

Bioinformatic prediction of putative conveyers of O-GlcNAc transferase intellectual disability

Received for publication, April 5, 2022, and in revised form, July 12, 2022. Published, Papers in Press, July 19, 2022.
<https://doi.org/10.1016/j.jbc.2022.102276>

Conor W. Mitchell^{1,2}, Ignacy Czajewski² , and Daan M. F. van Aalten^{1,2,*} 

From the ¹Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark; ²Division of Cell and Developmental Biology, School of Life Sciences, University of Dundee, Dundee, United Kingdom

Edited by Robert Haltiwanger

Protein O-GlcNAcylation is a dynamic posttranslational modification that is catalyzed by the enzyme O-GlcNAc transferase (OGT) and is essential for neurodevelopment and postnatal neuronal function. Missense mutations in OGT segregate with a novel X-linked intellectual disability syndrome, the OGT congenital disorder of glycosylation (OGT-CDG). One hypothesis for the etiology of OGT-CDG is that loss of OGT activity leads to hypo-O-GlcNAcylation of as yet unidentified, specific neuronal proteins, affecting essential embryonic, and postnatal neurodevelopmental processes; however, the identity of these O-GlcNAcylated proteins is not known. Here, we used bioinformatic techniques to integrate sequence conservation, structural data, clinical data, and the available literature to identify 22 candidate proteins that convey OGT-CDG. We found using gene ontology and PANTHER database data that these candidate proteins are involved in diverse processes including Ras/MAPK signaling, translational repression, cytoskeletal dynamics, and chromatin remodeling. We also identify pathogenic missense variants at O-GlcNAcylation sites that segregate with intellectual disability. This work establishes a preliminary platform for the mechanistic dissection of the links between protein O-GlcNAcylation and neurodevelopment in OGT-CDG.

Pathogenic missense variants in O-GlcNAc transferase (OGT) give rise to OGT congenital disorder of glycosylation (OGT-CDG), a rare, clinically heterogeneous intellectual disability (ID) syndrome where the underlying molecular etiology is unknown. By filtering the O-GlcNAc proteome for putatively functional O-GlcNAc sites and O-GlcNAc sites mutated in ID patients, we have identified a subset of the O-GlcNAcome, the hypo-O-GlcNAcylation of which may give rise to OGT-CDG.

O-GlcNAcylation, the covalent attachment of mono-saccharide units of β -linked GlcNAc to the hydroxyl groups of protein serine/threonine (Ser/Thr) residues, is a dynamic and highly conserved posttranslational modification (PTM) occurring on thousands of nucleocytoplasmic and mitochondrial proteins (1). This dynamic PTM is catalyzed by the OGT (2, 3). OGT is a type B glycosyltransferase, composed of 13.5

N-terminal tetratricopeptide (TPR) repeats and a C-terminal glycosyltransferase domain (4). The N-terminal TPRs confer both substrate specificity and serve as a molecular hub for the assembly of multiprotein complexes (5). The C-terminal glycosyltransferase domain possesses both sugar transfer activity and a noncanonical protease activity, the latter being required for proteolytic maturation of the cell cycle regulator HCF-1 (6). The antagonistic enzyme to OGT, O-GlcNAc hydrolase, removes the O-GlcNAc modification (7).

OGT is most abundant in the brain (8), where O-GlcNAc mass spectrometric studies have identified thousands of O-GlcNAc-modified proteins involved in synaptic plasticity, excitatory neurotransmission, and neuronal activity-dependent gene transcription (9–11). However, the site-specific roles of O-GlcNAc in regulating neurodevelopmental and postnatal neuronal processes remains poorly understood, with the functional consequences of O-GlcNAcylation understood for only a handful of neuronal OGT substrates. Despite this, previous work has established the role of O-GlcNAc in modulating dozens of neurodevelopmental processes, including axonal growth (12), dendritic spine maturation (13), and synaptic vesicle dynamics (14). Furthermore, hippocampal-specific depletion of OGT abrogates long-term potentiation and excitatory synapse maturation, key processes underpinning learning and memory, potentially due to disrupted α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor trafficking (13). Additionally, O-GlcNAcylation is essential for dopaminergic neuron survival, with loss of OGT activity leading to massive apoptosis and near-abolition of dopamine release as a second order consequence (15). At the organismal level, O-GlcNAcylation is protective against neurodegeneration and age-related decline in learning and memory (16). Strikingly, pan-neuronal KO of OGT results in perinatal lethality in mice, which is immediately preceded by locomotor dysfunction and hyperphosphorylation of the Alzheimer's disease associated protein Tau (17). Taken together, O-GlcNAcylation is an essential PTM for neurodevelopment and neuronal function, though a comprehensive mechanistic understanding of how O-GlcNAc functions in the brain and developing central nervous system is lacking.

Recently, several groups have independently reported missense variants in the *OGT* gene that segregate with ID (18–21). IDs are a group of clinically and genetically

* For correspondence: Daan M. F. van Aalten, dmfvanaalten@dundee.ac.uk.

Putative conveyers of OGT-CDG

heterogenous neurodevelopmental disorders, collectively characterized by an intelligence quotient of less than 70, as well as compromised adaptive and social behavior (22). ID can result from mutations in one of over 700 genes and typically co-occurs with failure to reach key milestones in social, motor, or language development (known as developmental delay) (22). The ID syndrome which segregates with *OGT* missense variants, recently named the OGT-CDG, is symptomatically highly heterogenous (23). Patients typically present with ID, developmental delay, hypotonia, hearing/visual impairments, facial dysmorphism, long-thin fingers, clinodactyly, short stature, and eye abnormalities (23). However, additional conditions such as epilepsy and microcephaly are comorbid with OGT-CDG. This clinical heterogeneity strongly suggests that multiple signaling cascades and developmental processes are disrupted in OGT-CDG, further complicating therapeutic strategies and dissection of molecular etiology. Moreover, exactly how missense mutations in *OGT* give rise to OGT-CDG is not clear. Pathogenic *OGT* mutations have been reported in both the substrate specificity conferring TPRs (20, 21) and the glycosyltransferase domain, leading to divergent hypotheses regarding molecular etiology. For example, it has been hypothesized that *OGT* missense variants abrogate proteolytic maturation of HCF1, preventing normal cell cycle progression (23). However, *OGT* modifies >4000 substrates, and reduced *OGT* activity toward a single substrate cannot explain the clinical heterogeneity observed in OGT-CDG patients. Alternatively, it has been hypothesized that the missense variants reduce *OGT* stability, indirectly leading to reduced protein O-GlcNAcylation. While this has been observed for the *OGT*^{L254F} variant, where the TPR domain displays reduced thermal stability *in vitro* (24), this has not been reported for other OGT-CDG mutations. The majority of reported *OGT* missense variants map to the TPRs (20, 21, 23), leading to the hypothesis that selective loss of *OGT* interactors may give rise to OGT-CDG. However, no data have surfaced to date which substantiates this hypothesis. Instead, characterization of the OGT-CDG N648Y and N567K variants revealed that mutations in the *OGT* glycosyltransferase domain lead to lower global O-GlcNAcylation levels in mouse embryonic stem cells (mESCs) and *Drosophila melanogaster*, impair proteolytic maturation of HCF1, and strongly attenuate O-GlcNAcylation of the physiological *OGT* substrate Table 1 *in vitro* (18, 19). Collectively, these data indicate that hypo-O-GlcNAcylation of *OGT* substrates at particular developmental time points, due to

catalytic deficiency of *OGT*, may give rise to OGT-CDG. The identity of these *OGT* substrates, referred to in this manuscript as “phenotypic conveyers,” is unknown. Identification of these phenotypic conveyers would allow elucidation of the molecular etiology of OGT-CDG and further develop an understanding of the roles of O-GlcNAcylation in neurodevelopment and neuronal function through dissection of the site-specific roles of O-GlcNAc on these proteins.

One of the challenges in identifying phenotypic conveyers of the OGT-CDG phenotype is the lack of depth of information provided by O-GlcNAc site-mapping studies. O-GlcNAc proteomic studies have expanded the size of the O-GlcNAcome dramatically in the past 15 years (9–11, 25, 26). However, these same studies typically identify O-GlcNAc sites without subsequently predicting or attempting to ascribe functions to identified O-GlcNAc sites. This is crucial, as the presence of O-GlcNAc on a protein does not necessarily imply a functional consequence. For example, of the 11 O-GlcNAc sites identified on Keap1, only one—S104—modulates downstream NRF2 signaling in response to stress (27). Similarly, the O-GlcNAcylation stoichiometry of the methionine cycle regulator adenosyl homocysteinase (AHCY) is predominantly contributed by a single O-GlcNAc site, whereas the remaining three sites make negligible contributions to AHCY O-GlcNAcylation stoichiometry and lack an identifiable function (28). Consequently, although the O-GlcNAcome stands at over 4000 proteins, it is difficult to predict if and how hypo-O-GlcNAcylation of specific O-GlcNAcylated neuronal proteins would alter protein function and subsequently contribute to the OGT-CDG phenotypes.

