

Identification and characterization of a strict and a promiscuous *N*-acetylglucosamine-1-P uridylyltransferase in *Arabidopsis*

Ting YANG^{*†}, Merritt ECHOLS[†], Andy MARTIN[†] and Maor BAR-PELED^{†‡1}

^{*}Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, U.S.A., [†]Complex Carbohydrate Research Center (CCRC), University of Georgia, Athens, GA 30602, U.S.A., and [‡]Department of Plant Biology, University of Georgia, Athens, GA 30602, U.S.A.

UDP-GlcNAc is an essential precursor for glycoprotein and glycolipid synthesis. In the present study, a functional nucleotidyltransferase gene from *Arabidopsis* encoding a 58.3 kDa GlcNAc1pUT-1 (*N*-acetylglucosamine-1-phosphate uridylyltransferase) was identified. In the forward reaction the enzyme catalyses the formation of UDP-*N*-acetylglucosamine and PP_i from the respective monosaccharide 1-phosphate and UTP. The enzyme can utilize the 4-epimer UDP-GalNAc as a substrate as well. The enzyme requires divalent ions (Mg²⁺ or Mn²⁺) for activity and is highly active between pH 6.5 and 8.0, and at 30–37°C. The apparent *K_m* values for the forward reaction were 337 μM (GlcNAc-1-P) and 295 μM (UTP) respectively. Another GlcNAc1pUT-2, which shares 86 % amino acid sequence identity with GlcNAc1pUT-1, was found to convert, in addition to GlcNAc-1-P and GalNAc-1-P, Glc-1-P into corresponding

UDP-sugars, suggesting that subtle changes in the UT family cause different substrate specificities. A three-dimensional protein structure model using the human AGX1 as template showed a conserved catalytic fold and helped identify key conserved motifs, despite the high sequence divergence. The identification of these strict and promiscuous gene products open a window to identify new roles of amino sugar metabolism in plants and specifically their role as signalling molecules. The ability of GlcNAc1pUT-2 to utilize three different substrates may provide further understanding as to why biological systems have plasticity.

Key words: GlcNAc1pUT-1, GlcNAc1pUT-2, *N*-acetylglucosamine-1-phosphate uridylyltransferase, UDP-GalNAc-PPase, UDP-GlcNAc-PPase, UDP-2-acetamido-2-deoxy-α-D-glucopyranose.

INTRODUCTION

GlcNAc (*N*-acetylglucosamine) is a major sugar residue found in different types of glycans across species. In the ER (endoplasmic reticulum), the attachment of GlcNAc is required to initiate *N*-glycan processing for the synthesis of glycoproteins [1], and separately to initiate the synthesis of the core for GPI (glycosylphosphatidylinositol)-anchor-linked protein synthesis [2], and the synthesis of glycolipids [3]. In certain bacteria, the GlcNAc residue is the major component of the peptidoglycan in periplasm and in the lipopolysaccharide of lipid A (endotoxin) [4]. Both in plants and mammals, the specific attachment and removal of the GlcNAc residue from cytosolic and nuclear proteins play a signalling function for the modification of regulatory proteins. Such post-translational modification by either an OGT (*O*-GlcNAc transferase) or protein kinase [5] is very common. Specific OGTs catalyse the transfer of the GlcNAc residue from UDP-GlcNAc to a serine or threonine residue within a substrate protein [6], and the mutation of plant OGT is crucial for gamete and seed development [7–9]. Lastly in plants, several proteins modified with terminal GlcNAc are present at the nuclear rim and at the nuclear pore complex of tobacco, the role of these sugar modifications at the pore remain elusive [10].

Very little is known about the 4-epimer of GlcNAc, GalNAc (*N*-acetylgalactosamine) in plants. In animals GalNAc is largely found in the core structure of *O*-linked glycan. GalNAc residues are reported, however, in the polysaccharide of green algae [11]. Certain plant UDP-D-glucose-4-epimerase isoforms can *in vitro* interconvert UDP-GlcNAc and UDP-GalNAc [12]. The metabolite UDP-GalNAc was isolated from dahlia and squash

plants [13], and more recently from *Arabidopsis* suspension culture and from the endosperm of fennugreek plants [14]. How these sugar residues are made still remain unknown, as the genes encoding these activities have not been identified.

The biosynthesis of UDP-GlcNAc differs in different organism. In prokaryotes the transamination of Fru-6-P (fructose-6-phosphate) and glutamine into GlcN-6-P (glucosamine-6-P) by a glutamine:Fru-6-P amidotransferase initiates the pathway. GlcN-6-P is subsequently phosphoisomerized into GlcN-1-P by GlcN phosphate mutase. A single bifunctional enzyme glmU, comprising an acetyltransferase domain, transfers acetate from acetyl-CoA to GlcN-1-P, producing GlcNAc-1-P, and the nucleotidyltransferase domain of glmU uridylylates GlcNAc-1-P with UTP to form UDP-GlcNAc [15] (Figure 1A). Unlike in prokaryotes, in eukaryotes (Figure 1B), GlcN-6-P is first *N*-acetylated by an acetylase to GlcNAc-6-P and then converted into GlcNAc-1-P by phospho-GlcNAc mutase. The final step, first identified in 1954 by Smith and Mills [16] is carried out by a specific nucleotidyltransferase, GlcNAc1pUT (UTP:*N*-acetylglucosamine-1-P uridylyltransferase) that in the forward reaction catalyses the formation of UDP-GlcNAc and PP_i from GlcNAc-1-P and UTP. This type of uridylation activity belongs to a large family of enzymes that also catalyse the formation of other specific NDP-sugars such as UDP-Glc, UDP-GalA, UDP-GlcA, ADP-Glc, GDP-Man and many more. Common for this family is the metal-dependent transfer of the nucleotidyl moiety and the specific phosphate acceptors. This catalysis involves the cleavage of the pyrophosphate linkage (P-α:P-β) in the nucleotide substrate [17]. In the reverse reaction, PP_i and NDP-sugar are converted into sugar-1-P and NTP [15,18]. The sugar-1-P specificities of

Abbreviations used: DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; GlcNAc1pUT, *N*-acetylglucosamine-1-phosphate uridylyltransferase; LB, Luria-Bertani; NB, nucleotide binding; OGT, *O*-GlcNAc transferase; T-DNA, transfer DNA; UB, uridine binding.

¹ To whom correspondence should be addressed (email peled@ccrc.uga.edu).

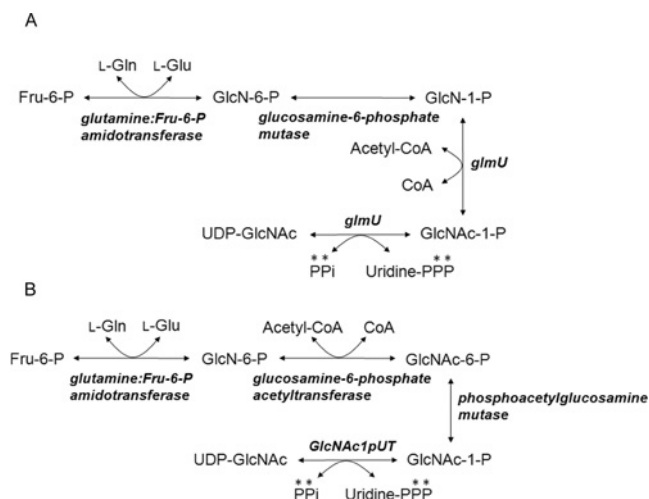


Figure 1 UDP-GlcNAc biosynthesis pathways in prokaryotes and eukaryotes

(A) UDP-GlcNAc biosynthesis pathway in prokaryotes. (B) UDP-GlcNAc biosynthesis pathway in eukaryotes. The enzymes and candidate genes are shown in italics.

GlcNAc1pUT across organisms are unresolved as some enzymes may use Glc-1-P and GlcNAc-1-P as substrates [15], whereas others may use GalNAc-1-P and GlcNAc-1-P [19]. The gene encoding the plant homologous activity has not been identified owing to low sequence identity, and the specificities of the plant enzyme are not clear.

In the present study we first report the identification and characterization of two different UDP-GlcNAc nucleotidyltransferases in *Arabidopsis* that convert GlcNAc-1-P and GalNAc-1-P into their corresponding UDP-sugars. The *Arabidopsis* GlcNAc1pUT provides an opportunity to understand how pools of UDP-GlcNAc and perhaps UDP-GalNAc are controlled and affect signalling events in the cytosol and nucleus.

EXPERIMENTAL

cDNA cloning of *Arabidopsis* GlcNAc1pUT

Total RNA from young rosette leaves of *Arabidopsis thaliana* Columbia ecotype was extracted using TRIzol[®] reagent (Invitrogen) and was used as a template for the reverse transcription reaction using 0.2 μ M Oligo-dT₁₇V primer [(V = dG, dA or dC)], dNTPs and 200 units of SuperScript II-reverse transcriptase (Invitrogen). BLAST analyses with human GlcNAc1pUT revealed several potential homologous proteins in plants and subsequently primers for cloning these putative genes were designed. The coding sequence of one gene corresponding to the At1g31070 loci (GlcNAc1pUT-1) was amplified by PCR using 1 unit of high-fidelity proof-reading Platinum DNA polymerase (Invitrogen), and 0.2 μ M of each forward and reverse primers: 5'-CACCATGGTAGAACCGTCGATGGAGAGAG-3' and 5'-GGATCCAGGGAAATTCACAAGGTGCATG-3' using cDNA of *Arabidopsis* as a template (the primer sequence that is underlined represents added bases to facilitate cloning). The PCR product was cloned to generate plasmid pCR4-topoTA:at1g31070.11#6; and DNA was sequenced (GenBank[®] accession number GU937393). The NcoI-BamHI fragment containing the full-length GlcNAc1pUT-1 gene without the stop codon was sub-cloned into a pET28b *Escherichia coli* expression vector so that upon expression the recombinant enzyme will have a

His₆ extension at its C-terminus. Similarly pET28c:at2g35020#13 (GlcNAc1pUT-2) was generated with a His₆ extension at the N-terminus.

