

Plant members of the $\alpha 1 \rightarrow 3/4$ -fucosyltransferase gene family encode an $\alpha 1 \rightarrow 4$ -fucosyltransferase, potentially involved in Lewis^a biosynthesis, and two core $\alpha 1 \rightarrow 3$ -fucosyltransferases¹

Hans Bakker^{a,*}, Elio Schijlen^a, Theodora de Vries^b, Wietske E.C.M. Schiphorst^c,
Wilco Jordi^a, Arjen Lommen^d, Dirk Bosch^a, Irma van Die^c

^aPlant Research International, Wageningen University and Research Centre, P.O. Box 16, 6700 AA Wageningen, The Netherlands

^bDepartment of Chemistry and Biochemistry, San Francisco State University, San Francisco, CA 94132, USA

^cDepartment of Molecular Cell Biology, Glycoimmunology Group, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

^dState Institute for Quality Control of Agricultural Products (RIKILT), Wageningen University and Research Centre, P.O. Box 230, 6700 AE Wageningen, The Netherlands

Received 27 July 2001; revised 4 October 2001; accepted 5 October 2001

First published online 17 October 2001

Edited by Marc Van Montagu

Abstract Three putative $\alpha 1 \rightarrow 3/4$ -fucosyltransferase ($\alpha 1 \rightarrow 3/4$ -FucT) genes have been detected in the *Arabidopsis thaliana* genome. The products of two of these genes have been identified in vivo as core $\alpha 1 \rightarrow 3$ -FucTs involved in *N*-glycosylation. An orthologue of the third gene was isolated from a *Beta vulgaris* cDNA library. The encoded enzyme efficiently fucosylates Gal $\beta 1 \rightarrow 3$ GlcNAc $\beta 1 \rightarrow 3$ Gal $\beta 1 \rightarrow 4$ Glc. Analysis of the product by 400 MHz ¹H-nuclear magnetic resonance spectroscopy showed that the product is $\alpha 1 \rightarrow 4$ -fucosylated at the *N*-acetylglucosamine residue. In vitro, the recombinant *B. vulgaris* $\alpha 1 \rightarrow 4$ -FucT acts efficiently only on neutral type 1 chain-based glycan structures. In plants the enzyme is expected to be involved in Lewis^a formation on *N*-linked glycans. © 2001 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: Fucose; Fucosyltransferase; Glycosyltransferase; Lewis^a

1. Introduction

Plants have great potential for the safe production of human proteins for research purposes or pharmaceutical applications. Many of such proteins are glycoproteins, and it is known that cell and species specific glycosylation modulates numerous protein characteristics [1,2]. Obviously, detailed knowledge of the glycosylation potential of plants is impor-

tant to devise strategies for genetic modification of plants to enable the production of proteins with a specific mammalian-like glycosylation. In many plants, Lewis^a containing glycans are found [3–5] whereas typical mammalian Gal $\beta 1 \rightarrow 4$ GlcNAc and Gal $\beta 1 \rightarrow 4$ [Fuc $\alpha 1 \rightarrow 3$]GlcNAc (Lewis^x) structures are lacking. The availability of the complete sequence of the *Arabidopsis thaliana* genome facilitates the identification of plant orthologues of human glycosyltransferase genes. Our aim in this study was to learn the potential of plants to fucosylate complex-type glycans by characterizing enzymes encoded by the putative members of the $\alpha 1 \rightarrow 3/4$ -fucosyltransferase ($\alpha 1 \rightarrow 3/4$ -FucT) gene family that were identified in the *A. thaliana* genome.

Members of the $\alpha 1 \rightarrow 3/4$ -FucT gene family are found in eukaryotes as well as prokaryotes [6]. In general, they catalyze the final and sometimes crucial step in the synthesis of a range of biologically active glycoconjugates. Six genes encoding $\alpha 1 \rightarrow 3/4$ -FucTs have been identified in humans, encoding FucT-III to FucT-VII and FucT-IX [7–16]. These FucTs have distinct acceptor specificities but in general catalyze the transfer of a fucose (Fuc) residue in α -linkage to the C3 of *N*-acetylglucosamine (GlcNAc) in Gal $\beta 1 \rightarrow 4$ GlcNAc (type 2) structures occurring on glycolipids and glycoproteins. FucT-III and FucT-V have a dual $\alpha 1 \rightarrow 3$ and $\alpha 1 \rightarrow 4$ specificity for type 1 or type 2 chains, respectively, and a higher affinity for type 1 (Gal $\beta 1 \rightarrow 3$ GlcNAc) containing structures than for type 2 structures [10,17]. In humans, the $\alpha 1 \rightarrow 3/4$ -FucTs are involved in the synthesis of the Lewis blood group determinants: Lewis^a, sialyl-Lewis^a, Lewis^b, Lewis^x, sialyl-Lewis^x and Lewis^y.

In plants, as well as in some invertebrates, a FucT activity (FucT-C3) has been described that does not occur in mammals and catalyzes the transfer of a Fuc in $\alpha 1 \rightarrow 3$ -linkage to the innermost asparagine-linked GlcNAc of *N*-glycans [18–20]. The immunogenic core $\alpha 1 \rightarrow 3$ -fucosylated *N*-glycan synthesized by this enzyme is an abundant carbohydrate epitope in plants and invertebrates [21–25], and contributes to the carbohydrate-based cross-reactivity that has long been observed in immunoassays of glycoproteins of plants, insects and mollusks [26,27]. Recently the cDNAs for two FucT-C3s have been cloned, from mung bean and *Drosophila me-*

*Corresponding author. Present address: Institut für Physiologische Chemie, Proteinstruktur, Medizinische Hochschule Hannover, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany.

Fax: (49)-511-532 3956..

E-mail address: bakker.hans@mh-hannover.de (H. Bakker).

¹ Nucleotide sequence data are available in the DDBJ/EMBL/GenBank databases under the accession numbers AJ315848 (Bv- $\alpha 1 \rightarrow 4$ FucT), AJ345084 (At-FucT-C3-1) and AJ345085 (At-FucT-C3-2).

