

A genetic and structural analysis of the *N*-glycosylation capabilities of rice and other monocotyledons

Renaud Léonard, Daniel Kolarich, Katharina Paschinger, Friedrich Altmann and Iain B. H. Wilson*

*Department für Chemie, Universität für Bodenkultur, Muthgasse 18, A-1190 Wien, Austria (*author for correspondence; e-mail iain.wilson@boku.ac.at)*

Received 3 April 2004; accepted in revised form 21 July 2004

Key words: fucosyltransferase, glycan, glucosyltransferase, monocotyledon, *N*-acetylglucosaminyltransferase, oligosaccharide, rice, xylosyltransferase

Abstract

The recent draft sequencing of the rice (*Oryza sativa*) genome has enabled a genetic analysis of the glycosylation capabilities of an agroeconomically important group of plants, the monocotyledons. In this study, we have not only identified genes putatively encoding enzymes involved in *N*-glycosylation, but have examined by MALDI-TOF MS the structures of the *N*-glycans of rice and other monocotyledons (maize, wheat and dates; *Zea mays*, *Triticum aestivum* and *Phoenix dactylifera*); these data show that within the plant kingdom the types of *N*-glycans found are very similar between monocotyledons, dicotyledons and gymnosperms. Subsequently, we constructed expression vectors for the key enzymes forming plant-typical structures in rice, *N*-acetylglucosaminyltransferase I (GlcNAc-TI; EC 2.4.1.101), core α 1,3-fucosyltransferase (FucTA; EC 2.4.1.214) and β 1,2-xylosyltransferase (EC 2.4.2.38) and successfully expressed them in *Pichia pastoris*. Rice GlcNAc-TI, FucTA and xylosyltransferase are therefore the first monocotyledon glycosyltransferases involved in *N*-glycan biosynthesis to be characterised in a recombinant form.

Abbreviations: MALDI-TOF MS, matrix-assisted laser desorption/ionisation time-of-flight spectrometry

Introduction

The most important food crops in the world are monocotyledons, in particular rice, wheat, sorghum and maize are key for the nutrition of man and farm animals. It is therefore not a surprise that rice, having the smallest monocotyledon genome, has, in addition to the many EST projects across the plant kingdom, been the second major target

for plant gene sequencing projects (Bancroft, 2002; Goff *et al.*, 2002; Yu *et al.*, 2002). However, not all the rice genome sequence information has been fully released to the public and the annotations are at an early stage, although these data are not just of commercial, but also of wider evolutionary, interest. This information also opens the way for the use of monocotyledons as protein factories, including the production of glycosylated proteins, and it is in this context that examining the glycosylation capabilities of these plants gains a new significance.

In plants in general, it is now known that the *N*-glycans range from oligomannose structures and core fucosylated/xylosylated paucimannosidic

The full rice GlcNAc-TI cDNA sequence, the partial rice core α 1,3-fucosyltransferase cDNA sequence and rice β 1,2-xylosyltransferase exon 1 genomic and partial cDNA sequences have been deposited in the EMBL nucleotide database under the accession numbers AJ457976, AJ582955, AJ621917 and AJ621918.

structures through to forms carrying not only the core modifications, but also Lewis a moieties (Lerouge *et al.*, 1998; Wilson *et al.*, 2001a). In addition, the genes encoding a variety of plant glycosyltransferases have recently been identified from dicotyledons. In particular, several enzymes involved in *N*-glycan biosynthesis have been characterised in recombinant forms: *N*-acetylglucosaminyl-1-phosphotransferase (Koizumi *et al.*, 1999), GlcNAc-TI (Bakker *et al.*, 1999; Strasser *et al.*, 1999a; Wenderoth and von Schaewen, 2000), GlcNAc-TII (Strasser *et al.*, 1999b), core α 1,3-fucosyltransferase (Leiter *et al.*, 1999; Wilson *et al.*, 2001b), β 1,2-xylosyltransferase (Strasser *et al.*, 2000; Pagny *et al.*, 2003) and Lewis-type α 1,4-fucosyltransferase (Bakker *et al.*, 2001; Wilson, 2001; Léonard *et al.*, 2002). In addition, α 1,2-fucosyl-, α 1,6-galactosyl and α 1,6-xylosyltransferases involved in xyloglucan or galactomannan biosynthesis (Edwards *et al.*, 1999; Sarria *et al.*, 2001; Faik *et al.*, 2002) and a number of UDP-glycosyltransferases involved in secondary metabolism (Lim *et al.*, 2003) have been identified. There are, however, many plant glycosyltransferase homologues of no known function and many linkages for which the genetic basis of their synthesis *in vivo* remains to be elucidated.

Although, we have previously examined some other monocotyledons such as asparagus, banana, coconut and onion (Wilson *et al.*, 2001a), our analyses of *N*-glycans isolated from members of the grass family were previously restricted to three grass pollens (Wilson and Altmann, 1998). Considering, however, the potential for carbohydrate-related allergy (e.g., related to wheat flour; Garcia-Casado *et al.*, 1996), the glycosylation capabilities of important food crops such as rice and wheat should be examined. However, other than studies on the glycosylation of individual glycoproteins from cereals (Hayashi *et al.*, 1990; Johansson *et al.*, 1992; Harthill and Thomsen, 1995; Costa *et al.*, 1997; Kotake *et al.*, 2001), there has been, until recently, little published knowledge on the *N*-glycan structures from these species. In the present study, we extend our analyses of glycans from monocotyledons as well as examine relevant glycosyltransferase homologues in rice. Specifically we have cloned cDNAs encoding rice *N*-acetylglucosaminyltransferase I (GlcNAc-TI), core α 1,3-fucosyltransferase and

β 1,2-xylosyltransferase and expressed them in *Pichia pastoris*.

Experimental procedures

Glycan analysis

Glycans from rice leaves and grains, maize leaves, wheat leaves, banana and date palm fruits were isolated by previously published procedures and analysed by MALDI-TOF MS (Wilson *et al.*, 2001a).

Identification and cloning of rice cDNAs

The Monsanto Rice Genome database (www.rice-research.org) was searched using the tBLASTn program using the sequence of *Arabidopsis N*-acetylglucosaminyltransferase I, core α 1,3-fucosyltransferase and β 1,2-xylosyltransferase. RNA and genomic DNA from rice plants (cv. Nipponbare), raised from seed kindly provided by the International Rice Research Centre (The Philippines), was prepared using Trizol reagent (Invitrogen).

In the case of GlcNAc-TI and the fucosyltransferase, two-step RT-PCR was performed using Superscript or ImmProm reverse transcriptase, with T18 as primer, for the first strand and Expand polymerase mix (Roche Biochemicals) for PCR. cDNA products were purified after gel electrophoresis using the QIAGEN gel purification kit and ligated into the pGEM-T vector (Promega). Positive clones were selected and sequenced using the BigDye kit. The following primer combinations were used to clone rice GlcNAc-TI cDNA: for the full length coding region, RGnTI/1 5'-GGGGAGATCCAAGAG-3' and RGnTI/2/KpnI 5'-CGGGTACCTCTCCACACCTCAACAAC-3'; for the soluble form, RGnTI/3/PmlI 5'-CCGGCACGTGTCCGCCTTTTTTCTACTC-3' and RGnTI/2/KpnI. To generate the expression vector, a pGEM-T clone containing the RGnTI/3/PmlI-RGnTI/2/KpnI fragment was digested with Pml I and Kpn I and ligated into pGAP α C cut with the same enzymes.

For rice FucTA soluble form, the primers RFTA/5/PmlI CCGGCACGTGGCCGCCTCGA CATG and RFTA/6/KpnI CGGGTACCGAAAA TGTCTAGAGCATAG were used. For engineer-

ing the expression vector, a pGEM-T clone containing the RFucTA/7/PmlI-RFucTA/6/KpnI PCR fragment corresponding to the catalytic and stem regions of FucTA was then digested with *Pml*I and *Kpn*I and ligated into pGAP α C cut with the same enzymes.

The open reading frame of rice xylosyltransferase was cloned in the two-step procedure using both appropriate fragments of rice cDNA and genomic DNA as PCR templates. cDNA prepared from 10 days old rice seedlings was used as the template for reverse transcription with the *C.therm.* polymerase GC-rich system (Roche; betain concentration 1 M). The DNA sequence encoding the *N*-terminal part of the soluble rice xylosyltransferase, corresponding to the first exon, but lacking the 5'-end encoding the putative cytoplasmic tail and the transmembrane domain, was obtained by PCR of genomic DNA with the primers RXT/5'/EcoRI, 5'-GGAATTCCAC-CAGTCGCCGCACCC-3', and RXTinv13', 5'-CTGCGAGCAGTGTAGCTCG-3'. The two last exons were amplified from cDNA with the primers RXTinv15', 5'-TGGGTTGAAGAACCAACACT-TC-3', and RXT/3'/KpnI, 5'-GGGGTACCCCTCAGCATCCAAAACCTTTCAG-3'. In both cases the PCR was performed with *Pfu* polymerase (Promega). After purification through GFX columns (Pharmacia), the two DNA fragments were respectively cut by *Eco*RI and *Kpn*I and ligated together with T4 ligase (Promega). The resulting ligation product was used as a template for a PCR using the primers RXT/5'/EcoRI and RXT/3'/KpnI. After digestion with *Eco*RI and *Kpn*I, the obtained DNA fragment was ligated into pGAPZ α A cut with the same enzymes.