Protein expression and purification

E. coli cells, BL21(de3)plysS-derived strain, harbouring the GlcNAc1pUT expression construct or an empty vector control, were cultured for 16 h at 37 °C in LB (Luria-Bertani) medium (20 ml) supplemented with kanamycin (50 μ g/ml) and chloramphenicol (34 μ g/ml). A portion (8 ml) of the cultured cells was transferred into fresh LB liquid medium (250 ml) supplemented with the same antibiotics, and the cells then grown at 37 °C at 250 rev./min until the cell density reached $A_{600} = 0.6$. The cultures were then transferred to 18 °C and gene expression was induced by the addition of IPTG (isopropyl β -D-thiogalactoside) to a final concentration of 0.5 mM. After 24 h growth while shaking (250 rev./min), the cells were centrifuged (6000 *g* for 10 min at 4 °C), resuspended in lysis buffer [10 ml of 50 mM Tris/HCl (pH 7.6), containing 10 % (v/v) glycerol, 50 mM NaCl, 0.5 mM EDTA, 1 mM DTT (dithiothreitol) and 0.5 mM PMSF] and lysed in an ice bath by 24 sonication cycles each (10 s pulse; 20 s rest) using a Misonix S-4000 equipped with 1/8" microtip probe. The lysed cells were centrifuged at 4 °C for 30 min at 20000 *g*, and the supernatant (termed s20) was recovered and kept at –20 °C. His-tagged proteins were purified on a column (10 mm internal diameter $\times$ 150 mm length) containing Ni-Sepharose (2 ml, Qiagen) that was previously equilibrated with a buffer [50 mM sodium phosphate (pH 7.6) and 0.3 M NaCl]. The bound His₆-tagged protein was eluted with the same buffer containing increasing concentrations of imidazole. The fractions containing GlcNAc1pUT activities were stored in aliquots at –80 °C. The concentration of protein was determined using BSA as a standard. The molecular mass of the recombinant protein was estimated by size-exclusion chromatography using a Superdex-200 column (1 cm internal diameter $\times$ 30 cm length; GE Healthcare) with buffer comprising 0.1 M sodium phosphate (pH 7.6) and 0.1 M NaCl. Separate solutions (0.5 ml) of recombinant GlcNAc1pUT or a mixture of standard proteins [10 mg each of alcohol dehydrogenase (157 kDa), BSA (66 kDa), carbonic anhydrase (29 kDa) and cytochrome C (12.4 kDa)] were separately injected via a Waters 626 LC HPLC system equipped with a photodiode array detector (PDA 996) and a Waters Millennium32 workstation, at 0.5 ml/min. The eluant was monitored at A_{280} and the fractions were collected every 15 s. Fractions containing enzyme activity were analysed and kept at –80 °C. The native molecular mass of recombinant GlcNAc1pUT was determined based on the elution time and the calibration curve of the standard proteins described above.

Enzyme assays

Unless otherwise mentioned, the typical forward HPLC-based reactions for the formation of UDP-GlcNAc were carried out in a final volume of 50 μ l and consisted of 1 mM GlcNAc-1-P, 1 mM UTP, 5 mM MgCl₂, 100 mM Tris/HCl (pH 7.6), 1 unit of yeast inorganic pyrophosphatase (Sigma) and recombinant GlcNAc1pUT (420 ng). After 10 min incubation at 37 °C, reactions were terminated (1 min at 100 °C). Chloroform (50 μ l) was added and after vortexing (30 s) and centrifugation [14000 *g* for 5 min, at room temperature (25 °C)], the entire upper aqueous phase was collected and subjected to chromatography on anion-exchange chromatography using a TSK-DEAD-5-PW column (7.5 mm internal diameter $\times$ 75 mm length; Bio-Rad) and an ammonium formate HPLC gradient system [20]. Nucleotides and

nucleotide sugars were detected by their UV absorbance using a photodiode array detector that was connected to the HPLC system. The maximum absorbance for uridine nucleotides and UDP-sugars were 261.8 nm in ammonium formate. The peak area of analytes was determined based on standard calibration curves. HPLC-based reverse reactions were carried out in a similar manner and included 1 mM PP_i , 1 mM UDP-GlcNAc, 5 mM MgCl_2 , 100 mM Tris/HCl (pH 7.6) and 420 ng of GlcNAc1pUT.

Real-time ^1H -NMR analysis of GlcNAc1pUT-1 activities

The forward reaction in the final volumes of 180 μl in a mixture of $^2\text{H}_2\text{O}/\text{H}_2\text{O}$ of 8:1 (v/v) consisted of 0.1 M sodium phosphate (pH 7.6), 5 mM MgCl_2 , 2 mM UTP, 2 mM GlcNAc-1-P and enzymes: 1.05 μg of recombinant GlcNAc1pUT-1 and 1 unit of yeast inorganic pyrophosphatase in H_2O buffer. For the reverse PPase reaction, 2.1 μg of recombinant GlcNAc1pUT-1 was incubated with 2 mM PP_i and 2 mM UDP-GalNAc (or UDP-GlcNAc). Immediately upon addition of the enzyme, the reaction mixture was transferred into a 3 mm NMR tube. Real-time ^1H -NMR spectra were obtained using a Varian Inova 600 MHz spectrometer equipped with a cryogenic probe. Data acquisition was not started until approx. 2 min after the addition of enzyme to the reaction mixture due to spectrometer set-up requirements (shimming). Sequential one-dimensional proton spectra were acquired over the course of the enzymatic reaction. All spectra were referenced to the water resonance at 4.765 p.p.m. downfield of DSS (2,2-dimethyl-2-silapentane-5-sulfonate). Processing of the data as covariance matrices was performed with MatLab (The Mathworks).

Enzyme properties and inhibition assays

The forward nucleotidyltransferase activity of GlcNAc1pUT-1 was measured with various buffers, at different temperatures, with different ions or with different potential inhibitors. For the optimal pH experiments, 420 ng of recombinant enzyme was first mixed with 5 mM MgCl_2 and 100 mM of each individual buffer (Tris/HCl, phosphate, Mes, Mops or Hepes). The optimal pH assays were initiated after the addition of specific GlcNAc-1-P and UTP. Inhibitor assays were performed under the standard assay conditions except for the addition of various additives (sugars, nucleotides and antibiotics) to the reaction buffer. These assays were incubated for 10 min at 37°C, and were subsequently terminated by heat (1 min at 100°C). The amount of UDP-sugar formed was calculated from a calibration curve of HPLC UV spectra of standards. For the experiments aimed at defining the optimal temperature, assays were performed under standard assay conditions except that reactions were incubated at different temperatures for 10 min. Subsequently, the activities were terminated (100°C). For the experiments aimed at determining whether GlcNAc1pUT-1 required metals, assays were performed with UTP and GlcNAc-1-P with a variety of ions. After 10 min at 37°C incubation, the activity was terminated by heat. The amount of UDP-GlcNAc formed was calculated from HPLC UV spectra of standards.

For the experiments aimed at determining the ability of GlcNAc1pUT to utilize other sugar-1-Ps, the standard assay was modified by substituting the GlcNAc-1-P with different sugar-1-Ps (for example, GlcA-1-P, Gal-1-P, etc.), and reactions were incubated for 60 min at 37°C.

Kinetics

The forward GlcNAc1pUT catalytic activity was determined at 37°C for 4 min using 0.1 M Tris/HCl (pH 7.6), 5 mM MgCl_2 ,

1 mM GlcNAc-1-P, recombinant GlcNAc1pUT-1 (0.9 pmol, 50 ng) or recombinant GlcNAc1pUT-2 (0.6 pmol, 38 ng), and variable concentrations of UTP (20 μM to 4 mM) or with fixed UTP (1 mM) and variable concentrations of GlcNAc-1-P (20 μM to 4 mM). The forward kinetic assays included 2 units of yeast inorganic pyrophosphatase to deplete PP_i . Kinetics for the reverse reactions were performed in the same conditions as above, with a fixed concentration of PP_i (1 mM) and variable concentrations of UDP-GlcNAc (20 μM to 4 mM), and recombinant GlcNAc1pUT-1 (0.9 pmol, 50 ng) or recombinant GlcNAc1pUT-2 (0.3 pmol, 19 ng). In a separate series of reverse-reaction experiments, assays were performed at 37°C for 10 min with a fixed concentration of PP_i (1 mM) and variable concentrations of UDP-GalNAc (200 μM to 10 mM), and recombinant GlcNAc1pUT-1 (1.8 pmol, 100 ng) or recombinant GlcNAc1pUT-2 (0.6 pmol, 38 ng). The study of the kinetics of GlcNAc1pUT-2 for Glc-1-P was essentially carried out as described above. Enzyme velocity data of the amount (μM) of UDP-sugar produced per s, as a function of substrate concentrations were plotted. The Solver tool (Excel version 11.5 program) was used to generate the best-fit curve calculated by non-linear regression analyses, and for the calculation of V_{max} and apparent K_m . The amount of enzyme used and the incubation time were adjusted so that the reactions were in the linear range.

Structural prediction for GlcNAc1pUT

Three-dimensional structure models for GlcNAc1pUT were predicted using both the PHYRE (Protein Homology/analogy Recognition Engine) (<http://www.sbg.bio.ic.ac.uk/~phyre/>) web-server [21] and the SWISS-MODEL comparative protein modelling server (<http://swissmodel.expasy.org/>). Structural models were superimposed with known crystal structures from human AGX1 (PDB number 1JY1) [19]. All models and protein structures were visualized using PyMOL (DeLano Scientific; <http://www.pymol.org>). Amino acid alignments of the GlcNAc1pUT sequence to fold library protein were performed using the PHYRE program.

RESULTS

Identification, cloning and characterization of *Arabidopsis* GlcNAc1pUT-1

The enzyme nomenclature for NDP-sugar pyrophosphorylase (abbreviated to PPase) is the same as for sugar-1-P nucleotidyltransferases. To identify a potential functional GlcNAc-1-P nucleotidyltransferase homologue in plants, we compared the human AGX1 protein sequence (GenBank® accession number NP_003106) with the NR sequence database. BLAST analyses of homologous proteins from different species revealed that *Arabidopsis* genome loci At1g31070 shares overall a low sequence identity with functional UDP-Glc PPase and UDP-sugar PPase (24% and 26% respectively), and approx. 34% and 40% amino acid sequence identity with human and yeast functional UDP-GlcNAc PPase respectively [15,19]. Interestingly, the functional bacterial glmU (UDP-GlcNAc PPase) shares sequence similarity with plant GDP-Man PPase (At2g39770), rather than At1g31070. This suggests that sequence alone is not sufficient to predict function, nevertheless after amino acid sequence alignment of these homologous proteins with At1g31070 (GlcNAc1pUT-1) we have identified two consensus motifs (Supplementary Figure S1 at <http://www.BiochemJ.org/bj/430/bj4300275add.htm>). The N-terminal region of GlcNAc1pUT-1 comprising a putative NB (nucleotide binding) motif

LXGG(L/Q)G(E/T)(R/T)(L/M)GX₃(I/P)K, with a predicted coil structure starting at amino acid 134 and a PXGXG motif probably involved in UB (uridine binding) starting at amino acid 250. Phylogenetic analysis indicated that the UDP-GlcNAc PPase-like proteins from different species are distinguished from other nucleotidyltransferases involved in UDP-Glc and UDP-sugar synthesis (results not shown). In addition to At1g31070, another gene, At2g35020, with 86 % identity with UDP-GlcNAc PPase was also identified (GlcNAc1pUT-2). To determine, however, whether the *Arabidopsis* GlcNAc1pUT encode such activity, despite the low sequence identity with the human, yeast and bacterial enzymes, and more specifically to establish whether the enzymes have a different or similar range of sugar-1-Ps and NTP specificity, the genes were cloned, and the recombinant proteins expressed in *E. coli* were analysed.

