

Construction of a Functional CMP-Sialic Acid Biosynthesis Pathway in Arabidopsis¹[OA]

Alexandra Castilho, Martin Pabst, Renaud Leonard, Christiane Veit, Friedrich Altmann, Lukas Mach, Josef Glössl, Richard Strasser*, and Herta Steinkellner

Institute of Applied Genetics and Cell Biology (A.C., C.V., L.M., J.G., R.S., H.S.) and Department of Chemistry (M.P., R.L., F.A.), University of Natural Resources and Applied Life Sciences, BOKU-Vienna, A-1190 Vienna, Austria

Previous studies have reported that plants contain negligible amounts of free or protein-bound *N*-acetylneuraminic acid (Neu5Ac). This is a major disadvantage for the use of plants as a biopharmaceutical expression system, since *N*-glycans with terminal Neu5Ac residues are important for the biological activities and half-lives of recombinant therapeutic glycoproteins in humans. For the synthesis of Neu5Ac-containing *N*-glycans, plants have to acquire the ability to synthesize Neu5Ac and its nucleotide-activated derivative, cytidine monophospho-*N*-acetylneuraminic acid. In this study, we have generated transgenic *Arabidopsis* (*Arabidopsis thaliana*) plants expressing three key enzymes of the mammalian Neu5Ac biosynthesis pathway: UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase, *N*-acetylneuraminic acid phosphate synthase, and CMP-*N*-acetylneuraminic acid synthetase. Simultaneous expression of UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase and *N*-acetylneuraminic acid phosphate synthase resulted in the generation of significant Neu5Ac amounts (1,275 nmol g⁻¹ fresh weight in leaves) in planta, which could be further converted to cytidine monophospho-*N*-acetylneuraminic acid (2.4 nmol g⁻¹ fresh weight in leaves) by coexpression of CMP-*N*-acetylneuraminic acid synthetase. These findings are a major step toward the production of Neu5Ac-containing glycoproteins in plants.

In recent years, plants have become an attractive alternative for the production of therapeutically relevant proteins (Saint-Jore-Dupas et al., 2007). Several studies have demonstrated that plants can efficiently produce complex mammalian proteins that need extensive posttranslational modifications (for a recent review, see Fischer et al., 2004). *N*-Glycosylation is one of the most common and important posttranslational modifications of proteins, and an advantage of using plants as an expression system is their ability to perform *N*-glycosylation similar to mammalian cells. However, besides a common core structure, plant *N*-glycans differ from their mammalian counterparts to some extent as they (1) carry plant-specific immunogenic β 1,2-Xyl and α 1,3-Fuc residues and (2) lack terminal β 1,4-Gal and *N*-acetylneuraminic acid (Neu5Ac) residues. Hence, the production of recombinant therapeutic glycoproteins in plants requires modifications

of the plant *N*-glycosylation pathway not just to overcome the possible immunogenicity of nonmammalian residues but also to enable the addition of terminal Neu5Ac. Engineering of the *N*-glycosylation pathway in plants by knockout and RNA interference approaches resulted in lines that synthesize complex *N*-glycan structures lacking the plant-specific glycan epitopes (Koprivova et al., 2004; Strasser et al., 2004, 2008; Cox et al., 2006; Schähs et al., 2007). The predominant *N*-glycan structure in these plants is GlcNAc₂Man₃GlcNAc₂ (GnGn), which is the required acceptor substrate for its sequential elongation with β 1,4-Gal and Neu5Ac residues. Functional expression of human β 1,4-galactosyltransferase in plants already has been achieved (Palacpac et al., 1999; Bakker et al., 2001).

The majority of *N*-glycans on mammalian glycoproteins are terminated by Neu5Ac and other sialic acids linked to terminal β 1,4- or β 1,3-Gal residues. These negatively charged sugars affect the biological activities and half-lives of many therapeutic glycoproteins (Schauer, 2000; Erbayraktar et al., 2003; Varki, 2007). For example, tobacco (*Nicotiana tabacum*)-derived follicle-stimulating hormone was demonstrated to be fully active in vitro but exhibited only limited activity in vivo, most probably due to its low circulatory stability as a consequence of the lack of Neu5Ac (Dirnberger et al., 2001). The sialylation of glycoproteins involves the enzymatic transfer of Neu5Ac from its activated form, cytidine monophospho-*N*-acetylneuraminic acid (CMP-Neu5Ac), to a β 1,4-Gal-terminated

¹ This work was supported by Grant LS 154 from the Vienna Science and Technology Fund and by a grant from the Austrian Science Fund (grant no. P18314 to H.S.).

* Corresponding author; e-mail richard.strasser@boku.ac.at.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Richard Strasser (richard.strasser@boku.ac.at).

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.108.117572

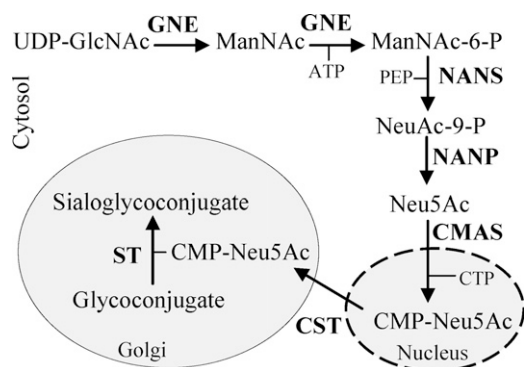


Figure 1. Schematic representation of the mammalian pathway for the sialylation of glycoconjugates. The enzymes involved in the process are indicated: GNE, NANS, NANP, CMAS, CMP-Neu5Ac transporter (CST), and sialyltransferase (ST).

acceptor glycoprotein (Fig. 1). A sialyltransferase-like activity was recently reported in rice (*Oryza sativa*; Takashima et al., 2006), and some *Arabidopsis* (*Arabidopsis thaliana*) proteins show homology to mammalian enzymes involved in sialylation (Shah et al., 2003). However, plants contain at best minute amounts of Neu5Ac and CMP-Neu5Ac (Seveno et al., 2004; Zeleny et al., 2006). Therefore, glycoengineering of plant N-glycans into Neu5Ac-terminated oligosaccharides requires the functional expression of heterologous enzymes and transporters able to catalyze (1) Neu5Ac synthesis, (2) Neu5Ac activation, (3) CMP-Neu5Ac transport into the Golgi apparatus, and (4) Neu5Ac transfer to a β 1,4-galactosylated acceptor substrate.

Initial attempts to introduce Neu5Ac residues into plant N-glycans have concentrated mainly on the heterologous expression of enzymes and proteins that act in a late stage of the biosynthetic pathway. Rat α 2,6-sialyltransferase expressed in transgenic *Arabidopsis* was found to be enzymatically active (Wee et al., 1998). In addition, human β 1,4-galactosyltransferase was functionally expressed in plants (Palacpac et al., 1999; Bakker et al., 2001), and successful expression of a mammalian CMP-sialic acid transporter and a CMP-sialic acid synthetase in cultured plant cells has been reported (Misaki et al., 2006). However, with the exception of β 1,4-galactosyltransferase, in planta activity has not been demonstrated for any of these proteins.

In this study, we report the in planta expression of three enzymes required for the synthesis of CMP-Neu5Ac in mammals: mouse UDP-GlcNAc 2-epimerase/*N*-acetylmannosamine kinase (GNE; Horstkorte et al., 1999), human *N*-acetylneuraminic acid phosphate synthase (NANS; Lawrence et al., 2000), and human CMP-sialic acid synthetase (CMAS; Lawrence et al., 2001). The genes were simultaneously expressed in *Arabidopsis*, which resulted in the generation of significant amounts of Neu5Ac and CMP-Neu5Ac from endogenous precursors.