Expression of rice glycosyltransferases

After transformation and screening, preparations of the expression vectors were linearised with *Avr*II and electroporated into *Pichia pastoris* strain GS115. Zeocin-resistant yeast colonies were grown in 2 ml YPD medium in the presence of 50 μ g/ml zeocin at 30 °C overnight prior to, in the case of GlcNAc-TI and xylosyltransferases, changing the temperature to 16 °C for 3–5 days. For the fucosyltransferase, expression was continued for a further 4 days at 30 °C. Aliquots of medium were collected, con-

centrated if required and assayed for the relevant glycosyltransferase activity.

Glycosyltransferase assays

HPLC-based *N*-acetylglucosaminyltransferase I assays, as well as HPLC- and MALDI-TOF MS based core α 1,3-fucosyltransferase and xylosyltransferases assays, were performed basically as previously described (Altmann *et al.*, 1993; Wilson *et al.*, 2001b; Bencúrová *et al.*, 2003). In brief, *N*-acetylglucosaminyltransferase I was assayed using 0.35 nmol pyridylaminated Man₅GlcNAc₂ (derived from *Aspergillus* amy-lase), 5 nmol UDP-GlcNAc, 100 nmol MnCl₂, 10 nmol AMP, 0.5 μ mol GlcNAc in 40 mM MES, pH 6 buffer together with an aliquot of concentrated culture medium of yeast expressing rice GlcNAc-TI (final volume 10 μ l). Incubations were performed at 16 °C and the products analysed by RP-HPLC (buffer A, 100 mM ammonium acetate, pH 4; buffer B, 30% methanol in water) using a gradient of 14–19% B over 20 min. Fractions corresponding to substrate and product were collected and analysed by MALDI-TOF MS.

In the case of rice core fucosyltransferase, dansylated-GnGn glycopeptide derived from human IgG or dansylated GnGn glycopeptide derived from bovine fibrin, carrying biantennary *N*-glycans with two non-reducing terminal *N*-acetylglucosamine residues, were used (Bencúrová *et al.*, 2003). A 0.25 nmol glycopeptide was mixed with 10 nmol GDP-Fuc, 100 nmol MnCl₂ in 40 mM MES, pH 7.5 buffer together with an aliquot of culture medium of yeast expressing rice FucTA (final volume 10 μ l) and incubated at 25 °C. The dansylated glycopeptide products were analysed by RP-HPLC and the dansylated glycopeptides by MALDI-TOF MS.

Recombinant xylosyltransferase was assayed using the dansylated-GnGn glycopeptide derived from bovine fibrin: 0.25 nmol glycopeptide was mixed with 10 nmol UDP-Xyl, 100 nmol MnCl₂ in 40 mM MES, pH 6 buffer together with an aliquot of culture medium of yeast expressing rice xylosyltransferase (final volume 10 μ l) and incubated at 16 °C. The reaction products were analysed by MALDI-TOF MS.

Results

Structural analysis of monocotyledon N-glycans

Previously we analysed N-glycans from a number of plants, including some monocotyledons (Wilson and Altmann, 1998; Wilson *et al.*, 2001a). We extended this analysis to a further species: rice (leaves and grains), maize (leaves), wheat (leaves),

banana (fruit) and date palm (fruit). The results (Figure 1) indicate that all these species synthesise a wide range of previously known plant N-glycans such as oligomannosidic, xylosylated and fucosylated paucimannosidic and complex structures with and without the Lewis a epitope (Lerouge *et al.*, 1998). In contrast to Bardor *et al.* (2003), our data suggest that in rice grains MMX (Figure 1B; m/z 1066) is the dominant glycan species,

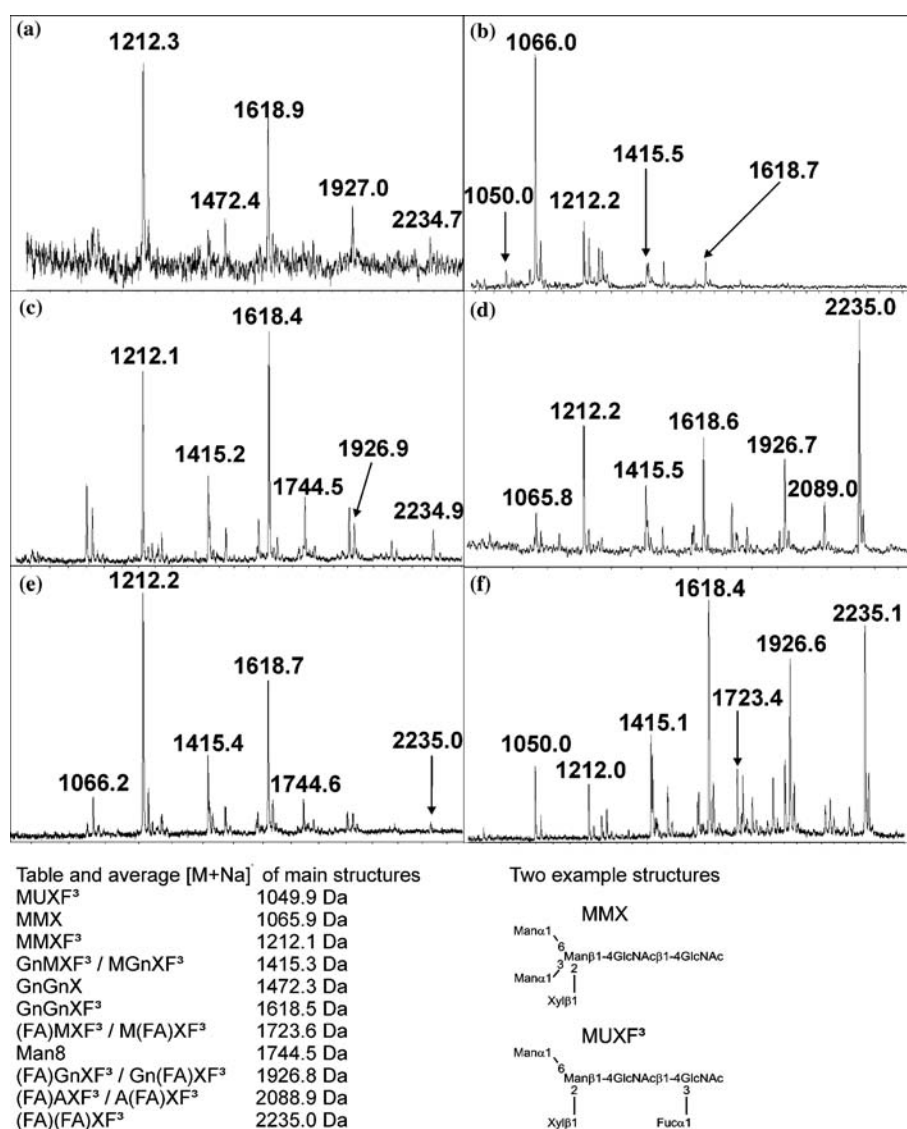


Figure 1. N-glycan analysis of monocotyledons by MALDI-TOF MS. (a), rice leaves; (b), rice seeds; (c), wheat plant; (d), dates (fruit); (e), maize leaves; (f), banana fruit. The m/z values for the major structures are annotated and two example structures are shown: MMX (with two terminal mannose residues and a xylose residue) and MUXF³ (with one α 1,6-linked mannose, no α 1,3-linked mannose, a xylose and a core α 1,3-linked fucose). Gn signifies terminal GlcNAc, A terminal galactose and FA terminal Le^a (previously abbreviated GF). A fuller account of nomenclature for the N-glycan abbreviations is given in Koprivova *et al.* (2004).

whereas we only find trace amounts of MUXF³ (m/z 1049) (see Figure 1 for the structures of MMX and MUXF³). Such dissimilarities may, of course, be due to differences in the exact method of preparation. As previously noted, fruits (Figures 1D and 1F) tend to have larger quantities of Le^a-containing glycans than leaves and seeds (Wilson *et al.*, 2001a), whereas in the leaves of the three cereals, smaller structures tend to predominate.