In this study, a filter-based bioinformatics approach was used to predict the function of O-GlcNAc sites on neuronal proteins that have already been reported to be linked to ID. Utilizing O-GlcNAc site-mapping data in combination with sequence conservation, structural data, clinical databases, and literature review, we predict the existence of 38 functional O-GlcNAc sites across 22 neuronal proteins. In most cases, the role of the identified O-GlcNAc site is unknown. As these proteins are pathogenically mutated in reported ID/developmental delay (DD) syndromes and harbor putatively functional O-GlcNAc sites, their hypo-O-GlcNAcylation may contribute to OGT-CDG etiology. Additionally, by screening the ClinVar database against a recently compiled O-GlcNAc site database, we identify eight missense variants at O-GlcNAcylation sites in

Table 1
Missense variants at conserved sites of O-GlcNAcylation identified in patients with ID/DD

Gene	Variant	Patient condition	Pathogenicity	PolyPhen score
<i>PUM1</i>	S802F	Spinocerebellar ataxia 47	Uncertain significance	0.921
<i>TUBB2B</i>	S172P	Cortical dysplasia, complex, with other brain malformations	Pathogenic	
<i>TUBB2B</i>	S172L	Cortical dysplasia, complex, with other brain malformations	Likely pathogenic	
<i>DOCK7</i>	S190N	Epileptic encephalopathy, early infantile, 23	Uncertain significance	0.995
<i>DOCK7</i>	S190G	Epileptic encephalopathy, early infantile, 23	Uncertain significance	0.989
<i>ATRX</i>	S594C	Alpha thalassemia-X-linked intellectual disability	Uncertain significance	0.993
<i>HCF1</i>	T556M	Mental retardation 3, X-linked	Uncertain significance	0.959
<i>KMT2A</i>	S1858I	Wiedemann–Steiner syndrome	Uncertain significance	0.855

The ClinVar database was screened for pathogenic missense mutations at O-GlcNAc sites reported in a recently published O-GlcNAc database (see (46)). Patient condition and the associated pathogenicity of the missense mutation are summarized. Where the significance of the missense mutation is uncertain, PolyPhen HumVar scores are provided, where default cut-offs of 0.15 for “likely benign” and 0.85 for “probably damaging” are applied.

ID/DD patients, further suggesting a convergence of the O-GlcNAc proteome and ID etiology.

Results and discussion

A filter-based approach identifies OGT-CDG candidate conveyers

To identify candidate conveyers of the OGT-CDG phenotypes, we used a filter-based approach integrating O-GlcNAc site-mapping data, available structural data, clinical databases, and the literature. O-GlcNAc site-mapping data were pooled manually from 14 studies that reported the O-GlcNAcomes of various cell lines (HEK293 (25), mESCs (29), HeLa (30), T-lymphocytes (26)) and tissue extracts (rat brain (31), mouse brain (32), and mouse synaptosomes (9, 11, 33, 34)). Next, we extracted O-GlcNAcylated proteins, where missense mutations in the corresponding coding gene were reported to give rise to ID/DD, using the list of ID/DD-associated genes compiled by Vissers *et al.* (22) and the DECIPHER database (35) as reference points. The DECIPHER database reports on the pathogenicity and associated phenotype of variants in a given gene, and proteins passed this filter if their corresponding gene was the sole pathogenic variant in a patient with ID or DD. About 105 O-GlcNAcylated proteins, segregating with X-linked ID passed this first filter (see File S1 for list of O-GlcNAcylated X-linked ID proteins). Sequence

conservation of O-GlcNAc sites was assessed from sequence alignments of higher eukaryotic orthologues (human, mouse, rat, rabbit, and sheep), leading to rejection of candidates where the O-GlcNAcylated Ser/Thr was poorly conserved (absent in ≥ 1 species). Subsequently, the available literature was used to investigate whether such conserved O-GlcNAc sites resided in functionally important domains or motifs of the protein. Examples of such functionally important domains or motifs include protein:protein interaction domains, protein:nucleic acid interaction domains, enzymatic active sites, or metal-binding sites. These filters reduced the number OGT-CDG candidate conveyers in the whole O-GlcNAcome to 22 candidate proteins (Fig. 1).

We further segregated O-GlcNAc sites on our 22 candidate proteins into two groups. Crystallographic studies of OGT in complex with substrate peptides have suggested that the active site and TPR domain of OGT accommodate disordered regions of proteins adopting an extended conformation (36–38). However, several examples exist of O-GlcNAc sites residing in ordered regions of the protein fold and these studies do not yet explain why (11, 28, 36). Importantly, Ser/Thr residues in ordered secondary structures would be inaccessible to the OGT active site. An example of this is AHCY, which is activated by O-GlcNAcylation at Thr136, despite this site residing in an α -helix that is theoretically inaccessible to the OGT active site (28). Recently, it has been discovered that

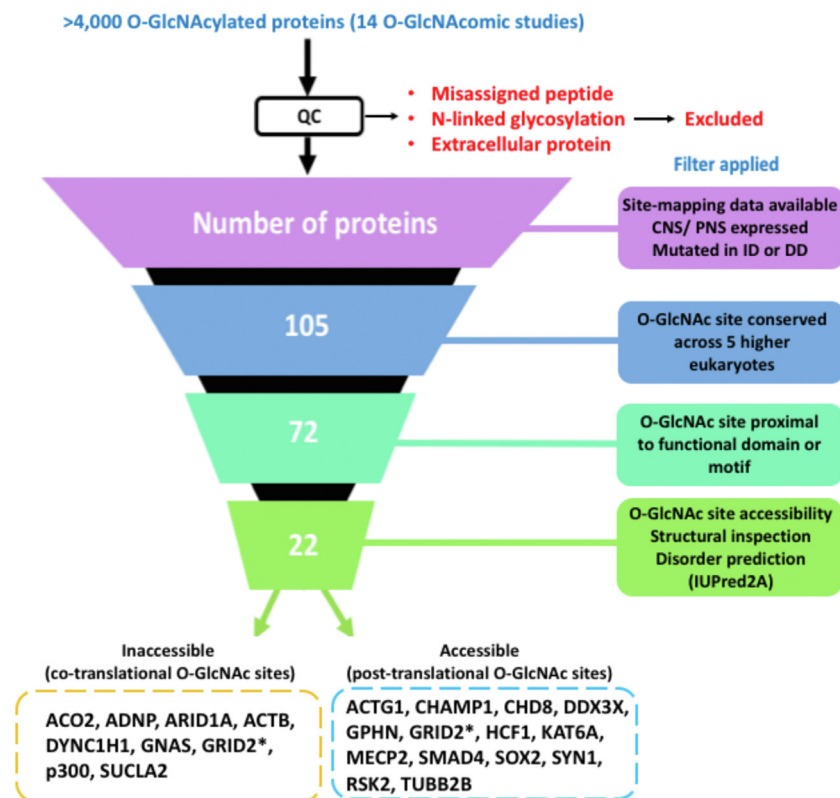


Figure 1. Illustration of the rational, filter-based approach for identification of candidate conveyers of the OGT-CDG phenotypes. Site-mapping data for O-GlcNAcylated proteins identified in 14 O-GlcNAc mass spectrometric (O-GlcNAcomics) studies of mouse, rat, and human CNS tissue samples and cell lines were pooled and applied to the above illustrated, filter-based approach (see Experimental procedures). The relevant filters applied at each level are annotated. GRID2 (*) was predicted to harbor both cotranslational and posttranslational O-GlcNAc sites on the basis of IUPred disorder prediction. CDG, congenital disorder of glycosylation; CNS, central nervous system; OGT, O-GlcNAc transferase.

Putative conveyers of OGT-CDG

O-GlcNAcylation can also occur cotranslationally on disordered peptidyl-tRNAs, as they emerge from the ribosome (39). It is possible that O-GlcNAc sites in inaccessible, ordered regions of the protein fold represent a subpopulation of O-GlcNAc sites that are cotranslationally modified by OGT. We therefore further segregated the list of candidate proteins/sites into those with accessible (in surface-exposed, disordered protein regions) and inaccessible (within sequestered, ordered regions of the protein fold) O-GlcNAc sites. To achieve this, we used available Protein Data Bank (PDB) structures of proteins or where such data were unavailable, the disorder predictor IUPred2A (40). Overall, our filter-based approach (illustrated in Fig. 1) suggests that of the 22 candidate conveyers of the OGT-CDG phenotypes, 13 are posttranslationally O-GlcNAcylated, eight cotranslationally O-GlcNAcylated, and one candidate (GRID2) harbors both cotranslational and posttranslational O-GlcNAc sites. For summaries of each OGT-CDG candidate and details regarding conserved O-GlcNAc site(s) of interest, see Table S1.

Of the 39 candidate O-GlcNAc sites, the majority (~58%) were predicted to be functional based on their localization to a protein:protein binding interface. Four candidate O-GlcNAc sites were either reported as being phosphorylated or SUMOylated or as proximal (± 10 residues) to a functional phosphorylation/SUMOylation site (Fig. 2A). The remaining sites were either located in enzyme active sites (RSK-2, GNAS), a regulatory domain/motif (KAT6A, SMAD4) or a protein:nucleic acid interface (ADNP, ARID1A; Fig. 2A). Classification of OGT-CDG candidate conveyers by protein class, using PANTHER (Protein ANalysis THrough Evolutionary Relationships) (42), showed a substantial portion of candidates were cytoskeletal or microtubule-associated proteins (19%; TUBB2B; ACTB; ACTG1; DYNC1H1; Fig. 2B). Other classes of protein present in our candidate list were transcriptional regulators (SOX2, SMAD4), chromatin binding or regulatory proteins (P300, KAT6A), and metabolic enzymes (ACO2; SUCLA2). Transmembrane signaling receptors (GRID2), protein modifying enzymes (RSK-2), and nucleic acid metabolism enzymes (MECP2) were also identified (Fig. 2B). Seven candidates were not annotated as belonging to a single protein class. Overall, our OGT-CDG candidate conveyers represent a diverse pool of proteins.

