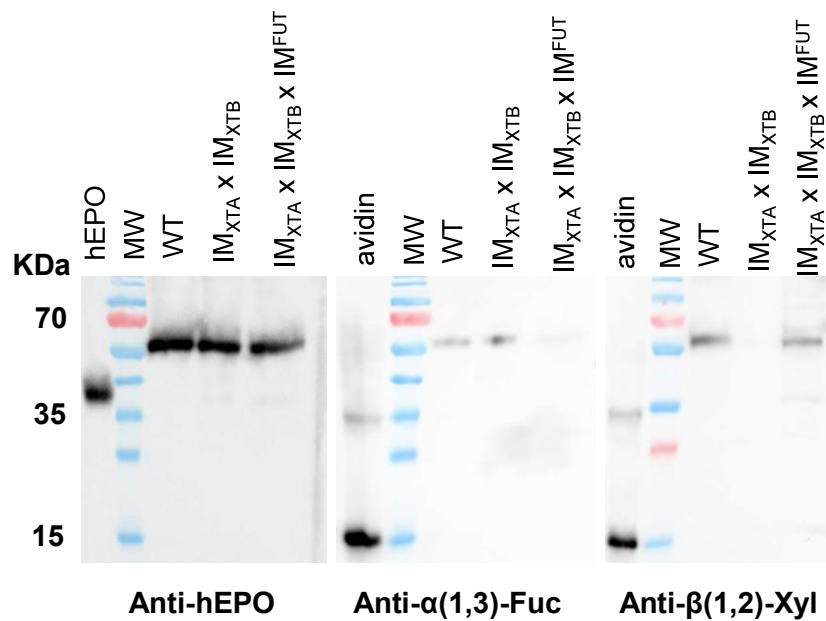
	Glyco-site	Peptide sequence	Peptide mass	glyco peptide	Relative intensity (%)	Charge	Therical glycan mass	Calculated glycan mass	Structure <sup>1</sup>			
CC5325	N44	K.EAENITIGCAE.H	1194.49	1292.52	19	2+	1388.51	1388.55	M4XF****			
				1358.54	12	2+	1522.54	1522.59	M5XF**			
				1439.57	29	2+	1684.65	1682.55	GM5XF**			
				1520.59	39	2+	1846.64	1844.69	G2M5XF**			
	N57	E.HCSLLENITVPDTK.V	1627.74	1001.74	4	3+	1374.50	1374.36	M4XF***			
				1006.10	6	3+	1388.51	1388.56	M4XF****			
				1050.79	6	3+	1522.54	1521.63	M5XF**			
				1054.79	5	3+	1536.55	1533.63	M5XF***			
				1104.47	12	3+	1684.65	1683.67	GM5XF**			
				1154.14	32	3+	1832.63	1832.58	G2M5XF*			
N103	R.GQALLVNSSQPWEPLQLHVDK.A	2359.24	1159.14	35	3+	1846.64	1846.68	G2M5XF**				
			1245.59	22	3+	1374.50	1374.53	M4XF***				
			1249.92	21	3+	1388.51	1388.52	M4XF****				
			1294.61	14	3+	1522.54	1522.19	M5XF**				
			1348.63	33	3+	1684.65	1684.55	GM5XF**				
			1402.97	11	3+	1846.64	1846.67	G2M5XF**				
			IM <sub>XTA</sub> xIM <sub>XTB</sub>	N44	K.EAENITIGCAE.H	1194.49	1205.47	21	2+	1214.44	1214.45	M4F*
							1212.48	35	2+	1228.45	1228.50	M4F**
							1278.50	7	2+	1244.45	1244.49	M5**
							1368.01	9	2+	1538.54	1539.53	GM5F*
N57	E.HCSLLENITVPDTK.V	1627.74		1448.55	28	2+	1700.58	1700.61	G2M5F*			
				948.07	5	3+	1214.44	1214.47	M4F*			
				952.75	12	3+	1228.45	1228.51	M4F**			
				1006.77	5	3+	1390.55	1389.57	M5F**			
				1056.44	10	3+	1538.54	1538.56	GM5F*			
				1110.13	52	3+	1700.58	1699.95	G2M5F*			
N103	R.GQALLVNSSQPWEPLQLHVDK.A	2359.24	1114.80	17	3+	1714.62	1713.65	G2M5F**				
			1191.95	10	3+	1214.44	1213.61	M4F*				
			1196.90	28	3+	1228.45	1228.46	M4F**				
			1241.25	41	3+	1360.50	1361.51	M4XF**				
			1245.69	11	3+	1376.49	1374.83	M5F*				
			1353.95	10	3+	1700.58	1699.81	G2M5F*				
			IM <sub>XTA</sub> xIM <sub>XTB</sub> xIM <sub>FUT</sub>	N44	K.EAENITIGCAE.H	1194.49	1125.35	23	2+	1054.37	1054.21	M4
							1206.34	28	2+	1216.42	1216.18	M5
							1272.48	47	2+	1348.46	1348.46	M5X
							894.95	26	3+	1054.37	1054.37	M4
N57	E.HCSLLENITVPDTK.V	1627.74		939.06	15	3+	1186.41	1186.46	M4X			
				948.96	21	3+	1216.42	1216.14	M5			
				993.01	30	3+	1348.46	1348.29	M5X			
				1057.10	8	3+	1540.51	1540.56	G2M5			
				N103	R.GQALLVNSSQPWEPLQLHVDK.A	2359.24	1089.90	11	3+	906.33	906.46	M3*
							1138.81	26	3+	1054.37	1054.19	M4
1192.52	21	3+	1216.42				1216.22	M5				
1236.87	42	3+	1348.46				1348.37	M5X				

**Table S2.** Distribution of *N*-glycans on the 3 glycosylation sites of rhEPOs produced in *C. reinhardtii* CC5325 strain, IM<sub>XTA</sub>xIM<sub>XTB</sub> and IM<sub>XTA</sub>xIM<sub>XTB</sub>xIM<sub>FUT</sub> mutants. Relative intensity for each glycopeptide was deduced from their relative ESI ion intensity. \**O*-methylation. <sup>1</sup>See Table S1 for structure of *N*-glycans.

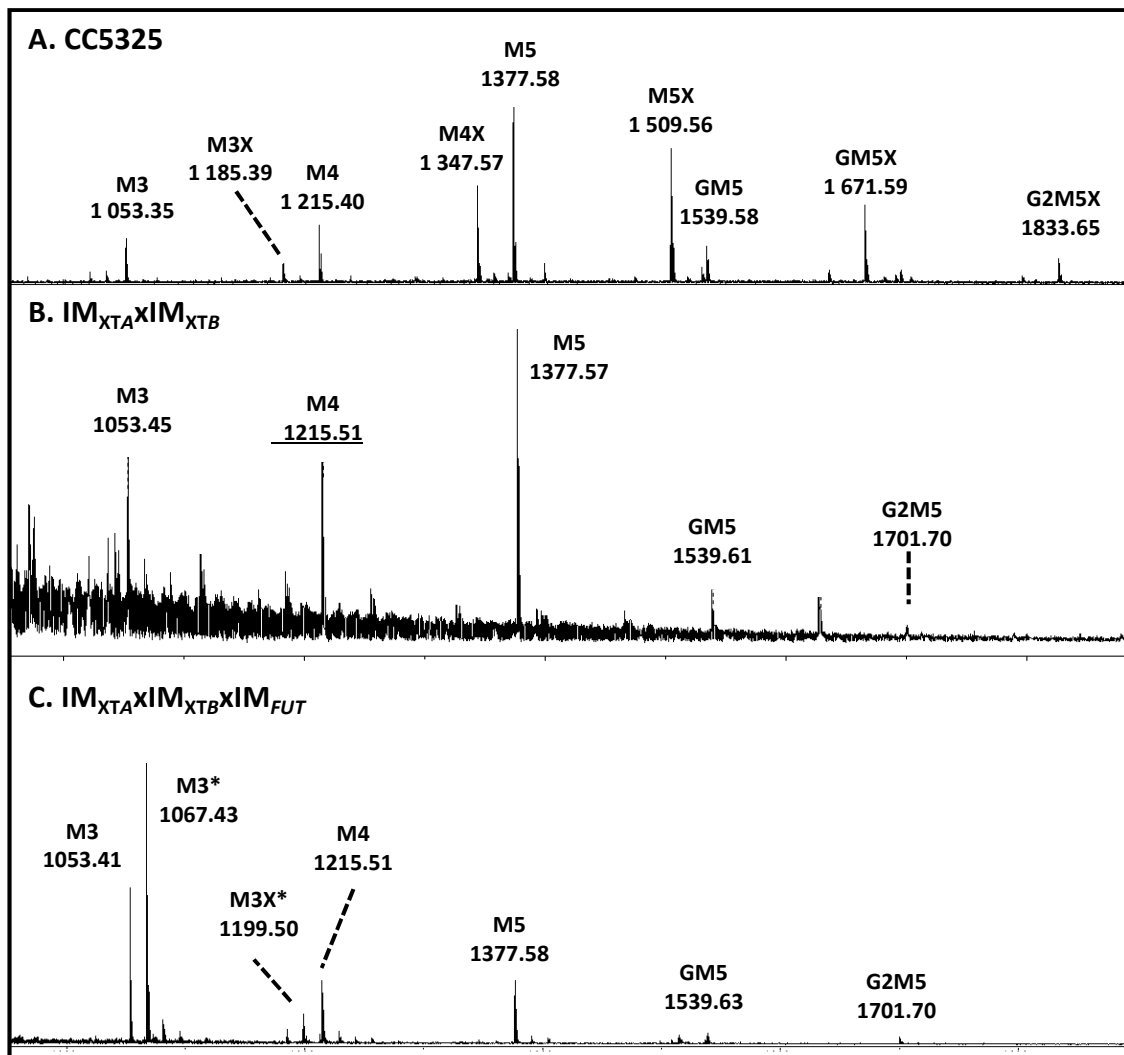




**Figure S1.** Sequence coverage determined by LC-ESI MS/MS analysis of the rhEPO-clover expressed in the *C. reihardtii* CC5325 strain. Black arrows point out the glycosylation sites.



**Figure S2. Western-blot analysis of secreted rhEPO expressed in *C. reinhardtii*.** Western-blot analysis were performed on rhEPO purified from CC5325,  $IM_{XTA} \times IM_{XTB}$  and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  mutants, using an anti-hEPO antibody (left panel) or anti- $\alpha(1,3)$ -Fuc (middle panel) or  $\beta(1,2)$ -Xyl antibodies (right panel).



**Figure S3.** MALDI-TOF MS of 2AB-labelled *N*-glycans released by PNGase A from proteins of *C. reinhardtii* CC5325 (A),  $IM_{XTA} \times IM_{XTB}$  (B) and  $IM_{XTA} \times IM_{XTB} \times IM_{FUT}$  (C). See Table S1 for structures.

# Chapter 2: Elimination of immunogenic epitopes by genome editing

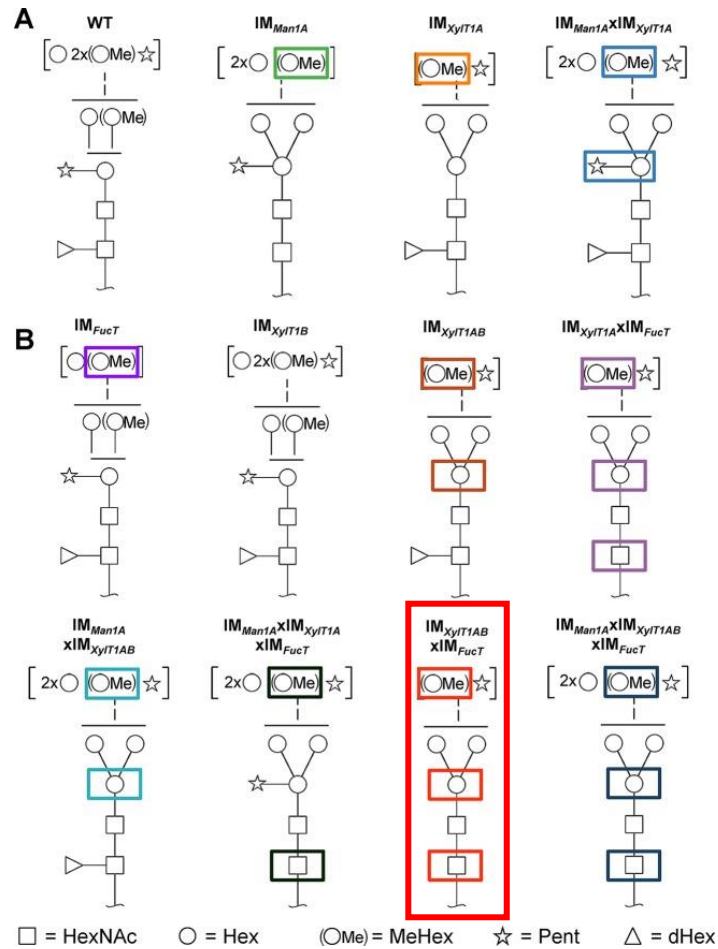


## I. Context

As mentioned in the introduction of this manuscript, the last ten years of research on the protein *N*-glycosylation in *C. reinhardtii* have made it possible to identify the nature of the glycans *N*-linked to proteins and in particular the complex *N*-glycans resulting from the Golgi steps which carry, respectively  $\beta(1,2)$ ,  $\beta(1,4)$ -xylose and  $\alpha(1,3)$ -fucose (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020; Schulze *et al.*, 2018) residues. Such specific residues are absent from human glycoproteins and are reputed to be immunogenic (Bardor *et al.*, 2003). Therefore, in the context of bioproduction, this microalga is incompatible in the purpose of using it for the expression of recombinant glycosylated biopharmaceuticals suitable for human therapy.

The previous work of Anne OLTMANNS and Pierre-Louis LUCAS, former PhD students at the University of Münster's Plant Biochemistry and Biotechnology Laboratory and the University of Rouen's Glyco-MEV Laboratory respectively, on various glycomutants has led to the proposal of the xylosylation pathway described in figure 18 of the introduction in which XTA is mainly responsible for the addition of core  $\beta(1,2)$ -xylose and XTB is mainly involved in the transfer of  $\beta(1,4)$ -xylose onto an  $\alpha(1,2)$ -mannose residue and, to a lesser extent, it performs a similar function to XTA.

The characterisation of mature *N*-glycans structures and of the effector enzymes involved in their synthesis was achieved by using insertion mutants from the CLiP library (Li *et al.*, 2016, 2019) and crossing them to generate multiple mutants (Schulze *et al.*, 2018; Lucas *et al.*, 2020; Oltmanns *et al.*, 2020; Schulze *et al.*, 2020). Among the various mutants studied so far shown on Figure 21, the  $IM_{XylIT1AB}IM_{FucT}$  triple mutant stood out for its reduced ability to fucosylate and xylosylate endogenous proteins (Oltmanns *et al.*, 2020). This phenotype was also confirmed when the *N*-glycosylation profile of human erythropoietin expressed in this strain has been analysed, as described in the first chapter of results of this manuscript. This confirms that knock-out strategies developed for the deletion of *N*-glycan immunogenic epitopes are suitable for the glyco-engineering of a therapeutic protein.



**Figure 21 : Schematic representation of the main N-glycan structures identified in glycoengineered mutants compared with a wild-type strain modified from (Oltmanns *et al.*, 2020).**

A, N-glycan compositions of WT and glycoengineered strains described in (Schulze *et al.*, 2018). B, N-glycan compositions of glycoengineered strains analyzed in (Oltmanns *et al.*, 2020). The monosaccharides represented above the continuous horizontal line can be found in the structures described. The Symbol Nomenclature for Glycans is described in (Varki *et al.*, 2009). XyIT1A, core xylosyltransferase; XyIT1B, second minority core xylosyltransferase; FucT, fucosyltransferase. Squared in red, the  $IM_{XyIT1AB}xIM_{FucT}$  ( $IM_{XTA}xIM_{XTB}xIM_{FucT}$ ) that has been chosen as strain of interest for this chapter.

Although xylosylation has decreased, a residual activity is still present in this triple mutant, indicating a potential redundancy and the presence of other genes involved in xylosylation as hypothesised in (Lucas *et al.*, 2020). In the same publication, Lucas and co-authors performed an *in-silico* research for genes encoding putative xylosyltransferases by sequence alignment with other endogenous or exogenous xylosyltransferases already identified. Three other amino acid sequences lacking the N-terminal CTS domain required for Golgi glycosyltransferases encoded by either Cre10.g458950, Cre13.g588750 or Cre08.g361250 were identified (Figure 19).

They show a homology, particularly in the C-terminal part, with XTA and XTB (Lucas *et al.*, 2020). Therefore, they are representing good candidates for xylosylation activity, especially the ones leading to the residual xylosylation in the double and triple mutants.

Interestingly, it has been shown that inactivating the Cre08.g361250 gene produces an acetate-requiring strain. It would therefore appear from this study that this gene could be also involved in photosynthesis (Wakao *et al.*, 2021).

In this context and since interesting insertion mutants were not available in the CLiP library, it was decided to inactivate by CRISPR/Cas9 the gene encoding the putative XTs independently in the triple mutant  $IM_{XTA}IM_{XTB}IM_{FucT}$  in order to characterise them. As xylosylation has already decreased in this strain, it is therefore expected that inactivation of the candidate genes will lead to a total loss of the xylosyltransferase activity and thus the synthesis of non xylosylated and non fucosylated glycoproteins.

This strategy is expected to be useful to create a strain that express non-xylosylated recombinant therapeutic glycoproteins in the culture medium, as it is already the case for current production models.

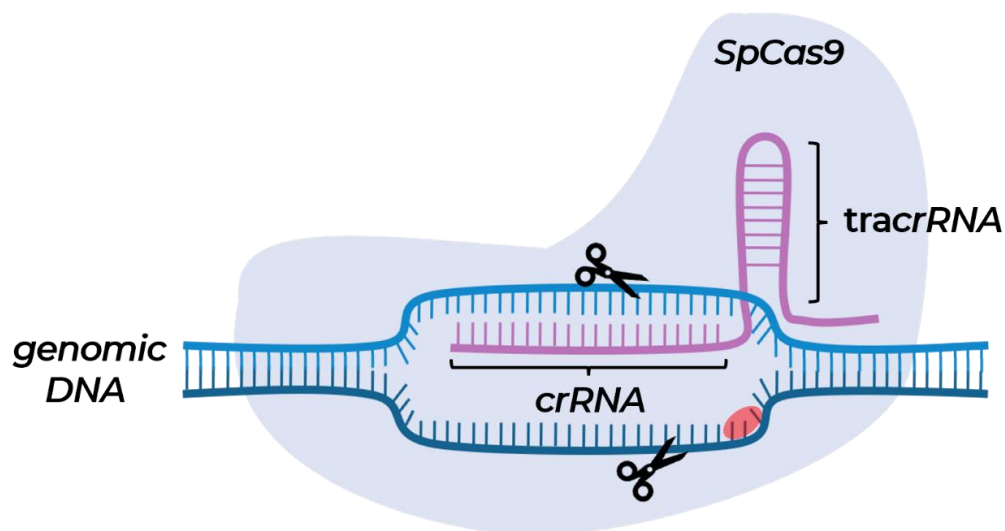


## II. Results

### A. CRISPR/Cas9 genome editing

Two targets per candidate gene were predicted *in silico* in the first third of the coding sequence, preferably in the middle of an exon, in accordance with the method adapted from (Kelterborn *et al.*, 2022) described in the Materials and Methods chapter.

The aim is to generate a mutation that leads to the insertion of a STOP codon as early as possible in the gene using ribonucleoproteins (RNPs). Briefly, a ribonucleoprotein is an assembly of a guide RNA (crRNA), a trans-activating RNA (tracrRNA) and a Cas9 protein as described in Figure 22. The resulting RNP will act on the gene by cutting it at the three base pair (NGG) protospacer adjacent motif (PAM) that follows the DNA sequence targeted.

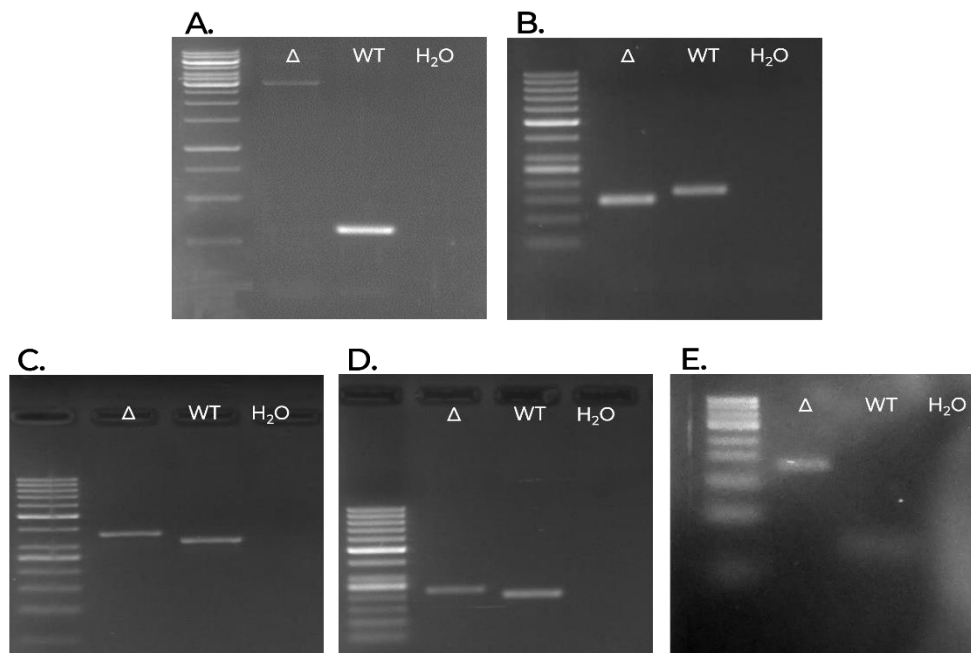


**Figure 22 : Scheme of a ribonucleoproteine (RNP) complex acting on genomic DNA.**

The crRNA contains the complementary sequence of the target in the gene and is partially complementary with the tracrRNA. Together, they form the two-pieces guide RNA that specifically recognises the targeted part of the gene within the genomic DNA and, in this case, recruits the *Streptococcus pyogenes* Cas9 (SpCas9). The PAM sequence (NGG) is represented by a red circle.

The  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  triple mutant strain has therefore been transformed, by electroporation with these RNPs, a FlagV3 sequence which is a 30 bp sequence containing STOP codons forward or reverse and a plasmid containing the hygromycin resistance cassette, at the Ruhr Universität of Bochum thanks to Dr. Anja HEMSCHMEIER who kindly provided us with the NEPA electroporator. The UVM4 strain has also been transformed following the same protocol as a control. Colonies selected on TAP medium supplemented with agar and 25 $\mu$ g/mL hygromycin were then screened by PCR.

Despite several attempts of transformation and hundreds of colonies screened, only one mutant per candidate gene (Cre10.g458950, Cre13.g588750 or Cre08.g361250) could be detected by PCR screening in both  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  and UVM4 strains, with the exception of Cre10.g458950 which could only be successfully mutated in strain UVM4 (Figure 23). The results of analyses carried out on mutants obtained in strain UVM4 were not exploitable, therefore this manuscript focuses on mutants obtained in the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain.



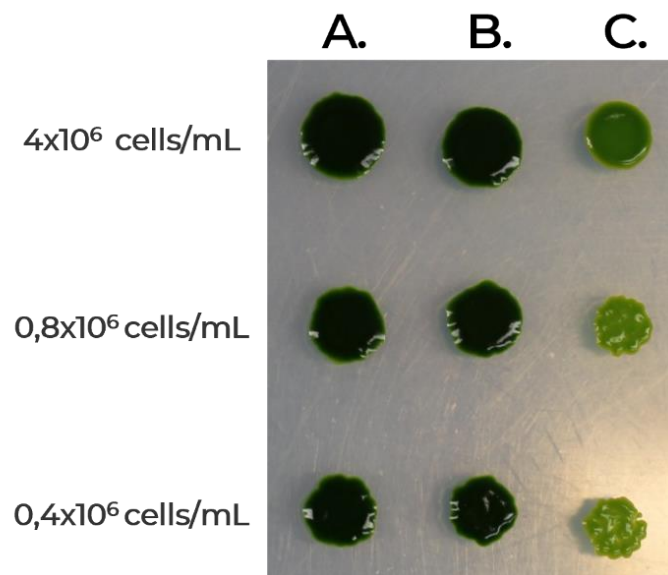
**Figure 23 : PCR Screening of the CRISPR/Cas9 mutants obtained in the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain (A. and B.) and the UVM4 strain (C., D. and E.).**

During the transformation process, a 30 bp sequence containing STOP codons and called FlagV3 is added to the transformation master mix as well as a plasmid containing a resistance cassette and the RNPs. After cleavage of the DNA a repair with the FlagV3 sequence insertion is expected. However other repairs can occur such as bigger insertions or deletions thus explaining the different electrophoretic profiles observed in the five panels. A.  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta$ Cre08.g361250 ; B.  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta$ Cre13.g588750 ; C. UVM4 $\Delta$ Cre08.g361250 ; D. UVM4 $\Delta$ Cre10.g458950 ; E. UVM4 $\Delta$ Cre13.g588750. WT stands for the control with the parental strain and H<sub>2</sub>O stands for the water control.

DNA electrophoresis gels show a variation in amplicon size resulting from amplification of the mutated region following cleavage by RNPs. The amplification of the native gene Cre08.g361250 with the primers we designed would result in an amplicon made of 280 base pair (bp). However, a difference of 1330 bp is observed (Figure 23 A) thus indicating an insertion in the targeted gene. In contrast, it appears that a deletion has occurred in Cre13.g588750, as shown by the 23 bp difference from the predicted 168 bp native amplicon (Figure 23 B).

Such a result is probably due to the random nature of the post-cut repair phenomenon. Genome editing using CRISPR/Cas9 exploits the cellular machinery capacity to repair DNA errors. The cell will either repair the gene identically, shift the reading frame or insert or remove a few base pairs. The latter events are hardly perceptible in PCR screening, which is why the 30bp FlagV3 sequence was added to the transformation mix. However, a more significant insertion or deletion can be observed, as shown in Figure 23. Finally, the mutation may be the result of a combination of events, such as a simultaneous deletion and insertion.

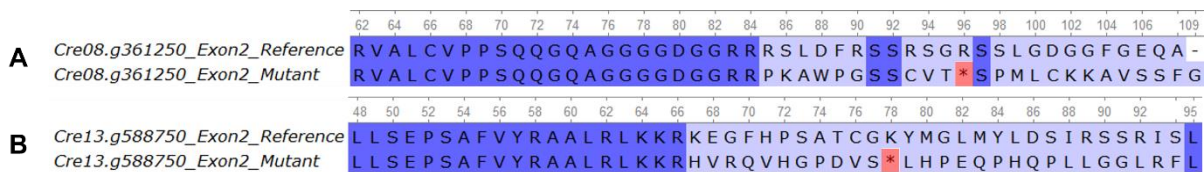
Furthermore, random insertions in the genome can occur during the electroporation process thus generating unwished mutations in other genes. The  $IM_{XTAx}IM_{XTBx}IM_{FucT} \Delta Cre13.588750$  strain has shown a pale green colour compared to the the parental strain dark green colour (Figure 24) when cultivated in normal growth conditions. This might be due to random insertions in genes involved in the chlorophyll content. As shown by the size of the spots of this mutant in spot culture, it would appear that the mutation has an impact on growth. This would be consistent with a mutation impacting the photosystem as it is one of the sources of energy of this mixotrophic organism. It can be assumed that the protein production could also be impacted.



**Figure 24 : Spot culture of A.  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  triple mutant, B.  $IM_{XTAx}IM_{XTBx}IM_{FucT} \Delta Cre08.g361250$  and C.  $IM_{XTAx}IM_{XTBx}IM_{FucT} \Delta Cre13.g588750$  on TAP agar medium in normal light at 25°C.**

The colour of strain  $IM_{XTAx}IM_{XTBx}IM_{FucT} \Delta Cre13.g588750$  is paler than that of the other two strains that grew on the plate, indicating a difference in chlorophyll composition that may have an impact on the whole metabolism. After 48h of liquid culture, 10 $\mu$ L of each strains have been inoculated on TAP medium supplemented with 2% agar with the following concentration A.  $4 \times 10^6$  cells/mL. B.  $0,8 \times 10^6$  cells/mL. C.  $0,4 \times 10^6$  cells/mL. The photo was taken after seven days of cultivation.

The amplicons resulting from the PCR screening were sent to Eurofins for Sanger sequencing in order to verify the nature of the mutation and the correct insertion of a STOP codon in the reading frame of the targeted gene. The genomic editing of the Cre08.g361250 candidate resulted in the insertion of the FlagV3 sequence, part of the selection plasmid and junk DNA concomitantly with a 4 bp deletion in exon 2 leading to the modification of the amino acid sequence and the insertion of a stop codon in the reading frame (asterisk highlighted in red in Figure 25). For the Cre13.g588750 candidate, a 23 bp deletion impacted exon 2 lead to a modification of the amino acid sequence and insertion of a stop codon in the reading frame (asterisk highlighted in red in Figure 25).



