A missense mutation in a patient with developmental delay affects the activity and structure of the hexosamine biosynthetic pathway enzyme AGX1

Xiping Chen, Olawale G. Raimi, Andrew T. Ferenbach and Daan M.F. van Aalten (iD)

Division of Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dundee, UK

Correspondence
D. M. F. van Aalten, Division of Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dow Street, Dundee, DD1 5EH, UK
Tel: +441382384979
E-mail: dmfvanaalten@dundee.ac.uk

O-GlcNAcylation is a post-translational modification catalysed by *O*-GlcNAc transferase (OGT). Missense mutations in *OGT* have been associated with developmental disorders, OGT-linked congenital disorder of glycosylation (OGT-CDG), which are characterized by intellectual disability. OGT relies on the hexosamine biosynthetic pathway (HBP) for provision of its UDP-GlcNAc donor. We considered whether mutations in UDP-*N*-acetylhexosamine pyrophosphorylase (UAP1), which catalyses the final step in the HBP, would phenocopy OGT-CDG mutations. A *de novo* mutation in *UAP1* (NM_001324114:c.G685A:p.A229T) was reported in a patient with intellectual disability. We show that this mutation is pathogenic and decreases the stability and activity of the UAP1 isoform AGX1 *in vitro*. X-ray crystallography reveals a structural shift proximal to the mutation, leading to a conformational change of the *N*-terminal domain. These data suggest that the UAP1^{A229T} missense mutation could be a contributory factor to the patient phenotype.

Keywords: enzyme mutation; neurodevelopment; *O*-GlcNAcylation; pathogenesis; protein structure

O-linked β-*N*-acetylglucosamine (*O*-GlcNAc) is a post-translational modification that occurs on thousands of nuclear and cytoplasmic proteins. The dynamic cycling of *O*-GlcNAc is regulated by a single pair of enzymes, the *O*-GlcNAc transferase (OGT), which transfers the GlcNAc moiety from UDP-GlcNAc to Ser/Thr of protein targets [1], and *O*-GlcNAcase, which removes *O*-GlcNAc from the modified proteins [2]. Protein *O*-GlcNAcylation is involved in various physiological processes, such as transcription, nutrient sensing and cell signalling [3–7]. In *Caenorhabditis elegans*, *O*-GlcNAc orchestrates metabolism through the insulin-signalling pathway to maximize neuron regeneration in response to neuronal injury [8]. The *Drosophila* OGT homologue is encoded by *super sex combs*, which is required for the repression of Polycomb genes [9], a cluster of genes essential for the development of animals and plants by dynamically regulating chromatin modification [10–12]. Deletion of either the *Ogt* or the *Oga* gene leads to embryonic lethality in mice [13,14]. *O*-GlcNAcylation is involved in neurodegenerative diseases, such as Alzheimer's disease [15], amyotrophic lateral sclerosis [16] and Parkinson's disease (PD) [17,18]. Recently, dysregulation of *O*-GlcNAcylation caused by *OGT* gene mutations has been associated with a developmental disorder termed OGT-linked

congenital disorder of glycosylation (OGT-CDG) with patients characterized by intellectual disability and developmental delay [19–23].

UDP-GlcNAc, the substrate of OGT, is the end product of the hexosamine biosynthetic pathway (HBP) (Fig. 1) that is present in all three kingdoms of life [24–26]. The first and rate-limiting step of the HBP is catalysed by glucosamine-fructose-6-phosphate aminotransferase 1 (GFAT) which transfers an amino group from glutamine to fructose-6-phosphate to produce glucosamine-6-phosphate. Glucosamine-6-phosphate then turns into GlcNAc-6-phosphate by the action of glucosamine-6-phosphate *N*-acetyltransferase (GNA1). The third step is catalysed by phosphoglucomutase 3 (PGM3) that converts GlcNAc-6-phosphate into GlcNAc-1-phosphate. Finally, the production of UDP-GlcNAc is reversibly catalysed by UDP-*N*-acetylhexosamine pyrophosphorylase (UAP1) which uses GlcNAc-1-phosphate and UTP as substrates in the forward reaction to produce UDP-GlcNAc and pyrophosphate.

There are several lines of evidence that HBP is essential for development. *GFAT* gene knockout cells cannot survive without supplementation of the enzyme product, GlcNAc [27]. *GFAT* gene mutations have been linked to congenital myasthenic syndromes [28], a group of conditions owing to the defect of signal transmission from nerve cells to muscles [29]. *GFAT* patients displayed reduced levels of GFAT protein and *O*-GlcNAcylation [28]. Inactivation of GNA1 in mouse embryonic fibroblasts led to defects in UDP-GlcNAc biosynthesis, *O*-GlcNAc modification and cell proliferation [30]. In addition, mutations of the *PGM3* gene that resulted in compromised enzymatic activity were identified as a cause of immunodeficiency, skeletal dysplasia and neurocognitive impairment [31–33]. A study in *Drosophila* showed that mutation of *nesthocker*, the orthologue of the *PGM3* gene, resulted in decreased levels of cytoplasmic UDP-GlcNAc, disrupted protein *O*-GlcNAcylation and finally blocked mesodermal development [34]. In *Drosophila*, the *UAP1* gene orthologue *mummy* modulates
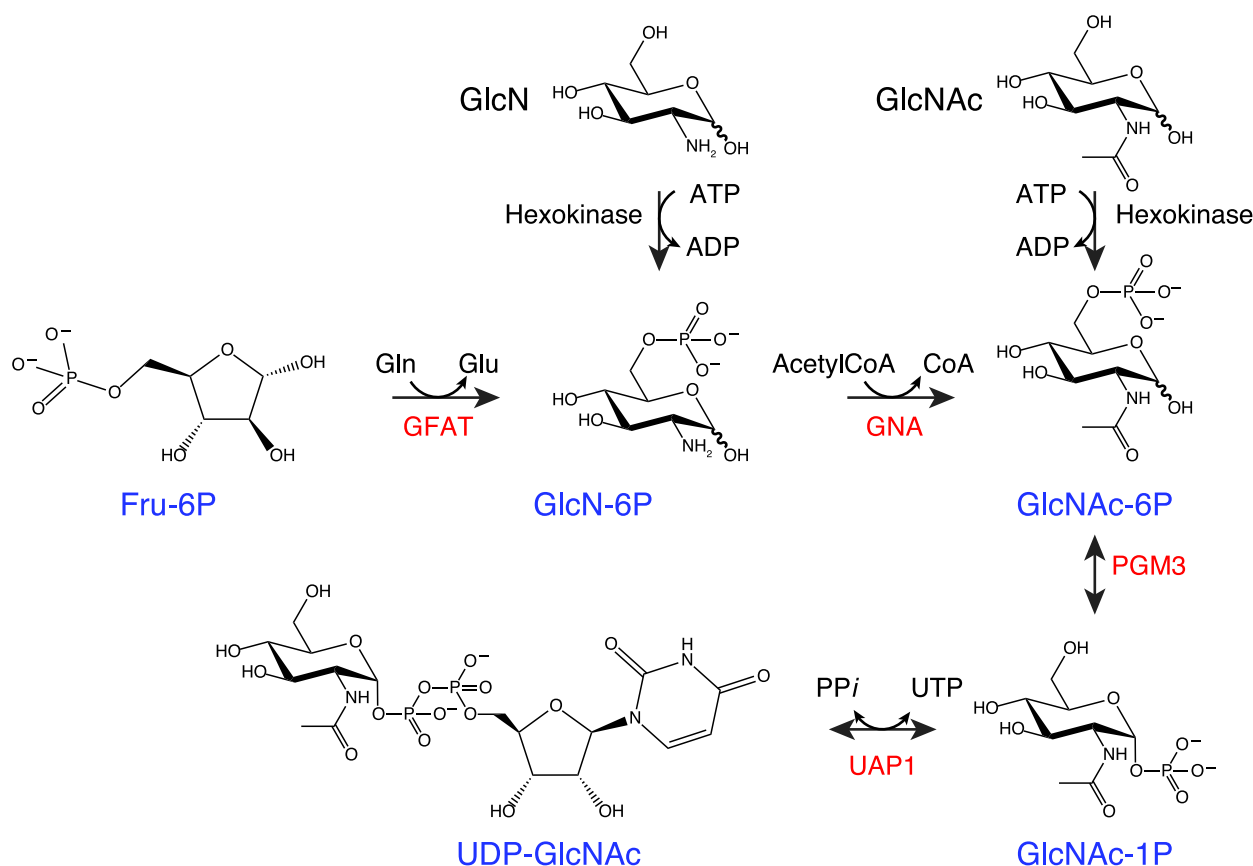

Fig. 1. The eukaryotic HBP. The four main enzymes of the pathway are coloured in red, and metabolites generated by the HBP are coloured in blue. Reversible and irreversible reactions are indicated with double- and single-headed arrows, respectively

decapentaplegic (Dpp) signalling in the embryonic epidermis and is required for epithelial morphogenesis and nervous system development [35–37]. These data collectively suggest a link between the perturbation of the HBP, protein *O*-GlcNAcylation and neurodevelopment.