In addition to conserved O-GlcNAc sites residing in functional domains or regions of ID/DD-associated proteins, we also investigated whether poorly conserved O-GlcNAc sites on ID/DD-associated proteins may reside in interesting regions of a given protein. Toward this end, we analyzed whether the 105 ID/DD-associated O-GlcNAc proteins identified after the first filter harbored poorly conserved O-GlcNAc sites (O-GlcNAc Ser/Thr not conserved in ≥ 1 species) in functional domains/regions of the protein. Four hits (NDUFS3, aprataxin, PDHA1, and ankyrin G) were identified due to O-GlcNAc sites either localizing to mitochondrial targeting sequences (PDHA1, NDUFS3), the interaction site for PARP-1 (aprataxin), or an isoform-specific domain associated with regulation of protein localization (ankyrin G). These lower confidence hits are summarized in Table S2.

We subsequently used gene ontology (GO) analysis to assess possible convergence of candidate conveyers on a small set of biological processes and signaling pathways. Most of the candidates were annotated as being involved broadly in development and neurogenesis (Fig. 2C). Intriguingly, we also observed that some of our candidate conveyers were annotated as forming part of the stress response (DDX3X, DYNC1H1; see discussion on DDX3X later), a process known to be regulated by O-GlcNAcylation (41). Additional biological processes included cell cycle regulation, chromatin organization, regulation of gene expression, and protein modification (acetylation/ubiquitination, Fig. 2C). For all these biological processes, previous work has strongly suggested a functional role for O-GlcNAcylation in their regulation (6, 27, 39, 41–43). We further carried out pathway analysis using PANTHER and found that the candidate conveyers were involved in a number of signaling pathways. Of particular interest, we noted that Dpp signaling, Wnt signaling, and synaptic vesicle trafficking were hits (Fig. 2D). In *D. melanogaster*, Dpp (BMP4 in vertebrates) signaling is negatively regulated at the level of the bone morphogenic protein (BMP) receptor Sax by O-GlcNAcylation (44), and the stability of Wnt signaling nodes CK2 and β -catenin are negatively and positively regulated by O-GlcNAc, respectively (45, 46). O-GlcNAcylation of Synapsin I, which is also identified as a candidate conveyer in this study (Table S1), regulates presynaptic vesicle dynamics and the size of the reserve pool of synaptic vesicles (14). Thus, the OGT-CDG candidate conveyers identified here are involved in diverse biological processes and signaling pathways, many of which are already known to be regulated by O-GlcNAcylation. Future work, characterizing the function of O-GlcNAc on these candidate proteins, guided by the hypotheses we have presented, may reveal additional levels of O-GlcNAc-dependent regulation in these processes. Later, we discuss selected examples of candidate conveyers and their associated O-GlcNAc sites of interest.

O-GlcNAcylation of cytoskeletal proteins suggests that dysregulated cytoskeletal dynamics may contribute to OGT-CDG pathology

Precise regulation of cytoskeletal dynamics is essential for the ability of initially spherical cell soma to give rise to distinct somatodendritic and axonal compartments (47). This process, termed neuronal polarization, is essential for the correct flow of electrochemical impulses and the trafficking of RNA, ribosomes, and ion channels to distal synapses (48). Indeed, disruption of cytoskeletal dynamics leads to neurodevelopmental disorders such as lissencephaly and cortical dysplasia (49). Regarding cytoskeletal proteins, our bioinformatic analysis predicted O-GlcNAc sites on both β -tubulin (isoform TUBB2B) and actin (isoforms ACTB1 and ACTG1) to be important for regulating their respective proteins' functions (see later).

In our bioinformatic analysis, we identified S172 as a highly conserved β -tubulin O-GlcNAc site. S172L and S172P mutations segregate with cortical dysplasia, complex, with other

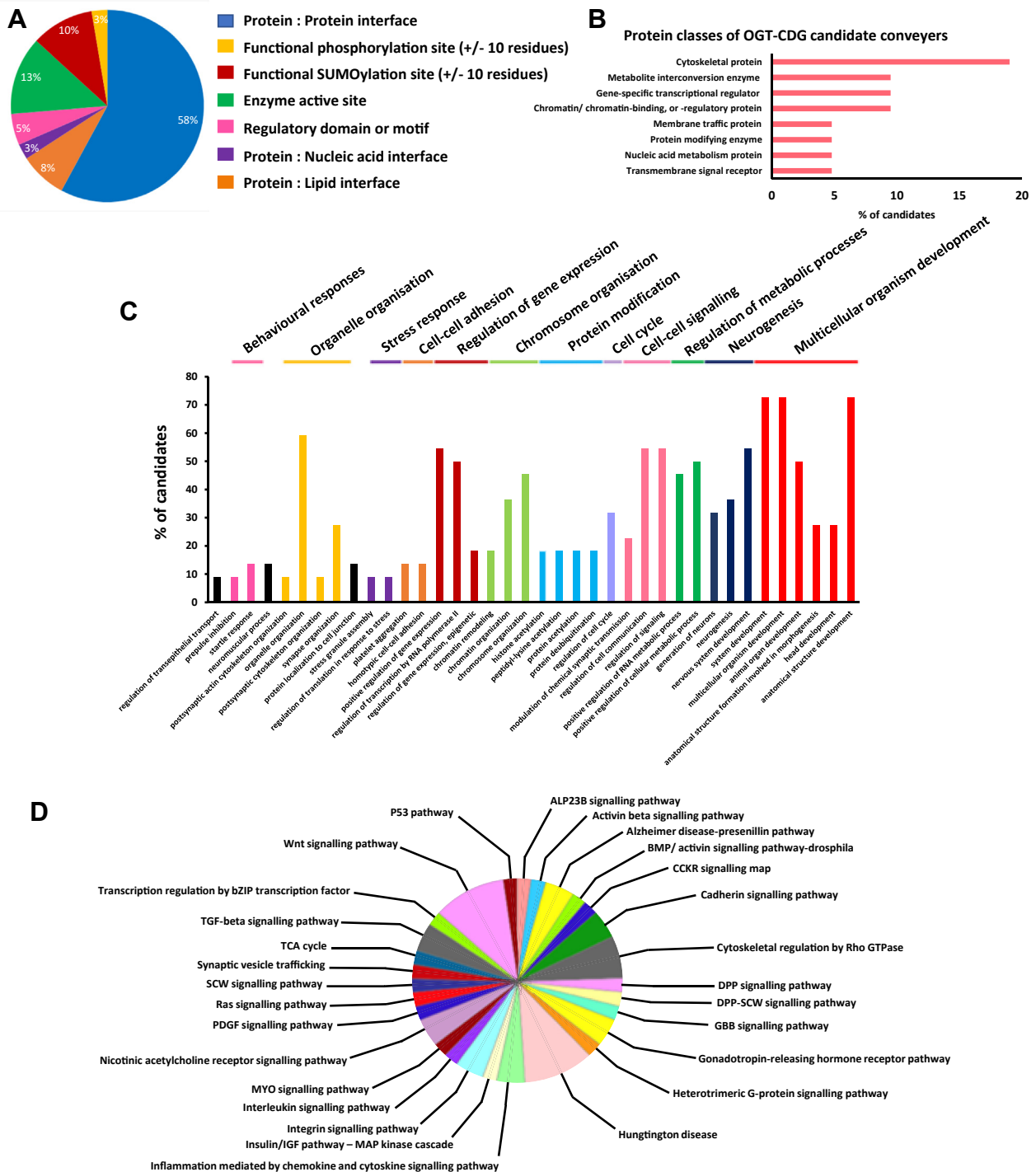


Figure 2. Classification of OGT-CDG candidate conveyers by protein class, positioning of O-GlcNAc site(s), and convergent signaling pathways and biological processes. *A*, percentage of O-GlcNAc sites on OGT-CDG candidate conveyers, which were predicted to be functional based on either: (I) present on a protein:protein interface; (II) adjacent (± 10 residues) from another PTM site, which has been validated as functional in previous work; (III) within or directly abutting against an enzyme active site; and (IV) present within a known regulatory region. *B*, stratification of OGT-CDG candidate conveyers by class of protein using PANTHER, UniProt accessions of human homologs were used (see [File S1](#)). *C*, classification of biological processes involving OGT-CDG candidate conveyers using the Gene Ontology (GO) resource. GO terms were clustered manually into common groups and highlighted. *D*, signaling pathways involving OGT-CDG candidate conveyers were extracted using PANTHER analysis of Gene IDs, which were manually retrieved from NCBI. CDG, congenital disorder of glycosylation; OGT, O-GlcNAc transferase.

brain malformations, seven in several ID patients (49) (CDCBM7; See section “[Identification of pathogenic missense variants at O-GlcNAcylation sites](#)” and [File S2](#)). Additionally,

S172 resides on a loop that forms part of the GTPase active site ([Fig. 3A](#)) (50, 51). Phosphorylation of S172 by CDK1, during mitosis, prevents GTP/GDP exchange, presumed to be due to