Abbreviations: GlcNAc, *N*-acetylglucosamine; Fuc, fucose; FucT-(-C3), (core $\alpha 1 \rightarrow 3$)-fucosyltransferase; EST, expressed sequence tag; LNT, lacto-*N*-tetraose; LNFP II, lacto-*N*-fucopentaose II

lanogaster, and identified as members of the same gene family as the mammalian $\alpha 1 \rightarrow 3/4$ -FucTs [28,29]. In the genome of *A. thaliana*, three putative FucT genes of the $\alpha 1 \rightarrow 3/4$ -FucT family can be detected. While this manuscript was in preparation, Wilson et al. [30] reported that one of these genes (FucTA, At-FucT-C3-1 in this manuscript) encodes a FucT-C3 activity in vitro, whereas expression of FucTB (At-FucT-C3-2 in this manuscript) did not result in any FucT activity in vitro and was suggested to be a redundant non-active FucT-C3. Here, we demonstrate that both *A. thaliana* enzymes show FucT-C3 activity in vivo when expressed in CHO cells. In addition, we report the isolation of an orthologue of the third *A. thaliana* FucT gene from a *Beta vulgaris* (beet) cDNA library. This cDNA encodes an $\alpha 1 \rightarrow 4$ -FucT, catalyzing the transfer of a Fuc residue to the C4 of GlcNAc, with a strict substrate specificity for neutral type 1 chain-based glycan structures.

2. Materials and methods

2.1. Materials

Unlabeled GDP-Fuc was purchased from Boehringer Mannheim. GDP-[14 C]Fuc (250 Ci/mol) was purchased from New England Nuclear Corp. (Boston, MA, USA) and diluted with the unlabeled nucleotide sugar to obtain a specific radioactivity of 5 Ci/mol. Dr. Monica Palcic (University of Alberta, Edmonton, AB, Canada) kindly provided the 8-methoxycarbonyloctyl glycoside acceptors.

2.2. Cloning of plant cDNAs

A cDNA clone encoding At-FucT-C3-1 (accession no. AJ345084) was isolated from an *A. thaliana* selique cDNA library in the mammalian expression vector pABE [31] by a PCR screening method and sibling selection as described previously for the isolation of a GlcNAcT-I cDNA [32]. The cDNA clone was completely sequenced and contained the complete open reading frame as was predicted from the genomic sequence (protein accession no. BAB02969). As the cDNA insert was in the wrong orientation for expression, the plasmid was digested with *Bst*XI and the insert religated in the same vector. An At-FucT-C3-2 cDNA clone (accession no. AJ345085) was isolated by PCR using primers covering the putative start and stop (ttgattcatctcttaggttaatg and tccagaggaatgaaactagacg) codons from selique cDNA. *Bst*XI linkers (Invitrogen) were ligated to the PCR fragment and inserted into *Bst*XI-digested pABE. Sequencing confirmed that the sequence was identical to the prediction from the *A. thaliana* genome (protein accession no. AAG13053). Sequences of At-FucT-C3-1 and At-FucT-C3-2 are also identical to those cloned by Wilson et al. [30] (DNA accession no. AJ404860 (FucTA) and AJ404861 (FucTB), respectively) except that At-FucT-C3-2 is not the alternative spliced form cloned by Wilson et al. [30], but is identical to the described correctly spliced form. Bv- $\alpha 1 \rightarrow 4$ FucT cDNA (accession no. AJ315848) was recognized as an orthologue of At-FucT-3 (protein accession no. AAG52222) in a *B. vulgaris* root expressed sequence tag (EST) library in SK⁺ made within Plant Research International (accession no. AJ315848). The cDNA was sequenced and appeared to contain the complete putative coding region. The cDNA insert was isolated from SK⁺ and introduced into pABE by *Hind*III/*Xba*I digestion and ligation (pABE-Bv- $\alpha 1 \rightarrow 4$ FucT).

2.3. Expression of the plant cDNAs in CHOP8 cells and immunohistochemistry

The putative FucT genes were transiently expressed in CHOP8 cells [33]. Cells were grown and transfected as described before [32,34]. Two days after transfection cells were washed with phosphate-buffered saline (PBS), fixed in 1% glutaraldehyde in PBS, washed with Tris-buffered saline (TBS) (150 mM NaCl, 10 mM Tris-HCl pH 7.5), blocked with 2% milkpowder/TBS and immunostained by a core $\alpha 1 \rightarrow 3$ -Fuc specific antibody in 2% milkpowder/0.1% Tween 20/TBS. This antibody was purified from commercially available rabbit anti-horseradish peroxidase (HRP) antiserum (Rockland, Gilbertsville, PA, USA), that has a dual specificity for the $\beta 1 \rightarrow 2$ -xylose and core $\alpha 1 \rightarrow 3$ -Fuc epitope of HRP N-glycans. By affinity chroma-

tography with honeybee venom phospholipase A2 (Sigma), that carries N-linked glycans with core $\alpha 1 \rightarrow 3$ -Fuc, but differs from HRP by lacking the $\beta 1 \rightarrow 2$ -xylose epitope, a core $\alpha 1 \rightarrow 3$ -Fuc specific antibody was obtained, as described previously by Faye et al. [35]. The secondary antibody was an alkaline phosphatase-conjugated goat anti-rabbit (Jackson ImmunoResearch Laboratories). For color development, Fast-red (Sigma) was used.

2.4. Expression of Bv- $\alpha 1 \rightarrow 4$ -FucT in COS-7M6 cells and FucT activity

COS-7 cells (ATCC 1651) were grown in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum, 2 mM L-glutamine and 1/50 penicillin/streptomycin solution (all from Life Technologies). The cells were transiently transfected with pABE-Bv- $\alpha 1 \rightarrow 4$ -FucT using the calcium phosphate precipitation method as described previously [36]. After 2 days, cells were washed with PBS and collected in 50 mM sodium-cacodylate buffer pH 7.2 at a concentration of 2.5 mg/ml, and assayed for FucT activity. The standard reaction mixture (50 μ l) contained 5 nmol GDP-[14 C]Fuc (4–5 Ci/mol), 2.5 μ mol sodium-cacodylate buffer pH 7.2, 1 μ mol MnCl₂, 0.2 μ mol ATP, 20 nmol acceptor substrate, 0.1% (v/v) Triton X-100 and 20 μ l homogenate. Reactions were incubated for 2 h at 37°C, in which period no more than 10% of the limiting substrate was converted. Incorporation of Fuc was determined as described [37]. Values were corrected for incorporation into endogenous acceptors. Three replicates were used per data point, and the assays were performed twice.