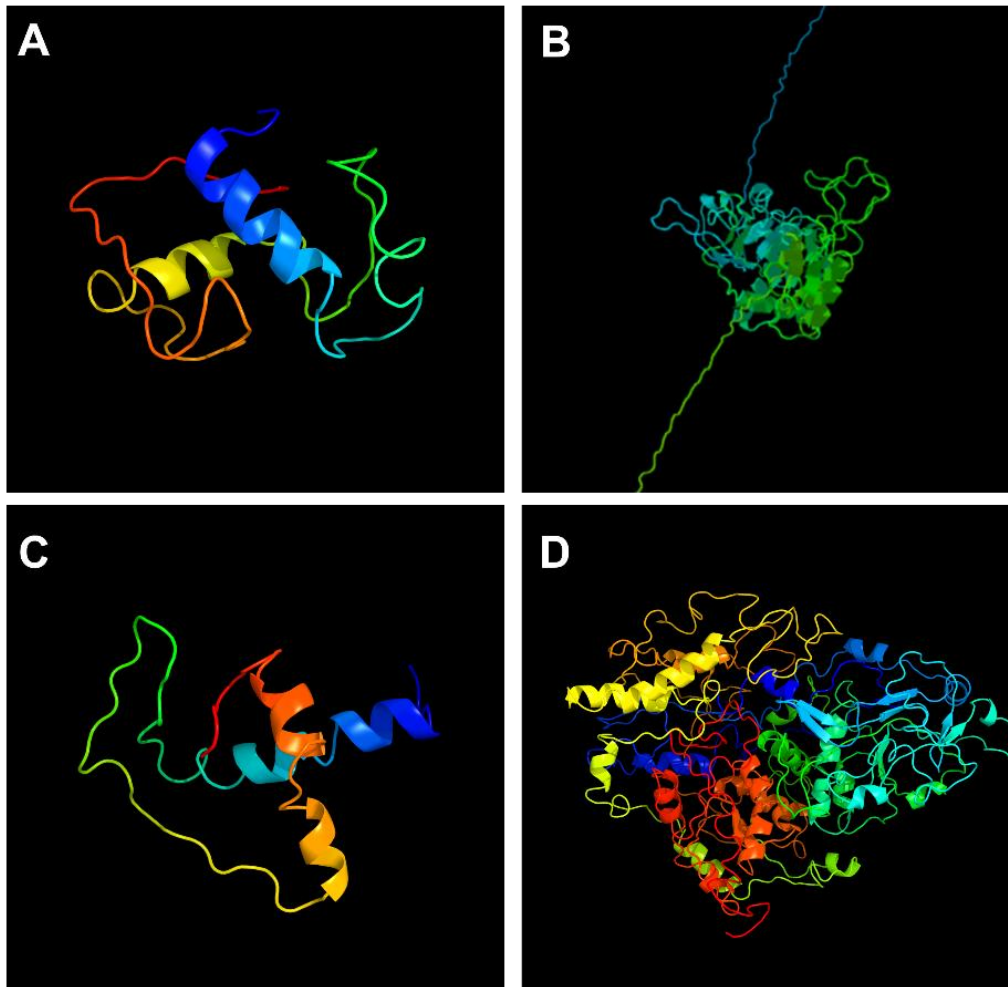
**Figure 25 : Sequence alignment of amino acid sequences**

Sequences from Sanger sequencing have been translated into one letter amino acid sequences. Sequences that are shown have the first STOP codon in the reading frame. A. Cre08.g361250 mutated amino acid sequence is compared to the reference. B. Cre13.g588750 mutated amino acid sequence is compared to the reference. Dark blue highlighted amino acids are strictly identical, light blue highlighted amino acids are different and asterisks highlighted in red stand for a STOP codon. The dash stands for a missing amino acid.

Therefore, the sequences resulting from the sequencing were translated into amino acid sequences for comparison with the reference sequences from the Phytozome bank and a 3D structure prediction has been done using the Phyre<sup>2</sup> webtool (Kelley *et al.*, 2015) (Figure 26).

The protein resulting from the translation of the Cre08.g361250 edited gene would be composed of 95 amino acids, compared with the 1063 amino acids of the native protein. Of the 95 amino acids that would finally be expressed, only the first 84 amino acids would be identical to the reference sequence of the native protein, i.e. 0.08% of the total protein.

In the case of the Cre13.g588750 edited gene, translation of such a sequence would only produce a protein composed of 77 amino acids, 66 of which are identical to the reference sequence, i.e. 0.08% of the reference native protein would be expressed.



**Figure 26 : 3D modelisation predictions of the truncated proteins compared to the corresponding native predicted 3D structure.**

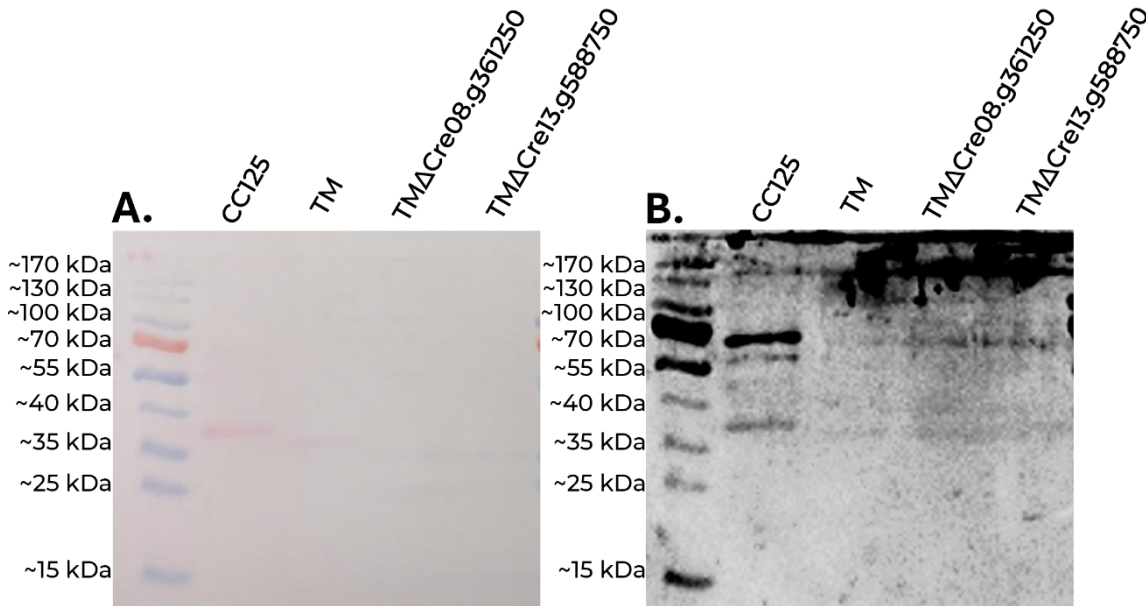
All these 3D structures have been predicted using the Phyre<sup>2</sup> webtool in intensive mode (Kelley *et al.*, 2015). A represent the truncated protein encoded by the gene Cre08.g361250 whereas B (highly zoomed) represent its putative native modelisation. C represent the truncated protein encoded by the gene Cre13.g588750 whereas D represent its putative native modelisation. Image coloured by rainbow N → C terminus.

Other stop codons were generated as a result of these modifications elsewhere in the genes and are not shown in Figure 25. It is therefore highly unlikely that the protein, and in particular its active site, was expressed. If this were the case, the sequence would have been greatly modified, thereby impacting the catalytic site.

Now that it had been established that a stop codon had been inserted preventing translation of the proteins encoded by the genes Cre08.g361250 and Cre13.g588750 respectively, it has been decided to analyse the proteins secreted in the culture medium by both mutants and to compare them with the parental strain using a biochemical approach.

## B. Biochemical analysis

Firstly, a biochemical analysis by western blot was carried out using the polyclonal HRP antibody mixture (P7899, Sigma-Aldrich) described in (Oltmanns *et al.*, 2020) which specifically recognises both  $\beta$ 1,2-xylose and  $\alpha$ 1,3-fucose residues. Twenty micrograms of proteins from the culture medium have been first separated by SDS-PAGE. Wild-type strain CC125 was used as a positive control (Figure 27).



**Figure 27 : Immunoblot analysis of proteins from the culture medium of CC125,  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  triple mutant (TM) and the two mutants obtained via genome editing  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta$ Cre08.g361250 (TM $\Delta$ Cre08.g361250) and  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta$ Cre13.g588750 (TM $\Delta$ Cre13.g588750).**

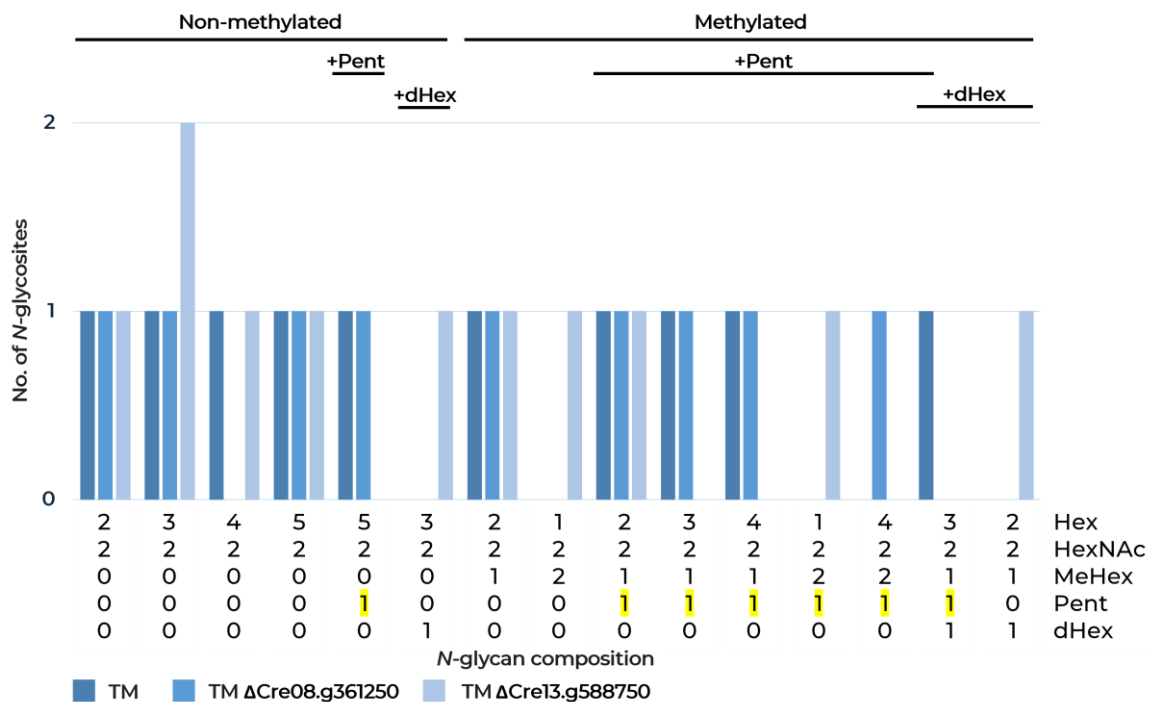
$\beta$ (1,2)-xylose and  $\alpha$ (1,3)-fucose are strongly recognised by the anti-HRP antibodies on the CC125 positive control. The recognition is decreased for the TM as well as for the two mutants. No significant difference is noticed between the mutants and the parent strain TM. A. Ponceau red staining of protein transferred onto the nitrocellulose membrane. B. Immunodetection of  $\beta$ (1,2)-xylose and  $\alpha$ (1,3)-fucose by anti-HRP antibodies.

In the cc125 reference strain track, xylosylated and/or fucosylated proteins are recognised by anti-HRP antibodies as expected. For the triple mutant, the western blot is once again consistent with the phenotype observed in the article by (Oltmanns *et al.*, 2020) as well as in the article in chapter I of this results section. Indeed, the signal is much weaker for this strain compared to cc125 reference strain, indicating a lower xylose and fucose content. When the mutants for the Cre08.g361250 and Cre13.g588750 genes were compared to their parental strain, the triple mutant, there was no significant difference in the intensity of the signals.

This analysis therefore suggests that the mutations had little or even no impact on the *N*-glycosylation profile of the secreted proteins. In order to verify this and take the analysis further, it was decided to carry out glycoproteomic analyses on secreted proteins and glycomic analyses on total intracellular proteins.

### C. Mass spectrometry analysis of *N*-glycans composition

Proteins from the culture media of the mutants generated and from the corresponding parental strains were prepared for glycoproteomic analysis. The experiments were carried out in technical triplicates and the raw results obtained were then processed using in-house developed Python tool SugarPy (Schulze *et al.*, 2020). Peptide sequences were identified at the MS2 level by analysis of intact *N*-glycopeptides. The composition of the *N*-glycans was deduced from a series of neutral losses in the corresponding MS1 spectra. SugarPy then generated all possible combinations for a given list of lost monosaccharides and maximum glycan length. In this study, the monosaccharides found predominantly in the *N*-glycans of *Chlamydomonas reinhardtii* (HexNAc, Hex, MeHex, dHex, Pent) were used. The data was then sorted manually to retain only the common peptides bearing predicted glycan structures likely to be found in *C. reinhardtii* in accordance with the literature (Figure 28) (Vanier *et al.*, 2017).



**Figure 28 : *N*-glycan composition of  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  (TM) strain and the  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta Cre08.g361250$  ( $TM\Delta Cre08.g361250$ ) and  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta Cre13.g588750$  mutants ( $TM\Delta Cre13.g588750$ ).**

Bar blots representing comparison of the *N*-glycan composition of *N*-glycopeptides that were identified in the 3 strains. Bars height represent the number of *N*-glycosites identified. Above the bars an indication can be found, to which *N*-glycan type the respective composition can be assigned to. Under the bars, pentosylated (xylosylated) species are highlighted in yellow. Hex : Hexose ; HexNAc : *N*-acetylhexosamine ; MeHex : MethylHexose ; Pent : Pentose ; dHex : deoxyHexose.



As expected, short glycans bearing one to three hexose (mannose) residues are mostly found in all of the three strains. Indeed, this phenotype has been reported for the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  in previous studies and in the first chapter of this manuscript. It can therefore be assumed that the mutations did not impact the glycan length in this genetic background.

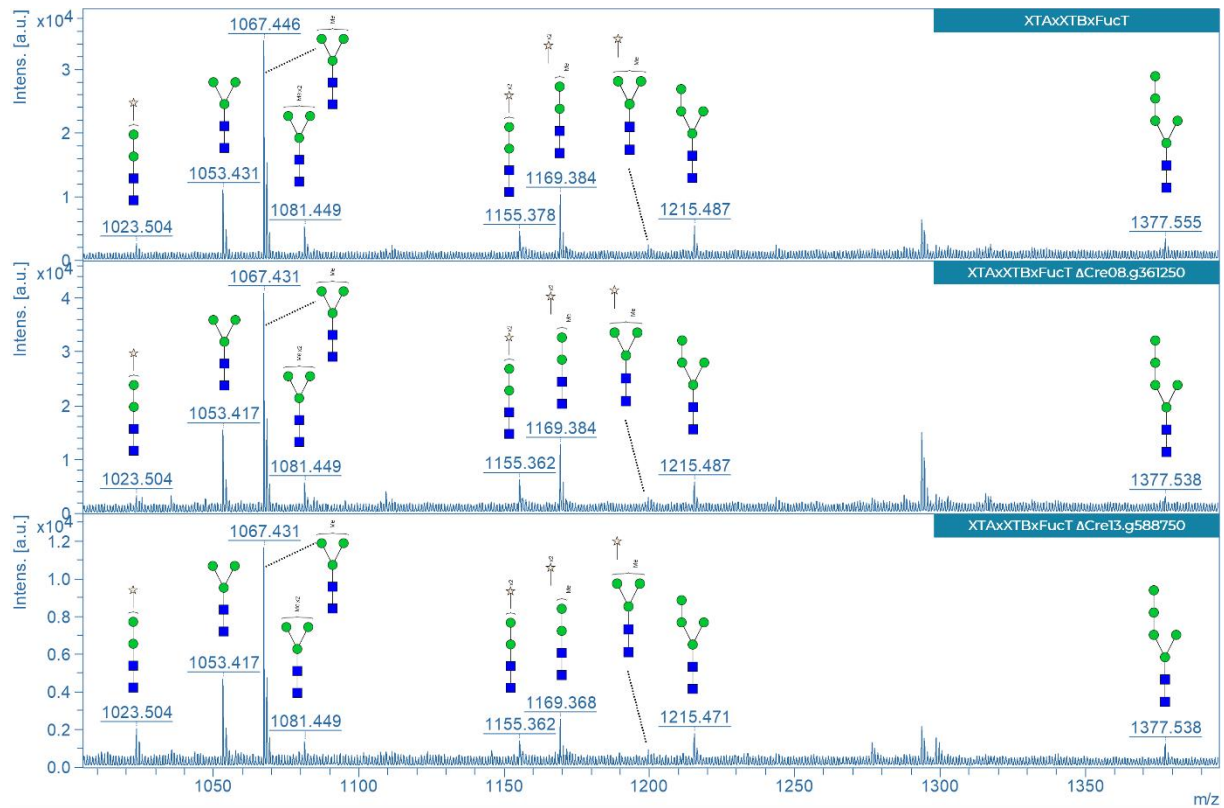
At first sight, glycoproteomic analyses (Figure 28) reveal a higher proportion of *N*-glycosites occupied by methylated glycans. Many of the latter are mono-pentosylated compared with *N*-glycosites occupied by non-methylated glycans. It is worth noting that no di-pentosylated species were detected, although a few species bearing deoxyhexoses were found.

Of the 10 *N*-glycosites identified in the reference strain, 5 are bearing mono-pentosylated glycans, including 1 mono-pentosylated and methylated. The same observation was made of the 9 *N*-glycosites identified in the strain in which the Cre08.g361250 gene is inactivated. For the strain in which the Cre13.g588750 gene is inactivated, 11 *N*-glycosites were identified, 2 of which are occupied by mono-pentosylated *N*-glycans.

It would therefore appear that, by observing the common *N*-glycosites of secreted proteins in the culture medium, xylosylation is reduced in  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta Cre13.588750$  but not completely eliminated. However, despite the replicates carried out, few common *N*-glycopeptides were identified when comparing the three strains, mainly due to the low number of peptides found in the  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta Cre13.588750$  strain.

It was also decided to analyse the glycan profile of total intracellular proteins. The proteins were deglycosylated with PNGase F, which is able to cleave matured glycans *N*-linked to Asn of glycosylation sites, except those having a  $\alpha 1,3$ -fucose linked to the proximal GlcNAc residue (Tretter *et al.*, 1991). The  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  spectrum shown in Figure 29 once again matched the results published in Oltmanns *et al* 2020. Unlike the wild-type strains, which contain more  $Man_5GlcNAc_2$ ,  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  contains more  $Man_3GlcNAc_2$ , some of which are methylated and/or xylosylated (Figure 29).

The *N*-glycan profiles of the two mutants obtained are very similar to the parental strain and do not show a reduction or disappearance of the xylose content (Figure 29). This would suggest that the mutations generated in putative XTs do not impact the xylosylation of total intracellular proteins.

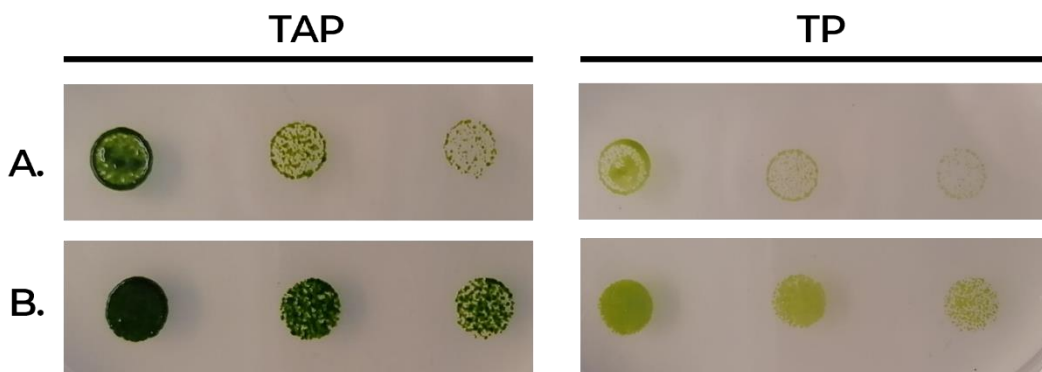


**Figure 29 : MALDI-TOF spectra of glycans isolated from whole cell proteins in the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain and the  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta Cre08.g361250$  and  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta Cre13.g588750$  mutants.**

Proteins have been extracted from a fresh cell pellet and loaded into an acrylamide gel in order to proceed with a *in gel* sample preparation. Reduction and alkylation have been done prior the trypsin/chemotrypsin digestion. *N*-Glycopeptides have been extracted from gel pieces and *N*-glycans have been released using a *N*-glycosidase F (PNGase F). After a 2 aminobenzamide labeling, *N*-glycans have been analysed using Matrix Assisted Laser Desorption Ionisation–Time Of Flight Mass Spectrometry (MALDI-TOF MS). *N*-glycan profiles of the mutants are similar to the corresponding mother strain showing non negligible xylosylated glycans.

#### D. Acetate requirement

As indicated in the context of this chapter, inactivation in the cc4051 strain of the Cre08.g361250 gene appears to lead to the creation of a strain requiring acetate. This would suggest that this gene is involved in photosynthesis (Wakao *et al.*, 2021). Although far from the main objective of this chapter, which is the elimination of immunogenic epitopes in *Chlamydomonas reinhardtii*, a simple growth test with (TAP) and without (TP) acetate was carried out on solid medium with the Cre08.g361250 mutant in the genetic background of the triple mutant (Figure 30).



**Figure 30 : Acetate requirement test comparing spot cultures of A.  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  and B.  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta Cre08.g361250$  on TAP and TP solid media.**

After 48h of liquid culture, 10 $\mu$ L of each strains have been inoculated on both TAP and TP media supplemented with 2% agar with the following concentration 4x10<sup>6</sup> cells/mL ; 0,8x10<sup>6</sup> cells/mL ; 0,4x10<sup>6</sup> cells/mL (From left to right). The photo was taken after three days of cultivation in normal light conditions (60-80  $\mu$ E) at 25°C.

When the two culture conditions are compared, it seems that growth is less important in the absence of acetate for both strains, based on the paler colour of the spot cultures. However, irrespective of the variation in medium, it would appear that the mutant strain grows faster than the parental strain based on the pattern of spot cultures, whatever the starting cell concentration. Our results therefore seem to contradict those obtained in the 2021 publication (Wakao *et al.*, 2021). Nevertheless, this test was only carried out once under a single light condition. It would therefore be interesting to repeat it under different growing conditions to explore this hypothesis further.

### III. Conclusion and discussion

#### A. *N*-glycosylation

Two mutants affecting the putative xylosyltransferases Cre08.g361250 and Cre13.g588750 were successfully generated in the  $IM_{XTA}IM_{XTB}IM_{FucT}$  parental strain with the *XTA*, *XTB* and *FucT* genes inactivated. It would have been interesting to obtain several clones with different mutations in the same gene in order to strengthen the results of the analyses by biological replicates. It would therefore be wise to repeat the experiment until a satisfactory number of clones have been obtained, in order to establish these results more firmly. This is all the truer for the Cre13.g588750 candidate gene, where the strain obtained by inactivating it appears to have a different phenotype to that expected and a growth deficiency (Figure 24).

Unfortunately, no mutants for the Cre10.g458950 gene could be generated in this genetic background. One hypothesis is that the RNPs designed for this gene are ineffective. However, this gene was mutated successfully in strain UVM4. The primary hypothesis was therefore abandoned in favor of the hypothesis that mutation of this gene would be lethal in the genetic background of the  $IM_{XTA}IM_{XTB}IM_{FucT}$  triple mutant. This conjecture still needs to be confirmed.

Immunoblot analysis of the proteins secreted into the culture medium using antibodies directed against  $\alpha(1,3)$ -fucoses and  $\beta(1,2)$ -xyloses did not reveal any major differences between the parental strain and the mutants. The conclusions that can be drawn from this technique are limited to a simple assessment of signal intensity. In addition, the detection of  $\alpha(1,3)$ -fucose epitopes does not allow any conclusions to be drawn regarding xylosylation.

Mass spectrometry analyses have allowed more conjectures. Glycoproteomic analyses revealed a lower content of pentosylated (xylosylated) species in the mutant whose Cre13.g588750 gene was inactivated. In contrast, no difference has been observed in the pentose (xylose) content of the Cre08.g361250 mutant. MALDI-TOF analysis of total intracellular proteins did not reveal any difference between the *N*-glycosylation profiles of the mutants and those of the parent strain. Altogether mass spectrometry analyses suggest that the Cre13.g588750 mutation only affects the xylosylation of proteins secreted into the culture medium whereas the Cre08.g361250 mutation has no impact on xylosylation. So far, analyses carried out on *N*-glycans of both total intracellular proteins and proteins secreted into the culture medium have not

revealed any difference in the *N*-glycosylation profile. This lends little support to the hypothesis that the mutation affects only excreted proteins.

Overall, the analyses did not reveal any disappearance of xylosylation. This seems to indicate that xylosylation is indeed under the control of a multigene family, as hypothesised by Pierre-Louis LUCAS in 2020. To verify this, it would be interesting to generate strains combining mutations in the three candidate genes identified. These multiple mutants could be generated by crossing mutants already obtained or by new genome editing. However, crosses have the advantage of generating progeny in which other mutation combinations are present, as was the case for the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  triple mutant, and reduce the risk of random insertion.

During glycoproteomic analysis, few *N*-glycopeptides in common with other strains were found in  $IM_{XTAx}IM_{XTBx}IM_{FucT}\Delta Cre13.588750$  strain. In addition, this mutant had a phenotype that had an impact on chlorophyll composition, as evidenced by its paler colour. It is therefore possible that this other phenotype, probably due to random insertion during electroporation transformation, could have an impact on protein production and secretion into the culture medium.

In order to verify whether the phenotypes are linked, the gene mutated by genomic engineering would have to be rescued, by transforming this strain with a plasmid containing the intact gene sequence, or by crossing.

The perspective of the creation of a strain synthesising proteins bearing non-xylosylated and non-fucosylated three mannose *N*-glycans could represent an opportunity to start further glycoengineering by expressing GnTI or GnT-like enzymes as those found in *Trypanosoma brucei*.

Finally, in the longer term, it would also be interesting to analyse *O*-glycosylation, which has been little studied to date. Indeed, the candidate genes have been identified as being glycosyltransferases, so it could be that the successfully mutated genes that do not lead to any significant phenotype encode proteins involved in protein *O*-glycosylation or xylosylation of other polymers. A second hypothesis would be that the Cre10.g458950 candidate gene that we did not manage to inactivate is responsible of the transfer of the residual xylose.

## B. Acetate requirement

Regarding the involvement of the Cre08.g361250 gene in photosynthesis, preliminary tests did not allow any conclusions to be drawn. Although our results are not consistent with those published in 2021 (Wakao *et al.*, 2021), the possibility that the mutant obtained is an acetate-requiring mutant cannot be ruled out. In addition, more independent mutants for this same gene should be analysed in order to support these results. The difference in results can also be explained by the genetic background used. Such an influence of the genetic background has already been observed in *Chlamydomonas reinhardtii*, notably in the study of Zinzius *et al.*, 2023 about calredoxin.

According to Wakao *et al.*, 2021, the mutant for Cre08.g361250 does indeed appear in the table listing the higher-confidence photosynthesis candidate genes, but so far, only two genes that are necessary for photoautotrophy have been experimentally identified. However, the Cre08.g361250 gene is not one of them. Based on the differences in acetate requirement phenotype shown in this article, it is tempting to wonder to what extent photoautotrophy is affected while this gene is inactivated. Furthermore, this gene is predicted to encode a glyco-enzyme. Nevertheless, a potential location within the chloroplast is shown which is not consistent with the fact that, so far, chloroplast does not allow this type of PTM.

Expression of the whole gene in fusion with a sequence encoding a fluorescent protein is a prospect that would make it possible to confirm its localisation within the cell. The results obtained could guide future research into the protein encoded by this gene.

# Chapter 3: Targeting the Golgi apparatus





## I. Context

It is now well known that the sequence of *N*-glycans depends on the distribution of glycoenzymes throughout the different Golgi cisternae (Wang *et al.*, 2017). Glycoenzymes involved in the maturation steps are type II membrane proteins, harbouring in their N-terminal part a particular domain called 'Cytoplasm, Transmembrane and Stem' (CTS) region. This domain has been demonstrated to be responsible for the anchorage of the enzymes in Golgi membranes as mentioned in the introduction of this manuscript. The composition in amino acids of the CTS domain and its length has been shown to be correlated to the thickness and the lipid composition of the organelle membrane thus influencing sublocation in the Golgian cisternae (Schoberer *et al.*, 2019; Sharpe *et al.*, 2010).

The Golgi stages of the *N*-glycosylation pathway in *Chlamydomonas reinhardtii* are independent of *N*-acetylglucosaminyltransferases and the resulting *N*-glycans are therefore different from those found in mammals (Mathieu-Rivet *et al.*, 2013, 2014; Vanier *et al.*, 2017). Furthermore, the presence of galactose or sialic acid residues in the *N*-glycans has never been demonstrated so far. These important differences, raise the question of the engineering of the *N*-glycosylation pathway to make this model microalga suitable for the expression of recombinant glycoproteins compatible with human therapeutic applications.

Successful efforts to humanise organisms such as plants and yeast have taken this parameter into account by using distinct CTS domains in order to express glycoenzymes in specific Golgi cisternae. As an example, the expression of a fusion protein composed of the CTS domain of the rat  $\alpha(2,6)$ -sialyltransferase in N-terminal position and the catalytic domain of human  $\beta(1,4)$ -galactosyltransferase enabled modification of the *N*-glycosylation pathway by anchoring this protein in the trans-Golgi. This approach improved the efficacy of antibodies produced in *N. benthamiana* against the HIV virus (Strasser *et al.*, 2009b). However, to reach this goal, it is necessary to first find the addressing sequences of glycoenzymes in the cisternae of the *C. reinhardtii* Golgi apparatus that remains unexplored.

We therefore set ourselves the challenge of initiating a study of the ability of CTS domains from heterologous enzymes as well as putative CTS domains from endogenous proteins to address glycoenzymes within the Golgi apparatus cisternae.

To achieve this, constructs in which the sequences encoding these CTS domains are fused to the sequence encoding the mCherry reporter protein have been used to transform strains expressing a compartment marker from either the ER or the Golgi (CSI strains, expressing mVenus-tagged proteins (Mackinder *et al.*, 2017)). Therefore, the co-regionalisation of the fusion proteins <sup>CTS</sup>mCherry with the mVenus-tagged proteins resident will be analysed.

