Journal of Biological Chemistry
February 2011, Vol. 286 (8), Pages 6152-6164
http//dx.doi.org/10.1074/jbc.M110.175711

N-Glycans of *Phaeodactylum tricornutum* Diatom and Functional Characterization of Its *N*-Acetylglucosaminyltransferase I Enzyme*

Bérengère Baïet[1], Carole Burel[1], Bruno Saint-Jean[2], Romain Louvet[1], Laurence Menu-Bouaouiche[1], Marie-Christine Kiefer-Meyer[1], Elodie Mathieu-Rivet[1], Thomas Lefebvre[3], Hélène Castel[3], Aude Carlier[2,4], Jean-Paul Cadoret[2], Patrice Lerouge[1] and Muriel Bardor[1, 5, *]

[1] Université de Rouen, Laboratoire Glycobiologie et Matrice Extracellulaire Végétale (Glyco-MEV), EA 4358, IFRMP23, Faculté des Sciences, 76821 Mont-Saint-Aignan Cédex, France.
[2] Laboratoire Physiologie et Biotechnologies des Algues, IFREMER, Centre de Nantes, Rue de l'Ile d'Yeu, BP 21105, 44311 Nantes Cedex 03, France.
[3] Université de Rouen, Laboratoire de Différenciation et Communication Neuronale et Neuroendocrine, Inserm U982, IFRMP23, Faculté des Sciences, 76821 Mont-Saint-Aignan Cedex, France.
[4] Present adress: Algenics S. A. S., Pôle Bio Ouest, rue du moulin de la rousselière, 44800 Saint Herblain, France.
[5] Present adress: Bioprocessing Technology Institute (BTI), Agency for Science Technology and Research (A*STAR), 20 Biopolis Way, #06-01, Centros, Singapore, 138668, Singapore.

[*] Corresponding author : Dr Muriel Bardor, Phone: (33) 2 35 14 67 58, Fax: (33) 2 35 14 66 15, email address : muriel.bardor@univ-rouen.fr

* This work was supported by funds from the University of Rouen (to P. L. and M. B.).

Abstract:

N-Glycosylation, a major co- and post-translational event in the synthesis of proteins in eukaryotes, is unknown in aquatic photosynthetic microalgae. In this paper, we describe the *N*-glycosylation pathway in the diatom *Phaeodactylum tricornutum*. Bio-informatic analysis of its genome revealed the presence of a complete set of sequences potentially encoding for proteins involved in the synthesis of the lipid-linked $Glc_3Man_9GlcNAc_2$-PP-dolichol *N*-glycan, some subunits of the oligosaccharyltransferase complex, as well as endoplasmic reticulum glucosidases and chaperones required for protein quality control and, finally, the α-mannosidase I involved in the trimming of the *N*-glycan precursor into Man-5 *N*-glycan. Moreover, one *N*-acetylglucosaminyltransferase I, a Golgi glycosyltransferase that initiates the synthesis of complex type *N*-glycans, was predicted in the *P. tricornutum* genome. We demonstrated that this gene encodes for an active *N*-acetylglucosaminyltransferase I, which is able to restore complex type *N*-glycans maturation in the Chinese hamster ovary Lec1 mutant, defective in its endogeneous *N*-acetylglucosaminyltransferase I. Consistent with these data, the structural analyses of *N*-linked glycans demonstrated that *P. tricornutum* proteins carry mainly high mannose type *N*-glycans ranging from Man-5 to Man-9. Although representing a minor glycan population, paucimannose *N*-glycans were also detected, suggesting the occurrence of an *N*-acetylglucosaminyltransferase I-dependent maturation of *N*-glycans in this diatom.

Keywords: Carbohydrate Biosynthesis, Carbohydrate Structure, Glycosylation, Golgi, Plant, N-Acetylglucosaminyltransferase, Phaeodactylum, Diatom, Glycosyltransferase, Microalgae

Microalgae are a group of aquatic photosynthetic microorganisms which are so diverse that they have been gathered in a "paraphylum". Among these microalgae, diatoms belong to the heterokont group and are responsible for approximatively 40% of marine primary productivity (1, 2). Despite their physiological relevance in the marine ecosystem, molecular and cellular processes in diatom remain widely unknown. For example, so far, little is known about secretion, post-translational modifications and intracellular trafficking of proteins in diatoms. Diatom species are usually classified into two major groups, the bi/multipolar centrics and the pennates. Recently, the genome of a pennate diatom, Phaeodactylum tricornutum, became available (3) revealing a wealth of information about diatom biology. Access to this data (Joint Genome Institute, Walnut Creek, CA, USA, http://genome.jgi-psf.org), together with the fact that P. tricornutum is easy to culture in vitro and can be genetically transformed (4, 5), provide the opportunity to perform comparative genomic studies and to dissect biosynthetic pathways.

N-glycosylation is a major co- and post-translational modification in the synthesis of proteins in eukaryotes. N-glycan processing occurs in the secretory pathway and is essential for glycoproteins destined to be secreted or integrated in the membranes. In this process, a $Man_5GlcNAc_2$-PP-dolichol oligosaccharide intermediate is assembled by the stepwise addition of monosaccharides to dolichol pyrophosphate on the cytosolic face of the Endoplasmic Reticulum (ER). This intermediate is then extended in the lumen of the ER until a $Glc_3Man_9GlcNAc_2$-PP-dolichol N-glycan precursor is completed (6). This precursor is transferred by the oligosaccharyltransferase (OST) multisubunit complex onto the asparagine residue of the consensus Asn-X-Ser/Thr sequences of a target nascent protein (6). The precursor is then deglucosylated/reglucosylated to ensure the quality control of the neosynthetized protein through the interaction with ER-resident chaperones such as calnexin and calreticulin. These ER events are crucial for the proper folding of the secreted proteins and are highly conserved in eukaryotes investigated so far (7). In contrast, evolutionary adaptation of N-glycan processing in the Golgi apparatus has given rise to a large variety of organism-specific complex structures that allow the protein to carry out diverse glycan-mediated biological functions. Alpha mannosidase I (α-Man I) located in the early compartment of the Golgi apparatus (Cis cisternae) first degrade the oligosaccharide precursor into high mannose-type N-glycans ranging from $Man_9GlcNAc_2$ (Man-9) to $Man_5GlcNAc_2$ (Man-5). N-acetylglucosaminyltransferase I (GnT I) then transfers the first N-acetylglucosamine (GlcNAc) residue on the α(1,3)-mannose arm of $Man_5GlcNAc_2$ enabling the initiation of the synthesis of multiple structurally different complex-type N-glycans. Following GnT I action, alpha mannosidase II (α-Man II) and N-acetylglucosaminyltransferase II (GnT II) give rise to the synthesis of the core $GlcNAc_2Man_3GlcNAc_2$ which is finally maturated into organism-specific complex N-glycans by transfer of various monomers by characteristic glycosyltransferases. GnT I and thus the GnT I-dependent maturation of N-glycans appeared during evolution at the same period as metazoans (8). Complex-type N-glycans were demonstrated to be engaged in crucial steps of the development of pluricellular organisms (8-11). For instance, GnT I-null embryos of mice die at about 10 days after fertilization, indicating that complex N-glycans are required for morphogenesis in mammals (9, 10). Similarly, inactivation of the GnT I in worm and fly reduces their viabilities (8, 11). In plants, *Arabidopsis cgl* mutant, defective in GnT I activity, was demonstrated to grow normally in standard culture conditions (12). However, these plants exhibited a strong phenotype in salt induced stress conditions for example, suggesting a role for mature plant N-glycans in specific physiological processes (13-15). In contrast, animal cultured cells having GnT I null mutations usually grow normally (16).

While major data regarding protein N-glycosylation have been established in yeast and higher eukaryotes, nothing is known on N-glycan biosynthesis and structures in microalgae. In this paper, we describe the N-glycosylation pathway in the diatom *P. tricornutum.* We also demonstrate that the predicted GnT I from *P. tricornutum* is able to complement the biosynthesis of complex-type N-glycans in the Chinese Hamster Ovary (CHO) Lec1 mutant which is defective in its endogenous GnT I. To the best of our knowledge, this is the first functional characterisation of a N-glycan glycosyltransferase from microalgae.

EXPERIMENTAL PROCEDURES

Materials – TRIzol reagent, anti-V5 antibodies, alpha Minimum Essential Medium (MEM), Fetal Bovine Serum (FBS), Geneticin and pcDNA3.1/V5-His-TOPO TA kit were purchased from Invitrogen (Uppsala, Sweden). Silica membrane spin column and RNeasy Mini Kit were obtained from Qiagen (Courtaboeuf, France). RQ1 DNAse was from Promega (Charbonnières-les-bains, France). High Capacity cDNA Reverse Transcription Kit, 2x *Power* SYBR® Green I PCR Master Mix, Sequazyme™ Peptide Mass Standards Kit as well as the Voyager DE-Pro MALDI-TOF instrument were from Applied Biosystems (Foster City, CA, USA). CHO K1 (wild-type) and CHO Lec 1 cell lines were obtained from American Type Culture Collection (Manassas, VA, USA). Nucleofactor device was from Amaxa (Köln, Germany). Pierce® BCA Protein Assay Kit and Extensor Hi-Fidelity PCR Enzyme Mix were purchased from Thermo Scientific (Rockford, IL, USA). Pepsin, Jack

bean alpha-mannosidase proteomic grade, alpha-L-fucosidase from bovine kidney, 2,5-dihydroxybenzoic acid (DHB), 1,2-diamino-4,5-methylene dioxybenzene (DMB), UDP-diphosphogalactose, 2-aminobenzamide (2-AB), β(1,4)-galactosyltransferase from bovine milk, Nonidet P40 (Igepal CA-630) were obtained from Sigma-Aldrich (Saint-Quentin Fallavier, France). Biotinylated lectins such as Phytohemagglutinin E and L (PHA-E and L), *Ricinus Communis* Agglutinin 120 (RCA 120), *Erythrina Crista Galli* Agglutinin (ECA) and peanut agglutinin (PNA) were from Vector Laboratories Inc. (Bulingame, CA, USA). Concanavalin A (ConA) and Streptavidine horseradish peroxidase conjugate were from GE Healthcare (Buckinghamshire, UK). ElectroChimioLuminescent (ECL) revelation Kit was purchased from Amersham Biosciences (Uppsala, Sweden). Antibodies directed against core anti-β(1,2)-xylose and core α(1,3)-fucose were from AgriSera (Vännäs, Sweden). Peptide N-glycosidase F (PNGase F), Peptide N-glycosidase A (PNGase A) and Endoglycosidase H (Endo H) were from Roche (Basel, Switzerland). C18 column (C18 monomeric S/N E000930-10-2) and C18 cartridge were from Varian Inc. (Palo Alto, CA, USA). Carbograph cartridges were from Alltech/Grace Davison Discovery Sciences (Templemars, France). NanoDrop® ND-1000 was from Thermo Fisher (Illkirch, France). Mx3000PTM Q-PCR system was from Stratagene (La Jolla, CA, USA). *Fastprep* 24 device and 50 mL BigD Lysing tubes were purchased from MP Biomedicals (Illkirch, France).