Putative conveyers of OGT-CDG

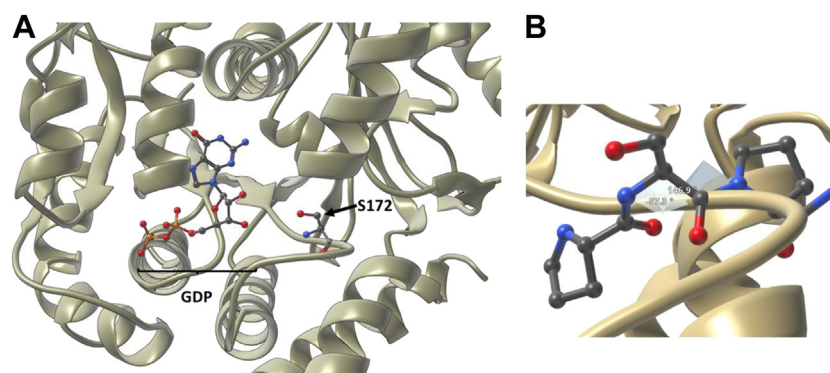


Figure 3. The β -tubulin O-GlcNAc site S172 and associated backbone angles. A, crystal structure of β -tubulin from Lowe *et al.* (PDB entry: 1JFF (51) with S172 and bound GDP highlighted as ball and stick depictions. Note how the S172 hydroxyl points directly toward the bound GDP in the GTPase active site. B, measured dihedral angles for S172 ($\Phi = -57^\circ$, $\Psi = 147^\circ$). P171, S172, and P173 are highlighted as ball and stick depictions.

hydrogen bonding interactions between the phosphate and GDP preventing GDP dissociation from the active site (50). The presence of a bulky GlcNAc moiety in the GTPase active site would likely sterically hinder GTP binding, thus reducing the pool of GTP-bound tubulin and inhibiting tubulin polymerization. Interestingly, Ji *et al.* (52) previously observed that O-GlcNAcylated tubulin is exclusively monomeric, supporting the hypothesis that tubulin O-GlcNAcylation prevents microtubule assembly, though the functional sites of tubulin O-GlcNAcylation were not identified in this study. If S172 O-GlcNAcylation does regulate tubulin polymerization indirectly, through modulation of GTP binding affinity to tubulin, then hypo-O-GlcNAcylation of tubulin in the context of OGT-CDG could lead to increased rates of tubulin polymerization. This increased rate of tubulin polymerization, in-turn, could deplete the pool of monomeric tubulin, altering the structural plasticity of the microtubules in neurons and potentially preventing morphological changes in the dendrites and axons that are crucial for neuronal polarization (47). This hypothesis could be investigated by an *in vitro* polymerization assay, comparing TUBB2B^{WT} and TUBB2B^{S172A} polymerization rates (53).

In addition to the microtubule networks in neurons, the actin cytoskeleton plays an essential role in regulating neuronal polarization, and fine-tuning actin cytoskeleton dynamics and function severely impacts both axon differentiation (54) and growth cone motility (55) prior to synaptogenesis. Pharmacological inhibition of actin polymerization induces spontaneous axon differentiation from immature neurites (54). Furthermore, heterozygous actin loss-of-function variants have been reported in patients displaying ID, developmental delay, internal organ malformation, and growth retardation (56). The highly conserved actin O-GlcNAc site S365 lies on the interface between actin and its regulatory protein profilin (57). The S365 hydroxyl group points toward a E364-K126 salt bridge on the actin:profilin interface, with the S365 hydroxyl and E364 carboxylate only 3.3 Å apart (see PDB entry 2BTF (57)). A bulky O-GlcNAc moiety on S365 could therefore disrupt formation of this salt bridge to inhibit actin:profilin binding. Profilins stimulate the exchange of GDP for GTP on actin, thus regulating the size of the pool of polymerizable

GTP-actin monomers (58). Knockdown of profilin IIa in mice leads to increased neurite outgrowth, ostensibly due to reduced actin polymerization and reduced F-actin density (59). Whether actin S365 O-GlcNAcylation inhibits actin polymerization through reduced affinity of actin for profilin and whether hypo-O-GlcNAcylation of actin at S365 alters actin polymerization in the context of OGT-CDG, will require further investigation.

In addition to S365, our bioinformatic analysis of the O-GlcNAcome identified S265 as a previously reported, conserved O-GlcNAc site on actin, which resides in a hydrophobic loop (residues 264–273). In the Holmes model of F-actin, this hydrophobic loop inserts into a pocket formed by an adjacent F-actin protomer in a manner analogous to a “socket” and a “plug,” with the loop constituting the plug and the socket formed by hydrophobic residues in the adjacent promoter (60). Holmes *et al.* (60) hypothesized that loss of this interaction between adjacent actin protomers would likely severely disrupt actin polymerization, as other electrostatic interactions along the length of F-actin may not be sufficient to compensate for loss of the hydrophobic interactions contributed by residues 264 to 273. The presence of O-GlcNAc on S265 therefore presents an interesting target for study and suggests that O-GlcNAc may regulate actin polymerization both directly (through disruption of actin protomer assembly) and indirectly (through reduced affinity for profilin).

O-GlcNAcylation of DDX3X implicates dysregulated RNA metabolism in OGT-CDG etiology

The filter approach identified a member of the DEAD box RNA helicase family—DEAD box RNA helicase 3 X-linked (DDX3X)—as a putative conveyer of the OGT-CDG phenotype. DDX3X is an ATP-dependent RNA helicase, which assembles into trimers capable of unwinding structured mRNA 5' UTRs (61, 62). This local strand separation is essential for docking and scanning of mRNA for translation initiation sites, and DDX3X-mediated RNA unwinding is crucial for the translation of cell cycle progression genes (63) and components of the Rac1-PKA signaling axis required for neurite outgrowth (See Fig. 4) (64). Furthermore, DDX3X

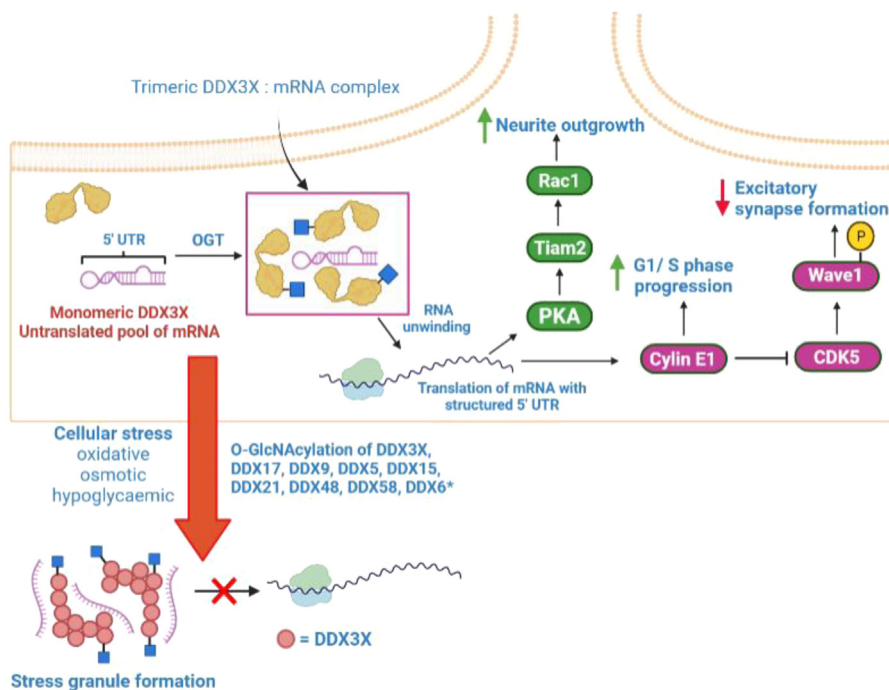


Figure 4. The DDX3X O-GlcNAc sites S584 and S588, and hypothesized functions of DDX3X O-GlcNAcylation in RNA metabolism and neurodevelopment. Illustrated summary of hypothesized function(s) of DDX3X O-GlcNAcylation. Briefly, DDX3X O-GlcNAcylation in the CTE at residues S584 and S588 may promote formation of functional DDX3X trimers. Functional DDX3X trimers subsequently promote the translation of signaling nodes in the Rac1/PKA and cyclin E1/CDK5 signaling axes. Additionally, in response to cellular stress, DDX3X and at least nine other DEAD-box RNA helicases (*) may be O-GlcNAcylated as part of the stress response. CTE, C-terminal extension.

knockdown in bone marrow-derived macrophages reduces stress granule formation in response to oxidative stress by ~50% (65). Stress granules are translationally repressive aggregates of ribonucleoproteins formed in response to cellular stresses such as glucose deprivation, osmotic stress, or oxidative stress to protect untranslated mRNAs from damage (66). Curiously, OGT knockdown in U2OS cells also abolishes stress granule formation following arsenite-induced osmotic stress (41), and mass spectrometric studies have identified at least nine DEAD box RNA helicase as being O-GlcNAcylated (26). However, exactly how OGT and stress granule formation are mechanistically linked is presently unknown (see Fig. 4 for illustrated hypothesis). Furthermore, missense mutations in the *ddx3x* gene are linked with a clinically heterogeneous ID syndrome comprising of, though not limited to, developmental delay, hypotonia, corpus callosum hypoplasia, epilepsy, behavioral problems, microcephaly, and abnormal gait (67).