## II. Results

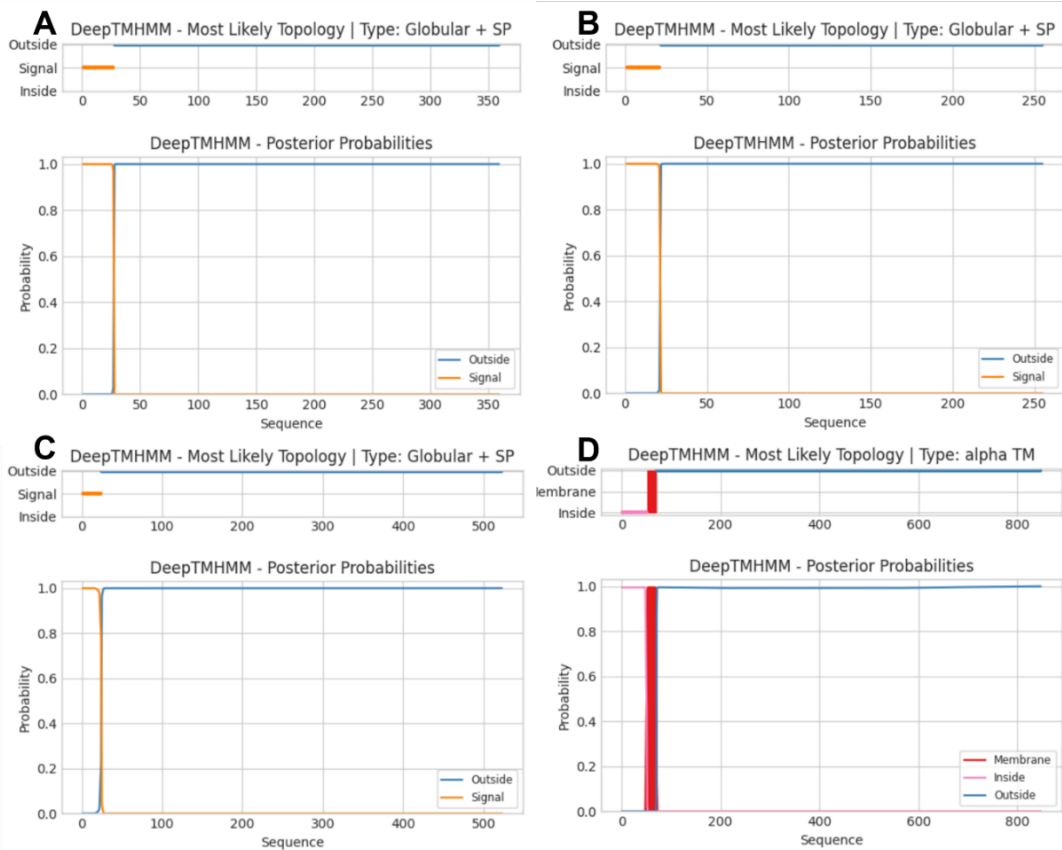
### A. Choice of the CTS domains

The strategy initially focused on identifying CTS domains, from endogenous proteins. Two genes, Cre10.g439500 and Cre16.g661700, have been predicted to encode putative Golgi proteins (Mackinder *et al.*, 2017) potentially having a CTS domain. Moreover, fluorescent signal resulting from the expression of each protein Venus-tagged exhibited a relatively punctate shape under confocal microscopy that would suggest a Golgian signal by analogy with other studies of this type in plant cells or algal cells (Vilarrasa-Blasi *et al.*, 2021; Gao *et al.*, 2012; NEUMANN *et al.*, 2003). A study by (Vilarrasa-Blasi *et al.*, 2021) also used the protein encoded by the gene Cre16.g661700, a remorin, as a Golgi marker of salt stress, thus further confirming the location of the protein encoded by this gene. The results give these genes good hope of finding N-terminal transmembrane domains.

Regarding endogenous candidate genes coding for type II membrane proteins, genes encoding XTA (Cre09.g391282) and XTB (Cre16.g678997) (Lucas *et al.*, 2020) were therefore considered, as xylosylation steps are thought to occur in the Golgi apparatus (Ropitiaux *et al.*, 2024).

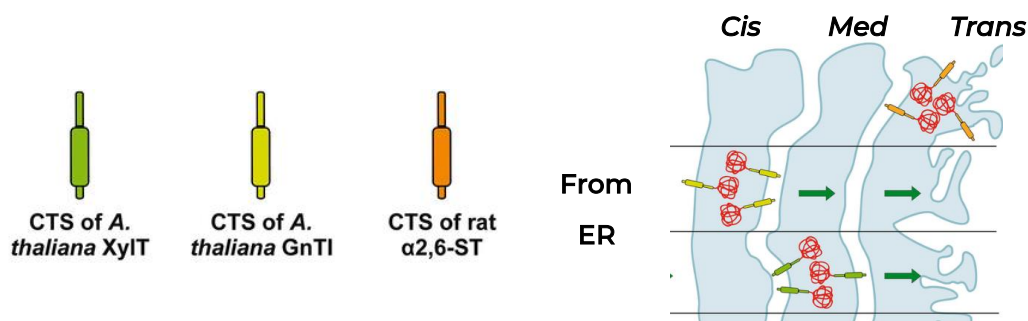
However, according to the updated DeepTMHMM1.0 transmembrane domain prediction tool (Hallgren *et al.*, 2022), only XTB is predicted to possess a transmembrane domain in its N-terminal part (Figure 31) corresponding to 70 AA covering almost the whole of exon 1 (72 AA) with 19 AA making up the alpha transmembrane helix. The size of the transmembrane helix predicted is consistent with other similar structures in plants composed of around 20 amino acids. Although it is difficult to delimit a CTS region *per se*, the prediction of such a transmembrane domain suggests the presence of a CTS region.

Furthermore, these predictions contradict those made with a previous version of the transmembrane helix prediction tool (TMHMM2.0) (Krogh *et al.*, 2001) about genes Cre10.g439500 and Cre16.g661700 (Supplemental Figure 40). Nevertheless, the degree of confidence in prediction tools remains debatable. Therefore, it has been decided to use them in our strategy.



**Figure 31 : Transmembrane domain prediction using deepTMHMM 1.0 (Hallgren *et al.*, 2022).** Amino acid sequences of A. Cre10.g439500, B. Cre16.g661700, C. Cre09.g391282 (XTA) and D. Cre16.g678997 (XTB) have been uploaded as a FASTA file. Graphics A, B and C are representing signal peptides prediction whereas graphic D is the only one representing a alpha transmembrane helix in the gene encoding the endogenous XTB.

In parallel, the coding sequences for the CTS domains of AtXylIT, AtGnT, RnST and AtMNSI (Schwientek *et al.*, 1995; Pagny *et al.*, 2003; Strasser *et al.*, 2005; Kajjura *et al.*, 2010), respectively located in the *median*, *cis*, *trans* (Figure 32) and early Golgi cisternae, were considered as well in this study.

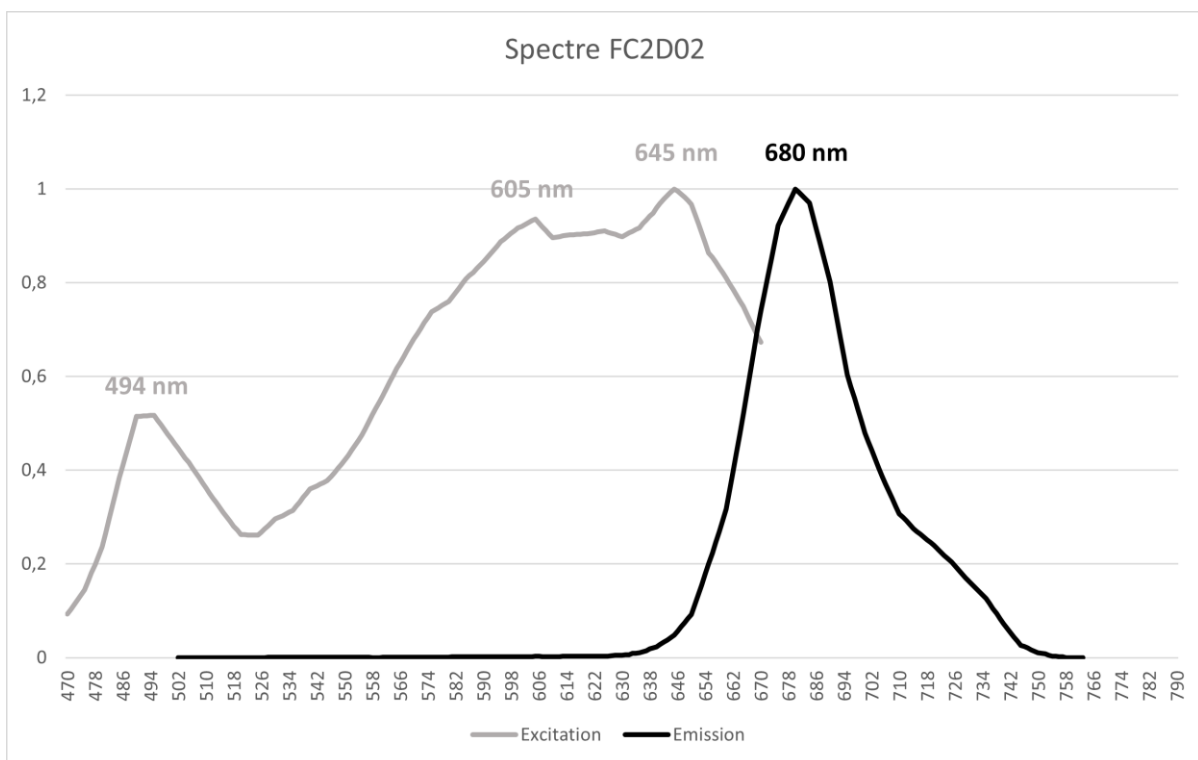


**Figure 32 : Localisation of the catalytic domain of the human  $\beta(1,4)$ -galactosyltransferase in fusion with different CTS domains within plant cells Golgi cisternae modified from (Loos and Steinkellner, 2014).**

The proteins fused to CTS domains of AtXylIT, AtGnT, RnST are respectively located in the median, cis and trans Golgi cisternae.

## B. Plasmid construction

To determine which fluorescent protein would be suitable as a reporter for tagging CTS domains, CSI strains intended for further expressing the fusion proteins were first submitted to a lambda scan analysis in order to determine the wavelength ranges of chlorophyll autofluorescence. The measurements have been done three times on three individual cells and repeated on three different days using a Leica SP5 confocal microscope. The data has been compiled to give the result shown in Figure 33.



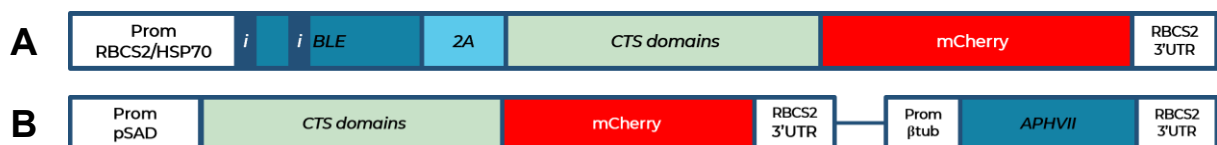
**Figure 33 : Lambda scan of CSI FC2D02**

Cells have been immobilised on a polylysine coated slide and measurement of excitation and emission have been done for each wavelength between 470 nm and 790 nm. The same experiment has been done on other CSI strains and gave similar results.

Despite the expression of a mVenus ( $\lambda_{Em}=527$  nm) fluorescent protein, no emission maximum corresponding to this one has been detected in CSI strains maybe due to a weaker signal compared to the chlorophyll. It has been concluded that whatever the excitation wavelength between 470 nm and 660 nm, the emission is close to zero from 502 nm to 630 nm and reaches its maximum at 680 nm corresponding to the chlorophyll autofluorescence. In other words, any fluorescent protein having a maximum of emission distinct from 680 nm can be used in *C. reinhardtii*.

In the light of these data, mCherry, whose maximum emission ( $\lambda_{Em}=610$  nm) is different from mVenus ( $\lambda_{Em}=527$  nm), has been chosen as a fluorescent tag.

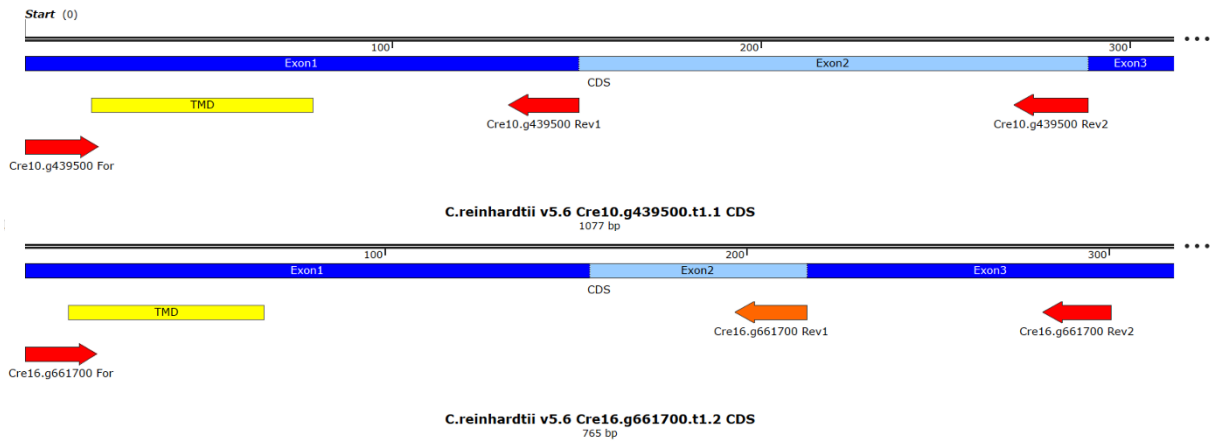
Therefore, two plasmids referenced in the literature have been used as backbone plasmids. Both of them contain the sequence encoding the mCherry fluorescent protein ( $\lambda_{Ex}=587$  nm,  $\lambda_{Em}=610$  nm). Two series of plasmids, one based on pLM006 (Mackinder *et al.*, 2016) and the other based on pBR09 (Rasala *et al.*, 2013) have been cloned (Figure 34).



**Figure 34 : Schematic representation of the constructs used for expressing proteins consisting in CTS domains fused to a mCherry ( $\lambda_{Ex}=587$ nm,  $\lambda_{Em}=610$ nm) fluorescent protein.**

In both constructs, the sequence encoding the CTS domain of interest is inserted and fused to the sequence coding for a mCherry fluorescent protein. A This plasmid is based on the pBR09 (Rasala *et al.*, 2013). The expression is allowed by the combination of the chimeric promoter RBCS2/HSP70A with RBCS2 3'UTR terminator. The selection is based on the zeocin resistance gene *BLE* disrupted by 2 introns and fused to the sequence encoding a FMDV2A self-cleaving peptide. B This plasmid is based on the pLM006 (Mackinder *et al.*, 2016). The expression is driven by the pSAD/RBCS2 3'UTR promoter/terminator combination. The selection is allowed by a separate expression unit allowing the expression of *APHVII* gene, conferring the resistance to hygromycin B.

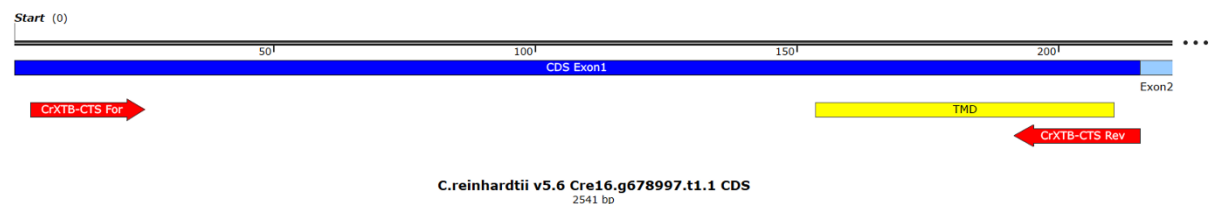
Two sets of primers were used to amplify the sequences corresponding to the putative CTS domains since it was not possible to predict them with confidence in Cre10.g439500 and Cre16.g661700. A first pair to amplify strictly the sequence predicted to encode the transmembrane domain and a second pair to amplify a broader sequence as shown in Figure 35. The sequences encoding CTS domain previously described (AtXylIT, AtGnT, RnST and AtMNSI) have been synthesised and amplified to be cloned by Gibson assembly or using restriction enzyme in pLM006, or pBR09, respectively (Figure 34).



**Figure 35 : Amplification scheme of the sequence corresponding to the predicted Cre10.g439500 and Cre16.g661700 CTS domains on gDNA.**

Exons are shown in blue with alternating shades to distinguish them from one another. The transmembrane region (TMD) of the CTS domain is shown in yellow. The primers are shown in red. Two pairs of primers per gene have been designed. CDS stands for CoDing Sequence.

Regarding the transmembrane domain from XTB, in accordance with the prediction exposed in Figure 31, the entire exon 1 of the gene *XTB* was amplified without its initiator codon as shown in Figure 36 and inserted into another backbone plasmid described in Figure 37 by Gibson assembly. The mClover fluorescent protein ( $\lambda_{Ex}=505$  nm,  $\lambda_{Em}=515$  nm) has been chosen for this construction. As it is incompatible with the concomitant expression of mVenus, this construct was designed to be used alone in a so-called wild-type strain with the aim of adopting a reverse approach to that initially proposed, i.e. transforming first with constructs containing the predicted CTS domains, studying the intensity and shape of the signal by confocal imaging and finally adding Golgian and/or ER markers via another plasmid that was also optimised.



**Figure 36 : Amplification scheme on gDNA of sequence corresponding to the N-terminal part of Cre16.g678997 (XTB) containing the predicted transmembrane helix.**

Exons are shown in blue with alternating shades to distinguish them from one another. The transmembrane region (TMD) of the gene is shown in yellow. The primers are shown in red. CDS stands for CoDing Sequence.

To optimise the expression of the fusion protein  $N\text{-ter}^{XTB}m\text{Clover}$ , the sequence was disrupted by three copies of the first RBCS2 intron and a single C-terminal copy of RBCS2 intron 2 using the online intronserter tool, in order to mimic the natural intron/exon distribution and thus facilitate recruitment of the transcriptional machinery. However, the choice of promoter/terminator combination has remained classic, with selection based on cleavage of the FMDV2A peptide between the nourseothricin resistance gene and the fusion protein of interest.

In addition to the fluorescent tag, a C-terminal 6xHN histidine affinity tag has been also added for further detection by electron microscopy (Figure 37).



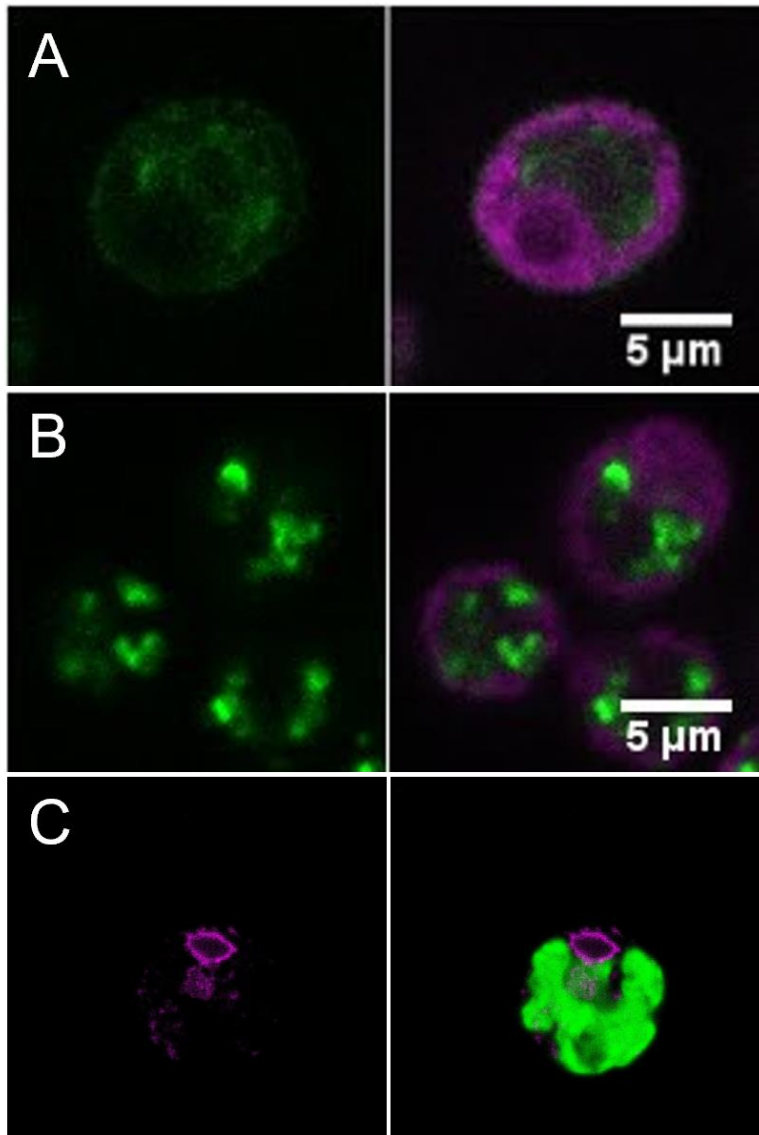
**Figure 37 : Schematic representation of the plasmid containing sequence corresponding to the N-terminal part of XTB containing the predicted transmembrane helix.**

The transmembrane domain of interest (XTB CTS in yellow) is inserted and fused to a Clover fluorescent protein. The expression is driven by the pAR (RBCS2/HSP70A) / RBCS2 3'UTR promoter/terminator combination. The selection is based on the nourseothricin resistance gene NAT fused to a FMDV2A self-cleaving peptide. A C-terminal HN-tag has been added. The whole sequence has been disrupted by 4 introns, the first three corresponding to intron 1 and the last corresponding to intron 2 of the RBCS2 gene, as proposed by the intronserter online tool. L stands as a linker between the fluorescent tag and the putative CTS region.



### C. Expression in strains with compartment markers

FC2D02 and FC1E02 CSI strains expressing mVenus-tagged Golgi markers were acquired through the Chlamydomonas resource centre. As a control, the FC1H12 strain with a fluorescent signal corresponding to the mVenus-tagged Sec61 endoplasmic reticulum resident protein was also acquired (Mackinder *et al.*, 2017). The signal shape associated to the expression of such proteins are shown in the Figure 38.



**Figure 38 : Confocal imaging of A. FC1E02, B. FC2D02 and C. FC1H12 strains from the CSI (Mackinder *et al.*, 2017).**

All the strains have the same cc4533 genetic background and are expressing a Venus-tagged *C. reinhardtii* endogenous protein. On the left, the signal corresponding to the fusion protein is represented and the merge of the fusion protein signal and chloroplast fluorescence is shown on the right. A. FC1E02 is expressing the Golgi predicted Cre10.g439500 protein giving a blurred punctated signal. B. FC2D02 is expressing the Golgi predicted Cre16.g661700 protein giving a relatively strong punctated signal. C. FC1H12 is expressing the ER resident Sec61 translocon encoded by the Cre03.g171350 gene and giving the characteristic reticulated shape surrounding the nucleus.

The respective signals from the two fluorescent reporters (mVenus, mCherry) will be analysed to assess whether there is co-regionalisation in one of the two subcellular compartments.

After transformation by electroporation, positive colonies selected on TAP medium supplemented with the antibiotic corresponding to the inserted resistance cassette were then screened using a fluorescence plate reader. Colonies exhibiting the highest expression levels based on fluorescence intensity relative to OD<sub>720nm</sub>, reflecting the cell density, were then observed using an SP5 confocal microscope.

Despite all the efforts and hundreds of colonies screened, no convincing results were obtained as little or no specific signal corresponding to mCherry was detected by microscopy. This result also called into question the specificity of the signal detected during plate reader screening and lead to a PCR screening on gDNA that revealed a transformation efficiency ranging from 21,5 to 46,2% in these strains whose genetic background is cc4533 (Table 21).

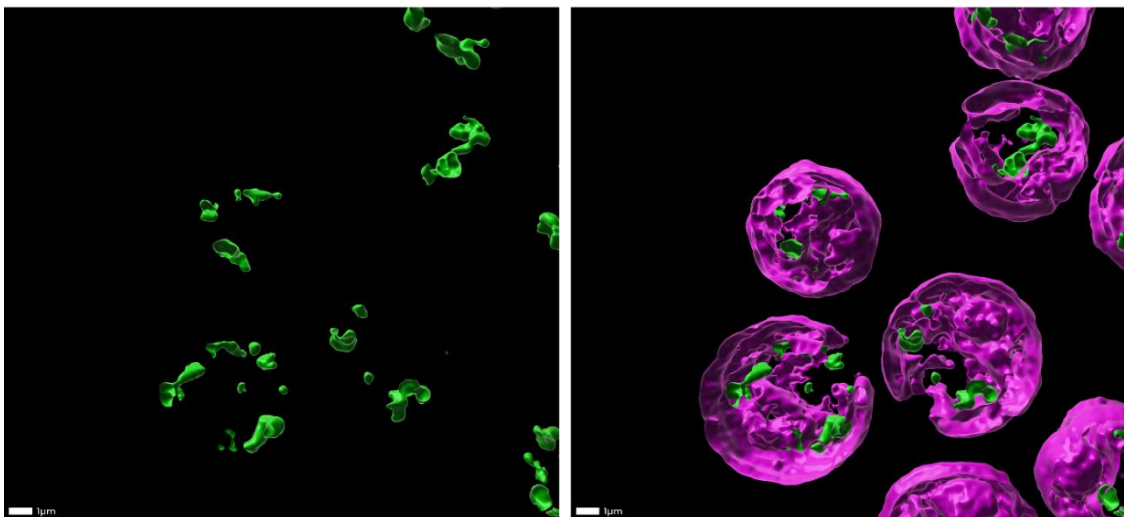
**Table 21 : Example of a PCR screening of some of the colonies transformed with the pLM006 based plasmids in the CSI strains.**

CTS domain	Screened colonies	gDNA screening positives	%
<b>FC2D02</b>			
Cre16L	93	20	21,5
Cre16S	93	24	25,8
Cre10L	93	23	24,7
Cre10S	93	20	21,5
<b>FC1H12</b>			
Cre16L	93	42	45,1
Cre16S	93	30	32,2
Cre10L	93	43	46,2
Cre10S	93	35	37,6

The weak expression may be due to epigenetic mechanisms that limit the expression of nuclear transgenes in *C. reinhardtii* strains such as the strains that has been used whose genetic background is cc4533 (Neupert *et al.*, 2020).

Therefore, the plasmid encoding the fusion protein N-terXTBmClover was introduced *via* electroporation in the UVM4 strain (Neupert *et al.*, 2009, 2020) allowing to reach transformation efficiency up to 90%. The resulting colonies were screened using a fluorescence plate reader as described previously in order to select only the best fluorescence expressors. Around ten colonies with a specific fluorescence signal at least three times greater than the background signal of the wild-type parent strain have been selected. To check whether the fluorescent protein was maintained in the cell and not secreted, the culture medium and the cell pellet have been separated by centrifugation and submitted to the fluorescence screening. The selected colonies were then observed using a SP8 confocal microscope.

Three-dimensional modelling and deconvolution were carried out on the collected images *via* the Imaris software (Figure 39). On the different analysed clones, the mClover fluorescence was associated to several punctuated signals that could correspond to Golgi cisternae or secretion vesicles (NEUMANN *et al.*, 2003; Gao *et al.*, 2012; Vilarrasa-Blasi *et al.*, 2021) It is also clear that there are several points forming the signal, which is consistent with the various demonstrations of the presence of multiple Golgi stacks within the same cell (Giddings, 2003). However, co-localisation with a marker of the Golgi apparatus remains essential in order to draw any further conclusions.



**Figure 39 : Example of 3D confocal imaging of the endogenous XT B N-terminal region in fusion with the mClover fluorescent protein in UVM4 (Scale bar : 1µm).**

Acquisitions have been processed with IMARIS. On the left, the signal corresponding to the fusion protein is represented by a green lut. On the right, the merge of the fusion protein signal and chloroplast fluorescence shown as a magenta lut. Confocal imaging have been done on several independant colonies with the same result.

### III. Discussion and Conclusion

#### A. Choice of the CTS domain

The first predictions revealed a transmembrane domain in the N-terminal part, which could indicate the presence of a CTS region in the endogenous Cre10.g439500 and Cre16.g661700 proteins predicted to be golgian. In two different studies (Mackinder *et al.*, 2017; Vilarrasa-Blasi *et al.*, 2021) a punctate form of the signal observed by confocal microscopy could suggest a Golgi-like signal or secretion vesicles. However, no signal could be observed by confocal microscopy. On the other hand, the problems encountered during prediction call into question the selection of candidate genes.

As for the exogenic CTS domains, no signal could be observed, as was the case for the putative endogenous CTS domains. However, as these domains from plants and rat have been used successfully in other studies (Loos and Steinkellner, 2014), it is assumed here that maybe the combination of non-optimised plasmids, an unsuitable genetic background and the choice of fluorescent protein in the CSI strains led to poor results.

It can be also assumed that the size of the ultrastructures of interest could also be a parameter that affects the quality of the signal obtained.

Finally, as all the results point to a different strategy, the N-terminal part of endogenous xylosyltransferase B from *C. reinhardtii* was expressed in fusion to mClover in UVM4. The specific fluorescent signal associated to vesicles could correspond to a localisation within the Golgi stacks. Since the sequence used is derived from a glycosyltransferase involved in maturation steps, results suggest that this sequence could encode a CTS domain. However, proof of its Golgi localisation and sublocalisation has yet to be provided.

## B. Plasmid construction

Two types of plasmid constructs were cloned on the basis of pLM006 and pBR09 (Mackinder *et al.*, 2016; Rasala *et al.*, 2013), but little or no expression was obtained. These results can be attributed to various factors.

It may be due to the fact that there is no intron included in the sequence. Indeed, it has since been demonstrated that the insertion of introns in the transgene sequence of interest allows better recruitment of the transcriptional machinery by mimicking the natural intron/exon distribution of the microalga (Baier, Wichmann, *et al.*, 2018; Jaeger *et al.*, 2019; Baier *et al.*, 2020).

It could also be due to the fluorescent protein itself. The photophysical parameters of fluorescent proteins are also parameters to be taken into account (Mukherjee and Jimenez, 2022). In addition to the excitation and emission wavelengths that need to be compared when choosing a fluorescent protein, the brightness and quantum yield have proved to be just as important. As previously indicated, the quantum yield corresponds to the following ratio:  $QY = \text{photons emitted} / \text{photons absorbed}$ . Intrinsic parameters of mCherry, such as the brightness (15,84) and quantum yield (0,22), revealed that mCherry is rather weak. For comparison, mClover ( $\lambda_{Ex}=505$  nm,  $\lambda_{Em}=515$  nm) used in fusion with the N-terminal sequence of XTB has a brightness 5,3 times greater (84,36) than that of mCherry and a quantum yield that is 3,5 times greater (0,76) (Lam *et al.*, 2012; Shaner *et al.*, 2004).

Finally, as suggested previously, all the above assumptions combined with a strain with a low level of expression might give less chance of detecting signal with a plate reader or a confocal microscope thus justifying the result described in this chapter.