The human *UAP1* gene encodes two different isoforms, named AGX1 and AGX2, with AGX1 being more abundant in testis and AGX2 in somatic tissues [38]. The two isoforms differ by a 17-amino acid insertion in the *C* terminus of AGX2. Compared to AGX2, AGX1 has similar efficiency in producing UDP-GlcNAc and 50 times more efficiency in producing UDP-GalNAc [39]. Structural characterization revealed that both AGX1 and AGX2 comprise three domains, that is *N*-terminal, catalytic and *C*-terminal domain, and the 17-amino acid insertion in the *C*-terminal domain of AGX2 ablated oligomeric assembly [40]. *UAP1* genes, designated *glmU* in bacteria, have been reported to be essential in both prokaryotes and eukaryotes [41,42]. For instance, inactivation of GlmU thermosensitive mutant in *Escherichia coli* resulted in a 37% reduction in peptidoglycan content and cell lysis [43]. The fungal *UAP1* gene orthologue *UAP* is essential in both *Saccharomyces cerevisiae* and *Aspergillus fumigatus* [44,45]. Because the action of OGT depends on the steady supply of cytoplasmic UDP-GlcNAc through the HBP, we considered whether *UAP1* gene mutations would phenocopy OGT-CDG mutations.

As a clinical genetics programme, the Deciphering Developmental Disorders Study (DDD) was recently established to investigate undiagnosed developmental disorders using exome sequencing and array-based detection of chromosomal rearrangements [46]. A *de novo* heterozygous mutation in the *UAP1* gene (NM_001324114:c.G685A:p.A229T, DDD entry: https://decipher.sanger.ac.uk/ddd/research-variant/c9774b30226f59f8c1f79a7578fe5fc3/overview), which encodes a missense variant UAP1[A229T], has been identified in a patient from the DDD study with reported cranial and skeletal abnormalities and intellectual disability. Here, we investigate possible effects of the A229T mutation on AGX1 using various *in vitro* approaches. First, clinical bioinformatics tools predict that the mutation is likely to be pathogenic, while biochemical analysis with recombinant AGX1[A229T] reveals that the A229T mutation causes a reduction of protein thermal stability. Compared to wild-type AGX1, AGX1[A229T] has lower activity in producing UDP-GlcNAc. X-ray crystallographic structural characterization demonstrates that the A229T mutation lies proximal to the active site. The mutation induces local structural shift which weakens the hydrogen bond network connecting the *N*-terminal and

catalytic domain, leading to conformational changes of the *N*-terminal domain that explains changes in catalytic activity. Together, these *in vitro* data suggest that the UAP1[A229T] missense mutation could contribute to the patient phenotype.

Materials and methods

Construct cloning, protein expression and purification

The construct with full-length AGX1 (UniProt Q16222) cloned into pGEX6P4 (expressing GST-tagged protein) was obtained from our lab stock [47]. The A229T mutation was introduced using site-directed mutagenesis with forward primer ggtggtctttatcggAcacttgcagcccag and reverse primer ctgggctgcaagtgTccgataaagaccacc. All constructs were verified by sequencing.

E. coli (DE3) pLysS was used for recombinant expression of *N*-terminally GST-tagged AGX1[wt] and AGX1[A229T]. Protein expression was induced at 18 °C by addition of 200 μM IPTG when absorbance at OD_{600} reached 0.6 and incubated for a further 16 h. Cells were harvested by centrifugation at 4200 r·min^{-1} at 4 °C for 30 min and then resuspended in lysis buffer (100 mM Tris, 150 mM NaCl, 0.5 mM TCEP pH 7.5, 0.1 mg·mL^{-1} lysozyme, 0.1 μg·mL^{-1} DNase, 1 mM benzamidine, 0.2 mM PMSF and 5 μM leupeptin). Cells were lysed using French Press with a pressure of 50 000 psi. Cell lysate was then spun down at 20 000 r·min^{-1} at 4 °C for 30 min in an Avanti J-25 centrifuge (Beckman). The supernatant was incubated with Glutathione-Sepharose beads (GE Healthcare) at 4 °C for 2 h, and beads were then washed with 4–5 column volumes of lysis buffer. Bound protein was cleaved overnight from GST beads by PreScission protease at a final concentration of 0.5 μM. Protein was pooled then concentrated to 5 mL before loading onto a pre-equilibrated Superdex 200 26/60 preparative column (AKTA Prime system used; GE Healthcare). Fractions containing pure protein were pooled and concentrated to 0.5 mL using a 50 kDa 20 mL Vivaspin concentrators. Protein was quantified using NanoDrop and then flash-frozen in liquid nitrogen.

Differential scanning fluorimetry assay

Assays were performed with 2 μM protein (1 mg·mL^{-1}), 5 × SYPRO Orange dye (Sigma) and 1 mM ligand in the assay buffer (50 mM Tris pH 7.5, 150 mM NaCl, 2 mM MgCl$_2$ and 0.5 mM TCEP) in a total volume of 25 μL. Fluorescence (λ_{ex} = 530 nm, λ_{em} = 560 nm) was monitored with a Bio-Rad (CFX Connect™) real-time system, while system temperature was increased from 20 to 90 °C in 1 °C increments. Data were fitted to Boltzmann sigmoidal curve using GRAPHPAD PRISM® 6 to obtain T_m, which is the

inflection point of the curve. Assays were performed with six biological replicates, and each biological replicate contained three technical replicates.

Enzymatic activity assay

A coupled enzyme assay was deployed to study steady-state kinetics. In the assay, AGX1^{wt} utilizes UTP and GlcNAc-1P to produce UDP-GlcNAc and pyrophosphate; pyrophosphate is then degraded by pyrophosphatase (coupling enzyme) into inorganic phosphate. BIOMOL Green® reagent (0.03% w/v malachite green, 0.2% w/v ammonium molybdate and 0.5% v/v Triton X-100 in 0.7 N HCL) was used to detect the inorganic phosphate produced [48]. To determine Michaelis–Menten parameters, the assay was performed in triplicate with each reaction in a total volume of 100 µL buffer containing 50 mM Tris, pH 7.5, 10 mM MgCl$_2$, 10% (v/v) glycerol, 1 mM DTT, 0.05 units pyrophosphatase and 20 nM AGX1, and one of the two substrates was kept at an excess concentration (500 µM), while the other one varied from 0 to 500 µM. Reactions were initiated by the addition of 20 nM AGX1^{wt}/ AGX1^{A229T} and then incubated at room temperature for 10 min. Reactions were terminated by the addition of 100 µL BIOMOL Green® reagent. The assay plate was kept at room temperature for 30 min to allow colour development. The absorbance at 620 nm was measured using a Spectra max 340 PC. The assay was performed with five biological replicates, and each biological replicate contained three technical replicates. The turnover of substrates was < 10% under all conditions tested. Data were analysed using GRAPHPAD PRISM® to calculate K_m, V_{max} and *P*-values.