Searching DNA databanks for glycosyltransferase homologues

By searching the initially Monsanto, and latterly the Syngenta, Rice Genome Databases (both accessible to registered users only), we found that rice has, as expected from the glycan analyses, a large number of homologues of eukaryotic glycosyltransferases (some relevant regions of the Monsanto genome sequences having now been deposited on Genbank). Of particular interest for the present study was the presence of homologues of all known glycosyltransferases required for the biosynthesis of the typical plant-type complex and paucimannosidic *N*-glycans: i.e., GlcNAc-TI, GlcNAc-TII (a putatively single exon gene), core α 1,3-fucosyltransferase and β 1,2-xylosyltransferase. Furthermore, two homologues of the plant Lewis-type α 1,4-fucosyltransferase in the rice genome have been previously noted (see also: Wilson *et al.*, 2001b). The rice genome also encodes a number of β 1,3-glycosyltransferase homologues, one or some of which could conceivably be the enzyme responsible for the addition of β 1,3-galactose to *N*-glycans (so facilitating the action of the α 1,4-fucosyltransferase) or to arabinogalactans. The sequences of some β 1,3-glycosyltransferase homologues from rice have been deposited by us in the Genbank/EMBL database (AJ506738 to AJ506741).

Other intriguing homologies include the presence of three sialyltransferase homologues (two are seemingly orthologous to *Arabidopsis* STL1 and STL2, the third being a paralogue and not an obvious orthologue of *Arabidopsis* STL3) and the presence of an *N*-acetylglucosaminyltransferase III homologue. The presence of sialylated *N*-glycans is yet to be proven in plants and it is possible that these sialyltransferase-like proteins utilise other donors; however, a recent report claims that sialylated glycoconjugates may be present in cul-

tured *Arabidopsis* cells (Shah *et al.*, 2003). Similarly, the present and other studies have not indicated that plants possess *N*-glycans with bisecting *N*-acetylglucosamine; it is, though, possible that the (widespread) plant GlcNAc-TIII homologues synthesise completely different β 1,4-glycosidic linkages.

Sequence analysis of rice GlcNAc-TI

The previously described analysis of the rice genome resulted in identification of a single GlcNAc-TI gene (see Monsanto sequences OJ990920 and OJ000313 as well as the recently deposited chromosome 2 accessions AP004082 and AP005290). A corresponding rice EST (C96709) covering (with frameshifts due to sequencing errors) 750 bp of the 3'-end of the predicted cDNA was also found. cDNA fragments encoding the full length and soluble forms of rice GlcNAc-TI were readily isolated by RT-PCR. The full-length rice GlcNAc-TI is predicted to be a protein of 442 residues (Figure 2) with no potential *N*-glycosylation site and is encoded across eighteen exons (ca. 5 kb of genome from ATG to TAG). This exon structure is entirely conserved by comparison to that of *Arabidopsis* and the first six exons are also conserved by comparison to a partial *Medicago* genomic sequence (AC121239), but in humans the coding region is on one exon only (AC022413). In *Drosophila*, the coding region is split over seven exons with most of the intron/exon junctions corresponding to those in plants, albeit the first fly exon is equivalent to the first six plant exons.

The GlcNAc-TI from rice is highly conserved as compared to the corresponding sequence from wheat (reconstructed from various ESTs) that is 90% identical at the protein level. The overall percentage identity to other plant GlcNAc-TI sequences is around 70% as compared to *Arabidopsis* (Bakker *et al.*, 1999), tobacco (Strasser *et al.*, 1999a), potato (Wenderoth and von Schaewen, 2000) and soybean (the latter sequence also being reconstructed from overlapping ESTs) and 50% as compared to residues 4–455 of the moss sequence (Koprivova *et al.*, 2003). In contrast, the identity to animal sequences was only about 40% for the region including rice residues 100–435. The only other significant homology was that to human POMGnT (GlcNAc-T1.2) with 34% identity to residues 93–333 of the rice protein.

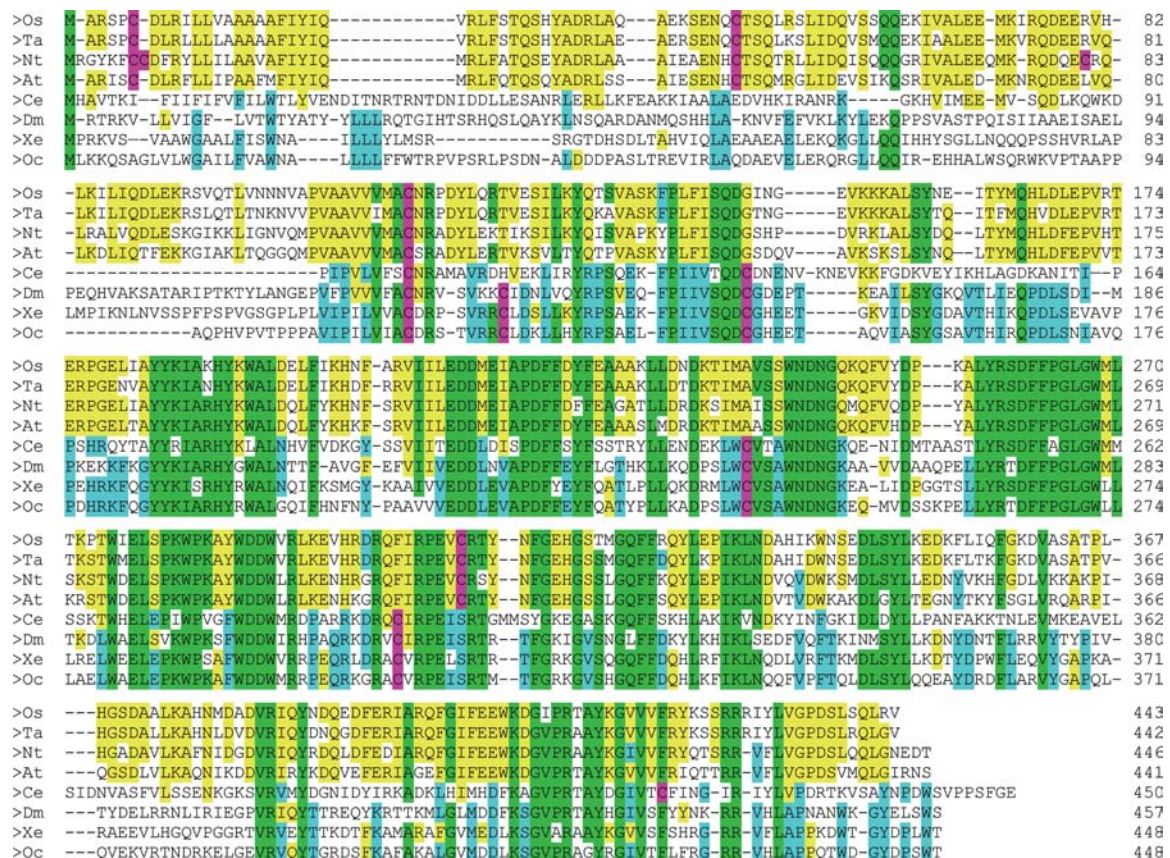


Figure 2. Alignment of selected GlcNAc-TI protein sequences. The aligned sequences are of two monocotyledons (Os, rice – *Oryza sativa*; Ta, wheat – *Triticum aestivum*), two dicotyledons (Nt, tobacco – *Nicotiana tabacum*; At, *Arabidopsis thaliana*), two invertebrates (Ce, *Caenorhabditis elegans* GLY-13; Dm, *Drosophila melanogaster*) and two vertebrates (Xe, frog – *Xenopus laevis*; Oc, rabbit – *Oryctolagus cuniculus*). Yellow indicates residues that are conserved in at least three of the four plant sequences, turquoise residues that are conserved in at least three of the four animal sequences, green are conserved in at least three plant and three animal sequences; pink indicates cysteine residues.

A phylogenetic analysis yields an evolutionary tree compatible with the relationship between plant and animals (data not shown). The divergence between mono- and dicotyledons is also obvious from sequence comparisons of this glycosyltransferase, with a slightly divergent, and net two residues shorter, C-terminus being present in the all grass family species for which GlcNAc-TI sequence information is available (Figure 2). However, in comparison to the conclusions from the crystal structure of rabbit GlcNAc-TI (Ünlügil *et al.*, 2000), many of the residues specifically mentioned as being involved in substrate binding or catalysis are conserved between plants and mammals, including the EDD cation-binding motif (rabbit 211–213) and

the aspartate predicted to be the general base (rabbit 291). Interestingly though, some of the Cys residues that form disulphide bonds in the rabbit structure are absent from plant sequences (e.g., rabbit bridge C115–C145), while a number of residues near to or forming a loop which becomes structured when UDP-GlcNAc binds (those corresponding to rabbit residues 315, 321, 322 and 332) are not conserved. Furthermore the mammalian C123 residue, mutated to arginine in an inactive GlcNAc-TI isolated from one Lec1 cell line (Puthalakath *et al.*, 1996), aligns with a threonine in plant sequences; this suggests that residues required for activity of GlcNAc-TI in one kingdom are not necessarily conserved in another.