We identified DDX3X as a candidate conveyer of OGT-CDG based on the presence of two highly conserved O-GlcNAc sites in the DDX3X C-terminal extension (CTE)—S584 and S588. The CTE is a highly conserved 26-residue stretch that is essential for DDX3X RNA unwinding activity (61). Deletion of this same region in the *Saccharomyces cerevisiae* ortholog Ded1p abolishes Ded1p trimerization (62), suggesting that the DDX3X CTE is required for association of DDX3X into functional trimers. Additionally, the CTE forms the majority of the region required for interaction with the mRNA export receptor TAP (residues 582–631) (68). The presence of two highly conserved O-GlcNAc sites in the CTE suggests that O-GlcNAcylation of DDX3X could modulate

DDX3X trimerization and/or RNA unwinding activity. Alternatively, DDX3X O-GlcNAcylation could modulate nuclear mRNA export by altering the affinity of DDX3X for TAP. Both S584, S588, and the surrounding sequence are highly conserved in eukaryotes despite the region being intrinsically disordered (see File S1 for IUPred2A scores), further suggesting these residues are important for regulating DDX3X function.

Hypo-O-GlcNAcylation of transcription factors may lead to transcriptomic dysregulation in OGT-CDG

The link between O-GlcNAcylation and regulation of eukaryotic gene expression has been a source of considerable interest, with several components of the transcription/translation machinery being O-GlcNAc modified (43, 69, 70). For example, O-GlcNAcylation of RNA polymerase II and eIF4G modulates transcription and translation, respectively (43, 70). Of the candidates annotated as signaling nodes in neurogenesis (see GO analysis, Fig. 2C), the transcriptional regulator Methyl CpG binding protein 2 (MECP2), is a particularly interesting hit. MECP2 mutations are the predominant cause of Rett's syndrome, an autistic spectrum disorder characterized by postnatal developmental regression, loss of motor skills, dystonia, loss of language/communication skills, and stereotypic hand movements (71). In resting neurons, MECP2 represses transcription at loci containing methylated CpG dinucleotides. This transcriptional repression is mediated *via* the scaffolding function of the MECP2 transcriptional repression domain, which recruits mSin3A, N-CoR, and c-Ski to silence transcription (72). Following neuronal depolarization, an unknown

Putative conveyers of OGT-CDG

kinase phosphorylates MECP2 at S423 and S426, leading to dissociation of MECP2 from methylated loci and up-regulated expression of the neurotrophin *BDNF* (73). S426 is also a conserved site of O-GlcNAcylation on MECP2. Mice expressing a MECP2^{S421A;S424A} (S423 and S426 in human) knock-in display enhanced hippocampal-dependent learning and memory in a Morris water maze test and altered gene expression patterns (73). MECP2 O-GlcNAcylation has already been reported to be antagonistic to phosphorylation, with glucosamine supplementation elevating O-GlcNAcylation stoichiometry and reducing total MECP2 phosphorylation (32). However, the extent to which MECP2 S426 O-GlcNAcylation antagonizes neuronal depolarization-dependent changes in MECP2 phosphorylation and *BDNF* expression has yet to be investigated (See Fig. 5 for an illustrated summary of the previous discussion). Furthermore, it remains to be investigated whether hypo-O-GlcNAcylation of S426 in mature neurons dysregulates neuronal depolarization-dependent gene expression in the context of OGT-CDG.

In addition to the genetic landscape of differentiated neurons, disrupted gene expression patterns early in development could also contribute to OGT-CDG etiology. SOX2, a key pluripotency marker in embryonic stem cells (ESCs), was identified as a candidate conveyer of the OGT-CDG phenotype on account of the proximity of the O-GlcNAc site S246 to the SUMOylation site K247. SUMOylation of K247 reduces SOX2-dependent gene expression from the *Fgf4* promoter due to reduced SOX2 DNA binding (74). Intriguingly, work by

Myers *et al.*(75) found that mutagenesis of S246 to alanine in mESCs altered SOX2 genomic occupancy and increased SOX2 interaction with PARP1 (see later). Importantly, SOX2 loss-of-function variants give rise to an ID syndrome characterized by anophthalmia/microphthalmia, as well as developmental delay and increased risk of intracranial teratomas, indicating that hypo-O-GlcNAcylation of this essential pluripotency factor could contribute to OGT-CDG (76).

During retinoic acid-induced ESC differentiation, Erk1/2 phosphorylates PARP1 (77). This phosphorylation event stimulates auto-PARYlation of PARP1 and promotes PARP1-SOX2 complex formation (78). In this complex, SOX2 is unable to bind Oct4-Sox enhancers and upregulate expression of key pluripotency genes including *Nanog*, *Oct4*, *Fgf4*, and *Sox2*, thus stimulating differentiation (78). These data are consistent with reports that SOX2 O-GlcNAcylation decreases with differentiation, and loss of O-GlcNAc on SOX2 increases PARP1 interaction, leading to reduced expression of genes required for maintenance of pluripotency (75). Whether PARP1-SOX2 complex formation is reduced during early stages of embryogenesis in OGT-CDG patients or in previously generated ESC models of the N567K and N648Y OGT mutations, has yet to be investigated. If complex stoichiometry is increased prematurely during embryogenesis, then premature ESC differentiation and dysregulated SOX2-dependent gene expression could contribute to the OGT-CDG phenotype.

In addition to transcription factors, a member of the histone acetyl transferase family of transcriptional

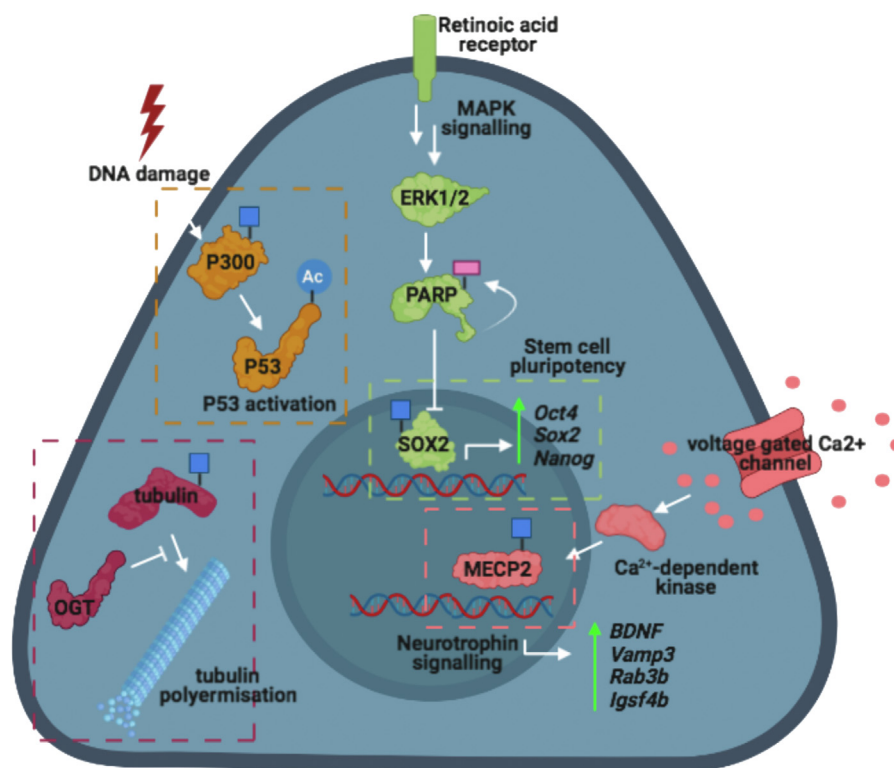


Figure 5. Summary of the proposed functions of O-GlcNAc on selected examples of OGT-CDG candidate conveyers. Hypothesized functions of O-GlcNAc on the OGT-CDG candidate conveyers are illustrated here. OGT-CDG candidate conveyers are annotated with a *blue square*, highlighting their O-GlcNAcylation status. Where expression of specific genes is expected to be affected by candidate O-GlcNAcylation status, these same genes are listed vertically in *italics* along with the expected change in expression following O-GlcNAcylation of the corresponding candidate protein. CDG, congenital disorder of glycosylation; OGT, O-GlcNAc transferase.

coregulators—P300—was identified as a candidate conveyer of OGT-CDG in our analysis. P300 acetylates lysine residues on histones to facilitate chromatin unwinding and upregulate gene expression. During DNA damage, P300 is recruited to P53, where P300 upregulates P53-dependent transcription *via* histone acetylation and acetylation of five C-terminal lysine residues on P53 (79). The latter event activates the DNA-binding activity of P53 (79). P300 mutations are the second most common cause of Rubinstein-Taybi syndrome, which manifests in mild-to-moderate ID, growth delay, skeletal abnormalities, and facial dysmorphism (80). This study identified S1734 as a conserved O-GlcNAc site residing on the interface of the P300 Taz2 domain with the P53 N-terminal transactivation domain (TAD) (81). The P300 Taz2–P53 TAD interaction is predominantly mediated by hydrophobic contacts, though electrostatic and polar interactions between several arginine, glutamine, and serine residues also contribute to complex formation (81). S1734 forms hydrogen bonds to both T18 and Q17 of P53, stabilizing the complex. Phosphorylation of T18 on P53 increases the affinity of the Taz2–TAD complex 11-fold, indicating this residue is particularly important for modulating the affinity of P53 for P300 (81). The presence of a bulky GlcNAc moiety could be hypothesized to act antagonistically to T18 phosphorylation, by inducing a steric clash between S1734 and T18/Q17 on the Taz2–TAD interface. This in turn would likely exert negative downstream effects on P53-dependent transcriptional activation due to reduced P300 recruitment and subsequent reductions in the levels of histone acetylation. Overall, P300 O-GlcNAcylation at S1734 may negatively regulate P300–P53 complex formation, as indicated by prior structural studies showing a close proximity of the S1734 hydroxyl group to the regulatory P53 phospho-site T18 and Q17 (See Figure 5 for an illustrated summary of the previous discussion). Hypo-O-GlcNAcylation of this site in the context of OGT-CDG could potentially lead to dysregulated P53 activity, which is essential for cell cycle arrest and apoptosis in response to DNA damage. To investigate the functional effects of O-GlcNAc on P300, *in vitro* acetylation of P53 by S1734 O-GlcNAcylated P300 and GlcNAc-deficient P300 could be analyzed as described previously (79).