Thus, based on these results, new constructs will be designed with introns distributed throughout the sequence to take account of the latest developments in molecular biology. The various sequences encoding CTS domains or predicted to encode transmembrane helices could be fused to the mScarlet fluorescent protein (Bindels *et al.*, 2017). These constructs could be introduced into a strain known for its ability to express transgenes without silencing, such as UVM4 (Neupert *et al.*, 2009, 2020) or the strain described in Dementyeva *et al.*, 2021. In addition, constructs containing the full sequences of compartment markers such as the ADP-ribosylation factor (ARF) proteins (Chavrier and Goud, 1999) for the Golgi apparatus and Sec61

(Greenfield and High, 1999) or BiP (Haas, 1994) for the endoplasmic reticulum should be synthesised in the hope of observing co-regionalisation of the different fluorescent reporters.

### C. Expression in strains with compartment markers

Initially, a low transformation efficiency was observed (46,2% maximum) when using CSI strains with the cc4533 genetic background. What's more, of the 45% or so that were positive, the expression level was zero. Although positive results have been obtained in this strain for expressing proteins fused to fluorescent tag (Mackinder *et al.*, 2017), we could not exclude that silencing phenomena could prevent the expression of transgenes in our study. As the results obtained with the UVM4 strain were more promising than those obtained with the cc4533 genetic background, it would be preferable to continue the study with the latter as suggested previously.

Despite the relatively punctate signal observed while expressing the N-terminal part of the sequence coding for the XTB in fusion with the mClover fluorescent protein, confocal microscopy of colocalisation with a protein known to reside in the Golgi membranes, such as ARF2, could confirm or refute its correct localisation. An additional experiment involving the use of brefeldin A (Singleton *et al.*, 1958) that causes dismantling of the Golgi stacks and accumulation of secretory proteins in the endoplasmic reticulum (Fujiwara *et al.*, 1988), might help in confirming the Golgi localisation. The use of such a chemical has already proved its worth in other organisms such as Tobacco (Ritzenthaler *et al.*, 2002)

Confocal microscopy remains a limiting technique for accurate localisation. Indeed, to determine precisely in which Golgi compartment(s) a protein is found, electron microscopy would have to be used. Fortunately, this technique is applicable to *C. reinhardtii* (Ropitiaux *et al.*, 2024) and the presence of distinct fluorescent proteins as well as an affinity tag in the C-terminal position of the plasmids could allow their immunolocalization using beads coated with a tag-specific antibody.

To sum up, these preliminary results allowed to develop a suitable strategy for expressing fusion proteins. A strain change combined with an intron-optimised plasmid containing a fluorescent protein whose intrinsic properties are optimal for this localisation work has already been undertaken. This work needs to be continued in this

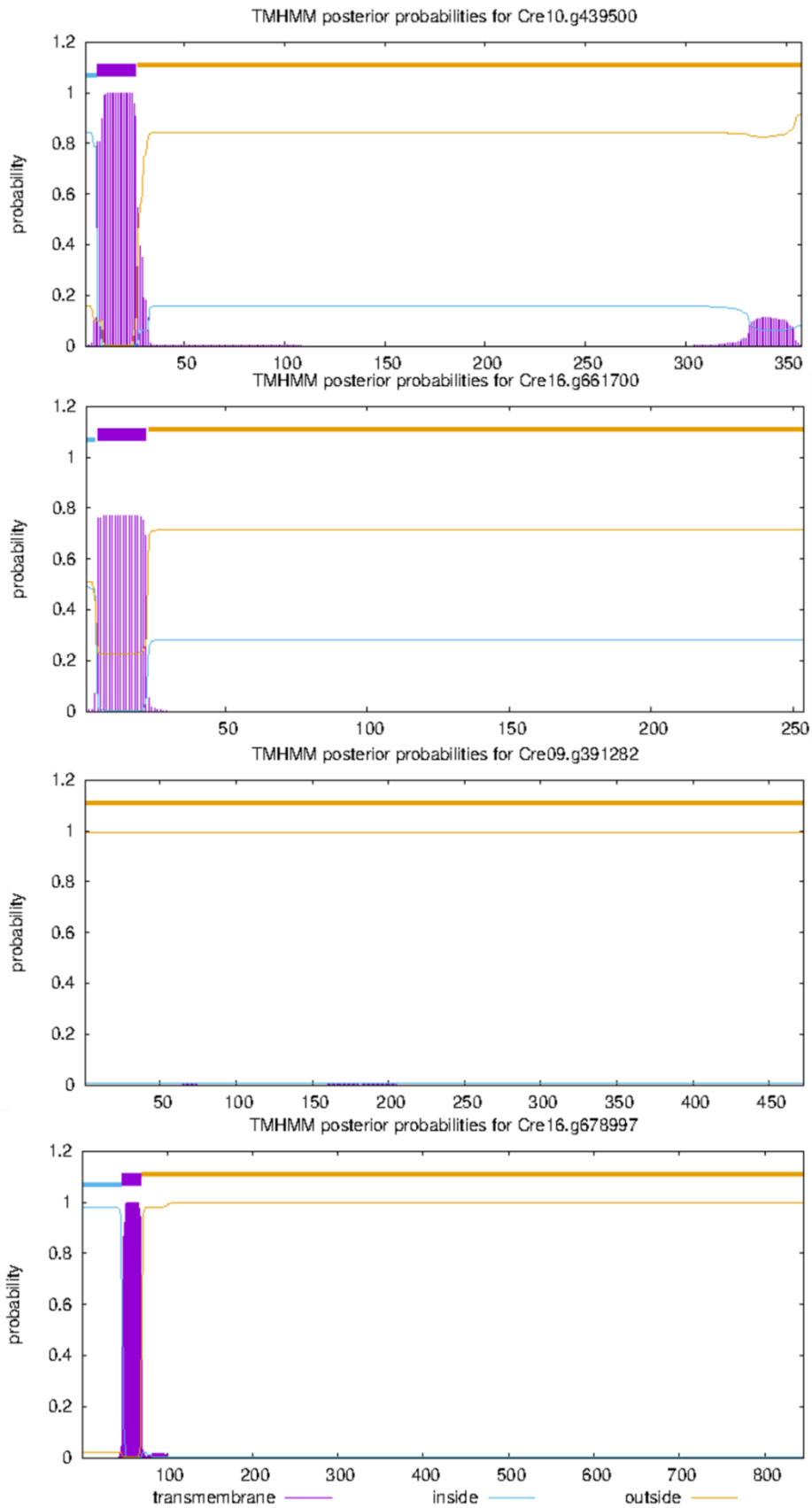
direction and supplemented by in-depth studies of sublocalisation within the cell using electron microscopy.

The output of this work should enable the identification of CTS regions able to anchor heterologous glycosylation enzymes involved in the golgian maturation stages in the near future, with the aim of reproducing in *C. reinhardtii* the distribution found in humans.

## Chapter 3

### Supplemental data





**Figure 40 : Transmembrane domain prediction of Cre10.g439500, Cre16.g661700, Cre09.g391282 (XTA) and Cre16.g678997 (XTB) using TMHMM-2.0 (Krogh et al., 2001).**

Each amino acid sequence has been uploaded as a FASTA file. Prediction have been realised using the extensive output format with graphics.



# Conclusion & Perspectives



## I. Conclusion

For the past ten years, studies aimed at improving our knowledge of the green microalga *Chlamydomonas reinhardtii* protein biosynthesis and secretion and its *N*-glycosylation processing. This is of specific interest when thinking about using it as an alternative system for producing biopharmaceutical glycoproteins.

Exploiting the potential of *C. reinhardtii* for this type of application remains challenging for two main reasons: the need to express nuclear transgenes at high yields compatible with the requirements of industrial profitability (Hellwig *et al.*, 2004; Langer and Rader, 2015) and the need to remodel the *N*-glycosylation pathway in order to produce proteins with humanized *N*-glycans that can be used in human therapy (Mathieu-Rivet *et al.*, 2014; Barolo *et al.*, 2020; Goshtasbi *et al.*, 2023).

As a result, until now, *C. reinhardtii* has mainly been used for the expression of non-glycosylated recombinant proteins, exploiting the large volume of the chloroplast and its genome, which is easier to transform and allow higher production yield (Taunt *et al.*, 2017). However, proteins expressed via chloroplast DNA transformation do not undergo the *N*-glycosylation step that is crucial for their activity, immunogenicity and lifespan (Mathieu-Rivet *et al.*, 2014). In the context of the production of biomedicines, most of which being glycoproteins, expression *via* this organelle is therefore unthinkable. In contrast, protein nuclear expression provides us with these post-translational modification steps. Nevertheless, it was necessary to first characterise the *N*-glycosylation pathway in *Chlamydomonas reinhardtii*.

Mathieu-Rivet *et al.* set out to meet this demand in 2013 by combining bioinformatics studies with structural analysis approaches, in order to propose a first version of the *N*-glycosylation pathway in *C. reinhardtii*. The characterisation of this metabolic pathway has evolved in tandem with the emergence of new strains, such as UVM4 and UVM11, capable of satisfactorily expressing transgenes, as well as molecular biology tools aimed at improving expression levels by a factor of 5 (10 $\mu$ M vs 50 $\mu$ M in culture medium) compared with a wild-type strain (Lauersen, Berger, *et al.*, 2013). In the light of advances in transgene expression, the MoClo toolbox (Crozet *et al.*, 2018) and the Intronserter online cloning tool (Baier, Wichmann, *et al.*, 2018; Jaeger *et al.*, 2019) have also opened up the field of possibilities by enabling the synthesis of optimised plasmids in record time.

The SweetBioPharm project, led by the two laboratories that have pioneered the study of the *N*-glycosylation pathway in this green microalga, is therefore a continuation of previous work and seeks to gain a better understanding of this metabolic pathway with a view to remodelling it, putting fundamental studies at the service of future applications.

## A. Chapter I

In 2009, Eichler Stahlberg and collaborators produced secreted EPO in *cw15arg<sup>-</sup> Chlamydomonas reinhardtii*'s culture medium using ARS2 secretion signal peptide with an observed MW of 33 kDa which is close to the 34 kDa in human. Those results allow the hypothesis of a glycosylation of this biopharmaceutical glycoprotein. Indeed, this hormone contains 3 *N*-glycosylation sites (Asn24, Asn38, Asn83) and 1 *O*-glycosylation site (Ser126). However, in this study the expression yields, up to 100 µg/L of secreted recombinant EPO, were too low to collect enough material for further analysis (Eichler-Stahlberg et al, 2009).

During this project, optimised plasmids have been successfully created enabling a better expression of a secreted human EPO in fusion with the mClover fluorescent protein in glycoengineered strains. Colonies have been screened by fluorescence and production yields have been assayed by ELISA on the best expressing clones. Affinity chromatography purification of the recombinant EPO protein has been performed. Biochemical analysis such as western blot and SDS-PAGE allowed the confirmation of the molecular weight (~60 kDa) corresponding to the potentially glycosylated EPO in fusion with the mClover fluorescent protein (33 kDa + 27 kDa). Furthermore, the specific binding of antibody directed against β(1,2)-xylose and α(1,3)-fucose on the purified EPO has been shown by immunoblotting confirming the presence of *N*-glycans in the purified recombinant EPO and the addition of β(1,2)-xylose and α(1,3)-fucose residues on the *N*-glycan's structures. These residues are reputed to be transferred during maturation stages of *C. reinhardtii* *N*-glycosylation pathway (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020) thus testifying of the presence of mature *N*-glycans on this glycoprotein. To the best of our knowledge, our work represents the first-time characterization of the glycosylation of a recombinant protein produced in *Chlamydomonas*. This has been performed by glycoproteomic analysis confirming that *C. reinhardtii* is able to glycosylate human erythropoietin in its 3 *N*-glycosites as suggested in 2009 and that all glycosites were occupied with *N*-glycans structures

accordingly to the *N*-glycosylation profile of the strain (Eichler-Stahlberg *et al.*, 2009). In contrast, no *O*-glycosylation have been detected in the recombinant hEPO produced in *C. reinhardtii*.

## B. Chapter II

The green microalga *C. reinhardtii* synthesizes non canonical linear oligomannosidic *N*-glycans ranging from Man<sub>5</sub>GlcNAc<sub>2</sub> to Man<sub>3</sub>GlcNAc<sub>2</sub> and bearing specific residues such as  $\beta$ (1,2),  $\beta$ (1,4)-xylose and  $\alpha$ (1,3)-fucose that are absent in mammalian glycoproteins (Lucas *et al.*, 2020; Oltmanns *et al.*, 2019; Schulze *et al.*, 2018).

Crossings of insertional mutants for the genes encoding the characterised xylosyltransferases (XTA, XTB) and fucosyltransferase (FucT) lead to a glycoengineered strain with a higher amount of Man<sub>3</sub>GlcNAc<sub>2</sub> *N*-glycan that represent the basic structure for further glycoengineering. However, analysis of this strain suggests the presence of xylosylated *N*-glycans that probably results from the activity of additional xylosyltransferases involved in the xylosylation process in *C. reinhardtii*.

Three additional putative candidates sharing a highly conserved C-terminal amino acid sequence with XTA and XTB have been identified. It has been decided to inactivate by CRISPR/Cas9 the gene encoding the putative XTs in the triple mutant IM<sub>XTAx</sub>IM<sub>XTBx</sub>IM<sub>FucT</sub> using ribonucleoproteins (RNP). Based on the output obtained from the prediction webtools, two targets have been selected. Mutants have been generated with a method adapted from Kelterborn *et al.*, 2022. Despite the various attempts carried out, it was not possible to generate a mutant for the putative XT Cre10.g458950 in the IM<sub>XTAx</sub>IM<sub>XTBx</sub>IM<sub>FucT</sub> strain. However, mutants have been successfully generated for the two other putative XTs. After confirmation by sequencing that a STOP codon has been inserted in the reading frame of each mutant, *N*-glycans from both intracellular and secreted proteins from glycosylation mutants have been analyzed. The glycoproteomic and glycomic results are consistent and do not show a decrease or absence of xylosylated residues in the *N*-glycosylation profile compared to the mother strain. Nevertheless, it might be possible that those predicted glycoenzymes, forming a multigenic family with redundant activity, are not involved in the *N*-glycosylation pathway but rather in the *O*-glycosylation which remains little studied to date or in xylosylation of other polymers. It could also be suggested that the

Cre10.g458950 gene we did not manage to inactivate is encoding for a putative xylosyltransferase responsible for the transfer of the residual xylose.

In parallel, the candidate gene Cre08.g361250 has been identified as potentially involved in photosynthesis (Wakao *et al.*, 2021). However, preliminary tests carried out on the mutant obtained for this gene in the genetic background of the triple mutant did not allow us to conclude whether it is involved in photosynthesis. A separate project could be devoted to exploring this hypothesis.

### C. Chapter III

*N*-glycans in *C. reinhardtii* do not contain terminal *N*-acetylglucosamine residues which is consistent with bioinformatic analyses that revealed a lack of the gene encoding *N*-acetylglucosaminyltransferase I (GnTI), a Golgi enzyme that is necessary for the subsequent construction of mammalian biantennary *N*-glycans (Mathieu-Rivet *et al.*, 2013). These important differences in the glycoenzyme repertoire raise the question of the engineering of the *N*-glycosylation pathway to make it suitable for the expression of recombinant glycoproteins compatible with human therapy. Such glycoengineering relies on the use of 'Cytoplasm, Transmembrane and Stem' (CTS) regions located in the *N*-terminal part of the glycoenzymes and that are responsible for their addressing and distribution in the Golgi cisternae. However, the targeting mechanisms of enzymes in the Golgi apparatus remain unexplored in *C. reinhardtii*. It was therefore necessary to first unravel this mechanism.

In order to understand the targeting of glycosyltransferases (GTs) in *C. reinhardtii*, different sequences encoding distinct CTS regions from either endogenous type II transmembrane proteins such as the characterised XTB or heterologous Golgi-located glycosyltransferases from *A. thaliana* or *R. norvegicus* (Loos & Steinkellner, 2014) have been fused with the sequence coding for Clover/mScarlet fluorescent proteins. Strains from the Chlamydomonas Spatial Interactome collection (MacKinder *et al.*, 2017) expressing ER and Golgi located YFP/RFP tagged marker proteins have been transformed with the aim of colocalising our fusion proteins within the secretory pathway compartments. However, we could never observe a significant fluorescent signal with our constructions. Hence, plasmids have been optimized and transformed into UVM4 strain (Neupert *et al.*, 2009). A screening based on fluorescence detection has been done to select several independent colonies. Then, the fluorescent markers are detected on each colony by confocal microscopy to evaluate the capacity of each

CTS region to target proteins in the *C. reinhardtii* Golgi apparatus. Imaging of independent colonies expressing the endogenous YFP tagged putative CTS from XT<sub>B</sub> revealed a punctuated signal which seems to be characteristic of a Golgi localisation. Nevertheless, this needs to be confirmed.

## II. Perspectives

Over these 4 years of research, it became clear that recombinant erythropoietin was an effective reporter protein for *N*-glycosylation in a microalgal strain, even when this metabolic pathway was modified, as reported in the first chapter of this manuscript. The plasmids developed and the techniques adapted to the purification and analysis of EPO have since proved their worth in the several projects. However, EPO remains a small molecule with only 3 *N*-glycosites. It would be now interesting to carry out similar work on more complex glycoproteins such as therapeutic antibodies. On another hand, we have also demonstrated that the IM<sub>XTA<sub>x</sub></sub>IM<sub>XTB<sub>x</sub></sub>IM<sub>FucT</sub> strain is capable of transferring paucimannosidic *N*-glycans to a recombinant glycoprotein. This type of glycosylation is sometimes sufficient to ensure the activity of a pharmaceutical glycoprotein. The prospect of expressing biologics with paucimannosidic type *N*-glycans can therefore also be envisaged.

The work and efforts carried out to gain a better understanding of the xylosylation process in this green microalga have not led to any solid conclusions but have justified the emergence of new projects in the two partner laboratories of the SweetBioPharm project. The mutants successfully generated by genome editing in Pr. Michael HIPPLER's laboratory described in the second chapter of this thesis should be the subject of in-depth analyses. The outlook of such analyses would be the crossing of the obtained mutants in order to look for a stronger phenotype affecting *N*-glycosylation. The candidate genes mutated during the SweetBioPharm project encode proteins that have been predicted to be involved in glycosylation. If no phenotype impacting the *N*-glycosylation is observed, it would be of interest to further investigate the mutants with the aim of finding a phenotype affecting the *O*-glycosylation that remains little studied so far or to analyse other glycomolecules. Moreover, higher amount of Man<sub>3</sub>GlcNAc<sub>2</sub> *N*-glycan has been found in those mutants. This basic structure could be used as a chassis for further glycoengineering such as the heterologous expression of glycosylation enzymes such as TbGT11 GnT-like enzyme from *T. brucei* that has a substrate specificity with the Man<sub>3</sub> residue



(Damerow et al, 2014). If this knock-in strategy proves effective, it will make it possible to compensate for the GnTI enzyme deficiency, thereby humanising a *C. reinhardtii* cell line.

Although the actions of XTA and XTB are known in *C. reinhardtii*, these enzymes have never been fully characterised from a biochemical point of view. It would be therefore interesting to expand once again the knowledge of *N*-glycosylation in *C. reinhardtii*, through a biochemical characterisation of the enzymes involved in the stages of Golgian maturation. As well as the biochemical characterisation of xylosyltransferases A and B is concerned, it would also be interesting to discover their localisation within subcellular structures thus initiating a mapping of the *N*-glycosylation enzymes.

The work detailed in the third chapter of this manuscript also served as preliminary work to map the players in the *N*-glycosylation pathway in *C. reinhardtii*. A detailed mapping coupled with characterisation of the various players in *N*-glycosylation would provide a better understanding and enable us to envisage more serenely customised glycoengineering as suggested previously with the complementation and achieve the ultimate goal of creating a humanised strain of *C. reinhardtii* for the production of biopharmaceuticals.

Finally, the development of a glycoengineered strain with expression levels far exceeding those achieved in this study would be a goal to achieve in order to see the emergence of a strain capable of competing with current production models. Work on optimising culture conditions will also be required to make progress in this direction.



# General discussion



This chapter complements the discussions proposed in the various results chapters. Based on the results obtained during this thesis work, the following paragraphs therefore raise questions about the future of *Chlamydomonas reinhardtii* as a model for the production of biopharmaceutical glycoproteins. The  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain, used several times in the work described in this manuscript, will be the common thread running through the different paragraphs hereafter.

### I. Complementation of *C. reinhardtii* N-glycosylation pathway.

Overall, the work in this thesis has confirmed once again that the repertoire of enzymes involved in the N-glycosylation pathway in *C. reinhardtii* is very different from that of other eukaryotes and in particular that of humans. This synthesis pathway in the green microalga is independent of the GnTI required for the subsequent formation of human-type complex glycans (Mathieu-Rivet *et al.*, 2013, 2014; Vanier *et al.*, 2017). In a context of bioproduction as it is described in introduction of this manuscript, it would appear that, as it stands, *C. reinhardtii* cannot be used for the production of therapeutic glycoproteins for human use. However, it has been demonstrated in Chapter 1 of this manuscript that when N-glycosylation is modified, these modifications are reflected in the glycan profile of the recombinant proteins expressed. Furthermore, if the study of addressing mechanisms in *C. reinhardtii*, as described in chapter 3 of this manuscript, were to reveal CTS domains capable of taking a glycoprotein into the various Golgi cisternae, it would then be easier to carry out glycoengineering. The  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain used in Chapters 1 and 2 of this manuscript has a higher proportion of paucimannosidic structures of the  $Man_3GlcNAc_2$  type and a reduced xylosylation and fucosylation capacity (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020). This would make it an interesting strain to use as a basis for glycoengineering.

A study of N-glycosylation in the parasite *Trypanosoma brucei* revealed two genes called TbGT11 and TbGT15 encoding glycosyltransferases with GnTI-like and a GnTII-like activities respectively (Damerow *et al.*, 2014, 2016). Both enzymes specific substrate is a  $Man_3GlcNAc_2$  such as that the major N-glycan found on the endogenous glycoproteins of the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  mutant. Provided that the results of the studies in Chapter 3 aiming at finding CTS domains with the capacity to anchor a protein into the *C. reinhardtii* Golgi apparatus prove successful, such GnT-like enzymes could be addressed to the Golgi cisternae.

The output of such engineering would lead to the graft of GlcNAc onto the paucimannosidic  $\text{Man}_3\text{GlcNAc}_2$  N-glycans found in the  $\text{IM}_{\text{XTAx}}\text{IM}_{\text{XTBx}}\text{IM}_{\text{FucT}}$  and therefore the synthesis of  $\text{GlcNAc}_1\text{Man}_3\text{GlcNAc}_2$  and  $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$  N-glycans paving the way for further humanisation steps, for instance the expression of a  $\beta(1,4)$ -galactosyltransferase as it has been done in plants (Strasser *et al.*, 2009a).

A similar experiment aiming at humanising *C. reinhardtii* N-glycosylation pathway has been carried out in 2017 via the heterologous expression of GnTI from *Phaeodactylum tricornutum* (PtGnTI) and *A. thaliana* (Vanier *et al.*, 2017). This study did not result in GlcNAc residues in terminal position of N-glycans but enabled to reassess the N-glycosylation pathway in *C. reinhardtii* and demonstrate that PtGnTI had no chance of working on a non-cannonical  $\text{Man}_5\text{GlcNAc}_2$ . Two additional hypotheses can be put forward: 1) the donor substrate was not present in the Golgi apparatus and 2) Golgian sublocalisation is not favourable to the activity of these enzymes.

In the light of the work carried out during this thesis project, it may therefore be possible to avoid obstacles encountered during the attempt of humanisation of the N-glycosylation pathway in *C. reinhardtii* in 2017. The combined use of the  $\text{IM}_{\text{XTAx}}\text{IM}_{\text{XTBx}}\text{IM}_{\text{FucT}}$  strain and the GnT-like enzymes from *T. brucei* would ensure compatibility between enzyme and acceptor substrate. The use of the endogenous putative CTS domain of XTB or any other CTS domain whose efficacy has been demonstrated in *C. reinhardtii* would ensure anchoring of the GnT-like in the Golgi membranes.

Another outcome could be that such remodelling would change the substrate specificity of the remaining endogenous unknown xylosyltransferases thus producing glycans devoid of immunogenic epitopes. These assumptions are based on the fact that the glycoengineering work is proving effective. However, although such work would aim to ensure the subcellular localisation of N-glycosylation effectors and their acceptor specificity, an activated donor such as UDP-GlcNAc is still required. Despite the presence of UDP-GlcNAc in the cytosolic compartment thus ensuring the synthesis of the chitobiose basis of the LLO on the ER membrane, the presence in the *C. reinhardtii* Golgi apparatus of a specific transporter of this nucleotide sugar has never

been reported so far (Vanier *et al.*, 2017). The heterologous expression of such a transporter is still conceivable, as it has been done in yeast (Baptista *et al.*, 2015).

## II. To humanise or not to humanise

The  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain seems to be a perfect candidate for a glycoengineering in order to humanise its *N*-glycosylation pathway. The operation seems straightforward due to the presence of large quantities of paucimannosidic *N*-glycans of the  $Man_3GlcNAc_2$  type with low levels of xylosylation and no fucosylation (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020). Moreover, such structures have been found on rhEPO as described in Chapter 1. But isn't this type of glycosylation already sufficient for the expression of therapeutic glycoproteins that don't require the presence of complex sialylated glycans with lactosamine? The aim would no longer be to make it a universal model, but a model for targeted applications.

Most glycoproteins of therapeutic interest, such as rhEPO presented in the article published in the Plant Biotechnology Journal, require the presence of sialylated complex glycans to ensure efficacy and prevent clearance (Fukuda *et al.*, 1989). However, the literature shows that certain diseases, such as Pompe disease, characterized by an accumulation of lysosomal glycogen caused by a mutation affecting the gene encoding the acid alpha-glucosidase (GAA) enzyme (Stevens *et al.*, 2022), can be cured with a supplementation of recombinant GAA. Such recombinant proteins harbouring  $Man_3GlcNAc_2$  *N*-glycans are subject to cellular delivery through mannose receptors thus proving their value and effectiveness (Van Patten *et al.*, 2007). Some of these molecules have already been produced in vegetal models, such as Taliglucerase alpha, used in the treatment of Gaucher's disease (Aviezer *et al.*, 2009) or the Human  $\alpha$ -galactosidase used to treat Fabry disease (Hennermann *et al.*, 2019). Following the example of glycoengineered plant lines for the production of glycoproteins harbouring paucimannosidic *N*-glycans (Liebminger *et al.*, 2011; Sariyatun *et al.*, 2021), it therefore seems possible to use the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain as it is. Moreover, residual xylosylation activity reported in this strain would not be a problem as in the case of de la Taliglucerase alpha, where these residues don't hinder its cellular uptake by macrophages and its stability (Tekoah *et al.*, 2013; Strasser, 2023).

### III. *C. reinhardtii* production yield

As explained in the preamble of this chapter, the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain is the main thread of this discussion, since it allows so much conjecture. However, whatever glycoengineering basis or model for the production of glycoproteins with paucimannosidic *N*-glycans it may be, this strain has relatively low production yields. A concentration of 0,02 ng/mL of rhEPO in the culture medium has been reported in the first chapter of results. This is consistent with the genetic background of this strain, which is not dedicated to the overexpression of recombinant proteins. However, the UVM4 strain, which has been used in Chapter 3, showed a better capacity for transgene expression with yields up to 15 mg/mL of a modified Venus fluorescent protein thus reaching the minimum of 10 mg/mL considered as promising for commercial purposes (Hellwig *et al.*, 2004; Neupert *et al.*, 2009; Ramos-Martinez *et al.*, 2017; Neupert *et al.*, 2020). Although the variation in yield from one recombinant protein to another means that we must qualify our comments, these values are still encouraging for the overexpression of recombinant proteins in this strain.

Additional analyses undertaken during this thesis showed that the profile of *N*-glycans carried by the total intracellular proteins of strain UVM4 does not appear to differ from that of other so-called wild-type strains. The MALDI-TOF mass spectra appended to this discussion show structures ranging from  $Man_3GlcNAc_2$  to  $Man_5GlcNAc_2$ , sometimes xylosylated, fucosylated and/or methylated, which is consistent with the data in the literature and the analytical results presented in this manuscript (Figure 41 and 42). It would therefore appear that the UV exposure that led to the creation of this strain did not generate any undesirable mutations impacting *N*-glycosylation.

The genome editing method used successfully in Chapter 2 could then be used to:

- either inactivate the gene encoding XTA, XTB and FucT in UVM4 in order to reproduce the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain but with an epigenetic mechanisms that doesn't limit the expression of nuclear transgenes (Neupert *et al.*, 2020, 2009).

- or to inactivate in the  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  strain the gene encoding the sirtuin-type histone deacetylase identified as being involved in silencing phenomena during the study of the UVM4 strain (Neupert *et al.*, 2020).

Nonetheless, this method is limited by the fact that random insertions may occur in the genome as observed during our genome editing experiment. Therefore, these uncontrolled phenomena during the genome editing process by CRISPR/Cas9 could then be an obstacle to obtaining a strain combining both the synthesis of Man<sub>3</sub>GlcNAc<sub>2</sub> *N*-glycans devoid of xylose and fucose residues and a satisfactory level of transgene expression.

On another hand, a knock-in strategy could also be envisaged in UVM4 strain by expressing a heterologous  $\alpha(1,2)$ -mannosidase, such as those found in *Aspergillus saitoi* (Fujita *et al.*, 1997), that would shorten non-canonical Man<sub>5</sub>GlcNAc<sub>2</sub> by trimming the two terminal  $\alpha(1,2)$ -mannose residues. Depending on the sublocalisation of the glycosidase within the secretory pathway, the natural substrate of xylosyltransferase and fucosyltransferase could then be modified, leading to a higher proportion of Man<sub>3</sub>GlcNAc<sub>2</sub>.

In the event of a successful outcome, a strain combining both the synthesis of Man<sub>3</sub>GlcNAc<sub>2</sub> *N*-glycans devoid of xylose and fucose residues and a satisfactory level of transgene expression would be created. However, this would once again bring us back to the dilemma set out in the two sections above. Besides, throughout the various discussions and perspectives, there was no question of a major obstacle to the use of *Chlamydomonas reinhardtii* as a potential alternative production model for human therapeutic glycoproteins. It would therefore seem that the future of *C. reinhardtii* in bioproduction remains to be written.