2.5. Product characterization

Lacto-N-tetraose (LNT) (250 nmol) was incubated with 500 nmol GDP-[14 C]Fuc (0.2 Ci/mol), and 250 μ l cell extract under otherwise standard incubation conditions and incubated for 18 h at 37°C. The product was separated from unincorporated nucleotide sugar by ion exchange on a 5 ml column Dowex 1-X8 (Cl⁻ form), isolated on a Spherisorb S5 NH₂ column (Waters, 4.6 \times 250 mm) run in acetonitrile/15 mM KH₂PO₄ (80:20) using a linear gradient (0.3%/min, flow 2 ml/min), and desalted on a column (0.7 \times 45 cm) of Bio-gel P-2 (200–400 mesh) run in water. The fucosylated product was exchanged in D₂O (99.75 atom%, Merck) three times with intermediate lyophilization. Finally, the sample was redissolved in 400 μ l D₂O (99.96 atom% D, Aldrich, Milwaukee, WI, USA). ¹H-nuclear magnetic resonance (NMR) spectroscopy was performed on a Bruker MSL 400 spectrometer operating at 400 MHz at a probe temperature of 300 K. Resolution enhancement was achieved by Lorentzian to Gaussian transformation. Chemical shifts are expressed in ppm downfield from internal sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS), but were actually measured by reference to internal acetone (δ =2.225 ppm in D₂O).

3. Results

3.1. Cloning of plant FucT cDNA clones

Screening of the *A. thaliana* genomic database with mammalian and putative bacterial sequences of the $\alpha 1 \rightarrow 3/4$ -FucT gene family resulted in the identification of three putative $\alpha 1 \rightarrow 3/4$ -FucT genes. All three showed a low overall homology with the known $\alpha 1 \rightarrow 3/4$ -FucTs, but did show a conserved sequence motif ($\alpha 3$ motif II) found in the mammalian and bacterial genes (Fig. 1, [6]). A cDNA clone encoding At-FucT-C3-1 was isolated from an *A. thaliana* selique cDNA library by a PCR screening method and sibling selection as described previously for the isolation of a GlcNAcT-I cDNA [32]. Another cDNA, At-FucT-C3-2, was isolated by PCR using primers covering the putative start and stop codons from selique cDNA. The third putative gene, At-FucT-3, could not be isolated from the same cDNA source using this approach. ESTs containing part of the At-FucT-3 sequence, however, can be found in GenBank. An orthologue of At-FucT-3 was found in an EST library from *B. vulgaris*. The EST clone contained a complete open reading frame with 64% identity to At-FucT-3 and was named Bv- $\alpha 1 \rightarrow 4$ -FucT.

```

At1 M-----GVFSNLRGPKIGLTHEELPVVANGSTSSSSSPSSFKRVSTFLPFCVALVVIIEIGFLCRLDNASLVD 69
At2 M-----GVFSNLRGPRAGATHDEFPATNGSPSSSSSSSKRKLNNLLPLCVALVVIIEIGFLGRDLQVALVD 69
At3 MPM-----RYL-NAMAALLMM-----FFTLILSFTGILEFPASSTSMHSIDPEPKLSDSTSDPFS-DVLVAYKKWDFEV 69
Bv MRS-----KPLINPISITIMLSISIFFILFFSGGFQVPSFASPSI-IDQTIKSN-----KPDFFT-DLLGAFRKWDSOM 67
H3 MPLGAAPQWPWRCLAAFLFOLLVAVC--FFSYLRVSRDDATGSPRAPSGSSRODTT-----PTRPTLLILLWTWPHFI 74

At1 TLTHFFTSSS-----DLKVGSGIEK-QEELRVDSVTYSRDFTKDPIFISGS-NKDFK-SCSVDCVMGFTSDKKPDAAF 142
At2 TLTDFFTQSPSLQSPPARSDRKKIGLFTDRSCDEEWMREDVTSYRDFTKDPIFISGG-EKDFQ-WCSVDCTFGDSSGKTPDAAF 153
At3 GCARFRENHKDAILGNVSSG-----SLOEFGGG-----K-LMKHKVVLKVGWTWIPDNLENLYSCRCMTCLWTKSSV-LADSPD 143
Bv GCLKFKREKYONWGFSNASVS--LQFSRIDEVGGD-----KELKHKHVNVLKVGWTWIPDNLENLYECKCGLSCFWTKSSV-LSDPPD 146
H3 PVA--LRSRCSEMPVGT-----DGHITADRVYQADTVIVHHWDIMSNPKSRL-----PPSPRPQGR 131

                                S      G
                                α-3 motif I
At1 GLSHOPTLSIIRSM-ESAQYYOENNLQAARRKGYDVMTTSLSSDVPVGY---F-----SWAEYDIM-----APVOPKTE 209
At2 GLGOKPGTLSIIRSM-ESAQYYPENDLAQARRRGYDVMTTSLSSDVPVGY---F-----SWAEYDIM-----SPVOPKTE 220
At3 ALLFETTTPLQRRYGDPLRVYMELEAGRKRSRGREDIFISYHAKDDVOTTYAGSLFHNNRNYHISPHKNNVDLYVWSSRCRLPHRD 229
Bv ALLFETTTPLQRRYGDPLRIYMDLEAGRKRSRGREDIFVSYHAKDDVOTTYAGALFHNNRNYVYSSWKKNDTLVYVWSSRCRLPERN 232
H3 WIWNLLEPPNCOHLEALDRYF-----NLMTSYRSDSIFTPYG-----WLEP-----WSGQPAHPPLN 185

                                Y F I EN YITEK
                                α-3 motif II
At1 KALAAAFISNCAARNFR-----LQALEALMKTNVKIDSYGGCHNRNDGVS---EKVE-ALKHYKFLAFENTNEEDYVTEKFF-OS 285
At2 RAIAAAFISNCGARNFR-----LQALEALMKTNIKIDSYGGCHNRNDGKV---DKVE-ALKRYKFLAFENTNEEDYVTEKFF-OS 296
At3 R-LAKSLDLLIPHSFGKC---LNNVGGDLSA---LSMYPECAVHAENAEAKWYDHLHCAMSHYKFLAIENTAVESYVTEKLF-YA 307
Bv D-IAKRLLSVLSHHSFGKC---LNNVGGDKA---LYMYPECAKDPNSAPKWDHLCAMSHYKFLAIENTMTESYVTEKLF-YA 310
H3 L-SAKTELVAWAVSNWPKDSARVRYOSLD-AHLKVDVYG---RSHKPLPK--GTMMETLSRYKFLAFENSLHPDYITEKLWRNA 264

IPV
At1 LVAGSVPVVVGAP--NIEEFAPSPDSFLHIKOMDDVKAOKMKYLAADNPAYNOTLRWKHEGSPDSFKAIDMAAVHSSCRLCIF 369
At2 LVAGSVPVVVGPP--NIEEFAPASDSFLHIKTMEDVPAKRMKYLAAANPAAAYNOTLRWKYEGSPDSFKAIDMAAVHSSCRLCIF 380
At3 LDGSGVPIYFGAS--NVQDFVP-PHSVIDGSKFGSMOELAAAYVKRLGDDPVAYSEYHAWRCGLMGNYGKTRAYSLDTLPCRLC-- 388
Bv LDGSAVPIYFGAP--NVWDFVP-PHSIIDGSKFSSLEELASYKALANDPVAYAEYHAWRCGLVGNYGKTRATSIDLTPCRLC-- 391
H3 LEAWAVPVVLGSPRSNYERFLP-PDAF IHVDDFQSPKDLARYLOELDKDHARYLSYFRWRETLRPSFSGWALDF-----CKAC-- 341

At1 VATRIREQEESPEFKRRPKCKCTRG-SETVYHLYVRERGRFDMESIFLKGDNLTLEAESAVLAKFMSLRYPEIWKKERPASLRGD 454
At2 LATRVREQEEESPFFKRPCKCSRGGSDTVYHVFRERGRFEMESVFLRGKSVTOEALSAVLAKFKSLKHEAVWKKERPGLNKG 466
At3 -----EESIRRGGKNA-GV 401
Bv -----EAVSRGGRSARAL 405
H3 -----WKLQESRYOTVRSIAAWFT 361

At1 GKLRVHGIYPIGLTORQALYNFKFEGNSSLSTHIQNPCKPFEVVFV 501
At2 KELKIHRIYPLGLTORQALYNFKFEGNSSLSSHIONNPCKAFEVVFV 513