An XBZ (xanthene-based Zn) assay was developed to probe the activity of AGX1^{wt} and AGX1^{A229T} catalysing the reverse reaction. In the XBZ assay, the fluorescent probe XBZ is activated once it binds to UTP [49]. The high sensitivity and selectivity of XBZ for NTP over NDP-sugars and sugar-phosphates has been demonstrated previously [49,50]. Since pyrophosphate is one of the substrates of AGX1 in the reverse reaction and, in addition to UTP, it also induces XBZ fluorescence [49], we showed that 50 µM pyrophosphate was thoroughly degraded into inorganic phosphate after incubation with 0.05 U pyrophosphatase at room temperature for 30 min (Fig. S5A). A UTP standard curve was obtained using a UTP standard that was purchased from Sigma (catalog no. U6625) and serially diluted in the buffer containing 50 mM Tris pH 7.5, 0.5 mM MgCl$_2$, 2% (v/v) glycerol, 20 µM pyrophosphate, 40 µM UDP-GlcNAc and 0.05 U pyrophosphatase in a total volume of 100 µL (Fig. S5B). Plates were placed at room temperature for 30 min to allow the development of background signals. For measuring the reverse reaction catalysed by AGX1^{wt} and AGX1^{A229T}, assays were performed in mixtures containing 50 mM Tris pH 7.5, 2% (v/v) glycerol, 0.5 mM MgCl$_2$, 40 µM UDP-GlcNAc and 20 µM

pyrophosphate in a total volume of 100 µL. Reactions were initiated by addition of 0.5 nM enzyme and then incubated at room temperature for 1 min. Reactions were terminated by boiling for 5 min. After cooling down to room temperature, reactions were then supplied with 0.05 U pyrophosphatase (in a volume of 5 µL) and incubated further at room temperature for 30 min to allow thorough degradation of unconsumed pyrophosphate. XBZ was prepared at 15 µM in a buffer containing 25 mM HEPES pH 7.5, 10 mM NaCl, 75 µM pyrocatechol violet and 50% (v/v) methanol. Two hundred microlitre XBZ solution was added to each assay, and fluorescence was measured using a Spectra max 340 PC with λ_{ex} and λ_{em} set as 485 and 530 nm, respectively. The turnover of both substrates was below 10%. Data analysis was performed using GRAPHPAD PRISM®.

Crystallization and structure solution

Before setting up crystal trays, AGX1^{A229T} (10 mg·mL^{-1}) was buffer exchanged into crystallization buffer containing 25 mM Tris, pH 7.5, which is similar to the crystallization buffer of wild-type AGX1 [40]. Crystallization was set up in sitting drop format with each drop containing 0.3 µL reservoir solution and 0.2 µL protein. Crystals appeared in 0.2 M MgCl$_2$, 0.1 M Tris, pH 8.5 and 30% PEG 4000 within 3 days. Crystals were flash-frozen in liquid nitrogen and sent to the I24 beamline of the UK National Synchrotron (Diamond Light Source) for data collection. Data sets were indexed and integrated with *iMOSFLM* [51]. The phase problem was solved by molecular replacement using *MOLREP* [52] and the PDB 1JV1 [40] as a phase donor. The structure was refined using *REFMAC*5 [53] and built manually using COOT [54]. PYMOL [55] was used to generate figures.

Bioinformatics analysis

UniProt IDs of protein sequences used to construct the phylogenetic tree are: F1S210 (*Sus scrofa*), F1MJP7 (*Bos taurus*), Q16222 (*Homo sapiens*), Q91YN5 (*Mus musculus*), F1QZD3 (*Danio rerio*), F6SUV5 (*Xenopus tropicalis*), F1NFV9 (*Gallus gallus*), Q9Y0Z0 (*Drosophila melanogaster*), Q54GN5 (*Dictyostelium discoideum*), Q940S3 (*Arabidopsis thaliana*), Q386Q8 (*Trypanosoma brucei brucei*), Q18493 (*C. elegans*) and P43123 (*S. cerevisiae*). A phylogenetic tree was calculated by CLUSTALX [56] and visualized using MEGA7 [57] with topology displayed. Protein sequences were aligned using JALVIEW [58].

Results

UAP1 A229T mutation is potentially pathogenic

In diploid organisms, haploinsufficiency is a phenomenon in which a single copy of a functional gene

is not sufficient to produce the normal/wild-type phenotype. Since the patient is only heterozygous for the UAP1 A229T missense mutation (DDD entry: https://decipher.sanger.ac.uk/ddd/research-variant/c9774b30226f59f8c1f79a7578fe5fc3/overview), we first investigated whether the *UAP1* gene is associated with haploinsufficiency. Haploinsufficiency score (HI index) is the predicted probability of a gene to be haploinsufficient and over 80% of protein-coding genes in the human genome have been scored ranging from 0 (haploinsufficient) to 100% (haplosufficient) [59]. Based on the analysis of protein-coding genetic variation in 60 706 humans, the loss intolerance (pLI) is computed to predict the probability of a gene to be loss-of-function (LoF) intolerant [60]. High pLI scores (pLI ≥ 0.9) indicate LoF intolerance, whereas low pLI scores (pLI ≤ 0.1) indicate LoF tolerance. The predicted HI and pLI of the *UAP1* gene are 42.14% and 0.91, respectively, which suggests that the *UAP1* gene is potentially haploinsufficient and LoF intolerant.

To investigate whether the UAP1 A229 mutation is tolerated in an evolutionary context, primary sequences of eukaryotic UAP1 orthologues were aligned. The data show that A229 is conserved in all eukaryotes (Figs 2 and S1), suggesting an important structural and/or functional role. In addition to
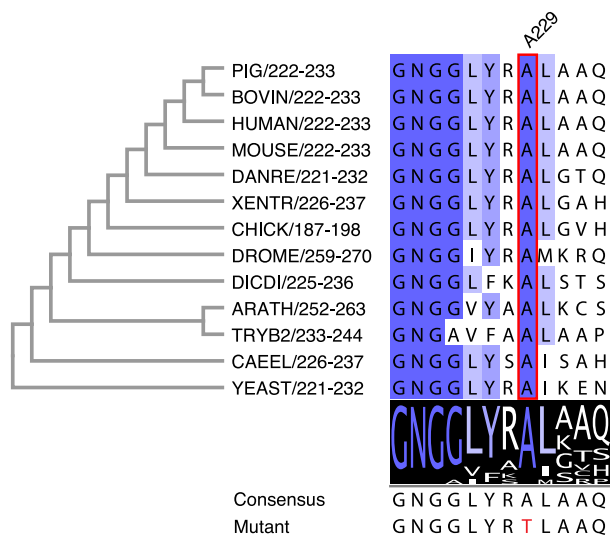
sequence analysis, the UAP1 A229T mutation was predicted to be pathogenic by *in silico* tools, such as SIFT [61] (score = 0.03) and Polyphen-2 [62] (score = 1) which evaluate effects of amino acid substitutions on proteins and MutationTaster [63] (*P* value = 1) which predicts the pathogenic potential of the DNA mutation. Together, these data suggest that the heterozygous UAP1 A229T mutation is potentially pathogenic.

The A229T mutation decreases AGX1 stability

We next investigated the possible effects of the A229T mutation on the biochemical properties of AGX1. AGX1^{A229T} was generated by site-directed mutagenesis, and GST-tagged AGX1^{wt} and AGX1^{A229T} were recombinantly expressed in *E. coli* and purified through glutathione-affinity and size exclusion chromatography. During protein purification, it was noticed that AGX1^{A229T} was prone to forming aggregates in solution, which may explain the lower protein yield (1.0 mg·L^{-1} of culture) compared to AGX1^{wt} (4.5 mg·L^{-1} of culture). To examine whether this was the result of intrinsic AGX1^{A229T} stability, we analysed the thermal stability of AGX1^{wt} and AGX1^{A229T} using a differential scanning fluorimetry assay (DSF). This revealed that, compared to AGX1^{wt}, AGX1^{A229T} was less stable as indicated by a reduction of the melting temperature (T_m) by approximately 5.3 °C (Fig. 3A). DSF can also be used to observe protein–ligand interactions, which often increase the melting temperature [64,65]. The melting temperatures of both AGX1^{wt} and AGX1^{A229T} were increased by incubation with the substrates/products, that is UTP, GlcNAc-1P and UDP-GlcNAc (Fig. 3B), suggesting that AGX1^{A229T} retains the ability to bind these substrates. Of note, AGX1^{A229T} was less stable than AGX1^{wt} regardless of which ligand was supplemented (Fig. 3B). Hence, these data suggest that the A229T mutation decreases AGX1 stability *in vitro*.

