Plant Biotechnology Journal

Plant Biotechnology Journal (2012) **10**, pp. 851–861 doi: 10.1111/j.1467-7652.2012.00704.x

Moss-based production of asialo-erythropoietin devoid of Lewis A and other plant-typical carbohydrate determinants

Juliana Parsons[1], Friedrich Altmann[2], Claudia K. Arrenberg[1], Anna Koprivova[1†], Anna K. Beike[1], Christian Stemmer[3‡], Gilbert Gorr[3‡], Ralf Reski[1,4,5] and Eva L. Decker[1]*

[1]*Plant Biotechnology, Faculty of Biology, University of Freiburg, Freiburg, Germany*

[2]*Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria*

[3]*Greenovation Biotech, Freiburg, Germany*

[4]*BIOSS Centre for Biological Signalling Studies, Freiburg, Germany*

[5]*FRIAS Freiburg Institute for Advanced Studies, Freiburg, Germany*

Received 29 September 2011;
revised 28 March 2012;
accepted 5 April 2012.
**Correspondence* (fax 49 761 2036945; email eva.decker@biologie.uni-freiburg.de)
†Present address: John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK.
‡Present address: Emcid Biotech, Hans-Bunte-Str. 19, 79108 Freiburg, Germany.
Accession numbers: JN711463, XM_001778308, XM_001773431.

Keywords: α1,4 fucosyltransferase, β1,3 galactosyltransferase, moss bioreactor, *Physcomitrella patens*, plant-made pharmaceutical, EPO.

Summary

Protein therapeutics represent one of the most increasing areas in the pharmaceutical industry. Plants gain acceptance as attractive alternatives for high-quality and economical protein production. However, as the majority of biopharmaceuticals are glycoproteins, plant-specific N-glycosylation has to be taken into consideration. In *Physcomitrella patens* (moss), glyco-engineering is an applicable tool, and the removal of immunogenic core xylose and fucose residues was realized before. Here, we present the identification of the enzymes that are responsible for terminal glycosylation (α1,4 fucosylation and β1,3 galactosylation) on complex-type N-glycans in moss. The terminal trisaccharide consisting of α1,4 fucose and β1,3 galactose linked to N-acetylglucosamine forms the so-called Lewis A epitope. This epitope is rare on moss wild-type proteins, but was shown to be enriched on complex-type N-glycans of moss-produced recombinant human erythropoietin, while unknown from the native human protein. Via gene targeting of moss galactosyltransferase and fucosyltransferase genes, we identified the gene responsible for terminal glycosylation and were able to completely abolish the formation of Lewis A residues on the recombinant biopharmaceutical.

Introduction

Recombinant protein therapeutics constitute the largest class of products currently developed by the pharmaceutical industry. As about 70% of the so-called biopharmaceuticals on the market or in clinical studies are glycoproteins (Durocher and Butler, 2009), the development of alternative expression systems with human-like glycosylation patterns is essential. Protein glycosylation in mammalian expression systems is most similar to human glycans. Likewise, plants, as higher eukaryotes, are able to produce complex-type N-glycans (Lerouge *et al.*, 1998; Wilson, 2002). The core structure of the complex N-glycans, a pentasaccharide of three mannoses and two N-acetylglucosamines (Man3GlcNAc2), is conserved between mammals and plants, as well as up to two terminally attached GlcNAc residues (reviewed in Gomord *et al.*, 2010). Differing from the human structure, most plant N-glycans carry a β1,2 xylose and an α1,3 fucose linked to the glycan core. These residues raised concerns about plant-produced biopharmaceuticals as they were shown to be highly immunogenic in mammals (Bardor *et al.*, 2003; Bencúrová *et al.*, 2004; Garcia-Casado *et al.*, 1996; Jin *et al.*, 2008; Van Ree *et al.*, 2000). The creation of transgenic plants, in which expression of the respective glycosyltransferase genes was either abolished (Koprivova *et al.*, 2004; Strasser *et al.*, 2004) or reduced (Cox *et al.*, 2006; Sourrouille *et al.*, 2008; Strasser *et al.*, 2008), increased the acceptance of plant-made biopharmaceuticals, and several products are currently studied in clinical trials (Davies, 2010; Paul and Ma, 2011). Another difference between human and plant complex-type N-glycans lies in their terminal elongation. While on many human N-glycans, β1,4 galactose (h1,4 Gal) is linked to the outer GlcNAc residues followed by terminal sialic acid, these two sugars are non-existing on plant glycoproteins. In contrast, β1,3 galactose (Gal) and α1,4 fucose (Fuc) may be attached to the terminal GlcNAc residues giving rise to the Fucα1-4(Galβ1-3)GlcNAc structure. This trisaccharide, which was originally described for Lewis-positive histo-blood groups in humans (Henry *et al.*, 1995), is known as Lewis A epitope (Le^a). Although Le^a determinants may be regarded as human-type glycosylation, this structure occurs only rarely on glycoproteins of healthy adult humans. Moreover, Le^a and Sialyl-Le^a were shown to be highly increased in the sera from patients with colon cancer (Zhang *et al.*, 1994). Therefore, the repression of Le^a attachment to recombinant plant-made biopharmaceuticals is desirable.

Le^a is widespread on glycoproteins of several plant species and was estimated to represent about 15% of the total N-glycan population in tobacco (Fitchette *et al.*, 1999; Wilson, 2001). In addition, the Le^a epitope was also described on recombinant proteins produced in tobacco (Petruccelli *et al.*, 2006). In contrast, the appearance of Le^a structures on

glycoproteins in the moss *Physcomitrella patens* or the model plant *Arabidopsis thaliana* was shown to be low (Koprivova *et al.*, 2003; Strasser *et al.*, 2007; Viëtor *et al.*, 2003). In *P. patens*, several biopharmaceuticals were successfully expressed (Baur *et al.*, 2005; Büttner-Mainik *et al.*, 2011; Gorr and Jost, 2005; Schuster *et al.*, 2007; Weise *et al.*, 2007); however, in one case, the human erythropoietin (hEPO), a significant proportion of Le^a structures attached to the N-glycan core was observed (Weise *et al.*, 2007).

Erythropoietin is a hormone, predominantly synthesized by the kidneys, which plays an important role in the maturation of red blood cells (erythropoiesis). It is commonly used to treat and prevent anaemia in nephrology and oncology patients (Jelkmann, 2000). Recombinant hEPO (rhEPO), mainly produced in chinese hamster ovary (CHO) cells, has a worldwide market of about 10 billion euros per year, making it the leading biopharmaceutical in the market (Jelkmann, 2007). Human EPO consists of 165 amino acids and has a molecular mass of approximately 34 kDa. Three N-linked (at Asn 24, Asn 38 and Asn 83) and one O-linked (at Ser 126) carbohydrate chains are attached to the amino acid chain, which in total represent 40% of the molecular mass (Takeuchi *et al.*, 1989). Yet, these oligosaccharides are not essential for the activity of EPO *in vitro*, indicating that solely the polypeptide is involved in binding to the EPO receptor (Dordal *et al.*, 1985). On the other hand, desialylated EPO (asialo-EPO) looses its erythropoietic activity *in vivo*, while it maintains, or even increases, its receptor affinity *in vitro* (Tsuda *et al.*, 1990). This indicates a major role of glycosylation in increasing the half-life of EPO in circulation (Takeuchi and Kobata, 1991).

More recently, EPO was shown to be produced in many tissues under hypoxia or metabolic stress, where it inhibits apoptosis and protects these tissues from collateral damage elicited by the injury (reviewed in Brines and Cerami, 2008). In this context, EPO was found to be neuroprotective, representing a potential treatment for ischaemic stroke, diabetes-induced eye damage and peripheral nerve injury (Brines *et al.*, 2008). For these indications, an increase in red blood cells is not desirable (Masuda *et al.*, 1993). However, asialo-EPO can provide tissue protection without the undesired erythropoietic side effect, making asialo-EPO an advantageous pharmaceutical (Carelli *et al.*, 2011; Erbayraktar *et al.*, 2003).

Recombinant hEPO was already successfully produced in moss (Weise *et al.*, 2007), and its activity was confirmed *in vitro*. Considering that plants naturally lack the required enzymes for glycoprotein sialylation, moss can be assumed as a suitable platform for the production of asialo-EPO. However, to reliably provide a safe pharmaceutical, Le^a residues detected on moss-produced rhEPO have to be abolished.

Here, we identified the genes encoding α1,4 fucosyl- and β1,3 galactosyltransferases responsible for terminal plant N-glycan elongation in *P. patens*. In this species, gene functions can be eliminated by precise gene targeting via homologous recombination with a unique efficiency (Kamisugi *et al.*, 2006; Khraiwesh *et al.*, 2010; Mosquna *et al.*, 2009; Strepp *et al.*, 1998). Using this method to target the corresponding glycosyltransferase genes in transgenic rhEPO-producing moss lines, we were able to completely abolish Le^a formation on recombinant as well as endogenous moss glycoproteins, thus providing the next generation of plant N-glycan humanization. Furthermore, we successfully demonstrated β1,4 galactosylation of rhEPO after introduction of human β1,4 galactosyltransferase to the Le^a-free moss lines.