A highly expressed protein band (58 kDa) was detected after SDS/PAGE analysis of *E. coli* cells expressing GlcNAc1pUT-1 (Figure 2A, lane 2; marked with an arrow) when compared with control cells expressing empty vector (Figure 2A, lane 3). The apparent mass of the column-purified enzyme as determined by SDS/PAGE analysis (Figure 2A, lane 4) is in agreement with the calculated mass of the translated gene product fused at the C-terminal portion to His₆. Preliminary HPLC-based forward assays (Figure 2B, panel 3) demonstrated that the recombinant GlcNAc1pUT-1 converts GlcNAc-1-P and UTP in the presence of Mg²⁺ into UDP-GlcNAc. The enzymatic product eluted at 11.7 min (Figure 2B, panel 3) had the same retention time as standard UDP-GlcNAc, whereas control cells expressing empty vector had no detectable activity (Figure 2B, panel 4). In the reverse assay, recombinant enzyme converts UDP-GlcNAc and PP_i into UTP (Figure 2B, panel 5), whereas the control had no activity (Figure 2B, panel 6). The HPLC peak marked #1 (Figure 2B) was collected from the column, and its structure and identity were confirmed by ¹H-NMR spectroscopy (Supplementary Figure S2 at <http://www.BiochemJ.org/bj/430/bj4300275add.htm>). The chemical-shift assignments for UDP-GlcNAc are summarized in Table 1. The diagnostic *J*_{1,2} coupling value of 3 Hz and *J*_{2,3}, *J*_{3,4}, *J*_{4,5} values of 10, 9.7, 9.7 Hz respectively indicates α-glucopyranose configuration. The linkage of the anomeric GlcNAc residue with the phosphate is given by the coupling constant values 7 and 3 Hz for *J*_{P-β, H1} and *J*_{P-β, H2} respectively, and also supported by the chemical shift of H1 (5.50 p.p.m.). In addition, the NAc-C=O resonance was assigned 2 p.p.m. and the H2 chemical shift (3.98 p.p.m.) is consistent with the diagnostic 2-acetamido-2-deoxy group. Collectively, these results provide unambiguous evidence that the enzyme product of GlcNAc1pUT-1 is UDP-α-D-GlcNAc.

Interestingly, the enzyme can also utilize the 4-epimer UDP-GalNAc as a substrate (Figure 2B, panel 7). The forward reaction with GalNAc-1-P could not be performed, as this amino-sugar-1-P is commercially unavailable. Based on these analyses we propose that GlcNAc1pUT-1 is a UT or depending on other enzyme nomenclature a PPase, and, like the human AGX1 enzyme, has the same specific activities with UDP-GlcNAc and UDP-GalNAc as substrates.

Characterization and properties of GlcNAc1pUT-1

GlcNAc1pUT-1 requires divalent cations such as Mg²⁺ or Mn²⁺ (Table 2), and the activity, as expected, is abolished in the presence of EDTA. However, cations such as Ca²⁺, for example, cannot substitute for magnesium. The recombinant GlcNAc1pUT-1 catalyses the conversion of GlcNAc-1-P into UDP-GlcNAc over a wide range of pH values (3.3–9.0, see Supplementary Figure S3 at <http://www.BiochemJ.org/bj/430/bj4300275add.htm>), with

maximum activity observed at pH 7.6–8.0 in either Tris/HCl, or at pH 7.6 in phosphate buffer. The enzyme is also active when reactions were performed in Hepes, Mops or Mes buffers (results not shown) at a similar range of neutral pH. GlcNAc1pUT-1 is also active over a broad range of temperatures (0–65 °C), but has maximum activity at 30–37 °C (Table 3).

We next investigated the NTP specificity of GlcNAc1pUT-1. Neither CTP, GTP, ITP nor ATP is the substrate for the recombinant enzyme when using GlcNAc-1-P as a substrate. Several commercially available sugar-1-Ps were tested as substrates for GlcNAc1pUT-1 with UTP. The reaction carried out with Glc-1-P, GlcA-1-P, Xyl-1-P, Gal-1-P, GalA-1-P, Fuc-1-P, Man-1-P, GalN-1-P, Fru-1-P, Man-1-P or Glc-6-P as substrates failed to show any conversion, even when assays were expanded for a longer incubation time. To determine whether GlcNAc1pUT-1 may recognize other NTPs we performed our standard assays in the presence of competing nucleotides (ATP, CTP, GTP and ITP at 0.5 mM each). In all cases GlcNAc-1-P was readily uridylated, suggesting that an enzyme other than UTP cannot recognize other NTPs. These experiments demonstrate the enzyme preferences for GlcNAc-1-P and GalNAc-1-P as its substrates, along with UTP. To determine whether the enzyme recognizes and binds NDPs such as UDP or ADP, prior to the standard assay, the enzyme was incubated with NDPs. These NDPs as well as other nucleotides tested (e.g. NMP, NAD and NADH) had no effect on GlcNAc1pUT-1 activity. In contrast, the forward enzyme activity was reduced by 56 % in the presence of 0.5 mM PP_i (Supplementary Table S1 at <http://www.BiochemJ.org/bj/430/bj4300275add.htm>). Many antimicrobial reagents are aminoglycoside derivatives. To determine whether the GlcNAc1pUT-1 activity is affected by these aminoglycosides, the enzyme was first incubated with individual antibiotics (gentamycin, kanamycin, hygromycin and streptomycin) and subsequently the standard assay was performed. Interestingly, the enzyme activity was dramatically reduced by 36 % and 74 % in the presence of 0.5 mg/ml hygromycin and streptomycin respectively (Supplementary Table S1), whereas on the other hand gentamycin and kanamycin had no effect on plant GlcNAc1pUT-1 activity.

Real-time ¹H-NMR analysis of GlcNAc1pUT-1

To monitor the dynamics of the enzymatic reaction and the substrate preference of GlcNAc1pUT-1, we used real-time ¹H-NMR spectroscopy (Figure 3). These assays were carried out in phosphate buffer to avoid the proton signals from the Tris reagent. As shown in the time-dependent enzymatic progression, conversion of the GlcNAc-1-P (5.34 p.p.m.) into UDP-α-D-GlcNAc (5.50 p.p.m.) is observed (Figure 3A). The NMR-based assay also provides unambiguous evidence that the enzyme can also utilize UDP-GalNAc in the presence of PP_i as a substrate [see conversion of the UDP-α-D-GalNAc (5.54 p.p.m.) into GalNAc-1-P (5.37 p.p.m.)] (Figure 3B).

GlcNAc1pUT-2 is promiscuous and converts GlcNAc-1-P, GalNAc-1-P and Glc-1-P into the corresponding nucleotide sugar

Similar to GlcNAc1pUT-1, a highly expressed protein band (60 kDa) was detected after SDS/PAGE analysis of *E. coli* cells expressing GlcNAc1pUT-2 (Figure 4A, lane 2; marked with an arrow) when compared with control cells expressing empty vector (Figure 4A, lane 3). The apparent mass of the column-purified enzyme as determined by SDS/PAGE analysis (Figure 4A, lane 4) is in agreement with the calculated mass of the translated gene product fused at the N-terminal portion to His₆. HPLC-based forward assays (Figure 4B, panel 3) demonstrated that