The A229T mutation affects AGX1 activity

Given that A229 is conserved in eukaryotic UAP1 orthologues (Figs 2 and S1), and the A229T mutation resides in the catalytic domain, we next investigated the potential effects of the A229T mutation on the steady-state kinetics of AGX1 catalysing the forward reaction. The reaction was monitored through a colorimetric assay using pyrophosphatase as the coupling enzyme, which hydrolyses the AGX1 pyrophosphate product leading to free inorganic phosphate that can be detected with Biomol Green [48]. The results show that, compared to AGX1^{wt}, AGX1^{A229T} has a twofold

Fig. 2. Sequence alignment of eukaryotic UAP1s. Sequences are named as organism abbreviation followed by sequence region. The phylogenetic tree was constructed based on full-length proteins using Clustalx [56] and then visualized with mega7 [57]. A229 is highlighted in a red box. DANRE, *Danio rerio*; XENTR, *Xenopus tropicalis*; DROME, *Drosophila melanogaster*; DICDI, *Dictyostelium discoideum*; ARATH, *Arabidopsis thaliana*; TRYB2, *Trypanosoma brucei brucei* (strain 927/4 GUTat10.1); CAEEL, *Caenorhabditis elegans*.
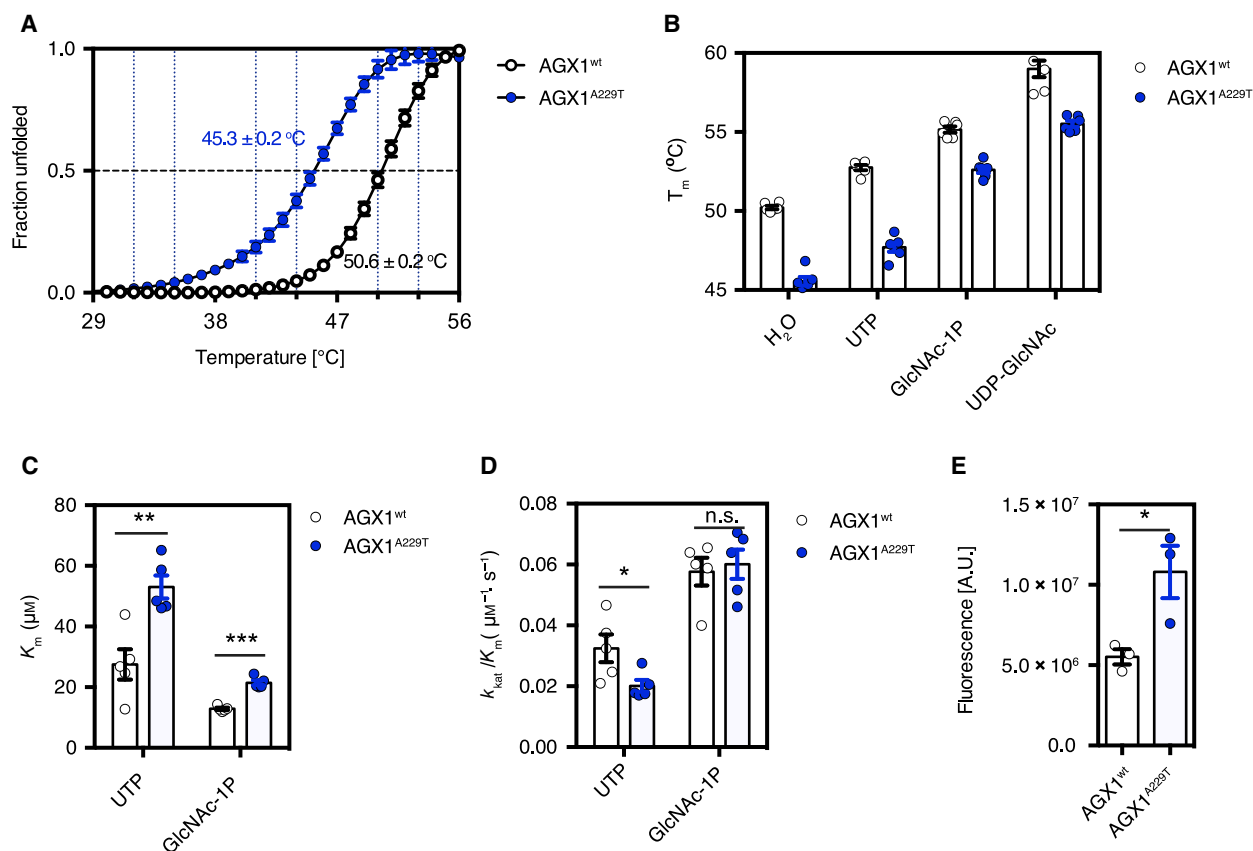
Fig. 3. *In vitro* characterization of recombinant AGX1[A229T]. (A) Thermal denaturation curve showing unfolded fraction of recombinant AGX1[wt] and AGX1[A229T] as a function of temperature. Data were fitted to Boltzmann sigmoidal curve equation, with error bars representing SEM of six biological replicates. (B) Melting temperature (T_m) of AGX1[wt] and AGX1[A229T] in the absence or presence of ligand. (C), (D) K_m and k_{cat}/K_m of AGX1[wt] and AGX1[A229T] utilizing UTP and GlcNAc-1P. Error bars represent SEM of five biological replicates. *$P = 0.0382$, **$P = 0.0036$, ***$P < 0.001$, *t*-test. (E) Activity assay for the AGX1[wt] and AGX1[A229T] reverse reaction. Error bars represent SEM of three biological replicates. *$P = 0.0359$, *t*-test.

increased K_m ($P < 0.01$, *t*-test) towards both substrates (i.e. UTP and GlcNAc-1P) and decreased catalytic efficiency in utilizing UTP ($P = 0.0382$, *t*-test) (Fig. 3C, D). To enable comparison of AGX1[wt] and AGX1[A229T] activity in catalysing the reverse reaction, we used a xanthene-based Zn (XBZ) fluorophore, which fluoresces upon binding of the reaction product UTP [49]. In this assay, AGX1[A229T] is more active than AGX1[wt] in consuming UDP-GlcNAc and pyrophosphate ($P = 0.0359$, *t*-test, Fig. 3E). Together, these data suggest that the A229T mutation affects AGX1 activity which may result in lower UDP-GlcNAc levels *in vivo*.

The A229T mutation induces structural changes

We next determined the crystal structure of AGX1[A229T] to investigate the structural consequences of substituting the conserved A229 with a threonine.

We were unsuccessful in obtaining the AGX1[A229T] crystals using the published AGX1[wt] crystallization condition [40]. However, crystals of AGX1[A229T] in complex with UDP-GlcNAc were obtained from screening commercial crystallization conditions, and synchrotron diffraction data were collected to 1.7 Å (Table 1). The structure was solved by molecular replacement using the published AGX1 structure (PDB: 1JV1 [40]) as a search model and refined to $R = 0.187$, $R_{free} = 0.226$. AGX1[A229T] crystallized with two molecules in a *P*1 unit cell. Structural superposition of AGX1[A229T] and AGX1[wt] revealed conformational differences (pairwise RMSDs of the A and B chains 0.4–1.0 Å, Fig. S2) in the *N*-terminal domain of both chains (Fig. 4A,B). In addition, the uridine moiety of UDP-GlcNAc binding in the active site of the AGX1[A229T] B chain points to the outside of the substrate binding pocket, which is different from the

Table 1. Scaling and model-building statistics of the AGX1^{A229T} crystal structure. Values in brackets are for the highest resolution shell.

Data collection	
Space group	P 1
Cell dimensions	
a, *b*, *c* (Å)	45.61, 64.57, 82.07
α, β, γ (°)	95.63, 101.09, 104.93
Resolution (Å)	29.26–1.70 (1.70–1.73)
R_{merge}	0.024 (0.429)
$I/\sigma I$	46.3 (2.6)
Completeness (%)	95.4 (90.9)
Redundancy	3.3 (3.2)
No. of reflections	93 239 (9102)
R_{work}/R_{free}	0.180/0.226
No. of nonhydrogen atoms	
Protein	7729
Ligand/ion	78
Water	715
B factors	
Protein	26.5
Ligand/ion	22.4
Water	32.9
RMSDs	
Bond lengths (Å)	0.011
Bond angles (°)	1.72
Ramachandran plot	
In preferred regions (%)	98.2
In allowed regions (%)	1.8
Outliers (%)	0.0
PDB code	6Z2F

conformation of UDP-GlcNAc binding in the active site of the AGX1^{A229T} A chain and AGX1^{wt} (Fig. 4B). We suggest that the conformational changes in the AGX1^{A229T} B chain are due to an alkylation of Cys251 in the active site that displaces UDP-GlcNAc (Fig. 4B). We consider this alkylation to be artefactual as carbamidomethylation on multiple cysteines of both AGX1^{wt} and AGX1^{A229T} was observed with peptide mass fingerprinting (Table S1). Since our aim was to investigate the structural consequences of substituting the conserved A229 with a threonine, we focused on a detailed comparison of AGX1^{wt} A chain with its AGX1^{A229T} equivalent.