#### IV. What about O-glycosylation?

As mentioned in the introduction of the manuscript, O-glycosylation, like N-glycosylation, is a post-translational modification of proteins that occurs in the Golgi apparatus. In the case of O-glycosylation, the linkage occurs between a N-acetylgalactosamine (GalNAc) residue and the oxygen (O) of the hydroxyl group of a serine or threonine of the protein sequence. These glycans can also be transferred onto a hydroxyproline (Varki *et al.*, 1999). Their roles are diverse and linked to the great diversity of O-glycan structures found in eukaryotic organisms. The roles attributed to O-glycans in mammals and plants include physical protection by mucin-type O-glycans, involvement in the ABO blood group system and mediation of host/pathogen interactions (Varki *et al.*, 1999; Wandall *et al.*, 2021; Mohideen and Mahal, 2024). Other studies on O-glycosylated proteins have demonstrated roles of this type of glycans in protection against protease activity but also in protein release and structural roles (Thompson and Wakarchuk, 2022). Furthermore, such O-glycans can be found on biologics and ensure similar functions (Zhang *et al.*, 2013).

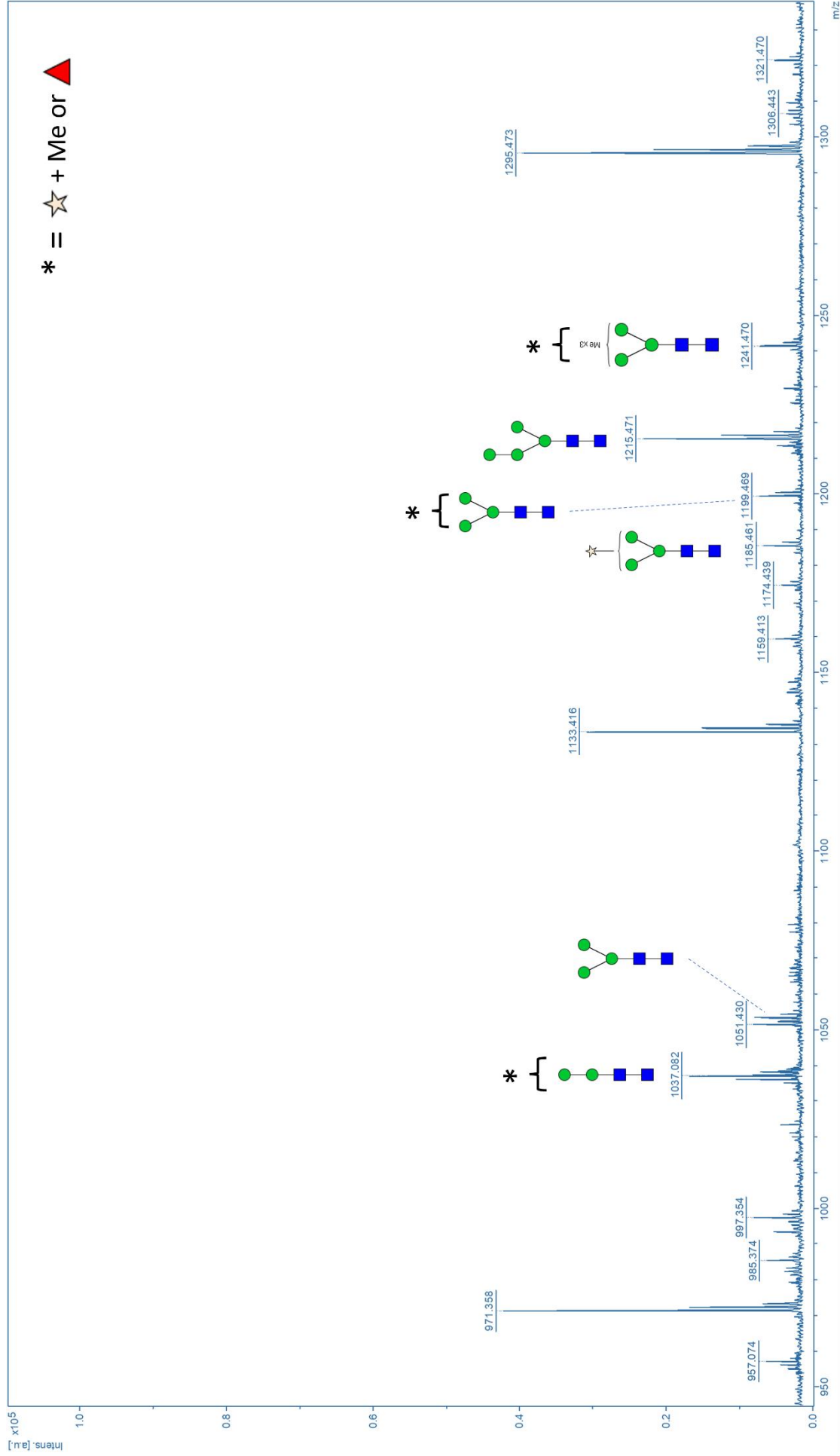
In *Chlamydomonas reinhardtii*, O-glycoproteins are mainly found in cell wall that have been described to be multilayers structures composed of hydroxyproline-rich glycoproteins (HRGPs), which play a protective and structural role (Adair and Apt, 1990; Bollig *et al.*, 2007; Poulhazan *et al.*, 2024). These O-glycans are mainly composed of arabinose and galactose as well as glucose, xylose, and mannose to a lesser extent (Bollig *et al.*, 2007; Poulhazan *et al.*, 2024). A study carried out in 1995 also demonstrated, using O-glycosylation mutants, that O-glycans were involved in sexual recognition in this microalga (Vallon and Wollman, 1995). However, to date, O-glycosylation remains a little-studied area in *C. reinhardtii*. Therefore, each chapter of this manuscript proposes to extend the studies to protein O-glycosylation. As mentioned above, xylose is one of the monosaccharides found in O-glycan-type structures in *C. reinhardtii*. It could therefore be that the genes identified to encode putative xylosyltransferases in Chapter 2 of this manuscript are not involved in N-glycosylation but rather in O-glycosylation. This would be consistent with the fact that no change in the N-glycosylation profile was observed when these genes were inactivated. In addition, if the results show that the proteins encoded by these genes are indeed involved in O-glycosylation processing, mapping work such as the one

proposed in Chapter 3 of this manuscript could also prove useful. Such results would enable the elucidation and a proposal of a O-glycosylation pathway in this microalga.

In addition, two studies have established the hypothesis that *C. reinhardtii* could carry out O-glycosylation of proteins on serine and threonine residues (Poulhazan *et al.*, 2024; Saito *et al.*, 2014). The results of our study carried out on rhEPO described in Chapter 1 of this manuscript run counter to these hypotheses. Indeed, despite the presence of a O-glycosylation site (Ser126) on this cytokine, no O-glycan could be identified on the microalgae-made rhEPO. Moreover, other data on the glycosylation of rhEPO and its engineered forms have not revealed any difference in activity on rhEPO lacking O-glycans (Thompson and Wakarchuk, 2022). It would therefore seem that in this particular case of therapeutic glycoprotein, O-glycosylation is of minor importance. Nevertheless, it has been shown that in other cases, the O-glycosylation of therapeutic glycoproteins is important, particularly in the case of highly O-glycosylated molecules such as etanercept, an antibody directed against the tumour necrosis factor alpha (TNF- $\alpha$ ) used in the treatment of inflammatory arthritis. The presence of sialylated O-glycans improves its lifespan in the human body and its binding to TNF- $\alpha$  (Biel *et al.*, 2022; Thompson and Wakarchuk, 2022). Although O-glycosylation was not previously considered by the health authorities to be an essential criterion for the validation and marketing authorisation of biologics, it is now considered to be important (Zhang *et al.*, 2013). Research supported by health authorities such as the FDA has demonstrated the importance of O-glycosylation and the need to develop O-glycoengineered biopharmaceutical production cell lines (Biel *et al.*; Commissioner, 2023).

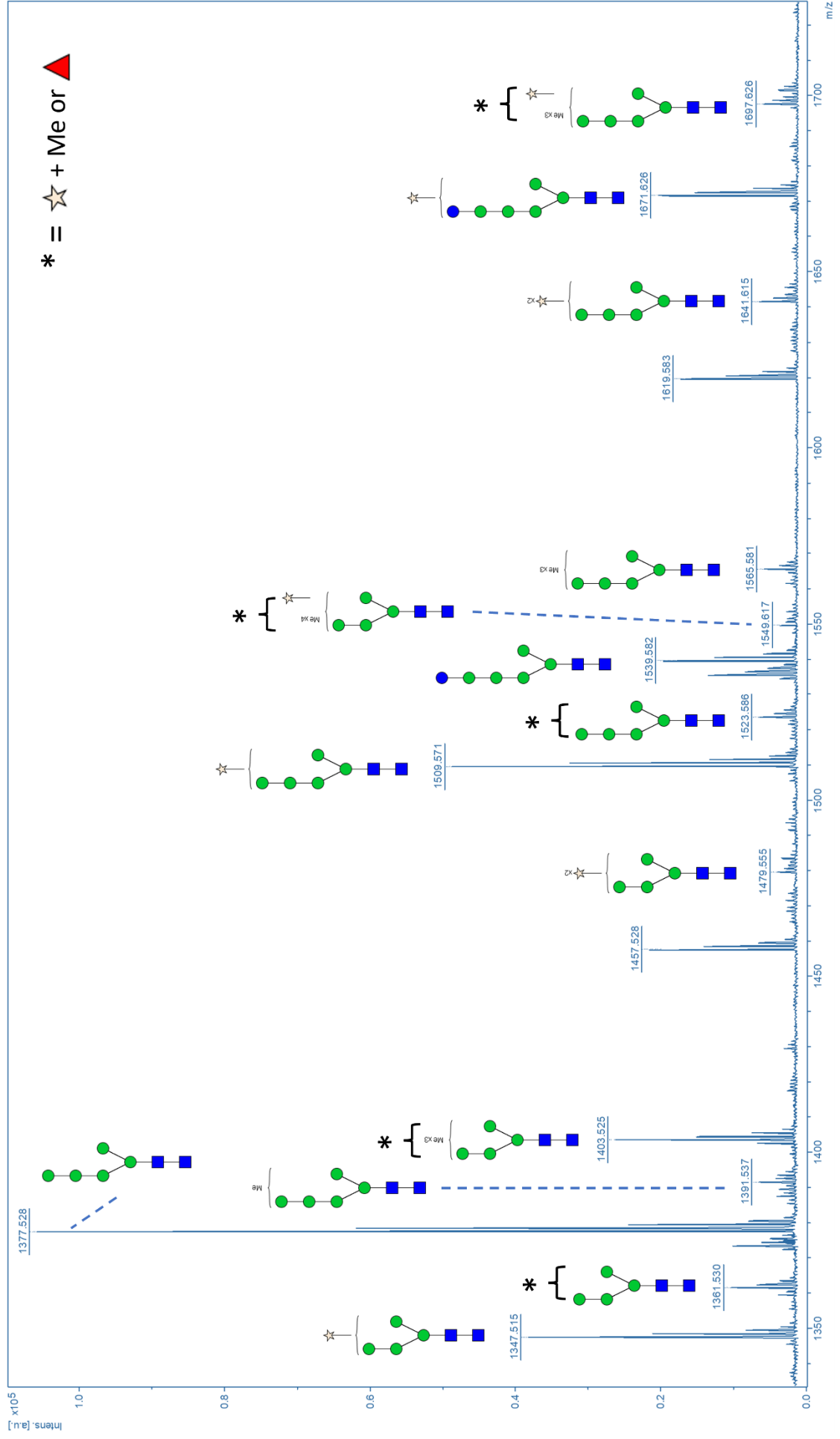
The current SweetBioPharm project did not provide any new information on O-glycosylation in *Chlamydomonas reinhardtii*, but it does point to prospects in this area. The results of the study of rhEPO glycosylation indicate that the only O-glycosylation site was not occupied. The absence of glycans at this single site does not allow any conclusion about the capacity of this microalga of carrying out this post-translational modification. However, the literature dealing with O-glycans in biopharmaceutical indicates the direction research should take and confirm the interest in O-glycosylation if we want to make *Chlamydomonas reinhardtii* a universal model for the production of glycoproteins of pharmaceutical interest. These prospects for elucidating this metabolic pathway could be realised through the expression of highly O-glycosylated proteins

such as etanercept, described earlier, which contains 13 O-glycosylation sites, or the Von Willebrand Factor required for hemostasis, which contains 10 O-glycosites (Biel *et al.*, 2022; S. Ward *et al.*, 2021).



**Figure 41 : MALDI-TOF spectrum of N-glycans isolated from whole cell proteins in the UVM4 strain with m/z ranging from 950 to 1330.**

Proteins have been extracted from a fresh cell pellet and loaded into an acrylamide gel in order to proceed with a *in gel* sample preparation. Reduction and alkylation have been done prior the trypsin/chemotrypsin digestion. N-Glycopeptides have been extracted from gel pieces and N-glycans have been released using a N-glycosidase F (PNGase F). After a 2 aminobenzamide labelling, N-glycans have been analysed using Matrix Assisted Laser Desorption Ionisation – Time Of Flight Mass Spectrometry (MALDI-TOF MS). The N-glycan profile of the UVM4 strain is similar to other wild-type strains.



**Figure 42 : MALDI-TOF spectrum of N-glycans isolated from whole cell proteins in the UVM4 strain with m/z ranging from 1340 to 1720.**

Proteins have been extracted from a fresh cell pellet and loaded into an acrylamide gel in order to proceed with a *in gel* sample preparation. Reduction and alkylation have been done prior the trypsin/chemotrypsin digestion. N-Glycopeptides have been extracted from gel pieces and N-glycans have been released using a N-glycosidase F (PNGase F). After a 2 aminobenzamide labelling, N-glycans have been analysed using Matrix Assisted Laser Desorption Ionisation – Time Of Flight Mass Spectrometry (MALDI-TOF MS). The N-glycan profile of the UVM4 strain is similar to other wild-type strains.



# Résumé étendu en français



## I. Introduction

Le marché des biomédicaments est prédit pour atteindre 389 milliards de dollars en 2024. Parmi les différentes catégories de biomédicaments, les glycoprotéines recombinantes représentent à ce jour environ 70% des biomédicaments commercialisés (Kesik-Brodacka, 2018; O'Flaherty *et al.*, 2020; Sethuraman and Stadheim, 2006; Walsh and Walsh, 2022).

Les glycanes portés par ces protéines recombinantes sont des modifications post-traductionnelles essentielles car elles ont un impact sur les caractéristiques fonctionnelles de ces molécules telles que l'activité, la durée de vie, la stabilité et l'immunogénicité (Ramazi and Zahiri, 2021).

Un système d'expression hétérologue pour la production de protéines recombinantes est donc choisi pour sa capacité à réaliser la glycosylation (Walsh, 2010). Cependant, bien que les modèles choisis soient capables de transférer des *N*-glycanes sur les protéines recombinantes et ainsi assurer leur conformation tridimensionnelle, l'utilisation de ces molécules dans le cadre d'une thérapie à visée humaine est problématique en raison des maturations spécifiques à chaque espèce. En effet, de tels groupements seront considérés comme du « non soi » par l'organisme humain et seront donc reconnus comme immunogènes (Bardor *et al.*, 2005; van Beers and Bardor, 2012; Ghaderi *et al.*, 2010).

Le développement de stratégies d'humanisation ayant pour but de corriger les différences de glycosylation observées entre protéines naturelles et recombinants a donc été nécessaire. Selon le modèle d'expression les différences peuvent être infimes comme dans le cas des cellules CHO. Cependant, des différences plus importantes ont été mises en évidence chez les cellules de levures et de plantes: glycanes hypermannosylés chez la levure, présence de xyloses et fucoses immunogènes et absence de certains résidus (galactoses et les acides sialiques) chez les plantes.

L'humanisation consiste donc en la reconstitution des séquences glycaniques humaines par inactivation de gènes codant pour des enzymes responsables de l'ajout de résidus immunogènes et/ou la complémentation par des gènes codant des enzymes absentes dans ces organismes. C'est ainsi que par l'inactivation de 4 gènes et l'expression de 14 autres, la voie de *N*-glycosylation de la levure a pu être



humanisée (De Pourcq *et al.*, 2010). Ces glycanes dits complexes résultent d'étapes de maturation golgienne impliquant des enzymes appelées *N*-acétylglucosaminyltransférases et mannosidases, qui sont essentielles au transfert de GlcNAc pour former un résidu de type GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>. Cette structure est ensuite décorée de galactose et d'acides sialiques pour conférer à la protéine thérapeutique son activité, comme dans le cas de l'érythropoïétine (EPO).

De nos jours, les glycoprotéines recombinantes à visée thérapeutique humaine sont produites en grande majorité dans des cellules mammifères telles que les cellules CHO (Chinese Hamster Ovary cells) capable d'un rendement maximale de 8 g/L (Al-Majmaie *et al.*, 2021; Walsh and Walsh, 2022). Cependant, l'utilisation de tels systèmes d'expression engendre un coût de production important de par les conditions de culture exigeantes et les étapes de purification nécessaires qui s'expliquent par leur sensibilité aux contaminations virales (Merten, 2002). Ces éléments ont poussé la communauté scientifique à chercher des modèles de production alternatifs.

Ces dernières années ont vu l'émergence des microalgues dans le domaine industriel en tant que systèmes alternatifs (Dehghani *et al.*, 2022). Ces organismes sont déjà valorisés dans de nombreux domaines (Parameswari. and Lakshmi, 2022; Wang *et al.*, 2024).

La microalgue verte *Chlamydomonas reinhardtii* (*C. reinhardtii*) s'est démarquée des autres microalgues comme système d'expression alternatif des produits biopharmaceutiques de par les nombreux outils moléculaires et biochimiques développés par la communauté scientifique pour l'ingénierie des voies métaboliques (Crozet *et al.*, 2018; Gao *et al.*, 2015; Jaeger *et al.*, 2019; Sizova *et al.*, 2013; Zhang *et al.*, 2014). Ainsi, ces données laissent entrevoir la perspective d'une ingénierie de la voie de *N*-glycosylation chez cet organisme unicellulaire en vue d'une production de produits biopharmaceutiques à visée thérapeutique humaine

Deux défis majeurs accompagnent la perspective d'exploiter le potentiel de *C. reinhardtii* pour ce type d'application: L'expression des transgènes nucléaires à des niveaux compatibles avec les exigences de rentabilité industrielle et la nécessité d'humaniser la voie de *N*-glycosylation de façon à produire des glycoprotéines thérapeutiques compatibles avec une utilisation chez l'homme.

À ce jour, les enzymes effectrices de la voie de *N*-glycosylation chez *C. reinhardtii* ont été en partie identifiées. De plus, une stratégie combinant des analyses biochimiques et des analyses par spectrométrie de masse ont permis la caractérisation des *N*-glycanes portés par les protéines solubles et membranaires. Cette microalgue synthétise des *N*-glycanes oligomannosidiques non canoniques allant de Man<sub>5</sub>GlcNAc<sub>2</sub> à Man<sub>3</sub>GlcNAc<sub>2</sub> et portant des résidus spécifiques tels que  $\beta$ (1,2),  $\beta$ (1,4)-xylose et  $\alpha$ (1,3)-fucose qui sont absents dans les glycoprotéines des mammifères et potentiellement immunogènes (Lucas *et al.*, 2020; Mathieu-Rivet *et al.*, 2014; Oltmanns *et al.*, 2020). D'autre part, ces études n'ont pas permis de mettre en évidence la présence de résidus *N*-acétylglucosamine terminaux nécessaires à la construction ultérieure de *N*-glycanes biantennés analogues à ceux des mammifères (Vanier *et al.*, 2017). Ces données permettent d'estimer les modifications à apporter à la voie de *N*-glycosylation chez *C. reinhardtii* afin de produire des glycoprotéines à visée thérapeutique humaine.

Le projet SWEETBIOPHARM s'est inscrit dans ce contexte à travers trois objectifs principaux. Dans un premier temps, l'EPO, une glycoprotéine rapportrice, a été exprimée dans différents mutants de *N*-glycosylation en quantité suffisante pour en caractériser sa *N*-glycosylation. Puis, une élimination par édition génomique des résidus immunogènes que sont les xyloses et fucoses a été entreprise. Enfin, une étude de l'adressage des glycoenzymes de maturation dans l'appareil de Golgi *via* les domaines "Cytoplasm, Transmembrane and Stem" (CTS) a été réalisée dans le but de développer une glycoingénierie sur mesure.

## II. Chapitre I

En 2009, Eichler Stahlberg et ses collaborateurs ont produit de l'érythropoïétine humaine recombinante (rhEPO) dans le milieu de culture dans la souche cw15arg de *Chlamydomonas reinhardtii* en utilisant le peptide signal de sécrétion ARS2 avec une masse moléculaire observée de 33kDa qui est proche des 34kDa chez l'homme. Ces résultats permettent d'émettre l'hypothèse d'une glycosylation de cette glycoprotéine biopharmaceutique. En effet, cette hormone contient 3 sites de *N*-glycosylation (Asn24, Asn38, Asn83) et 1 site de *O*-glycosylation (Ser126). Cependant, dans cette étude, les rendements d'expression étaient trop faibles pour recueillir suffisamment de protéines pour une analyse plus approfondie (Eichler-Stahlberg *et al.*, 2009).

Par conséquent, il a donc été entrepris d'exprimer la rhEPO dans des mutants déficients en activité xylosyltransférase et/ou fucosyltransférase, nommés  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  et  $IM_{XTAx}IM_{XTB}$  de *C. reinhardtii* dans lesquels la capacité de xylosylation et de fucosylation a été affectée. Les résultats obtenus dans ces souches ont été comparés à ceux obtenus dans la souche de type sauvage cc5325 afin d'évaluer l'efficacité des stratégies knock-out développées pour la suppression des épitopes immunogènes des *N*-glycanes chez *C. reinhardtii*. Après purification, des analyses biochimiques combinées à une étude glycoprotéomique ont montré que la rhEPO possède des *N*-glycanes matures, dans lesquels des résidus  $\beta(1,2)$ -xylose et  $\alpha(1,3)$ -fucose peuvent être identifiées, selon les caractéristiques de la souche, sur les 3 *N*-glycosites. Ces *N*-glycanes sont cohérents avec ceux trouvés sur les glycoprotéines endogènes dans les différentes souches utilisées dans ce travail. Par ailleurs, ces analyses ont apporté un éclairage nouveau en mettant en évidence la présence de résidus *O*-méthyl fucose liés au GlcNAc proximal du noyau chitobiose des *N*-glycanes dans les deux souches ayant une capacité de fucosylation conservée.

Ces travaux ont fait l'objet d'une publication en 2024 dans Plant Biotechnology Journal.

### III. Chapitre II

Les études menées dans le but d'élucider les processus de maturation Golgienne chez *C. reinhardtii* ont requis l'utilisation de mutants d'insertion affectant les gènes codant les xylosyltransferase et fucosyltransférase. Ainsi, il a été possible de déterminer que la xylosyltransférase A est responsable de l'ajout de  $\beta(1,2)$ -xylose sur le core du glycane et la xylosyltransférase B est majoritairement responsable de l'ajout de  $\beta(1,4)$ -xylose sur un  $\alpha(1,2)$ -mannose du *N*-glycane de type  $\text{Man}_5\text{GlcNAc}_2$  non canonique retrouvé chez *C. reinhardtii*. Une activité mineure de transfert de  $\beta(1,2)$ -xylose lui est également attribuée. Cependant, en dépit de l'inactivation combinée de ces deux enzymes, une activité de xylosylation résiduelle a été détectée. Il semblerait donc que la xylosylation soit assurée par une famille multigénique avec une redondance d'activité (Lucas *et al.*, 2020; Oltmanns *et al.*, 2020).

Trois candidats putatifs supplémentaires ont été identifiés, partageant une séquence d'acide amines C-terminale hautement conservée avec XTA et XTb. Il a été décidé d'inactiver par CRISPR/Cas9 le gène codant pour les xylosyltransférases putatives dans le triple mutant  $\text{IM}_{\text{XTA}}\text{IM}_{\text{XTb}}\text{IM}_{\text{FucT}}$  en utilisant des ribonucléoprotéines (RNP). Malgré les différentes tentatives effectuées, il n'a pas été possible de générer un mutant dans lequel le gène candidat Cre10.g458950 a été inactivé dans la souche  $\text{IM}_{\text{XTA}}\text{IM}_{\text{XTb}}\text{IM}_{\text{FucT}}$ . Cependant, des mutants pour les deux autres gènes codant des xylosyltransférases putatives ont été générés avec succès. Après confirmation par séquençage qu'un codon STOP a été inséré dans le cadre de lecture de chaque mutant, les glycanes des protéines intracellulaires et sécrétées des mutants de glycosylation ont été analysés.

Les résultats glycoprotéomiques et glycomiques sont cohérents et ne montrent pas de diminution ou d'absence de résidus xylosylés dans le profil de *N*-glycosylation par rapport à la souche mère. Néanmoins, il est possible que ces glycoenzymes prédites ne soient pas impliquées dans la voie de la *N*-glycosylation mais plutôt dans la *O*-glycosylation ou la synthèse d'autres glycomolécules qui restent peu étudiée à ce jour.

Les perspectives de ce travail seraient le croisement des mutants obtenus afin de rechercher un phénotype plus fort affectant la *N*-glycosylation. Il serait également judicieux d'approfondir les recherches sur le gène Cre10.g458950 qui n'a pu être

inactivé. Enfin, d'autres recherches sur les mutants pourraient être effectuées dans le but de trouver un phénotype affectant la O-glycosylation.

En parallèle, le gène candidat Cre08.g361250 a été identifié comme potentiellement impliqué dans la photosynthèse (Wakao *et al.*, 2021). Cependant, les tests préliminaires menés sur le mutant obtenu pour ce gene dans le fond génétique du triple mutant n'ont pas permis de conclure quant à son implication dans la photosynthèse. Un projet annexe pourrait être dédié à l'exploration de cette hypothèse.

#### IV. Chapitre III

Les *N*-glycanes de *C. reinhardtii* ne contiennent pas de résidus *N*-acétylglucosamine terminaux, ce qui est cohérent avec les analyses bio-informatiques qui ont révélé l'absence du gène codant pour la *N*-acétylglucosaminyltransférase I (GnTI), une enzyme de l'appareil de Golgi nécessaire à la construction ultérieure des *N*-glycanes biantennaires des mammifères (Mathieu-Rivet *et al.*, 2013). Ces différences importantes dans le répertoire des glycoenzymes soulèvent la question de l'ingénierie de la voie de *N*-glycosylation pour la rendre adaptée à l'expression de glycoprotéines recombinantes compatibles avec la thérapie humaine. Cette ingénierie repose sur l'utilisation des régions CTS situées dans la partie *N*-terminale des glycoenzymes et qui sont responsables de leur adressage et de leur distribution dans les citernes de l'appareil de Golgi (Rabouille *et al.*, 1995; Schoberer and Strasser, 2011). Cependant, les mécanismes de ciblage des enzymes dans l'appareil de Golgi restent inexplorés chez *C. reinhardtii*. Il était donc nécessaire d'élucider d'abord ces mécanismes.

Afin de comprendre le ciblage des glycosyltransférases (GTs) chez *C. reinhardtii*, nous avons fusionné différentes séquences codant pour des régions distinctes « Cytoplasm, Transmembrane and Stem » (CTS) provenant soit de protéines transmembranaires endogènes de type II telles que la XTB caractérisée, soit de glycosyltransférases hétérologues localisées au niveau du Golgi provenant d'*A. thaliana* ou de *R. norvegicus* (Loos and Steinkellner, 2014) avec la séquence codant pour des protéines fluorescentes de type Clover/mScarlet. Des souches de la collection Chlamydomonas Spatial Interactome (Mackinder *et al.*, 2017) exprimant des protéines marquées YFP/RFP localisées dans le RE et le Golgi ont été transformées dans le but de colocaliser nos protéines de fusion au sein des compartiments de la

voie sécrétoire. Cependant, nous n'avons jamais pu observer un signal fluorescent significatif avec nos constructions. Les plasmides ont donc été optimisés puis utilisés pour transformer la souche UVM4 (Neupert *et al.*, 2009). Un criblage basé sur la détection de fluorescence a été effectué pour sélectionner plusieurs colonies indépendantes. Ensuite, les marqueurs fluorescents ont été détectés sur chaque colonie par microscopie confocale afin d'évaluer la capacité de chaque région CTS à cibler des protéines dans l'appareil de Golgi de *C. reinhardtii*. L'imagerie de colonies indépendantes exprimant le potentiel CTS endogène marqué par YFP de XTB a révélé un signal ponctué qui est caractéristique d'une localisation dans le Golgi. Néanmoins, cette localisation doit être confirmée par des analyses microscopiques plus précises comme la microscopie électronique à transmission.

## V. Conclusion et perspectives

Nous avons, à travers ce projet, démontré que l'érythropoïétine est une protéine rapportrice de la *N*-glycosylation même lorsque cette dernière a été modifiée. Pour autant que nous le sachions, la publication émanant des travaux décrits en Chapitre I de ce manuscrit est la première relatant la caractérisation détaillée de la *N*-glycosylation d'une protéine thérapeutique recombinante exprimée chez *C. reinhardtii*. Il serait judicieux à présent d'envisager de produire des glycoprotéines plus complexes comme les anticorps monoclonaux. Il est également possible d'exprimer des protéines recombinantes qui requièrent pour leur activité, une glycosylation proche de l'une de souches utilisées lors des travaux décrits dans ce manuscrit. En effet, la Taliglucerase alpha est une glycoprotéine pharmaceutique dont l'activité est favorisée par la présence de *N*-glycanes paucimannosidiques tels que ceux retrouvés dans le triple mutant  $IM_{XTA}IM_{XTB}IM_{FucT}$ .

Les deux autres chapitres de résultats constituent des travaux préliminaires ouvrant la voie vers des projets en cours de développement. En effet, les travaux sur l'inactivation des gènes codant des xylosyltransférases putatives n'ont pas permis d'identifier un phénotype impactant la *N*-glycosylation. Cependant, il serait intéressant de réaliser des croisements des mutants obtenus durant le projet SWEETBIOPHARM afin d'obtenir des mutants multiples qui auraient un phénotype plus marqué. Ces travaux permettraient d'élucider les processus de xylosylation qui, vraisemblablement, sont sous l'influence d'une famille multigénique présentant une redondance des activités de transferts des résidus xylose.

A plus long terme, une complémentation de la voie de *N*-glycosylation serait envisageable dans le triple mutant  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  compte tenu de la présence d'un résidu majoritaire  $Man_3GlcNAc_2$ . En effet, la présence de ce motif glycanique combiné à l'absence de résidus immunogènes que sont le xylose et le fucose dans ce fond génétique en font un candidat idéal pour l'expression d'une enzyme ayant une activité de type GnTI et ainsi obtenir une souche dont la *N*-glycosylation a été humanisée.