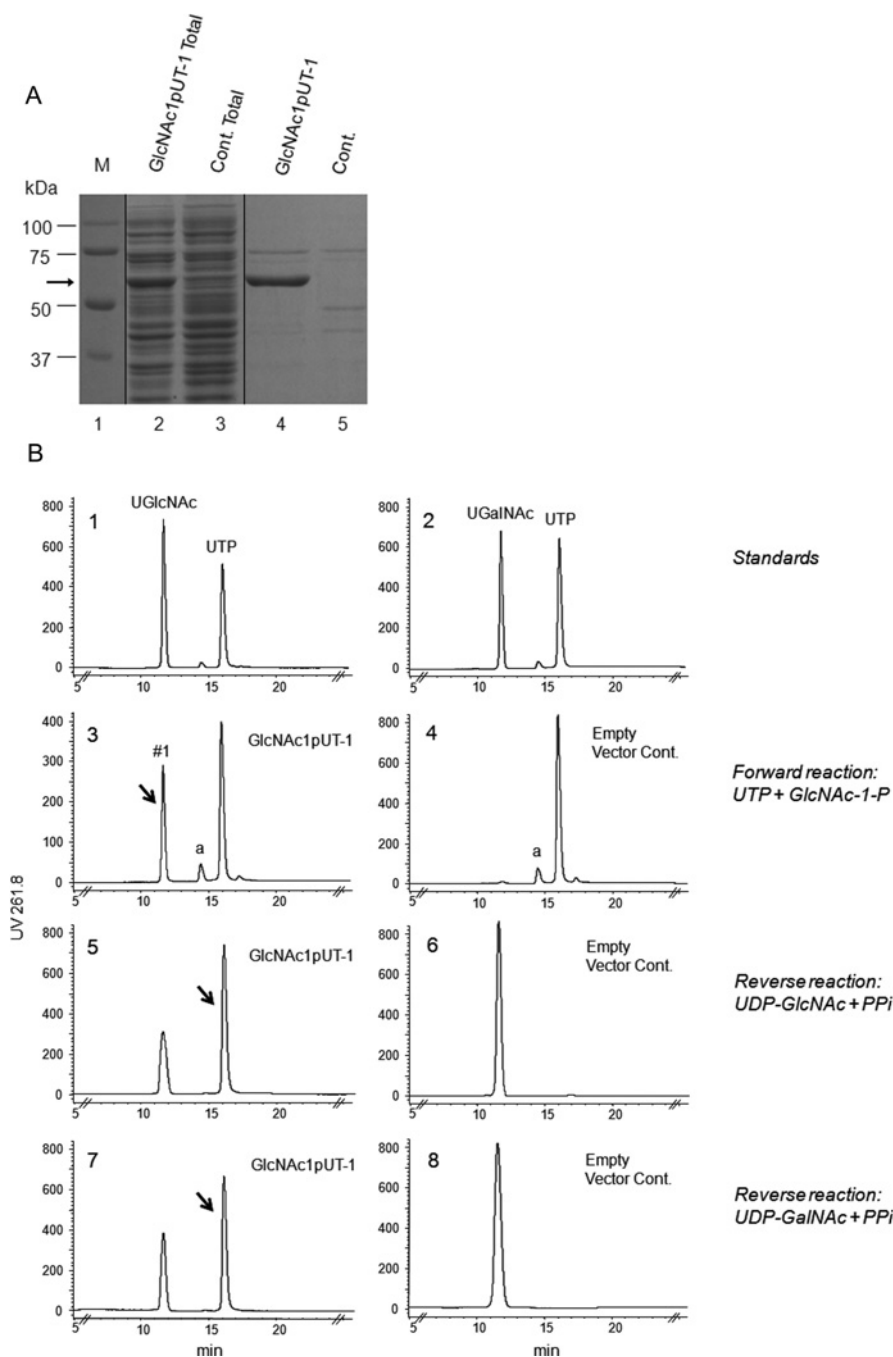


Figure 2 Recombinant GlcNAc1pUT-1 protein purification and HPLC-based assays

(A) SDS/PAGE of total soluble protein isolated from *E. coli* cells expressing recombinant GlcNAc1pUT-1 (lane 2), control empty vector (lane 3), Ni-column-purified recombinant GlcNAc1pUT-1 (lane 4) and control empty vector (lane 5). Molecular mass standards are in lane 1. (B) HPLC-based assays of GlcNAc1pUT-1. Forward assays included UTP and GlcNAc-1-P with either recombinant GlcNAc1pUT-1 (panel 3) or with protein extracted from cells expressing empty vector (panel 4). Reverse assays included PP_i and UDP-GlcNAc or UDP-GalNAc with either GlcNAc1pUT-1 (panels 5 and 7) or control empty vector protein (panels 6 and 8). Panel 1 shows chromatography of UDP-GlcNAc and UTP standards; panel 2 shows UDP-GalNAc and UTP standards. The HPLC peaks in the panels, based on retention time, are UDP-GlcNAc (11.7 min, marked as #1), UDP-GalNAc (11.7 min) and UTP (16.1 min). The minor peak marked as 'a' at 14.5 min is UDP contamination from the UTP reagent. The peaks labelled with arrows represent the nucleotide-based enzymatic reaction products. PP_i and GlcNAc-1-P are not UV visible.

the recombinant GlcNAc1pUT-2 converts GlcNAc-1-P and UTP in the presence of Mg²⁺ into UDP-GlcNAc. The enzymatic product eluted at 11.7 min (Figure 4B, panel 3) and had the same retention time as standard UDP-GlcNAc, whereas control cells expressing empty vector had no detectable activity (Figure 4B, panel 4). In the reverse assay, recombinant enzyme converts

UDP-GalNAc and PP_i into UTP (Figure 4B, panel 5), whereas the control had no activity (Figure 4B, panel 6). Interestingly, and unlike GlcNAc1pUT-1, recombinant GlcNAc1pUT-2 can also utilize Glc-1-P as a substrate, converting Glc-1-P and UTP in the presence of Mg²⁺ into UDP-Glc (Figure 4B, panel 7). The HPLC-peak marked #2 (Figure 4B) was collected from the

Table 1 Proton chemical shifts and coupling constants of UDP-GlcNAc formed from GlcNAc-1-P by recombinant GlcNAc1pUT

GlcNAc1pUT enzymatic product was purified from the HPLC column, essentially as described in Figure 2. The product peak was collected, freeze-dried, resuspended in ²H₂O and analysed by proton NMR. Chemical shifts are in p.p.m. relative to the internal DSS signal set at 0 p.p.m. GlcNAc proton–proton coupling constants in Hz are indicated, as well as the *J*_{1,P} and *J*_{2,P} coupling values between phosphate and the H1 and H2 protons of GlcNAc.

Proton	H1	H2	H3	H4	H5	H6, 6
GlcNAc						
Chemical shift, δ (p.p.m.)	5.50	3.98	3.82	3.54	3.92	3.86, 3.79
<i>J</i> coupling constants (Hz)	<i>J</i> _{1,P} 7, <i>J</i> _{1,2} 3	<i>J</i> _{2,3} 10, <i>J</i> _{2,P} 3	<i>J</i> _{3,4} 9.7	<i>J</i> _{4,5} 9.7	<i>J</i> _{5,6} 3.5	<i>J</i> _{6,6} 12
Rib						
Chemical shift, δ (p.p.m.)	5.97	4.37	4.36	4.28	4.24, 4.18	-
<i>J</i> coupling constants (Hz)	<i>J</i> _{1',2'} 4.5	-	-	-	-	-
Uracil						
Chemical shift, δ (p.p.m.)	-	-	-	-	5.96	7.95
<i>J</i> coupling constants (Hz)	-	-	-	-	<i>J</i> _{5',6'} 8.2	-

Table 2 GlcNAc1pUT-1 requires metal for activity

GlcNAc1pUT-1 was mixed with additive (metal, EDTA or water control) for 10 min on ice. Subsequently, UTP and GlcNAc-1-P were added and assays were carried out under standard conditions. Each value is the mean of duplicate reactions, and the values varied by no more than ±5 %.

Additive (5 mM)	Relative GlcNAc1pUT activity, forward reaction (%)
MgCl ₂	100
MnCl ₂	126
CaCl ₂	8
ZnSO ₄	9
CuCl ₂	8
EDTA	11
Water	8

Table 3 The effect of temperature on GlcNAc1pUT-1 activity

The enzymatic reactions were performed under standard conditions except for the reaction temperature. Each value is the mean of duplicate reactions, and the values varied by no more than ±5 %.

Temperature (°C)	Relative GlcNAc1pUT activity (%)
4	29
25	69
30	100
37	100
42	60
55	24
65	15

column, and its structure and identity was confirmed as UDP-α-D-glucose by ¹H-NMR spectroscopy (Supplementary Figure S4 at <http://www.BiochemJ.org/bj/430/bj4300275add.htm>).

Determination of kinetic parameters of GlcNAc1pUT

Kinetic analyses of the two enzymes are summarized in Table 4. The data from experiments with GlcNAc-1-P as the variable substrate and fixed concentration of UTP (or vice versa) fit well to the Michaelis–Menten model. For GlcNAc1pUT-1, the apparent *K*_m values for the forward reaction were 337 μM (GlcNAc-1-P) and 295 μM (UTP), with *V*_{max} values of 0.69 μM · s⁻¹, and *k*_{cat}/*K*_m (μM⁻¹ · s⁻¹) values were 0.11 (GlcNAc-1-P) and 0.12 (UTP).

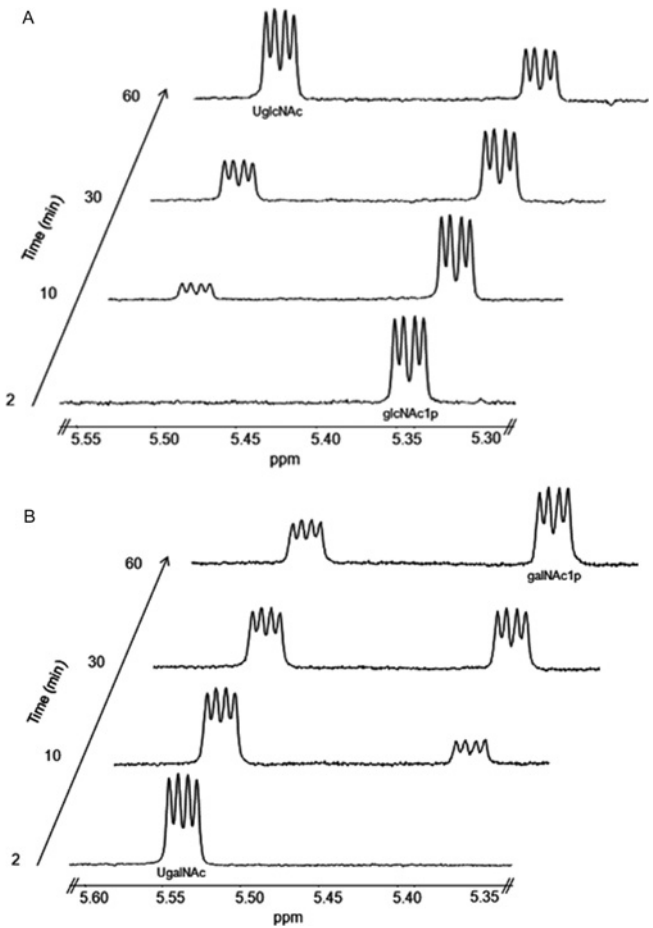


Figure 3 Real-time ¹H-NMR-based GlcNAc1pUT-1 assays

(A) GlcNAc1pUT-1 forward activity. The protein was mixed with 2 mM GlcNAc-1-P, phosphate buffer and 2 mM UTP. Approx. 2 min after enzyme addition and NMR shimming, data were collected. Progressions of enzyme activity covering the anomeric region of the proton NMR spectrum are shown. The ‘peak’ shape for sugar-1-P and UDP-sugar has a quadruple form: the chemical shift for GlcNAc-1-P is 5.54 p.p.m. and 5.50 p.p.m. for UDP-GlcNAc. (B) GlcNAc1pUT-1 reverse activity. The protein was mixed with 2 mM UDP-GalNAc, phosphate buffer and 2 mM PP_i. Approx. 2 min after enzyme addition and NMR shimming, data were collected. Progressions of enzyme activity covering the anomeric region of the proton NMR spectrum are shown. The chemical shifts are: GalNAc-1-P at 5.37 p.p.m. and UDP-GalNAc at 5.54 p.p.m.