In line with the decreased catalytic efficiency of AGX1^{A229T} in consuming UTP (Fig. 3D), the A229T mutation is in the middle of the α-helix that coordinates the uridine moiety of UDP-GlcNAc (Fig. 4C). In AGX1^{A229T} structure, there is a shift of 1.2 Å in the position of R228 (the residue next to A229) towards the *N* terminus and a shift of 1.8 Å in the *C* terminus of the helix encompassing the mutation to the opposite face of A229 (pushing effect) (Fig. 4C).

In the A and B chains of AGX1^{wt}, R228 interacts with E44 through two hydrogen bonds (3.0 and 3.1 Å in A chain; 3.3 and 2.6 Å in B chain), which contributes to the interaction between the *N*-terminal and catalytic domain (Fig. S3). The R228-E44 interaction is abolished in the AGX1^{A229T} structure caused by the position shift of R228 (Fig. 4F). The pushing effect is likely due to the bulkier side chain of threonine compared to that of alanine. In AGX1^{wt}, the distances between the side chain of A229 and the main chain of Q195 and Q196 are 3.7 and 3.6 Å, respectively. However, In AGX1^{A229T} structure these distances are reduced to 2.7 and 2.2 Å, respectively, which is below the C-O Van der Waals diameter (3.27 Å [66]). Since alanine has high helix propensity while threonine has high β-sheet propensity [67], we then investigated whether the A229T mutation could cause distortion of this key active site α-helix. We aligned the AGX1^{A229T} A chain with eukaryotic UAP1 orthologues at the point of the mutation. The results showed that, in the aligned sequences, even though there are sequence variations at positions apart from A229 and L230, the AGX1^{A229T} A chain retains the α-helical conformation around position 229 (Fig. 4E). Together, the A229T mutation induced a local structural shift without disrupting the secondary structure of AGX1.

In a global structural view, the A229T mutation causes the *N*-terminal domain of AGX1^{A229T} A chain to adopt an open/relaxed conformation compared to that of AGX1^{wt} (Fig. 4A,F). In AGX1^{wt} B chain, a distant interaction between Q112 in the catalytic loop and M218 in the *N*-terminal domain is mediated by R169 through hydrogen bonds (3.3 and 3.1 Å). The Q112-R169-M218 interaction was observed in the AGX1^{wt} B chain but not A chain (Fig. S3), which probably indicates the interaction is dynamic in solution or catalysis. Along with the conformational change of the *N*-terminal domain in the AGX1^{A229T} structure, is M218 shifted by 0.8 Å away from R169, weakening the Q112-R169-M218 interaction (Fig. 4F). Together, the data suggest considerable conformational changes induced by the A229T mutation.

Discussion

UAP1 is the last enzyme in the HBP, reversibly converting UTP and GlcNAc-1P to UDP-GlcNAc and pyrophosphate. The steady supply of cytoplasmic UDP-GlcNAc is not only required for normal cellular physiology [43–45], but also essential for development [28,31–33]. *O*-GlcNAcylation is one of the PTMs that depends on the steady supply of cytoplasmic UDP-GlcNAc. There is accumulating evidence that
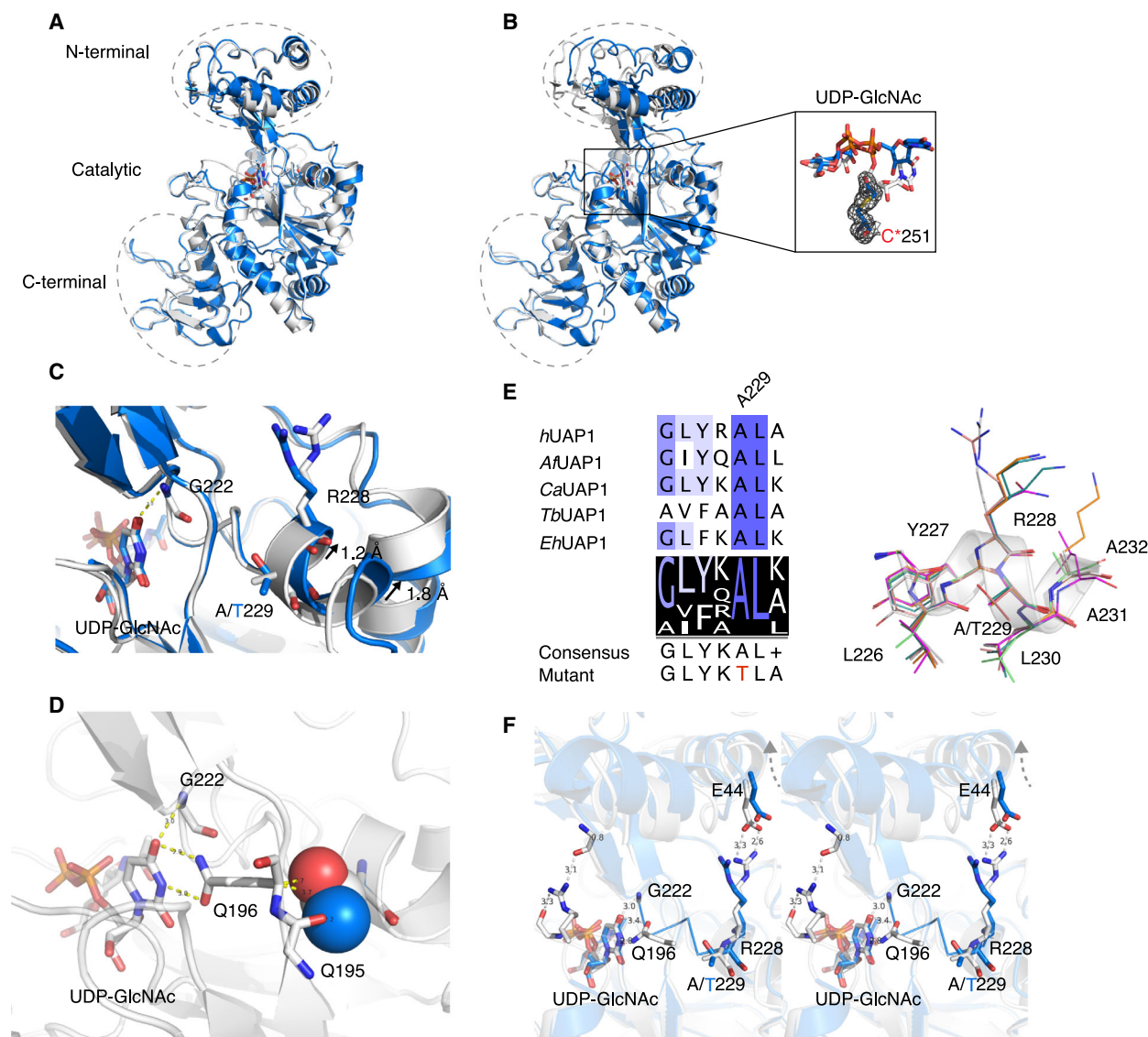
Fig. 4. Structural analysis of AGX1[A229T]. (A) Superposition of AGX1[A229T] A chain (blue cartoon) and AGX1[wt] A chain (grey cartoon; PDB: 1JV1 [40]). (B) Superposition of AGX1[A229T] B chain (blue cartoon) and AGX1[wt] A chain (grey cartoon; PDB: 1JV1 [40]). UDP-GlcNAc is highlighted as stick model in the active site of AGX1[wt] A chain (C_α blue) and AGX1[A229T] B chain (C_α grey). The electron density of carbamidomethylated Cys251 in AGX1[A229T] B chain is shown as black mesh contoured to 2.5 σ. (C) Structure superposition showing the difference between AGX1[A229T] A chain (blue cartoon) and the AGX1[wt] B chain (grey cartoon; PDB: 1JV1 [40]) at the point of the A229T mutation. R228, A/T229, G222 and UDP-GlcNAc are shown as a stick model. Structural shift induced by the mutation is indicated by the arrow. (D) Structure presentation of AGX1[wt] showing the interaction between UDP-GlcNAc, Q196 and A229. A229 is replaced by threonine and the side chain of the threonine is shown as sphere model. (E) Sequence and structure alignment of AGX1[A229T] and eukaryotic UAP1 orthologues at the point of the A229T mutation. The helix in AGX1[A229T] is shown in transparent cartoon representation with residues shown as a line model. (F) Stereoscopic view of the structural changes of AGX1 caused by the A229T mutation. The AGX1[A229T] A chain (blue cartoon) is superposed to the AGX1[wt] A chain (grey cartoon). The movement in the *N*-terminal domain caused by the A229T mutation is indicated by the arrow. The main chain of M218 and G222 and the side chain of E44, R228 and A/T229 are shown as stick models. The segment which connects A229, UDP-GlcNAc and the *N*-terminal domain is shown as blue ribbon.