Toutefois, pour envisager une telle complémentation, il faut s'assurer du bon adressage des glycoenzymes dans l'appareil de Golgi. Le dernier chapitre s'est focalisé sur la compréhension des mécanismes d'adressage *via* les domaines CTS. Les travaux d'optimisation de la stratégie mise en place ainsi que l'identification de domaines CTS endogènes et exogènes ont justifié la poursuite de cette étude dans un objectif de cartographier les acteurs de la *N*-glycosylation chez *C. reinhardtii* afin de mieux la comprendre et ainsi mieux la remodeler. Ces travaux laissent alors entrevoir une potentielle glycoingénierie sur mesure des voies de *N*-glycosylation chez la microalgue *Chlamydomonas reinhardtii* pour la production de glycoprotéines recombinantes thérapeutiques.

L'ensemble de ces travaux laissent présager un avenir pour *Chlamydomonas reinhardtii* dans la bioproduction de demain. Cependant, les rendements de production actuels restent assez faibles et demandent à être augmentés pour pouvoir un jour concurrencer les rendements des modèles de production actuels.

Les travaux nécessaires pour arriver à ce but ultime sont nombreux et ambitieux. Par ailleurs, la *O*-glycosylation reste à ce jour une voie métabolique non négligeable et peu étudiée chez cette microalgue. Si toutefois, les résultats des nombreux projets à venir ne garantissaient pas un avenir biopharmaceutique à ce modèle des plus intéressants, *Chlamydomonas reinhardtii* restera encore pour bien des années un modèle qui justifie un engouement certain pour la compréhension de ses voies métaboliques d'un point de vue fondamental.

## VI. Discussion

Les résultats obtenus au cours de ce travail de thèse permettent dans la partie intitulée « discussion générale » de discuter de l'avenir de *Chlamydomonas reinhardtii* en tant que modèle pour la production de glycoprotéines à intérêt thérapeutique. La souche mutante  $IM_{XTAx}IM_{XTBx}IM_{FucT}$  utilisée dans les travaux décrits dans ce manuscrit représente un bon châssis pour initier les travaux d'humanisation d'une lignée de *C. reinhardtii* permettant la production de protéines thérapeutiques portant des *N*-glycanes « humanisés » adaptés à leur utilisation en santé humaine. La problématique des rendements de production est également évoquée dans un contexte de compétition avec les modèles de production industrielle actuels. Enfin, la perspective d'étendre les recherches à l'étude des processus de *O*-glycosylation des protéines chez *C. reinhardtii* est ensuite amorcée.





# References



- Adair, W.S. and Apt, K.E. (1990) Cell wall regeneration in *Chlamydomonas*: accumulation of mRNAs encoding cell wall hydroxyproline-rich glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.*, **87**, 7355–7359.
- Aebi, M. (2013) N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta BBA - Mol. Cell Res.*, **1833**, 2430–2437.
- Al-Majmaie, R., Kuystermans, D., and Al-Rubeai, M. (2021) Biopharmaceuticals Produced from Cultivated Mammalian Cells. In: *Cell Culture Engineering and Technology: In appreciation to Professor Mohamed Al-Rubeai Cell Engineering (Pörtner, R., ed)*, pp. 3–52. Cham: Springer International Publishing.
- Alvisi, N., van Noort, K., Dwiani, S., Geschiere, N., Sukarta, O., Varossieau, K., et al. (2021)  $\beta$ -Hexosaminidases Along the Secretory Pathway of *Nicotiana benthamiana* Have Distinct Specificities Toward Engineered Helminth N-Glycans on Recombinant Glycoproteins. *Front. Plant Sci.*, **12**.
- An, H.J., Froehlich, J.W., and Lebrilla, C.B. (2009) Determination of Glycosylation Sites and Site-specific Heterogeneity in Glycoproteins. *Curr. Opin. Chem. Biol.*, **13**, 421–426.
- Angata, T. and Varki, A. (2002) Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem. Rev.*, **102**, 439–469.
- Arrêté du 23 juillet 2021 modifiant la liste des spécialités pharmaceutiques remboursables aux assurés sociaux.
- Aviezer, D., Brill-Almon, E., Shaaltiel, Y., Hashmueli, S., Bartfeld, D., Mizrachi, S., et al. (2009) A Plant-Derived Recombinant Human Glucocerebrosidase Enzyme—A Preclinical and Phase I Investigation. *PLoS ONE*, **4**, e4792.
- Baier, T., Jacobebbinghaus, N., Einhaus, A., Lauersen, K.J., and Kruse, O. (2020) Introns mediate post-transcriptional enhancement of nuclear gene expression in the green microalga *Chlamydomonas reinhardtii*. *PLoS Genet.*, **16**, e1008944.
- Baier, T., Kros, D., Feiner, R.C., Lauersen, K.J., Müller, K.M., and Kruse, O. (2018) Engineered Fusion Proteins for Efficient Protein Secretion and Purification of a Human Growth Factor from the Green Microalga *Chlamydomonas reinhardtii*. *ACS Synth. Biol.*, **7**, 2547–2557.
- Baier, T., Wichmann, J., Kruse, O., and Lauersen, K.J. (2018) Intron-containing algal transgenes mediate efficient recombinant gene expression in the green microalga *Chlamydomonas reinhardtii*. *Nucleic Acids Res.*, **46**, 6909–6919.
- Baïet, B., Burel, C., Saint-Jean, B., Louvet, R., Menu-Bouaouiche, L., Kiefer-Meyer, M.-C., et al. (2011) N-glycans of *Phaeodactylum tricornutum* diatom and functional characterization of its N-acetylglucosaminyltransferase I enzyme. *J. Biol. Chem.*, **286**, 6152–6164.
- Bakker, H., Bardor, M., Molthoff, J.W., Gomord, V., Elbers, I., Stevens, L.H., et al. (2001) Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 2899–2904.
- Bakker, H., Rouwendal, G.J.A., Karnoup, A.S., Florack, D.E.A., Stopen, G.M., Helsper, J.P.F.G., et al. (2006) An antibody produced in tobacco expressing a hybrid  $\beta$ -1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 7577–7582.
- Baptista, C.G., Rodrigues, E.C., Morking, P., Klinke, A., Zardo, M.L., Soares, M.J., et al. (2015) Identification of a Golgi-localized UDP-N-acetylglucosamine transporter in *Trypanosoma cruzi*. *BMC Microbiol.*, **15**, 269.
- Bardor, M., Faveeuw, C., Fitchette, A.-C., Gilbert, D., Galas, L., Trottein, F., et al. (2003) Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology*, **13**, 427–434.
- Bardor, M., Nguyen, D.H., Diaz, S., and 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.

- Barlowe, C., Orci, L., Yeung, T., Hosobuchi, M., Hamamoto, S., Salama, N., et al. (1994) COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell*, **77**, 895–907.
- Barolo, L., Abbriano, R.M., Commault, A.S., George, J., Kahlke, T., Fabris, M., et al. (2020) Perspectives for Glyco-Engineering of Recombinant Biopharmaceuticals from Microalgae. *Cells*, **9**, 633.
- Basnet, N.B., Ide, K., Tahara, H., Tanaka, Y., and Ohdan, H. (2010) Deficiency of N-glycolylneuraminic acid and Gal $\alpha$ 1-3Gal $\beta$ 1-4GlcNAc epitopes in xenogeneic cells attenuates cytotoxicity of human natural antibodies. *Xenotransplantation*, **17**, 440–448.
- van Beers, M.M.C. and Bardor, M. (2012) Minimizing immunogenicity of biopharmaceuticals by controlling critical quality attributes of proteins. *Biotechnol. J.*, **7**, 1473–1484.
- Bencúr, P., Steinkellner, H., Svoboda, B., Mucha, J., Strasser, R., Kolarich, D., et al. (2005) *Arabidopsis thaliana*  $\beta$ 1,2-xylosyltransferase: an unusual glycosyltransferase with the potential to act at multiple stages of the plant N-glycosylation pathway. *Biochem. J.*, **388**, 515–525.
- Berlec, A. and Strukelj, B. (2013) Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. *J. Ind. Microbiol. Biotechnol.*, **40**, 257–274.
- Berndt, A.J., Smalley, T.N., Ren, B., Simkovsky, R., Badary, A., Sproles, A.E., et al. (2021) Recombinant production of a functional SARS-CoV-2 spike receptor binding domain in the green algae *Chlamydomonas reinhardtii*. *PloS One*, **16**, e0257089.
- Bibikova, M., Golic, M., Golic, K.G., and Carroll, D. (2002) Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics*, **161**, 1169–1175.
- Biel, T., Faison, T., Matthews, A., Zou, G., Ortega-Rodriguez, U., Pegues, M., et al. Engineering O-glycovariants of Etanercept with Heightened Potency Using an O-Glycoengineered CHO Cell Line Platform.
- Biel, T.G., Faison, T., Matthews, A.M., Zou, G., Ortega-Rodriguez, U., Pegues, M.A., et al. (2022) An etanercept O-glycovariant with enhanced potency. *Mol. Ther. Methods Clin. Dev.*, **25**, 124–135.
- Bindels, D.S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K.E., Mastop, M., et al. (2017) mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat. Methods*, **14**, 53–56.
- Blifernez-Klassen, O., Klassen, V., Doebbe, A., Kersting, K., Grimm, P., Wobbe, L., and Kruse, O. (2012) Cellulose degradation and assimilation by the unicellular phototrophic eukaryote *Chlamydomonas reinhardtii*. *Nat. Commun.*, **3**, 1214.
- Bobrowicz, P., Davidson, R.C., Li, H., Potgieter, T.I., Nett, J.H., Hamilton, S.R., et al. (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.
- Bohlender, L.L., Parsons, J., Hoernstein, S.N.W., Rempfer, C., Ruiz-Molina, N., Lorenz, T., et al. (2020) Stable Protein Sialylation in *Physcomitrella*. *Front. Plant Sci.*, **11**, 610032.
- Bollig, K., Lamshöft, M., Schweimer, K., Marner, F.-J., Budzikiewicz, H., and Waffenschmidt, S. (2007) Structural analysis of linear hydroxyproline-bound O-glycans of *Chlamydomonas reinhardtii*—conservation of the inner core in *Chlamydomonas* and land plants. *Carbohydr. Res.*, **342**, 2557–2566.
- Borys, M.C., Dalal, N.G., Abu-Absi, N.R., Khatkhat, S.F., Jing, Y., Xing, Z., and Li, Z.J. (2010) Effects of culture conditions on N-glycolylneuraminic acid (Neu5Gc) content of a recombinant fusion protein produced in CHO cells. *Biotechnol. Bioeng.*, **105**, 1048–1057.
- Bosques, C.J., Collins, B.E., Meador, J.W., Sarvaiya, H., Murphy, J.L., Dellorusso, G., et al. (2010) Chinese hamster ovary cells can produce galactose- $\alpha$ -1,3-galactose antigens on proteins. *Nat. Biotechnol.*, **28**, 1153–1156.

- Bragonzi, A., Distefano, G., Buckberry, L.D., Acerbis, G., Foglieni, C., Lamotte, D., et al. (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.
- Breitling, J. and Aebi, M. (2013) N-Linked Protein Glycosylation in the Endoplasmic Reticulum. *Cold Spring Harb. Perspect. Biol.*, **5**, a013359.
- Butler, M. (2005) Animal cell cultures: recent achievements and perspectives in the production of biopharmaceuticals. *Appl. Microbiol. Biotechnol.*, **68**, 283–291.
- Cadoret, J.-P., Bardor, M., Lerouge, P., Cabiglieri, M., Henriquez, V., and Carlier, A. (2008) Les microalgues : Usines cellulaires productrices de molécules commerciales recombinantes. *MS Médecine Sci. ISSN Pap. 0767-0974 ISSN Numér. 1958-5381 2008 Vol 24 N° 4 P 375-382*.
- Cantagrel, V. and Lefeber, D.J. (2011) From glycosylation disorders to dolichol biosynthesis defects: a new class of metabolic diseases. *J. Inherit. Metab. Dis.*, **34**, 859–867.
- Castilho, A., Gattinger, P., Grass, J., Jez, J., Pabst, M., Altmann, F., et al. (2011) N-glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans. *Glycobiology*, **21**, 813–823.
- Castilho, A., Neumann, L., Gattinger, P., Strasser, R., Vorauer-Uhl, K., Sterovsky, T., et al. (2013) Generation of biologically active multi-sialylated recombinant human EPOFc in plants. *PLoS One*, **8**, e54836.
- Castilho, A., Pabst, M., Leonard, R., Veit, C., Altmann, F., Mach, L., et al. (2008) Construction of a Functional CMP-Sialic Acid Biosynthesis Pathway in Arabidopsis. *Plant Physiol.*, **147**, 331–339.
- Castilho, A., Strasser, R., Stadlmann, J., Grass, J., Jez, J., Gattinger, P., et al. (2010) In planta protein sialylation through overexpression of the respective mammalian pathway. *J. Biol. Chem.*, **285**, 15923–15930.
- Castilho, A., Windwarder, M., Gattinger, P., Mach, L., Strasser, R., Altmann, F., and Steinkellner, H. (2014) Proteolytic and N-Glycan Processing of Human  $\alpha$ 1-Antitrypsin Expressed in *Nicotiana benthamiana*. *Plant Physiol.*, **166**, 1839–1851.
- Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., and Haslam, S.M. (2008) GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.*, **7**, 1650–1659.
- Chai, Y.-R., Cao, X.-X., Ge, M.-M., Mi, C.-L., Guo, X., and Wang, T.-Y. (2020) Knockout of cytidine monophosphate-N-acetylneuraminic acid hydroxylase in Chinese hamster ovary cells by CRISPR/Cas9-based gene-editing technology. *Biochem. Eng. J.*, **161**, 107663.
- Chance, R.E. and Frank, B.H. (1993) Research, development, production, and safety of biosynthetic human insulin. *Diabetes Care*, **16 Suppl 3**, 133–142.
- Chapitre 1er : Dispositions générales. (Articles L5121-1 à L5121-21) - Légifrance.
- Chávez, M.N., Schenck, T.L., Hopfner, U., Centeno-Cerdas, C., Somlai-Schweiger, I., Schwarz, C., et al. (2016) Towards autotrophic tissue engineering: Photosynthetic gene therapy for regeneration. *Biomaterials*, **75**, 25–36.
- Chavier, P. and Goud, B. (1999) The role of ARF and Rab GTPases in membrane transport. *Curr. Opin. Cell Biol.*, **11**, 466–475.
- Chen, Q. (2016) Glycoengineering of plants yields glycoproteins with polysialylation and other defined N-glycoforms. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, 9404–9406.
- Cheng, X., Liu, G., Ke, W., Zhao, L., Lv, B., Ma, X., et al. (2017) Building a multipurpose insertional mutant library for forward and reverse genetics in *Chlamydomonas*. *Plant Methods*, **13**, 36.
- Chenu, S., Grégoire, A., Malykh, Y., Visvikis, A., Monaco, L., Shaw, L., et al. (2003) Reduction of CMP-N-acetylneuraminic acid hydroxylase activity in engineered Chinese hamster ovary cells using an antisense-RNA strategy. *Biochim. Biophys. Acta BBA - Gen. Subj.*, **1622**, 133–144.

- Chiba, Y., Suzuki, M., Yoshida, S., Yoshida, A., Ikenaga, H., Takeuchi, M., et al. (1998) Production of Human Compatible High Mannose-type (Man5GlcNAc2) Sugar Chains in *Saccharomyces cerevisiae*\*. *J. Biol. Chem.*, **273**, 26298–26304.
- Choi, B.-K., Bobrowicz, P., Davidson, R.C., Hamilton, S.R., Kung, D.H., Li, H., et al. (2003) Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 5022–5027.
- Chojnacki, T. and Dallner, G. (1988) The biological role of dolichol. *Biochem. J.*, **251**, 1–9.
- Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., et al. (2010) Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics*, **186**, 757–761.
- Colley, K.J. (1997) Golgi localization of glycosyltransferases: more questions than answers. *Glycobiology*, **7**, 1–13.
- Commault, A.S., Walia, N.K., Fabris, M., Barolo, L., Siboni, N., Adriaans, J., et al. (2020) Effect of biphasic temperature regime on therapeutic recombinant protein production in the green alga *Chlamydomonas reinhardtii*. *Algal Res.*, **50**.
- Commissioner, O. of the (2023) Profound Regulation of EGFR Activation and Degradation Through O-glycosylation—Lessons learned from the study of colorectal cancer cells. FDA.
- Cooper, D.K., Good, A.H., Koren, E., Oriol, R., Malcolm, A.J., Ippolito, R.M., et al. (1993) Identification of alpha-galactosyl and other carbohydrate epitopes that are bound by human anti-pig antibodies: relevance to discordant xenografting in man. *Transpl. Immunol.*, **1**, 198–205.
- Cox, K.M., Sterling, J.D., Regan, J.T., Gasdaska, J.R., Frantz, K.K., Peele, C.G., et al. (2006) Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol.*, **24**, 1591–1597.
- Craig, R.J., Gallaher, S.D., Shu, S., Salomé, P.A., Jenkins, J.W., Blaby-Haas, C.E., et al. (2022) The *Chlamydomonas* Genome Project, version 6: Reference assemblies for mating-type plus and minus strains reveal extensive structural mutation in the laboratory. *Plant Cell*, **35**, 644–672.
- Crozet, P., Navarro, F.J., Willmund, F., Mehrshahi, P., Bakowski, K., Lauersen, K.J., et al. (2018) Birth of a Photosynthetic Chassis: A MoClo Toolkit Enabling Synthetic Biology in the Microalga *Chlamydomonas reinhardtii*. *ACS Synth. Biol.*, **7**, 2074–2086.
- Damerow, M., Graalfs, F., Güther, M.L.S., Mehlert, A., Izquierdo, L., and Ferguson, M.A.J. (2016) A Gene of the  $\beta$ 3-Glycosyltransferase Family Encodes N-Acetylglucosaminyltransferase II Function in *Trypanosoma brucei*. *J. Biol. Chem.*, **291**, 13834–13845.
- Damerow, M., Rodrigues, J.A., Wu, D., Güther, M.L.S., Mehlert, A., and Ferguson, M.A.J. (2014) Identification and functional characterization of a highly divergent N-acetylglucosaminyltransferase I (TbGnTI) in *Trypanosoma brucei*. *J. Biol. Chem.*, **289**, 9328–9339.
- Dauvillée, D., Delhay, S., Gruyer, S., Slomianny, C., Moretz, S.E., d’Hulst, C., et al. (2010) Engineering the chloroplast targeted malarial vaccine antigens in *Chlamydomonas* starch granules. *PloS One*, **5**, e15424.
- Davidson, R.C., Nett, J.H., Renfer, E., Li, H., Stadheim, T.A., Miller, B.J., et al. (2004) Functional analysis of the ALG3 gene encoding the Dol-P-Man: Man5GlcNAc2-PP-Dol mannosyltransferase enzyme of *P. pastoris*. *Glycobiology*, **14**, 399–407.
- De Pourcq, K., De Schutter, K., and Callewaert, N. (2010) Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Appl. Microbiol. Biotechnol.*, **87**, 1617–1631.
- De Pourcq, K., Tiels, P., Van Hecke, A., Geysens, S., Vervecken, W., and Callewaert, N. (2012) Engineering *Yarrowia lipolytica* to produce glycoproteins homogeneously modified with the universal Man3GlcNAc2 N-glycan core. *PloS One*, **7**, e39976.

- De Schutter, K., Lin, Y.-C., Tiels, P., Van Hecke, A., Glinka, S., Weber-Lehmann, J., et al. (2009) Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat. Biotechnol.*, **27**, 561–566.
- De Wachter, C., Van Landuyt, L., and Callewaert, N. (2021) Engineering of Yeast Glycoprotein Expression. In: *Advances in Glycobiotechnology Advances in Biochemical Engineering/Biotechnology* (Rapp, E. and Reichl, U., eds), pp. 93–135. Cham: Springer International Publishing.
- Dehghani, J., Adibkia, K., Movafeghi, A., Pourseif, M.M., and Omid, Y. (2020) Designing a new generation of expression toolkits for engineering of green microalgae; robust production of human interleukin-2. *BiolImpacts BI*, **10**, 259–268.
- Dehghani, J., Movafeghi, A., Mathieu-Rivet, E., Mati-Baouche, N., Calbo, S., Lerouge, P., and Bardor, M. (2022) Microalgae as an Efficient Vehicle for the Production and Targeted Delivery of Therapeutic Glycoproteins against SARS-CoV-2 Variants. *Mar. Drugs*, **20**, 657.
- Dementyeva, P., Freudenberg, R.A., Baier, T., Rojek, K., Wobbe, L., Weisshaar, B., and Kruse, O. (2021) A novel, robust and mating-competent *Chlamydomonas reinhardtii* strain with an enhanced transgene expression capacity for algal biotechnology. *Biotechnol. Rep. Amst. Neth.*, **31**, e00644.
- Demurtas, O.C., Massa, S., Ferrante, P., Venuti, A., Franconi, R., and Giuliano, G. (2013) A *Chlamydomonas*-derived Human Papillomavirus 16 E7 vaccine induces specific tumor protection. *PLoS One*, **8**, e61473.
- Dent, R.M., Haglund, C.M., Chin, B.L., Kobayashi, M.C., and Niyogi, K.K. (2005) Functional Genomics of Eukaryotic Photosynthesis Using Insertional Mutagenesis of *Chlamydomonas reinhardtii*. *Plant Physiol.*, **137**, 545–556.
- Doudna, J.A. and Charpentier, E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, **346**, 1258096.
- Dreesen, I.A.J., Charpin-El Hamri, G., and Fussenegger, M. (2010) Heat-stable oral alga-based vaccine protects mice from *Staphylococcus aureus* infection. *J. Biotechnol.*, **145**, 273–280.
- Dubé, S., Fisher, J.W., and Powell, J.S. (1988) Glycosylation at specific sites of erythropoietin is essential for biosynthesis, secretion, and biological function. *J. Biol. Chem.*, **263**, 17516–17521.
- Dubey, K.K., Kumar, A., Baldia, A., Rajput, D., Kateriya, S., Singh, R., et al. (2023) Biomanufacturing of glycosylated antibodies: Challenges, solutions, and future prospects. *Biotechnol. Adv.*, **69**, 108267.
- Eichler-Stahlberg, A., Weisheit, W., Ruecker, O., and Heitzer, M. (2009) Strategies to facilitate transgene expression in *Chlamydomonas reinhardtii*. *Planta*, **229**, 873–883.
- Eidenberger, L., Kogelmann, B., and Steinkellner, H. (2023) Plant-based biopharmaceutical engineering. *Nat. Rev. Bioeng.*, **1**, 426–439.
- Einhaus, A., Baier, T., Rosenstengel, M., Freudenberg, R.A., and Kruse, O. (2021) Rational Promoter Engineering Enables Robust Terpene Production in Microalgae. *ACS Synth. Biol.*, **10**, 847–856.
- Ellgaard, L. and Helenius, A. (2003) Quality control in the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.*, **4**, 181–191.
- Faye, L., Gomord, V., Fitchettelaine, A.C., and Chrispeels, M.J. (1993) Affinity Purification of Antibodies Specific for Asn-Linked Glycans Containing  $\alpha 1 \rightarrow 3$  Fucose or  $\beta 1 \rightarrow 2$  Xylose. *Anal. Biochem.*, **209**, 104–108.
- Felsenstein, J. (1985) CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evol. Int. J. Org. Evol.*, **39**, 783–791.
- Fischer, N. and Rochaix, J.D. (2001) The flanking regions of *PsaD* drive efficient gene expression in the nucleus of the green alga *Chlamydomonas reinhardtii*. *Mol. Genet. Genomics* **MGG**, **265**, 888–894.

- Fischer, R., Stoger, E., Schillberg, S., Christou, P., and Twyman, R.M. (2004) Plant-based production of biopharmaceuticals. *Curr. Opin. Plant Biol.*, **7**, 152–158.
- Fisher, P., Thomas-Oates, J., Wood, A.J., and Ungar, D. (2019) The N-Glycosylation Processing Potential of the Mammalian Golgi Apparatus. *Front. Cell Dev. Biol.*, **7**.
- Fitchette-Lainé, A.C., Gomord, V., Cabanes, M., Michalski, J.C., Saint Macary, M., Foucher, B., et al. (1997) N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J. Cell Mol. Biol.*, **12**, 1411–1417.
- Flowers, J.M., Hazzouri, K.M., Pham, G.M., Rosas, U., Bahmani, T., Khraiwesh, B., et al. (2015) Whole-Genome Resequencing Reveals Extensive Natural Variation in the Model Green Alga *Chlamydomonas reinhardtii*. *Plant Cell*, **27**, 2353–2369.
- Frey, A.D., Karg, S.R., and Kallio, P.T. (2009) Expression of rat beta(1,4)-N-acetylglucosaminyltransferase III in *Nicotiana tabacum* remodels the plant-specific N-glycosylation. *Plant Biotechnol. J.*, **7**, 33–48.
- Fuhrmann, M., Hausherr, A., Ferbitz, L., Schödl, T., Heitzer, M., and Hegemann, P. (2004) Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Mol. Biol.*, **55**, 869–881.
- Fuhrmann, M., Oertel, W., and Hegemann, P. (1999) A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. *Plant J. Cell Mol. Biol.*, **19**, 353–361.
- Fujita, A., Yoshida, T., and Ichishima, E. (1997) Five crucial carboxyl residues of 1,2-alpha-mannosidase from *Aspergillus saitoi* (A. phoenicis), a food microorganism, are identified by site-directed mutagenesis. *Biochem. Biophys. Res. Commun.*, **238**, 779–783.
- Fujiwara, T., Oda, K., Yokota, S., Takatsuki, A., and Ikehara, Y. (1988) Brefeldin A causes disassembly of the Golgi complex and accumulation of secretory proteins in the endoplasmic reticulum. *J. Biol. Chem.*, **263**, 18545–18552.
- Fukuda, M.N., Sasaki, H., Lopez, L., and Fukuda, M. (1989) Survival of Recombinant Erythropoietin in the Circulation: The Role of Carbohydrates. *Blood*, **73**, 84–89.
- Fukuta, K., Yokomatsu, T., Abe, R., Asanagi, M., and Makino, T. (2000) Genetic engineering of CHO cells producing human interferon-gamma by transfection of sialyltransferases. *Glycoconj. J.*, **17**, 895–904.
- Galili, U., Shohet, S.B., Kobrin, E., Stults, C.L., and Macher, B.A. (1988) Man, apes, and Old World monkeys differ from other mammals in the expression of alpha-galactosyl epitopes on nucleated cells. *J. Biol. Chem.*, **263**, 17755–17762.
- Gao, C., Yu, C.K.Y., Qu, S., San, M.W.Y., Li, K.Y., Lo, S.W., and Jiang, L. (2012) The Golgi-Localized Arabidopsis Endomembrane Protein12 Contains Both Endoplasmic Reticulum Export and Golgi Retention Signals at Its C Terminus[C][W]. *Plant Cell*, **24**, 2086–2104.
- Gao, H., Wang, Y., Fei, X., Wright, D.A., and Spalding, M.H. (2015) Expression activation and functional analysis of HLA3, a putative inorganic carbon transporter in *Chlamydomonas reinhardtii*. *Plant J. Cell Mol. Biol.*, **82**, 1–11.
- Geisler, K., Scaife, M.A., Mordaka, P.M., Holzer, A., Tomsett, E.V., Mehrshahi, P., et al. (2021) Exploring the Impact of Terminators on Transgene Expression in *Chlamydomonas reinhardtii* with a Synthetic Biology Approach. *Life Basel Switz.*, **11**, 964.
- Gemmill, T.R. and Trimble, R.B. (1999) Overview of N- and O-linked oligosaccharide structures found in various yeast species. *Biochim. Biophys. Acta BBA - Gen. Subj.*, **1426**, 227–237.
- Geng, D., Wang, Y., Wang, P., Li, W., and Sun, Y. (2003) Stable expression of hepatitis B surface antigen gene in *Dunaliella salina* (Chlorophyta). *J. Appl. Phycol.*, **15**, 451–456.
- Georgianna, D.R., Hannon, M.J., Marcuschi, M., Wu, S., Botsch, K., Lewis, A.J., et al. (2013) Production of recombinant enzymes in the marine alga *Dunaliella tertiolecta*. *Algal Res.*, **2**, 2–9.
- Ghaderi, D., Taylor, R.E., Padler-Karavani, V., Diaz, S., and Varki, A. (2010) Implications of the presence of N-glycolylneuraminic acid in recombinant therapeutic glycoproteins. *Nat. Biotechnol.*, **28**, 863–867.