Functional expression of rice GlcNAc-TI

In order to show that the rice GlcNAc-TI gene encodes a functional protein, a soluble form of rice GlcNAc-TI was expressed in the yeast *P. pastoris* under the control of the constitutive glyceralde-

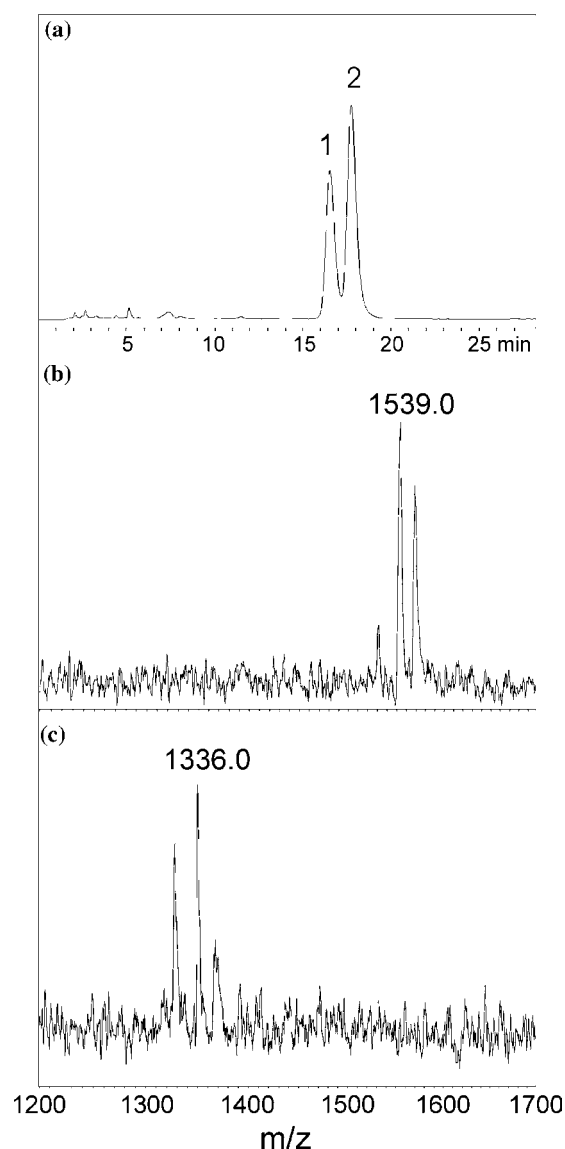


Figure 3. Assay of recombinant rice GlcNAc-TI. $\text{Man}_5\text{GlcNAc}_2\text{-PA}$ was incubated with culture supernatant of *Pichia* transformed with rice GlcNAc-TI in the presence of UDP-GlcNAc and the reaction products were analysed by RP-HPLC (panel A). The collected peaks were then separately analysed by MALDI-TOF MS (peak 1, panel C and peak 2, panel D).

hyde-3-phosphate dehydrogenase promoter; the assay was based on transfer of GlcNAc to a pyridylaminated $\text{Man}_5\text{GlcNAc}_2$ substrate with the measurements performed by HPLC. As compared to controls, only supernatants of *Pichia* transformed with the rice GlcNAc-TI converted the $\text{Man}_5\text{GlcNAc}_2$ substrate as judged by appearance of a slightly earlier eluting peak (Figure 3a). Product and substrate peaks were collected and subject to MALDI-TOF MS analysis (Figure 3b and c, respectively), the product peak displaying a mass increase of 203 in comparison to the substrate. This difference corresponds to the addition of a single *N*-acetylglucosamine residue to $\text{Man}_5\text{GlcNAc}_2$. Control experiments with supernatants of yeast expressing rice FucTA showed no such activity (data not shown).

In initial experiments using unconcentrated culture supernatant, expression (2 days) and assay at 30 °C resulted in 1.8% conversion of substrate after 48 h, whereas expression (4 days) and assay at 16 °C resulted in 4.8% conversion of substrate after 48 h (2.5 mU/l). The improvement in expression and activity at 16 °C, however, is not so pronounced as with *Arabidopsis* β 1,2-xylosyltransferase expressed in the same system (Bencúrová *et al.*, 2003). Due to the relatively low expression, the supernatants of yeast expressing rice GlcNAc-TI were always concentrated after expression in order to reduce assay times and increase enzymatic product amount.

The enzyme appeared to have a broad pH optimum between 4.5 and 7.5; the K_m value for the acceptor was determined to be 50 μM . This is more than ten-fold lower than the values (0.7–5.4 mM) obtained for other recombinant GlcNAc-TI enzymes using a $\text{Man}_3\text{-octyl}$ substrate (Sarkar and Schachter, 2001). However, K_m values for $\text{Man}_3\text{-type}$ substrates can be up to ten-times higher than those with $\text{Man}_5\text{-type}$ substrates as indicated by the values obtained for the native GlcNAc-TI activity from insect Mb-0503 cells (0.2 mM with $\text{Man}_5\text{GlcNAc}_2\text{-PA}$, 2.1 mM with $\text{Man}_3\text{GlcNAc}_2\text{-PA}$; Altmann *et al.*, 1993). Indeed in the present study transfer to $\text{Man}_3\text{GlcNAc}_2\text{-PA}$ was one order of magnitude lower than to $\text{Man}_5\text{GlcNAc}_2\text{-PA}$ (data not shown). A preference for $\text{Man}_5\text{GlcNAc}_2$ is compatible with the normal order of reactions in the Golgi in which GlcNAc-TI acts before α -mannosidase II (Kornfeld and Kornfeld, 1985).

Cloning of partial core α 1,3-fucosyl- and β 1,2-xylosyltransferases

A major characteristic of plant *N*-glycans is the presence of β 1,2-xylose and α 1,3-fucose residues on the trimannosyl core region. cDNAs encoding the two relevant enzymes have been cloned previously from *Arabidopsis* (Strasser *et al.*, 2000; Wilson *et al.*, 2001). Thus it was of interest to clone the orthologous cDNAs from rice. The relevant Monsanto genomic sequences, used to identify the single identifiable orthologues, were OSM130406 for FucTA and OSM19052 for xylosyltransferase. The more recently deposited Genbank sequences have both been mapped to chromosome 8.

For the core α 1,3-fucosyltransferase (FucTA, encoded by the *OsFUT11* gene) it proved possible to isolate a cDNA fragment corresponding to the whole stem and catalytic region: the very 5'-end seemed refractory to being isolated by RT-PCR. The partial fragments showed that the homology-based predictions of the intron/exon structure in these regions (Wilson *et al.*, 2001b) were indeed correct and verify that this gene is transcribed and spliced. The percentage identity is 66% to residues 88–501 of the *Arabidopsis* FucTA and 54% identical to residues 122–533 of the partial moss FucTA (AJ 429145) sequences. All observable intron/exon boundaries are conserved between *Arabidopsis* and rice. An alignment of the rice FucTA sequence with those of other plant core α 1,3-fucosyltransferases has already been published (Wilson *et al.*, 2001b).

In the case of β 1,2-xylosyltransferase, despite various attempts with different primer combinations and 5'-RACE, it was only possible to clone a fragment corresponding to the 3'-end of the coding region of this cDNA: this fragment was considered too short to contain the entire catalytic region – the GC-rich 5'-region was therefore cloned from genomic DNA and then spliced *in vitro* to the 3'-region. The translated *in vitro* reconstructed rice xylosyltransferase fragment is homologous to amino acids 56–535 of *Arabidopsis* β 1,2-xylosyltransferase (59% identity) and 46% identical to 180–583 of the moss protein sequence (Figure 4). The first *N*-glycosylation site in the *Arabidopsis* xylosyltransferase (N53) not required for enzyme activity is absent from all other aligned sequences, whereas the second site (N303) important for full activity is present in the rice enzyme, but absent

from barley, moss, soybean and wheat. In comparison to the relevant rice genomic sequence (now accessible under Genbank number AP004190; chromosome 8), the rice xylosyltransferase cDNA fragment contains both the penultimate and the final exon, a splice site conserved also in the *Arabidopsis* gene. The significance of the GXHGAGLTH motif shared with mammalian cDNAs encoding proteins (see, e.g., Genbank accession AK091089) of unknown function, but carrying KDEL-like sequences at the C-terminus, is unclear.

Functional expression of the core α 1,3-fucosyltransferase

In order to verify that the partial rice FucTA, lacking the transmembrane domain, was indeed enzymatically active, the corresponding cDNA was used to transform *P. pastoris*. The enzyme was assayed by both HPLC and MALDI-TOF based assays (Figure 5). The former assay indicate the conversion of the substrate to a species of lower retention time, compatible with the previously documented effect of the addition of a core α 1,3-fucose residue (Figure 5A). The latter assay shows that an increase in mass (146) only in the presence of both GDP-Fuc and supernatant of yeast transfected with rice FucTA which is also a further indication of the addition of a single fucose residue by the enzyme (Figure 5C). Unlike the rice GlcNAc-TI, rice FucTA displays no particular temperature sensitivity. Its pH optimum is around eight and it has a K_m value estimated to 20 μ M using dansylated GnGn as a substrate. This compares with 67 μ M for *Arabidopsis* FucTA using the same substrate (Wilson *et al.*, 2001b) and with 119 and 200 μ M using an underivatised GnGn fibrin glycopeptide for the recombinant and native forms of the mung bean enzyme (Staudacher *et al.*, 1995; Leiter *et al.*, 1999).