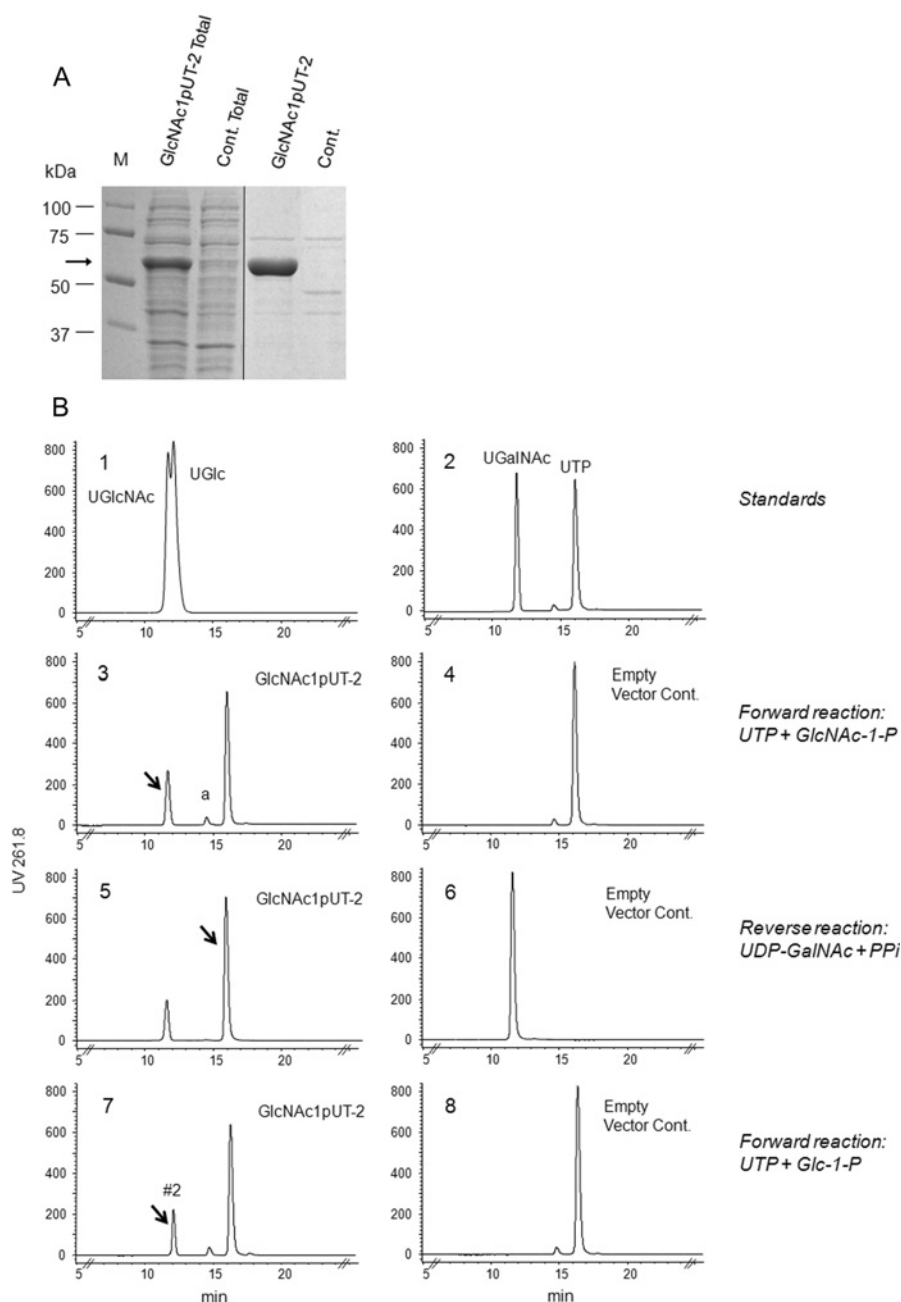


Figure 4 Recombinant GlcNAc1pUT-2 protein purification and HPLC-based assays

(A) SDS/PAGE of total soluble protein isolated from *E. coli* cells expressing recombinant GlcNAc1pUT-2 (lane 2), control empty vector (lane 3), Ni-column-purified recombinant GlcNAc1pUT-2 (lane 4) and control empty vector (lane 5). (B) HPLC-based assays of GlcNAc1pUT-2. Forward assays included UTP and GlcNAc-1-P or Glc-1-P with either recombinant GlcNAc1pUT-2 (panels 3 and 7) or with protein extracted from cells expressing empty vector (panels 4 and 8). Reverse assays included PP_i and UDP-GalNAc with either GlcNAc1pUT-2 (panel 5) or control empty vector protein (panel 6). Panel 1 shows UDP-GlcNAc and UDP-Glc standards; panel 2 shows UDP-GalNAc and UTP standards. The HPLC peaks in the panels, based on retention time, are UDP-GlcNAc (11.7 min), UDP-Glc (12.3 min, marked as #2), UDP-GalNAc (11.7 min) and UTP (16.1 min). The minor peak marked as 'a' at 14.5 min is UDP contamination from the UTP reagent. The peaks labelled by the arrows represent the enzymatic products.

The kinetics for the reverse reaction had apparent K_m values of $219 \mu\text{M}$ (UDP-GlcNAc) and $2768 \mu\text{M}$ (UDP-GalNAc), with $V_{\max}$ values of $0.93 \mu\text{M} \cdot \text{s}^{-1}$ (UDP-GlcNAc) and $0.75 \mu\text{M} \cdot \text{s}^{-1}$ (UDP-GalNAc). The k_{cat}/K_m ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$) values are 0.22 (UDP-GlcNAc) and 0.015 (UDP-GalNAc). For GlcNAc1pUT-2, the apparent K_m values for the forward reaction were $180 \mu\text{M}$ (GlcNAc-1-P) and $203 \mu\text{M}$ (UTP), with $V_{\max}$ values of $0.36 \mu\text{M} \cdot \text{s}^{-1}$, and k_{cat}/K_m ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$) values were 0.16 (GlcNAc-1-P) and 0.14 (UTP). The kinetics for the reverse reaction, had apparent K_m values of $65 \mu\text{M}$

(UDP-GlcNAc) and $808 \mu\text{M}$ (UDP-GalNAc), with $V_{\max}$ values of $0.19 \mu\text{M} \cdot \text{s}^{-1}$ (UDP-GlcNAc) and $0.18 \mu\text{M} \cdot \text{s}^{-1}$ (UDP-GalNAc). The k_{cat}/K_m ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$) values are 0.46 (UDP-GlcNAc) and 0.018 (UDP-GalNAc).

Structure prediction for GlcNAc1pUT

Chromatography of GlcNAc1pUT-1 on a gel-filtration column showed that the recombinant enzyme has an apparent molecular

Table 4 Enzyme kinetics of GlcNAc1pUT

The forward kinetic GlcNAc1pUT reactions were measured for 4 min with various concentrations of GlcNAc-1-P (0.02–4 mM) and UTP (0.02–4 mM). In the reverse reaction, the enzymatic activity was measured with various concentrations of UDP-GlcNAc (0.02–4 mM) for 4 min and UDP-GalNAc (0.2–10 mM) for 10 min. Enzyme velocities were plotted and Solver software was used to generate a best-fit curve and for the calculation of V_{max} and apparent K_m . Each value is the mean of triplicate reactions, and the values varied by no more than $\pm 10\%$. The K_m of GlcNAc1pUT-2 for UTP is 928 μM when Glc-1-P is the substrate, with a k_{cat}/K_m value of $3.1 \times 10^{-4} \mu\text{M}^{-1} \cdot \text{s}^{-1}$.

(a)				
Forward reaction				
Enzyme	K_m (GlcNAc-1-P) (μM)	K_m (UTP) (μM)	k_{cat}/K_m (GlcNAc-1-P) ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	k_{cat}/K_m (UTP) ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)
GlcNAc1pUT-1	337	295	0.11	0.12
GlcNAc1pUT-2	180	203	0.16	0.14
(b)				
Reverse reaction				
Enzyme	K_m (UDP-GlcNAc) (μM)	K_m (UDP-GalNAc) (μM)	k_{cat}/K_m (UDP-GlcNAc) ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	k_{cat}/K_m (UDP-GalNAc) ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)
GlcNAc1pUT-1	219	2768	0.22	0.015
GlcNAc1pUT-2	65	808	0.46	0.018

Table 5 The predicted UDP-GlcNAc binding site of GlcNAc1pUT

The protein atoms that are presumably hydrogen-bonded to the nucleotide sugar in the structure of GlcNAc1pUT are listed.

Interacting atoms	Amino acids (GlcNAc1pUT-1)	Amino acids (GlcNAc1pUT-2)
Ribose O2'	Gly ¹³⁷	Gly ¹³³
	Leu ¹³⁴	Lys ¹³⁰
Ribose O3'	Leu ¹³⁴	Lys ¹³⁰
	Val ²⁸²	Val ²⁷⁹
Uracil O2''	Met ¹⁹⁵	Met ¹⁹²
	Gln ²²⁶	Gln ²²³
Uracil O4''	Gly ¹³⁶	Gly ¹³²
	Gly ²⁵²	Gly ²⁴⁹
	Gln ²²⁶	Gln ²²³
	Pro ²⁵⁰	Pro ²⁴⁷
α -Phosphate	Asn ²⁵³	Asn ²⁵⁰
β -Phosphate	Tyr ³³⁶	Tyr ³³³
	Lys ⁴³²	Lys ⁴²⁹
	Asn ²⁵³	Asn ²⁵⁰
GlcNAc O-2	Glu ³³⁵	Glu ³³²
GlcNAc O-3	Glu ³³⁵	Glu ³³²
	Asn ³⁶⁰	Asn ³⁵⁷
GlcNAc O-4	Gly ³²⁰	Phe ³¹⁹
	Asn ³⁶⁰	Asn ³⁵⁷
	Gly ³²⁰	Gly ³¹⁷
	Val ³¹⁹	Val ³¹⁸
GlcNAc O-6		Phe ³¹⁹
	Lys ⁴³²	Lys ⁴²⁹

mass of 60 kDa, indicating that the native enzyme is a monomer in solution, since the predicted molecular mass for monomeric protein is 58.3 kDa; whereas the human AGX1 and yeast QRI1 appears to form a complex as a dimer in solution. A three-dimensional model of *Arabidopsis* GlcNAc1pUT-1 was generated using SWISS-MODEL [22] and PHYRE prediction programs, using human GlcNAc1p nucleotidyltransferase (AGX1) complexed with UDP-GlcNAc (PDB code 1JV1) as the most appropriate template. The overall structure of the model is similar to the fold of the human AGX1. Based on the model, the predicted amino acid residues of the plant GlcNAc1pUT-1 involved in binding of the uracil and ribose, as well as GlcNAc are shown in Table 5. The positions of those amino acids are identical with

the corresponding amino acids in the crystal structure of AGX1. This suggests that, although GlcNAc1pUT-1 shares only 32 % amino acid identity with AGX1, its structure is conserved, despite significant divergence of amino acids between the two enzymes. A three-dimensional model of *Arabidopsis* GlcNAc1pUT-2 was also generated using AGX1 as the template. The two *Arabidopsis* GlcNAc1pUTs have a very similar structural fold (Figure 5B), with some changes in various loop regions (Figure 5B, a–e). For example, the loop designated ‘e’ (between amino acids Pro³¹² and Gly³¹⁷) on GlcNAc1pUT-2 in approximation with the GlcNAc sugar moiety is modelled differently compared with the GlcNAc1pUT-1. Other loop variations are also observed between the two *Arabidopsis* proteins.