- Giddings, T.H. (2003) Freeze-substitution protocols for improved visualization of membranes in high-pressure frozen samples. *J. Microsc.*, **212**, 53–61.
- Goetze, A.M., Liu, Y.D., Zhang, Z., Shah, B., Lee, E., Bondarenko, P.V., and Flynn, G.C. (2011) High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology*, **21**, 949–959.
- Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., et al. (1996) Life with 6000 genes. *Science*, **274**, 546, 563–567.
- Gomord, V., Fitchette, A.-C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., and Faye, L. (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol. J.*, **8**, 564–587.
- Göritzer, K., Grandits, M., Grünwald-Gruber, C., Figl, R., Mercx, S., Navarre, C., et al. (2022) Engineering the N-glycosylation pathway of *Nicotiana tabacum* for molecular pharming using CRISPR/Cas9. *Front. Plant Sci.*, **13**, 1003065.
- Gorman, D.S. and Levine, R.P. (1965) Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. U. S. A.*, **54**, 1665–1669.
- Goshtasbi, H., Okolodkov, Y.B., Movafeghi, A., Awale, S., Safary, A., Barar, J., and Omid, Y. (2023) Harnessing microalgae as sustainable cellular factories for biopharmaceutical production. *Algal Res.*, **74**, 103237.
- Greenfield, J.J.A. and High, S. (1999) The Sec61 complex is located in both the ER and the ER-Golgi intermediate compartment. *J. Cell Sci.*, **112**, 1477–1486.
- Gregory, J.A., Li, F., Tomosada, L.M., Cox, C.J., Topol, A.B., Vinetz, J.M., and Mayfield, S. (2012) Algae-produced Pfs25 elicits antibodies that inhibit malaria transmission. *PloS One*, **7**, e37179.
- Gregory, J.A., Topol, A.B., Doerner, D.Z., and Mayfield, S. (2013) Alga-Produced Cholera Toxin-Pfs25 Fusion Proteins as Oral Vaccines. *Appl. Environ. Microbiol.*, **79**, 3917–3925.
- Grunewald, S., Matthijs, G., and Jaeken, J. (2002) Congenital disorders of glycosylation: a review. *Pediatr. Res.*, **52**, 618–624.
- Gryz, E., Perlińska-Lenart, U., Gawarecka, K., Jozwiak, A., Pitsyk, S., Lipko, A., et al. (2019) Poly-Saturated Dolichols from Filamentous Fungi Modulate Activity of Dolichol-Dependent Glycosyltransferase and Physical Properties of Membranes. *Int. J. Mol. Sci.*, **20**, 3043.
- Gutiérrez, S. and Lauersen, K.J. (2021) Gene Delivery Technologies with Applications in Microalgal Genetic Engineering. *Biology*, **10**, 265.
- Gutternigg, M., Kretschmer-Lubich, D., Paschinger, K., Rendić, D., Hader, J., Geier, P., et al. (2007) Biosynthesis of Truncated N-Linked Oligosaccharides Results from Non-orthologous Hexosaminidase-mediated Mechanisms in Nematodes, Plants, and Insects. *J. Biol. Chem.*, **282**, 27825–27840.
- Haas, I.G. (1994) BiP (GRP78), an essential hsp70 resident protein in the endoplasmic reticulum. *Experientia*, **50**, 1012–1020.
- Haddley, K. (2012) Taliglucerase alfa for the treatment of Gaucher's disease. *Drugs Today Barc. Spain* 1998, **48**, 525–532.
- Hager, K.J., Pérez Marc, G., Gobeil, P., Diaz, R.S., Heizer, G., Llapur, C., et al. (2022) Efficacy and Safety of a Recombinant Plant-Based Adjuvanted Covid-19 Vaccine. *N. Engl. J. Med.*, **386**, 2084–2096.
- Hallgren, J., Tsigirgos, K.D., Pedersen, M.D., Armenteros, J.J.A., Marcatili, P., Nielsen, H., et al. (2022) DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. 2022.04.08.487609.
- Hamilton, S.R., Bobrowicz, P., Bobrowicz, B., Davidson, R.C., Li, H., Mitchell, T., et al. (2003) Production of complex human glycoproteins in yeast. *Science*, **301**, 1244–1246.
- Hamilton, S.R., Davidson, R.C., Sethuraman, N., Nett, J.H., Jiang, Y., Rios, S., et al. (2006) Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science*, **313**, 1441–1443.

- Hamilton, S.R. and Gerngross, T.U. (2007) Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr. Opin. Biotechnol.*, **18**, 387–392.
- Han, Y., Kanbe, T., Cherniak, R., and Cutler, J.E. (1997) Biochemical characterization of *Candida albicans* epitopes that can elicit protective and nonprotective antibodies. *Infect. Immun.*, **65**, 4100–4107.
- Harduin-Lepers, A. (2023) The vertebrate sialylation machinery: structure-function and molecular evolution of GT-29 sialyltransferases. *Glycoconj. J.*, **40**, 473–492.
- Harris, E.H. (2001) CHLAMYDOMONAS AS A MODEL ORGANISM. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **52**, 363–406.
- He, D.-M., Qian, K.-X., Shen, G.-F., Zhang, Z.-F., Li, Y.-N., Su, Z.-L., and Shao, H.-B. (2007) Recombination and expression of classical swine fever virus (CSFV) structural protein E2 gene in *Chlamydomonas reinhardtii* chloroplasts. *Colloids Surf. B Biointerfaces*, **55**, 26–30.
- Hebert, D.N. and Molinari, M. (2007) In and Out of the ER: Protein Folding, Quality Control, Degradation, and Related Human Diseases. *Physiol. Rev.*, **87**, 1377–1408.
- Helenius, A. and Aebi, M. (2004) Roles of N-Linked Glycans in the Endoplasmic Reticulum. *Annu. Rev. Biochem.*, **73**, 1019–1049.
- Helenius, J., Ng, D.T.W., Marolda, C.L., Walter, P., Valvano, M.A., and Aebi, M. (2002) Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature*, **415**, 447–450.
- Hellwig, S., Drossard, J., Twyman, R.M., and Fischer, R. (2004) Plant cell cultures for the production of recombinant proteins. *Nat. Biotechnol.*, **22**, 1415–1422.
- Hempel, F., Lau, J., Klingl, A., and 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.
- Hempel, F. and Maier, U.G. (2012) An engineered diatom acting like a plasma cell secreting human IgG antibodies with high efficiency. *Microb. Cell Factories*, **11**, 126.
- Hennermann, J.B., Arash-Kaps, L., Fekete, G., Schaaf, A., Busch, A., and Frischmuth, T. (2019) Pharmacokinetics, pharmacodynamics, and safety of moss- $\alpha$ Galactosidase A in patients with Fabry disease. *J. Inher. Metab. Dis.*, **42**, 527–533.
- Hernández-Ramírez, J., Wong-Arce, A., González-Ortega, O., and Rosales-Mendoza, S. (2020) Expression in algae of a chimeric protein carrying several epitopes from tumor associated antigens. *Int. J. Biol. Macromol.*, **147**, 46–52.
- Herscovics, A. (2001) Structure and function of Class I  $\alpha$ 1,2-mannosidases involved in glycoprotein synthesis and endoplasmic reticulum quality control. *Biochimie*, **83**, 757–762.
- Hood, E.E., Kusnadi, A., Nikolov, Z., and Howard, J.A. (1999) Molecular farming of industrial proteins from transgenic maize. *Adv. Exp. Med. Biol.*, **464**, 127–147.
- Hou, Q., Qiu, S., Liu, Q., Tian, J., Hu, Z., and Ni, J. (2013) Selenoprotein-transgenic *Chlamydomonas reinhardtii*. *Nutrients*, **5**, 624–636.
- Houeix, B. and Cairns, M.T. (2019) Engineering of CHO cells for the production of vertebrate recombinant sialyltransferases. *PeerJ*, **7**, e5788.
- Huang, C.-J., Lin, H., and Yang, X. (2012) Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J. Ind. Microbiol. Biotechnol.*, **39**, 383–399.
- Hutner, S.H., Provasoli, L., Schatz, A., and Haskins, C.P. (1950) Some Approaches to the Study of the Role of Metals in the Metabolism of Microorganisms. *Proc. Am. Philos. Soc.*, **94**, 152–170.
- Jaeger, D., Baier, T., and Lauersen, K.J. (2019) Intronserter, an advanced online tool for design of intron containing transgenes. *Algal Res.*, **42**, 101588.
- Jansing, J., Sack, M., Augustine, S.M., Fischer, R., and Bortesi, L. (2019) CRISPR/Cas9-mediated knockout of six glycosyltransferase genes in *Nicotiana benthamiana* for the production

- of recombinant proteins lacking  $\beta$ -1,2-xylose and core  $\alpha$ -1,3-fucose. *Plant Biotechnol. J.*, **17**, 350–361.
- Jarquín-Cordero, M., Chávez, M.N., Centeno-Cerdas, C., Bohne, A.-V., Hopfner, U., Machens, H.-G., et al. (2020) Towards a biotechnological platform for the production of human pro-angiogenic growth factors in the green alga *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.*, **104**, 725–739.
- Jez, J., Castilho, A., Grass, J., Vorauer-Uhl, K., Sterovsky, T., Altmann, F., and Steinkellner, H. (2013) Expression of functionally active sialylated human erythropoietin in plants. *Biotechnol. J.*, **8**, 371–382.
- Jiang, X. and Stern, D. (2009) Mating and Tetrad Separation of *Chlamydomonas reinhardtii* for Genetic Analysis. *JoVE J. Vis. Exp.*, e1274.
- Jones, C.S., Luong, T., Hannon, M., Tran, M., Gregory, J.A., Shen, Z., et al. (2013) Heterologous expression of the C-terminal antigenic domain of the malaria vaccine candidate Pfs48/45 in the green algae *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.*, **97**, 1987–1995.
- Jones, M.B., Rosenberg, J.N., Betenbaugh, M.J., and Krag, S.S. (2009) Structure and synthesis of polyisoprenoids used in N-glycosylation across the three domains of life. *Biochim. Biophys. Acta*, **1790**, 485–494.
- Jungmann, J. and Munro, S. (1998) Multi-protein complexes in the cis Golgi of *Saccharomyces cerevisiae* with  $\alpha$ -1,6-mannosyltransferase activity. *EMBO J.*, **17**, 423–434.
- Jungmann, J., Rayner, J.C., and Munro, S. (1999) The *Saccharomyces cerevisiae* protein Mnn10p/Bed1p is a subunit of a Golgi mannosyltransferase complex. *J. Biol. Chem.*, **274**, 6579–6585.
- Kajiura, H., Koiwa, H., Nakazawa, Y., Okazawa, A., Kobayashi, A., Seki, T., and Fujiyama, K. (2010) Two *Arabidopsis thaliana* Golgi  $\alpha$ -mannosidase I enzymes are responsible for plant N-glycan maturation. *Glycobiology*, **20**, 235–247.
- Kajiura, H., Okamoto, T., Misaki, R., Matsuura, Y., and Fujiyama, K. (2012) *Arabidopsis*  $\beta$ 1,2-xylosyltransferase: Substrate specificity and participation in the plant-specific N-glycosylation pathway. *J. Biosci. Bioeng.*, **113**, 48–54.
- Kallolimath, S., Castilho, A., Strasser, R., Grünwald-Gruber, C., Altmann, F., Strubl, S., et al. (2016) Engineering of complex protein sialylation in plants. *Proc. Natl. Acad. Sci.*, **113**, 9498–9503.
- Kamionka, M. (2011) Engineering of Therapeutic Proteins Production in *Escherichia coli*. *Curr. Pharm. Biotechnol.*, **12**, 268–274.
- Karaoglu, D., Kelleher, D.J., and Gilmore, R. (2001) Allosteric regulation provides a molecular mechanism for preferential utilization of the fully assembled dolichol-linked oligosaccharide by the yeast oligosaccharyltransferase. *Biochemistry*, **40**, 12193–12206.
- Keeling, P.J. (2013) The Number, Speed, and Impact of Plastid Endosymbioses in Eukaryotic Evolution. *Annu. Rev. Plant Biol.*, **64**, 583–607.
- Kelleher, D.J., Banerjee, S., Cura, A.J., Samuelson, J., and Gilmore, R. (2007) Dolichol-linked oligosaccharide selection by the oligosaccharyltransferase in protist and fungal organisms. *J. Cell Biol.*, **177**, 29–37.
- Kelleher, D.J. and Gilmore, R. (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology*, **16**, 47R–62R.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.
- Kelterborn, S., Boehning, F., Sizova, I., Baidukova, O., Evers, H., and Hegemann, P. (2022) Gene Editing in Green Alga *Chlamydomonas reinhardtii* via CRISPR-Cas9 Ribonucleoproteins. *Methods Mol. Biol. Clifton NJ*, **2379**, 45–65.
- Kesik-Brodacka, M. (2018) Progress in biopharmaceutical development. *Biotechnol. Appl. Biochem.*, **65**, 306–322.

- Khoury, G.A., Baliban, R.C., and Floudas, C.A. (2011) Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.*, **1**, 90.
- Kiefer, A.M., Niemeyer, J., Probst, A., Erkel, G., and Schroda, M. (2022) Production and secretion of functional SARS-CoV-2 spike protein in *Chlamydomonas reinhardtii*. *Front. Plant Sci.*, **13**, 988870.
- Kornfeld, R. and Kornfeld, S. (1985) Assembly of Asparagine-Linked Oligosaccharides. *Annu. Rev. Biochem.*, **54**, 631–664.
- Kostova, Z. and Wolf, D.H. (2003) NEW EMBO MEMBER'S REVIEW. *EMBO J.*, **22**, 2309–2317.
- Krainer, F.W., Gmeiner, C., Neutsch, L., Windwarder, M., Pletzenauer, R., Herwig, C., et al. (2013) Knockout of an endogenous mannosyltransferase increases the homogeneity of glycoproteins produced in *Pichia pastoris*. *Sci. Rep.*, **3**, 3279.
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.*, **305**, 567–580.
- Kulagina, N., Besseau, S., Godon, C., Goldman, G.H., Papon, N., and Courdavault, V. (2021) Yeasts as Biopharmaceutical Production Platforms. *Front. Fungal Biol.*, **2**.
- Kumar, A., Falcao, V.R., and Sayre, R.T. (2013) Evaluating nuclear transgene expression systems in *Chlamydomonas reinhardtii*. *Algal Res.*, **2**, 321–332.
- Kurosaka, A., Yano, A., Itoh, N., Kuroda, Y., Nakagawa, T., and Kawasaki, T. (1991) The structure of a neural specific carbohydrate epitope of horseradish peroxidase recognized by anti-horseradish peroxidase antiserum. *J. Biol. Chem.*, **266**, 4168–4172.
- Lam, A.J., St-Pierre, F., Gong, Y., Marshall, J.D., Cranfill, P.J., Baird, M.A., et al. (2012) Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat. Methods*, **9**, 1005–1012.
- Lam, C. and Krasnewich, D.M. (1993) PMM2-CDG. In: *GeneReviews®* (Adam, M.P., Feldman, J., Mirzaa, G.M., Pagon, R.A., Wallace, S.E., Bean, L.J., et al., eds). Seattle (WA): University of Washington, Seattle.
- Lambert, T.J. (2019) FPbase: a community-editable fluorescent protein database. *Nat. Methods*, **16**, 277–278.
- Lang, S., Pfeffer, S., Lee, P.-H., Cavalié, A., Helms, V., Förster, F., and Zimmermann, R. (2017) An Update on Sec61 Channel Functions, Mechanisms, and Related Diseases. *Front. Physiol.*, **8**.
- Langer, E. and Rader, R. (2015) Biopharmaceutical Manufacturing: Historical and Future Trends in Titrations, Yields, and Efficiency in Commercial-Scale Bioprocessing. *Bioprocess. J.*, **13**, 47–54.
- Lauersen, K.J., Berger, H., Mussgnug, J.H., and Kruse, O. (2013) Efficient recombinant protein production and secretion from nuclear transgenes in *Chlamydomonas reinhardtii*. *J. Biotechnol.*, **167**, 101–110.
- Lauersen, K.J., Vanderveer, T.L., Berger, H., Kaluza, I., Mussgnug, J.H., Walker, V.K., and Kruse, O. (2013) Ice recrystallization inhibition mediated by a nuclear-expressed and -secreted recombinant ice-binding protein in the microalga *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.*, **97**, 9763–9772.
- Lee, E.U., Roth, J., and 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.
- Lee, J., Lee, S.-K., Park, J.-S., and Lee, K.-R. (2023) Plant-made pharmaceuticals: exploring studies for the production of recombinant protein in plants and assessing challenges ahead. *Plant Biotechnol. Rep.*, **17**, 53–65.
- Léonard, R., Costa, G., Darrambide, E., Lhernould, S., Fleurat-Lessard, P., Carlué, M., et al. (2002) The presence of Lewis a epitopes in *Arabidopsis thaliana* glycoconjugates depends on an active alpha4-fucosyltransferase gene. *Glycobiology*, **12**, 299–306.

- Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.C., Gomord, V., and Faye, L. (1998) N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol. Biol.*, **38**, 31–48.
- Les médicaments biosimilaires dans l'UE - Guide d'information destiné aux professionnels de la santé.
- Levy-Ontman, O., Fisher, M., Shotland, Y., Weinstein, Y., Tekoah, Y., and 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.
- Li, X., Patena, W., Fauser, F., Jinkerson, R.E., Saroussi, S., Meyer, M.T., et al. (2019) A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. *Nat. Genet.*, **51**, 627–635.
- Li, X., Zhang, R., Patena, W., Gang, S.S., Blum, S.R., Ivanova, N., et al. (2016) An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in *Chlamydomonas reinhardtii*. *Plant Cell*, **28**, 367–387.
- Liebmingier, E., Veit, C., Pabst, M., Batoux, M., Zipfel, C., Altmann, F., et al. (2011)  $\beta$ -N-Acetylhexosaminidases HEXO1 and HEXO3 Are Responsible for the Formation of Paucimannosidic N-Glycans in *Arabidopsis thaliana*\*. *J. Biol. Chem.*, **286**, 10793–10802.
- Lin, N., Mascarenhas, J., Sealover, N.R., George, H.J., Brooks, J., Kayser, K.J., et al. (2015) Chinese hamster ovary (CHO) host cell engineering to increase sialylation of recombinant therapeutic proteins by modulating sialyltransferase expression. *Biotechnol. Prog.*, **31**, 334–346.
- Lobato Gómez, M., Huang, X., Alvarez, D., He, W., Baysal, C., Zhu, C., et al. (2021) Contributions of the international plant science community to the fight against human infectious diseases – part 1: epidemic and pandemic diseases. *Plant Biotechnol. J.*, **19**, 1901–1920.
- Loos, A. and Castilho, A. (2015) Transient Glyco-Engineering of *N. benthamiana* Aiming at the Synthesis of Multi-antennary Sialylated Proteins. *Methods Mol. Biol. Clifton NJ*, **1321**, 233–248.
- Loos, A., Gruber, C., Altmann, F., Mehofer, U., Hensel, F., Grandits, M., et al. (2014) Expression and glycoengineering of functionally active heteromultimeric IgM in plants. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 6263–6268.
- Loos, A. and Steinkellner, H. (2014) Plant glyco-biotechnology on the way to synthetic biology. *Front. Plant Sci.*, **5**, 523.
- López-Paz, C., Liu, D., Geng, S., and Umen, J.G. (2017) Identification of *Chlamydomonas reinhardtii* endogenous genic flanking sequences for improved transgene expression. *Plant J. Cell Mol. Biol.*, **92**, 1232–1244.
- Löw, P., Dallner, G., Mayor, S., Cohen, S., Chait, B.T., and Menon, A.K. (1991) The mevalonate pathway in the bloodstream form of *Trypanosoma brucei*. Identification of dolichols containing 11 and 12 isoprene residues. *J. Biol. Chem.*, **266**, 19250–19257.
- Lucas, P.-L. (2019) Etude et ingénierie de la N-glycosylation des protéines chez la microalgue verte *chlamydomonas reinhardtii*.
- Lucas, P.-L., Dumontier, R., Loutelier-Bourhis, C., Mareck, A., Afonso, C., Lerouge, P., et al. (2018) User-friendly extraction and multistage tandem mass spectrometry based analysis of lipid-linked oligosaccharides in microalgae. *Plant Methods*, **14**, 107.
- Lucas, P.-L., Mathieu-Rivet, E., Song, P.C.T., Oltmanns, A., Loutelier-Bourhis, C., Plasson, C., et al. (2020) Multiple xylosyltransferases heterogeneously xylosylate protein N-linked glycans in *Chlamydomonas reinhardtii*. *Plant J.*, **102**, 230–245.
- Mackinder, L.C.M., Chen, C., Leib, R.D., Patena, W., Blum, S.R., Rodman, M., et al. (2017) A Spatial Interactome Reveals the Protein Organization of the Algal CO<sub>2</sub>-Concentrating Mechanism. *Cell*, **171**, 133–147.e14.
- Mackinder, L.C.M., Meyer, M.T., Mettler-Altmann, T., Chen, V.K., Mitchell, M.C., Caspari, O., et al. (2016) A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, 5958–5963.

- Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., et al. (2022) Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.*, **50**, W276–W279.
- Madhavan, A., Arun, K.B., Sindhu, R., Krishnamoorthy, J., Reshmy, R., Sirohi, R., et al. (2021) Customized yeast cell factories for biopharmaceuticals: from cell engineering to process scale up. *Microb. Cell Factories*, **20**, 124.
- Magnusdottir, A., Vidarsson, H., Björnsson, J.M., and Örvar, B.L. (2013) Barley grains for the production of endotoxin-free growth factors. *Trends Biotechnol.*, **31**, 572–580.
- Maley, F., Trimble, R.B., Tarentino, A.L., and Plummer, T.H. (1989) Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal. Biochem.*, **180**, 195–204.
- Manuell, A.L., Beligni, M.V., Elder, J.H., Siefker, D.T., Tran, M., Weber, A., et al. (2007) Robust expression of a bioactive mammalian protein in *Chlamydomonas* chloroplast. *Plant Biotechnol. J.*, **5**, 402–412.
- Mathieu-Rivet, E., Kiefer-Meyer, M.-C., Vanier, G., Ovide, C., Burel, C., Lerouge, P., and Bardor, M. (2014) Protein N-glycosylation in eukaryotic microalgae and its impact on the production of nuclear expressed biopharmaceuticals. *Front. Plant Sci.*, **5**, 359.
- Mathieu-Rivet, E., Mati-Baouche, N., Walet-Balieu, M.-L., Lerouge, P., and Bardor, M. (2020) N- and O-Glycosylation Pathways in the Microalgae Polyphyletic Group. *Front. Plant Sci.*, **11**, 609993.
- Mathieu-Rivet, E., Scholz, M., Arias, C., Dardelle, F., Schulze, S., Le Mauff, F., et al. (2013) Exploring the N-glycosylation pathway in *Chlamydomonas reinhardtii* unravels novel complex structures. *Mol. Cell. Proteomics MCP*, **12**, 3160–3183.
- Mayfield, S.P., Franklin, S.E., and Lerner, R.A. (2003) Expression and assembly of a fully active antibody in algae. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 438–442.
- McATEER, M., Donnan, L., and John, P.C.L. (1985) The Timing of Division in *Chlamydomonas*. *New Phytol.*, **99**, 41–56.
- Melo, N.S., Nimtz, M., Conradt, H.S., Fevereço, P.S., and Costa, J. (1997) Identification of the human Lewis(a) carbohydrate motif in a secretory peroxidase from a plant cell suspension culture (*Vaccinium myrtillus* L.). *FEBS Lett.*, **415**, 186–191.
- Mercx, S., Smargiasso, N., Chaumont, F., De Pauw, E., Boutry, M., and 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**, 403.
- Merten, O.-W. (2002) Virus contaminations of cell cultures – A biotechnological view. *Cytotechnology*, **39**, 91–116.
- Mohideen, F.I. and Mahal, L.K. (2024) Infection and the Glycome—New Insights into Host Response. *ACS Infect. Dis.*, **10**, 2540–2550.
- Mohorko, E., Glockshuber, R., and Aebi, M. (2011) Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. *J. Inherit. Metab. Dis.*, **34**, 869–878.
- Molinari, M. and Helenius, A. (1999) Glycoproteins form mixed disulphides with oxidoreductases during folding in living cells. *Nature*, **402**, 90–93.
- Molino, J.V.D., de Carvalho, J.C.M., and Mayfield, S.P. (2018) Comparison of secretory signal peptides for heterologous protein expression in microalgae: Expanding the secretion portfolio for *Chlamydomonas reinhardtii*. *PLoS ONE*, **13**, e0192433.
- Moremen, K.W., Tiemeyer, M., and Nairn, A.V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.*, **13**, 448–462.
- Mukherjee, S. and Jimenez, R. (2022) Photophysical Engineering of Fluorescent Proteins: Accomplishments and Challenges of Physical Chemistry Strategies. *J. Phys. Chem. B*, **126**, 735–750.
- Nakanishi-Shindo, Y., Nakayama, K., Tanaka, A., Toda, Y., and 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.
- Nakano, M., Mishra, S.K., Tokoro, Y., Sato, K., Nakajima, K., Yamaguchi, Y., et al. (2019) Bisecting GlcNAc Is a General Suppressor of Terminal Modification of N-glycan\*, [S]. *Mol. Cell. Proteomics*, **18**, 2044–2057.
- Nakayama, K., Nagasu, T., Shimma, Y., Kuromitsu, J., and Jigami, Y. (1992) OCH1 encodes a novel membrane bound mannosyltransferase: outer chain elongation of asparagine-linked oligosaccharides. *EMBO J.*, **11**, 2511–2519.
- Narimatsu, Y., Büll, C., Chen, Y.-H., Wandall, H.H., Yang, Z., and Clausen, H. (2021) Genetic glycoengineering in mammalian cells. *J. Biol. Chem.*, **296**, 100448.
- NEUMANN, U., BRANDIZZI, F., and HAWES, C. (2003) Protein Transport in Plant Cells: In and Out of the Golgi. *Ann. Bot.*, **92**, 167–180.
- Neupert, J., Gallaher, S.D., Lu, Y., Strenkert, D., Segal, N., Barahimipour, R., et al. (2020) An epigenetic gene silencing pathway selectively acting on transgenic DNA in the green alga *Chlamydomonas*. *Nat. Commun.*, **11**, 6269.
- Neupert, J., Karcher, D., and Bock, R. (2009) Generation of *Chlamydomonas* strains that efficiently express nuclear transgenes. *Plant J. Cell Mol. Biol.*, **57**, 1140–1150.
- Nielsen, J. (2013) Production of biopharmaceutical proteins by yeast. *Bioengineered*, **4**, 207–211.
- Odani, T., Shimma, Y., Wang, X.-H., and Jigami, Y. (1997) Mannosylphosphate transfer to cell wall mannan is regulated by the transcriptional level of the MNN4 gene in *Saccharomyces cerevisiae*. *FEBS Lett.*, **420**, 186–190.
- O’Flaherty, R., Bergin, A., Flampouri, E., Mota, L.M., Obaidi, I., Quigley, A., et al. (2020) Mammalian cell culture for production of recombinant proteins: A review of the critical steps in their biomanufacturing. *Biotechnol. Adv.*, **43**, 107552.
- Olden, K., Bernard, B.A., Humphries, M.J., Yeo, T.-K., Yeo, K.-T., White, S.L., et al. (1985) Function of glycoprotein glycans. *Trends Biochem. Sci.*, **10**, 78–82.
- Oltmanns, A., Hoepfner, L., Scholz, M., Zinzus, K., Schulze, S., and Hippler, M. (2020) Novel Insights Into N-Glycan Fucosylation and Core Xylosylation in *C. reinhardtii*. *Front. Plant Sci.*, **10**.
- Oriol, R., Ye, Y., Koren, E., and Cooper, D.K. (1993) Carbohydrate antigens of pig tissues reacting with human natural antibodies as potential targets for hyperacute vascular rejection in pig-to-man organ xenotransplantation. *Transplantation*, **56**, 1433–1442.
- Paccalet, T., Bardor, M., Rihouey, C., Delmas, F., Chevalier, C., D’Aoust, M.-A., et al. (2007) Engineering of a sialic acid synthesis pathway in transgenic plants by expression of bacterial Neu5Ac-synthesizing enzymes. *Plant Biotechnol. J.*, **5**, 16–25.
- Pagny, S., Bouissonnie, F., Sarkar, M., Follet-Gueye, M.L., Driouich, A., Schachter, H., et al. (2003) Structural requirements for *Arabidopsis* beta1,2-xylosyltransferase activity and targeting to the Golgi. *Plant J. Cell Mol. Biol.*, **33**, 189–203.
- Palacpac, N.Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T., and Seki, T. (1999) Stable expression of human  $\beta$ 1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *Proc. Natl. Acad. Sci. U. S. A.*, **96**, 4692–4697.
- Parameswari., R.P. and Lakshmi, T. (2022) Microalgae as a potential therapeutic drug candidate for neurodegenerative diseases. *J. Biotechnol.*, **358**, 128–139.
- Parkins, D.A. and Lashmar, U.T. (2000) The formulation of biopharmaceutical products. *Pharm. Sci. Technol. Today*, **3**, 129–137.
- Paulson, J.C. and Colley, K.J. (1989) Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J. Biol. Chem.*, **264**, 17615–17618.
- Pederson, S., Biondi, D.M., Allan, B., Cady, R., Schaeffler, B., Baker, B., and Latham, J. (2021) Clinical Immunogenicity Evaluation of Eptinezumab, a Therapeutic Humanized Monoclonal Antibody Targeting Calcitonin Gene-Related Peptide (CGRP) for the Preventive Treatment of Migraine. *Front. Immunol.*, **12**, 765822.

- Peri, S., Kulkarni, A., Feyertag, F., Berninsone, P.M., and Alvarez-Ponce, D. (2018) Phylogenetic Distribution of CMP-Neu5Ac Hydroxylase (CMAH), the Enzyme Synthetizing the Proinflammatory Human Xenoantigen Neu5Gc. *Genome Biol. Evol.*, **10**, 207–219.
- Piirainen, M.A., Salminen, H., and Frey, A.D. (2022) Production of galactosylated complex-type N-glycans in glycoengineered *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.*, **106**, 301–315.
- Poulhazan, A., Arnold, A.A., Mentink-Vigier, F., Muszyński, A., Azadi, P., Halim, A., et al. (2024) Molecular-level architecture of *Chlamydomonas reinhardtii*'s glycoprotein-rich cell wall. *Nat. Commun.*, **15**, 986.
- Qiu, X., Wong, G., Audet, J., Bello, A., Fernando, L., Alimonti, J.B., et al. (2014) Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp. *Nature*, **514**, 47–53.
- Rabouille, C., Hui, N., Hunte, F., Kieckbusch, R., Berger, E.G., Warren, G., and Nilsson, T. (1995) Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. *J. Cell Sci.*, **108**, 1617–1627.
- Rademacher, T.W., Parekh, R.B., and Dwek, R.A. (1988) Glycobiology. *Annu. Rev. Biochem.*, **57**, 785–838.
- Ramazi, S. and Zahiri, J. (2021) Post-translational modifications in proteins: resources, tools and prediction methods. *Database J. Biol. Databases Curation*, **2021**, baab012.
- Ramírez, A.S., de Capitani, M., Pesciullesi, G., Kowal, J., Bloch, J.S., Irobalieva, R.N., et al. (2022) Molecular basis for glycan recognition and reaction priming of eukaryotic oligosaccharyltransferase. *Nat. Commun.*, **13**, 7296.
- Ramos-Martinez, E.M., Fimognari, L., and Sakuragi, Y. (2017) High-yield secretion of recombinant proteins from the microalga *Chlamydomonas reinhardtii*. *Plant Biotechnol. J.*, **15**, 1214–1224.
- Rasala, B.A., Barrera, D.J., Ng, J., Plucinak, T.M., Rosenberg, J.N., Weeks, D.P., et al. (2013) Expanding the spectral palette of fluorescent proteins for the green microalga *Chlamydomonas reinhardtii*. *Plant J.*, **74**, 545–556.
- Rasala, B.A., Lee, P.A., Shen, Z., Briggs, S.P., Mendez, M., and Mayfield, S.P. (2012) Robust expression and secretion of Xylanase 1 in *Chlamydomonas reinhardtii* by fusion to a selection gene and processing with the FMDV 2A peptide. *PLoS One*, **7**, e43349.
- Rasala, B.A. and Mayfield, S.P. (2011) The microalga *Chlamydomonas reinhardtii* as a platform for the production of human protein therapeutics. *Bioeng. Bugs*, **2**, 50–54.
- Rasala, B.A., Muto, M., Lee, P.A., Jager, M., Cardoso, R.M.F., Behnke, C.A., et al. (2010) Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of *Chlamydomonas reinhardtii*. *Plant Biotechnol. J.*, **8**, 719–733.
- Raschke, W.C., Kern, K.A., Antalis, C., and Ballou, C.E. (1973) Genetic Control of Yeast Mannan Structure: ISOLATION AND CHARACTERIZATION OF MANNAN MUTANTS. *J. Biol. Chem.*, **248**, 4660–4666.
- Rayner, J.C. and Munro, S. (1998) Identification of the MNN2 and MNN5 mannosyltransferases required for forming and extending the mannose branches of the outer chain mannans of *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **273**, 26836–26843.
- Research, C. for B.E. and (2019) What Are “Biologics” Questions and Answers. FDA.
- Ritzenthaler, C., Nebenführ, A., Movafeghi, A., Stussi-Garaud, C., Behnia, L., Pimpl, P., et al. (2002) Reevaluation of the Effects of Brefeldin A on Plant Cells Using Tobacco Bright Yellow 2 Cells Expressing Golgi-Targeted Green Fluorescent Protein and COPI Antisera. *Plant Cell*, **14**, 237–261.
- Ropitiaux, M., Bernard, S., Boulogne, I., Goux, D., Mollet, J.-C., Lerouge, P., et al. (2024) Subcellular localization of core beta(1,2)-xylosylated N-glycoproteins in the green microalgae *Chlamydomonas reinhardtii*. *Algal Res.*, **77**, 103366.
- Rosano, G.L. and Ceccarelli, E.A. (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.*, **5**, 172.