dysregulation of *O*-GlcNAcylation caused by muta-
tions in the *OGT* gene is associated with a develop-
mental disorder, termed OGT-CDG [23]. This study
was aimed at identifying other genes that are required
for *O*-GlcNAcylation that are mutated in patients with
developmental disorders. In the recent DDD Study

[46], a *de novo*, heterozygous *UAP1* gene missense mutation (NM_001324114:c.G685A:p.A229T) was identified in a patient with abnormalities of cranium, skeleton and nervous system. In this study, we demonstrated that UAP1 A229 is stringently conserved in eukaryotic UAP1 orthologues and the UAP1 A229T mutation is predicted to be pathogenic by clinical bioinformatic tools. Biochemical characterization revealed that, compared to AGX1^{wt}, AGX1^{A229T} has decreased stability and activity in producing UDP-GlcNAc while pyrophosphorolysis activity is increased. The stability and activity changes of AGX1 caused by the A229T mutation could collectively result in decreased UDP-GlcNAc and *O*-GlcNAcylation levels *in vivo*, giving a phenotype in the patient that resembles aspects of OGT-CDG. However, experiments with CRISPR/Cas9 knock-in cell lines or animal model systems are required to investigate this.

Structural analysis revealed that the A229T mutation induced both local and global structural changes. R228 shifted by 1.2 Å when A229 is mutated to threonine, followed by the disruption of the hydrogen bonds formed between R228 and E44. An allosteric inhibitor targeting *Trypanosoma brucei* UAP1 was discovered binding to the same site where R228 and E44 interact in human UAP1 (Fig. S4), and it was proposed that the inhibition was achieved by stabilizing the *N*-terminal domain in a conformation that prevents UTP binding [47], suggesting that the R228-E44 interaction in AGX1 could be contributing to regulation of activity. The R228-E44 interaction contributes to the interaction between the *N*-terminal and catalytic domains of AGX1^{wt}, and disruption in the AGX1^{A229T} structure leads to a shift of the *N*-terminal domain away from the catalytic domain. Accompanied by the shift of the *N*-terminal domain, is the disruption of the hydrogen-bonding interactions established through Q112-R169-M218 between the *N*-terminal domain and the key catalytic loop. Given both the R228-E44 and the Q112-R169-M218 interactions are human AGX1 specific [40,68,69] and AGX1 is over ten times more catalytically efficient than its fungal orthologues [45], they could be one of the factors that contribute to the high efficiency of AGX1, and disruption of them together in the AGX1^{A229T} structure could thus explain the stability and activity differences between AGX1^{wt} and AGX1^{A229T}. Threonine is one of the common substitutions of alanine in nonsynonymous coding single nucleotide polymorphisms (ncSNPs) database [70], and alanine-to-threonine substitutions have been associated with human conditions, such as amyloid diseases [70] and PD [71]. It seems that the mutation intolerance of UAP1 A229 in the evolutionary context is not determined by the R228-A229 interaction, as R228 is substituted to alanine in plants (Fig. 2), but more dependent on the interaction between the side chain of A229 and the main chain of Q195/Q196 (Fig. 4D). Together with the observation in fly models that the *Drosophila* UAP1 orthologue is required for epithelial morphogenesis and nervous system development [35–37], the UAP1^{A229T} missense mutation that results in considerable structure shifts that underpin the observed stability and activity changes of AGX1^{A229T} could be a contributory factor to the developmental delay observed in the patient.

Apart from being required for *O*-GlcNAcylation, UDP-GlcNAc is also the precursor for glycosylphosphatidylinositol anchor synthesis, heparan sulphate synthesis and *N*-linked glycosylation. Dysregulation of these pathways has also been associated with developmental disorders [72]. For instance, *N*-glycosylation is a common post-translational modification in all three kingdoms of life that regulates protein stability, protein processing and function [73]. DPAGT1/GPT (Dolichol phosphate *N*-acetylglucosamine-phosphotransferase) catalyses the first committed step of *N*-linked glycosylation, using dolichol phosphate and UDP-GlcNAc as substrates [74]. Mutations in the *DPAGT1* gene, which result in alteration of either protein activity, stability or structure, have been identified as causing congenital disorders [75–78]. Mutations of all the genes upstream of *UAP1* in the HBP have been identified as causing developmental disorders, and decreases in *O*-GlcNAcylation levels are the common and predominant molecular phenotype observed in either patient-derived cells, mouse or fly models [28,31–34]. *N*-linked glycosylation levels were unchanged in cells carrying patient *GFAT* mutations [33] and in mouse embryonic fibroblasts carrying a *GNA1* gene disruption [30]; however, they were decreased in cells carrying patient *PGM3* mutations [33]. The pathways affected by the UAP1 A229T mutation require further dissection using appropriate models.

Data accessibility

The structure reported in this manuscript has been deposited in the Protein Data Bank with accession

code 6Z2F. All remaining data are included within the manuscript and Supporting information.

Author contributions

DMFA conceived the study XC and ATF performed experiments. XC, OR and ATF analysed and interpreted the data. XC and DMFA wrote the manuscript with input from all authors.

Funding

References

1 Haltiwanger RS, Holt GD and Hart GW (1990) Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins: identification of a uridine diphospho-N-acetylglucosamine:Peptide beta-N-acetylglucosaminyltransferase. J Biol Chem 265, 2563–2568.

2 Gao Y, Wells L, Comer FI, Parker GJ and Hart GW (2001) Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic β-N-acetylglucosaminidase from human brain. J Biol Chem 276, 9838–9845.

3 Cheung WD and Hart GW (2008) AMP-activated protein kinase and p38 MAPK activate O-GlcNAcylation of neuronal proteins during glucose deprivation. J Biol Chem 283, 13009–13020.

4 Taylor RP, Geisler TS, Chambers JH and McClain DA (2009) Up-regulation of O-GlcNAc transferase with glucose deprivation in HepG2 cells is mediated by decreased hexosamine pathway flux. J Biol Chem 284, 3425–3432.

5 Taylor RP, Parker GJ, Hazel MW, Soesanto Y, Fuller W, Yazzie MJ and McClain DA (2008) Glucose deprivation stimulates O-GlcNAc modification of proteins through up-regulation of O-linked N-acetylglucosaminyltransferase. J Biol Chem 283, 6050–6057.

6 Zhang X, Shu XE and Qian S-B (2018) O-GlcNAc modification of eIF4GI acts as a translational switch in heat shock response. Nat Chem Biol 14, 909–916.

7 Yi W, Clark PM, Mason DE, Keenan MC, Hill C, Goddard WA, Peters EC, Driggers EM and Hsieh-Wilson LC (2012) Phosphofructokinase 1 glycosylation regulates cell growth and metabolism. Science 337, 975–980.

8 Taub DG, Awal MR and Gabel CV (2018) O-GlcNAc signaling orchestrates the regenerative response to neuronal injury in Caenorhabditis elegans. Cell Rep 24, 1931–1938.e3.

9 Gambetta MC, Oktaba K and Muller J (2009) Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. Science 325, 93–96.

10 Oktaba K, Gutierrez L, Gagneur J, Girardot C, Sengupta AK, Furlong EE and Muller J (2008) Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. Dev Cell 15, 877–889.

11 Aichinger E, Villar CB, Farrona S, Reyes JC, Hennig L and Kohler C (2009) CHD3 proteins and polycomb group proteins antagonistically determine cell identity in Arabidopsis. PLOS Genet 5, e1000605.