In silico genome analysis - In the *P. tricornutum* genome (http://genome.jgi-psf.org/Phatr2/Phatr2.home.html), annotation of genes involved in the N-glycan pathway was carried out by BLASTP or TBLASTN analyses with genomic sequences from *Homo sapiens, Mus musculus, Arabidopsis thaliana, Drosophila melanogaster, Saccharomyces cerevisiae, Physcomitrella patens, Medicago truncatula, Zea mays, Nicotiana plumbaginifolia* and *Oryza sativa*. Searches for signal peptides and cell localisation/targeting of mature proteins were done using Signal P, Signal-BLAST and Target P. Transmembrane domains were predicted using TMHMM, TOPPRED and HMMTOP. Pfam domains were identified using Pfam (Wellcome Trust Sanger Institute, Cambridge, UK). The phylogenetic tree was drawn using the *Phylogeny.fr* platform (17) and following three steps: i) complete sequences were aligned with ClustalW (v2.0.3) (18), ii) after alignment, ambiguous regions (i.e. containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b) (19) and iii) the phylogenetic tree was built using the maximum likelihood method implemented in the PhyML program (v3.0 aLRT) (20, 21). Graphical representation and edition of the phylogenetic tree were performed with TreeDyn (v198.3) (22). Thirty one sequences were selected from the CAZy GT13 family (23) or retrieved from the JGI database (http://www.jgi.doe.gov/). Origin and accession numbers for the different GnT I protein sequences are as follows: *Arabidopsis thaliana*, MGAT1_ARATH; *Bos Taurus*, Q5E9I4_BOVIN; *Caenorhabditis elegans*, MGAT1_CAEEL; *Cricetulus griseus*, P70680_CRIGR; *Danio rerio* (1), A0JMK2_DANRE; *Danio rerio* (2), Q6PE41_ DANRE; *Drosophila melanogaster*, Q60GL7_DROME; *Glycine max*, C6T9Z3_SOYBN; *Homo sapiens*, MGAT1_HUMAN; *Mesaucricetus auratus*, Q9QWX4_MESAU; *Mus musculus*, MGAT1_MOUSE; *Nicotiana tabacum*, Q8VXD3_TOBAC; *Oryctolagus cuniculus*, MGAT1_RABIT; *Oryza sativa*, Q8RW24_ORYSA; *Populus trichocarpa*, B9HPX2_POPTR; *Physcomitrella patens*, Q8L5D3_PHYPA; *Pongo abelii*, Q5RB13_PONPY; *Rattus norvegicus*, MGAT1_RAT; *Salmo salar*, B5X2V8_SALSA; *Schistosoma mansoni*, C4Q4I7_SCHMA; *Solanum tuberosum*, Q8VX56_SOLTU; *Sorghum bicolor*, C5YS64_SORBI; *Sus scrofa*, Q0KKC2_PIG; *Xenopus laevis*, Q90W56_XENLA; *Zea mays*, B6TLH4_MAIZE. GenBank accession number: *Triticum aestivum*, ABT33168. JGI accessions: *Micromonas pusilla* CCMP1545 jgi|MicpuC2|32201|estExt_Genewise1Plus.C_30312); *Emiliania huxleyi* (jgi|Emihu1|460566|estExtDG_Genemark1.C_13800003); *Fragilariopsis cylindrus* (jgi|Fracy1|189180|e_gw1.9.153.1); *Phaeodactylum tricornutum* (GenBank TPA BK007891); *Thalassosira pseudonana* (jgi|Thaps3|1600| fgenesh1_pg.C_chr_1000734).

Microalgal strain and culture conditions - The strain of *P. tricornutum* P.t1.8.6 (CCAP1055/1) was grown in batch culture method using two liter flat bottom flasks. The nutritive medium used for this experiment consisted of natural

seawater, sterilized by filtration through a 0.22 µm filter, enriched in Conway medium (24) and containing 40 $mg.L^{-1}$ sodium metasilicate. Diatom cells were maintained at 20°C under continuous illumination (280–350 µmol photons $m^{-2}.s^{-1}$). Cells (20.10^6 $cells.mL^{-1}$) were then centrifuged at 5,000 g for 20 min at 4°C and the resulting pellet was freeze-dried before biochemical analyses. *P. tricornutum* were grown in continuous culture conditions as previously described (25) for real-time quantitative PCR experiments. At the steady-state (15 to 20.10^6 $cells.mL^{-1}$), five samples (30 mL of each in triplicate) were harvested at 4,000 g for 20 min at 4°C. The supernatant was removed rapidly and cell pellets were re-suspended in 1 mL of TRIzol and immediately frozen and stored at -80°C until RNA extraction.

Real-time quantitative PCR experiments – Total RNA was extracted from cells using TRIzol method and then treated with RQ1 DNase to avoid DNA contamination and finally purified using RNeasy Mini Kit. cDNA templates for PCR amplification were synthesized from 350 ng of total RNA using the High Capacity cDNA Reverse Transcription Kit. Quantitative PCR was performed using *Power* SYBR® Green I PCR Master Mix in a final volume reaction of 25 µL. All reactions were performed following the instructions of the manufacturer with 5 µL of diluted cDNA (1/10) and 0.1 µM of specific primers. Quantitative measurements were performed in duplicate with a Stratagene Mx3000P™ Q-PCR system. Cycling parameters were one cycle of 10 min at 95°C, followed by 40 cycles (30 s at 95°C, 60 s at 60°C). The results were represented as the relative gene expression normalized to reference genes encoding for Ribosomal Protein Small subunit 30 S (RPS) and histone H4 (26). Specific primers for catalytic domain of the *P. tricornutum* GnT I gene (GNT1-Q-Fwd: 5'-CGTACGAATCGCCCTTACTC-3' and GNT1-Q-Rev: 5'-TTGCCGTCTTGTGAAATTACC-3') were designed using Primer3Plus program (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). The relative GnT I gene expression analysis was performed using the method already described (27, 28) where the comparative C_T method (29) and standard curve method were combined to calculate RNA molar ratio between the target and housekeeping genes (28).

Expression of P. tricornutum GnT I in CHO cells – Genomic DNA was extracted from *P. tricornutum* cell pellets as described in (30). The GnT I gene was amplified from *P. tricornutum* genomic DNA using primers 5' – ATGCGGTTGTGGAAACGTAC – 3' and 5'-TCTTTTCGGTGACGGAATG -3' and Extensor Hi-Fidelity PCR Enzyme Mix. Then, this GnT I gene was cloned according to the supplier's instructions in the pcDNA3.1/V5-His-TOPO vector, leading to the expression of the *P. tricornutum* GnT I fused with the V5 epitope under the control of the T7 promoter. This construction pcDNA3.1/V5-His-TOPO-GnT I was sequenced and presented the following mutations: Gln130→Arg; Val148→Ile; Gly159→Ser; Lys196→Gln; Val337→Ala; Asn422→Tyr (GenBank BankIt 1370344). CHO Lec 1 mutant cells were transfected by electroporation with this construct. Cells were trypsinized (0.05%), triturated in alpha MEM containing 10% FBS, pelleted by centrifugation, resuspended in 100 µl of solution V for nucleofection by an Amaxa Nucleofector Device set to program U-016 with the linearized pcDNA3.1/V5-His-TOPO-GnT I and 500 µg of sterile sonicated salmon sperm DNA. Then, the transfection was followed by repetitive rounds of limiting dilution of cells in 400 $µg.mL^{-1}$ of geneticin for selection. CHO wild-type and CHO Lec 1 mutant were grown in alpha MEM medium supplemented with 10% FBS at 37°C in a humidified incubator with an atmosphere of 5% CO_2. CHO Lec1 mutant complemented with *P. tricornutum* GnT I was grown in the same conditions with 600 $µg.mL^{-1}$ of geneticin.

Extraction of proteins from P. tricornutum - Two g of lyophilised *P. tricornutum* cells were lysed in 750 mM Tris-HCl pH 8 buffer containing 15% (w/v) of sucrose, 2% (v/v) of β-mercaptoethanol and 1 mM phenylmethylsulfonylfluoride (extraction buffer). The cell lysis was done in a 50 mL BigD Lysing tubes and assisted by the *Fastprep* 24 (15 times 30 sec at 6 M.s-1). The mixture was then centrifuged at 4°C for 5 min at 4,000g. The pellet was washed once with 10mL of extraction buffer and spin again at 4°C for 5 min at 4,000g. The resulting supernatants were pooled prior to centrifugation at 15,000g during 30 minutes at 4°C. The supernatant was then dialysed against water during 48h at 4°C prior to lyophilisation. Protein quantification was then performed using the Pierce® BCA Protein Assay Kit and bovine serum albumin as protein standard. Proteins from green onion were prepared in parallel and used as a positif control for affino- and

immunodetection analyses.

Immuno- and affinoblotting analysis - Fifty µg of total protein extract from *P. tricornutum* were separated by SDS-PAGE using a 12% polyacrylamide gel. Proteins were transferred onto nitrocellulose membrane and stained with Ponceau Red in order to control the transfer efficiency. Affinodetections were carried out as previously described (31) using ConA or biotinylated lectins such as PHA-E and L, ECA, RCA 120 and PNA. Biotinylated lectins were detected using streptavidin coupled with horseradish peroxidase and concanavalin A was directly detected by horseradish peroxidase. Final developments of the blots were obtained using 4-chloro-1-naphtol or ECL as substrate. Immunodetections were performed using specific core β(1,2)-xylose and core α(1,3)-fucose antibodies as previously reported (32). Oxidation of the glycan moiety of glycoproteins was carried out on the blots using sodium periodate according to (33). Immunodetection with anti-V5 antibodies was performed following the instructions of the supplier for dilution of the antibody and revelation (ECL Kit).

Deglycosylation by PNGase F or Endo H – For deglycosylation with PNGase F, 0.5 mg of proteins was dissolved in 2 mL of a 0.1 M Tris-HCl buffer pH 7.5 containing 0.1% SDS. The sample was then heated for 5 min at 100°C for protein denaturation. After cooling down, 2 mL of 0.1 M Tris-HCl buffer pH 7.5, containing 0.5% Nonidet P-40 were added to the sample. Digestion was performed with 10 U of PNGase F for 24 h at 37°C. For deglycosylation by Endo H, 0.5 mg of protein extract was dissolved in 1% SDS and denaturated by heating for 5 min at 100°C. The sample was then diluted five times in 500 µL of 150 mM sodium acetate buffer pH 5.7 and incubated overnight at 37°C with 10 mU of Endo H. Finally, proteins digested from either PNGase F or Endo H were precipitated by addition of 4 volumes of ethanol overnight at -20°C, separated by SDS-PAGE and affinodetected with concanavalin A as previously reported (31).

In vitro galactosylation - The *in vitro* galactosylation was performed by treating 50 µg of protein at 37°C for 24 h with 50 mU of β(1,4)-galactosyltransferase from bovine milk in 1 mL of 100 mM sodium cacodylate buffer pH 6.4, supplemented with 5 µmol of UDP-galactose and 5 µmol of $MnCl_2$ (34). The sample was then freeze-dried. Proteins and glycoproteins were separated by SDS-PAGE and electro-blotted onto nitrocellulose membrane. Glycoproteins were then affinodetected using biotinylated RCA 120 lectin (34).

Sialic acid analysis – Two mg of proteins were treated as described in (35). Then, the sample was submitted to DMB derivatization according to (36). DMB derivatives were separated by high performance liquid chromatography using a C18 column and detected by fluorescence using excitation and emission wavelengths of 373 and 448 nm, respectively (37). Neu5Ac was also coupled to DMB and used as a standard.

Isolation of N-linked glycans from P. tricornutum – For N-glycan profiling, both PNGase A and PNGase F were used. In contrast to PNGase F, PNGase A is able to release N-linked oligosaccharides carrying a fucose α(1,3)-linked to the proximal glucosamine residue (38). Proteins (above 13 mg) were deglycosylated by PNGase F as described for CHO N-glycan profiling. For PNGase A, five mg of freeze-dried proteins were resuspended in 3 mL of 4 M Tris-Guanidine HCl pH 8.5 prior to denaturation with a $2mg.mL^{-1}$ dithiothreitol (DTT) solution. After a short 30 second sonication, the sample was incubated at 50°C during 2h. One and half mL of iodoacetamide, prepared in 0.6 M Tris buffer pH 8.5 at 12 $mg.mL^{-1}$, was then added to the sample which was incubated in the dark during 2h at room temperature. The sample was dialysed 72h against water. The proteins were digested at 37°C for 48 h with 10 mg of pepsin dissolved in 2 mL of 10 mM HCl, pH 2.2. After neutralization with 1 M ammonium hydroxide, the solution was heated for 5 min at 100°C and lyophilized. Glycopeptides were then deglycosylated overnight at 37°C with 1.5 mU of PNGase A in a 50 mM sodium acetate buffer, pH 5.5. N-glycans either released by PNGase A or F were purified by successive elutions through a C18 and a Carbograph cartridges according to (39).