Results

Identification of α1,4 fucosyltransferase and β1,3 galactosyltransferase genes from *Physcomitrella patens*

Via BLAST (Basic Local Alignment Search Tool) searches in the Cosmoss database (http://www.cosmoss.org) using human and *Arabidopsis* α1,4 fucosyltransferases (FucT) and β1,3 galactosyltransferases (GalT), we were able to identify respective homologous sequences from the *P. patens* genome. The full-length cDNA of Pp*fuct* was obtained by 5′ and 3′ RACE (rapid amplification of cDNA ends) PCR. In *Physcomitrella*, α1,4 fucosyltransferase is a single-copy gene (Pp1s151_12V6.1), encoding a protein of 437 amino acids. The PpFucT protein showed 50.1% identity to FucT from *A. thaliana* (FUT13, AAK11728) and 19.6% identity to the human FucT (FUT3; NP_001091110), an enzyme with α1,3 fucosyltransferase and α1,4 fucosyltransferase activities, which is responsible for the addition of fucose in the last step of Le^a antigen biosynthesis.

Eleven amino acid sequences of putative galactosyltransferases from *P. patens* resulted from the BLASTP search against the gene models derived from the moss genome sequence: Pp1s251_37V6.1, Pp1s237_65V6.1, Pp1s163_84V6.1, Pp1s65_62V2.1, Pp1s34_213V2.1, Pp1s101_246V0.1, Pp1s152_166V0.1, Pp1s313_19V0.1, Pp1s34_146V0.1, Pp1s14_277V0.1, Pp1s263_36V0.1. All 11 amino acid sequences show sequence identities of 19%–31% and sequence similarities of 39%–49% to all six human sequences, within the galactosyltransferase domain (Pfam 01762). Subsequently, a protein family search in the Pfam 24.0 database (http://pfam.sanger.ac.uk/) was performed to identify further galactosyltransferase genes (Pfam 01762) in *P. patens*. Two additional putative galactosyltransferases were found: Pp1s121_83V0.1 and Pp1s111_119V2.1. Three of the GalT homologues, Pp1s251_37V6.1, Pp1s237_65-V6.1 and Pp1s163_84V6.1, were predicted to contain a signal peptide for the secretory pathway and a transmembrane domain, as well as a galactoside-binding lectin domain (Pfam 0037). These features were considered as potential prerequisites for the physiologic function of the GalT on glycoproteins in the Golgi apparatus. A phylogenetic analysis (Figure 1), performed with the putative *Physcomitrella* GalT sequences as well as six lectin-binding domain-containing *Arabidopsis* GalT and six human GalT homologues, shows distinct clades for the putative galactosyltransferases from the moss. One PpGalT sequence (Pp1s251_37V6.1) forms a distinct clade with two sequences from *A. thaliana* (At1g26810 and At3g06440). As Pp1s251_37V6.1 clusters together with GalT1 from *A. thaliana* (Strasser *et al.*, 2007), it was named GalT1 for *Physcomitrella* (PpGalT1). Two additional PpGalT sequences (Pp1s237_65V6.1 and Pp1s163_84V6.1) cluster with four AtGalT sequences (At1g74800, At4g21060, At1g27120 and At5g62620) and were named PpGalT2 and PpGalT3, respectively.

Generation of *fuct* and *galt* single- and double-knockout plants

To reveal their specific function on Le^a formation in moss glycoproteins, we created plasmid constructs for the targeted knockout (KO) of the single *fuct* gene as well as for the three *galt* homologues 1, 2 and 3 (Figure S1a). These constructs were transferred into *P. patens* protoplasts to generate single deletions for *fuct*, *galt1* and *galt3* via homologous recombination, respectively. Owing to the high similarity of GalT2 and GalT3

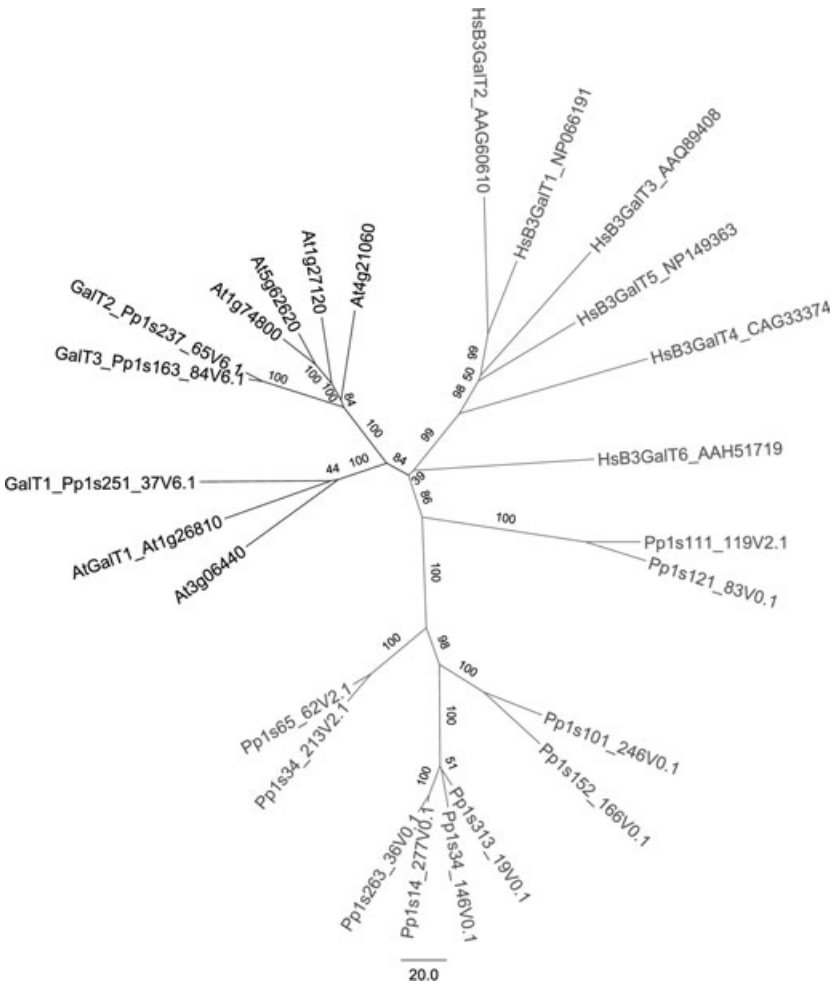

Figure 1 Phylogeny of putative galactosyltransferases from *Physcomitrella patens*. The phylogenetic tree was constructed from 13 *P. patens* amino acid sequences sharing the galactosyltransferase domain (Pfam 01762), six *Arabidopsis thaliana* sequences (At1g26810, At3g06440, At1g74800, At5g62620, At1g27120 and At4g21060) and six *Homo sapiens* sequences (HsB3GalT1, HsB3GalT2, HsB3GalT3, HsB3GalT4, HsB3GalT5 and HsB3GalT6) with the neighbour-joining algorithm implemented in Quicktree (Howe *et al.*, 2002). The tree was visualised with FigTree V 1.1.2. The clades of GalT1_Pp1s251_37V6.1 and its homologues in *A. thaliana*, and GalT2_Pp1s237_65V6.1, GalT3_Pp1s163_84V6.1 and their homologues in *A. thaliana* are shown in black. Bootstrap values were determined from 1000 replicates and are shown at the respective branches in percentage. Bar = 20.0% sequence divergence.

proteins, which share 71.5% identity, additionally Δ*galt2*/Δ*galt3* double KOs were created. For this purpose, a Δ*galt3* moss line was transformed with the *galt2* KO construct. Because of the observation that the moss-produced rhEPO displays a substantial amount of Le^a, the rhEPO-producing line 174.16 (Weise *et al.*, 2007) was used as parent plant for the *galt* KOs. Plants that survived two cycles of antibiotic selections were screened for genomic integration of the KO construct in the correct genomic locus. Plants with positive PCR for both 5′ and 3′ integration were further analysed by RT-PCR to prove the disappearance of the respective transcript (Figure S1b–e). By these means, seven lines were confirmed as targeted KOs for the *fuct* gene as well as nine Δ*galt1* lines. In addition, one Δ*galt3* line and two Δ*galt2*/Δ*galt3* double-KO lines were confirmed. One or two lines, respectively, of each genetic modification were chosen for further analysis and stored in the International Moss Stock Center (http://www.moss-stock-center.org; see Table S4).

Terminal N-glycosylation is affected in *fuct* and *galt* KO lines

To explore whether the disruption of the *galt* and *fuct* genes indeed results in changes of the glycosylation pattern, N-glycans of the cellular protein fraction from *Physcomitrella* wild type (WT), the rhEPO-synthesizing parent plant 174.16 as well as the different glycosyltransferase KO lines were analysed by MALDI-TOF mass spectrometry.