Identification of pathogenic missense variants at O-GlcNAcylation sites

So far, this study has focused on predicting functional O-GlcNAc sites on ID/DD-associated proteins using the filter-based approach outlined in Figure 1. Because mutations in these proteins segregate with ID/DD, it is also possible that hypo-O-GlcNAcylation (*i.e.*, mutations of Ser/Thr to other amino acids) of functional O-GlcNAc sites on these proteins may contribute to OGT-CDG pathology. We explored this possibility across the entire O-GlcNAcome (*i.e.*, not limited to the candidates in Fig. 1).

To identify mutations at O-GlcNAc sites that segregate with ID/DD, the ClinVar database (82) was screened for missense variants using the O-GlcNAc database (83), a recently compiled database of O-GlcNAc sites reported across multiple eukaryotic species. We also screened for missense variants ± 5

residues from the O-GlcNAc site, as mutations proximal to the O-GlcNAc site may affect the affinity of OGT for a given substrate. The ClinVar screen identified 125 missense variants on O-GlcNAc sites, eight of which segregated with ID/DD (See File S2 for the list of ClinVar hits). The remaining 117 O-GlcNAc site missense variants were either pathogenic for a non-ID/DD syndrome or scored as likely benign or possibly damaging by PolyPhen scoring (PolyPhen HumVar score < 0.85). Additionally, we identified 34 pathogenic and likely pathogenic variants proximal (± 5 residues) to the O-GlcNAc site, which segregate with ID/DD (See File S2). Two O-GlcNAcylation site mutations (TUBB2B^{S172P} and TUBB2B^{S172L}) were pathogenic or likely pathogenic in patients with cortical dysplasia, complex, with other brain malformations (CDCBM7) (49). Six O-GlcNAcylation site missense variants (PUM1^{S802F}, DOCK7^{S190N}, DOCK7^{S190G}, KMT2A^{S1858I}, ATRX^{S594C}, and HCF1^{T566M}) were of uncertain significance, although PolyPhen scoring predicted all six variants as probably pathogenic (PolyPhen HumVar score > 0.85). For summaries of ID/DD missense variants and associated pathogenicity/PolyPhen scores, see File S2.

Of the hits identified by the ClinVar search, the pathogenic TUBB2B^{S172P} mutation was particularly interesting. The TUBB2B^{S172P} variant segregates with CDCBM7 and displays impaired heterodimerization with α -tubulin and is not capable of polymerizing into microtubules (49). A similar effect would be predicted from the presence of a bulky O-GlcNAc moiety at S172 due to steric occlusion of the GTP/GDP-binding site (see earlier discussion and Fig. 3A), further substantiating the proposed role of this site in regulating tubulin polymerization. Additionally, an S172L variant was also reported for the CDCBM7 syndrome (Table 1); however, the effects of this mutation on tubulin function have not been investigated biochemically or structurally. It is of course possible that these mutations affect tubulin function independent of O-GlcNAcylation. Indeed, mutagenesis from Ser to Pro (a “helix breaker” with restricted backbone conformations) may disrupt tubulin function independently of the loss of O-GlcNAc. However, the backbone conformation of S172 is compatible with that preferred by a proline (Fig. 3B). Future work could explore the effects of proline mutagenesis and loss of S172 O-GlcNAcylation on tubulin function in ID/DD.

From the six missense variants predicted to be pathogenic ID/DD mutations by PolyPhen scoring, the HCF1^{T566M} variant is of particular interest. The HCF1^{T566M} variant is present in a patient with mental retardation 3, X-linked. T566 was not identified as a candidate O-GlcNAc site in our initial filter-based approach as it is not in a (known) functional region/domain, although it is proximal to the binding site for OGT (84). It is possible, however, that T566 O-GlcNAcylation stabilizes HCF1; HCF1 is cotranslationally O-GlcNAcylated (85), and the sole reported function of cotranslational O-GlcNAcylation is stabilization of nascent protein chains during translation (39). Thus, the HCF1^{T566M} variant may represent a destabilized hypomorph due to loss of T566 O-GlcNAcylation. Whether this is the case and whether HCF1 hypo-O-GlcNAcylation at T566 contributes to OGT-CDG etiology, are avenues for future research.

Conclusions

We have identified 22 ID/DD-associated proteins, the hypo-O-GlcNAcylation of which may contribute to OGT-CDG pathology. We arrived at these candidates by inferring functionality for conserved O-GlcNAc sites on these ID/DD-associated proteins through a filter-based approach integrating inputs from proteomic, clinical, structural, and literature databases. Additionally, we identified several conserved O-GlcNAc sites, which are pathogenically mutated in syndromic ID patients, adding further impetus to study the function of O-GlcNAc on these neuronal proteins. We anticipate that the results of this study will further highlight the convergence between the O-GlcNAc proteome and ID and galvanize further research into specific O-GlcNAc signaling nodes and how hypo-O-GlcNAcylation of these same proteins may give rise to OGT-CDG.

Experimental procedures

Filter-based approach for the identification of OGT-CDG candidate conveyers

O-GlcNAc proteomic (O-GlcNAcomics) data were pooled manually from 14 publications into a Microsoft Excel spreadsheet (version 16.46) (9–11, 25, 26, 29–34, 86–88). As a quality control step, O-GlcNAc sites were excluded if the reported O-GlcNAc site mapped to asparagine (N-linked glycosylation), if the corresponding peptide was misannotated (reported O-GlcNAcylated peptide absent from the UniProt FASTA sequence of the full-length protein) or if the protein was exclusively extracellular (determined by manual inspection of UniProt entries for protein localization). To assess whether an O-GlcNAcylated protein was ID/DD associated, the list of site-mapped O-GlcNAcylated proteins was then compared to a recently reported list of all known ID/DD-associated genes (22) and to the DECIPHER database of clinical genomic variation (35). If the gene encoding an O-GlcNAcylated protein was reported as being the sole “pathogenic” variant in a patient with “ID” or “developmental delay,” the O-GlcNAcylated protein passed this filter. Subsequently, to assess whether O-GlcNAc sites were conserved in vertebrates, amino acid sequences in FASTA format for human, mouse, rat, rabbit, and sheep homologs of a given O-GlcNAcylated protein were extracted from the UniProt database (see File S1 for UniProt IDs used). Clustal Omega was used to align amino acid sequences, with O-GlcNAc sites not conserved in ≥ 1 species excluded. In 3.84% of cases, one of rat, rabbit, or sheep sequences were absent for a given protein. In such cases, bovine, porcine, or chimpanzee sequences (in descending order of preference) were used instead, due to their availability in the UniProt database. Structural inspection of O-GlcNAc sites, to assess whether O-GlcNAc sites were surface exposed and thus accessible to the OGT active site, was carried out on available PDB structures using PyMOL (<https://pymol.org/2/>) v 1.7.4.5. Where structural information was not available, the IUPred2A (<https://iupred2a.elte.hu>) disorder predictor was used to predict whether an O-GlcNAc site resided in a disordered (and therefore accessible) region of the

protein, using the long disorder (default) setting (40). O-GlcNAc sites were classified as ordered if IUPred scores were higher than 0.5. Disorder predictions were carried out using FASTA sequences from human homologs. IUPred scores for individual O-GlcNAc sites are detailed in File S1.

GO and PANTHER pathway analysis

GO analysis for biological processes involving OGT-CDG candidate conveyers was carried out using the Gene Ontology resource (<http://geneontology.org>) (89). For investigating pathways in which OGT-CDG candidate conveyers are signaling nodes and classifying candidate proteins into classes, PANTHER version 16.0 was used (90). For both GO and PANTHER pathway analysis, UniProt identifiers for human homologs were used (see File S1). Outputs from classification of shared “biological processes,” “protein class,” and “pathways” were exported as .txt files into Microsoft Excel and used to generate the associated figures.