Isolation of N-linked glycans from CHO cells - Cells from wild-type CHO, CHO Lec 1 mutant and CHO Lec1 GnT I complemented cells were lysed by sonication (4 x 20 s) in 1 mL of Tris-HCL buffer 100 mM pH 7.5, SDS 0.1 %. After centrifugation at 100 g, proteins from the supernatant were deglycosylated by PNGase F as described above and then, N-glycans were purified as already mentioned. Then, N-glycan samples were concentrated and finally analysed by Matrix Assisted Laser Desorption Ionisation-Time Of Flight (MALDI-TOF) mass

spectrometer.

Preparation and exoglycosidase digestion of 2-AB oligosaccharides - N-glycans were labelled with 2-aminobenzamide (2-AB) using the protocol described by (40). After incubation at 60°C for 2h, 2-AB labelled N-glycans were purified by paper chromatography according to (41). The 2-AB labelled N-glycans were finally analyzed by MALDI-TOF mass spectrometry before and after jack bean α-mannosidase or α-L-fucosidase treatments following the principle described in (39). For the α-mannosidase digestion, 2.5 µL of 2-AB labelled N-glycans were incubated with 1 µL of water and 214 mU of enzyme during 48h at 37°C. The α-L-fucosidase from bovine kidney was desalted prior to use and resuspended in 40 µM of sodium acetate pH 5.5. Then, 80 mU of enzyme has been incubated with 2.5 µL of 2-AB labelled glycans at 37°C during 48h. Both digested samples were freeze dried and resuspended in 10 µL of water / 0.1% TFA prior to mass spectrometry analysis.

MALDI-TOF mass spectrometry analysis - Mass spectra were acquired on a Voyager DE-Pro MALDI-TOF instrument equipped with a 337 nm nitrogen laser. Mass spectra were performed in the reflector delayed extraction mode using DHB. This matrix, freshly dissolved at 5 $mg.mL^{-1}$ in a 70:30 acetonitrile / 0.1% TFA, was mixed with the water solubilized oligosaccharides in a ratio 1:1 (v/v). These spectra were recorded in a positive mode, using an acceleration voltage of 20,000 V with a delay time of 100 ns. They were smoothed once and externally calibrated using commercially available mixtures of peptides and proteins. In this study, the spectra were externally calibrated using des-Arg^1-bradykinin (904.4681 Da), angiotensin I (1,296.6853 Da), Glu^1-fibrinopeptide B (1,570.6774 Da) and $ACTH_{18-39}$ (2,465.1989 Da). Laser shots were accumulated for each spectrum, at least 5 times 1000 laser shots. The mass accuracy obtained is 0.011% in average which is in agreement with the specifications of the instrument used in this study.

RESULTS

In silico analysis of the P. tricornutum genome revealed a set of genes encoding for proteins involved in the N-glycosylation pathway - In eukaryotes, the N-glycan biosynthetic pathway can be divided into three steps: i) the synthesis of the $Glc_3Man_9GlcNAc_2$-PP-dolichol precursor and its transfer by the OST onto asparagine residues of nascent polypeptides entering the lumen of the rough ER, ii) deglucosylation/reglucosylation of the precursor N-glycan in the ER allowing the interaction with chaperones responsible for proper folding and oligomerization and finally iii) maturation in the Golgi apparatus of the high mannose-type N-linked oligosaccharides into complex-type N-glycans. Based on sequence homologies, we identified in the genome of *P. tricornutum* a set of putative sequences that are likely involved in the different steps of the N-glycan biosynthesis and maturation (Fig. 1, Table 1). Most of these identified genes have EST support (Fig. 1, Table 1).

All the genes encoding for enzymes involved in the biosynthesis of dolichol pyrophosphate-linked oligosaccharide on the cytosolic face and in the lumen of the ER were identified in the genome of *P. tricornutum* (Fig. 1, Table 1). The sequences and topologies of the predicted proteins are highy similar to the corresponding asparagine-linked glycosylation (ALG) orthologs described in other eukaryotes (42), except for ALG10 for which a *P. tricornutum* candidate sequence was not clearly identified. Putative transferases, abling to catalyse the formation of dolichol-activated mannose and glucose, were also found. Those two activated sugars are required for the elongation steps arising in the ER lumen. In addition to sequences involved in the biosynthesis of the dolichol pyrophosphate-linked oligosaccharide, two putative genes encoding for orthologs of the STT3 catalytic subunit of OST multisubunit complex were identified (Table 1). These multi-spanned sequences, sharing respectively 34% and 37% of identities with *A. thaliana* and *Homo sapiens* STT3 subunits, contain the conserved WWDYG domain required for the STT3 transferase activity (43).

Genes encoding for polypeptides involved in the quality control of proteins in the ER were also found in the *P. tricornutum* genome. Indeed, α and β subunits of α-glucosidase II were identified. The α subunit contains the characteristic DMNE sequence (44) and a C-type lectin domain involved in mannose binding (45). A putative UDP-glucose: glycoprotein glucosyltransferase (UGGT) and a

calreticulin, two molecules ensuring the quality control of the glycoproteins in the ER, are also predicted. Calreticulin is a soluble Ca^{2+} binding protein of the ER lumen involved in the retention of incorrectly or incompletely folded proteins. Putative *P. tricornutum* calreticulin exhibits more than 50% of identity with orthologs from *Nicotiana plumbaginifolia* (56%) and *A. thaliana* (53%). Structurally, the *P. tricornutum* calreticulin contains the three specific domains required for its biological function: a N-terminal domain of about 180 aminoacids, a central domain of about 70 residues containing three repeats of an acidic 17 aminoacid motif and a C-terminal domain rich in acidic and lysine residues, both responsible for Ca^{2+} binding (46). *P. tricornutum* calreticulin also exhibited a predicted signal peptide and a C-terminal YDEF tetrapeptide that may ensure its retention in the ER as HDEL, KDEL or YDEL signals that are known to play this function in higher eukaryotes (47-49).

In regard to Golgi enzymes involved in N-glycan biosynthesis, *P. tricornutum* genome contains two sequences encoding for proteins that belong to the Glycosyl Hydrolase family GH47 (http://www.cazy.org/) which catalyze the hydrolysis of the terminal α(1,2)-mannose residues of high mannose-type N-glycans. The first predicted sequence (seq n° 52346) encodes for a protein sharing respectively 32% and 30% of identity with MNS4 and MNS5 from *A. thaliana*, which were characterized as being ER degradation-enhancing mannosidases (50). The second sequence (Fig. 1, Table 1) encodes for a protein sharing respectively 35% and 34% of identity with MNS1 and MNS2, two *A. thaliana* α-Man I located in the Golgi apparatus and able to perform the trimming of Man-9 into Man-5 (50, 51). Furthermore, this putative *P. tricornutum* mannosidase exhibits the three conserved catalytic motifs of α-Man I, the threonine residue of the motif III and the two cysteine residues (C301 and C333) essential for the mannosidase activity (51, 52). A signal anchor is also predicted in the N-terminal part of the protein as required for a type II transmembrane protein (Table 1). Moreover, there is some EST supporting an expression for this enzyme in *P. triconutum* (table 1).

In addition, one putative GnT I and one putative α-Man II were also identified in *P. tricornutum* genome (Table 1). These enzymes are involved in the N-glycan maturation into complex oligosaccharides by transfering a terminal GlcNAc on the α(1,3)-mannose arm of Man-5 and then removing the two mannose residues located on the α(1,6)-mannose arm. The putative GnT I sequence is predicted to be a typical type II membrane protein. Its cytoplasmic tail contains 3 basic amino acids which could promote ER exit as demonstrated for *N. tabacum* GnT I (53). This sequence also possess a luminal part sharing 37% of identity with the rabbit GnT I (Fig. 2). From the crystal structure of this mammalian transferase, 22 amino acid residues in the catalytic domain were shown to form direct or water-mediated interactions with the UDP-GlcNAc nucleotide-sugar and the Mn^{2+} ion (54, 55). Fourteen of these residues are strictly identical in the *P. tricornutum* GnT I whereas the other residues are closely conserved (Fig. 2). Moreover, the SQD motif, which has been demonstrated to be important since this motif interacts with the uracil ring of the donor substrate, is also present in the *P. tricornutum* GnT I sequence (54). This conserved motif is present in all GnT I characterized so far (58). The EDD motif that is a variation of the canonical acidic metal binding DxD motif is conserved in the *P. tricornutum* sequence. This motif has been demonstrated to have critical interaction with UDP-GlcNAc and the metal ion in rabbit GnT I (55). The predicted α-Man II consisted of a large protein containing the three Pfam domains of CAZy GH38 glycosylhydrolases and the conserved residues involved in the Zn^{2+} binding in the catalytic site of *Drosophila melanogaster* α-Man II (59, 60).

In eukaryotes, α(1,3)- and α(1,6)-fucosyltransferases transfer fucose residues onto the proximal GlcNAc unit of the N-linked glycan core. *In silico* analysis of the *P. tricornutum* genome revealed the presence of three genes encoding for putative fucosyltransferases (FucT). These candidates exhibit the appropriate type II membrane protein topology (Table 1) and respectively 23% (seq n°46109), 28% (seq n° 46110) and 25% (seq n°54559) of identity with *A. thaliana* FucTA. These FucT candidates exhibited the motifs I and II of α(1,3)-FucT (61, 62), the $SNC^{G}/_{A}A^{R}/_{H}N$ sequence, specific for plant and *Drosophila melanogaster* α(1,3)-FucT (63-65) as well as the CXXC motif located at the C-terminal sequence involved in the formation of disulfide bonds (66). A putative type II xylosyltransferase is also predicted in the genome. This sequence shares 24% of identity

with the luminal part of *A. thaliana* β(1,2)-xylosytransferase (β(1,2)-XylT) involved in the transfer of a β(1,2)-xylose residue onto the β-Man of the N-glycan core (67, 68). Nevertheless, in the absence of reported motifs specific for β(1,2)-XylT activity, the involvement of this putative transferase in *P. tricornutum* N-glycan pathway remains highly hypothetical.

Searches for sequences encoding other N-glycan-maturating transferases, such as N-acetylglucosaminyltransferases ranging from GnT II to GnT VI that allow the formation of polyantennary N-glycans, or sialyltransferases, did not reveal any ortholog in the *P. tricornutum* genome.

P. tricornutum proteins mainly carry high mannose-type N-glycans - Analysis of glycans N-linked to *P. tricornutum* proteins was first investigated by western-blot on a total protein extract using probes specific for glycan epitopes. As illustrated in Fig. 3A, some *P. tricornutum* proteins were affinodetected by concanavalin A, a lectin specific for high mannose sequences (69). This affinodetection was largely suppressed upon treatment with Endo H or PNGase F (Fig. 3B), two enzymes able to cleave N-glycans. In contrast, affinodetections with ECA, PNA or PHA-E and L, four lectins specific for respectively, the lactosamine motif, the galβ1,3-GalNAc epitope, the bisected di-triantennary complex type N-glycans and tri- and tetraantennary complex-type N-glycans, did not show anything (data not shown). Immunodetections with antibodies raised against plant glycoepitopes (32) were then carried out (Fig. 3A). Antibodies specific for xylose epitope β(1,2)-linked to the core $Man_3GlcNAc_2$ did not detect proteins from *P. tricornutum* whereas proteins from green onion used as positive control were labeled (Fig. 3A). In contrast, total proteins were weakly immunodetected using core-α(1,3)-fucose antibodies (Fig. 3A). These signals disappeared after oxidation of the proteins with sodium periodate, revealing that the antibody recognition involved a glycan sequence (data not shown).