RESULTS

Heterologous Expression of Mouse GNE

In mammals, one of the initial steps in the Neu5Ac biosynthesis pathway is the conversion of the nucleotide sugar UDP-GlcNAc, which is naturally present in plant cells (Jiang et al., 2005), first to ManNAc and subsequently to ManNAc-6-P by the bifunctional enzyme GNE. To investigate whether mouse GNE recombinantly expressed in plants is functional, *Arabidopsis* plants were stably transformed with a binary vector harboring a myc-tagged GNE cDNA construct (35S:GNE; Fig. 2A). Expression of GNE was tested by immunoblotting of leaf extracts, and a high-expressing line was selected for further propagation. Western-blot analysis with anti-myc antibody showed a band of the expected size for the GNE protein (80 kD), which was not present in nontransformed plants. An additional weak band of approximately 40 kD visible in both transformed and nontransformed plants is due to unspecific binding of the anti-myc antibody (Fig. 3).

For activity assays, total soluble proteins from seedling cultures were extracted and incubated with UDP-GlcNAc in the presence of ATP. The reaction products were derivatized with 2-aminobenzoic acid (AA) and analyzed by reverse-phase liquid chromatography-electrospray ionization-tandem mass spectrometry (RP-LC-ESI-MS/MS). A peak eluting at a similar position as the standard (GlcNAc-6-P-AA) was found in 35S:GNE plants but not in wild-type plants. The fragmentation pattern of this peak was characteristic of HexNAc-P-AA (Fig. 4A), showing that the enzyme efficiently converts UDP-GlcNAc to ManNAc-6-P in vitro (generating 528 pmol ManNAc-6-P h⁻¹ mg⁻¹ total soluble protein). We also investigated the in vivo functionality of GNE by determining the amounts of ManNAc-6-P produced in *Arabidopsis* seedling cultures by ESI-MS/MS analysis. To corroborate the in vitro findings with a different methodology, unde-

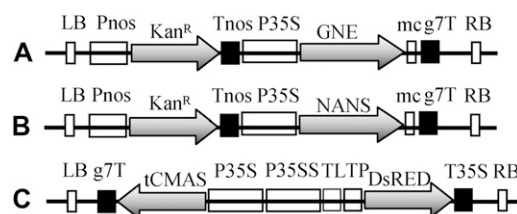


Figure 2. Schematic representation of the different plant expression cassettes used in this study. A, 35S:GNE. B, 35S:NANS. C, 35S:tCMAS. g7T, *Agrobacterium tumefaciens* gene 7 terminator; Kan^R, neomycin phosphotransferase II; LB, left border; mc, c-myc epitope tag; Pnos, nopaline synthase gene promoter; P35S, promoter of the 35S transcript of the *Cauliflower mosaic virus*; P35SS, promoter of the 35S transcript of the *Cauliflower mosaic virus* with double enhancer; RB, right border; TL, translational enhancer (5' untranslated region of *Tobacco etch virus*); Tnos, nopaline synthase gene terminator; T35S, terminator of the 35S transcript of the *Cauliflower mosaic virus*; TP, transit peptide from barley (*Hordeum vulgare*) granule-bound starch synthase I.

derivatized samples were analyzed in this case. The fragmentation pattern of the HexNAc-6-P peak obtained for 35S:GNE-expressing plants was indicative of ManNAc-6-P (Fig. 4B). Significant amounts of ManNAc-6-P were found (10 nmol g⁻¹ fresh weight seedlings). These results demonstrate that mouse GNE can synthesize ManNAc-6-P in planta entirely from endogenous precursors, presumably UDP-GlcNAc.

Heterologous Expression of Human NANS

In mammals, ManNAc-6-P is converted by NANS into Neu5Ac-9-P through condensation with phosphoenolpyruvate (PEP). In order to generate plants that express an active form of human NANS, transgenic Arabidopsis plants carrying a 35S:NANS construct (Fig. 2B) were generated. The integration of the transgene was confirmed by genomic PCR (data not shown). Positive lines were analyzed for the expression of the recombinant protein by western blotting using anti-myc antibodies. A line exhibiting strong NANS expression was chosen for further analysis. Immunoblotting exhibited a prominent band with the expected size of approximately 42 kD in this line (Fig. 3).

To test the *in vitro* activity of recombinant NANS, soluble protein extracts from seedling cultures were incubated with ManNAc-6-P and PEP. The reaction products were then derivatized with 1,2-diamino-4,5-methylene dioxybenzene (DMB). RP-HPLC analysis of the samples obtained using extracts from 35S:NANS plants showed a peak eluting at the same retention time as the standard (DMB-Neu5Ac; Fig. 5). This peak was not detected in samples incubated without ManNAc-6-P or in samples derived from extracts of nontransformed plants incubated with ManNAc-6-P and PEP (WT+ManNAc-6-P). RP-LC-ESI-MS/MS analysis revealed that only the reaction product obtained with 35S:NANS extracts exhibited a fragmentation

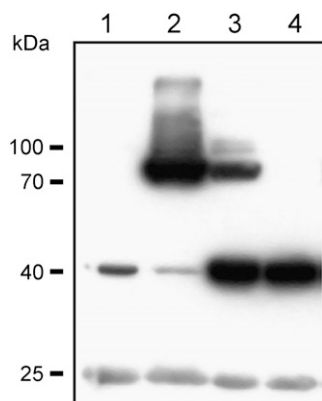


Figure 3. Immunoblot of Arabidopsis leaf extracts using anti-myc antibodies. Proteins were subjected to SDS-PAGE under reducing conditions. Lane 1, extract from the wild type; lane 2, extract from the 35S:GNE transgenic line; lane 3, extract from the 35S:GNE/35S:NANS transgenic line; lane 4, extract from the 35S:NANS transgenic line. Protein sizes are indicated.