DISCUSSION

We have cloned and biochemically characterized *Arabidopsis* GlcNAc1pUT-1 which in the presence of Mg^{2+} and UTP specifically uridylylates GlcNAc-1-P and GalNAc-1-P. The GlcNAc1pUT-1 enzyme is reversible and in the presence of UDP-GlcNAc and PP_i , for example, will form GlcNAc-1-P and UTP. We also cloned and characterized At2g35020, GlcNAc1pUT-2, that encodes an enzyme which uridylylates GlcNAc-1-P, GalNAc-1-P and in addition Glc-1-P (Figure 4). Although the two *Arabidopsis* genes share 86 % sequence identity, the latter GlcNAc1pUT-2 has broader substrate specificity. Comparing the activities of GlcNAc1pUT from different organisms, yeast GlcNAc1pUT (i.e. QRI1) can also utilize Glc-1-P as a substrate and is able to form UDP-Glc [15], just like *Arabidopsis* GlcNAc1pUT-2. The human AGX1 is more specific for the hexosamines and has both GlcNAc1p-UT and GalNAc1p-UT activity, just like *Arabidopsis* GlcNAc1pUT-1 [19].

In *Arabidopsis*, two genes encoding UDP-Glc PPases (UGP1, At3g03250, and UGP2, At5g17310) that share 96 % amino acid sequence similarity have been described [23,24]. Single or double T-DNA (transfer DNA) knockout mutants of these UGP genes revealed no phenotype. This may be explained, in part, by the overexpression of a gene encoding Sloppy, a UDP-sugar PPase [25,26], as suggested by Meng et al. [27]. However, we cannot exclude the possibility that in this mutant background, the promiscuous activity of GlcNAc1pUT-2 can contribute to the formation of UDP-glucose as well.

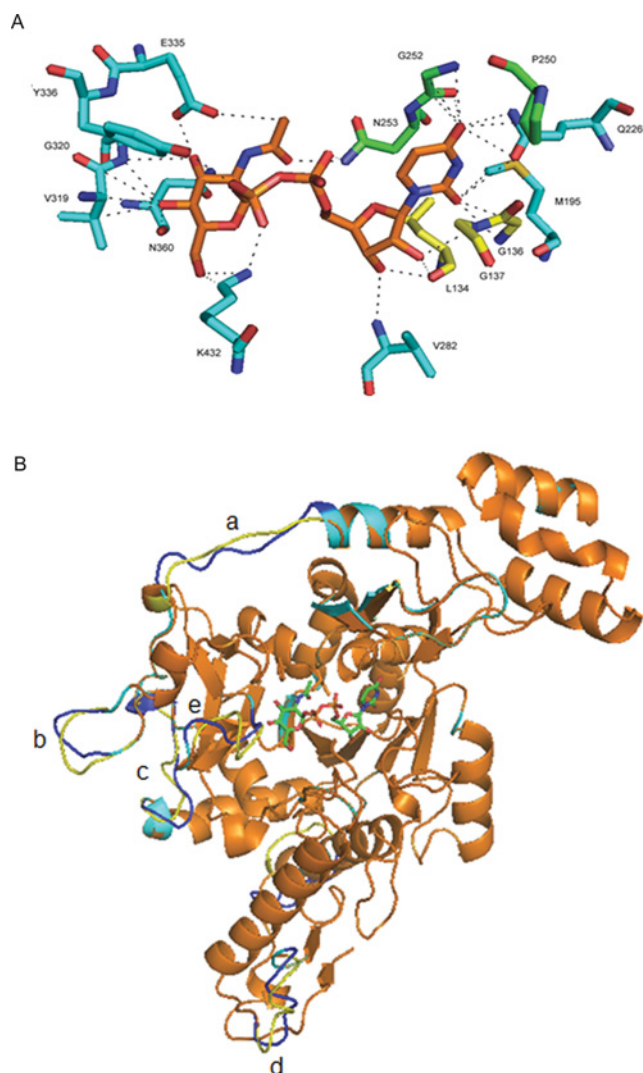


Figure 5 Homology three-dimensional structural model of GlcNAc1pUT-1 and GlcNAc1pUT-2

(A) Predicted amino acid residues in the active site of GlcNAc1pUT-1 complexed with UDP-GlcNAc. Amino acid residues involved in the NB loop are coloured yellow; amino acids involved in the UB loop are coloured green; and amino acids involved in the binding of hexosamine are coloured blue. Magnesium probably co-ordinates the transition by interacting with the α and β phosphate of UDP-GlcNAc. (B) Ribbon represents the superimposed structure of the UDP-GlcNAc-complexed form of GlcNAc1pUT-1 and GlcNAc1pUT-2. GlcNAc1pUT-1 is coloured in cyan and GlcNAc1pUT-2 is coloured in orange. The non-conserved structure between GlcNAc1pUT-1 and GlcNAc1pUT-2 are highlighted in blue and yellow respectively, and labelled a–e.

In the context of plant metabolism and physiology, the ability of the *Arabidopsis* enzyme to utilize UDP-GalNAc *in vivo* as a substrate requires further studies. There are a few reports that GalNAc may be a metabolite in plants. For example, *in vitro* studies have shown that some plant UDP-4-glucose epimerases, including the UGE from barley [12], interconvert UDP-GlcNAc and UDP-GalNAc. Relative large amounts of UDP-GalNAc have been reported to be present in the leaves of squash [28] and in dahlia tubers [13]. Recent metabolomic studies clearly demonstrate the existence of both UDP-GlcNAc and UDP-GalNAc in *Arabidopsis* and fenugreek [14]. An *in vitro* assay also demonstrated that some plant hexosaminidases have a minor activity towards *p*-nitrophenyl-*N*-acetylgalactosaminide as a substrate [29]. However, the existence of a GalNAc residue

in the glycans from land plants has not been reported to our knowledge. On the other hand, reports from Desmidiaceae, a green algae, described a GalNAc moiety in the polysaccharide as determined by lectin binding [11]. It is possible that the levels of GalNAc residues are very low in plants. As GalNAc and GlcNAc have an identical mass, MS of plant oligosaccharides released from glycoprotein would not distinguish these isomers. Therefore more detailed sugar analysis studies will be required to address the occurrence of these glycosyl residues in plants.

The symbiotic bacteria and plant pathogens (bacteria, insects and nematodes) synthesize glycans that contains GalNAc and GlcNAc residues. Thus it is feasible that plants produce hydrolytic enzymes that cleave such glycans and re-use the monosaccharides released. This may explain the need for dual plant GalNAc-1-P and GlcNAc-1-P UTs as described in the present report (Figure 2). Clearly further exploration to elucidate the metabolic pathway that may utilize GalNAc residues in plants is needed. Another important aspect of normal plant metabolism is the recycling of sugars including hexosamines from salvaged plant glycoproteins. How a plant recycles the glycoproteins, cleaves the sugars and re-utilizes the hexosamines remain to be studied. Work is currently underway to understand this cellular process.

GlcNAc1pUT-like sequences are found in the genome of prokaryotes and eukaryotes. Amino acid sequence alignment of several GlcNAc1pUT-like sequences along with UDP-Glc PPases and UDP-sugar PPases shares very low sequence identity, indicating that sequence alone is not the best criterion to determine the specific function of this class of enzymes. The biochemical data provided in the present study also demonstrate that sequence alone is not even sufficient to predict if an enzyme is selective or has broader substrate specificities. Although the amino acid sequence is highly divergent in the UT family of proteins, they appear to have conserved binding and catalytic motifs, suggesting that these UT enzymes have maintained a conserved fold and most likely a similar catalytic mechanism throughout evolution. Putative structural models were created using human AGX1 as a template (PDB code: 1JV1). The two enzymes share a highly conserved structural fold consisting of a central domain with an α/β structure resembling the Rossmann fold, with the N-terminal and C-terminal domains flanking at each end of the polypeptide chain.

Common consensus regions are the glycine-rich (NB) motif for nucleotide binding and the UB motif (Supplementary Figure S1). The amino acids in these two motifs that are involved in binding are shown in the three-dimensional structure homology modelling of GlcNAc1pUT-1 using AGX1 as a template. The amino acids probably involved in the recognition of the hexosamine moiety are shown in the homology model (Figure 5A). The conserved Glu³⁵⁵, Asn³⁶⁰ and Gly³²⁰ are likely to be involved in recognition of C-3, and the amino acids Asn³⁶⁰, Gly³²⁰ and Val³¹⁹ are involved in recognition of C-4 (see Table 5). The NAc-sugar residues of hexosamine linked via C-2 are likely to be critical for the recognition of GlcNAc-1-P, and be different than UT which recognized, for example, Glc-1-P. This is in agreement with the conserved Asn²²³ in AGX1 UAP [19]; in *Arabidopsis* UT-1 and UT-2, Asn²⁵³ (UT1) and Asn²⁵⁰ (UT2) respectively could be one of the amino acids that is involved in the recognition of NAc and α -phosphate. The residues involved in the C-6 recognition include the conserved Lys⁴³² in UT1 and Lys⁴²⁹ in UT2. We have tried to use UDP-GlcNAc-uronic acid or UDP-XylNAc as substrates, but the *Arabidopsis* GlcNAc1pUT-1 and GlcNAc1pUT-2 failed in utilizing them indicating that these *Arabidopsis* enzymes have a strict binding for the gluco-C-6 sugar moiety. Work is underway to mutate and to crystallize the enzyme with different ligands (e.g. GalNAc-1-P, GlcNAc-1-P, UTP, PP_i and Mg²⁺) to

identify how subtle changes in amino acids allow it to be able to accept the C4-epimer (i.e. gluco- compared with galacto-configurations). Comparison of the homology model between GlcNAc1pUT-1 and GlcNAc1pUT-2 indicated that both enzymes share a similar fold, except for some variations in some loop region (Figure 5B). The loop in approximation to the sugar moiety between amino acids Pro³¹² and Gly³¹⁷ on GlcNAc1pUT-2 appears different to GlcNAc1pUT-1. Whether this is the actual cause of the GlcNAc1pUT-2 to take UDP-Glc as a substrate remains unknown. The crystal structure is required to solve the promiscuity of GlcNAc1pUT-2. The promiscuity of GlcNAc1pUT-2 to both Glc-1-P and GlcNAc-1-P as substrates is not a unique feature of *Arabidopsis*, as the yeast PPase [15] utilizes both substrates as well.