12 Sauvageau M and Sauvageau G (2008) Polycomb group genes: keeping stem cell activity in balance. PLoS Biol 6, e113.

13 Shafi R, Iyer SP, Ellies LG, O'Donnell N, Marek KW, Chui D, Hart GW and Marth JD (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. Proc Natl Acad Sci USA 97, 5735–5739.

14 Yang YR, Song M, Lee H, Jeon Y, Choi E-J, Jang H-J, Moon HY, Byun H-Y, Kim E-K, Kim DH et al. (2012) O-GlcNAcase is essential for embryonic development and maintenance of genomic stability. Aging Cell 11, 439–448.

15 Liu F, Iqbal K, Grundke-Iqbal I, Hart GW and Gong CX (2004) O-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease. Proc Natl Acad Sci USA 101, 10804–10809.

16 Hsieh YL, Su FY, Tsai LK, Huang CC, Ko YL, Su LW, Chen KY, Shih HM, Hu CM and Lee WH (2019) NPGPx-mediated adaptation to oxidative stress protects motor neurons from degeneration in aging by directly modulating O-GlcNAcase. Cell Rep 29, 2134–2143.e7.

17 Marotta NP, Lin YH, Lewis YE, Ambroso MR, Zaro BW, Roth MT, Arnold DB, Langen R and Pratt MR (2015) O-GlcNAc modification blocks the aggregation and toxicity of the protein α-synuclein associated with Parkinson's disease. Nat Chem 7, 913–920.

18 Levine PM, Galesic A, Balana AT, Mahul-Mellier AL, Navarro MX, De Leon CA, Lashuel HA and Pratt MR (2019) α-Synuclein O-GlcNAcylation alters aggregation and toxicity, revealing certain residues as potential inhibitors of Parkinson's disease. Proc Natl Acad Sci USA 116, 1511–1519.

19 Pravata VM, Muha V, Gundogdu M, Ferenbach AT, Kakade PS, Vandadi V, Wilmes AC, Borodkin VS, Joss S, Stavridis MP et al. (2019) Catalytic deficiency of O-GlcNAc transferase leads to X-linked intellectual disability. Proc Natl Acad Sci USA 116, 14961–14970.

20 Selvan N, George S, Serajee FJ, Shaw M, Hobson L, Kalscheuer V, Prasad N, Levy SE, Taylor J, Aftimos S

et al. (2018) *O*-GlcNAc transferase missense mutations linked to X-linked intellectual disability deregulate genes involved in cell fate determination and signaling. *J Biol Chem* **293**, 10810–10824.

21 Vaidyanathan K, Niranjan T, Selvan N, Teo CF, May M, Patel S, Weatherly B, Skinner C, Opitz J, Carey J *et al.* (2017) Identification and characterization of a missense mutation in the *O*-linked β-*N*-acetylglucosamine (*O*-GlcNAc) transferase gene that segregates with X-linked intellectual disability. *J Biol Chem* **292**, 8948–8963.

22 Willems AP, Gundogdu M, Kempers MJE, Giltay JC, Pfundt R, Elferink M, Loza BF, Fuijkschot J, Ferenbach AT, van Gassen KLI *et al.* (2017) Mutations in *N*-acetylglucosamine (*O*-GlcNAc) transferase in patients with X-linked intellectual disability. *J Biol Chem* **292**, 12621–12631.

23 Pravata VM, Omelková M, Stavridis MP, Desbiens CM, Stephen HM, Lefeber DJ, Gecz J, Gundogdu M, Õunap K, Joss S *et al.* (2020) An intellectual disability syndrome with single-nucleotide variants in *O*-GlcNAc transferase. *Eur J Hum Genet* **28**, 706–714.

24 Milewski S, Gabriel I and Olchowy J (2006) Enzymes of UDP-GlcNAc biosynthesis in yeast. *Yeast* **23**, 1–14.

25 Mengin-Lecreulx D and van Heijenoort J (1994) Copurification of glucosamine-1-phosphate acetyltransferase and *N*-acetylglucosamine-1-phosphate uridyltransferase activities of *Escherichia coli*: characterization of the glmU gene product as a bifunctional enzyme catalyzing two subsequent steps in the pat. *J Bacteriol* **176**, 5788–5795.

26 Namboori SC and Graham DE (2008) Acetamido sugar biosynthesis in the euryarchaea. *J Bacteriol* **190**, 2987–2996.

27 Darabedian N, Gao J, Chuh KN, Woo CM and Pratt MR (2018) The metabolic chemical reporter 6-Azido-6-deoxy-glucose further reveals the substrate promiscuity of *O*-GlcNAc transferase and catalyzes the discovery of intracellular protein modification by *O*-glucose. *J Am Chem Soc* **140**, 7092–7100.

28 Senderek J, Müller JS, Dusl M, Strom TM, Guergueltcheva V, Diepolder I, Laval SH, Maxwell S, Cossins J, Krause S *et al.* (2011) Hexosamine biosynthetic pathway mutations cause neuromuscular transmission defect. *Am J Hum Genet* **88**, 162–172.

29 Engel AG, Shen XM, Selcen D and Sine SM (2015) Congenital myasthenic syndromes: pathogenesis, diagnosis, and treatment. *Lancet Neurol* **14**, 420–434.

30 Boehmelt G, Wakeham A, Elia A, Sasaki T, Plyte S, Potter J, Yang Y, Tsang E, Ruland J, Iscove NN *et al.* (2000) Decreased UDP-GlcNAc levels abrogate proliferation control in EMeg32-deficient cells. *EMBO J* **19**, 5092–5104.

31 Sassi A, Lazaroski S, Wu G, Haslam SM, Fliegauf M, Mellouli F, Patiroglu T, Unal E, Ozdemir MA, Jouhadi

Z *et al.* (2014) Hypomorphic homozygous mutations in phosphoglucomutase 3 (PGM3) impair immunity and increase serum IgE levels. *J Allergy Clin Immunol* **133**, 1410–1419.e13.

32 Stray-Pedersen A, Backe PH, Sorte HS, Mørkrid L, Chokshi NY, Erichsen HC, Gambin T, Elgstøen KBP, Bjørås M, Wlodarski MW *et al.* (2014) PGM3 mutations cause a congenital disorder of glycosylation with severe immunodeficiency and skeletal dysplasia. *Am J Hum Genet* **95**, 96–107.

33 Zhang Y, Yu X, Ichikawa M, Lyons JJ, Datta S, Lamborn IT, Jing H, Kim ES, Biancalana M, Wolfe LA *et al.* (2014) Autosomal recessive phosphoglucomutase 3 (PGM3) mutations link glycosylation defects to atopy, immune deficiency, autoimmunity, and neurocognitive impairment. *J Allergy Clin Immunol* **133**, 1400–1409.e5.

34 Mariappa D, Sauert K, Mariño K, Turnock D, Webster R, van Aalten DMF, Ferguson MAJ and Müller HAJ (2011) Protein *O*-GlcNAcylation is required for fibroblast growth factor signaling in *Drosophila*. *Sci Signal* **4**, ra89.

35 Humphreys GB, Jud MC, Monroe KM, Kimball SS, Higley M, Shipley D, Vrablik MC, Bates KL and Letsou A (2013) Mummy, A UDP-*N*-acetylglucosamine pyrophosphorylase, modulates DPP signaling in the embryonic epidermis of *Drosophila*. *Dev Biol* **381**, 434–445.

36 Schimmelpfeng K, Strunk M, Stork T and Klämbt C (2006) Mummy encodes an UDP-*N*-acetylglucosamine-dipohosphorylase and is required during *Drosophila* dorsal closure and nervous system development. *Mech Dev* **123**, 487–499.

37 Tonning A, Helms S, Schwarz H, Uv AE and Moussian B (2006) Hormonal regulation of mummy is needed for apical extracellular matrix formation and epithelial morphogenesis in *Drosophila*. *Development* **133**, 331–341.

38 Diekman AB and Goldberg E (1994) Characterization of a human antigen with sera from infertile patients1. *Biol Reprod* **50**, 1087–1093.

39 Wang-Gillam A, Pastuszak I and Elbein AD (1998) A 17-amino acid insert changes UDP-*N*-Acetylhexosamine pyrophosphorylase specificity from UDP-GalNAc to UDP-GlcNAc. *J Biol Chem* **273**, 27055–27057.