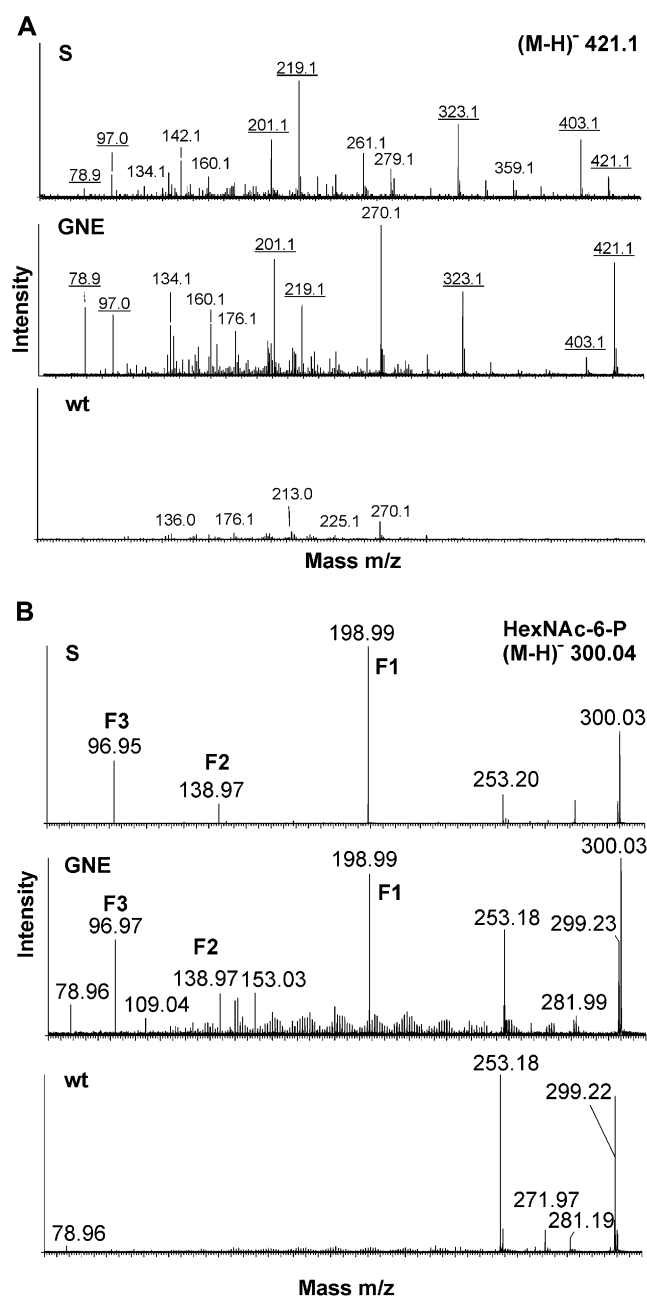


Figure 4. ManNAc-6-P production *in vitro* and *in planta* by GNE expression in Arabidopsis seedling cultures. A, Determination of ManNAc-6-P production *in vitro*. MS/MS spectra of the HexNAc-6-P peak region acquired from the GlcNAc-6-P standard (S), from extracts of 35S:GNE transgenic plants (GNE), and from wild-type plants (wt) incubated in the presence of UDP-GlcNAc and ATP. Masses of fragments arising from AA-labeled HexNAc-6-P are underlined. B, Demonstration of ManNAc-6-P production *in planta*. The presence of HexNAc-6-P in Arabidopsis extracts was analyzed by direct infusion ESI-MS/MS analysis of the [M-H]⁻ parent ion of mass 300.0. Characteristic fragmentation products are labeled F1, F2, and F3. Spectra were obtained with GlcNAc-6-P standard (S) and extracts from 35S:GNE transgenic plants (GNE) and wild-type plants (wt).

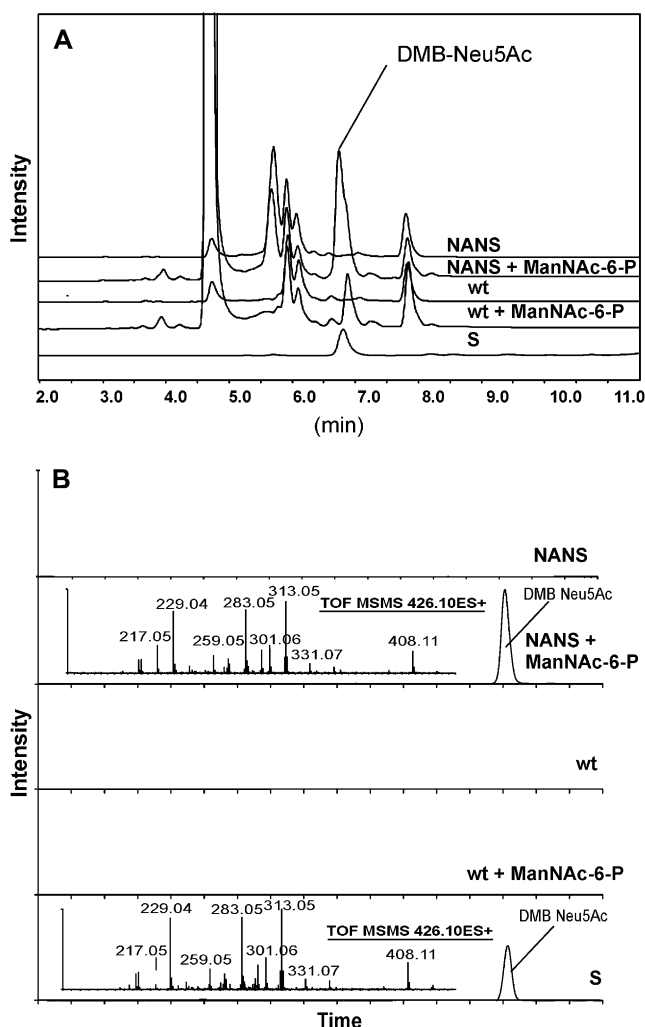


Figure 5. A, In vitro NANS activity assay analyzed by RP-HPLC of DMB-Neu5Ac in extracts from Arabidopsis seedlings: DMB-Neu5Ac standard (S), DMB-labeled extract from wild-type plants after incubation in the presence (wt + ManNAc-6-P) or absence (wt) of ManNAc-6-P, and 35S:NANS transgenic plants after incubation in the presence (NANS + ManNAc-6-P) or absence (NANS) of ManNAc-6-P. The elution position of DMB-Neu5Ac is indicated. B, The eluate in the region of DMB-Neu5Ac was collected, concentrated, and subjected to RP-LC-ESI-MS/MS to acquire a fragment spectrum of the presumed DMB-Neu5Ac peaks. All traces are shown with the same y axis setting.

pattern corresponding to DMB-Neu5Ac, whereas the peak seen in WT+ManNAc-6-P samples was clearly not DMB-Neu5Ac (Fig. 5B). This also demonstrates that Arabidopsis plants are essentially devoid of endogenous NANS activity.

Our results show that crude extracts of Arabidopsis plants expressing human NANS are able to efficiently convert ManNAc-6-P to Neu5Ac. In mammals, Neu5Ac-9-P formed by NANS is dephosphorylated by a Neu5Ac-9-P phosphatase (NANP; Maliekal et al., 2006). The accumulation of Neu5Ac rather than Neu5Ac-9-P in our assays suggests that plants contain a NANP-like enzyme.

In Vivo Synthesis of Neu5Ac in Plants

To test whether ManNAc-6-P produced by GNE can be further converted into Neu5Ac by NANS in planta, the 35S:GNE and 35S:NANS Arabidopsis lines were crossed. Western-blot analysis of the protein extracts from the resulting F₁ progeny revealed the presence of the 80- and 42-kD bands (Fig. 3), indicating the expression of both GNE and NANS. After derivatization with DMB, the presence of Neu5Ac in extracts from leaves of 35S:GNE/35S:NANS and wild-type plants was analyzed by RP-HPLC and ESI-time of flight (TOF)-MS/MS (Fig. 6). No Neu5Ac (detection limit, 10 pmol g⁻¹ fresh weight) was detected in wild-type plants; however, significant amounts (1,275 nmol g⁻¹ fresh weight) were detected in leaves from the transgenic Arabidopsis line expressing both GNE and NANS (Table I). Together, these data show that Neu5Ac can be produced in quantitative amounts in plants co-expressing mammalian GNE and NANS.