The major peak in the glycan spectrum of WT moss represented the complex-type N-glycan structure GnGnXF with terminally attached GlcNAc residues (Gn) and core xylose (X) and fucose (F), respectively (Figure 2a top). In addition, Le^a structures (FA) on one or both of the antennae were detected [$Gn(FA)XF_{iso}$ and (FA)(FA)XF, respectively]. The Δ*fuct* plant displays structures lacking the terminal fucose residues; however, β1,3-linked galactose residues (A) are still present (lower Figure 2a, $GnAXF_{iso}$ and AAXF). The rhEPO-expressing parent plant (174.16), in which core xylose and fucose residues had been abolished by targeted KO (Weise *et al.*, 2007), presents the GnGn structure and additional Le^a-containing glycans, $Gn(FA)_{iso}$ or (FA)(FA) (upper Figure 2b). The Δ*galt2*/Δ*galt3* double-KO plants did not show any alteration in the terminal antennae glycosylation in comparison with the parent plant (middle Figure 2b). The same was true for the Δ*galt3* single KO (data not shown), suggesting that these GalT homologues are not involved in the synthesis of Le^a epitopes in the gametophytic tissue. In contrast, in the Δ*galt1* plants, the major peak that appears in the spectrum corresponds to GnGn glycans, and no peak was observed around the Le^a-indicative masses of 1648 or 1956 Da, proving that Δ*galt1* is completely devoid of terminal β1,3 galactosyl residues. Moreover, any terminal fucose is also missing. This finding supports the observation that the α1,4 fucosyltransferase transfers a fucose residue to the preformed Galβ1,3GlcNAc substrate (Léonard *et al.*,

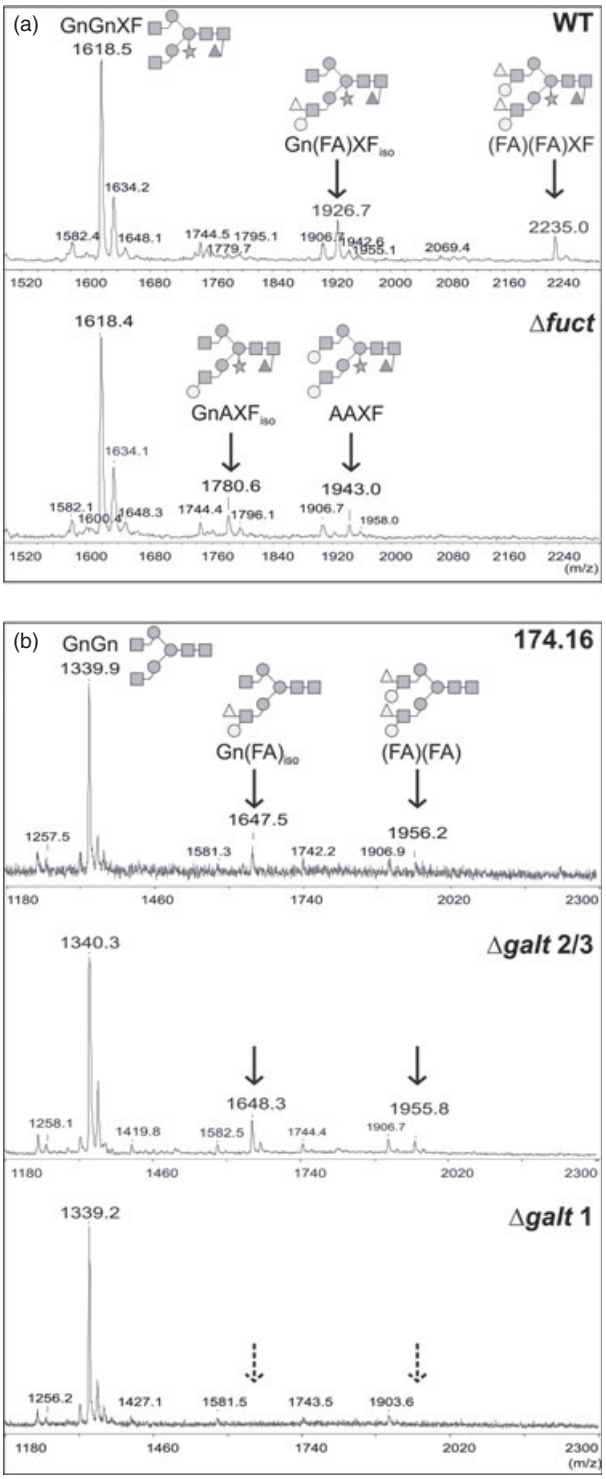

Figure 2 MALDI-TOF mass spectra of the endogenous N-linked glycans from moss. (a) Comparison of the MALDI-TOF mass spectra for the wild-type moss and the Δ*fuct* plant. (b) Comparison of the MALDI-TOF mass spectra for the recombinant human erythropoietin-producing plant 174.16, the double-knockout Δ*galt2*/Δ*galt3* No. 23 and the Δ*galt1* plant No. 58. For the indicated masses, please refer to Table 1. ○: galactose, ☆: xylose, □: N-acetylglucosamine, ●: mannose, ▲/△: fucose.

2002; Wilson, 2001), that is, the FucT acts as the last enzyme in the biosynthesis of plant glycoproteins, and preceding GalT activity is essential to provide the appropriate substrate for the FucT.

Moss-derived Lewis A-free recombinant hEPO

Via glycan analysis of the different moss strains, we proved the disappearance of Le^a residues on total intracellular N-glycans of moss plants lacking GalT1. Most important for biopharmaceutical production, however, is to confirm that the modified plants also produce Le^a-free rhEPO. Therefore, extracellular proteins from the parent plant 174.16, Δ*galt2*/Δ*galt3* and Δ*galt1* lines were precipitated from the culture supernatant, separated via SDS–PAGE and analysed by Western blotting.

Using anti-EPO antibodies, rhEPO could be detected in the medium of all analysed plants (Figure 3a). Different from CHO-derived rhEPO, which was chosen as a standard and migrated with an apparent molecular weight (MW) of about 34 kDa, moss-produced rhEPO migrated with a molecular weight (MW) of approximately 25 kDa. This size difference is because of different glycosylation patterns. CHO-produced rhEPO displays mainly tetra-antennary sialylated N-glycans, while moss-produced rhEPO displays exclusively di-antennary N-glycans as was also observed on human serum EPO (Skibeli *et al.*, 2001). The estimated different amounts of mainly mannose, galactose and sialic acid residues result in a calculated size difference of more than 7 kDa.

Using the anti-Le^a antibody JIM 84 (Horsley *et al.*, 1993), we detected terminal Le^a structures, accumulating at the size of rhEPO, on total secreted proteins of the parent plant (174.16) as well as the Δ*galt2*/Δ*galt3* line. In contrast, we observed the complete absence of any Le^a signal in the Δ*galt1* plants (Figure 3b).

To investigate the structure of the glycans attached to our protein of interest, the moss-produced rhEPO was further analysed via mass spectrometry. For this purpose, total soluble proteins were precipitated from the culture supernatant of the parent plant and Δ*galt1* lines and separated by SDS–PAGE. Subsequently, the main EPO-containing band was cut out from the Coomassie-stained gel (Figure S2) and subjected to mass spectrometry for an analysis of the EPO glycopeptides.

While the overall proportion of Le^a residues on moss glycoproteins is very small (Koprivova *et al.*, 2003), it was obviously enriched on the moss-produced rhEPO. Around three-quarters of the rhEPO molecules displayed complex glycans containing Le^a residues (Figure 4). With 3 N-glycosylation sites bound by complex-type bi-antennary glycans, there are several possible combinations where at least one Le^a motif is present. The possible glycosylation pattern for the two tryptic glycopeptides is listed in Table 2. By knocking out the *galt1* gene, the Le^a motif disappeared completely from the rhEPO glycosylation pattern, while the core glycan GnGn remains intact (Figure 4). As already observed for intracellular total moss glycoproteins, the main terminal sugar on moss N-glycans is GlcNAc, while terminal mannoses (M) are far less prominent (Figure 2 and Table 2). This was even more striking in the specific case of secreted rhEPO glycosylation which exhibits a tiny residual GnM peak in addition to the prominent GnGn structure. Therefore, we achieved an outstandingly homogenous product quality, having our protein of interest with almost exclusively the desired glycan pattern.