```

Fig. 1. Amino acid sequence alignment of three *A. thaliana* members of the $\alpha 1 \rightarrow 3/4$ -FucT gene family (At1 and At2 are FucT-C3-1 and -2, At3 is the *A. thaliana* orthologue of the *B. vulgaris* $\alpha 1 \rightarrow 4$ -FucT), the *B. vulgaris* $\alpha 1 \rightarrow 4$ -FucT (Bv) and the human FucT-III (H3, protein accession no. BAA96390). Amino acids that are conserved in all five members are shaded. Above the alignment, two conserved boxes ($\alpha 3$ motifs) are indicated that were previously identified based on sequence alignments of animal and bacterial FucTs [6]. Within the motifs, the previously indicated most conserved amino acid residues are shown. An * indicates the position of the tryptophan (W) residue, that determines the $\alpha 1 \rightarrow 4$ -FucT specificity of mammalian FucT-III [42]. The presented *A. thaliana* sequences are identical to the predictions from the genome sequence, except that At-FucT-3 was extended with eight amino acids at the N-terminus, as these aligned well with Bv- $\alpha 1 \rightarrow 4$ -FucT.

At-FucT-C3-1 and At-FucT-C3-2 each show 77% mutual similarity and are considered orthologues of the mung bean core $\alpha 1 \rightarrow 3$ -FucT. At-FucT-3 and Bv- $\alpha 1 \rightarrow 4$ -FucT are about 17% identical to At-FucT-C3-1 and At-FucT-C3-2. The similarity was mainly found around a conserved FucT motif ($\alpha 3$ motif II) that was previously identified in animal and bacterial $\alpha 1 \rightarrow 3/4$ -FucTs (Fig. 1, [6]). These data indicate that At-FucT-3 and Bv- $\alpha 1 \rightarrow 4$ -FucT form a distinct subfamily of plant FucTs within the $\alpha 1 \rightarrow 3/4$ -FucT family.

3.2. Expression of the plant cDNAs in CHOP8 cells

The three cDNAs have been inserted in the mammalian expression vector pABE [31] and transiently expressed in CHOP8 cells [33]. CHOP8 is a Chinese hamster ovary (CHO) cell line of the Lec8 complementation group. Lec8 cells are deficient in UDP-galactose transport and were chosen since their cell surface glycans, which lack galactose, are expected to be efficient acceptor substrates for core $\alpha 1 \rightarrow 3$ -FucTs [38]. Transfected cells were stained with an affinity-purified antibody specific for the core $\alpha 1 \rightarrow 3$ -Fuc structure. Expression of either At-FucT-C3-1 or At-FucT-C3-2 cDNA in CHOP8 cells resulted in cells that were recognized by the anti-core Fuc antibody (Fig. 2), indicating that both cDNAs encode core $\alpha 1 \rightarrow 3$ -FucTs. In vitro FucT assays of the transfected CHOP8 cells, using GDP-Fuc as donor substrate and glycopeptides of asialo-agalacto IgG as acceptor as described previously [20], demonstrated that At-FucT-C3-1 indeed showed core $\alpha 1 \rightarrow 3$ -FucT activity, whereas no enzymatic ac-

tivity could be detected using cells transfected with At-FucT-C3-2 cDNA (results not shown).