In Vivo Synthesis of CMP-Neu5Ac in Plants

In mammals, Neu5Ac is finally converted to CMP-Neu5Ac, a reaction catalyzed by CMAS. Since we failed to functionally express a full-length version of human CMAS in Arabidopsis, a construct encoding a truncated form of the enzyme lacking its 40 N-terminal amino acids (35S:tCMAS) was used to transform Arabidopsis. This construct lacks a Pro/Gly-rich domain that has been described to destabilize the protein (Krapp et al., 2003). Additionally, for screening purposes, a red fluorescent protein (DsRed) marker was included in the binary vector. This vector was used to transform 35S:GNE/35S:NANS plants. Seeds were screened for DsRed expression and kanamycin resistance, and the presence of the three transgenes was confirmed by PCR (data not shown). CMP-Neu5Ac synthesis was analyzed by ESI-TOF-MS/MS in extracts from 35S:GNE/35S:NANS/35S:tCMAS leaves and compared with that in wild-type plants. No CMP-Neu5Ac peak was detected in wild-type plants. However, in transgenic Arabidopsis expressing all three mammalian enzymes (GNE, NANS, and CMAS), a single peak was detected that coeluted with the standard CMP-Neu5Ac (Fig. 7A). Fragmentation of the corresponding peak led to a pattern that was identical with that of the standard (Fig. 7B). Quantification of the peak revealed the presence of significant amounts (2.4 nmol g⁻¹ fresh weight; Table I) of CMP-Neu5Ac. Together, these data show that CMP-Neu5Ac can be produced in plants by introducing three genes derived from the mammalian sialic acid biosynthesis pathway.

Total N-Glycan Analysis in Neu5Ac- and CMP-Neu5Ac-Producing Plants

Since our ultimate aim is to generate plants capable of synthesizing N-glycans with terminal Neu5Ac residues, we tested whether the coexpression of GNE,

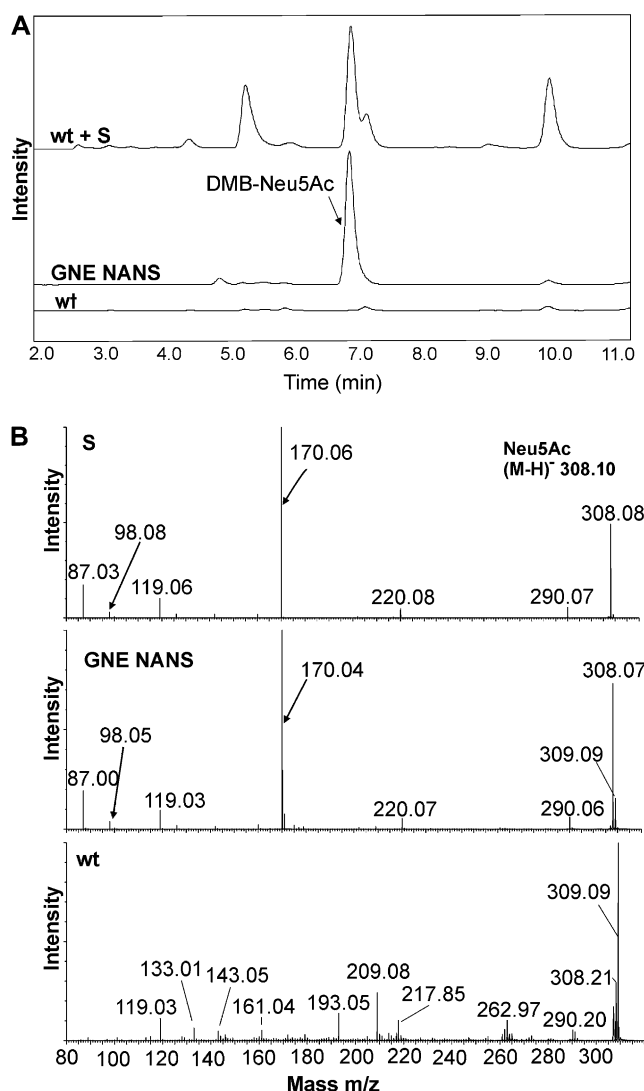


Figure 6. In planta synthesis of Neu5Ac. A, RP-HPLC of DMB-Neu5Ac in wild-type extract spiked with DMB-Neu5Ac (wt + S), extract from GNE- and NANS-expressing transgenic Arabidopsis (GNE NANS), and wild-type control (wt). The DMB-Neu5Ac peak is indicated. The background peaks are higher in the spiked wild-type extract because 10 times more sample was injected there. B, Mass spectrometric detection (ESI-MS/MS on mass 308.1) of underivatized Neu5Ac standard (S), direct infusion of partially purified extract from GNE- and NANS-expressing transgenic Arabidopsis (GNE NANS), and wild-type control (wt). Characteristic peaks are marked by arrows.

NANS, and CMAS interferes with the *N*-glycosylation machinery (e.g. by depriving the UDP-GlcNAc-dependent Golgi enzymes GlcNAc transferase I and II of their donor substrate). We first analyzed total protein extracts from leaves of different transgenic lines and wild-type plants by western blot using antibodies recognizing plant complex *N*-glycans carrying β 1,2-Xyl and core α 1,3-Fuc (Fig. 8A). No obvious differences were detected in the staining intensities between wild-type plants and the transgenic lines, indicating that the *N*-glycosylation pathway is not severely af-

ected by the production of high amounts of Neu5Ac and CMP-Neu5Ac. To detect even subtle changes in the *N*-glycan pattern, total *N*-glycans were isolated from leaves and analyzed by matrix-assisted laser-desorption ionization (MALDI)-TOF-MS. The relative amounts of complex *N*-glycans terminating with GlcNAc residues (GnGnXF and GnMXF), whose synthesis depends on GlcNAc transferases I and II, did not differ between wild-type plants and the transgenic

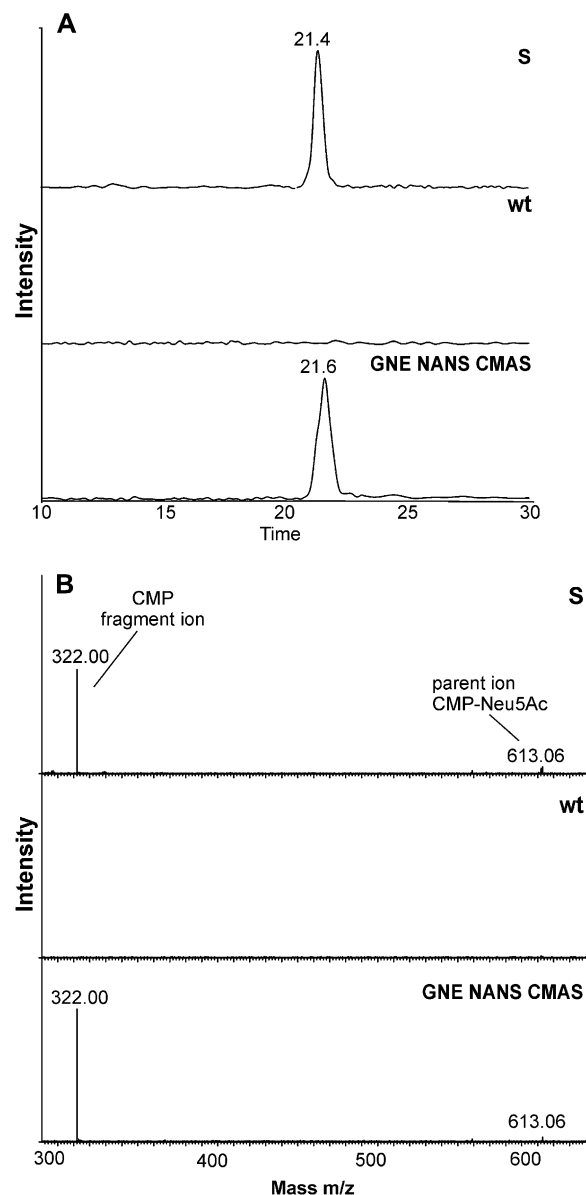


Figure 7. In planta synthesis of CMP-Neu5Ac. A, LC-ESI-MS/MS analysis. CMP-Neu5Ac standard (S), extract from wild-type control (wt), and extract from GNE-, NANS-, and CMAS-expressing transgenic Arabidopsis (GNE NANS CMAS) were analyzed with parent ion mass set to 613.1 D. B, Mass spectra at the elution position of CMP-Neu5Ac: standard CMP-Neu5Ac (S), wild-type control (wt), and GNE-, NANS-, and CMAS-expressing transgenic Arabidopsis (GNE NANS CMAS). All runs were set to identical absolute intensity.