ClinVar database screening for missense mutations in ID/DD-associated proteins

To retrieve variants in genes encoding O-GlcNAc modified proteins, the ClinVar short variants dataset (82) (https://ftp.ncbi.nlm.nih.gov/pub/clinvar/vcf_GRCh38/clinvar.vcf.gz, accessed 03/01/2022) was parsed using the PyVCF library and was cross compared with a recently published O-GlcNAcome (83) and an independent list of 146 conserved O-GlcNAc sites (File S1) parsed using the csv library (91). Identified variants were filtered to only include missense mutations. Amino acid substitutions for each variant were retrieved using the requests library (92) to interface with the esummary ClinVar API, parsed using the json library. Missense mutations affecting O-GlcNAcylated residues or sites ± 5 residues were then filtered based on pathogenicity and retained for manual follow up. Code can be accessed at https://github.com/IggyCz/ClinVar_O-GlcNAcome_search/blob/main/clinvar_searcher_v1.py.

PolyPhen scoring

PolyPhen v2 (93) was used to predict the pathogenicity of missense variants at conserved sites of O-GlcNAcylation. UniProt FASTA sequences for human homologs of a given candidate conveyer were manually entered into the search bar. HumVar scores, which factor in the presence of single nucleotide polymorphisms at a given site in healthy individuals, were used, along with default cut-offs of 0.85 for “probably damaging” and 0.15 for “likely benign.”

Data availability

All data are contained within the article.

Supporting information—This article contains supporting information (14, 49, 50, 56, 59–61, 68, 72, 74, 75, 81, 84, 94–110).

Acknowledgments—We would like to thank Dr Sergio Galan Bartual and Dr Florence Authier for insightful discussions and feedback.

Conor Mitchell was funded by the BBSRC EASTBIO Doctoral Training Program. Ignacy Czajewski was funded by the National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), T001682

Author contributions—C. W. M. and D. M. F. v. A. conceptualization; C. W. M. and I. C. methodology; I. C. software; I. C. validation; C. W. M. and I. C. investigation; C. W. M. and I. C. data curation; C. W. M. writing—original draft; C. W. M., I. C., and D. M. F. v. A. writing—review & editing; C. W. M. visualization; D. M. F. v. A. supervision; D. M. F. v. A. funding acquisition.

Funding and additional information—This work was funded by a Wellcome Trust Investigator Award (grant no. 110061 to D. M. F. v. A.) and a Novo Nordisk Foundation Laureate Award (to D. M. F. v. A.).

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AHCY, adenosyl homocysteinase; CDG, congenital disorder of glycosylation; CTE, C-terminal extension; DD, developmental delay; EESC, embryonic stem cell; GO, gene ontology; ID, intellectual disability; mESC, mouse embryonic stem cell; OGT, O-GlcNAc transferase; PDB, Protein Data Bank; PTM, posttranslational modification; TPR, tetratricopeptide.

References

- Yang, X., and Qian, K. (2017) Protein O-GlcNAcylation: emerging mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* **18**, 452–465
- Holt, G. D., and Hart, G. W. (1986) The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharide linkage, O-linked GlcNAc. *J. Biol. Chem.* **261**, 8049–8057
- Haltiwanger, R. S., Blomberg, M. A., and Hart, G. W. (1992) Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-N-acetylglucosamine:polypeptide beta-N-acetylglucosaminyltransferase. *J. Biol. Chem.* **267**, 9005–9013
- Martinez-Fleites, C., He, Y., and Davies, G. J. (2010) Structural analyses of enzymes involved in the O-GlcNAc modification. *Biochim. Biophys. Acta* **1800**, 122–133
- Stephen, H. M., Praissman, J. L., and Wells, L. (2021) Generation of an interactome for the tetratricopeptide repeat domain of O-GlcNAc transferase indicates a role for the enzyme in intellectual disability. *J. Proteome Res.* **20**, 1229–1242
- Capotosti, F., Guernier, S., Lammers, F., Waridel, P., Cai, Y., Jin, J., et al. (2011) O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1. *Cell* **144**, 376–388
- Vocadlo, D. J. (2012) O-GlcNAc processing enzymes: Catalytic mechanisms, substrate specificity, and enzyme regulation. *Curr. Opin. Chem. Biol.* **16**, 488–497
- Okuyama, R., and Marshall, S. (2003) UDP-N-acetylglucosaminyl transferase (OGT) in brain tissue: temperature sensitivity and subcellular distribution of cytosolic and nuclear enzyme. *J. Neurochem.* **86**, 1271–1280
- Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A., Lynn, A. J., et al. (2006) O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol. Cell Proteomics* **5**, 923–934
- Alfaro, J. F., Gong, C. X., Monroe, M. E., Aldrich, J. T., Clauss, T. R., Purvine, S. O., et al. (2012) Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain-specific O-GlcNAc transferase targets. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 7280–7285
- Trinidad, J. C., Barkan, D. T., Gullledge, B. F., Thalhammer, A., Sali, A., Schoepfer, R., et al. (2012) Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Mol. Cell Proteomics* **11**, 215–229
- Rexach, J. E., Clark, P. M., Mason, D. E., Neve, R. L., Peters, E. C., and Hsieh-Wilson, L. C. (2012) Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nat. Chem. Biol.* **8**, 253–261
- Lagerlöf, O., Hart, G. W., and Haganir, R. L. (2017) O-GlcNAc transferase regulates excitatory synapse maturity. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1684–1689
- Skorobogatko, Y., Landicho, A., Chalkley, R. J., Kossenkov, A. V., Gallo, G., and Vosseller, K. (2014) O-linked β -N-acetylglucosamine (O-GlcNAc) site thr-87 regulates synapsin I localization to synapses and size of the reserve pool of synaptic vesicles. *J. Biol. Chem.* **289**, 3602–3612
- Lee, B. E., Kim, H. Y., Kim, H. J., Jeong, H., Kim, B. G., Lee, H. E., et al. (2020) O-GlcNAcylation regulates dopamine neuron function, survival and degeneration in Parkinson disease. *Brain* **143**, 3699–3716
- Wheatley, E. G., Albarran, E., White, C. W., Bieri, G., Sanchez-Diaz, C., Pratt, K., et al. (2019) Neuronal O-GlcNAcylation improves cognitive function in the aged mouse brain. *Curr. Biol.* **29**, 3359–3369.e4
- O'Donnell, N., Zachara, N. E., Hart, G. W., and Marth, J. D. (2004) Ogt-dependent X-chromosome-linked protein glycosylation is a requisite modification in somatic cell function and embryo viability. *Mol. Cell Biol.* **24**, 1680–1690
- Pravata, V. M., Muha, V., Gundogdu, M., Ferenbach, A. T., Kakade, P. S., Vandadi, V., et al. (2019) Catalytic deficiency of O-GlcNAc transferase leads to X-linked intellectual disability. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 14961–14970
- Pravata, V. M., Gundogdu, M., Bartual, S. G., Ferenbach, A. T., Stavridis, M., Öunap, K., et al. (2020) A missense mutation in the catalytic domain of O-GlcNAc transferase links perturbations in protein O-GlcNAcylation to X-linked intellectual disability. *FEBS Lett.* **594**, 717–727
- Vaidyanathan, K., Niranjana, T., Selvan, N., Teo, C. F., May, M., Patel, S., 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
- Willems, A. P., Gundogdu, M., Kempers, M. J. E., Giltay, J. C., Pfundt, R., Elferink, M., et al. (2017) Mutations in N-acetylglucosamine (O-GlcNAc) transferase in patients with X-linked intellectual disability. *J. Biol. Chem.* **292**, 12621–12631
- Vissers, L. E., Gilissen, C., and Veltman, J. A. (2016) Genetic studies in intellectual disability and related disorders. *Nat. Rev. Genet.* **17**, 9–18
- Pravata, V. M., Omelková, M., Stavridis, M. P., Desbiens, C. M., Stephen, H. M., Lefeber, D. J., et al. (2020) An intellectual disability syndrome with single-nucleotide variants in O-GlcNAc transferase. *Eur. J. Hum. Genet.* **28**, 706–714
- Gundogdu, M., Llabrés, S., Gorelik, A., Ferenbach, A. T., Zachariae, U., and van Aalten, D. M. F. (2018) The O-GlcNAc transferase intellectual disability mutation L254F distorts the TPR helix. *Cell Chem. Biol.* **25**, 513–518.e4
- Hahne, H., Sobotzki, N., Nyberg, T., Helm, D., Borodkin, V. S., van Aalten, D. M., et al. (2013) Proteome wide purification and identification of O-GlcNAc-modified proteins using click chemistry and mass spectrometry. *J. Proteome Res.* **12**, 927–936
- Woo, C. M., Lund, P. J., Huang, A. C., Davis, M. M., Bertozzi, C. R., and Pitteri, S. J. (2018) Mapping and quantification of over 2000 O-linked glycopeptides in activated human T cells with isotope-targeted glycoproteomics (isotag). *Mol. Cell Proteomics* **17**(4), 764–775
- Chen, P. H., Smith, T. J., Wu, J., Siesser, P. F., Bisnett, B. J., Khan, F., et al. (2017) Glycosylation of KEAP1 links nutrient sensing to redox stress signaling. *EMBO J.* **36**, 2233–2250
- Zhu, Q., Cheng, X., Cheng, Y., Chen, J., Xu, H., Gao, Y., et al. (2020) O-GlcNAcylation regulates the methionine cycle to promote pluripotency of stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7755–7763
- Myers, S. A., Panning, B., and Burlingame, A. L. (2011) Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9490–9495
- Nandi, A., Sprung, R., Barma, D. K., Zhao, Y., Kim, S. C., Falck, J. R., et al. (2006) Global identification of O-GlcNAc-modified proteins. *Anal. Chem.* **78**, 452–458