Functional expression of the β 1,2-xylosyltransferase

The two-step cloning procedure previously described allowed the reconstruction of an open reading frame encoding a soluble form of the rice xylosyltransferase. This truncated soluble enzyme started at the 47th amino acid and included the proline rich region described as essential for *Arabidopsis* xylosyltransferase activity (Pagny *et al.*,

```

>Os RVVVEADPSSCKPWRLPSELFWSSASVRF--PASHSCDLYFGNGFSRLVDVLEA----- 82
>Hv HFEVQDFSPCKPWRLPSELFWADPSIF----PSSCEAYFGNGFSRLVVFLEF-----
>Ta -----
>So DGVGASDSSCKPWRLPSELFWSSA-----SEPHLCEAYFGNGFSRLVDVLEA-----
>At HVSNNQSSCKPWRLSELFWPPCRN----LPTGSCCYFGNGFTRVDFLAF----- 97
>Gm -----
>Gh LSNPARFLIPCKEWFTHPSLFWGFSPT----VALNSCEAYFGNGFTRKISNILEV----- 105
>Pt -----
>Pp FQGSGRQSLGKFWELGSAFTSWVDFDLDPD----SCEAFGNGFTTQATLLDPGPHSHARGSVVKLDSRVNS 127

>Os -----EGGGGGWFRCHHSETLRSSICEGARVR 110
>Hv -----GR-GGGGWFRCHHSETLRSSICEGARVR
>Ta -----
>So -----CRGGGGWFRCHHSETLRSSICEGARVR
>At -----RIGGGGCSWFRCHHSETLRSSICEGARVR 127
>Gm -----
>Gh -----KAALRGCSWFRCHHSETLRSSICEGARVR 135
>Pt -----
>Pp KRMQITVRSWRFRFVKQIGGVQVKEEGVSSRCDFQAACMGKQNKVQPSRGGCTACPYGATLRSSICEGNMI 202

>Os LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDAAPKRTG-----PLVAGETDAY 166
>Hv LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR-----
>Ta -----
>So LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR-----PMTGRSGSTTS
>At HVEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFVSSRLGFKRRRFGGEGGSAVSRRLVNDENNEY 202
>Gm -----
>Gh NDESKTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR-----RLVYKFLPKT 190
>Pt -----GGEDLGSVLERNEELPSYSPGAFPIESPHEAIGD-----GRVNPENLENI
>Pp MYEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR-----RALFNKMSRL 259

>Os VPTGGIGMHTMRLLLSRRVPEELHCSQWVEEPTLLVTRFEYANLEHTTIDWYSAYVSSRVTLNPRENVVFV 241
>Hv -----LLESRVPEELHCSQWVEEPTLLVTRFEYANLEHTTIDWYSAYVSSRVTLNPRENVVFV
>Ta -----WEEPTLLVTRFEYANLEHTTIDWYSAYVSSRVTLNPRENVVFV
>So -----
>At MQGGIGMHTMRLLVASIRAVDTNDFVCEWVEEPTLLVTRFEYANLEHTTIDWYSAYVSSRVTLNPRENVVFV 277
>Gm -----RIVRGKDFQDDEWEEPTLLVTRFEYANLEHTTIDWYSAYVSSRVTLNPRENVVFV
>Gh ILEN-VIRHPIRESVRSIVIVGONDFQCEWVEEPTLLVTRFEYANLEHTTIDWYSAYVSSRVTLNPRENVVFV 264
>Pt MPOCAVETHHTMRLLLELRHVESNDISCSQWEEPTLLVTRFEYANLEHTTIDWYSAYVSSRVTLNPRENVVFV
>Pp IFAKSTIRHHTMRLLLELRHVESNDISCSQWEEPTLLVTRFEYANLEHTTIDWYSAYVSSRVTLNPRENVVFV 334

>Os DGHCKAPLEETWEALFNSVIVYAKNFGSPVCFRHAILSPGLGYETALFKGLSESTFCGASAESETRKPDHKTARI 316
>Hv DGHCKAPLEETWEALFNSVIVYAKNFGSPVCFRHAILSPGLGYETALFKGLSESTFCGASAESETRKPDHKTARI
>Ta DGHCKAPLEETWEALFNSVIVYAKNFGSPVCFRHAILSPGLGYETALFKGLSESTFCGASAESETRKPDHKTARI
>So -----
>At DGHCLEETWEALFNSVIVYAKNFGSPVCFRHAILSPGLGYETALFKGLSESTFCGASAESETRKPDHKTARI 352
>Gm DGHCKAPLEETWEALFNSVIVYAKNFGSPVCFRHAILSPGLGYETALFKGLSESTFCGASAESETRKPDHKTARI
>Gh DGHCLEETWEALFNSVIVYAKNFGSPVCFRHAILSPGLGYETALFKGLSESTFCGASAESETRKPDHKTARI 339
>Pt DGHCSPLEDEALFNSVIVYAKNFGSPVCFRHAILSPGLGYETALFKGLSESTFCGASAESETRKPDHKTARI
>Pp DGHCSPLEDEALFNSVIVYAKNFGSPVCFRHAILSPGLGYETALFKGLSESTFCGASAESETRKPDHKTARI 408

>Os SEFGEMTASFTL-----LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDAAPKRTG----- 382
>Hv SEFGEMTASFTL-----LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR-----
>Ta SEFGEMTASFTL-----LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR-----
>So -----
>At SEFGEMTASFTL-----LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR----- 426
>Gm SEFGEMTASFTL-----LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR-----
>Gh SEFGEMTASFTL-----LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR----- 405
>Pt SEFGEMTASFTL-----LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR-----
>Pp SEFGEMTASFTL-----LDEPLTMSRGGEPLDVMGRPEEELPKVDFGALQVDFPAAVLPR----- 473

>Os -----CKCKINIVNGLFAHMSMKQOLRAICASVITGAHGAGLTHIVSATFDTKVLITISSVYRRPHEF 447
>Hv -----CKCKINIVNGLFAHMSMKQOLRAICASVITGAHGAGLTHIVSATFDTKVLITISSVYRRPHEF
>Ta -----CKCKINIVNGLFAHMSMKQOLRAICASVITGAHGAGLTHIVSATFDTKVLITISSVYRRPHEF
>So -----CKCKINIVNGLFAHMSMKQOLRAICASVITGAHGAGLTHIVSATFDTKVLITISSVYRRPHEF
>At -----GSTGTLCKCKINIVNGLFAHMSMKQOLRAICASVITGAHGAGLTHIVSATFDTKVLITISSVYRRPHEF 495
>Gm NY-----CKCKINIVNGLFAHMSMKQOLRAICASVITGAHGAGLTHIVSATFDTKVLITISSVYRRPHEF
>Gh CH-----CKCKINIVNGLFAHMSMKQOLRAICASVITGAHGAGLTHIVSATFDTKVLITISSVYRRPHEF
>Pt -----
>Pp SRSCMKRELGVELSVTHVGLFAHMSMKQOLRAICASVITGAHGAGLTHIVSATFDTKVLITISSVYRRPHEF 548

>Os LISWKSLEYHAINLPGSYARVTDVINELSNILKGLGC 485
>Hv LISWKSLEYHAINLPGSYASPDVTSEISNILKGLGC
>Ta LISWKSLEYHAINLPGSYASPDVTSEISNILKGLGC
>So LISWKSLEYHAINLPGSYARVTDVINELSNILKGLGC
>At LIAWKGLEYHAINLPGSYARVTDVINELSNILKGLGC 533
>Gm LISWKSLEYHAINLPGSYARVTDVINELSNILKGLGC
>Gh LIAWKGLEYHAINLPGSYARVTDVINELSNILKGLGC 509
>Pt -----
>Pp AMSWKSLEYHAINLPGSYARVTDVINELSNILKGLGC 595

```

Figure 4. Alignment of selected xylosyltransferases sequences. The partial reconstructed rice sequence was aligned with other full or partial translated cDNAs or ESTs. Residues conserved in a majority of available sequences are highlighted; gaps due to an absence of sequence information are indicated by dots and gaps due to misalignment are shown by dashes. The following four monocotyledonous, three dicotyledonous, one gymnosperm and one moss sequences were aligned: Os, *Oryza sativa* (rice); Hv, *Hordeum vulgare* (barley); Ta, *Triticum aestivum* (bread wheat); So, *Saccharum officinarum* (noble cane); At, *Arabidopsis thaliana*; Gm, *Glycine max* (Soybean); Gh, *Gossypium hirsutum* (cotton); Pt, *Pinus taeda* (loblolly pine); Pp, *Physcomitrella patens* (moss). The numbering shown commences either from the start codon (*Arabidopsis* and moss) or from the beginning of selected partial sequences (cotton and rice). No numbering is shown in the case of non-contiguous fragmented sequences (barley, cane, pine, soybean and wheat).