Table 1. Amounts of Neu5Ac and CMP-Neu5Ac determined in rosette leaves from 6-week-old *Arabidopsis* plants

The data represent average values of two independent samples.

Plant	Neu5Ac	CMP-Neu5Ac
	<i>nmol g⁻¹ fresh weight</i>	
Columbia	n.d. ^a	n.d.
GNE NANS	1,275	n.d.
GNE NANS CMAS	41	2.4

^an.d., not detectable (<10 pmol g⁻¹ fresh weight).

lines (Fig. 8B), despite the production of Neu5Ac and CMP-Neu5Ac. Notably, the transgenic plants producing high amounts of Neu5Ac and CMP-Neu5Ac in leaves (Table I) did not show any obvious phenotype under standard growth conditions.

DISCUSSION

Previous studies in our laboratory have shown that Neu5Ac is essentially undetectable in plants (Zeleny et al., 2006). These findings suggested that metabolic engineering of the Neu5Ac biosynthesis pathway is required to enable plants to synthesize significant amounts of Neu5Ac-containing glycoproteins. As an important step toward this goal, we have now successfully expressed three key enzymes in *Arabidopsis*: mouse GNE for the conversion of endogenous UDP-GlcNAc into ManNAc-6-P, human NANS for the synthesis of Neu5Ac-9-P from ManNAc-6-P, and human CMAS for the subsequent transformation of Neu5Ac into the sugar nucleotide CMP-Neu5Ac.

While Neu5Ac is generated in bacteria by sialic acid synthase through the condensation of ManNAc with PEP (Vann et al., 1997), mammalian cells use NANS to convert ManNAc-6-P into Neu5Ac-9-P. Neu5Ac-9-P is then dephosphorylated by NANP, which was identified only recently (Maliekal et al., 2006). The accumulation of Neu5Ac rather than Neu5Ac-9-P indicates that *Arabidopsis* contains a NANP homolog that catalyzes the dephosphorylation step. In fact, a putative plant NANP sharing approximately 30% identity with the human NANP is present in the genome of *Arabidopsis* (Q9ZVB6; gene identifier At2g41250).

Previous attempts to generate sialic acid in plants by expressing microbial sialic acid synthase showed a relatively low expression level of the heterologous protein, and although the bacterial enzyme was able to synthesize Neu5Ac in vitro, no in vivo accumulation of sialic acid was detected (Paccalet et al., 2007). Our data suggest that engineering of the mammalian Neu5Ac biosynthesis pathway into plants instead of using the bacterial enzymes is more efficient.

We show that the amounts of Neu5Ac produced in transgenic GNE-, NANS-, and CMAS-expressing *Arabidopsis* are quantitatively significant (41 nmol g⁻¹ fresh weight, equivalent to approximately 3 nmol mg⁻¹ total soluble protein) and comparable to the

levels found in mammalian cells (1 nmol mg⁻¹ protein; Tietze et al., 1989). Notably, the amount of Neu5Ac detected in leaves of GNE- and NANS-expressing plants was approximately 30 times higher compared with that in plants expressing all three mammalian

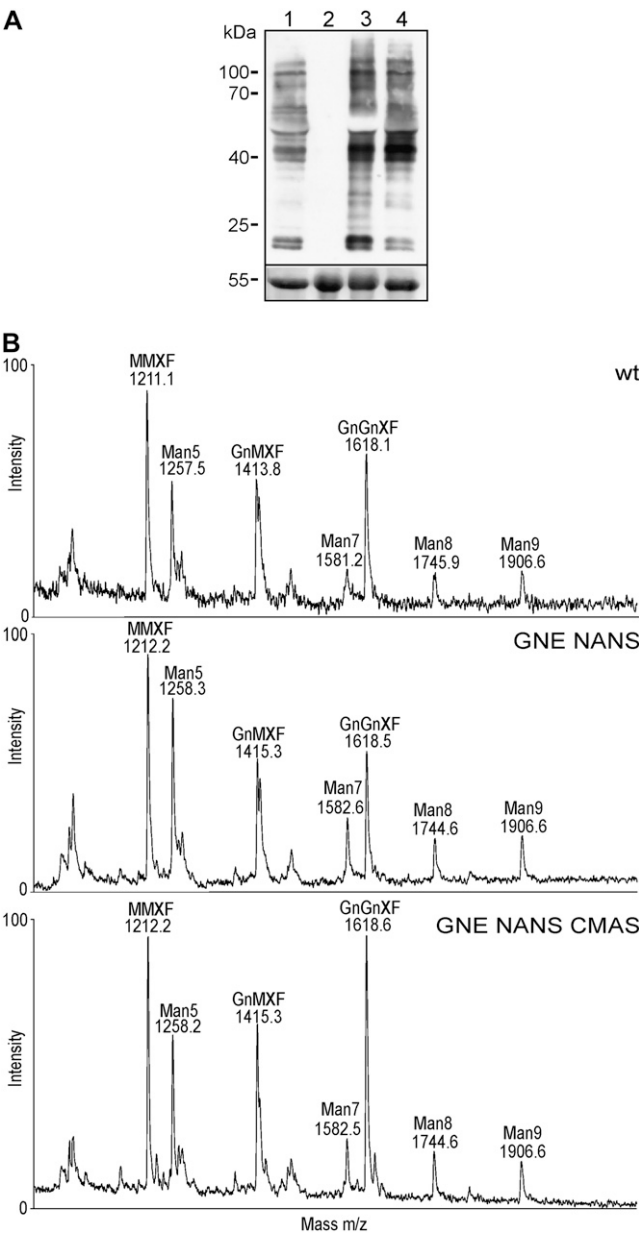


Figure 8. Analysis of changes in overall N-glycosylation in Neu5Ac- and CMP-Neu5Ac-producing *Arabidopsis* plants. A, Proteins were extracted from leaves and subjected to SDS-PAGE under reducing conditions. Immunoblotting was performed using anti-horseradish peroxidase antibodies. Lane 1, *Arabidopsis* wild type; lane 2, *xylt fuct* mutant plant (negative control; Strasser et al., 2004); lane 3, GNE NANS plant; lane 4, GNE NANS CMAS plant. The bottom panel shows Ponceau S staining of the membrane. B, MALDI-TOF-MS analysis of total N-glycans present in leaves from wild-type (wt), GNE NANS, and GNE NANS CMAS *Arabidopsis* plants. For abbreviations of N-glycan structures, see <http://www.proglycan.com/>.

proteins. This could be due to one of the following reasons. Neu5Ac is efficiently converted to CMP-Neu5Ac, and the amounts of CMP-Neu5Ac generated are actually much higher than the amounts detected in the soluble extract. Alternatively, the production of the nucleotide sugar CMP-Neu5Ac could lead to allosteric feedback inhibition of GNE activity (Kornfeld et al., 1964). This feedback inhibition could be avoided by using a mutated form of GNE, as reported for insect cells (Viswanathan et al., 2005). In addition, we cannot rule out the possibility that in the transgenic lines expressing the three enzymes, the amount of the GNE is lower due to gene-silencing effects resulting, for example, from the presence of additional copies of the 35S promoter (Fig. 2).