Phenotype and growth rate of the mutants were comparable to the parent lines, indicating that the overall metabolism is not affected by the mutations (data not shown). The rhEPO produced in different moss lines was measured by means of

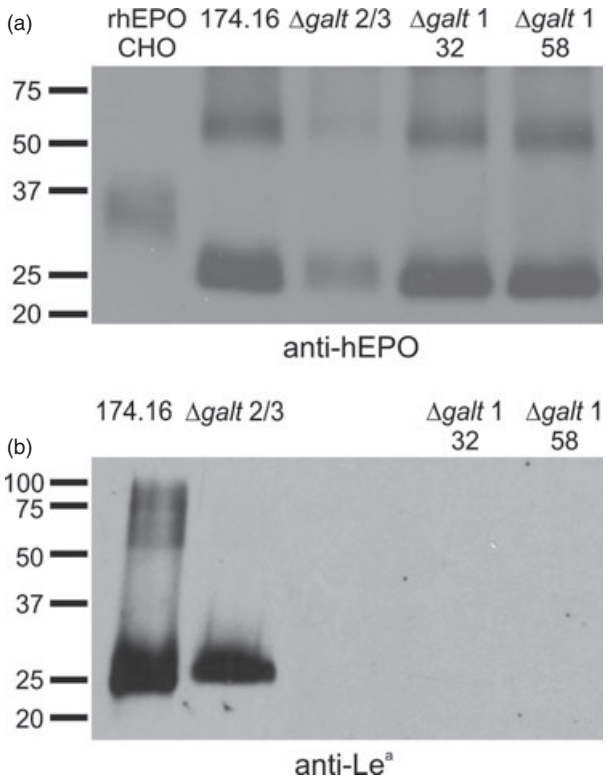

Figure 3 Western blot analysis proved disappearance of Le^a structures from moss-derived recombinant human erythropoietin (rhEPO). Total soluble proteins (TSP) were precipitated from culture supernatant of plants 174.16, Δ*galt2*/Δ*galt3* No. 23, Δ*galt1* No. 32 and 58. (a) Twenty-five nanograms of extracellular TSP and 5 ng of chinese hamster ovary-produced rhEPO standard were loaded, and detection was performed with anti-hEPO antibodies. (b) Thirty micrograms of TSP were loaded, and detection was performed with the anti-Lewis A (Le^a) antibody JIM 84.

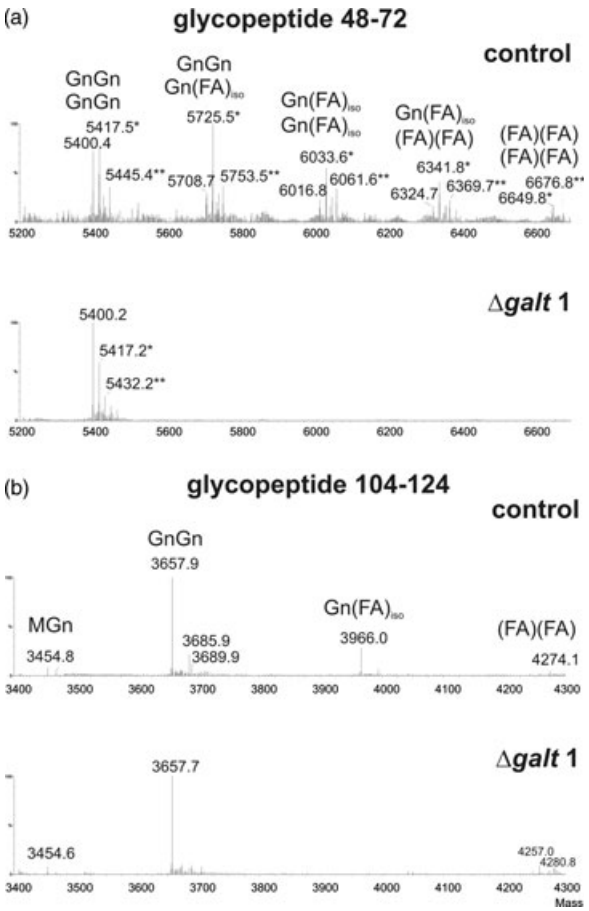

Figure 4 Glycopeptides of moss-produced recombinant human erythropoietin (rhEPO) from plants Δ*galt1* No. 32 and control plant 174.16. Tryptic digestion of rhEPO generated two glycopeptides. The peptide harbouring two glycosylation sites (EAENITTGCAEHCSLNENITVPDTK) is shown in panel a. The tryptic peptide containing one glycan (GQALLVNSSQPWEPLQLHVDK) is depicted in panel b. Glycopeptides containing Le^a structures are totally absent in the Δ*galt1* line. For the indicated masses, please refer to Table 2. Salt adducts are marked with asterisks.

ELISA. The rhEPO production levels observed in the Δ*galt1* plants were comparable to that of the parent plant 174.16 (Table S3), indicating that the genetic modifications did not affect the productivity of the rhEPO or stability of the product, respectively.

Although *in vitro* activity of moss-derived rhEPO has already been proven (Weise *et al.*, 2007), the addition of terminal human-type β1,4 galactose to the rhEPO molecule might be useful for tissue-protective EPO activity. Therefore, we created rhEPO-producing Δ*galt1* lines that additionally expressed the human β1,4 galactosyltransferase gene (*h1,4galt*). Some of the resulting plants that showed *h1,4galt* transcription (Figure S3) were further analysed for their ability to produce β1,4 galactosylated rhEPO. Total soluble proteins from the culture supernatants of *h1,4galt* insertion lines (*h1,4galt* i) were separated via SDS–PAGE and analysed by lectin affinity blotting with *Datura stramonium* agglutinin (DSA) which specifically detects Galβ1,4GlcNAc residues (Figure 5a). As a control, an anti-EPO Western blot was performed (Figure 5b). While no signal was observed for Δ*galt1* plants, we clearly detected terminal β1,4 galactose structures accumulating at the size of rhEPO in two *h1,4galt* i lines, proving the activity of the introduced human β1,4GalT. Comparison with the EPO bands detected by anti-EPO Western blotting suggests that a substantial proportion of the moss-derived rhEPO contains human-type β1,4 galactose residues.

Discussion

N-glycosylation is one of the most important post-translational modifications of proteins, which may have impact on functionality, stability or compatibility of a recombinant pharmaceutical. It starts in the endoplasmic reticulum with the transfer of the oligosaccharide precursor $Glc_3Man_9GlcNAc_2$ from a dolichol carrier to specific Asn-X-Ser (or Thr) tripeptides on the polypeptide chain. Subsequently, processing takes place in the Golgi apparatus, where complex-type N-glycans are finally synthesized. Plant complex-type N-glycans are characterized by the attachment of plant-specific β1,2 xylose and α1,3 fucose to the glycan core. Besides these modifications, plant complex-type N-glycans may be further processed by addition of terminal fucose and galactose residues yielding Fucα1-4(Galβ1-3)GlcNAc sequences (Fitchette-Lainé *et al.*, 1997; Melo *et al.*, 1997). These residues are synthesized by β1,3 galactosyltransferase (GalT) and α1,4 fucosyltransferase (FucT) and named Lewis A antigens (Le^a). Le^a-containing N-glycans are present in a wide spectrum of plants including mosses, ferns, and seed plants (Fitchette *et al.*, 1999;

Table 1 Structures and corresponding calculated masses of the major N-glycans, for the wild-type moss (WT), single-KO Δ*1,4fuct*, double-KO Δ*1,3fuct/*Δ*xylt* (glycan background from the plant 174.16) and Δα*1,3fuct/*Δ*xylt/*Δ*galt* triple-KO plants

WT		Δα*1,4fuct*		Δα*1,3fuct* + Δ*xylt*		Δα*1,3fuct* + Δ*xylt* + Δ*galt*	
Mass	Structure	Mass	Structure	Mass	Structure	Mass	Structure
1582.4	M7	1582.4	M7	1258.2	M5	1258.2	M5
1618.3	GnGnXF	1618.3	GnGnXF	1340.2	GnGn	1340.2	GnGn
1744.5	M8	1744.5	M8	1420.3	M6	1420.3	M6
1906.7	M9	1906.7	M9	1582.4	M7	1582.4	M7
1926.8	$Gn(FA)XF_{iso}$	1780.6	$Gn(A)XF_{iso}$	1648.5	$Gn(FA)_{iso}$	1486.5	$Gn(F)_{iso}$*
2235.0	(FA)(FA)XF	1942.8	(A)(A)XF	1744.5	M8	1744.5	M8
				1906.7	M9	1906.7	M9
				1955.7	(FA)(FA)	1631.7	(F)(F)*

*Masses and structures for an hypothetical case, where only terminal galactose disappears but terminal fucose remains in the Δ*galt* plants, in addition to the absence of xylose and α1,3 fucose.
KO, knockout.

Table 2 Average masses of ions that could be assigned to tryptically digested glycopeptides of recombinant human erythropoietin

Glycopeptide 48–72				Glycopeptide 104–124	
EAE**NIT**TGCAEHCSLNE**NIT**VPDTK (glycosylation sites A and B)				GQALLV**NSS**QPWEPLQLHVDK (glycosylation site C)	
A	B	M+H	$M+NH_4^+$	C	M+H
GnGn	GnM	5197.12	5214	GnM	3454.64
GnGn	GnGn	5400.20	5417	GnGn	3657.72
GnGn	(FA)Gn	5708.31	5725	AGn	3819.77
(FA)Gn	(FA)Gn	6016.42	6033	(FA)Gn	3965.83
(FA)Gn	(FA)(FA)	6324.53	6341	AA	3981.82
(FA)(FA)	(FA)(FA)	6632.65	6649	(FA)(FA)	4273.94

One of the three N-glycosylation sites lays on peptide 104–124, while the other two lay on the same peptide (48–72), giving raise to different combinations of glycopeptides.
Glycosylation sites underlined.