In order to investigate the presence in *P. tricornutum* proteins of complex glycans carrying terminal GlcNAc, we treated the protein extract with a β(1,4)-galactosyltransferase, an enzyme able to transfer a galactose residue onto terminal GlcNAc residues, and then analyzed by affinoblotting the resulting protein preparation with RCA 120, a lectin that binds specifically to Galβ1-4GlcNAc sequences (34). In contrast to a plant-derived IgG used as a positive control of galactose transfer, no signal was detected in *P. tricornutum* sample after this treatment (Fig. 3C), thus indicating that this diatom does not exhibit terminal GlcNAc onto its proteins at detectable level (0.5 μg). Moreover, the presence of N-acetylneuraminic acid (Neu5Ac), the main sialic acid found in mammals, was investigated by coupling to DMB (37) and analysis of the resulting DMB-derivatives by liquid chromatography. Although low intensity peaks were detected by fluorescence, none of them co-migrated with a standard of DMB-Neu5Ac (data not shown).

To investigate their detailed N-glycan structures, N-glycans were released from proteins by PNGase A treatment (31). The resulting N-glycans were then coupled to 2-amidobenzamide (2-AB) (40, 41) to facilitate their detection and analysis by MALDI-TOF mass spectrometry. As illustrated in Fig. 4A, major ions correspond to $(M+Na)^+$ adducts of 2-AB derivatives of $Hexose_{5-9}GlcNAc_2$. Other minor ions were also detected in the mass spectrum profile with *m/z* values corresponding to $Hexose_{3-4}GlcNAc_2$ oligosaccharides. Moreover, ion at *m/z* 1199 was assigned to $Hexose_3DeoxyhexoseGlcNAc_2$ N-linked glycan. The pool of glycans was then submitted to exoglycosidase digestions. The oligosaccharide mixture was converted to $HexoseGlcNAc_2$ and $HexoseDeoxyHexoseGlcNAc_2$ upon a treatment with Jack bean α-mannosidase (Fig. 4B) demonstrating the presence of α-linked mannose residues in PNGase A-released N-glycans. Furthermore, a treatment of the sample with α-L-fucosidase resulted in the suppression of ion at *m/z* 1199 (not shown). As a consequence, main ions detected in MALDI-TOF mass spectrum (Fig. 4A) were assigned to high mannose N-glycans ranging from Man-5 to Man-9 and minor ions were assigned to Man-3, Man-4 and $Man_3FucGlcNAc_2$.To investigate the location of the fucose residue onto the core N-glycan in this later N-glycan, proteins were submitted to a deglycosylation experiment by PNGase F, a deglycosylating enzyme that is not able to cleave N-linked oligosaccharides harbouring a fucose α(1,3)-linked to the proximal glucosamine residue (38). The ion assigned to $Man_3FucGlcNAc_2$ was not observed in the mass spectrum indicating that this glycan carries a core α(1,3)-fucose residue (Fig. 4C).

Putative P. tricornutum GnT I gene is able to complement the deficiency in N-glycan maturation in CHO Lec1 mutant - Bio-informatic analysis of *P. tricornutum* genome revealed a gene potentially encoding for a GnT I glycosyltransferase. The expression of this gene was monitored in a continuous culture over a one-month period by real time quantitative PCR. The relative gene expression was normalized to two reference genes encoding for Ribosomal Protein Small subunit 30 S (RPS) and histone H4, that have been recently described as appropriate housekeeping genes for real time quantitative PCR in *P. tricornutum* (26). The results show that *P. tricornutum* GnT I is steadily expressed in standard culture conditions over 32 days (Fig. 5A).

To demonstrate that this putative transferase is able to transfer *in vivo* a GlcNAc residue onto the proteins carrying Man-5 N-linked glycans, the *P. tricornutum* GnT I (GenBank BankIt ID 1370344) was expressed in fusion with a V5 tag into CHO Lec1 mutant deficient in its endogenous GnT I (16). On the basis of the immunodetection of the V5 epitope, CHO transformants were found to efficiently express the recombinant GnT I. Two transformants (2 and 4) were selected for N-linked glycan analysis (Fig 5B). Proteins from these two clones, from the wild-type CHO and from the CHO Lec1 mutant, were isolated and their N-linked glycans were released by treatment with PNGase F. As illustrated in Fig. 6A, MALDI-TOF mass spectrum of N-glycans released from the CHO Lec1 mutant showed that it accumulates high mannose-type N-glycans, contrasting with the wild-type CHO cell exhibiting both high mannose and complex-type N-glycans (Fig. 6B). In Fig. 6C, glycoproteins from the complemented line (clone 4) displayed both high mannose oligosaccharides and a complete set of complex-type N-glycans identical to the one observed in wild-type CHO cells (Fig. 6B). The same N-glycan profile was obtained with the clone 2 (data not shown). This demonstrates that the expression of the *P. tricornutum* GnT I was able to restore the biosynthesis of complex N-glycans in the mammalian cell mutant.

DISCUSSION

Bio-informatic analysis of the *P. tricornutum* genome revealed the presence of a completed set of sequences potentially encoding for proteins involved in the synthesis of the lipid-linked $Glc_3Man_9GlcNAc_2$-PP-dolichol N-glycan, the subunits of the OST complex that catalyses its transfer onto the asparagine residues of target proteins, as well as ER glucosidases and chaperones (6, 7). This suggests that this diatom possesses the ER machinery required for glycoprotein quality control previously characterized for other eukaryotes (71). The genome analysis also revealed the presence of one Golgi mannosidase I involved in the trimming of the N-glycan precursor into Man-5 high mannose-type N-glycan (50, 51). Consistent with these sequence predictions, two lines of biochemical evidence strongly suggest that proteins from the *P. tricornutum* Pt 1.8.6 strain are mainly N-glycosylated by high mannose-type oligosaccharides. Firstly, proteins from this strain are affinodetected by concanavalin A, a lectin specific for high mannose sequences. This detection is largely suppressed upon treatment with Endo H or PNGase F, two deglycosylating enzymes specific for N-linked glycans. Furthermore, affino- and immunodetection with other glycan-specific probes as well as search for sialic acids were unsuccessful. Secondly, MALDI-TOF mass spectrometry of the N-glycan population released by PNGase A allowed the detection of $Hexose_{5\text{-}9}GlcNAc_2$ oligosaccharides, sensitive to an α-mannosidase treatment, as major oligosaccharide species. We concluded that proteins from the *P. tricornutum* mainly carry Man-5 to Man-9 high mannose-type N-glycans. Other minor glycan species *i.e.* Man-3, Man-4 and $Man_3FucGlcNAc_2$ carrying a fucose α(1,3)-linked to the proximal glucosamine residue were also detected. The presence of this later sequence corroborates the weak detection of core α(1,3)-fucose epitopes on western-blot, potential product of the putative FucT sequences predicted in the *P. tricornutum* genome.

In eukaryotes, the first steps of the N-glycan processing into complex N-glycans are controlled by GnT I, α-Man II and GnT II. The resulting core N-glycan is modified by the action of a wide variety of glycosyltransferases giving rise to mature N-linked glycans involved in various biological processes (8-11). Putative GnT I and α-Man II are predicted in the *P. tricornutum* genome. We mainly focussed on GnT I characterization since this transferase is

the first enzyme initiating the complex-type maturation of oligosaccharides N-linked to secreted proteins. In standard culture conditions, we demonstrated that the gene encoding this putative transferase is expressed. However, no glycan carrying terminal GlcNAc residues has been detected on *P. tricornutum* proteins by either a galactosyltransferase assay or by MALDI-TOF mass spectrometry analysis of the PNGase A-released oligosaccharides. Therefore, the *in vivo* activity of this putative GnT I was investigated by expressing the full-length protein in CHO Lec1 mutant lacking its endogenous GnT I activity (16). Wild-type like N-glycosylation profiles were detected in transformed cell lines thus demonstrating that the putative GnT I from this diatom was able to restore the biosynthesis of complex-type N-glycans in GnT I-null CHO cells. This shows that *P. tricornutum* gene encodes for a protein able to perform *in vivo* the processing of oligomannosides into complex-type N-glycans, thus corresponding to a GnT I activity. To our knowledge, this work is the first functional characterization of a microalgal N-glycan glycosyltransferase. These data also suggest that both the targeting and the retention mechanism of Golgi enzymes are conserved between mammals and diatoms.

Search for putative GnT I was carried out in genomes of species belonging to the three main microalgae lineages, i.*e. Viridiplantae, Heteroconta* and *Haptophyta* of the tree of live. Genes encoding for putative GnT I were identified in two other heterokonts, i.e. *Fragilariopsis cylindrus* and *Thalassosira pseudonana*, as well as in the haptophyte *Emiliania huxleyi*. Search for GnT I in *Viridiplantae* only revealed one sequence in *Micromonas pusilla* and not in other species (Fig. 7). The microalgal GnT I complete sequences share about 25% of identity with plant and mammal GnT I. *Heteroconta* and *Haptophyta* GnT I are gathered in a distinct lineage which could be clearly separated from other GnT I as seen in the phylogenetic tree (Fig. 7). So far, the processing of N-linked glycans into complex oligosaccharide has been mainly described in multicellular higher eukaryotes such as animals and land plants and has been demonstrated to be required for normal morphogenesis in animals (9, 10). Our results show that this key Golgi transferase is also involved in the processing of N-linked glycans in unicellular microalgae species.

Mainly high mannose-type N-glycans were detected onto *P. tricornutum* proteins which suggests that GnT I possess a limited *in vivo* impact on glycans N-linked to secreted proteins of this diatom. No glycan carrying terminal GlcNAc residue has been detected onto *P. tricornutum* proteins. However, although representing a minor glycan population, small size N-glycans Man-3 and Man-4, as well as $Man_3FucGlcNAc_2$ were detected in the *P. tricornutum* N-glycan mass profile. These glycans, named paucimannose structures, have been previously found in invertebrates and plants and result from the degradation of GlcNAc-terminated complex glycans ($GlcNAcMan_{3–4}GlcNAc_2$) by N-acetylglucosaminidases after their biosynthesis in the Golgi apparatus. Indeed, in the GnT I-dependent pathway, $GlcNAcMan_5GlcNAc_2$, the product of GnT I (Fig. 1), is successively converted in the Golgi apparatus into $GlcNAcMan_4GlcNAc_2$ and then into $GlcNAcMan_3GlcNAc_2$ by action of the α-Man II and then into $GlcNAcMan_3FucGlcNAc_2$ by α-FucT. Elimination of terminal GlcNAc by β-N-acetylglucosaminidases in the secretory system or in compartments where proteins accumulate can then degrade these oligosaccharides into Man-3, Man-4 and $Man_3FucGlcNAc_2$. Such a processing was demonstrated in insect cells (72), *C. elegans* (73) and plants (74). Two putative processing β-N-acetylglucosaminidases belonging to the CAZy GH20 family are predicted in the *P. tricornutum* genome (sequences n° 49563 and 45073). These glycosidases share respectively 43% and 36% of identity with DmFDL, a β-N-acetylglucosaminidase from *Drosophila* that is able to specifically hydrolyse a GlcNAc residue located onto the α(1,3)-antenna of N-glycans giving rise to paucimannose oligosaccharides (75, 76). Taken together, these data suggest that such a processing may also occur in diatoms.

The biochemical characterization of the core-modifying FucTs and of the processing β-N-acetylglucosaminidases are currently studied. The N-glycosylation patterns of proteins from *P. tricornutum* grown in different conditions are currently under investigation in order to study this major post-translational modification in relation with pleomorphism and/or stress environmental conditions and the function of complex N-glycans in the diatom physiology.