40 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.

41 Stokes MJ, Güther MLS, Turnock DC, Prescott AR, Martin KL, Alphey MS and Ferguson MAJ (2008) The synthesis of UDP-*N*-acetylglucosamine is essential for

bloodstream form *Trypanosoma brucei in vitro* and *in vivo* and UDP-*N*-acetylglucosamine starvation reveals a hierarchy in parasite protein glycosylation. *J Biol Chem* **283**, 16147–16161.

42 Zhang W, Jones VC, Scherman MS, Mahapatra S, Crick D, Bhamidi S, Xin Y, McNeil MR and Ma Y (2008) Expression, essentiality, and a microtiter plate assay for mycobacterial GlmU, the bifunctional glucosamine-1-phosphate acetyltransferase and *N*-acetylglucosamine-1-phosphate uridyltransferase. *Int J Biochem Cell Biol* **40**, 2560–2571.

43 Mengin-Lecreulx D and Van Heijenoort J (1993) Identification of the glmU gene encoding *N*-acetylglucosamine-1-phosphate uridyltransferase in *Escherichia coli*. *J Bacteriol* **175**, 6150–6157.

44 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.

45 Fang W, Du T, Raimi OG, Hurtado-Guerrero R, Urbaniak MD, Ibrahim AFM, Ferguson MAJ, Jin C and van Aalten DMF (2013) Genetic and structural validation of *Aspergillus fumigatus* UDP-*N*-acetylglucosamine pyrophosphorylase as an antifungal target. *Mol Microbiol* **89**, 479–493.

46 Deciphering Developmental Disorders Study, Fitzgerald TW, Gerety SS, Jones WD, van Kogelenberg M, King DA, McRae J, Morley KI, Parthiban V, Al-Turki S *et al.* (2015) Large-scale discovery of novel genetic causes of developmental disorders. *Nature* **519**, 223–228.

47 Urbaniak MD, Collie IT, Fang W, Aristotelous T, Eskilsson S, Raimi OG, Harrison J, Navratilova IH, Frearson JA, van Aalten DMF *et al.* (2013) A novel allosteric inhibitor of the uridine diphosphate *N*-Acetylglucosamine pyrophosphorylase from *Trypanosoma brucei*. *ACS Chem Biol*. **8**, 1981–1987.

48 Mok MT and Edwards MR (2005) Critical sources of error in colorimetric assay for UDP-*N*-acetylglucosamine pyrophosphorylase. *Anal Biochem* **343**, 341–343.

49 Ojida A, Takashima I, Kohira T, Nonaka H and Hamachi I (2008) Turn-on fluorescence sensing of nucleoside polyphosphates using a xanthene-based Zn (II) complex chemosensor. *J Am Chem Soc* **130**, 12095–12101.

50 Lee HS and Thorson JS (2011) Development of a universal glycosyltransferase assay amenable to high-throughput formats. *Anal Biochem* **418**, 85–88.

51 Battye TGG, Kontogiannis L, Johnson O, Powell HR and Leslie AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271–281.

52 Vagin A and Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22–25.

53 Murshudov GN, Vagin AA and Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.

54 Emsley P, Lohkamp B, Scott WG and Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.

55 Wang ZV, Deng Y, Gao N, Pedrozo Z, Li DL, Morales CR, Criollo A, Luo X, Tan W, Jiang N *et al.* (2014) Spliced X-box binding protein 1 couples the unfolded protein response to hexosamine biosynthetic pathway. *Cell* **156**, 1179–1192.

56 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.

57 Kumar S, Stecher G and Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* **33**, 1870–1874.

58 Waterhouse AM, Procter JB, Martin DMA, Clamp M and Barton GJ (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191.

59 Huang N, Lee I, Marcotte EM and Hurles ME (2010) Characterising and predicting haploinsufficiency in the human genome. *PLOS Genet* **6**, e1001154.

60 Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB *et al.* (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291.

61 Vaser R, Adusumalli S, Leng SN, Sikic M and Ng PC (2016) SIFT missense predictions for genomes. *Nat Protoc* **11**, 1–9.

62 Adzhubei I, Jordan DM and Sunyaev SR (2013) Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* **76**, 7.20.1–7.20.41.

63 Schwarz JM, Cooper DN, Schuelke M and Seelow D (2014) MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* **11**, 361–362.

64 Dai R, Geders TW, Liu F, Park SW, Schnappinger D, Aldrich CC and Finzel BC (2015) Fragment-based exploration of binding site flexibility in *Mycobacterium tuberculosis* BioA. *J Med Chem* **58**, 5208–5217.

65 Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ, Wasney GA, Yeung R, Arrowsmith C, Ball LJ, Berglund H *et al.* (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci USA* **103**, 15835–15840.

66 Alvarez S (2013) A cartography of the van der Waals territories. *Dalt Trans* **42**, 8617–8636.

67 Chou PY and Pasman GD (1974) Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. *Biochemistry* **13**, 211–222.

68 Maruyama D, Nishitani Y, Nonaka T, Kita A, Fukami TA, Mio T, Yamada-Okabe H, Yamada-Okabe T and Miki K (2007) Crystal structure of uridine-diphospho-*N*-acetylglucosamine pyrophosphorylase from *Candida albicans* and catalytic reaction mechanism. *J Biol Chem* **282**, 17221–17230.

69 Raimi OG, Hurtado-Guerrero R, Borodkin V, Ferenbach A, Urbaniak MD, Ferguson MAJ and van Aalten DMF (2020) A mechanism-inspired UDP-*N*-acetylglucosamine pyrophosphorylase inhibitor. *RSC Chem Biol* **1**, 13–25.

70 Podoly E, Hanin G and Soreq H (2010) Alanine-to-threonine substitutions and amyloid diseases: butyrylcholinesterase as a case study. *Chem Biol Interact* **187**, 64–71.

71 Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R *et al.* (1997) Mutation in the α-synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047.

72 Péanne R, de Lonlay P, Foulquier F, Kornak U, Lefeber DJ, Morava E, Pérez B, Seta N, Thiel C, Van Schaftingen E *et al.* (2018) Congenital disorders of glycosylation (CDG): Quo vadis? *Eur J Med Genet* **61**, 643–663.

73 Dell A, Galadari A, Sastre F and Hitchen P (2010) Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *Int J Microbiol* **2010**, 148178.

74 Rine J, Hansen W, Hardeman E and Davis RW (1983) Targeted selection of recombinant clones through gene dosage effects. *Proc Natl Acad Sci USA* **80**, 6750–6754.

75 Basiri K, Belaya K, Liu WW, Maxwell S, Sedghi M and Beeson D (2013) Clinical features in a large Iranian family with a limb-girdle congenital myasthenic syndrome due to a mutation in DPAGT1. *Neuromuscul Disord* **23**, 469–472.

76 Carrera IA, Matthijs G, Perez B and Cerdá CP (2012) DPAGT1-CDG: report of a patient with fetal hypokinesia phenotype. *Am J Med Genet A* **158A**, 2027–2030.

77 Dong YY, Wang H, Pike ACW, Cochrane SA, Hamedzadeh S, Wyszyński FJ, Bushell SR, Royer SF, Widdick DA, Sajid A *et al.* (2018) Structures of DPAGT1 explain glycosylation disease mechanisms and advance TB antibiotic design. *Cell* **175**, 1045–1058.e16.

78 Iqbal Z, Shahzad M, Vissers LELM, van Scherpenzeel M, Gilissen C, Razzaq A, Zahoor MY, Khan SN, Kleefstra T, Veltman JA *et al.* (2013) A compound heterozygous mutation in *DPAGT1* results in a congenital disorder of glycosylation with a relatively mild phenotype. *Eur J Hum Genet* **21**, 844–849.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Sequence alignment of eukaryotic UAP1s.

Fig. S2. Heatmap showing the RMSDs of superposition between the two chains of AGX1wt and AGX1A229T.

Fig. S3. Structural representation of AGX1^{wt} in complex with UDP-GlcNAc.

Fig. S4. Structural superposition of AGX1^{wt} and *Tb*UAP1 crystal structures.

Fig. S5. Development of the XBZ assay to measure the reverse reaction catalysed by AGX1^{A229T}.

Table S1. Analysis of AGX1^{wt} and AGX1^{A229T} carbamidomethylation by protein fingerprinting.