Wilson *et al.*, 2001). They were estimated to represent about 15% of the total N-glycan population in tobacco (Fitchette *et al.*, 1999). As demonstrated by immunolabelling experiments, the Le^a epitope is associated with extracellular and plasma membrane-bound glycoproteins (Fitchette *et al.*, 1999), those compartments in which the majority of recombinant pharmaceutical glycoproteins are located. As Le^a structures are rare on human glycoproteins and increased in connection with certain types of cancer (Zhang *et al.*, 1994), they should be avoided on recombinant therapeuticals.

All of the biopharmaceuticals that have already been successfully expressed in the moss *P. patens*, human vascular endothelial growth factor, complement regulatory factor H, IgG immunoglobulins and hEPO (Baur *et al.*, 2005; Büttner-Mainik *et al.*, 2011; Gorr and Jost, 2005; Schuster *et al.*, 2007; Weise *et al.*, 2007) are extracellular glycoproteins. One of them, hEPO, showed a high proportion of N-glycans bearing the Le^a epitope (Weise *et al.*, 2007).

The features that make moss an attractive system for biopharmaceutical production comprise its growth as tissue cultures in strictly contained and scalable photobioreactors, the possibility of product secretion to the culture supernatant and high flexibility for medium pH and temperature (Decker and Reski, 2007, 2012). Its unique amenability for targeted gene replacements by homologous recombination has been employed in adapting the moss glycosylation pattern to the requirements for biopharmaceutical production by knocking out the core β1,2 xylosyltransferase and α1,3 fucosyltransferase genes (Koprivova *et al.*, 2004).

Here, we aimed to identify the enzymes responsible for the biosynthesis of terminal Le^a structures on *P. patens* glycoproteins and to eliminate their function in rhEPO-producing moss lines in order to provide a Le^a-free asialo-EPO as a new interesting pharmaceutical target protein.

While α1,4FucT is encoded by a single-copy gene in *P. patens*, 13 putative β1,3GalT homologues were identified from the moss genome sequence. On the basis of domain features and sequence similarity to *A. thaliana* GalTs, we took three of the *P. patens galt* genes under closer consideration. By generation of single Δ*fuct*, Δ*galt1* and Δ*galt3* as well as Δ*galt2/*Δ*galt3* double-KO lines, we intended to identify those enzymes relevant for Le^a formation in *Physcomitrella*.

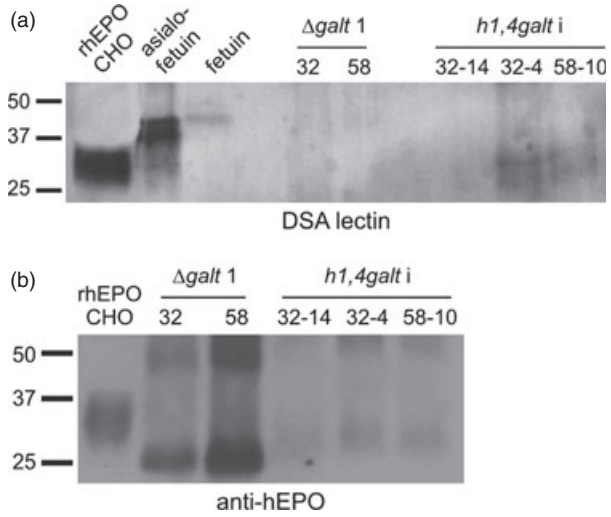

Figure 5 Lectin affinity blot analysis demonstrated specific terminal β1,4 galactosylation of recombinant human erythropoietin (rhEPO) from *h1,4galt* insertion plants. (a) TSP were precipitated from culture supernatant of plants Δ*galt1* No. 32 and 58 and insertion lines *h1,4galt* i 32-14, 32-4 and 58-10. As positive controls, 200 ng of chinese hamster ovary (CHO)-produced rhEPO, 500 ng of β1,4 galactosylated asialo-fetuin and 200 ng of β1,4 galactosylated fetuin were loaded, and detection was performed with lectin *Datura stramonium* agglutinin. (b) As a control, CHO-produced rhEPO and culture supernatant of plants △*galt1* No. 32 and 58 and insertion lines *h1,4galt* i 32-14, 32-4 and 58-10 were analysed by anti-EPO Western blotting.

MALDI-TOF mass spectra analysis confirmed changes in the N-glycosylation pattern in the mutants generated in this work. In comparison to WT with Le^a epitopes on the N-glycans $Gn(FA)XF_{iso}$ and (FA)(FA)XF (Figure 2a), Δ*fuct* plants exhibited the corresponding structures with the terminal galactose residues only, $GnAXF_{iso}$ and AAXF, respectively. These structures have been only rarely observed on plant glycoproteins (Wilson, 2001). Steinkellner *et al.* (Strasser *et al.*, 2007) suggested two possible explanations: either the FucT enzyme modifies quickly and quantitatively the glycans with terminal galactose, or the Galβ1-3GlcNAc oligosaccharide is rapidly degraded by the action of β-galactosidases. In the Δ*fuc*t plants analysed here, we could clearly show an increase in terminal galactose compared to the WT (Figure 2a). This indicates that FucT modifies the Galβ1-3GlcNAc substrate very efficiently. However, as the $GnAXF_{iso}$/AAXF proportion in Δ*fuct* is not identical to the Gn(FA)XFiso/(FA)(FA)XF fractions of WT, we cannot exclude additional cleavage of terminal Gal residues.

While the disruption of *galt2* and *galt3* was obviously without any influence on terminal galactosylation, MALDI-TOF mass spectra from Δ*galt1* plant material revealed the absence of the complete Le^a epitope, confirming that the FucT enzyme needs the presence of the galactose to transfer a fucose to the GlcNAc residue.

In addition, immunoblotting performed using an antibody binding specifically to Le^a, JIM 84, showed the absence of this antigen from total secreted proteins of the Δ*galt1* plants. In contrast, Le^a epitopes were considerably high in the rhEPO-producing parent plant and the Δ*galt2*/Δ*galt3* double-KO plants at the mass corresponding to moss-produced rhEPO.

A closer look at the glycopeptides of secreted rhEPO from the Δ*galt1* plants confirmed the lack of any Le^a epitopes. Moreover, the high homogeneity of the product was striking, with almost exclusively one type of glycosylation, the desired GnGn structure. Here, a pharmaceutical product is presented with an excellent degree of uniformity.

In vitro activity of moss-derived rhEPO has already been proven (Weise *et al.*, 2007). However, moss-produced rhEPO lacked the terminal β1,4 galactose present on asialo-EPO which was tested for neuro-protection (Erbayraktar *et al.*, 2003). Therefore, we engineered the moss genome for human β1,4 galactosyltransferase activity (Figure 5) in a proof-of-concept approach and demonstrated substantial β1,4 galactosylation of the secreted rhEPO in the resulting insertion lines. However, the degree of human-type galactosylation might be further increased using a hybrid instead of the native human β1,4GalT (Strasser *et al.*, 2009).

The glycan engineering performed in the production lines presented in this work did not affect either biomass accumulation or rhEPO productivity; therefore, we suggest the moss bioreactor as the system of choice for the production of asialo-EPO as a new exciting biopharmaceutical target.

Experimental procedures

Identification of α1,4 fucosyltransferase and β1,3 galactosyltransferase in *Physcomitrella patens*

For the identification of α1,4 fucosyltransferase homologues in *P. patens*, the amino acid sequence of the human galactoside 3(4)-L-fucosyltransferase (FUT3, NP_001091110.1) and the *A. thaliana* homologue At1g71990.1 were used to perform a BLASTP search in the *P. patens* resource (http://www.cosmoss.org). The sequence of the complete cDNA of the Pp*fuct* was obtained via RACE (GeneRacer™; Invitrogen, Karlsruhe, Germany) according to the manufacturer's protocol.

For the identification of Ppβ1,3 galactosyltransferase homologues, amino acid sequences of six human β1,3 galactosyltransferases HsB3GalT1 (NP066191), HsB3GalT2 (AAG60610), HsB3GalT3 (AAQ89408), HsB3GalT4 (CAG33374), HsB3GalT5 (NP149363) and HsB3GalT6 (AAH51719) were used to perform a BLASTP search in http://www.cosmoss.org. Galactosyltransferase amino acid sequences from *A. thaliana* (At1g26810, At3g06440, At1g74800, At5g62620, At1g27120 and At4g21060) were taken from The Arabidopsis Information Resource database (TAIR, http://www.arabidopsis.org).