REFERENCES

1. Nelson, D.M., Treguer, P., Brzezinski, M.A., Leynaert, A., & Queguiner, B., (1995) *Globalal. Biogeochemical Cycles* **9**, 359–372
2. Raven, J.A., and Waite, A.M. (2004) *New Phytologist* **162**, 45-61
3. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, Fournet J, Haruta M, Huysman MJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert Y, Kaplan A, Kröger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-Jézéquel V, Lopez PJ, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A, Oudot-Le Secq MP, Napoli C, Obornik M, Parker MS, Petit JL, Porcel BM, Poulsen N, Robison M, Rychlewski L, Rynearson TA, Schmutz J, Shapiro H, Siaut M, Stanley M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van de Peer Y, Grigoriev IV (2008) *Nature* **456**, 239–244
4. Falciatore, A., and Bowler, C. (2002) *Annu Rev Plant Biol.* **53**, 109-130
5. Scala, S., Carels, N., Falciatore, A., Chiusano, M.L., Bowler, C. (2002) *Plant Physiol.* **129**, 993-1002
6. Burda, P., and Aebi, M. (1999) *Biochim Biophys Acta.* **1426**, 239-257
7. Helenius, A., and Aebi, M. (2001) *Science.* **291**, 2364-2369
8. Schachter H (2009) *Carbohydr Res.* **344**, 1391-1396
9. Ioffe, E., and Stanley, P. (1994) *Proc Natl Acad Sci U S A.* **91**, 728-732
10. Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J.W., and Marth, J.D. (1994) *EMBO J.* **13**, 2056-2065
11. Sarkar, M., Leventis, P.A., Silvescu, C.I., Reinhold, V.N., Schachter, H., and Boulianne, G.L. (2006) *J. Biol. Chem.* **281**, 12776-12785
12. von Schaewen, A., Sturm, A., O'Neill, J., and Chrispeels, M.J. (1993) *Plant Physiol.* **102,** 1109-1118
13. Kang, J.S., Frank, J., Kang, C.H., Kajiura, H., Vikram, M., Ueda, A., Kim, S., Bahk, J.D., Triplett, B., Fujiyama, K., Lee, S.Y., von Schaewen, A., and Koiwa, H. (2008) *Proc. Natl. Acad. Sci. U S A.* **105**, 5933-5938
14. von Schaewen, A., Frank, J., and Koiwa, H. (2008) *Plant Signal Behav.* **3**, 871-873
15. Frank, J., Kaulfurst-Soboll, H., Rips, S., Koiwa, H. and von Schaewen, A. (2008) *Plant Physiol.* **148**, 1354-1367
16. Stanley, P., Narasimhan, S., Siminovitch, L., and Schachter, H. (1975) *Proc. Natl. Acad. Sci. U S A.* **72**, 3323-3327
17. Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., and Gascuel, O. (2008) *Nucleic Acids Res.* **36** (Web Server issue), W465-9
18. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) *Nucleic Acids Res.* **22**, 4673-4680
19. Castresana J (2000) *Mol. Biol. Evol.* **17**, 540-552
20. Guindon, S., and Gascuel, O. (2003) *Syst. Biol.* **52**, 696-704
21. Anisimova, M., and Gascuel, O. (2006) *Syst. Biol.* **55**, 539-552
22. Chevenet, F., Brun, C., Bañuls, A.L., Jacq, B., and Christen, R. (2006) *BMC Bioinformatics* **7**, 439

23. Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard T., Lombard V., and Henrissat, B. (2009) *Nucleic Acids Res.* **37**, D233-238
24. Walne PR (1966) *Fisheries investigations* **2**, 1-52
25. Bougaran, G., Le Dean, L., Lukomska, E., Kaas, R., and Baron, R. (2003) *Aquat. Living Resour.* **16**, 389-394
26. Siaut, M., Heijde, M., Mangogna, M., Montsant, A., Coesel, S., Allen, A., Manfredonia, A., Falciatore, A., and Bowler, C. (2007) *Gene* **406**, 23-35
27. Chung, C.C., Hwang, S.P., and Chang, J. (2005) *Appl. Environ Microbiol.* **71**, 8744-8751
28. Kang, L. K., Hwang, S. P. L., Lin, H. J., Chen, P. C., and Chang, J. (2009) *J. Phycol.* **45**, 864-872
29. Livak, K.J., and Schmittgen, T.D. (2001) *Methods* **25**, 402-408
30. Falciatore, A., Casotti, R., Leblanc, C., Abrescia, C., and Bowler, C. (1999) *Mar. Biotechnol.* **1**, 239-251
31. Fitchette, A.C., Tran Dinh, O., Faye, L., and Bardor, M. (2007) *Methods in Molecular Biology* , 317-342
32. Faye, L., Gomord, V., Fitchette-Lainé, A.C., and Chrispeels, M.J. (1993) *Anal. Biochem.* **209**, 104-108
33. Fitchette-Lainé, A-C., Denmat, L-A., Lerouge, P., and Faye, L. (1998) *Methods in Biotechnology* **3**, 271-290
34. Bardor, M., Loutelier-Bourhis, C., Paccalet, T., Cosette, P., Fitchette, A.C., Vézina, L.P., Trépanier, S., Dargis, M., Lemieux, R., Lange, C., Faye, L., and Lerouge, P. (2003) *Plant Biotechnol. J.* **1**, 451-462
35. Bardor, M., Nguyen, D.H., Diaz, S., and Varki, A. (2005) *J. Biol. Chem.* **280**, 4228-4237
36. Hara, S., Yamaguchi, M., Takemori, Y., Furuhata, K., Ogura, H., and Nakamura, M. (1989) *Anal. Biochem.* **179**, 162-166
37. Klein, A., Diaz, S., Ferreira, I., Lamblin, G., Roussel, P., and Manzi, A.E. (1997) *Glycobiology* **7**, 421-432
38. Tretter, V., Altmann, F., and März, L. (1991) *Eur. J. Biochem.* **199**, 647-652
39. Bakker, H., Bardor, M., Molthoff, J.W., Gomord, V., Elbers, I., Stevens, L.H., Jordi, W., Lommen, A., Faye, L., Lerouge, P., and Bosch, D. (2001) *Proc. Natl. Acad. Sci. U S A.* **98**, 2899-2904
40. Bigge, J.C., Patel, T.P., Bruce, J.A., Goulding, P.N., Charles, S.M., and Parekh, R.B. (1995) *Anal. Biochem.* **230**, 229-238
41. Séveno, M., Cabrera, G., Triguero, A., Burel, C., Leprince, J., Rihouey, C., Vézina, L.P., D'Aoust, M.A., Rudd, P.M., Royle, L., Dwek, R.A., Harvey, D.J., Lerouge, P., Cremata, J.A., and Bardor, M. (2008) *Anal. Biochem.* **379**, 66-72
42. Weerapana, E., and Imperiali, B. (2006) *Glycobiology* **16**, 91R-101R
43. Yan, Q., and Lennarz, W.J. (2002) *J. Biol. Chem.* **277**, 47692-47700
44. Feng, J., Romaniouk, A.V., Samal, S.K., and Vijay, I.K. (2004) *Glycobiology* **14**, 909-921
45. Wilkinson, B.M., Purswani, J., and Stirling, C.J. (2006) *J. Biol. Chem.* **281**, 6325-6333
46. Kwiatkowski, B.A., Zielińska-Kwiatkowska, A.G., Migdalski, A., Kleczkowski, L.A., and Wasilewska, L.D. (1995) *Gene* **165**, 219-222
47. Matsushima, R., Kondo, M., Nishimura, M., and Hara-Nishimura, I. (2003) *Plant J.* **33**, 493-502
48. Pagny, S., Denmat-Ouisse, L.A., Gomord, V., and Faye, L. (2003) *Plant Cell Physiol.* **44**, 173-182
49. Gnanasambandam, A., Polkinghorne, I.G., and Birch, R.G. (2007) *Plant Biotechnol. J.* **5**, 290-296

50. Liebminger, E., Hüttner, S., Vavra, U., Fischl, R., Schoberer, J., Grass, J., Blaukopf, C., Seifert, G.J., Altmann, F., Mach, L., and Strasser, R. (2009) *Plant Cell* **21**, 3850-3867
51. Kajiura, H., Koiwa, H., Nakazawa, Y., Okazawa, A., Kobayashi, A., Seki, T., and Fujiyama, K. (2010) *Glycobiology* **20**, 235-247
52. Tempel, W., Karaveg, K., Liu, Z.J., Rose, J., Wang, B.C., and Moremen, K.W. (2004) *J. Biol. Chem.* **279**, 29774-29786
53. Schoberer, J., Vavra, U., Stadlmann, J., Hawes, C., Mach, L., Steinkellner, H., and Strasser, R. (2009) *Traffic* **10**, 101-115
54. Unligil, U.M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J.M. (2000) *EMBO J.* **19**, 5269-5280
55. Gordon, R.D., Sivarajah, P., Satkunarajah, M., Ma, D., Tarling, C.A., Vizitiu, D., Withers, S.G., and Rini, J.M. (2006) *J. Mol. Biol.* **360**, 67-79
56. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18**, 2714-2723
57. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* **15**, 305-308
58. Strasser, R., Staldmann, J., Svoboda, B., Altmann, F., Glossl, J., and Mach, L. (2005) *Biochem. J.* **387**, 385-391.
59. Strasser, R., Schoberer, J., Jin, C., Glössl, J., Mach, L., and Steinkellner, H. (2006) *Plant J.* **45**, 789-803
60. Van den Elsen, J.M., Kuntz, D.A., and Rose, D.R. (2001) *EMBO J.* **20**, 3008-3017
61. Breton, C., and Imberty, A. (1999) *Curr. Opin. Struct. Biol.* **9**, 563-571
62. Oriol, R., Mollicone, R., Cailleau, A., Balanzino, L., and Breton, C. (1999) *Glycobiology* **9**, 323-334
63. Bakker H, Schijlen E, de Vries T, Schiphorst WE, Jordi W, Lommen A, Bosch D, van Die I. (2001) *FEBS Lett.* **507**, 307-12
64. Wilson, I.B., Rendić, D., Freilinger, A., Dumić, J., Altmann, F., Mucha, J., Müller, S., and Hauser, M.T. (2001) *Biochim. Biophys. Acta.* **1527**, 88-96
65. Fabini, G., Freilinger, A., Altmann, F., and Wilson, I.B. (2001) *J. Biol. Chem.* **276**, 28058-28067
66. Holmes, E.H., Yen, T.Y., Thomas, S., Joshi, R., Nguyen, A., Long, T., Gallet, F., Maftah, A., Julien, R., and Macher, B.A. (2000) *J. Biol. Chem.* **275**, 24237-24245
67. Strasser, R., Mucha, J., Mach, L., Altmann, F., Wilson, I.B., Glössl, J., Steinkellner, H. (2000) *FEBS Lett.* **472**, 105-108
68. Pagny, S., Bouissonnie, F., Sarkar, M., Follet-Gueye, M.L., Driouich, A., Schachter, H., Faye, L., and Gomord, V. (2003) *Plant J.* **33**, 189-203
69. Faye, L., and Chrispeels, M.J. (1985) *Anal. Biochem.* **149**, 218-224
70. Varki, A., Cummings, R. D., Esko, J.D., Freeze, H. H., Stanley, P., Marth, J. D., Bertozzi, C.R., Hart, G. W. and Etzler, M. E. (2009) *Proteomics* **9**, 5398-5399
71. Banerjee, S., Vishwanath, P., Cui, J., Kelleher, D.J., Gilmore, R., Robbins, P.W., and Samuelson, J. (2007) *Proc. Natl. Acad. Sci. U S A.* **104,** 11676-11681.
72. Altmann, F., Schwihla, H., Staudacher, E., Glössl, J., and März, L. (1995) *J. Biol. Chem.* **270**, 17344-17349
73. Zhang, W., Cao, P., Chen, S., Spence, A.M., Zhu, S., Staudacher, E., and Schachter, H. (2003) *Biochem. J.* **372**, 53-64
74. Vitale, A., and Chrispeels, M.J. (1984) *J. Cell Biol.* **99**, 133-140
75. Léonard, R., Rendic, D., Rabouille, C., Wilson, I.B., Préat, T., and Altmann, F. (2006) *J. Biol. Chem.* **281**, 4867-4875
76. Geisler, C., Aumiller, J.J., and Jarvis, D.L. (2008) *J. Biol. Chem.* **283**, 11330-11339

FOOTNOTES

B.B. is a fellow of the french ministry for research and higher education. Authors gratefully thank C. Bowler, ENS, Paris and the Joint Genome Institute for giving priority access to the *P. tricornutum* genome. P.L. and M.B. acknowledge the University of Rouen for its support. Authors also thank S. Calbo, INSERM U905, University of Rouen, C. Loutelier-Bourhis, CNRS UMR614, University of Rouen, C. Rihouey, CNRS UMR 6522, University of Rouen and R. Kaas, Laboratoire Physiologie et Biotechnologies des Algues, Nantes, IFREMER for technical help and advices as well as Margot Lerouge for help with producing figures. Authors would like to give our thanks to J.-C. Mollet, A. Driouich and S. Calbo from the University of Rouen, for critical readings of our manuscript.