The amount of detected CMP-Neu5Ac (2.4 nmol g^{-1} fresh weight, equivalent to approximately 200 pmol mg^{-1} total soluble protein) is 13 times lower than the amount found in rat liver (32 nmol g^{-1} wet weight; Pels Rijcken et al., 1990) but almost as high as in Chinese hamster ovary cells ($0.12 \text{ nmol per } 10^{-6}$ cells, corresponding to approximately 600 pmol mg^{-1} total soluble protein; Hooker et al., 1999).

It has been reported for rice roots that a reduction of their UDP-GlcNAc content interferes with protein *N*-glycosylation and causes cell shrinkage (Jiang et al., 2005). Since UDP-GlcNAc is the starting metabolite of Neu5Ac synthesis, high-level Neu5Ac production in plants could cause a depletion of endogenous UDP-GlcNAc pools and thus affect plant growth and development. However, we did not observe any obvious phenotypes in our transgenic Arabidopsis plants producing large amounts of Neu5Ac. Furthermore, the overall *N*-glycosylation pattern did not show any significant alterations compared with wild-type plants. This strongly indicates that our metabolic engineering approach did not affect endogenous nucleotide sugar pools, especially UDP-GlcNAc levels.

The generation of Arabidopsis plants able to synthesize significant amounts of CMP-Neu5Ac, in combination with the already achieved in planta expression of the CMP-Neu5Ac transporter (Misaki et al., 2006) as well as the two Golgi-located transferases (1,4-galactosyltransferase and α 2,6-sialyltransferase; Wee et al., 1998; Palacpac et al., 1999; Bakker et al., 2001), suggest the feasibility of fully reconstructing the mammalian *N*-glycosylation pathway in plants.

MATERIALS AND METHODS

Construction of myc and DsRed Plant Expression Vectors

The binary plant expression vector p19 was generated by insertion of a myc tag linker fragment containing a stop codon into the *Bam*HI site of pPT2 (Strasser et al., 2005). To use the DsRed (Jach et al., 2001) as a reporter for plant gene expression, the pPT2 expression cassette was excised with *Eco*RI and *Hind*III and inserted into the pPZP200 vector (Hajdukiewicz et al., 1994) cut with the same restriction enzymes. The DsRed expression cassette was isolated from the pTRAP-2G12-Ds vector (Schähs et al., 2007) as two separate PCR products in order to remove the internal *Bam*HI restriction site: the primer pair CA355F/dsRed3R (*Hind*III, 5'-TATATGGAAGCTTCGGGTGG-AGGGGATCAG-3'/*Bgl*II, 5'-ATATAGATCTCCGCTAAAGGAACAGATG-

GTGGC-3') amplifies a fragment including the promoter and DsRed gene. The second PCR uses the Ca35PF/Ca35PR primers (*Bgl*II, 5'-TATAAGATCT-GTCCGCAAAATCACCAGTCTC-3'/*Hind*III, 5'-GGTAAAGCTTCTCGA-GGGTCACTGGATTTTGGTTT-3') to amplify the terminator. The PCR products, digested with *Hind*III and *Bgl*II, were inserted into the *Hind*III site of the pPZP200 already containing the pPT2 cassette. The resulting vector, pMS, has two expression cassettes inserted head to head: one containing the DsRed gene for screening and one empty for cloning target genes.

Bacterial strains containing the cDNAs of mouse GNE (clone IRAKp961N12108Q), human NANS (clone IRALp962F017Q), and human CMAS (clone DKFZp666I142Q) were purchased from the German Resource Centre for Genome Research (<http://www.rzpd.de>). All clones were completely sequenced to confirm the presence of the full-length open reading frame.

The coding regions lacking the stop codons of the GNE and NANS genes were amplified by PCR with primers GNE3F (*Xba*I, 5'-TATATCTAGAAATG-GAGAAGAACGGGAACAAC-3')/GNE4R (*Bgl*II, 5'-TATAAGATCTGTGG-ATCTGCGCGTTGTGTA-3') and NANS3F (*Xba*I, 5'-TATATCTAGAAATG-CGCTGGAGCTGGAGCTG-3')/NANS4R (*Bam*HI, 5'-TATAGGATCCAGA-CTTGATTTTGGCCATG-3'). A truncated form of human CMAS lacking the first 40 amino acids was amplified by PCR with primers CMAS12F (*Xba*I, 5'-TATATCTAGAAATGAAGCCCCGACCTGGCAGCCCTA-3') and CMAS5R (*Bam*HI, 5'-TATAGGATCCCTATTTTGGCATGAATTATTAAC-3'). The PCR products were purified from the gel using the Invisorb Spin PCRapid kit (Invitek), digested with *Xba*I/*Bgl*II (GNE) or with *Xba*I/*Bam*HI (NANS and CMAS), and cloned into the *Xba*I/*Bam*HI-digested p19 (GNE and NANS) and pMS (CMAS) vectors. The resulting plant expression vectors were named p19GNE (35S:GNE), p19NANS (35S:NANS), and pMS-tCMAS (35S:tCMAS). Individual clones were sequenced using the PRISM BigDye terminator cycle sequence kit and ABI 3100 genetic analyzer (Applied Biosystems).

Generation of Arabidopsis Transgenic Lines

Transgenic Arabidopsis (*Arabidopsis thaliana*) plants were generated by floral dipping as described by Strasser et al. (2005). Kanamycin-resistant plants were screened by PCR to confirm the presence of GNE, NANS, and CMAS using the primer pairs GNE5F (5'-GGATGCCCTGATCTCGTTTA-3')/GNE6R (5'-AACGATTTACCCCTTCATGC-3'), NANS1F (5'-CTGCGGGT-CATCTCGGAATA-3')/NANS2R (5'-CCCAGGGCAGCTCCACA-3'), and CMAS8F (5'-AGGGGCTTCCAGAGTGAT-3')/CMAS9R (5'-TATGTTCA-GCTTCGATTCG-3'), respectively.

Individual transgenic plants carrying the 35S:GNE and 35S:NANS constructs were crossed for the coexpression of both proteins in a single plant. The presence of the cDNA sequences in *F*₁ hybrid plants was confirmed by PCR from genomic DNA. The *F*₂ generation of transgenic plants was propagated on soil and as seedling cultures in liquid Murashige and Skoog basal medium (Sigma-Aldrich), pH 5.8, containing 3% (w/v) Suc. The flasks were incubated for 2 d at 4°C in the dark and then transferred to constant light under agitation (130 rpm) at 20°C for 12 d. For coexpression of the three mammalian genes, *F*₂ hybrid plants growing on soil were transformed with pMS-tCMAS by floral dipping, and potential CMAS transgenic Arabidopsis seeds were selected for DsRed expression with a Leica MZ FLIII stereomicroscope equipped with a DsRed filter set. The presence of CMAS was confirmed by PCR with the primers CMAS8F/CMAS9R.