Expression of the Bv- $\alpha 1 \rightarrow 4$ -FucT cDNA in CHOP8 cells did not result in positive staining of transfected cells, indicating that the Bv- $\alpha 1 \rightarrow 4$ -FucT cDNA does not encode a functionally active core $\alpha 1 \rightarrow 3$ -FucT.

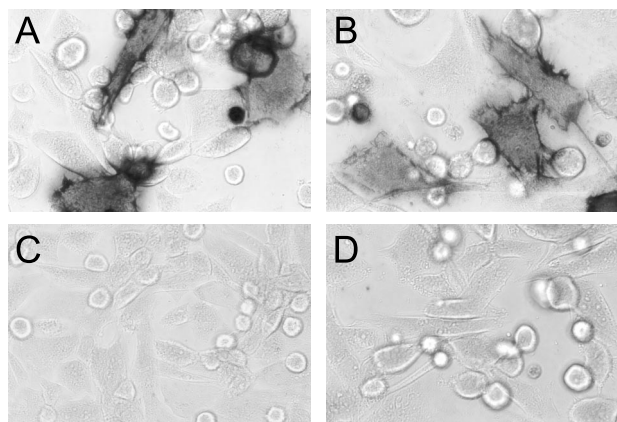


Fig. 2. CHOP8 cells transiently expressed with At-FucT-C3-1 (A), At-FucT-C3-2 (B) and Bv- $\alpha 1 \rightarrow 4$ -FucT (C) and an empty vector control (D). Cells were stained with an antibody specific for core $\alpha 1 \rightarrow 3$ -linked Fuc.

Table 1

Acceptor specificity of recombinant Bv- α 1 \rightarrow 4-FucT towards reducing oligosaccharides and 8-methoxycarbonyloctyl glycoside acceptors

Substrate	FucT activity (%)
Type 1 chain-based oligosaccharides	
1. Gal β 1 \rightarrow 3GlcNAc	42
2. Gal β 1 \rightarrow 3GlcNAc β 1 \rightarrow 3Gal β 1 \rightarrow 4Glc (LNT)	59
3. Gal β 1 \rightarrow 3GlcNAc β -R ^a	59
4. Fuc α 1 \rightarrow 2Gal β 1 \rightarrow 3GlcNAc β -R	100
5. NeuAc α 2 \rightarrow 3Gal β 1 \rightarrow 3GlcNAc β -R	< 0.1
Type 2 chain-based oligosaccharides	
6. Gal β 1 \rightarrow 4GlcNAc	1
7. Gal β 1 \rightarrow 4GlcNAc β 1 \rightarrow 3Gal β 1 \rightarrow 4Glc	0.7
8. Gal β 1 \rightarrow 4GlcNAc-R	2
9. Fuc α 1 \rightarrow 2Gal β 1 \rightarrow 4GlcNAc-R	4
10. NeuAc α 2 \rightarrow 3Gal β 1 \rightarrow 4GlcNAc-R	3

Relative rates for each acceptor (final concentration reducing oligosaccharides: 1 mM and 8-methoxycarbonyloctyl glycoside acceptors 0.4 mM) are expressed as percentage of the incorporation with compound #4, 100% activity: 90 pmol/min/mg.

^aR, O-(CH₂)₈-COOCH₃.

3.3. Characterization of Bv- α 1 \rightarrow 4-FucT

To investigate the enzymatic properties of Bv- α 1 \rightarrow 4-FucT, homogenates of COS-7 cells transfected with pABE-Bv- α 1 \rightarrow 4-FucT were incubated with radiolabeled GDP-Fuc and a panel of oligosaccharide acceptor substrates. COS-7, a monkey kidney cell line, was chosen for these experiments as they facilitate plasmid replication, usually resulting in a higher expression level of recombinant proteins than CHO cells. The results (Table 1) show that the enzyme efficiently acts on acceptor substrates containing neutral type 1 chain-based structures (Gal β 1 \rightarrow 3GlcNAc), presumably producing Lewis^a and Lewis^b determinants, respectively. Bv- α 1 \rightarrow 4-FucT failed to act on sialylated type 1 substrates and all type 2 chain-based acceptor substrates were used with poor efficiency.

To establish the structure of the product synthesized with the recombinant Bv- α 1 \rightarrow 4-FucT, 200 nmol of fucosylated LNT was prepared by incubation of 250 nmol LNT, the transfected COS cells and GDP-Fuc. The product from this large-scale enzyme reaction appeared to be a pentasaccharide, as was observed by high-performance liquid chromatography.

Table 2

400 MHz ¹H-NMR chemical shift values and *J* constants of structural reporter group protons for the constituent monosaccharides of the product synthesized by enzymatic transfer of Fuc by Bv- α 1 \rightarrow 4-FucT to LNT (#F-LNT)

Residue	Reporter group	Anomer	Chemical shifts (<i>J</i> constants) (ppm) (Hz)		
			#F-LNT	LNT	LNFP II
Fuc(1-4)	H-1	α,β	5.026 (3.6)	–	5.026 (4.0)
	H-5	α,β	4.872	–	4.873
	H-6	α,β	1.180 (6.5)	–	1.180 (6.6)
Gal(1-4)	H-1	α	4.438 (7.8)	4.441 (7.7)	4.441 (7.8)
	H-1	β	4.438 (7.8)	4.441 (7.7)	4.436 (7.9)
	H-4	α,β	4.155 (3.0)	4.153 (3.4)	4.154 (3.3)
Gal(1-3)	H-1	α,β	4.506 (7.7)	4.480 (7.8)	4.505 (7.7)
	H-1	α	4.704 (8.3)	4.737 (8.3)	4.707 (8.5)
	H-1	β	4.704 (8.3)	4.732 (8.4)	4.703 (8.5)
GlcNAc	H-3	α,β	4.082 (10.1)	N.D.	4.082 (9.1)
	NAc	α,β	2.033	2.028	2.030
	H-1	α	5.221 (3.7)	5.220 (3.9)	5.220 (3.8)
Glc	H-1	β	4.663 (7.9)	4.662 (7.9)	4.662 (8.0)
	H-2	β	3.280	3.278	3.279

Reference compounds are LNT and LNFP II (LNFP II was purified from human milk according to Donald and Feeney, 1988). Chemical shifts are given in ppm downfield from DSS in D₂O, at 300 K, acquired at 400 MHz (but were actually measured relative to internal acetone, δ =2.225 ppm). N.D., not determined; –, not applicable.