The abbreviations used are: α-Man, alpha mannosidase; GnT, N-acetylglucosaminyltransferase; DHB, 2,5-DiHydroxyBenzoic acid; MEM, alpha Minimum Essential Medium; FBS, Fetal Bovine Serum; DMB, 1,2-Diamino-4,5-Methylene Dioxybenzene; 2-AB, 2-AminoBenzamide; PHA-E and L, Phytohemagglutinin E and L; RCA 120, *Ricinus Communis* Agglutinin 120; ECA, *Erythrina Crista Galli* Agglutinin; PNA, peanut agglutinin; ConA, Concanavalin A; ECL, ElectroChimioLuminescent; PNGase F, Peptide N-glycosidase F; PNGase A, Peptide N-glycosidase A; Endo H, Endoglycosidase H; CHO, Chinese Hamster Ovary; SDS, Sodium Dodecyl Sulfate; ALG, Asparagine-Linked Glycosylation; GH, GlycosylHydrolase; GT, GlycosylTransferase; MALDI-TOF, Matrix Assisted Laser Desorption Ionisation-Time Of Flight; Man, Mannose; GlcNAc, N-acetylglucosamine; OST, Oligosaccharyltransferase; ER, Endoplasmic Reticulum; Glc, glucose; PP-Dol, dolicholpyrophosphate; Asn, asparagine; Ser, serine; Thr, threonine; SP/SA, Signal Peptide or Signal Anchor; EST, Expressed Sequence Tag; TMD, TransMembrane Domain; Y, yes; N, no; GlcNAcT, N-acetylglucosaminyltransferase; ManT, mannosyltransferase; GlcT, glucosyltransferase; P-Dol GlcT, Dolichol-phosphate β-glucosyltransferase; P-Dol ManT, Dolichol-phosphate mannosyltransferase; α-Glc II, α-glucosidase II; FucT, fucosyltransferase.

FIGURE LEGENDS

Fig. 1. **N-glycosylation pathway in *P. tricornutum* based on bioinformatic analysis of the genome database**. Sequences of the N-glycosylation pathway identified in the *P. tricornutum* genome, are numbered in bold. Man, mannose; GlcNAc, N-acetylglucosamine; Glc, glucose; PP-Dol, dolicholpyrophosphate; ALG, Asparagine-Linked Glycosylation.

Fig. 2. **Catalytic amino acids are very conserved in the putative GnT I protein from *P. tricornutum*.** Protein sequences alignment between rabbit (1FOA) and *P. tricornutum*, as proposed by the Swiss-Pdb viewer program (56). Secondary structural elements are represented above the alignment for the *P. tricornutum* GnT I and below the alignment for the rabbit GnT I with ➡ as β strand and (ααα) as α helix. Essential residues for the binding of the donor substrate (UDP-GlcNAc) are indicated by arrowheads above the alignment: in black when identicals and in white when not. Rabbit GnT I disulphide bridges are also numbered. The figure was created with the Espript program (http://espript.ibcp.fr/ ; 57).

Fig. 3. ***P. tricornutum* glycoproteins harbour N-linked oligosaccharides**. (A) Affinodetection using Con A and immunodetection using antibodies raised against the core β(1,2)-xylose (anti-Xyl) and core α(1,3)-fucose (anti-Fuc) epitopes of proteins isolated from green onion used as a positive control (Lane 1) and from *P. tricornutum* (Lane 2). (B) Affinodetection by Con A of proteins extracted from *P. tricornutum* treated (+) or not (-) with Endoglycosidase H (Endo H) and Peptide N-glycosidase F (PNGase F). (C) Affinodetection with RCA 120 of *P. tricornutum* proteins treated (+) or not (-) by

bovine β(1,4)-galactosyltransferase. Plant-derived IgG was used as a positive control of the galactose transfer efficiency (34). Arrows indicate the migration of heavy (H) and light (L) chains.

Fig. 4. **High mannose-type N-glycans are the main oligosaccharides N-linked to *P. tricornutum* proteins**. (A) MALDI-TOF mass spectrum of N-linked glycans released by PNGase A from glycoproteins of *P. tricornutum* and labelled with 2-aminobenzamide (2-AB). (B) MALDI-TOF mass spectrum of the pool of N-glycans after treatment with Jack bean α-mannosidase. (C) MALDI-TOF mass spectrum of 2-AB-labelled N-linked glycans released by PNGase F from glycoproteins of *P. tricornutum*. Man-3 to Man-9: paucimannose and high mannose-type N-glycans $Man_3GlcNAc_2$ to $Man_9GlcNAc_2$. *: contaminants. ■: potassium adducts.

Fig. 5. **Expression of the GnT I gene in *P. tricornutum* and in CHO Lec1 mutant**. (A) Expression of the GnT I gene of *P. tricornutum* by real time quantitative-PCR over a one month period. The relative gene expression was normalized to two reference genes encoding for histone H4 and Ribosomal Protein Small subunit 30 S (RPS). (B) Western-blot analysis using anti-V5 antibodies of proteins isolated from two transformants (2 and 4) of CHO Lec1 expressing the *P. tricornutum* GnT I fused to a V5 tag. GnT I was detected at the expected molecular mass of 56 kDa.

Fig. 6. ***P. tricornutum* GnT I complements N-glycan maturation deficiency in CHO Lec1 mutant.** MALDI-TOF mass spectra of glycans N-linked to proteins extracted from CHO cells. (A) CHO Lec1 mutant, (B) CHO wild-type and (C) transformant 4 of CHO Lec1 mutant complemented with *P. tricornutum* GnT I gene. Man-4 to Man-9: high mannose-type N-glycans $Man_4GlcNAc_2$ to $Man_9GlcNAc_2$. Symbols (70): black square with black outline: GlcNAc, grey circle with black outline: Man, white circle with black outline: Gal, grey triangle with black outline: Fucose.

Fig. 7. **GnT I are predicted in other microalgae.** Phylogenetic tree of GnT I from algae, plants and animals based on the maximum likelihood method. The scale bar (0.4) represents the number of amino acid residue substitutions per site. Microalgal GnT I sequences are indicated in bold. Accession numbers for the different reference sequences are indicated in the experimental procedure section.

Table 1: **References of predicted proteins involved in the N-glycan biosynthesis and quality control of secreted proteins in *P. tricornutum***. SP/SA: Signal Peptide or Signal Anchor, EST: Expressed Sequence Tag, TMD: TransMembrane Domain, Y: yes, N: no, GlcNAcT: N-acetylglucosaminyltransferase, ManT: mannosyltransferase, GlcT: glucosyltransferase, P-Dol GlcT: Dolichol-phosphate β-glucosyltransferase, P-Dol ManT: Dolichol-phosphate mannosyltransferase, OST: Oligosaccharyltransferase, α-Glc II: α-glucosidase II, α-Man I: α-mannosidase I, GnT I: N-acetylglucosaminyltransferase I, α-Man II: α-mannosidase II, FucT: fucosyltransferase.

Protein N°	Gene location	Predicted protein function	Length	EST	SP/ SA	TMD	Pfam
9724	chr_2:964854-966112	GlcNAc-P-transferase ALG7	440	N	Y	9	PF00953
9427	chr_1:864467-864835	β(1,4)-GlcNAcT ALG 13	170	N	Y	0	PF04101
14444	chr_15:95429-95971	β(1,4)-GlcNAcT ALG 14	180	Y	N	2	PF08660
14002	chr_13:187936-189201	β(1,4)-ManT ALG 1	448	Y	N	2	PF00534
22554	chr_18:299673-301001	α(1,3)-ManT ALG 2	503	Y	N	1	PF00534
54621	chr_11:102045-103589	α(1,2)-ManT ALG 11	433	Y	Y	4	PF00534
10976	chr_4:589265-590525	α(1,3)-ManT ALG 3	414	Y	Y	9	PF05208
44574	chr_4:944634-947670	α(1,2)-ManT ALG 9	556	Y	N	10	PF03901
44425	chr_4:509561-511432	α(1,2)-ManT ALG 12	581	Y	Y	6	PF03901
44117	chr_3:1027730-1029328	α(1,3)-GlcT ALG 6	532	Y	Y	11	PF03155
44905	chr_5:604679-606258	α(1,3)-GlcT ALG 8	436	Y	Y	9	PF03155
45980	chr_8:845268-846314	P-Dol GlcT ALG 5	348	Y	Y	1	PF00535
19705	chr_6:554632-555416	P-Dol ManT (DPM1)	236	Y	N	0	PF00535
55197	chr_30:46605-50076	OST (STT3 subunit)	911	Y	Y	10	PF02516
55198	chr_30:50287-53582	OST (STT3 subunit)	894	Y	Y	10	PF02516
50836	chr_5:359528-361787	α-Glc II, subunit α	712	Y	N	0	PF01055
54169	chr_3:489007-491615	α-Glc II, subunit β	803	N	Y	0	PF07915
41172	chr_28:129840-131182	Calreticulin	421	Y	Y	0	PF00262
54787	chr_14:637216-637863	UDP-glucose : glycoprotein glucosyltransferase	499	Y	N	0	PF06427
1815	bd_32x35 : 113900-1161195	α-Man I	666	Y	Y	1	PF01532
54844	chr_16:89066-90505	GnT I	444	Y	Y	1	PF03071
52248	chr_14:60509-61027	α-Man II	1498	Y	N	1	PF01074 PF09261 PF07748
54599	chr_10:692088-693836	α(1,3)-FucT	481	Y	Y	1	PF00852
46109	chr_9:288291-290659	α(1,3)-FucT	770	Y	Y	1	PF00852
46110	chr_9:291303-293777	α(1,3)-FucT	718	Y	Y	1	PF00852

Figure 1

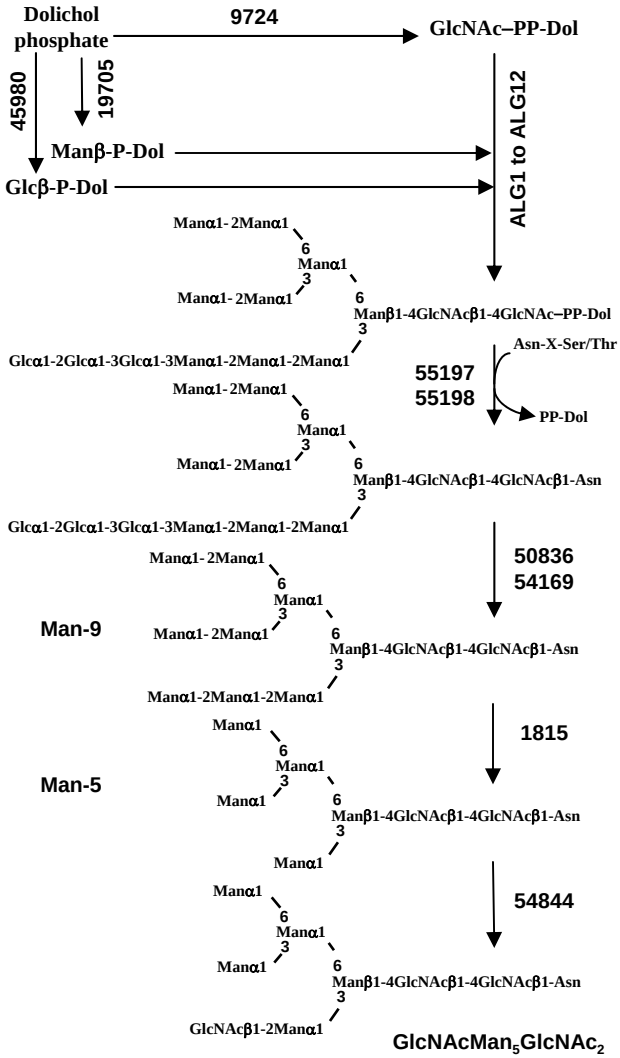

Dolichol phosphate
9724
GlcNAc–PP-Dol
45980
19705
Manβ-P-Dol
Glcβ-P-Dol
ALG1 to ALG12
Manα1- 2Manα1
6
Manα1
3
Manα1- 2Manα1
6
Manβ1-4GlcNAcβ1-4GlcNAc–PP-Dol
3
Glcα1-2Glcα1-3Glcα1-3Manα1-2Manα1-2Manα1
Asn-X-Ser/Thr
55197
55198
PP-Dol
Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn
50836
54169
Man-9
Man-5
1815
54844
Manα1-2Manα1-2Manα1
GlcNAcβ1-2Manα1
GlcNAcMan5GlcNAc2