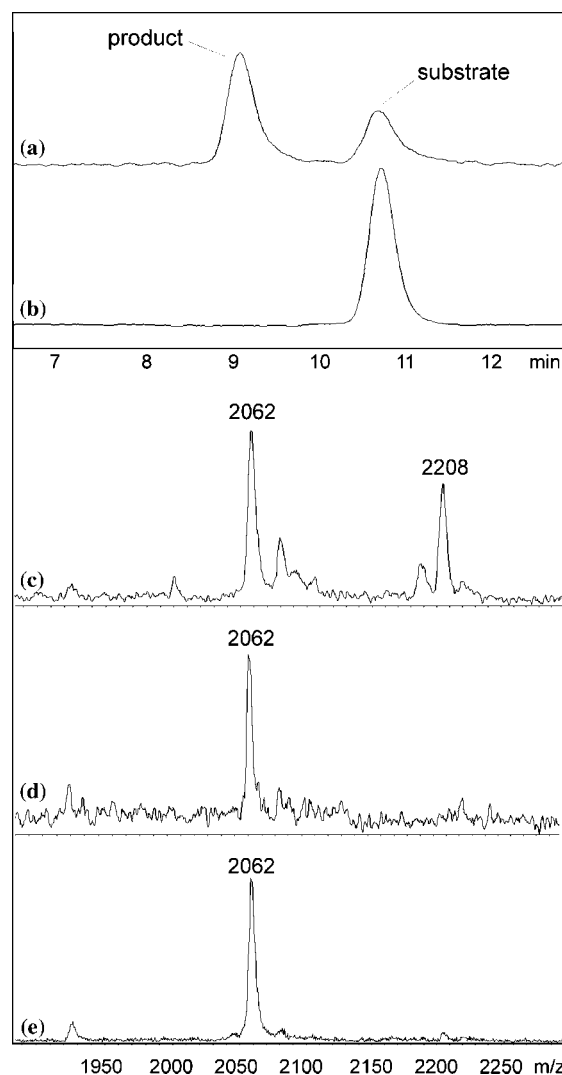


Figure 5. Assay of recombinant rice core α 1,3-fucosyltransferase. Fucosyltransferase assays were performed by either RP-HPLC or MALDI-TOF MS. Panels a and b show RP-HPLC chromatograms with fluorescence detection of the products of dansyl-GnGn incubated in the presence of GDP-Fuc with supernatant of *Pichia* transformed with vectors encoding either FucTA (a) or rice GlcNAc-TI (b; as a negative control). Panels c, d and e show MALDI-TOF mass spectra of the products of dansyl-GnGn incubated in the presence (c,e) and absence (d) of GDP-Fuc with supernatant of *Pichia* transformed with either a vector encoding FucTA (c,d) or a vector lacking an insert (e).

2003). It was expressed in *Pichia* and assayed by MALDI using dansylated GnGn as substrate. Soluble rice xylosyltransferase was found to be secreted in an active form into the culture supernatant when the expression induction was per-

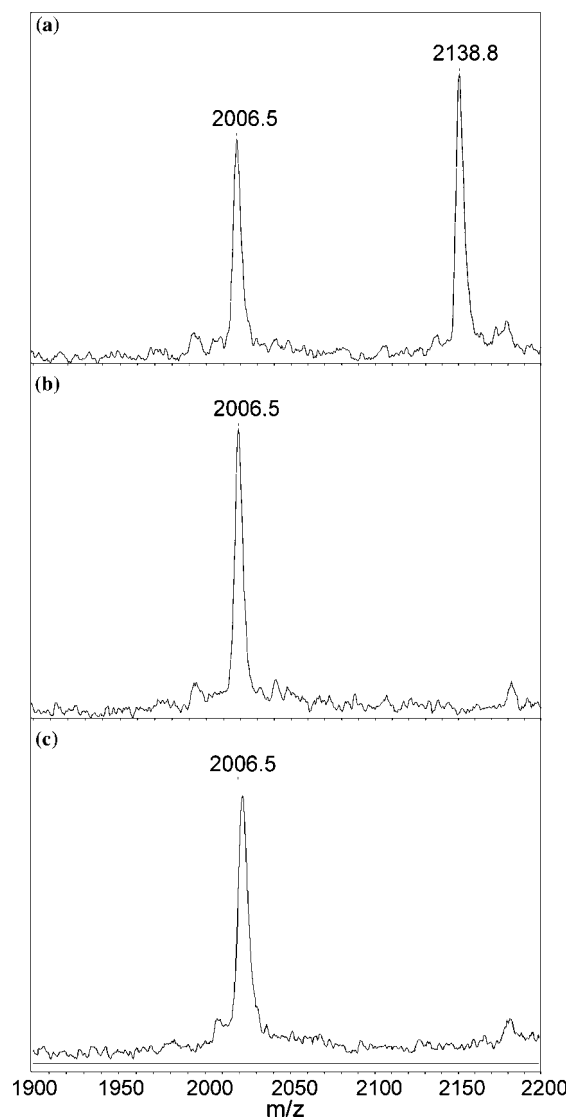


Figure 6. Assay of recombinant rice β 1,2-xylosyltransferase. Dansyl-GnGn was incubated with supernatant of *Pichia* transformed with either vector encoding xylosyltransferase (a, b) or vector lacking an insert (c) in the presence (a, c) and absence (b) of UDP-Xyl. The reaction products were analysed by MALDI-TOF MS.

formed at 16 °C (Figure 6), as judged by the mass increase (132) compatible with addition of a single xylose residue. The induction temperature was also found to be the optimal temperature for the enzyme activity. On the other hand, using a higher temperature, e.g., 30 °C, during induction or assay led to a dramatic decrease in activity. The enzyme is stable for at least 3 months when stored at 4 °C and can be lyophilised. No transfer of xylose was

seen with supernatants of either yeast transformed with vector lacking insert (Figure 6c) or wild-type *Pichia*.

MALDI based assays were also performed in order to analyse the physicochemical parameters affecting the activity of this enzyme. The K_m value for the dansylated GnGn was determined to be 55 μM , which is comparable to that obtained for the purified soybean enzyme (75 μM with a GnGn tyrosinamide oligosaccharide; Zeng *et al.*, 1997). The optimum pH value of the rice xylosyltransferase using MOPS and MES buffer is 6.7, which is compatible with the Golgi cell localisation already described for *Arabidopsis* xylosyltransferase (Dirnberger *et al.*, 2002; Pagny *et al.*, 2003). The enzyme requires the presence of divalent cations such as Mn^{2+} , Mg^{2+} , Ca^{2+} , and Fe^{2+} for full activity whereas Co^{2+} , Zn^{2+} and Cu^{2+} reduce enzyme activity to levels even below those in the presence of EDTA or in the absence of divalent cations.

Discussion

By examining the glycome of monocotyledons, we show by analytical means the presence of xylose, fucose and Lewis a in the grass family – similar to most higher plants, including gymnosperms, analysed to date (Fitchette *et al.*, 1999; Wilson *et al.*, 2001a). One exception to this may be the model organism *Arabidopsis thaliana* in which no Lewis a groups have been detected by mass spectrometry or HPLC (Rayon *et al.*, 1999), although the use of antibodies would suggest that this moiety is indeed present in this plant (Léonard *et al.*, 2002). The ancient origin of these types of *N*-glycans is also shown by their presence in the moss *Physcomitrella patens* (Koprivova *et al.*, 2003; Viëtor *et al.*, 2003) and ginkgo (Kimura and Matsuo, 2000).

Specifically regarding monocotyledons crops, one recent paper on the glycosylation of recombinant human lactoferrin in maize did not show the presence of Lewis a epitopes (Samyn-Petit *et al.*, 2001), but our structural data, as well as the presence of ESTs homologous to α 1,4-fucosyltransferases, indicate that maize can indeed synthesise this structure in its leaves. Furthermore, while this manuscript was in preparation, a paper was published which included data on rice and

maize seed *N*-glycans, as well as recombinant avidin produced in maize seeds, and which also showed the presence of Lewis a in these plants (Bardor *et al.*, 2003) – we, however, not only examined rice seeds, but also maize and rice leaves.

The structural data indicate that monocotyledons should possess the capability to synthesise complex *N*-glycans. In other plants as well as in animals this requires the prior action of GlcNAc-TI, even though many fucosylated and xylosylated glycans in plants (and in invertebrates) lack any non-reducing terminal *N*-acetylglucosamine residues. Previous studies suggest that the requisite xylosyl- and fucosyltransferases only use *N*-glycan substrates which contain a non-substituted *N*-acetylglucosamine residue on the α 1,3-antenna – i.e., the residue which results from the action of GlcNAc-TI – but this GlcNAc residue may be later removed by β -*N*-acetylhexosaminidase(s). In order to examine the glycosylation capabilities of rice at a genetic level, we then searched for relevant genes in the various databases, some of which were or are only with restrictions available to the public.