- Rouwental, G.J.A., Wuhrer, M., Florack, D.E.A., Koeleman, C.A.M., Deelder, A.M., Bakker, H., et al. (2007) Efficient introduction of a bisecting GlcNAc residue in tobacco N-glycans by expression of the gene encoding human N-acetylglucosaminyltransferase III. *Glycobiology*, **17**, 334–344.
- Rudd, P.M. and Dwek, R.A. (1997) Glycosylation: Heterogeneity and the 3D Structure of Proteins. *Crit. Rev. Biochem. Mol. Biol.*, **32**, 1–100.
- Ruddock, L.W. and Molinari, M. (2006) N-glycan processing in ER quality control. *J. Cell Sci.*, **119**, 4373–4380.
- Ruiz-Canada, C., Kelleher, D.J., and Gilmore, R. (2009) Cotranslational and Posttranslational N-Glycosylation of Polypeptides by Distinct Mammalian OST Isoforms. *Cell*, **136**, 272–283.
- Saito, F., Suyama, A., Oka, T., Yoko-o, T., Matsuoka, K., Jigami, Y., and Shimma, Y. (2014) Identification of Novel Peptidyl Serine  $\alpha$ -Galactosyltransferase Gene Family in Plants  $\blacklozenge$ . *J. Biol. Chem.*, **289**, 20405–20420.
- Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
- Samuelson, J., Banerjee, S., Magnelli, P., Cui, J., Kelleher, D.J., Gilmore, R., and Robbins, P.W. (2005) The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 1548–1553.
- Sandrin, M.S., Vaughan, H.A., Dabkowski, P.L., and McKenzie, I.F. (1993) Anti-pig IgM antibodies in human serum react predominantly with Gal( $\alpha$ 1-3)Gal epitopes. *Proc. Natl. Acad. Sci. U. S. A.*, **90**, 11391–11395.
- Sariyatun, R., Florence, Kajiura, H., Ohashi, T., Misaki, R., and Fujiyama, K. (2021) Production of Human Acid-Alpha Glucosidase With a Paucimannose Structure by Glycoengineered Arabidopsis Cell Culture. *Front. Plant Sci.*, **12**.
- Sasso, S., Stibor, H., Mittag, M., and Grossman, A.R. (2018) From molecular manipulation of domesticated *Chlamydomonas reinhardtii* to survival in nature. *eLife*, **7**, e39233.
- von Schaewen, A., Sturm, A., O'Neill, J., and Chrispeels, M.J. (1993) Isolation of a mutant Arabidopsis plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. *Plant Physiol.*, **102**, 1109–1118.
- Schähs, M., Strasser, R., Stadlmann, J., Kunert, R., Rademacher, T., and Steinkellner, H. (2007) Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. *Plant Biotechnol. J.*, **5**, 657–663.
- Schneider, J.D., Castilho, A., Neumann, L., Altmann, F., Loos, A., Kannan, L., et al. (2014) Expression of human butyrylcholinesterase with an engineered glycosylation profile resembling the plasma-derived orthologue. *Biotechnol. J.*, **9**, 501–510.
- Schnell, R.A. and Lefebvre, P.A. (1993) Isolation of the *Chlamydomonas* regulatory gene NIT2 by transposon tagging. *Genetics*, **134**, 737–747.
- Schoberer, J., Liebming, E., Vavra, U., Veit, C., Grünwald-Gruber, C., Altmann, F., et al. (2019) The Golgi Localization of GnTI Requires a Polar Amino Acid Residue within Its Transmembrane Domain1[OPEN]. *Plant Physiol.*, **180**, 859–873.
- Schoberer, J. and Strasser, R. (2011) Sub-compartmental organization of Golgi-resident N-glycan processing enzymes in plants. *Mol. Plant*, **4**, 220–228.
- Schroda, M., Blöcker, D., and Beck, C.F. (2000) The HSP70A promoter as a tool for the improved expression of transgenes in *Chlamydomonas*. *Plant J. Cell Mol. Biol.*, **21**, 121–131.
- Schroda, M. and Remacle, C. (2022) Molecular Advancements Establishing *Chlamydomonas* as a Host for Biotechnological Exploitation. *Front. Plant Sci.*, **13**.
- Schulze, S., Oltmanns, A., Fufezan, C., Krägenbring, J., Mormann, M., Pohlschröder, M., and Hippler, M. (2020) SugarPy facilitates the universal, discovery-driven analysis of intact glycopeptides. *Bioinformatics*.

- Schulze, S., Oltmanns, A., Machnik, N., Liu, G., Xu, N., Jarmatz, N., et al. (2018) N-Glycoproteomic Characterization of Mannosidase and Xylosyltransferase Mutant Strains of *Chlamydomonas reinhardtii*. *Plant Physiol.*, **176**, 1952–1964.
- Schwarz, F. and Aebi, M. (2011) Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.*, **21**, 576–582.
- Schwientek, T., Lorenz, C., and Ernst, J.F. (1995) Golgi Localization in Yeast Is Mediated by the Membrane Anchor Region of Rat Liver Sialyltransferase (\*). *J. Biol. Chem.*, **270**, 5483–5489.
- Selas Castiñeiras, T., Williams, S.G., Hitchcock, A.G., and Smith, D.C. (2018) *E. coli* strain engineering for the production of advanced biopharmaceutical products. *FEMS Microbiol. Lett.*, **365**, fny162.
- Sethuraman, N. and Stadheim, T.A. (2006) Challenges in therapeutic glycoprotein production. *Curr. Opin. Biotechnol.*, **17**, 341–346.
- Shailubhai, K., Pukazhenti, B.S., Saxena, E.S., Varma, G.M., and Vijay, I.K. (1991) Glucosidase I, a transmembrane endoplasmic reticular glycoprotein with a luminal catalytic domain. *J. Biol. Chem.*, **266**, 16587–16593.
- Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.G., Palmer, A.E., and Tsien, R.Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.*, **22**, 1567–1572.
- Shao, N. and Bock, R. (2008) A codon-optimized luciferase from *Gaussia princeps* facilitates the in vivo monitoring of gene expression in the model alga *Chlamydomonas reinhardtii*. *Curr. Genet.*, **53**, 381–388.
- Sharpe, H.J., Stevens, T.J., and Munro, S. (2010) A Comprehensive Comparison of Transmembrane Domains Reveals Organelle-Specific Properties. *Cell*, **142**, 158–169.
- Shin, Y.-J., Castilho, A., Dicker, M., Sádio, F., Vavra, U., Grünwald-Gruber, C., et al. (2017) Reduced paucimannosidic N-glycan formation by suppression of a specific  $\beta$ -hexosaminidase from *Nicotiana benthamiana*. *Plant Biotechnol. J.*, **15**, 197–206.
- Shrimal, S., Cherepanova, N.A., and Gilmore, R. (2017) DC2 and KCP2 mediate the interaction between the oligosaccharyltransferase and the ER translocon. *J. Cell Biol.*, **216**, 3625–3638.
- Sijmons, P.C., Dekker, B.M., Schrammeijer, B., Verwoerd, T.C., van den Elzen, P.J., and Hoekema, A. (1990) Production of correctly processed human serum albumin in transgenic plants. *Biotechnol. Nat. Publ. Co.*, **8**, 217–221.
- Singleton, V.L., Bohonos, N., and Ullstrup, A.J. (1958) Decumbin, a New Compound from a Species of *Penicillium*. *Nature*, **181**, 1072–1073.
- Sizova, I., Greiner, A., Awasthi, M., Kateriya, S., and Hegemann, P. (2013) Nuclear gene targeting in *Chlamydomonas* using engineered zinc-finger nucleases. *Plant J. Cell Mol. Biol.*, **73**, 873–882.
- Sizova, I., Kelterborn, S., Verbenko, V., Kateriya, S., and Hegemann, P. (2021) *Chlamydomonas* POLQ is necessary for CRISPR/Cas9-mediated gene targeting. *G3 GenesGenomesGenetics*, **11**, jkab114.
- Slattery, S.S., Giguere, D.J., Stuckless, E.E., Shrestha, A., Briere, L.-A.K., Galbraith, A., et al. (2022) Phosphate-regulated expression of the SARS-CoV-2 receptor-binding domain in the diatom *Phaeodactylum tricorutum* for pandemic diagnostics. *Sci. Rep.*, **12**, 7010.
- Sousa, M.C., Ferrero-Garcia, M.A., and Parodi, A.J. (1992) Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *Biochemistry*, **31**, 97–105.
- Specht, E., Miyake-Stoner, S., and Mayfield, S. (2010) Micro-algae come of age as a platform for recombinant protein production. *Biotechnol. Lett.*, **32**, 1373–1383.
- Steentoft, C., Bennett, E.P., Schjoldager, K.T.-B.G., Vakhrushev, S.Y., Wandall, H.H., and Clausen, H. (2014) Precision genome editing: A small revolution for glycobiology. *Glycobiology*, **24**, 663–680.

- Stevens, D., Milani-Nejad, S., and Mozaffar, T. (2022) Pompe Disease: a Clinical, Diagnostic, and Therapeutic Overview. *Curr. Treat. Options Neurol.*, **24**, 573–588.
- Stoker, M. and Macpherson, I. (1964) Syrian Hamster Fibroblast Cell Line BHK21 and its Derivatives. *Nature*, **203**, 1355–1357.
- Stolz, J. and Munro, S. (2002) The Components of the *Saccharomyces cerevisiae* Mannosyltransferase Complex M-Pol I Have Distinct Functions in Mannan Synthesis \*. *J. Biol. Chem.*, **277**, 44801–44808.
- Strasser, R. (2014) Biological significance of complex N-glycans in plants and their impact on plant physiology. *Front. Plant Sci.*, **5**, 363.
- Strasser, R. (2023) Plant glycoengineering for designing next-generation vaccines and therapeutic proteins. *Biotechnol. Adv.*, **67**, 108197.
- Strasser, R. (2016) Plant protein glycosylation. *Glycobiology*, **26**, 926–939.
- Strasser, R., Altmann, F., Mach, L., Glössl, J., and Steinkellner, H. (2004) Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking beta1,2-linked xylose and core alpha1,3-linked fucose. *FEBS Lett.*, **561**, 132–136.
- Strasser, R., Bondili, J.S., Schoberer, J., Svoboda, B., Liebminger, E., Glössl, J., et al. (2007) Enzymatic Properties and Subcellular Localization of *Arabidopsis*  $\beta$ -N-Acetylhexosaminidases. *Plant Physiol.*, **145**, 5–16.
- Strasser, R., Bondili, J.S., Vavra, U., Schoberer, J., Svoboda, B., Glössl, J., et al. (2007) A unique beta1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell*, **19**, 2278–2292.
- Strasser, R., Castilho, A., Stadlmann, J., Kunert, R., Quendler, H., Gattinger, P., et al. (2009a) Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous beta1,4-galactosylated N-glycan profile. *J. Biol. Chem.*, **284**, 20479–20485.
- Strasser, R., Castilho, A., Stadlmann, J., Kunert, R., Quendler, H., Gattinger, P., et al. (2009b) Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous beta1,4-galactosylated N-glycan profile. *J. Biol. Chem.*, **284**, 20479–20485.
- Strasser, R., Schoberer, J., Jin, C., Glössl, J., Mach, L., and Steinkellner, H. (2006) Molecular cloning and characterization of *Arabidopsis thaliana* Golgi alpha-mannosidase II, a key enzyme in the formation of complex N-glycans in plants. *Plant J. Cell Mol. Biol.*, **45**, 789–803.
- Strasser, R., Stadlmann, J., Schähs, M., Stiegler, G., Quendler, H., Mach, L., et al. (2008) Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.*, **6**, 392–402.
- Strasser, R., Stadlmann, J., Svoboda, B., Altmann, F., Glössl, J., and Mach, L. (2005) Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. *Biochem. J.*, **387**, 385–391.
- Su, H., van Eerde, A., Rimstad, E., Bock, R., Branza-Nichita, N., Yakovlev, I.A., and Clarke, J.L. (2023) Plant-made vaccines against viral diseases in humans and farm animals. *Front. Plant Sci.*, **14**.
- Su, Z.-L., Qian, K.-X., Tan, C.-P., Meng, C.-X., and Qin, S. (2005) Recombination and heterologous expression of allophycocyanin gene in the chloroplast of *Chlamydomonas reinhardtii*. *Acta Biochim. Biophys. Sin.*, **37**, 709–712.
- Sun, M., Qian, K., Su, N., Chang, H., Liu, J., Shen, G., and Chen, G. (2003) Foot-and-mouth disease virus VP1 protein fused with cholera toxin B subunit expressed in *Chlamydomonas reinhardtii* chloroplast. *Biotechnol. Lett.*, **25**, 1087–1092.
- Surzycki, R., Greenham, K., Kitayama, K., Dibal, F., Wagner, R., Rochaix, J.-D., et al. (2009) Factors effecting expression of vaccines in microalgae. *Biol. J. Int. Assoc. Biol. Stand.*, **37**, 133–138.

- Suzuki, T., Cummings, R.D., Aebi, M., and Parodi, A. (2022) Glycans in Glycoprotein Quality Control. In: *Essentials of Glycobiology* (Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., et al., eds). Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
- Swiezewska, E. and Danikiewicz, W. (2005) Polyisoprenoids: Structure, biosynthesis and function. *Prog. Lipid Res.*, **44**, 235–258.
- Takouridis, S.J., Tribe, D.E., Gras, S.L., and Martin, G.J.O. (2015) The selective breeding of the freshwater microalga *Chlamydomonas reinhardtii* for growth in salinity. *Bioresour. Technol.*, **184**, 18–22.
- Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., and Muchmore, E. (2003) Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 12045–12050.
- Taunt, H.N., Stoffels, L., and Purton, S. (2017) Green biologics: The algal chloroplast as a platform for making biopharmaceuticals. *Bioengineered*, **9**, 48–54.
- Tearle, R.G., Tange, M.J., Zannettino, Z.L., Katerelos, M., Shinkel, T.A., Van Denderen, B.J., et al. (1996) The alpha-1,3-galactosyltransferase knockout mouse. Implications for xenotransplantation. *Transplantation*, **61**, 13–19.
- Tekoah, Y., Tzaban, S., Kizhner, T., Hainrichson, M., Gantman, A., Golembo, M., et al. (2013) Glycosylation and functionality of recombinant  $\beta$ -glucocerebrosidase from various production systems. *Biosci. Rep.*, **33**, e00071.
- Thompson, N. and Wakarchuk, W. (2022) O-glycosylation and its role in therapeutic proteins. *Biosci. Rep.*, **42**, BSR20220094.
- Torzillo, G., Scoma, A., Faraloni, C., and Giannelli, L. (2015) Advances in the biotechnology of hydrogen production with the microalga *Chlamydomonas reinhardtii*. *Crit. Rev. Biotechnol.*, **35**, 485–496.
- Tran, M., Henry, R.E., Siefker, D., Van, C., Newkirk, G., Kim, J., et al. (2013) Production of anti-cancer immunotoxins in algae: ribosome inactivating proteins as fusion partners. *Biotechnol. Bioeng.*, **110**, 2826–2835.
- Tran, M., Van, C., Barrera, D.J., Pettersson, P.L., Peinado, C.D., Bui, J., and Mayfield, S.P. (2013) Production of unique immunotoxin cancer therapeutics in algal chloroplasts. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, E15–22.
- Tran, M., Zhou, B., Pettersson, P.L., Gonzalez, M.J., and Mayfield, S.P. (2009) Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. *Biotechnol. Bioeng.*, **104**, 663–673.
- Tretter, V., Altmann, F., and März, L. (1991) Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached alpha 1----3 to the asparagine-linked N-acetylglucosamine residue. *Eur. J. Biochem.*, **199**, 647–652.
- Trombetta, E.S., Simons, J.F., and Helenius, A. (1996) Endoplasmic reticulum glucosidase II is composed of a catalytic subunit, conserved from yeast to mammals, and a tightly bound noncatalytic HDEL-containing subunit. *J. Biol. Chem.*, **271**, 27509–27516.
- Vallon, O. and Wollman, F.A. (1995) Mutations Affecting O-Glycosylation in *Chlamydomonas reinhardtii* Cause Delayed Cell Wall Degradation and Sex-Limited Sterility. *Plant Physiol.*, **108**, 703–712.
- Van Patten, S.M., Hughes, H., Huff, M.R., Piepenhagen, P.A., Waire, J., Qiu, H., et al. (2007) Effect of mannose chain length on targeting of glucocerebrosidase for enzyme replacement therapy of Gaucher disease. *Glycobiology*, **17**, 467–478.
- Vanier, G., Lucas, P.-L., Loutelier-Bourhis, C., Vanier, J., Plasson, C., Walet-Balieu, M.-L., et al. (2017) Heterologous expression of the N-acetylglucosaminyltransferase I dictates a reinvestigation of the N-glycosylation pathway in *Chlamydomonas reinhardtii*. *Sci. Rep.*, **7**, 10156.
- Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., and Marth, J. (1999) O-Glycans. In: *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press.

- Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., et al. eds. (2009) *Essentials of Glycobiology*, 2nd edn. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
- Vassilakos, A., Michalak, M., Lehrman, M.A., and Williams, D.B. (1998) Oligosaccharide binding characteristics of the molecular chaperones calnexin and calreticulin. *Biochemistry*, **37**, 3480–3490.
- van der Veen, S.J., Hollak, C.E.M., van Kuilenburg, A.B.P., and Langeveld, M. (2020) Developments in the treatment of Fabry disease. *J. Inherit. Metab. Dis.*, **43**, 908–921.
- Vervecken, W., Kaigorodov, V., Callewaert, N., Geysens, S., De Vusser, K., and Contreras, R. (2004) In vivo synthesis of mammalian-like, hybrid-type N-glycans in *Pichia pastoris*. *Appl. Environ. Microbiol.*, **70**, 2639–2646.
- Vilarrasa-Blasi, J., Velloso, T., Jinkerson, R.E., Fauser, F., Xiang, T., Minkoff, B.B., et al. (2021) Identification of green lineage osmotic stress pathways. 2021.07.19.453009.
- Wakao, S., Shih, P.M., Guan, K., Schackwitz, W., Ye, J., Patel, D., et al. (2021) Discovery of photosynthesis genes through whole-genome sequencing of acetate-requiring mutants of *Chlamydomonas reinhardtii*. *PLoS Genet.*, **17**, e1009725.
- Walsh, G. (2002) Biopharmaceuticals and biotechnology medicines: an issue of nomenclature. *Eur. J. Pharm. Sci. Off. J. Eur. Fed. Pharm. Sci.*, **15**, 135–138.
- Walsh, G. (2010) Post-translational modifications of protein biopharmaceuticals. *Drug Discov. Today*, **15**, 773–780.
- Walsh, G. and Walsh, E. (2022) Biopharmaceutical benchmarks 2022. *Nat. Biotechnol.*, **40**, 1722–1760.
- Wandall, H.H., Nielsen, M.A.I., King-Smith, S., de Haan, N., and Bagdonaite, I. (2021) Global functions of O-glycosylation: promises and challenges in O-glycobiology. *FEBS J.*, **288**, 7183–7212.
- Wang, M., Ye, X., Bi, H., and Shen, Z. (2024) Microalgae biofuels: illuminating the path to a sustainable future amidst challenges and opportunities. *Biotechnol. Biofuels Bioprod.*, **17**, 10.
- Wang, P., Wang, H., Gai, J., Tian, X., Zhang, X., Lv, Y., and Jian, Y. (2017) Evolution of protein N-glycosylation process in Golgi apparatus which shapes diversity of protein N-glycan structures in plants, animals and fungi. *Sci. Rep.*, **7**, 40301.
- Wang, X., Brandsma, M., Tremblay, R., Maxwell, D., Jevnikar, A.M., Huner, N., and Ma, S. (2008) A novel expression platform for the production of diabetes-associated autoantigen human glutamic acid decarboxylase (hGAD65). *BMC Biotechnol.*, **8**, 87.
- Wang, X.-H., Nakayama, K., Shimma, Y., Tanaka, A., and Jigami, Y. (1997) MNN6, a Member of the KRE2/MNT1 Family, Is the Gene for Mannosylphosphate Transfer in *Saccharomyces cerevisiae* \*. *J. Biol. Chem.*, **272**, 18117–18124.
- Wannathong, T., Waterhouse, J.C., Young, R.E.B., Economou, C.K., and Purton, S. (2016) New tools for chloroplast genetic engineering allow the synthesis of human growth hormone in the green alga *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.*, **100**, 5467–5477.
- Ward, B.J., Gobeil, P., Séguin, A., Atkins, J., Boulay, I., Charbonneau, P.-Y., et al. (2021) Phase 1 randomized trial of a plant-derived virus-like particle vaccine for COVID-19. *Nat. Med.*, **27**, 1071–1078.
- Ward, S., O’Sullivan, J.M., and O’Donnell, J.S. (2021) The Biological Significance of von Willebrand Factor O-Linked Glycosylation. *Semin. Thromb. Hemost.*, **47**, 855–861.
- Welch, L.G. and Munro, S. (2019) A tale of short tails, through thick and thin: investigating the sorting mechanisms of Golgi enzymes. *FEBS Lett.*, **593**, 2452–2465.
- Wiggins, C.A.R. and Munro, S. (1998) Activity of the yeast MNN1  $\alpha$ -1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc. Natl. Acad. Sci.*, **95**, 7945–7950.

- Wilson, I.B.H., Rendić, D., Freilinger, A., Dumić, J., Altmann, F., Mucha, J., et al. (2001) Cloning and expression of cDNAs encoding  $\alpha$ 1,3-fucosyltransferase homologues from *Arabidopsis thaliana*. The cDNA sequences referred to in this publication have been deposited with the EMBL database under the numbers AJ404860 (FucTA), AJ404861 (FucTB) and AJ404862 (FucTC). *Biochim. Biophys. Acta BBA - Gen. Subj.*, **1527**, 88–96.
- Yang, Z., Li, Yinü, Chen, F., Li, D., Zhang, Z., Liu, Y., et al. (2006) Expression of human soluble TRAIL in *Chlamydomonas reinhardtii* chloroplast. *Chin. Sci. Bull.*, **51**, 1703–1709.
- Yoon, S.-M., Kim, S.Y., Li, K.F., Yoon, B.H., Choe, S., and Kuo, M.M.-C. (2011) Transgenic microalgae expressing *Escherichia coli* AppA phytase as feed additive to reduce phytate excretion in the manure of young broiler chicks. *Appl. Microbiol. Biotechnol.*, **91**, 553–563.
- Zhang, M.-P., Wang, M., and Wang, C. (2020) Nuclear transformation of *Chlamydomonas reinhardtii*: A review. *Biochimie*.
- Zhang, P., Wang, T., Bardor, M., and Song, Z. (2013) Deciphering O-glycomics for the development and production of biopharmaceuticals. *Pharm. Bioprocess.*, **1**, 89–104.
- Zhang, R., Patena, W., Armbruster, U., Gang, S.S., Blum, S.R., and Jonikas, M.C. (2014) High-Throughput Genotyping of Green Algal Mutants Reveals Random Distribution of Mutagenic Insertion Sites and Endonucleolytic Cleavage of Transforming DNA. *Plant Cell*, **26**, 1398–1409.
- Zhang, Y.-K., Shen, G.-F., and Ru, B.-G. (2006) Survival of human metallothionein-2 transplastomic *Chlamydomonas reinhardtii* to ultraviolet B exposure. *Acta Biochim. Biophys. Sin.*, **38**, 187–193.
- Zimran, A., Gonzalez-Rodriguez, D.E., Abrahamov, A., Cooper, P.A., Varughese, S., Giraldo, P., et al. (2018) Long-term safety and efficacy of taliglucerase alfa in pediatric Gaucher disease patients who were treatment-naïve or previously treated with imiglucerase. *Blood Cells. Mol. Dis.*, **68**, 163–172.
- Zinzius, K., Marchetti, G.M., Fischer, R., Milrad, Y., Oltmanns, A., Kelterborn, S., et al. (2023) Calredoxin regulates the chloroplast NADPH-dependent thioredoxin reductase in *Chlamydomonas reinhardtii*. *Plant Physiol.*, **193**, 2122–2140.



## Résumé

Le marché des biomédicaments atteindra 389 milliards de dollars en 2024. Sur ce marché, les glycoprotéines recombinantes représentent à ce jour environ 70% des biomédicaments commercialisés. La production de ces molécules avec les modèles mammifères actuels tels que les cellules CHO (Chinese Hamster Ovary) s'avère onéreuse. En effet, le risque important de contamination des cultures de cellules mammifères par des agents pathogènes et leurs conditions de croissance exigeantes justifient la recherche et le développement de modèles de production alternatifs. Cependant, il est bien connu que les modifications post-traductionnelles, comme la *N*-glycosylation, ont un impact sur les paramètres pharmacocinétiques et la demi-vie des protéines thérapeutiques. Ainsi, il est nécessaire de s'assurer que les modifications post-traductionnelles engendrées par le système hôte soient compatibles avec une utilisation thérapeutique humaine des protéines produites de manière hétérologue. Parmi les microalgues étudiées jusqu'à présent comme systèmes d'expression alternatifs des produits biopharmaceutiques, la microalgue verte *Chlamydomonas reinhardtii* est un modèle qui justifie les nombreux moyens déployés pour le développement d'outils moléculaires pour l'ingénierie de ses voies métaboliques. Bien qu'à ce jour la majorité des protéines recombinantes produites en microalgue l'a été dans ce modèle, leur glycosylation n'a pas été étudiée. De plus, l'analyse des glycanes portés par les glycoprotéines endogènes a montré que la voie de *N*-glycosylation se distingue de celle des mammifères. Dans ce contexte, une tentative de glycoingénierie de la microalgue verte *C. reinhardtii* a été réalisée afin d'optimiser la production de glycoprotéines recombinantes thérapeutiques. Dans un premier temps, l'érythropoïétine humaine, un biomédicament utilisé dans le cadre de ce travail comme glycoprotéine rapportrice, a été exprimée dans différents mutants de *N*-glycosylation puis caractérisée. Cette étude a permis de mettre en évidence que *C. reinhardtii* est capable de glycosyler ce type de molécules en accord avec le profil *N*-glycannique de la souche. Par la suite, une élimination par édition génomique des gènes codant des glycosyltransférases responsables de l'ajout des résidus immunogènes que sont les xyloses et fucoses a été entreprise. Enfin, une étude de l'adressage des glycoenzymes impliquées dans les étapes de maturation dans l'appareil de Golgi via les domaines CTS a été amorcée dans le but de compléter la voie de *N*-glycosylation en introduisant et exprimant des glycoenzymes hétérologues. Les résultats obtenus dans les deux dernières parties constituent des travaux préliminaires ouvrant la voie vers des projets en cours de développement.

**Mots-clés:** Biomédicaments, Protéines, *Chlamydomonas reinhardtii*, *N*-glycosylation, glyco-ingénierie.



## **Abstract**

The market for biomedicines will reach \$389 billion by 2024. Recombinant glycoproteins currently account for around 70% of the biomedicines on the market. The production of these molecules using current mammalian models such as CHO (Chinese Hamster Ovary) cells is proving expensive. The significant risk of contamination of mammalian cell cultures by pathogens and their demanding growth conditions justify the research and development of alternative production models. However, it is well known that post-translational modifications, such as *N*-glycosylation, have an impact on the pharmacokinetic parameters and half-life of therapeutic proteins. It is therefore necessary to ensure that the post-translational modifications generated by the host system are compatible with the therapeutic use of heterologously produced proteins in humans. Among the microalgae studied to date as alternative expression systems for biopharmaceuticals, the green microalga *Chlamydomonas reinhardtii* is a model that justifies the many resources deployed to develop molecular tools for engineering its metabolic pathways. Although the majority of recombinant proteins produced in microalgae have been produced in this model, their glycosylation has not been studied. Furthermore, analysis of the glycans carried by endogenous glycoproteins has shown that the *N*-glycosylation pathway differs from that of mammals. In this context, an attempt was made to glycoengineer the green microalga *C. reinhardtii* in order to optimise the production of therapeutic recombinant glycoproteins. Firstly, human erythropoietin, a biomedicine used in this work as a glycoprotein reporter, was expressed in different *N*-glycosylation mutants and then characterised. This study demonstrated that *C. reinhardtii* is capable of glycosylating this type of molecule in accordance with the strain's *N*-glycan profile. Genomic editing was then used to eliminate the genes encoding the glycosyltransferases responsible for adding the immunogenic residues xyloses and fucoses. Finally, a study of the addressing of glycoenzymes involved in maturation steps in the Golgi apparatus via CTS domains was initiated with the aim of complementing the *N*-glycosylation pathway by introducing and expressing heterologous glycoenzymes. The results obtained in the last two sections constitute preliminary work paving the way for projects currently under development.

**Key words:** Biopharmaceuticals, Proteins, *Chlamydomonas reinhardtii*, *N*-glycosylation, glycoengineering.