Figure 2

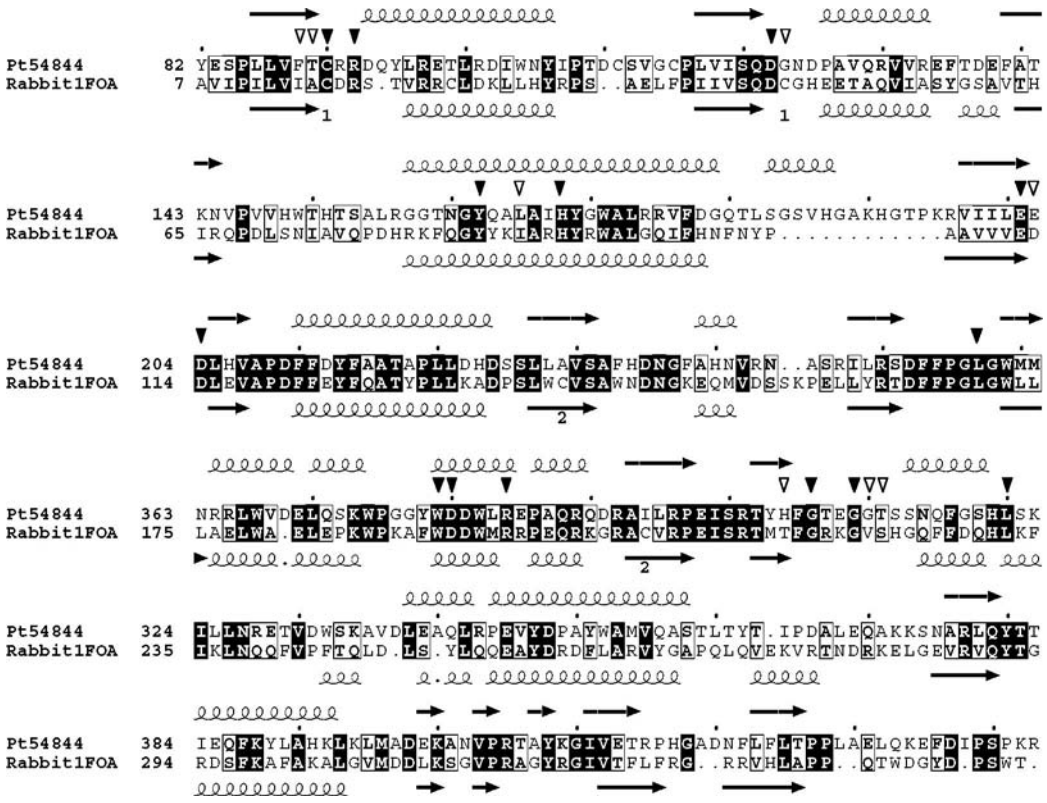

Pt54844 82 YESPLLVFTCRRDQYLRETLRDIWNYIPTDCSVGCPLVISQDGNDPAVQRVVREFTDEFAT
Rabbit1FOA 7 AVIPILVIACDRS.TVRRCLDKLLHYRPS..AELFPIIVSQDCGHEETAQVIASYGSAVTH
Pt54844 143 KNVPVVHWTHTSALRGGTNGYQALAIHYGWALRRVFDGQTLSGSVHGAKHGTPKRVIILEE
Rabbit1FOA 65 IRQPDLSNIAVQPDHRKFQGYYKIARHYRWALGQIFHNFNYP.....AAVVVED
Pt54844 204 DLHVAPDFFDYFAATAPLLDHDSSLLAVSAFHDNGFAHNVRN..ASRILRSDFFPGLGWMM
Rabbit1FOA 114 DLEVAPDFFEYFQATYPLLKADPSLWCVSAWNDNGKEQMVDSSKPELLYRTDFFPGLGWLL
Pt54844 363 NRRLWVDELQSKWPGGYWDDWLREPAQRQDRAILRPEISRTYHFGTEGGTSSNQFGSHLSK
Rabbit1FOA 175 LAELWA.ELEPKWPKAFWDDWMRRPEQRKGRACVRPEISRTMTFGRKGVSHGQFFDQHLKF
Pt54844 324 ILLNRETVDWSKAVDLEAQLRPEVYDPAYWAMVQASTLTYT.IPDALEQAKKSNARLQYTT
Rabbit1FOA 235 IKLNQQFVPFTQLD.LS.YLQQEAYDRDFLARVYGAPQLQVEKVRTNDRKELGEVRVQYTG
Pt54844 384 IEQFKYLAHKLKLMADEKANVPRTAYKGIVETRPHGADNFLFLTPPLAELQKEFDIPSPKR
Rabbit1FOA 294 RDSFKAFAKALGVMDDLKSGVPRAGYRGIVTFLFRG..RRVHLAPP..QTWDGYD.PSWT.

Figure 3

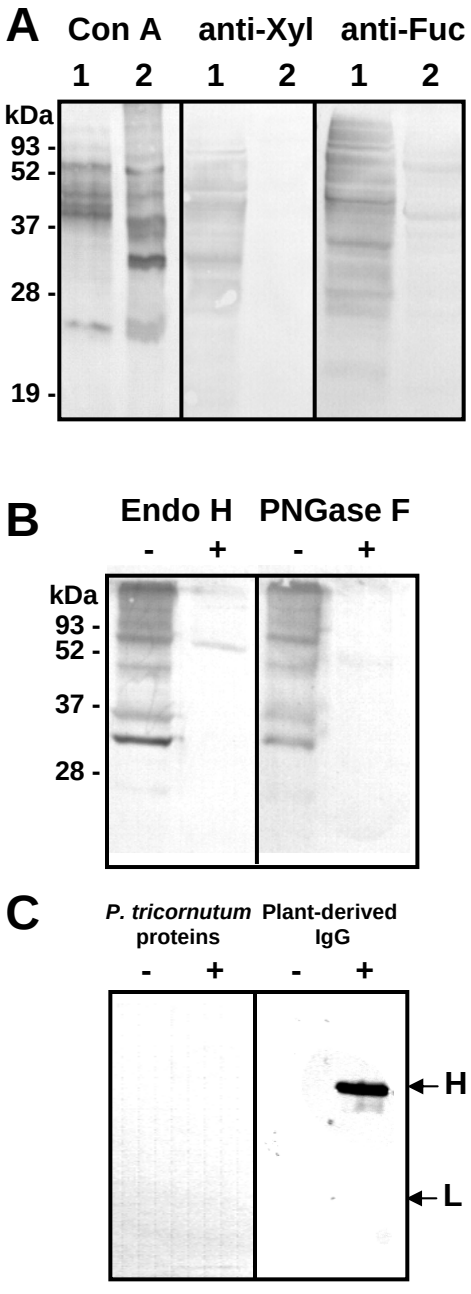

A
Con A
anti-Xyl
anti-Fuc
1 2 1 2 1 2
kDa
93 -
52 -
37 -
28 -
19 -
B
Endo H
PNGase F
- + - +
kDa
93 -
52 -
37 -
28 -
C
P. tricornutum proteins
Plant-derived IgG
- + - +
H
L

Figure 4

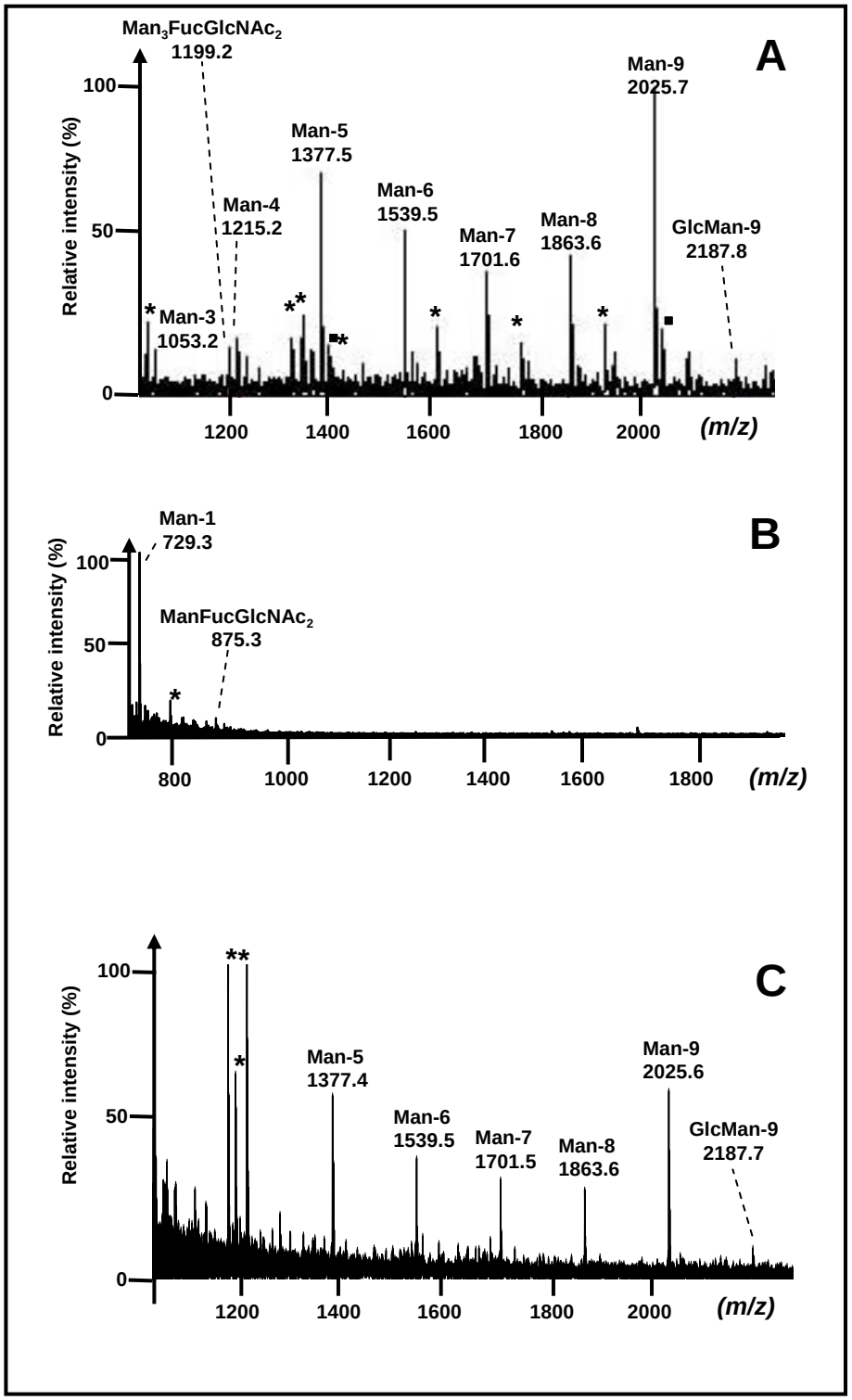

A
Man$_3$FucGlcNAc$_2$
1199.2
Man-9
2025.7
Man-5
1377.5
Man-6
1539.5
Man-4
1215.2
Man-7
1701.6
Man-8
1863.6
GlcMan-9
2187.8
Man-3
1053.2
Relative intensity (%)
100
50
0
1200
1400
1600
1800
2000
(m/z)
B
Man-1
729.3
ManFucGlcNAc$_2$
875.3
Relative intensity (%)
100
50
0
800
1000
1200
1400
1600
1800
(m/z)
C
Man-5
1377.4
Man-9
2025.6
Man-6
1539.5
Man-7
1701.5
Man-8
1863.6
GlcMan-9
2187.7
Relative intensity (%)
100
50
0
1200
1400
1600
1800
2000
(m/z)