The presence of a GlcNAc-TI homologue is compatible with the occurrence of complex *N*-glycans in rice – in order to prove the function of this gene, a truncated form of the cDNA was engineered into a *Pichia* expression vector and the thereby expressed protein was found to indeed be active. At the protein sequence level, a high degree of conservation is also discernible. Furthermore, since this gene is orthologous to the GlcNAc-TI sequences of animals, a complete eukaryotic phylogeny can be built up and comparisons with the 3-D structure of rabbit GlcNAc-TI can be made. The determination of the rice GlcNAc-TI sequence may therefore aid the design of mutant GlcNAc-TI forms created to verify mechanistic models of this enzyme's action.

We assume the difficulties in cloning the 5'-fragments of the core fucosyltransferase and xylosyltransferase cDNAs were due to a very heavy bias towards G+C in these parts of the sequences (the phenomenon of such G+C gradients in rice having been reported elsewhere; Yu *et al.*, 2002), whereas the 3'-ends displayed a more balanced distribution of A+T and G+C. Thus even though we did not isolate full length cDNAs encoding rice α 1,3-fucosyl and β 1,2-xylosyltransferases, sequence comparisons to orthologues in other plants can be made. By comparison

to *Arabidopsis thaliana*, the percentage identities within the putative catalytic domains are around 60–70%, whereas in comparison to moss core fucosyl- and xylosyltransferases this figure is around 50%. In addition, the intron/exon structures are conserved between rice and *Arabidopsis*. Furthermore, the rice core α 1,3-fucosyltransferase is the first monocotyledonous one to be described and is the third plant one (and fourth one in any eukaryote) for which there is demonstrated enzymatic activity. The rice β 1,2-xylosyltransferase is only the second such recombinant enzyme for which enzymatic activity has been shown.

In conclusion, we have performed the first concerted combined genomic and structural analysis of the *N*-glycosylation potential of a monocotyledonous crop plant, rice, and find that the glycosylation machinery of a grass family member is rather similar to that of other plants. Indeed, not only rice has the full complement of the relevant genes required for plant-type complex glycosylation, but an analysis of ESTs and short genomic sequences shows that wheat, barley, *Saccharum* and maize all contain gene sequences with homology to those of GlcNAc-TI, FucTA and xylosyltransferase. Considering that some of these species are tetra- or hexaploid, a full analysis of homologies awaits the cloning of full-length cDNAs. Certainly in the case of wheat, there may be multiple GlcNAc-TI genes since at least two types of *N*-terminal region appear to be encoded by various ESTs (data not shown). We would also assume that some of these species also contain multiple core α 1,3-fucosyltransferase and xylosyltransferase genes. In the case of *Arabidopsis*, single GlcNAc-TI and xylosyltransferase genes have been found, but two core α 1,3-fucosyltransferase genes are present due to the partial duplication events that took place during the evolution of brassicas (Blanc *et al.*, 2000; Vision *et al.*, 2000). In rice, however, we only detected single homologues of the three genes.

Our data have obvious repercussions for the use of grass family members as expression systems, due to core xylose and α 1,3-fucose being immunogenic. The recent realisation of a multiple glycosylation knockout in *Arabidopsis* (Strasser *et al.*, 2004) indicates the approach which may, dependent on consumer acceptance and/or on proof that plants that lack complex glycosylation survive

under natural environmental conditions, in the future be applicable to monocotyledons. Indeed, an ambitious long-term aim of reducing the immunogenic or allergenic potential of endogenous and recombinant glycoproteins produced by cereals using ablation or silencing of glycosyltransferase expression is aided by our identification of the relevant genes from rice.

Acknowledgements

This work was supported in part by a Neose Technologies Glycoscience Research Award to I.B.H.W. and by EU grant (contract QLKI-1999-00765) and FWF grant (S8803) to F.A. The authors thank Angela Linder, Thomas Iskratsch, Georg Hinterkörner, David Schuster and Jasmin Drexler, project students in the laboratory, for performing some of the transferase assays and HPLC analyses and to Monika Bencúrová for preliminary experiments on the cloning of xylosyltransferase and for the preparation of rice cDNA. The use of the Monsanto Rice Genome Database (www.rice-research.org) and Syngenta rice database (www.tmmri.org) is gratefully acknowledged.

References

- Altmann, F., Kornfeld, G., Dalik, T., Staudacher, E., and Glössl, J. 1993. Processing of asparagine-linked oligosaccharides in insect cells: *N*-acetylglucosaminyltransferase I and II activities in cultured lepidopteran cells. *Glycobiology* 3: 619–625.
- Bakker, H., Lommen, A., Jordi, W., Stiekema, W. and Bosch, D. 1999. An *Arabidopsis thaliana* cDNA complements the *N*-acetylglucosaminyltransferase I deficiency of CHO Lec1 cells. *Biochem. Biophys. Res. Commun.* 261: 829–832.
- Bakker, H., Schijlen, E., de Vries, T., Schiphorst, W.E., Jordi, W., Lommen, A., Bosch, D. and van Die, I. 2001. Plant members of the α 1 $\rightarrow$ 3/4-fucosyltransferase gene family encode an α 1 $\rightarrow$ 4-fucosyltransferase, potentially involved in Lewis^a biosynthesis, and two core α 1 $\rightarrow$ 3-fucosyltransferases. *FEBS Lett.* 507: 307–312.
- Bancroft, I. 2002. Insights into cereal genomes from two draft genome sequences of rice. *Genome Biol.* 3:Rev. 1015.1–1015.3.
- Bardor, M., Faveeuw, C., Fitchette, A., Gilbert, D., Galas, L., Trottein, F., Faye, L. and Lerouge, P. 2003. Immunoreactivity in mammals of two typical plant glyco-epitopes, core α (1,3)-fucose and core xylose. *Glycobiology* 13: 427–434.
- Bencúrová, M., Rendiač, D., Fabini, G., Kopecky, E.-M., Altmann, F. and Wilson, I.B.H. 2003. Expression of eukaryotic glycosyltransferases in the yeast *Pichia pastoris*. *Biochimie* 85: 413–422.