The biological significance of GlcNAc1p-UT is critical, not only as post-translational modification of regulatory proteins, but also as a component of glycolipid and glycoprotein. So far T-DNA insertion in At1g31070 loci has no visible phenotype (T. Yang, unpublished work). Considering the key role of UDP-GlcNAc in development and regulation, and the fact that two genes encoding this activity appears to be expressed in most *Arabidopsis* tissues, suggest that both genes need to be knockout for functional analyses. However, in many organisms, the lack of UDP-GlcNAc is lethal.

AUTHOR CONTRIBUTION

Ting Yang was involved in all aspects of the study, including experimental design, performing the research, data analysis and manuscript preparation. Merritt Echols and Andy Martin were involved in performing the research. Maor Bar-Peled directed the study and was involved in all aspects of experimental design, data analysis and manuscript revision.

ACKNOWLEDGEMENTS

We thank Dr John Glushka for his guidance and assistance with the NMR experiments.

FUNDING

This work was supported by the National Science Foundation [grant number IOB-0453664 (to M.B.-P.)] and the BioEnergy Science Center, which is supported by the Office of Biological and Environmental Research in the DOE Office of Science. This research also benefited from activities at the Southeast Collaboratory for High-Field Biomolecular NMR, a research resource at the University of Georgia, funded by the National Institute of General Medical Sciences [grant number GM66340] and the Georgia Research Alliance.

REFERENCES

- Stanley, P., Schachter, H. and Taniguchi, N. (2008) N-Glycans. In *Essentials of Glycobiology* (Varki, A., Cummings, R., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E., eds), Cold Spring Harbor Laboratory Press, NY
- Hancock, J. F. (2004) GPI-anchor synthesis: Ras takes charge. *Dev. Cell* **6**, 743–745
- Schnaar, R. L., Suzuki, A. and Stanley, P. (2008) Glycosphingolipids. In *Essentials of Glycobiology* (Varki, A., Cummings, R., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E., eds), Cold Spring Harbor Laboratory Press, NY
- Raetz, C. R. and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635–700
- Wells, L., Kreppel, L. K., Comer, F. I., Wadzinski, B. E. and Hart, G. W. (2004) O-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits. *J. Biol. Chem.* **279**, 38466–38470
- Hart, G. W., Haltiwanger, R. S., Holt, G. D. and Kelly, W. G. (1989) Glycosylation in the nucleus and cytoplasm. *Annu. Rev. Biochem.* **58**, 841–874
- Jacobsen, S. E., Binkowski, K. A. and Olszewski, N. E. (1996) SPINDLY, a tetratricopeptide repeat protein involved in gibberellin signal transduction in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9292–9296
- Silverstone, A. L., Tseng, T. S., Swain, S. M., Dill, A., Jeong, S. Y., Olszewski, N. E. and Sun, T. P. (2007) Functional analysis of SPINDLY in gibberellin signaling in *Arabidopsis*. *Plant Physiol.* **143**, 987–1000
- Hartweck, L. M., Scott, C. L. and Olszewski, N. E. (2002) Two O-linked N-acetylglucosamine transferase genes of *Arabidopsis thaliana* L. Heynh. have overlapping functions necessary for gamete and seed development. *Genetics* **161**, 1279–1291
- Heese-Peck, A. and Raikhel, N. V. (1998) A glycoprotein modified with terminal N-acetylglucosamine and localized at the nuclear rim shows sequence similarity to aldose-1-epimerases. *Plant Cell* **10**, 599–612
- Brosch-Salomon, S., Hoftberger, M., Holzinger, A. and Lutz-Meindl, U. (1998) Ultrastructural localization of polysaccharides and N-acetyl-D-galactosamine in the secretory pathway of green algae (Desmidiaceae). *J. Exp. Bot.* **49**, 9
- Zhang, Q., Hrmova, M., Shirley, N. J., Lahnstein, J. and Fincher, G. B. (2006) Gene expression patterns and catalytic properties of UDP-D-glucose 4-epimerases from barley (*Hordeum vulgare* L.). *Biochem. J.* **394**, 115–124
- Gonzalez, N. S. and Pontis, H. G. (1963) Uridine diphosphate fructose and uridine diphosphate acetylgalactosamine from dahlia tubers. *Biochim. Biophys. Acta* **69**, 179–181
- Alonso, A. P., Piasecki, R. J., Wang, Y., Laclair, R. W. and Shachar-Hill, Y. (2010) Quantifying the labeling and the levels of plant cell wall precursors using ion chromatography tandem mass spectrometry. *Plant Physiol.* **153**, 915–924
- Mio, T., Yabe, T., Arisawa, M. and Yamada-Okabe, H. (1998) The eukaryotic UDP-N-acetylglucosamine pyrophosphorylases. Gene cloning, protein expression, and catalytic mechanism. *J. Biol. Chem.* **273**, 14392–14397
- Smith, E. E. and Mills, G. T. (1954) Uridine nucleotide compounds of liver. *Biochim. Biophys. Acta* **13**, 386–400
- Feingold, D. S. and Avigad, G. (1980) Sugar nucleotide transformations in plants. In *The Biochemistry of Plants* (Preiss, J., ed.), pp. 101–170, Academic Press, NY
- Milewski, S., Gabriel, I. and Olchow, J. (2006) Enzymes of UDP-GlcNAc biosynthesis in yeast. *Yeast* **23**, 1–14
- Peneff, C., Ferrari, P., Charrier, V., Taburet, Y., Monnier, C., Zamboni, V., Winter, J., Harnois, M., Fassy, F. and Bourne, Y. (2001) Crystal structures of two human pyrophosphorylase isoforms in complexes with UDPGlc(Gal)NAC: role of the alternatively spliced insert in the enzyme oligomeric assembly and active site architecture. *EMBO J.* **20**, 6191–6202
- Yang, T., Bar-Peled, L., Gebhart, L., Lee, S. G. and Bar-Peled, M. (2009) Identification of galacturonic acid-1-phosphate kinase, a new member of the GHMP kinase superfamily in plants, and comparison with galactose-1-phosphate kinase. *J. Biol. Chem.* **284**, 21526–21535
- Kelley, L. A. and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371
- Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201
- McCoy, J. G., Bitto, E., Bingman, C. A., Wesenberg, G. E., Bannen, R. M., Kondrashov, D. A. and Phillips, Jr, G. N. (2007) Structure and dynamics of UDP-glucose pyrophosphorylase from *Arabidopsis thaliana* with bound UDP-glucose and UTP. *J. Mol. Biol.* **366**, 830–841
- Meng, M., Wilczynska, M. and Kleczkowski, L. A. (2008) Molecular and kinetic characterization of two UDP-glucose pyrophosphorylases, products of distinct genes, from *Arabidopsis*. *Biochim. Biophys. Acta* **1784**, 967–972
- Litterer, L. A., Schnurr, J. A., Plaisance, K. L., Storey, K. K., Gronwald, J. W. and Somers, D. A. (2006) Characterization and expression of *Arabidopsis* UDP-sugar pyrophosphorylase. *Plant Physiol. Biochem.* **44**, 171–180
- Kotake, T., Hojo, S., Yamaguchi, D., Aohara, T., Konishi, T. and Tsumura, Y. (2007) Properties and physiological functions of UDP-sugar pyrophosphorylase in *Arabidopsis*. *Biosci. Biotechnol. Biochem.* **71**, 761–771
- Meng, M., Geisler, M., Johansson, H., Harholt, J., Scheller, H. V., Mellerowicz, E. J. and Kleczkowski, L. A. (2009) UDP-glucose pyrophosphorylase is not rate limiting, but is essential in *Arabidopsis*. *Plant Cell Physiol.* **50**, 998–1011
- Tolstikov, V. V. and Fiehn, O. (2002) Analysis of highly polar compounds of plant origin: combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry. *Anal. Biochem.* **301**, 298–307
- Gutternigg, M., Kretschmer-Lubich, D., Paschinger, K., Rendic, D., Hader, J., Geier, P., Ranftl, R., Jantsch, V., Lochnit, G. and Wilson, I. B. (2007) Biosynthesis of truncated N-linked oligosaccharides results from non-orthologous hexosaminidase-mediated mechanisms in nematodes, plants, and insects. *J. Biol. Chem.* **282**, 27825–27840

SUPPLEMENTARY ONLINE DATA

Identification and characterization of a strict and a promiscuous *N*-acetylglucosamine-1-P uridylyltransferase in *Arabidopsis*

Ting YANG^{*†}, Merritt ECHOLS[†], Andy MARTIN[†] and Maor BAR-PELED^{†‡1}

^{*}Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, U.S.A., [†]Complex Carbohydrate Research Center (CCRC), University of Georgia, Athens, GA 30602, U.S.A., and [‡]Department of Plant Biology, University of Georgia, Athens, GA 30602, U.S.A.

Table S1 The effect of potential inhibitors on GlcNAc1pUT-1 activity

Sugar-1-Ps and nucleotides (at a final concentration of 0.5 mM), antibiotics (at 1 mM) or control (water) were mixed with GlcNAc1pUT-1 in 100 mM Tris/HCl (pH 7.6) for 10 min on ice prior to performing the enzymatic reaction. Each value is the mean of duplicate reactions, and the values varied by no more than $\pm 10\%$.

Additive	Relative GlcNAc1pUT activity, forward reaction (%)
UMP	97
NADH	96
CTP	99
TTP	98
ATP	99
NAD ⁺	96
GTP	100
ADP	100
NADP ⁺	96
PP _i	56
ITP	96
UDP	101
Glc-1-P	99
Gal-1-P	99
GlcA-1-P	97
GalA-1-P	100
GalN-1-P	98
Xyl-1-P	100
Fuc-1-P	98
GalNAc	97
GlcNAc	101
Man-1-P	99
Gentamycin	93
Kanamycin	102
Streptomycin	74
Hygromycin	37
Control	100

¹ To whom correspondence should be addressed (email peled@ccrc.uga.edu).