The structure of this pentasaccharide (#F-LNT) was analyzed by 400 MHz ¹H-NMR spectroscopy. Relevant chemical shifts are listed in Table 2, together with the chemical shifts of reference compounds LNT and lacto-*N*-fucopentaose II (LNFP II). Chemical shifts of LNT and LNFP II were essentially as reported before [39,40]. The chemical shifts of the fucosylated product (#F-LNT) appeared to be similar to those of LNFP II. Thus, the Fuc residue appears to be attached to the GlcNAc residue in an α 1 \rightarrow 4-linkage. In the spectrum of #F-LNT no signals were found typical for a Fuc in α 3-linkage to Glc (two doublets of a H-1 of a residue in α -linkage at δ =5.427 and at δ =5.371 ppm; [17]). The H-1 residue of the Gal of the Gal β 1 \rightarrow 4GlcNAc element is shifted downfield in the spectrum of #F-LNT (δ =4.506 ppm), due to the presence of a Fuc residue at the GlcNAc. Furthermore, the H-1 of the α -anomeric Glc residue (δ =5.221 ppm) is not shifted upfield, compared to the H-1 of a Glc which is fucosylated (compounds #F3 and #F10 in [17]). Similarly, in the spectrum of #F-LNT there are no signals indicative for a Fuc in α 2-linkage to Gal, indicating that the Bv- α 1 \rightarrow 4-FucT lacks any α 1 \rightarrow 2-FucT activity.

Thus, the Bv- α 1 \rightarrow 4-FucT can be identified as a GDP-Fuc:GlcNAc α 1 \rightarrow 4-FucT capable of synthesizing Lewis^a and Lewis^b epitopes.

4. Discussion

Three plant members of the α 1 \rightarrow 3/4-FucT gene family have been cloned and, based on immunohistochemical and in vitro enzymatic studies, their function in plant glycosylation is proposed. Two cDNAs with high sequence similarity to the mung bean core α 1 \rightarrow 3-FucT gene were isolated from an *A. thaliana* cDNA library and designated At-FucT-C3-1 and At-FucT-C3-2. We were not successful in isolating the third cDNA from *A. thaliana*, but found and cloned instead an orthologous EST of this gene from the sugar beet, *B. vulgaris*.

Transient transfection of the plant cDNAs in CHOP8 cells resulted in binding of an antiserum specific for core α 1 \rightarrow 3-Fuc to the At-FucT-C3-1 and At-FucT-C3-2, but not to the Bv- α 1 \rightarrow 4-FucT-transfected cells. These data indicate expression of core α 1 \rightarrow 3-fucosylated *N*-glycans on the cell surface

of the antiserum positive cells, suggesting that both cDNAs encode functional core $\alpha 1 \rightarrow 3$ -FucTs. The identity of At-FucT-C3-1 could be established by in vitro FucT assays using a homogenate of the transfected cells. In contrast and for reasons unknown we were not able to demonstrate in vitro core $\alpha 1 \rightarrow 3$ -FucT activity using the At-FucT-C3-2-transfected cells as enzyme source, although staining intensity of transfected cells appeared to be similar. Similarly, Wilson et al. [30] reported the in vitro core $\alpha 1 \rightarrow 3$ -FucT activity of FucTA (=At-FucT-C3-1) and the in vitro inactivity of the correctly spliced form of FucTB (=At-FucT-C3-2) and concluded that FucTB could possibly be a redundant product of duplication of an original FucT-C3 gene. Our immunohistochemistry results, however, strongly favor the possibility that At-FucT-C3-2 is a functionally active core $\alpha 1 \rightarrow 3$ -FucT. The inability to show in vitro activity might be caused by a variety of reasons, i.e. the enzyme may have a low affinity towards the substrates used in vitro, or may be unstable in the assay conditions used. Furthermore, it should be considered that in the CHOP8 cells overexpression of both enzymes is achieved. Possibly, the activity of At-FucT-C3-2 is low compared to that of At-FucT-C3-1, and under the detection limit of the in vitro assay. This activity, however, might be sufficient to fucosylate in vivo. To get insight in the contribution of At-FucT-C3-1 and At-FucT-C3-2 to the fucosylation of *N*-glycans in *A. thaliana*, however, the fucosylation capacity of selective knock-outs of both the respective genes in *A. thaliana* should be investigated.

To define the enzymatic properties of the putative FucT encoded by the Bv- $\alpha 1 \rightarrow 4$ -FucT cDNA, we used COS-7 cells transiently transfected with this cDNA as enzyme source. Our data clearly show that Bv- $\alpha 1 \rightarrow 4$ -FucT has similar properties as the human Lewis-type FucT (FucT-III). Out of a panel of type 1 and type 2 chain containing structures, Bv- $\alpha 1 \rightarrow 4$ -FucT has a strict preference for neutral type 1-based acceptor substrate. Still the substrate specificity of the plant enzyme is distinct from the specificity of the human Lewis enzyme FucT-III [37], since the sialylated type 1 acceptor #5 was not a substrate for the plant enzyme and type 2 acceptors were very poor substrates. The enzyme also lacked the ability to transfer Fuc to the C3 of Glc in LNT [17], as appeared from the structural identification of the product by $^1\text{H-NMR}$ spectroscopy. The enzymatic properties of Bv- $\alpha 1 \rightarrow 4$ -FucT resemble those of a FucT activity purified from mung bean seedling [41], indicating that in mung bean a similar gene is expressed. Interestingly, in human FucT-III, a tryptophan residue (W^{111}) has been identified that is responsible for $\alpha 1 \rightarrow 4$ -FucT specificity [42]. The same residue can be found in FucT-V, the second mammalian enzyme with $\alpha 1 \rightarrow 4$ -FucT activity, but an arginine residue (R) is found at the corresponding position in the mammalian $\alpha 1 \rightarrow 3$ -FucTs. A single amino acid substitution $\text{W}^{111} \rightarrow \text{R}$ in FucT-III was sufficient to change the specificity of Fuc transfer from H-type 1 to H-type 2 acceptors [42]. In agreement with the proposed function of this region in acceptor binding, W^{114} in Bv- $\alpha 1 \rightarrow 4$ -FucT and W^{111} in its *A. thaliana* orthologue align with W^{111} of FucT-III, whereas the two *A. thaliana* core $\alpha 1 \rightarrow 3$ -FucTs did not have either W or R in the corresponding position. Comparison of the Bv- $\alpha 1 \rightarrow 4$ -FucT area around W^{114} showed several more conserved amino acids (see Fig. 1) with FucT-III, that may contribute to the $\alpha 1 \rightarrow 4$ structural motif, although they cannot be found in FucT-V. We were not successful in isolating a full