Plant material and transformation procedure

Physcomitrella patens (Hedw.) B.S.G. was cultivated as described previously (Frank *et al.*, 2005). The rhEPO-producing moss line 174.16 (Weise *et al.*, 2007) was used as starting material for the *galt* KOs. In this moss line, the α1,3 fucosyltransferase and the β1,2 xylosyltransferase genes are disrupted (Koprivova *et al.*, 2004). Protoplast isolation and polyethylene glycol-mediated transfection were performed as described previously (Frank *et al.*, 2005). For mutant selection, Knop media were supplemented with 50 μg/mL G418, 50 μg/mL Zeocin™ (Invitrogen) or 100 μg/mL sulphadiazine (Sigma). The selection pressure started 14 days after transformation for G418 resistance and 9 days after transformation in the case of Zeocin™ and sulphadiazine selection, respectively.

For rhEPO production, *P. patens* was cultivated in 500 mL flasks with 200 mL basal medium prepared as follows: a solution containing 250 mg/L KCl, 250 mg/L $MgSO_4 \cdot 7H_2O$, 1 g/L

$Ca(NO_3)_2 \cdot 4H_2O$ and 1 g/L 2-(N-morpholino)-ethanesulphonic acid (MES) was adjusted to pH 5.9 with 1 M KOH and autoclaved. Subsequently the following solutions were added per litre of this solution: 10 mL 25 g/L KH_2PO_4 (autoclaved separately), 1 mL 50 mM Fe-EDTA (sterile filtrated) and 1 mL microelement solution (614 mg/L H_3BO_3, 25 mg/L $CoCl_2 \cdot 6H_2O$, 25 mg/L $CuSO_4 \cdot 5H_2O$, 830 mg/L KI, 390 mg/L $MnCl_2 \cdot 4H_2O$, 250 mg/L $Na_2MoO_4 \cdot 2H_2O$, 25 mg/L $NiCl_2 \cdot 6H_2O$, 10 mg/L $Na_2SeO_3 \cdot 5H_2O$, 42 mg/L Zn acetate$\cdot 2H_2O$, pH 7.0) (sterile filtrated). The cultures were subcultivated weekly as described earlier (Reski and Abel, 1985) for at least 3 weeks. Afterwards, the cultures were kept 16 days without subculturing before the supernatant was harvested.

Creation of α1,4 fucosyltransferase and β1,3 galactosyltransferase KO constructs and human β1,4 galactosyltransferase knockin constructs

To generate the *galt1* and *galt2* KO constructs, *P. patens* genomic DNA fragments corresponding to the respective genes were amplified using the primers listed in Table S1 and cloned into pCR®4-TOPO® (Invitrogen). The *galt1* genomic fragment was first linearized using *Nhe*I and *Nde*I, thus deleting a 1481-bp fragment, and recircularized by ligating a double-stranded oligonucleotide containing restriction sites for *Bam*HI and *Hin*dIII. These sites were used for the insertion of the Zeocin™ selection cassette, obtained from the plasmid pUC-zeo. For creation of pUC-zeo, the *ble* gene was amplified from the vector pMS171 (Schroda *et al.*, 2002) with the primers listed in Table S1 and cloned into the vector pRT101-neo (Girke *et al.*, 1998) using *Xho*I and *Xba*I restriction sites to replace the *nptII* gene, resulting in pRT101-zeo. Subsequently, the Zeocin™ selection cassette, including the 35S promoter and the 35S terminator, was removed from pRT101-zeo and ligated into the pUC vector (Vieira and Messing, 1982) via *Pst*I digestion, giving rise to the pUC-zeo vector. For the *galt2* KO construct, a 1093-bp fragment was deleted from the genomic sequence cloned into the pCR®4-TOPO® using *Hin*dIII and *Eco*RV sites and replaced by the Zeocin™ selection cassette obtained from the pUC-Zeo plasmid using the flanking restriction sites *Hin*dIII and *Ecl*136II. In both KO constructs, the homologous regions flanking the selection marker were around 700 bp. The *galt3* KO construct was generated by PCR performed with genomic DNA of *P. patens*. Two fragments were amplified, representing the 5′ part and the 3′ part of the KO construct, respectively. Both amplified constructs were digested with *Hin*dIII followed by ligation using T4 DNA ligase. The resulting product comprised a deletion of 270 bp with regard to the genomic *galt3* sequence. Moreover, a stop codon had been introduced within the early 5′ part of the corresponding cDNA. To generate the *Δgalt2/Δgalt3* double KOs, the rhEPO-producing *Δgalt3* mutant was transformed with the *galt2* KO construct.

The KO construct for *fuct* was synthesized by PCR amplification of a 2860-bp genomic fragment with the primers listed in Table S1. Then, the *nptII* cartridge was amplified by PCR introducing the *Bst*Z17I restriction sites and was cloned into the *Bst*Z17I in the *fuct* genomic fragment.

The human *β1,4galt* knockin construct was created by replacement of the Zeocin™ selection cassette from the *galt1* KO construct with the human *β1,4galt* expression cassette from plasmid pCR4-XTko-GTki (Huether *et al.*, 2005), resulting in the *galt1*KO/*h1,4galt*KI construct. To generate *h1,4galt* insertion lines, *Δgalt1* lines 32 and 58 were co-transfected with either the *galt1*KO/*h1,4galt*KI construct or pCR4-XTko-GTki together with pUCsul, which confers sulphadiazine resistance (Markmann-Mulisch *et al.*, 2007).

Screening of transgenic plants

Screening for transformants with the 5′ and 3′ integration of the KO construct in the correct genomic locus was performed via direct PCR (Schween *et al.*, 2002) with genomic DNA extracted using a protocol modified from Edwards *et al.* (1991). Briefly, approximately 5 mg of plant material was homogenized for 1 min in a bead mill (Qiagen, Hilden, Germany) with 200 μL of extraction buffer (200 mM Tris–HCl, pH 7.5; 250 mM NaCl; 25 mM EDTA; and 0.5% SDS), and the sample was then shortly vortexed and centrifuged at 20 000 ***g*** for 1 min. Subsequently, 150 μL of the supernatant was transferred to a fresh tube, mixed with 150 μL cold (–20 °C) isopropanol and incubated for 2 min. After centrifugation at 20 000 ***g*** for 5 min at 4 °C, the pellets were washed with 200 μL cold 70% ethanol, dried at room temperature and resuspended in 30 μL TE buffer. From these extracts, 2 μL was used as template for PCR. Plants that in this genomic DNA screening showed the expected PCR products were considered as putative KOs and selected for further analysis.

The absence of the glycosyltransferase transcripts in the KO lines and expression of *h1,4galt* in the *h1,4galt* i lines was analysed via RT-PCR. For this purpose, RNA was extracted from approximately 100 mg fresh weight plant material with TRIzol® reagent (Invitrogen) following the manufacturer's recommendations. Synthesis of cDNA was performed with M-MuLV reverse transcriptase (Fermentas, St Leon-Rot, Germany) using 2 μg RNA previously treated with DNAse I (Fermentas), and cDNA quality was checked by PCR with primers C45f and c45r which amplify the constitutively expressed gene of the ribosomal protein L21. The specific transcripts were amplified with the primers listed in Table S2.

Protein and N-glycan analysis

Total soluble proteins were recovered from 160 mL of 7- to 16-day-old culture supernatant by precipitation with 10% (w/v) trichloroacetic acid (TCA; Sigma-Aldrich, Deisenhofen, Germany) as described earlier (Büttner-Mainik *et al.*, 2011). The pellet was resuspended in sample loading buffer (Laemmli, 1970) without bromophenol blue, and total proteins were measured by means of the bicinchoninic acid assay (BCA protein assay kit; Pierce, Rockford, IL), following the manufacturer's recommendations. Bromophenol blue was added later to the samples, and electrophoretic separation of proteins was carried out in 12% SDS–polyacrylamide gels (Ready Gel Tris-HCl; BioRad, Munich, Germany) at 150 V for 1 h under non-reducing conditions.

For glycopeptide analysis, the proteins in the gels were stained with PageBlue® Protein Staining Solution (Fermentas), and the bands corresponding to 25 kDa were cut out, S-alkylated and digested with trypsin. Analysis by reversed-phase liquid chromatography coupled to electrospray ionization mass spectrometry (LC–ESI-MS) was performed as described (Grass *et al.*, 2011).