Immunoblot Analysis

Total proteins were extracted from leaves with $1 \times$ PBS supplemented with 1% (v/v) Triton X-100. Proteins from young leaves or from four to five Arabidopsis seedlings were extracted in the same way. The protein extracts were centrifuged twice at 8,000 rpm for 10 min at 4°C and subjected to 12.5% SDS-PAGE under reducing conditions. The fractionated proteins were blotted to Hybond-ECL nitrocellulose membranes (GE Healthcare) and blocked for 1 h in $1 \times$ PBS supplemented with 3% (w/v) bovine serum albumin and 1% (v/v) Tween 20, and detection was performed using anti-myc antibodies (9E10; Santa Cruz Biotechnology) and SuperSignal West Pico Chemiluminescent substrate (Pierce). Immunoblots for the detection of complex *N*-glycans carrying β 1,2-Xyl and core α 1,3-Fuc were made using anti-horseradish peroxidase antibodies as described previously (Strasser et al., 2004).

Determination of in Vitro GNE Activity

Seedlings (2 g) were homogenized in 1 mL of 100 mM Tris-HCl, pH 7.5, and 20 mM MgCl₂ in the presence of protease inhibitors (Complete Mini, EDTA

free; Roche). After centrifugation at 13,000 rpm for 10 min at 4°C, 93 μL of supernatant was incubated, in a final volume of 100 μL , with 50 nmol of UDP-GlcNAc (Sigma-Aldrich) in the presence of 100 nmol of ATP and 1 mM CaCl_2 either for 4 h or overnight. The reaction mixture was centrifuged at 13,000 rpm for 10 min. The liquid was applied to a C18-RP SPE cartridge, whereby the flow through was collected and dried by vacuum. The dried sample was derivatized with AA (Anumula, 1994). Twenty-five percent (v/v) of the reaction mixture was then injected to an Aquasil RP column (0.32×150 mm; Thermo Scientific) using a 65 mM ammonium formate buffer, pH 3.0, as the aqueous solvent. After sample injection, which was performed with buffer only, a gradient from 3.8% to 25% acetonitrile (AcCN) was developed over 30 min followed by a gradient to 50% AcCN in 10 min. The flow rate of 5 $\mu\text{L min}^{-1}$ was maintained by an UltiMate 3000 LC system (Dionex). Eluting ManNAc-6-P (or GlcNAc-6-P) was detected by ESI-MS/MS with a fixed-parent mass of 421.1 D on a Q-TOF Ultima Global apparatus (Waters Micromass) operated in negative ion mode (capillary voltage = -3.1 kV, collision energy = 29%). GlcNAc-6-P (Sigma-Aldrich) was used as a standard.

In Vivo ManNAc-6-P Analysis

Seedlings (0.5 g) were homogenized with an Ultra-Turrax homogenizer in 1 mL of water. After adding 80 μL of concentrated acetic acid, the mixture was centrifuged at 20,000 rpm for 20 min. The supernatant was applied to a C18-RP SPE cartridge. The flow through was directly applied to a gel filtration column (1×75 cm) filled with Biogel P2 fine (Bio-Rad) and eluted with 20 mM ammonium acetate, pH 5.5, at a flow rate 1 mL min^{-1} . Fractions were analyzed for the presence of HexNAc-6-P by direct infusion into the ESI source (in 50% [v/v] AcCN and 0.1% [v/v] formic acid). Fragments of the parent mass of 300.0 D were measured with a Q-TOF Ultima Global apparatus operated in the negative ion mode.

In Vitro NANS Activity Assay

Seedlings (2 g) were homogenized with an Ultra-Turrax homogenizer in 1 mL of 50 mM Bicine hydrochloride, pH 8.0, in the presence of protease inhibitors (Complete Mini, EDTA free; Roche) and further centrifuged at 13,000 rpm for 10 min at 4°C. The assay was performed by incubating 83 μL of supernatant with 5 nmol of ManNAc-6-P and 5 nmol of PEP in the presence of 12.5 mM MnCl_2 in a final volume of 100 μL . After overnight incubation at 37°C, the reaction mixture was centrifuged and dried as above. The dried sample was derivatized with DMB (Sigma-Aldrich; Hara et al., 1987). The reaction mixture was analyzed by RP-HPLC with fluorescence detection as described (Zeleny et al., 2006). The eluate in the region of DMB-Neu5Ac was collected, concentrated, and subjected to RP-LC-ESI-MS/MS to acquire a fragment spectrum. Characteristic fragments from DMB-Neu5Ac were obtained in MS/MS mode from the parent ion with mass-to-charge ratio = 426.0 ($[\text{M}+\text{H}]^+$; Zeleny et al., 2006). For both HPLC and MS analyses, results were quantified by comparison with DMB-Neu5Ac standards.

In Vivo Neu5Ac Analysis

The first steps were performed as described for the ManNAc-6-P analysis using Arabidopsis leaves (0.5 g) from 6-week-old plants instead of seedlings. Fractions obtained after separation by gel filtration (see above) were brought to 50% (v/v) AcCN and 0.1% (v/v) formic acid and analyzed for the presence of Neu5Ac by direct infusion into the ESI source. Analytes were detected by MS/MS ($[\text{M}-\text{H}]^-$ = 308.1) in negative ion mode. Additionally, quantification was done by HPLC after DMB derivatization (see above).

In Vivo CMP-Neu5Ac Analysis

Leaves (0.1 g) from 6-week-old Arabidopsis plants were homogenized with an Ultra-Turrax homogenizer (in 100 μL of 1% NH_4 solution). The sample was centrifuged at 13,000 rpm for 10 min, and the supernatant was applied to a C18-RP SPE cartridge. The flow through was brought to 40% AcCN (v/v) and applied to a 10-mg HyperSep Hypercarb SPE cartridge (Thermo Scientific). The flow through was collected and freeze-dried, and equivalents of 1% were analyzed on a Hypercarb column (0.32×50 mm; Thermo Scientific) using a 65 mM ammonium formate buffer, pH 3.0, as the aqueous solvent. After sample injection, which was performed with buffer only, a gradient from 4% to 22% AcCN was developed in 30 min with a flow rate of 8 $\mu\text{L/min}$.

Analytes were detected with an ESI-Q-TOF Ultima Global apparatus in MS/MS mode with MS1 set on mass-to-charge ratio = 613.1, the mass of the $[\text{M}-\text{H}]^-$ ion of CMP-Neu5Ac.

N-Glycan Analysis

N-Glycans were extracted from 500 mg of leaves from 6-week-old transgenic Arabidopsis plants and analyzed by MALDI-TOF-MS as described previously (Strasser et al., 2004).

ACKNOWLEDGMENTS

We thank Pia Gatteringer and Karin Polacek (University of Natural Resources and Applied Life Sciences, Vienna) for excellent technical assistance.

Received February 7, 2008; accepted March 3, 2008; published March 7, 2008.