length cDNA of the putative *A. thaliana* $\alpha 1 \rightarrow 4$ -FucT, although EST fragments have been detected in the databases. Wilson et al. [30], however, have isolated the full coding region of the putative *A. thaliana* $\alpha 1 \rightarrow 4$ -FucT, but were unable to assign a function to the gene, which may be in accordance with previous findings that Lewis^a has not been detected in *Arabidopsis* [3–5].

Lewis^a epitopes, however, are found on many other plants [3–5], and we could show that this structure is readily synthesized in vitro by Bv- $\alpha 1 \rightarrow 4$ -FucT, indicating that the Bv- $\alpha 1 \rightarrow 4$ -FucT gene described here may have a function in the formation of Lewis^a glycan epitopes on *N*-linked glycans in beet. It is not clear yet if the capacity of Bv- $\alpha 1 \rightarrow 4$ -FucT to catalyze the synthesis of the Lewis^b epitope has a functional relevance. So far, the Lewis^b epitope has not been described in plants. A prerequisite for the formation of Lewis^b structures is the coordinate expression of an $\alpha 1 \rightarrow 2$ -FucT that catalyzes the transfer of Fuc in $\alpha 1 \rightarrow 2$ -linkage to Gal in $\text{Gal}\beta 1 \rightarrow 3\text{GlcNAc-R}$, i.e. an orthologue of the human FucTs I and II, belonging to the gene family of $\alpha 1 \rightarrow 2$ -FucTs. So far, such an enzyme has not been described in plants, but several sequences that putatively encode plant members of this gene family can be detected in the databases. Recently, plant $\alpha 1 \rightarrow 2$ -FucTs belonging to this family have been cloned [43,44]. These enzymes, however, seem to catalyze specifically the synthesis of $\alpha 1 \rightarrow 2$ -fucosylated xyloglucan structures.

In conclusion, we have identified two core $\alpha 1 \rightarrow 3$ -FucT genes and one $\alpha 1 \rightarrow 4$ -FucT gene in plants. So far, no other $\alpha 1 \rightarrow 3/4$ -FucT homologues have been found in plants, suggesting that none of the plant FucTs identified so far encodes an enzyme that is able to $\alpha 1 \rightarrow 3$ -fucosylate type 2 chain-based ($\text{Gal}\beta 1 \rightarrow 4\text{GlcNAc}$, LacNAc) structures, thus forming Lewis^x or Lewis^y determinants. So far, no LacNAc structures have been found on plant glycoproteins. We have shown previously that the glycosylation potential of plants can be modified, by introducing the human $\beta 1 \rightarrow 4$ -galactosyltransferase gene in tobacco. This resulted in the formation of *N*-glycans containing typical mammalian LacNAc structures [45], an achievement that is promising for future genetic glycan engineering of plants. Knowing the specificities of the endogenous plant FucTs is important as it facilitates the prediction of the fucosylation of recombinant glycoproteins produced in these plants, as well as the design of additional genetic modifications that will be required to produce for example Lewis^x, Lewis^y and Lewis^b structures in plants.

Acknowledgements: The technical assistance of Ben Bruyneel and Christien Bank in part of the experiments is greatly acknowledged. We thank Dr. Monica Palcic for providing us with the 8-methoxycarbonyloctyl glycoside acceptors.

References

- [1] Cumming, D.A. (1991) *Glycobiology* 1, 115–130.
- [2] Varki, A. (1993) *Glycobiology* 3, 97–130.
- [3] Fitchette-Lainé, A.C., Gomord, V., Cabanes, M., Michalski, J.C., Saint, M.M., Foucher, B., Cavelier, B., Hawes, C., Lerouge, P. and Faye, L. (1997) *Plant J.* 12, 1411–1417.
- [4] Melo, N.S., Nimtz, M., Conradt, H.S., Fevereiro, P.S. and Costa, J. (1997) *FEBS Lett.* 415, 186–191.
- [5] Wilson, I.B., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C.J., Kamberling, J.P. and Altmann, F. (2001) *Glycobiology* 11, 261–274.
- [6] Oriol, R., Mollicone, R., Cailleau, A., Balanzino, L. and Breton, C. (1999) *Glycobiology* 9, 323–334.