For hEPO Western blot analyses, 100 ng of proteins was loaded for SDS–PAGE as described above, and for Lewis A Western blots, 30 μg of proteins was used. Gels were then blotted

to polyvinylidene fluoride (PVDF) membranes (Immobilon-P; Millipore, Bedford, MA) in a Trans-Blot SD Semi-Dry Electrophoretic Cell (Bio-Rad) for 1 h with 1 mA/cm^2 membrane. For the hEPO Western blot, a mouse anti-hEPO monoclonal antibody (MAB2871; R&D Systems, Minneapolis, MN) and an anti-mouse antibody coupled with peroxidase (NA 931V; GE Healthcare, München, Germany) in 1 : 4000 and 1 : 150 000 dilutions, respectively, were used. As standard, 20 ng rhEPO (cat. no 287-TC; R&D Systems) was loaded. For the anti-Le^a immunoblot, a 1 : 10 dilution from the JIM 84 antibody (kindly provided by Prof. Hawes, Oxford Brookes University) and a 1 : 50 000 dilution from a rabbit anti-rat antibody (ab 6250; Abcam, Cambridge, MA) were used. Detection was performed using the ECL Advance Western Blotting Detection kit (GE Healthcare) following the manufacturer's instructions.

For lectin affinity blots, SDS–PAGE and blotting to PVDF membranes were performed as described above. Terminal Galβ1,4GlcNAc residues on blotted glycoproteins were detected with digoxigenin-labelled DSA (DIG Glycan Differentiation kit; Roche, Mannheim, Germany) according to the manufacturer's instructions.

Quantification of the moss-produced rhEPO was performed using a hEPO Quantikine IVD ELISA kit (cat. no DEP00; R&D Systems) according to the manufacturer's protocol.

Structural analysis of N-glycans was performed by MALDI-TOF MS of enzymatically released N-glycans that were isolated as described in detail elsewhere (Altmann *et al.*, 2001).

Acknowledgements

This work was supported by a grant from the German Federal Ministry of Education and Research (BMBF) No. 0313852C, by contract research 'Glykobiologie/Glykomik' of the Baden-Württemberg Stiftung and the Excellence Initiative of the German Federal and State Governments (EXC294 to R.R.). We are grateful to Dr. Johannes Stadlmann for performing the LC–ESI-MS analysis. We thank Prof. Dr. Hawes (Oxford Brookes University) for providing the JIM 84 antibody; Dagmar Krischke, Agnes Novakovic and Ingrid Heger for excellent technical assistance; and Anne Katrin Prowse for proofreading of the manuscript.

References

Altmann, F., Fabini, G., Ahorn, H. and Wilson, I.B. (2001) Genetic model organisms in the study of N-glycans. *Biochimie*, **83**, 703–712.

Bardor, M., Faveeuw, C., Fitchette, A.C., Gilbert, D., Galas, L., Trottein, F., Faye, L. and Lerouge, P. (2003) Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology*, **13**, 427–434.

Baur, A., Reski, R. and Gorr, G. (2005) Enhanced recovery of a secreted recombinant human growth factor using stabilizing additives and by co-expression of human serum albumin in the moss *Physcomitrella patens*. *Plant Biotechnol. J.* **3**, 331–340.

Bencúrová, M., Hemmer, W., Focke-Tejkl, M., Wilson, I.B. and Altmann, F. (2004) Specificity of IgG and IgE antibodies against plant and insect glycoprotein glycans determined with artificial glycoforms of human transferrin. *Glycobiology*, **14**, 457–466.

Brines, M. and Cerami, A. (2008) Erythropoietin-mediated tissue protection: reducing collateral damage from the primary injury response. *J. Intern. Med.* **264**, 405–432.

Brines, M., Patel, N.S., Villa, P., Brines, C., Mennini, T., De Paola, M., Erbayraktar, Z., Erbayraktar, S., Sepodes, B., Thiemermann, C., Ghezzi, P., Yamin, M., Hand, C.C., Xie, Q.W., Coleman, T. and Cerami, A. (2008) Nonerythropoietic, tissue-protective peptides derived from the tertiary structure of erythropoietin. *Proc. Natl Acad. Sci. U.S.A.* **105**, 10925–10930.

Büttner-Mainik, A., Parsons, J., Jérôme, H., Hartmann, A., Lamer, S., Schaaf, A., Schlosser, A., Zipfel, P.F., Reski, R. and Decker, E.L. (2011) Production of biologically active recombinant human factor H in *Physcomitrella*. *Plant Biotechnol. J.* **9**, 373–383.

Carelli, S., Marfia, G., Di Giulio, A.M., Ghilardi, G. and Gorio, A. (2011) Erythropoietin: recent developments in the treatment of spinal cord injury. *Neurol. Res. Int.* **2011**, 453179.

Cox, K.M., Sterling, J.D., Regan, J.T., Gasdaska, J.R., Frantz, K.K., Peele, C.G., Black, A., Passmore, D., Moldovan-Loomis, C., Srinivasan, M., Cuison, S., Cardarelli, P.M. and Dickey, L.F. (2006) Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol.* **24**, 1591–1597.

Davies, H.M. (2010) Review article: commercialization of whole-plant systems for biomanufacturing of protein products: evolution and prospects. *Plant Biotechnol. J.* **8**, 845–861.

Decker, E.L. and Reski, R. (2007) Moss bioreactors producing improved biopharmaceuticals. *Curr. Opin. Biotechnol.* **18**, 393–398.

Decker, E.L. and Reski, R. (2012) Glycoprotein production in moss bioreactors. *Plant Cell Rep.* **31**, 453–460.

Dordal, M.S., Wang, F.F. and Goldwasser, E. (1985) The role of carbohydrate in erythropoietin action. *Endocrinology*, **116**, 2293–2299.

Durocher, Y. and Butler, M. (2009) Expression systems for therapeutic glycoprotein production. *Curr. Opin. Biotechnol.* **20**, 700–707.

Edwards, K., Johnstone, C. and Thompson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* **19**, 1349.

Erbayraktar, S., Grasso, G., Sfacteria, A., Xie, Q.W., Coleman, T., Kreilgaard, M., Torup, L., Sager, T., Erbayraktar, Z., Gokmen, N., Yilmaz, O., Ghezzi, P., Villa, P., Fratelli, M., Casagrande, S., Leist, M., Helboe, L., Gerwein, J., Christensen, S., Geist, M.A., Pedersen, L.O., Cerami-Hand, C., Wuerth, J.P., Cerami, A. and Brines, M. (2003) Asialoerythropoietin is a nonerythropoietic cytokine with broad neuroprotective activity in vivo. *Proc. Natl Acad. Sci. U.S.A.* **100**, 6741–6746.

Fitchette, A.C., Cabanes-Macheteau, M., Marvin, L., Martin, B., Satiat-Jeunemaitre, B., Gomord, V., Crooks, K., Lerouge, P., Faye, L. and Hawes, C. (1999) Biosynthesis and immunolocalization of Lewis a-containing N-glycans in the plant cell. *Plant Physiol.* **121**, 333–344.

Fitchette-Lainé, A.C., Gomord, V., Cabanes, M., Michalski, J.C., Saint Macary, M., Foucher, B., Cavelier, B., Hawes, C., Lerouge, P. and Faye, L. (1997) N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J.* **12**, 1411–1417.

Frank, W., Decker, E.L. and Reski, R. (2005) Molecular tools to study *Physcomitrella patens*. *Plant Biol.* **7**, 220–227.

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.

Girke, T., Schmidt, H., Zahringer, U., Reski, R. and Heinz, E. (1998) Identification of a novel delta 6-acyl-group desaturase by targeted gene disruption in *Physcomitrella patens*. *Plant J.* **15**, 39–48.

Gomord, V., Fitchette, A.C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D. and Faye, L. (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol. J.* **8**, 564–587.

Gorr, G. and Jost, W. (2005) Glycosylation design in transgenic moss for better product efficacy. *Bioprocess. J.* **4**, 26–30.

Grass, J., Pabst, M., Chang, M., Wozny, M. and Altmann, F. (2011) Analysis of recombinant human follicle-stimulating hormone (FSH) by mass spectrometric approaches. *Anal. Bioanal. Chem.* **400**, 2427–2438.

Henry, S., Oriol, R. and Samuelsson, B. (1995) Lewis histo-blood group system and associated secretory phenotypes. *Vox Sang.* **69**, 166–182.

Horsley, D., Coleman, J., Evans, D., Crooks, K., Peart, J., Satiat-Jeunemaitre, B. and Hawes, C. (1993) A monoclonal antibody, JIM 84, recognizes the Golgi apparatus and plasma membrane in plant cells. *J. Exp. Bot.* **44**(Suppl.), 223–229.

Howe, K., Bateman, A. and Durbin, R. (2002) QuickTree: building huge Neighbour-Joining trees of protein sequences. *Bioinformatics*, **18**, 1546–1547.

Huether, C.M., Lienhart, O., Baur, A., Stemmer, C., Gorr, G., Reski, R. and Decker, E.L. (2005) Glyco-engineering of moss lacking plant-specific sugar residues. *Plant Biol.* **7**, 292–299.

Jelkmann, W. (2000) Use of recombinant human erythropoietin as an antianemic and performance enhancing drug. *Curr. Pharm. Biotechnol.* **1**, 11–31.

Jelkmann, W. (2007) Recombinant EPO production – points the nephrologist should know. *Nephrol. Dial. Transplant.* **22**, 2749–2753.