- Blanc, G., Barakat, A., Guyot, R., Cooke, R. and Delseny, M. 2000. Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* 12: 1093–1101.
- Costa, J., Ashford, D.A., Nimtz, M., Bento, I., Frazão, C., Esteves, C.L., Faro, C.J., Kervinen, J., Pires, E., Veríssimo, P., Wlodawer, A. and Carrondo, M.A. 1997. The glycosylation of the aspartic proteinases from barley (*Hordeum vulgare* L.) and cardoon (*Cynara cardunculus* L.). *Eur. J. Biochem.* 243: 695–700.
- Dirnberger, D., Bencur, P., Mach, L. and Steinkellner, H. 2002. The Golgi localization of *Arabidopsis thaliana* β 1,2-xylosyltransferase in plant cells is dependent on its cytoplasmic and transmembrane sequences. *Plant Mol. Biol.* 50: 273–281.
- Edwards, M.E., Dickson, C.A., Chengappa, S., Sidebottom, C., Gidley, M.J. and Reid, J.S. 1999. Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. *Plant J.* 19: 691–697.
- Faik, A., Price, N.J., Raikhel, N.V. and Keegstra, K. 2002. An *Arabidopsis* gene encoding an α -xylosyltransferase involved in xyloglucan biosynthesis. *Proc. Natl. Acad. Sci. USA* 99: 7797–7802.
- Fitchette, A.-C., Cabanes-Macheteau, M., Marvin, L., Martin, B., Satait-Jeunemaitre, B., Gomord, V., Crooks, K., Lerouge, P., Faye, L. and Hawes, C. 1999. Biosynthesis and immunolocalisation of Lewis a-containing *N*-glycans in the plant cell. *Plant Physiol.* 121: 333–343.
- Garcia-Casado, G., Sanchez-Monge, R., Chrispeels, M.J., Armentia, A., Salcedo, G. and Gomez, L. 1996. Role of complex asparagine-linked glycans in the allergenicity of plant glycoproteins. *Glycobiology* 6: 471–477.
- Goff, S.A., Ricke, D., Lan, T.-H., Presting, G., Wang, R. and Dunn, M. *et al.* 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296: 92–100.
- Harthill, J.E. and Thomsen, K.K. 1995. Analysis of glycan structures of barley (1-3,1-4)- β -D-glucan 4-glucanohydrolase isoenzyme EII. *Plant Physiol. Biochem.* 33: 9–18.
- Hayashi, M., Tsuru, A., Mitsui, T., Takahashi, N., Hanzawa, H., Arata, Y. and Akazawa, T. 1990. Structure and biosynthesis of the xylose-containing carbohydrate moiety of rice α -amylase. *Eur. J. Biochem.* 191: 287–295.
- Johansson, A., Rasmussen, S.K., Harthill, J.E. and Welinder, K.G. 1992. cDNA, amino acid and carbohydrate sequence of barley seed-specific peroxidase BP 1. *Plant Mol. Biol.* 18: 1151–1161.
- Kimura, Y. and Matsuo, S. 2000. Changes in *N*-linked oligosaccharides during seed development of *Ginkgo biloba*. *Biosci. Biotechnol. Biochem.* 64: 562–568.
- Koizumi, N., Ujino, T., Sano, H. and Chrispeels, M.J. 1999. Overexpression of a gene that encodes the first enzyme in the biosynthesis of asparagine-linked glycans makes plants resistant to tunicamycin and obviates the tunicamycin-induced unfolded protein response. *Plant Physiol.* 121: 353–361.
- Kornfeld, R. and Kornfeld, S. 1985. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 54: 631–664.
- Koprivova, A., Altmann, F., Gorr, G., Kopriva, S., Reski, R. and Decker, E.L. 2003. *N*-glycosylation in the moss *Physcomitrella patens* is organised similarly to that in higher plants. *Plant Biol.* 5: 582–591.
- Koprivova, A., Stemmer, C., Altmann, F., Hoffmann, A., Kopriva, S., Gorr, G., Reski, R. and Decker, E.L. 2004. Targeted knockouts of *Physcomitrella* lacking plant-specific immunogenic *N*-glycans. *Plant Biotech. J.* in press.
- Kotake, T., Tonari, A., Ohta, M., Matsuura, F. and Sakurai, N. 2001. Small complex-type *N*-linked glycans are attached to cell-wall bound exo- β -glucanases of both mung bean and barley seedlings. *Physiol. Plant* 112: 308–314.
- Leiter, H., Mucha, J., Staudacher, E., Grimm, R., Glössl, J. and Altmann, F. 1999. Purification, cDNA cloning, and expression of GDP-L-Fuc:Asn-linked GlcNAc α 1,3-fucosyltransferase from mung beans. *J. Biol. Chem.* 274: 21830–21839.
- Léonard, R., Costa, G., Darrambide, E., Lhernould, S., Fleurat-Lessard, P., Carlue, M., Gomord, V., Faye, L. and Maftah, A. 2002. The presence of Lewis a epitopes in *Arabidopsis thaliana* glycoconjugates depends on an active α 4-fucosyltransferase gene. *Glycobiology* 12: 299–306.
- Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fitchette-Lainé, A., Gomord, V. and Faye, L. 1998. *N*-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol. Biol.* 38: 31–48.
- Lim, E.K., Baldauf, S., Li, Y., Elias, L., Worrall, D., Spencer, S.P., Jackson, R.G., Taguchi, G., Ross, J. and Bowles, D.J. 2003. Evolution of substrate recognition across a multigene family of glycosyltransferases in *Arabidopsis*. *Glycobiology* 13: 139–145.
- Pagny, S., Bouissonnie, F., Sarkar, M., Follet-Gueye, M.L., Driouch, A., Schachter, H., Faye, L. and Gomord, V. 2003. Structural requirements for *Arabidopsis* β 1,2-xylosyltransferase activity and targeting to the Golgi. *Plant J.* 33: 189–203.
- Puthalakath, H., Burke, J. and Gleeson, P.A. 1996. Glycosylation defect in Lec1 Chinese hamster ovary mutant is due to a point mutation in *N*-acetylglucosaminyltransferase I gene. *J. Biol. Chem.* 271: 27818–27822.
- Rayon, C., Cabanes-Macheteau, M., Loutelier-Bourhis, C., Salliot-Maire, I., Lemoine, J., Reiter, W.D., Lerouge, P. and Faye, L. 1999. Characterization of *N*-glycans from *Arabidopsis*. Application to a fucose-deficient mutant. *Plant Physiol.* 119: 725–734.
- Samyn-Petit, B., Gruber, V., Flahaut, C., Wajda-Dubos, J.-P., Farrer, S., Pons, A., Desmaizieres, G., Slomianny, M.-C., Theisen, M. and Delannoy, P. 2001. *N*-glycosylation potential of maize: the human lactoferrin used as a model. *Glycoconjugate J.* 18: 519–527.
- Sarkar, M. and Schachter, H. 2001. Cloning and expression of *Drosophila melanogaster* UDP-GlcNAc: α -3-D-mannoside β 1,2-*N*-acetylglucosaminyltransferase I. *Biol. Chem.* 382: 209–217.
- Sarria, R., Wagner, T.A., O'Neill, M.A., Faik, A., Wilkerson, C.G., Keegstra, K. and Raikhel, N.V. 2001. Characterization of a family of *Arabidopsis* genes related to xyloglucan fucosyltransferase I. *Plant Physiol.* 127: 1595–1606.
- Staudacher, E., Dalik, T., Wawra, P., Altmann, F. and März, 1995. Functional purification and characterisation of a GDP-fucose: β -*N*-acetylglucosamine (Fuc to Asn linked GlcNAc) α 1,3-fucosyltransferase from mung beans. *Glycoconjugate J.* 12: 780–786.
- Shah, M.M., Fujiyama, K., Flynn, C.R. and Joshi, L. 2003. Sialylated endogenous glycoconjugates in plant cells. *Natl. Biotechnol.* 21: 1470–1471.
- Strasser, R., Mucha, J., Schwihla, H., Altmann, F., Glössl, J. and Steinkellner, H. 1999a. Molecular cloning and characterisation of cDNA coding for β 1,2-*N*-acetylglucosaminyltransferase I (GlcNAc-TI) from *Nicotiana tabacum*. *Glycobiology* 9: 779–785.
- Strasser R., Steinkellner, H., Borén, M., Altmann, F., Mach, L., Glössl, J. and Mucha, J. 1999b. Molecular cloning of cDNA encoding *N*-acetylglucosaminyltransferase II from *Arabidopsis thaliana*. *Glycoconjugate J.* 16: 787–791.

- Strasser, R., Mucha, J., Mach, L., Altmann, F., Wilson, I.B.H., Glössl, J. and Steinkellner, H. 2000. Molecular cloning and functional expression of β 1,2-xylosyltransferase cDNA from *Arabidopsis thaliana*. *FEBS Lett.* 472: 105–108.
- Strasser, R., Altmann, F., Mach, L., Glössl, J. and Steinkellner, H. 2004. Generation of *Arabidopsis thaliana* plants with complex *N*-glycans lacking β 1,2-linked xylose and core α 1,3-linked fucose. *FEBS Lett.* 561: 132–136.
- Ünlügil, U.M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H. and Rini, J.M. 2000. X-ray crystal structure of rabbit *N*-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J.* 19: 5269–5280.
- Viêtor, R., Loutelier-Bourhis, C., Fitchette, A.C., Margerie, P., Gonneau, M., Faye, L. and Lerouge, P. 2003. Protein *N*-glycosylation is similar in the moss *Physcomitrella patens* and in higher plants. *Planta* 218: 269–275.
- Vision, T.J., Brown, D.G. and Tanksley, S.D. 2000. The origins of genomic duplications in *Arabidopsis*. *Science* 290: 2114–2117.
- Wenderoth, I. and von Schaewen, A. 2000. Isolation and characterisation of plant *N*-acetyl glucosaminyltransferase I (GntI) cDNA sequences. Functional analyses in the *Arabidopsis* cgl mutant and in antisense plants. *Plant Physiol.* 123: 1097–1108.
- Wilson, I.B.H. and Altmann, F. 1998. Structural analysis of *N*-glycans from allergenic grass, ragweed and tree pollens: core α 1,3-linked fucose and xylose present in all pollens examined. *Glycoconjugate J.* 15: 1055–1070.
- Wilson, I.B.H., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C.J.M., Kamerling, J.P. and Altmann, F. 2001a. Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core α 1,3-linked fucose and xylose substitutions. *Glycobiology* 11: 261–274.
- Wilson, I.B.H., Rendić, D., Freilinger, A., Dumić, J., Altmann, F., Mucha, J., Müller, S. and Hauser, M-T. 2001b. Cloning and expression of cDNAs encoding α 1,3-fucosyltransferase homologues from *Arabidopsis thaliana*. *Biochim. Biophys. Acta* 1527: 88–96.
- Wilson, I.B.H. 2001. Identification of a cDNA encoding a plant Lewis-type α 1,4-fucosyltransferase. *Glycoconjugate J.* 18: 439–447.
- Yu, J., Hu, S., Wang, J., Wong, G.K-S., Li, S. and Liu, B. *et al.* 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296: 79–92.
- Zeng, Y., Bannon, G., Thomas, V.H., Rice, K., Drake, R. and Elbein, A. 1997. Purification and specificity of β 1,2-xylosyltransferase, an enzyme that contributes to the allergenicity of some plant proteins. *J. Biol. Chem.* 272: 31340–31347.