- [7] Goelz, S.E., Hession, C., Goff, D., Griffiths, B., Tizard, R., Newman, B., Chi-Rosso, G. and Lobb, R. (1990) *Cell* 63, 1349–1356.
- [8] Kaneko, M., Kudo, T., Iwasaki, H., Ikehara, Y., Nishihara, S., Nakagawa, S., Sasaki, K., Shiina, T., Inoko, H., Saitou, N. and Narimatsu, H. (1999) *FEBS Lett.* 452, 237–242.
- [9] Koszdin, K.L. and Bowen, B.R. (1992) *Biochem. Biophys. Res. Commun.* 187, 152–157.
- [10] Kukowska-Latallo, J.F., Larsen, R.D., Nair, R.P. and Lowe, J.B. (1990) *Genes Dev.* 4, 1288–1303.
- [11] Kumar, R., Potvin, B., Muller, W.A. and Stanley, P. (1991) *J. Biol. Chem.* 266, 21777–21783.
- [12] Lowe, J.B., Kukowska-Latallo, J.F., Nair, R.P., Larsen, R.D., Marks, R.M., Macher, B.A., Kelly, R.J. and Ernst, L.K. (1991) *J. Biol. Chem.* 266, 17467–17477.
- [13] Natsuka, S., Gersten, K.M., Zenita, K., Kannagi, R. and Lowe, J.B. (1994) *J. Biol. Chem.* 269, 20806.
- [14] Sasaki, K., Kurata, K., Funayama, K., Nagata, M., Watanabe, E., Ohta, S., Hanai, N. and Nishi, T. (1994) *J. Biol. Chem.* 269, 14730–14737.
- [15] Weston, B.W., Nair, R.P., Larsen, R.D. and Lowe, J.B. (1992) *J. Biol. Chem.* 267, 4152–4160.
- [16] Weston, B.W., Smith, P.L., Kelly, R.J. and Lowe, J.B. (1992) *J. Biol. Chem.* 267, 24575–24584.
- [17] de Vries, T., Srnka, C.A., Palcic, M.M., Swiedler, S.J., van den Eijnden, D.H. and Macher, B.A. (1995) *J. Biol. Chem.* 270, 8712–8722.
- [18] Staudacher, E., Altmann, F., Glössl, J., März, L., Schachter, H., Kamerling, J.P., Hard, K. and Vliegthart, J.F. (1991) *Eur. J. Biochem.* 199, 745–751.
- [19] Staudacher, E., Dalik, T., Wawra, P., Altmann, F. and März, L. (1995) *Glycoconjug. J.* 12, 780–786.
- [20] van Tetering, A., Schiphorst, W.E., van den Eijnden, D.H. and van Die, I. (1999) *FEBS Lett.* 461, 311–314.
- [21] Haslam, S.M., Coles, G.C., Munn, E.A., Smith, T.S., Smith, H.F., Morris, H.R. and Dell, A. (1996) *J. Biol. Chem.* 271, 30561–30570.
- [22] Khoo, K.H., Chatterjee, D., Caulfield, J.P., Morris, H.R. and Dell, A. (1997) *Glycobiology* 7, 663–677.
- [23] Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.C., Gomord, V. and Faye, L. (1998) *Plant Mol. Biol.* 38, 31–48.
- [24] Staudacher, E., Altmann, F., März, L., Hard, K., Kamerling, J.P. and Vliegthart, J.F. (1992) *Glycoconjug. J.* 9, 82–85.
- [25] van Die, I., Gomord, V., Kooyman, F.N., van den Berg, T.K., Cummings, R.D. and Vervelde, L. (1999) *FEBS Lett.* 463, 189–193.
- [26] Aalberse, R.C. and van Ree, R. (1997) *Allergy Immunol.* 15, 375–387.
- [27] Faye, L. and Chrispeels, M.J. (1988) *Glycoconjug. J.* 5, 245–256.
- [28] Fabini, G., Freilinger, A., Altmann, F. and Wilson, I.B. (2001) *J. Biol. Chem.*, (in press).
- [29] Leiter, H., Mucha, J., Staudacher, E., Grimm, R., Glössl, J. and Altmann, F. (1999) *J. Biol. Chem.* 274, 21830–21839.
- [30] Wilson, I.B., Rendic, D., Freilinger, A., Domic, J., Altmann, F., Mucha, J., Muller, S. and Hauser, M. (2001) *Biochim. Biophys. Acta* 1527, 88–96.
- [31] Eckhardt, M., Mühlenhoff, M., Bethe, A. and Gerardy-Schahn, R. (1996) *Proc. Natl. Acad. Sci. USA* 93, 7572–7576.
- [32] Bakker, H., Lommen, A., Jordi, W., Stiekema, W. and Bosch, D. (1999) *Biochem. Biophys. Res. Commun.* 261, 829–832.
- [33] Cummings, L., Warren, C.E., Granovsky, M. and Dennis, J.W. (1993) *Biochem. Biophys. Res. Commun.* 195, 814–822.
- [34] Bakker, H., Friedmann, I., Oka, S., Kawasaki, T., Nifant'ev, N., Schachner, M. and Mantei, N. (1997) *J. Biol. Chem.* 272, 29942–29946.
- [35] Faye, L., Gomord, V., Fitchette-Laine, A.C. and Chrispeels, M.J. (1993) *Anal. Biochem.* 209, 104–108.
- [36] Agterberg, M., van, Die, I., Yang, H., Andriessen, J.A., van Tetering, A., van den Eijnden, D.H. and Ploegh, H.L. (1993) *Eur. J. Biochem.* 217, 241–246.
- [37] de Vries, T., Palcic, M.P., Schoenmakers, P.S., van den Eijnden, D.H. and Joiasse, D.H. (1997) *Glycobiology* 7, 921–927.
- [38] Johnson, K.D. and Chrispeels, M.J. (1987) *Plant Physiol.* 84, 1301–1308.
- [39] Breg, J., Romijn, D., Vliegthart, J.F., Strecker, G. and Montreuil, J. (1988) *Carbohydr. Res.* 183, 19–34.
- [40] Platzter, N., Davoust, D., Lhermitte, M., Bauvy, C., Meyer, D.M. and Derappe, C. (1989) *Carbohydr. Res.* 191, 191–207.
- [41] Crawley, S.C., Hindsgaul, O., Ratcliffe, R.M., Lamontagne, L.R. and Palcic, M.M. (1989) *Carbohydr. Res.* 193, 249–256.
- [42] Dupuy, F., Petit, J.-M., Mollicone, R., Oriol, R., Julien, R. and Maftah, A. (1999) *J. Biol. Chem.* 274, 12257–12262.
- [43] Faik, A., Bar-Peled, M., DeRocher, A.E., Zeng, W., Perrin, R.M., Wilkerson, C., Raikhel, N.V. and Keegstra, K. (2000) *J. Biol. Chem.* 275, 15082–15089.
- [44] Perrin, R.M., DeRocher, A.E., Bar-Peled, M., Zeng, W., Norambuena, L., Orellana, A., Raikhel, N.V. and Keegstra, K. (1999) *Science* 284, 1976–1979.
- [45] 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. USA* 98, 2899–2904.