Jin, C., Altmann, F., Strasser, R., Mach, L., Schahs, M., Kunert, R., Rademacher, T., Glossl, J. and Steinkellner, H. (2008) A plant-derived human monoclonal antibody induces an anti-carbohydrate immune response in rabbits. *Glycobiology*, **18**, 235–241.

Kamisugi, Y., Schlink, K., Rensing, S.A., Schween, G., von Stackelberg, M., Cuming, A.C., Reski, R. and Cove, D.J. (2006) The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration. *Nucleic Acids Res.* **34**, 6205–6214.

Khraiwesh, B., Arif, M.A., Seumel, G.I., Ossowski, S., Weigel, D., Reski, R. and Frank, W. (2010) Transcriptional control of gene expression by microRNAs. *Cell*, **140**, 111–122.

Koprivova, A., Altmann, F., Gorr, G., Kopriva, S., Reski, R. and Decker, E.L. (2003) N-Glycosylation in the moss *Physcomitrella patens* is organized similarly to 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 Biotechnol. J.* **2**, 517–523.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.

Léonard, R., Costa, G., Darrambide, E., Lhernould, S., Fleurat-Lessard, P., Carlué, M., Gomord, V., Faye, L. and Maftah, A. (2002) The presence of Lewis a epitopes in *Arabidopsis thaliana* glycoconjugates depends on an active alpha4-fucosyltransferase gene. *Glycobiology*, **12**, 299–306.

Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.C., Gomord, V. and Faye, L. (1998) N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol. Biol.* **38**, 31–48.

Markmann-Mulisch, U., Wendeler, E., Zobell, O., Schween, G., Steinbiss, H.H. and Reiss, B. (2007) Differential requirements for RAD51 in *Physcomitrella patens* and *Arabidopsis thaliana* development and DNA damage repair. *Plant Cell*, **19**, 3080–3089.

Masuda, S., Nagao, M., Takahata, K., Konishi, Y., Gallyas Jr, F., Tabira, T. and Sasaki, R. (1993) Functional erythropoietin receptor of the cells with neural characteristics. Comparison with receptor properties of erythroid cells. *J. Biol. Chem.* **268**, 11208–11216.

Melo, N.S., Nimtz, M., Conradt, H.S., Fevereiro, P.S. and Costa, J. (1997) Identification of the human Lewis(a) carbohydrate motif in a secretory peroxidase from a plant cell suspension culture (*Vaccinium myrtillus* L.). *FEBS Lett.* **415**, 186–191.

Mosquna, A., Katz, A., Decker, E.L., Rensing, S.A., Reski, R. and Ohad, N. (2009) Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution. *Development*, **136**, 2433–2444.

Paul, M. and Ma, J.K. (2011) Plant-made pharmaceuticals: leading products and production platforms. *Biotechnol. Appl. Biochem.* **58**, 58–67.

Petruccelli, S., Otegui, M.S., Lareu, F., Tran Dinh, O., Fitchette, A.C., Circosta, A., Rumbo, M., Bardor, M., Carcamo, R., Gomord, V. and Beachy, R.N. (2006) A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. *Plant Biotechnol. J.* **4**, 511–527.

Reski, R. and Abel, W.O. (1985) Induction of budding on chloronemata and caulonemata of the moss, *Physcomitrella patens*, using isopentenyladenine. *Planta*, **165**, 354–358.

Schroda, M., Beck, C.F. and Vallon, O. (2002) Sequence elements within an HSP70 promoter counteract transcriptional transgene silencing in Chlamydomonas. *Plant J.* **31**, 445–455.

Schuster, M., Jost, W., Mudde, G.C., Wiederkum, S., Schwager, C., Janzek, E., Altmann, F., Stadlmann, J., Stemmer, C. and Gorr, G. (2007) In vivo glyco-engineered antibody with improved lytic potential produced by an innovative non-mammalian expression system. *Biotechnol. J.* **2**, 700–708.

Schween, G., Fleig, S. and Reski, R. (2002) High-throughput-PCR screen of 15,000 transgenic *Physcomitrella* plants. *Plant Mol. Biol. Rep.* **20**, 43–47.

Skibeli, V., Nissen-Lie, G. and Torjesen, P. (2001) Sugar profiling proves that human serum erythropoietin differs from recombinant human erythropoietin. *Blood*, **98**, 3626–3634.

Sourrouille, C., Marquet-Blouin, E., D'Aoust, M.A., Kiefer-Meyer, M.C., Seveno, M., Pagny-Salehabadi, S., Bardor, M., Durambur, G., Lerouge, P., Vezina, L. and Gomord, V. (2008) Down-regulated expression of plant-specific glycoepitopes in alfalfa. *Plant Biotechnol. J.* **6**, 702–721.

Strasser, R., Altmann, F., Mach, L., Glössl, J. and Steinkellner, H. (2004) Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking beta1,2-linked xylose and core alpha1,3-linked fucose. *FEBS Lett.* **561**, 132–136.

Strasser, R., Bondili, J.S., Vavra, U., Schoberer, J., Svoboda, B., Glossl, J., Leonard, R., Stadlmann, J., Altmann, F., Steinkellner, H. and Mach, L. (2007) A unique beta1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell*, **19**, 2278–2292.

Strasser, R., Stadlmann, J., Schahs, M., Stiegler, G., Quendler, H., Mach, L., Glossl, J., Weterings, K., Pabst, M. and Steinkellner, H. (2008) Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* **6**, 392–402.

Strasser, R., Castilho, A., Stadlmann, J., Kunert, R., Quendler, H., Gattinger, P., Jez, J., Rademacher, T., Altmann, F., Mach, L. and Steinkellner, H. (2009) Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous beta1,4-galactosylated N-glycan profile. *J. Biol. Chem.* **284**, 20479–20485.

Strepp, R., Scholz, S., Kruse, S., Speth, V. and Reski, R. (1998) Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proc. Natl Acad. Sci. U.S.A.* **95**, 4368–4373.

Takeuchi, M. and Kobata, A. (1991) Structures and functional roles of the sugar chains of human erythropoietins. *Glycobiology*, **1**, 337–346.

Takeuchi, M., Inoue, N., Strickland, T.W., Kubota, M., Wada, M., Shimizu, R., Hoshi, S., Kozutsumi, H., Takasaki, S. and Kobata, A. (1989) Relationship between sugar chain structure and biological activity of recombinant human erythropoietin produced in Chinese hamster ovary cells. *Proc. Natl Acad. Sci. U.S.A.* **86**, 7819–7822.

Tsuda, E., Kawanishi, G., Ueda, M., Masuda, S. and Sasaki, R. (1990) The role of carbohydrate in recombinant human erythropoietin. *Eur. J. Biochem.* **188**, 405–411.

Van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J.P., Loutelier-Bourhis, C., Rayon, C., Villalba, M., Koppelman, S., Aalberse, R., Rodriguez, R., Faye, L. and Lerouge, P. (2000) Beta(1,2)-xylose and alpha(1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J. Biol. Chem.* **275**, 11451–11458.

Vieira, J. and Messing, J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*, **19**, 259–268.

Viëtor, 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.

Weise, A., Altmann, F., Rodriguez-Franco, M., Sjoberg, E.R., Bäumer, W., Launhardt, H., Kietzmann, M. and Gorr, G. (2007) High-level expression of secreted complex glycosylated recombinant human erythropoietin in the *Physcomitrella* Delta-fuc-t Delta-xyl-t mutant. *Plant Biotechnol. J.* **5**, 389–401.

Wilson, I.B. (2001) Identification of a cDNA encoding a plant Lewis-type alpha1,4-fucosyltransferase. *Glycoconj. J.* **18**, 439–447.

Wilson, I.B. (2002) Glycosylation of proteins in plants and invertebrates. *Curr. Opin. Struct. Biol.*, **12**, 569–577.

Wilson, I.B., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C.J., Kamerling, J.P. and Altmann, F. (2001) Analysis of Asn-linked glycans from vegetable

foodstuffs: widespread occurrence of Lewis a, core alpha1,3-linked fucose and xylose substitutions. *Glycobiology*, **11**, 261–274.
Zhang, K., Baeckstrom, D. and Hansson, G.C. (1994) A secreted mucin carrying sialyl-Lewis a from colon carcinoma cells binds to E-selectin and inhibits HL-60 cell adhesion. *Int. J. Cancer*, **59**, 823–829.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Schematic representation of the *fuct* and *galt* corresponding knockout constructs and molecular analysis of the knockout mutants.
Figure S2 Extracellular proteins from rhEPO producing moss.
Figure S3 Expression analysis of human *β1,4galt* in the putative *h1,4galt* insertion lines.
Table S1 Oligonucleotides used for DNA amplification for KO constructs cloning.
Table S2 Oligonucleotides used for glycosyltransferases transcript screening.
Table S3 Quantification of moss produced rhEPO.
Table S4 International Moss Stock Center numbers of plants used in this work.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.