Figure 5

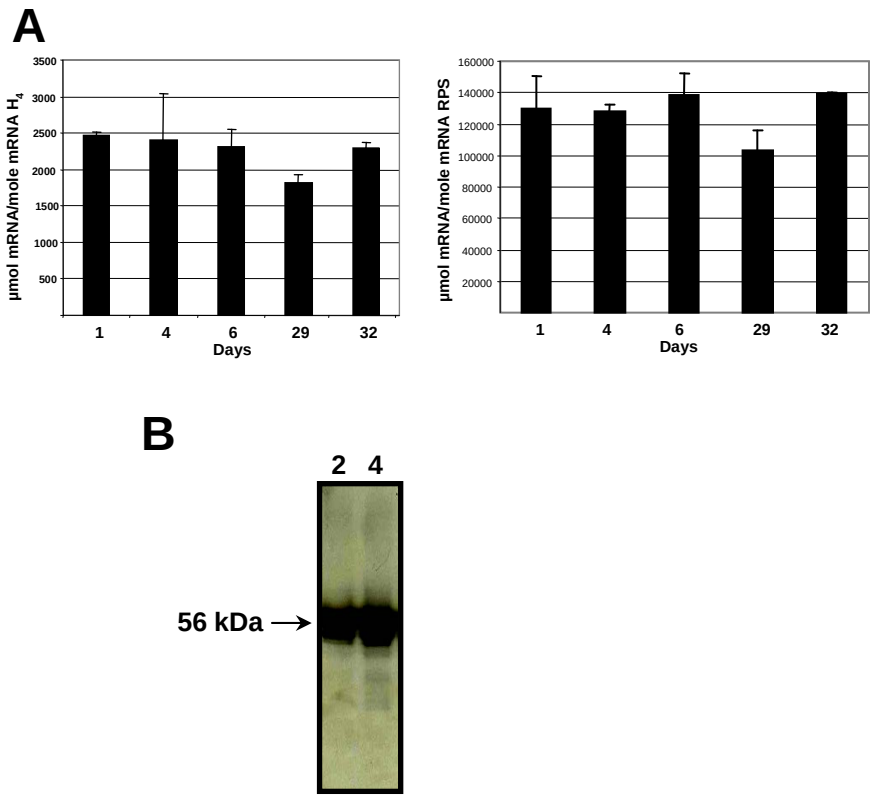

A
µmol mRNA/mole mRNA H4
3500
3000
2500
2000
1500
1000
500
1
4
6
29
32
Days
µmol mRNA/mole mRNA RPS
160000
140000
120000
100000
80000
60000
40000
20000
1
4
6
29
32
Days
B
2
4
56 kDa

Figure 6

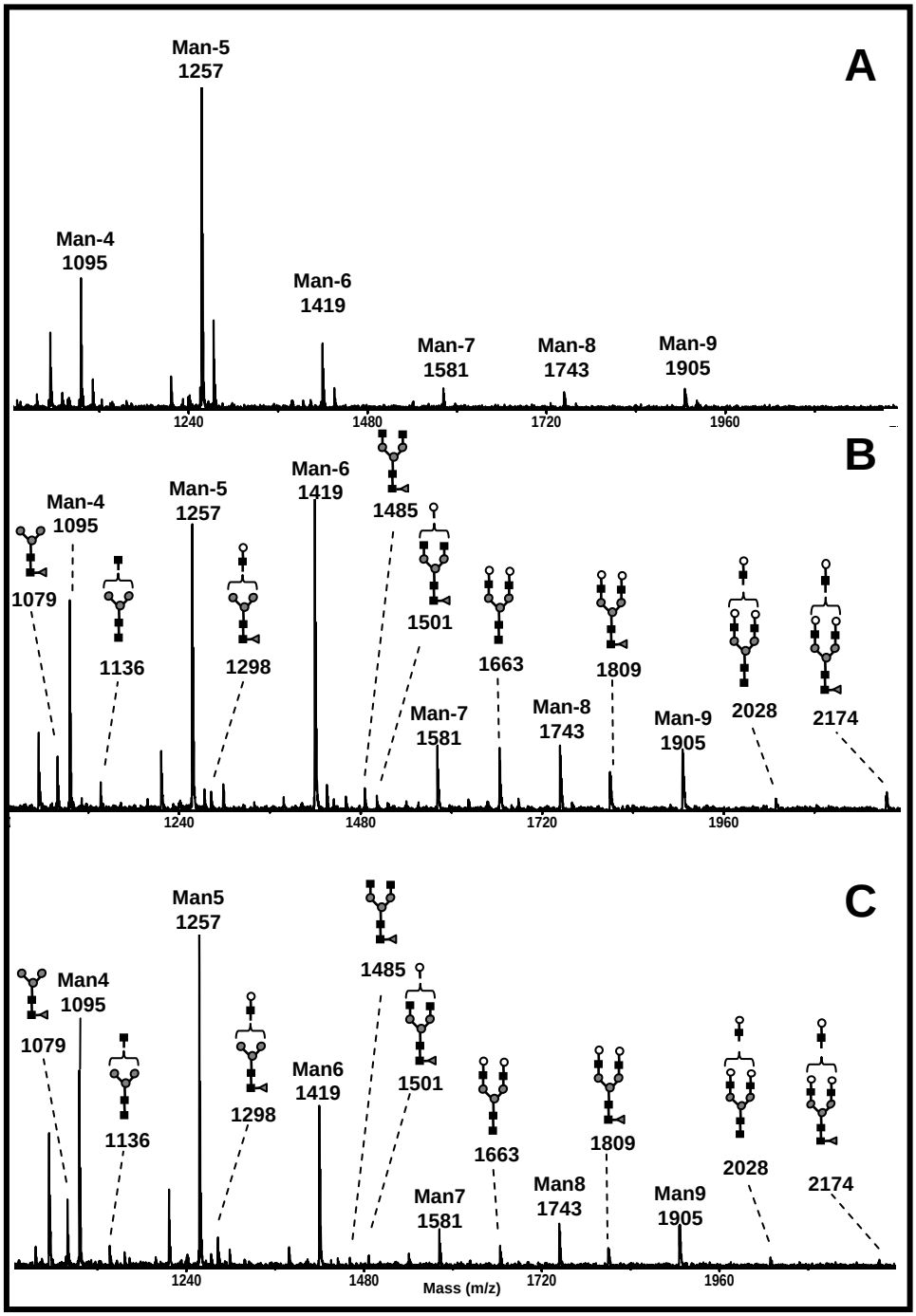
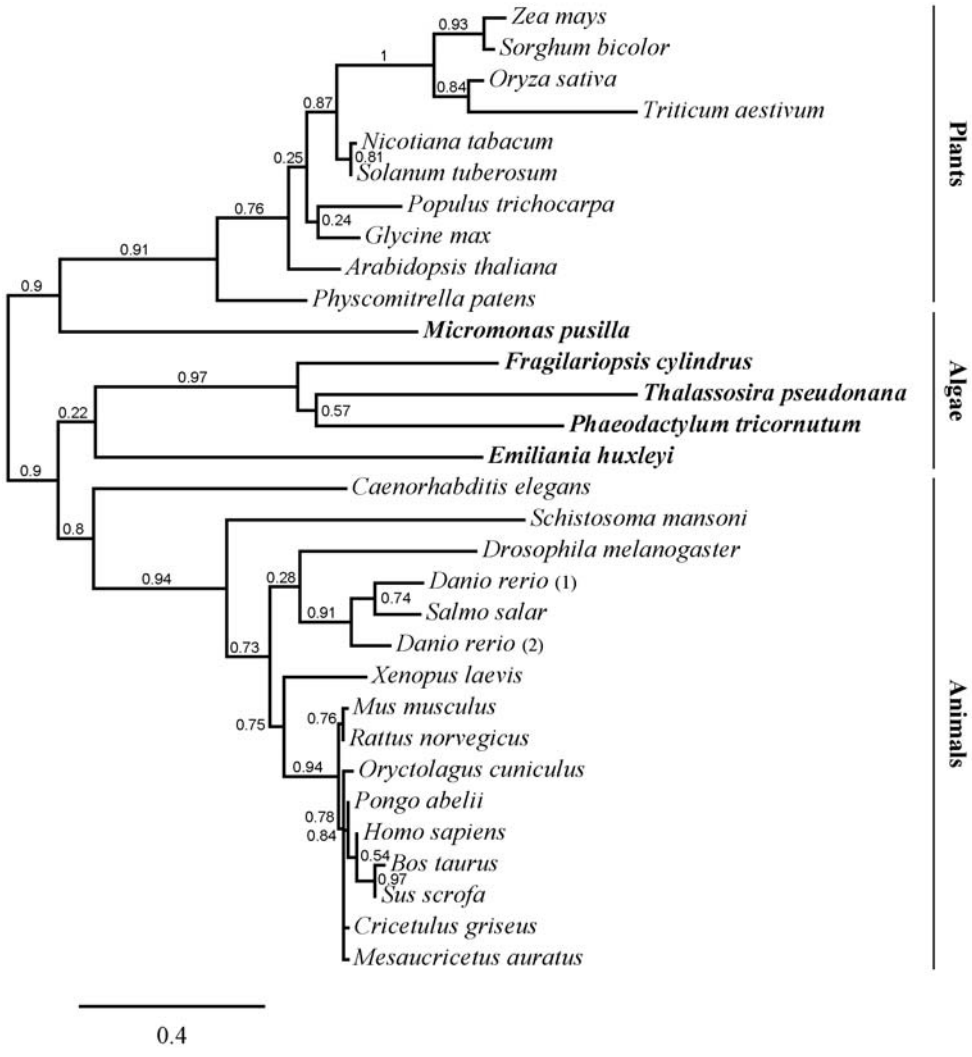

A
Man-5
1257
Man-4
1095
Man-6
1419
Man-7
1581
Man-8
1743
Man-9
1905
1240
1480
1720
1960
B
Man-6
1419
1485
Man-4
1095
Man-5
1257
1079
1136
1298
1501
1663
1809
2028
2174
Man-7
1581
Man-8
1743
Man-9
1905
1240
1480
1720
1960
C
Man5
1257
1485
Man4
1095
1079
1136
1298
Man6
1419
1501
1663
1809
2028
2174
Man7
1581
Man8
1743
Man9
1905
1240
1480
Mass (m/z)
1720
1960

Figure 7


Zea mays
Sorghum bicolor
Oryza sativa
Triticum aestivum
Nicotiana tabacum
Solanum tuberosum
Populus trichocarpa
Glycine max
Arabidopsis thaliana
Physcomitrella patens
Micromonas pusilla
Fragilariopsis cylindrus
Thalassosira pseudonana
Phaeodactylum tricornutum
Emiliania huxleyi
Caenorhabditis elegans
Schistosoma mansoni
Drosophila melanogaster
Danio rerio (1)
Salmo salar
Danio rerio (2)
Xenopus laevis
Mus musculus
Rattus norvegicus
Oryctolagus cuniculus
Pongo abelii
Homo sapiens
Bos taurus
Sus scrofa
Cricetulus griseus
Mesaucricetus auratus
Plants
Algae
Animals
0.4