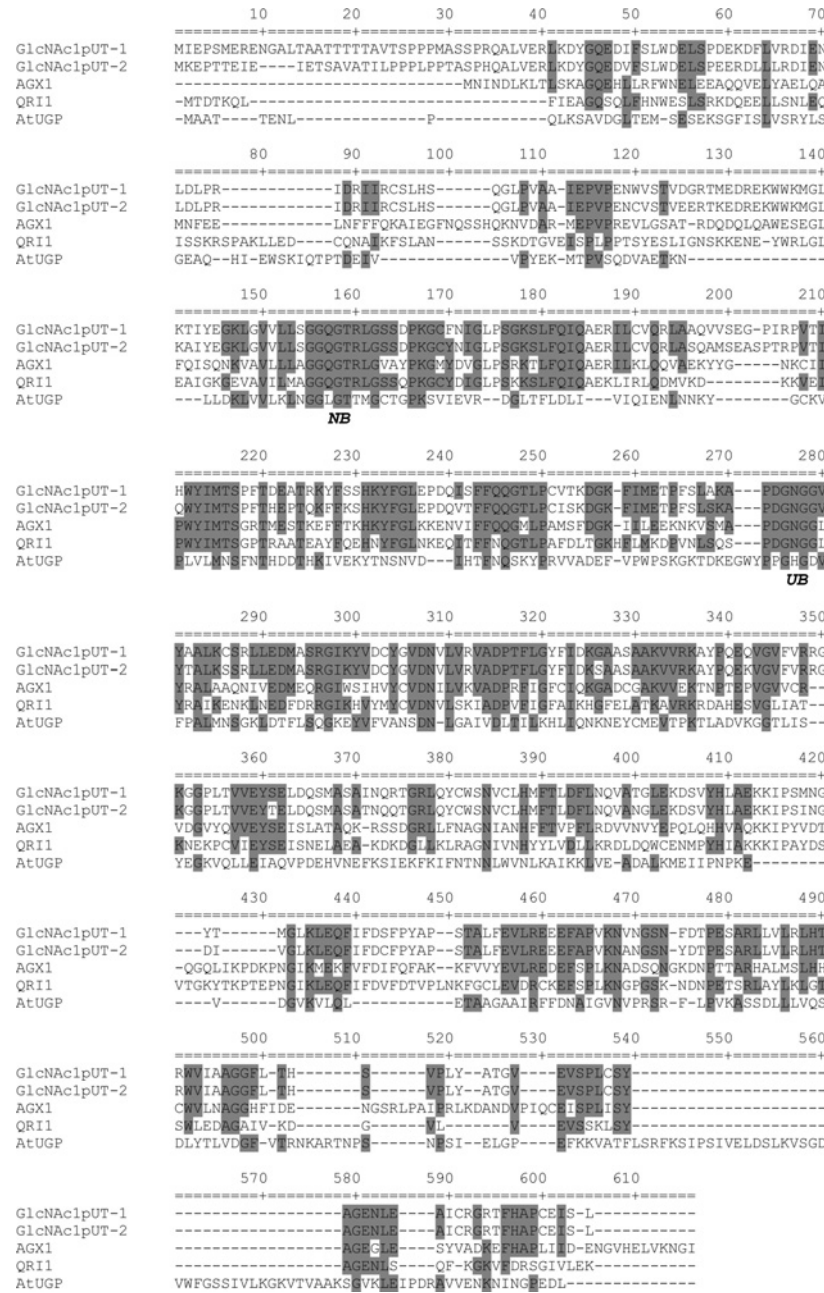


Figure S1 Sequence alignments of GlcNAc1pUT from different organisms

Sequences of GlcNAc1pUT and *Arabidopsis* UDP-Glc pyrophosphorylase (see gene names below) were aligned using T-coffee software [1] with G-block [2]. The conserved motifs presumably involved in nucleotide sugar binding (NB) and the uracil binding (UB) are indicated in bold. Potential amino acids that are conserved in GlcNAc1pUT and AtUGP are highlighted in grey, based on sequence alignment. *Arabidopsis thaliana* GlcNAc1pUT-1 (GenBank® accession number GU937393, At1g31070); *Arabidopsis thaliana* GlcNAc1pUT-2 (GenBank® accession number GU937394, At2g35020); *Homo sapiens* GlcNAc1pUT (AGX1, PDB code 1JV1); *Saccharomyces cerevisiae* GlcNAc1pUT (QR11, GenBank® accession number NP_010180.1); UDP-Glc PPase from *Arabidopsis thaliana* (AtUGP, GenBank® accession number NP_186975).

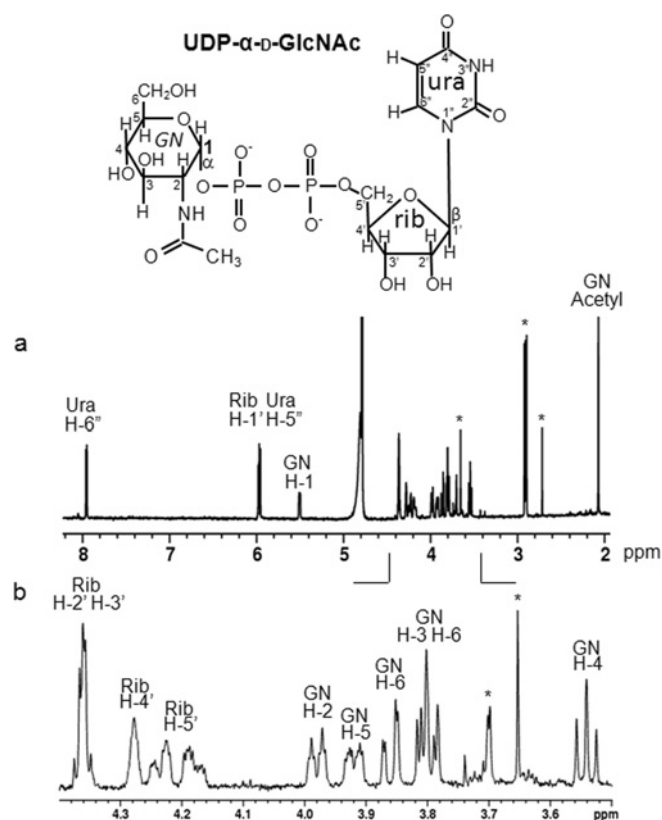


Figure S2 Product analyses by NMR confirm the specific activity of recombinant GlcNAc1pUT-1

Proton NMR spectra of GlcNAc1pUT-1 enzymatic products. Peak eluted from the Bio-Rad column (see Figure 2B of the main text, #1) was collected, freeze-dried, dissolved in ²H₂O and analysed by ¹H-NMR. The full NMR spectrum (2–8 p.p.m.) of the GlcNAc1pUT-1 reaction product is shown. A more detailed spectrum that covers the sugar anomeric region (5–6 ppm) and the NDP-sugar carbon ring (3.5–4.4 p.p.m.) are shown in panel 'a' and panel 'b'. The ¹H chemical-shift values of specific protons along the UDP-α-D-GlcNAc structure are indicated by underlining. H refer to protons belonging to the GlcNAc ring; H' refer to protons belonging to the ribose ring, and H" refer to protons belonging to the uracil ring. The peak labeled with '*' is the signal from column contaminants.

REFERENCES

- Notredame, C., Higgins, D. G. and Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217
- Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552

Received 1 March 2010/15 June 2010; accepted 17 June 2010

Published as BJ Immediate Publication 17 June 2010, doi:10.1042/BJ20100315

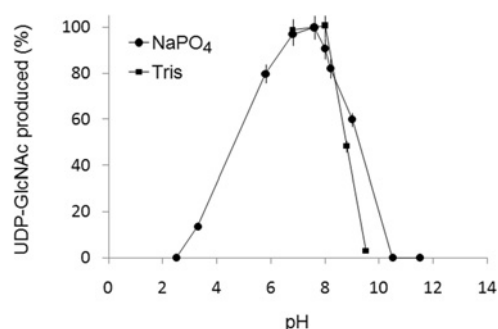


Figure S3 The effects of buffer and pH on GlcNAc1pUT-1 activity

GlcNAc1pUT-1 activity was determined at different buffers at different pH. Each value is the mean of duplicate reactions, and the values varied by no more than ±5%.

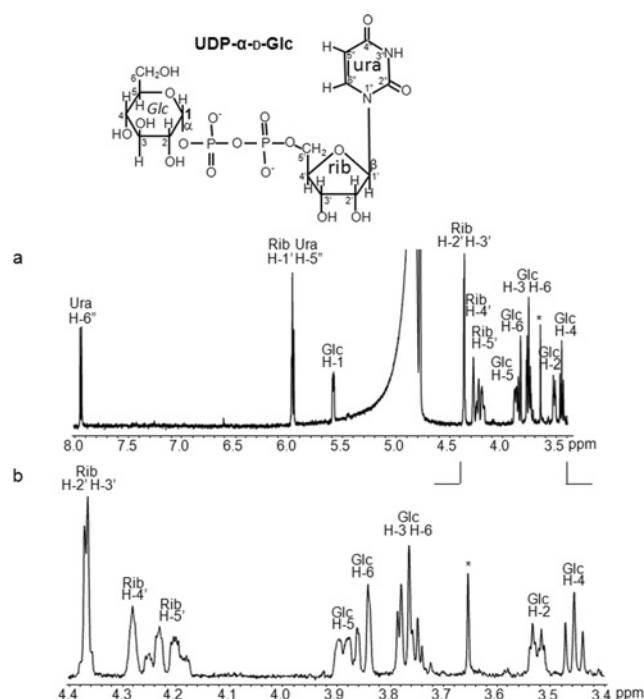


Figure S4 Product analyses by NMR confirm the specific activity of recombinant GlcNAc1pUT-2

Proton NMR spectra of GlcNAc1pUT-2 enzymatic products. Peak eluted from the Bio-Rad column (see Figure 4B of the main text, #2) was collected, freeze-dried, dissolved in ²H₂O and analysed by ¹H-NMR. The full NMR spectrum (3.5–8 p.p.m.) of the GlcNAc1pUT-2 reaction product is shown. A more detailed spectrum that covers the sugar anomeric region (5–6 p.p.m.) and the NDP-sugar carbon ring (3.4–4.4 p.p.m.) are shown in panel 'a' and panel 'b'. The ¹H chemical-shift values of specific protons along the UDP-α-D-Glc structure are indicated by underlining. H refer to protons belonging to the Glc ring; H' refer to protons belonging to the ribose ring, and H" refer to protons belonging to the uracil ring. The peak labelled with '*' is the signal from column contaminants.