LITERATURE CITED

- Anumula KR (1994) Quantitative determination of monosaccharides in glycoproteins by high-performance liquid chromatography with highly sensitive fluorescence detection. *Anal Biochem* **220**: 275–283
- Bakker H, Bardor M, Molthoff JW, Gomord V, Elbers I, Stevens LH, Jordi W, Lommen A, Faye L, Lerouge P, et al (2001) Galactose-extended glycans of antibodies produced by transgenic plants. *Proc Natl Acad Sci USA* **98**: 2899–2904
- Cox KM, Sterling JD, Regan JT, Gasdaska JR, Frantz KK, Peele CG, Black A, Passmore D, Moldovan-Loomis C, Srinivasan M, et al (2006) Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat Biotechnol* **24**: 1591–1597
- Dirnberger D, Steinkellner H, Abdennebi L, Remy JJ, van de Wiel D (2001) Secretion of biologically active glycoforms of bovine follicle stimulating hormone in plants. *Eur J Biochem* **268**: 4570–4579
- Erbayraktar S, Grasso G, Sfacteria A, Xie QW, Coleman T, Kreilgaard M, Torup L, Sager T, Erbayraktar Z, Gokmen N, et al (2003) Asialoerythropoietin is a nonerythropoietic cytokine with broad neuroprotective activity in vivo. *Proc Natl Acad Sci USA* **100**: 6741–6746
- Fischer R, Stöger E, Schillberg S, Christou P, Twyman RM (2004) Plant-based production of biopharmaceuticals. *Curr Opin Plant Biol* **7**: 152–158
- Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol Biol* **25**: 989–994
- Hara S, Takemori Y, Yamaguchi M, Nakamura M, Ohkura Y (1987) Fluorometric high-performance liquid chromatography of N-acetyl- and N-glycolylneuraminic acids and its application to their microdetermination in human and animal sera, glycoproteins and glycolipids. *Anal Biochem* **164**: 138–145
- Hooker AD, Green NH, Baines AJ, Bull AT, Jenkins N, Strange PG, James DC (1999) Constraints on the transport and glycosylation of recombinant IFN-gamma in Chinese hamster ovary and insect cells. *Biotechnol Bioeng* **63**: 559–572
- Horstkorte R, Nohring S, Wiechens N, Schwarzkopf M, Danker K, Reutter W, Lucka L (1999) Tissue expression and amino acid sequence of murine UDP-N-acetylglucosamine-2-epimerase/N-acetylmannosamine kinase. *Eur J Biochem* **260**: 923–927
- Jach G, Binot E, Frings S, Luxa K, Schell J (2001) Use of red fluorescent protein from *Discosoma* sp. (DsRED) as a reporter for plant gene expression. *Plant J* **28**: 483–491
- Jiang H, Wang S, Dang L, Wang S, Chen H, Wu Y, Jiang X, Wu P (2005) A novel short-root gene encodes a glucosamine-6-phosphate acetyltransferase required for maintaining normal root cell shape in rice. *Plant Physiol* **138**: 232–242
- Koprivova A, Stemmer C, Altmann F, Hoffmann A, Kopriva S, Gorr G, Reski R, Decker EL (2004) Targeted knockouts of *Physcomitrella* lacking plant specific immunogenic N-glycans. *Plant Biotechnol J* **2**: 517–523
- Kornfeld S, Kornfeld R, Neufeld EF, O'Brien PJ (1964) The feedback control of sugar nucleotide biosynthesis in liver. *Proc Natl Acad Sci USA* **52**: 371–379
- Krapp S, Münster-Kühnel AK, Kaiser JT, Huber R, Tiralongo J, Gerardy-

- Schahn R, Jacob U (2003) The crystal structure of murine CMP-5-N-acetylneuraminic acid synthetase. *J Mol Biol* **334**: 625–637
- Lawrence SM, Huddleston KA, Pitts LR, Nguyen N, Lee YC, Vann WF, Coleman TA, Betenbaugh MJ (2000) Cloning and expression of the human N-acetylneuraminic acid phosphate synthase gene with 2-keto-3-deoxy-D-glycero-D-galactonononic acid biosynthetic ability. *J Biol Chem* **275**: 17869–17877
- Lawrence SM, Huddleston KA, Tomiya N, Nguyen N, Lee YC, Vann WF, Coleman TA, Betenbaugh MJ (2001) Cloning and expression of human sialic acid pathway genes to generate CMP-sialic acids in insect cells. *Glycoconj J* **18**: 205–213
- Maliekal P, Vertommen D, Delpierre G, Van Schaftingen E (2006) Identification of the sequence encoding N-acetylneuraminic acid-9-phosphate phosphatase. *Glycobiology* **16**: 165–172
- Misaki R, Fujiyama K, Seki T (2006) Expression of human CMP-N-acetylneuraminic acid synthetase and CMP-sialic acid transporter in tobacco suspension-cultured cell. *Biochem Biophys Res Commun* **339**: 1184–1189
- Paccalet T, Bardor M, Rihouey C, Delmas F, Chevalier C, D'Aoust MA, Faye L, Vézina L, Gomord V, Lerouge P (2007) Engineering of a sialic acid synthesis pathway in transgenic plants by expression of bacterial Neu5Ac-synthesizing enzymes. *Plant Biotechnol J* **5**: 16–25
- Palacpac NQ, Yoshida S, Sakai H, Kimura Y, Fujiyama K, Yoshida T, Sek T (1999) Stable expression of human beta1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *Proc Natl Acad Sci USA* **96**: 4692–4697
- Pels Rijcken WR, Hooghwinkel GJ, Ferwerda W (1990) Pyrimidine metabolism and sugar nucleotide synthesis in rat liver. *Biochem J* **266**: 777–783
- Saint-Jore-Dupas C, Faye L, Gomord V (2007) From planta to pharma with glycosylation in the toolbox. *Trends Biotechnol* **25**: 317–323
- Schähs M, Strasser R, Stadlmann J, Kunert R, Rademacher T, Steinkellner H (2007) Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. *Plant Biotechnol J* **5**: 657–663
- Schauer R (2000) Achievements and challenges of sialic acid research. *Glycoconj J* **17**: 485–499
- Seveno M, Bardor M, Paccalet T, Gomord V, Lerouge P, Faye L (2004) Glycoprotein sialylation in plants? *Nat Biotechnol* **22**: 1351–1352
- Shah MM, Fujiyama K, Flynn CR, Joshi L (2003) Sialylated endogenous glycoconjugates in plant cells. *Nat Biotechnol* **21**: 1470–1471
- Strasser R, Altmann F, Mach L, Glössl J, 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, Stadlmann J, Schähs M, Stiegler G, Quendler H, Mach L, Glössl J, Weterings K, Pabst M, 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* (in press)
- Strasser R, Stadlmann J, Svoboda B, Altmann F, Glössl J, Mach L (2005) Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. *Biochem J* **387**: 385–391
- Takashima S, Abe T, Yoshida S, Kawahigashi H, Saito T, Tsuji S, Tsujimoto M (2006) Analysis of sialyltransferase-like proteins from *Oryza sativa*. *J Biochem* **139**: 279–287
- Tietze E, Seppala R, Renlund M, Hopwood JJ, Harper GS, Thomas GH, Gahl WA (1989) Defective lysosomal egress of free sialic acid (N-acetylneuraminic acid) in fibroblasts of patients with infantile free sialic acid storage disease. *J Biol Chem* **264**: 15316–15322
- Vann WF, Tavarez JJ, Crowley J, Vimr E, Silver RP (1997) Purification and characterization of the *Escherichia coli* K1 neuB gene product N-acetylneuraminic acid synthetase. *Glycobiology* **7**: 697–701
- Varki A (2007) Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature* **446**: 1023–1029
- Viswanathan K, Narang S, Hinderlich S, Lee YC, Betenbaugh MJ (2005) Engineering intracellular CMP-sialic acid metabolism into insect cells and methods to enhance its generation. *Biochemistry* **44**: 7526–7534
- Wee EQ, Sherrier DJ, Prime TA, Dupree P (1998) Targeting of active sialyltransferase to the plant Golgi apparatus. *Plant Cell* **10**: 1759–1768
- Zeleny R, Kolarich D, Strasser R, Altmann F (2006) Sialic acid concentrations in plants are in the range of inadvertent contamination. *Planta* **5**: 1–6