31. Khidekel, N., Ficarro, S. B., Peters, E. C., and Hsieh-Wilson, L. C. (2004) Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13132–13137
32. Rexach, J. E., Rogers, C. J., Yu, S. H., Tao, J., Sun, Y. E., and Hsieh-Wilson, L. C. (2010) Quantification of O-glycosylation stoichiometry and dynamics using resolvable mass tags. *Nat. Chem. Biol.* **6**, 645–651
33. Chalkley, R. J., Thalhammer, A., Schoepfer, R., and Burlingame, A. L. (2009) Identification of protein O-GlcNAcylation sites using electron transfer dissociation mass spectrometry on native peptides. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8894–8899
34. Zaro, B. W., Yang, Y. Y., Hang, H. C., and Pratt, M. R. (2011) Chemical reporters for fluorescent detection and identification of O-GlcNAc-modified proteins reveal glycosylation of the ubiquitin ligase NEDD4-1. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 8146–8151
35. Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., *et al.* (2009) DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. *Am. J. Hum. Genet.* **84**, 524–533
36. Lazarus, M. B., Jiang, J., Kapuria, V., Bhuiyan, T., Janetzko, J., Zandberg, W. F., *et al.* (2013) HCF-1 is cleaved in the active site of O-GlcNAc transferase. *Science* **342**, 1235–1239
37. Rafie, K., Raimi, O., Ferenbach, A. T., Borodkin, V. S., Kapuria, V., and van Aalten, D. M. F. (2017) Recognition of a glycosylation substrate by the O-GlcNAc transferase TPR repeats. *Open Biol.* **7**
38. Pathak, S., Alonso, J., Schimpl, M., Rafie, K., Blair, D. E., Borodkin, V. S., *et al.* (2015) The active site of O-GlcNAc transferase imposes constraints on substrate sequence. *Nat. Struct. Mol. Biol.* **22**, 744–750
39. Zhu, Y., Liu, T. W., Cecioni, S., Eskandari, R., Zandberg, W. F., and Vocadlo, D. J. (2015) O-GlcNAc occurs cotranslationally to stabilize nascent polypeptide chains. *Nat. Chem. Biol.* **11**, 319–325
40. Erdos, G., and Dosztanyi, Z. (2020) Analyzing protein disorder with IUPred2A. *Curr. Protoc. Bioinform.* **70**, e99
41. Ohn, T., Kedersha, N., Hickman, T., Tisdale, S., and Anderson, P. (2008) A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nat. Cell Biol.* **10**, 1224–1231
42. Sakabe, K., Wang, Z., and Hart, G. W. (2010) Beta-N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 19915–19920
43. Ranuncolo, S. M., Ghosh, S., Hanover, J. A., Hart, G. W., and Lewis, B. A. (2012) Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription *in vitro* and *in vivo*. *J. Biol. Chem.* **287**, 23549–23561
44. Moulton, M. J., Humphreys, G. B., Kim, A., and Letsou, A. (2020) O-GlcNAcylation dampens Dpp/BMP signaling to ensure proper *Drosophila* embryonic development. *Dev. Cell* **53**, 330–343.e3
45. Olivier-Van Stichelen, S., Dehennaut, V., Buzy, A., Zachayus, J. L., Guinez, C., Mir, A. M., *et al.* (2014) O-GlcNAcylation stabilizes β -catenin through direct competition with phosphorylation at threonine 41. *FASEB J.* **28**(8), 3325–3328
46. Tarrant, M. K., Rho, H. S., Xie, Z., Jiang, Y. L., Gross, C., Culhane, J. C., *et al.* (2012) Regulation of CK2 by phosphorylation and O-GlcNAcylation revealed by semisynthesis. *Nat. Chem. Biol.* **8**, 262–269
47. Flynn, K. C., and Bradke, F. (2020) Role of the cytoskeleton and membrane trafficking in axon–dendrite morphogenesis. In *Cellular Migration and Formation of Axons and Dendrites*, Elsevier, Amsterdam, Netherlands: 21–56
48. Craig, A. M., and Banker, G. (1994) Neuronal polarity. *Annu. Rev. Neurosci.* **17**, 267–310
49. Jaglin, X. H., Poirier, K., Saillour, Y., Buhler, E., Tian, G., Bahi-Buisson, N., *et al.* (2009) Mutations in the beta-tubulin gene TUBB2B result in asymmetrical polymicrogyria. *Nat. Genet.* **41**, 746–752
50. Fourest-Lieuvin, A., Peris, L., Gache, V., Garcia-Saez, I., Juillan-Binard, C., Lantéz, V., *et al.* (2006) Microtubule regulation in mitosis: tubulin phosphorylation by the cyclin-dependent kinase Cdk1. *Mol. Biol. Cell* **17**, 1041–1050
51. Löwe, J., Li, H., Downing, K. H., and Nogales, E. (2001) Refined structure of alpha beta-tubulin at 3.5 Å resolution. *J. Mol. Biol.* **313**, 1045–1057
52. Ji, S., Kang, J. G., Park, S. Y., Lee, J., Oh, Y. J., and Cho, J. W. (2011) O-GlcNAcylation of tubulin inhibits its polymerization. *Amino Acids* **40**, 809–818
53. Shelanski, M. L., Gaskin, F., and Cantor, C. R. (1973) Microtubule assembly in the absence of added nucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 765–768
54. Bradke, F., and Dotti, C. G. (1999) The role of local actin instability in axon formation. *Science* **283**, 1931–1934
55. Gomez, T. M., and Letourneau, P. C. (2014) Actin dynamics in growth cone motility and navigation. *J. Neurochem.* **129**, 221–234
56. Cuvertino, S., Stuart, H. M., Chandler, K. E., Roberts, N. A., Armstrong, R., Bernardini, L., *et al.* (2017) ACTB loss-of-function mutations result in a pleiotropic developmental disorder. *Am. J. Hum. Genet.* **101**, 1021–1033
57. Schutt, C. E., Myslik, J. C., Rozycki, M. D., Goonesekere, N. C., and Lindberg, U. (1993) The structure of crystalline profilin-beta-actin. *Nature* **365**, 810–816
58. Davey, R. J., and Moens, P. D. (2020) Profilin: many facets of a small protein. *Biophys. Rev.* **12**, 827–849
59. Da Silva, J. S., Medina, M., Zuliani, C., Di Nardo, A., Witke, W., and Dotti, C. G. (2003) RhoA/ROCK regulation of neurogenesis *via* profilin Ila-mediated control of actin stability. *J. Cell Biol.* **162**, 1267–1279
60. Holmes, K. C., Popp, D., Gebhard, W., and Kabsch, W. (1990) Atomic model of the actin filament. *Nature* **347**, 44–49
61. Floor, S. N., Condon, K. J., Sharma, D., Jankowsky, E., and Doudna, J. A. (2016) Autoinhibitory interdomain interactions and subfamily-specific extensions redefine the catalytic core of the human DEAD-box protein DDX3. *J. Biol. Chem.* **291**, 2412–2421
62. Putnam, A. A., Gao, Z., Liu, F., Jia, H., Yang, Q., and Jankowsky, E. (2015) Division of labor in an oligomer of the DEAD-box RNA helicase Ded1p. *Mol. Cell* **59**, 541–552
63. Lai, M. C., Chang, W. C., Shieh, S. Y., and Tarn, W. Y. (2010) DDX3 regulates cell growth through translational control of cyclin E1. *Mol. Cell Biol.* **30**, 5444–5453
64. Chen, H. H., Yu, H. L., and Tarn, W. Y. (2016) DDX3 modulates neurite development *via* translationally activating an RNA regulon involved in Rac1 activation. *J. Neurosci.* **36**, 9792–9804
65. Samir, P., Kesavardhana, S., Patmore, D. M., Gingras, S., Malireddi, R. K. S., Karki, R., *et al.* (2019) DDX3X acts as a live-or-die checkpoint in stressed cells by regulating NLRP3 inflammasome. *Nature* **573**, 590–594
66. Protter, D. S. W., and Parker, R. (2016) Principles and properties of stress granules. *Trends Cell Biol.* **26**, 668–679
67. Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M. R., *et al.* (2015) Mutations in DDX3X are a common cause of unexplained intellectual disability with gender-specific effects on Wnt signaling. *Am. J. Hum. Genet.* **97**(2), 343–352
68. Lai, M. C., Lee, Y. H., and Tarn, W. Y. (2008) The DEAD-box RNA helicase DDX3 associates with export messenger ribonucleoproteins as well as tip-associated protein and participates in translational control. *Mol. Biol. Cell* **19**, 3847–3858
69. Zeidan, Q., Wang, Z., De Maio, A., and Hart, G. W. (2010) O-GlcNAc cycling enzymes associate with the translational machinery and modify core ribosomal proteins. *Mol. Biol. Cell* **21**, 1922–1936
70. Zhang, X., Shu, X. E., 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
71. Gold, W. A., Krishnaraj, R., Ellaway, C., and Christodoulou, J. (2018) Rett syndrome: a genetic update and clinical review focusing on comorbidities. *ACS Chem. Neurosci.* **9**, 167–176
72. Bellini, E., Pavesi, G., Barbiero, I., Bergo, A., Chandola, C., Nawaz, M. S., *et al.* (2014) MeCP2 post-translational modifications: a mechanism to control its involvement in synaptic plasticity and homeostasis? *Front. Cell Neurosci.* **8**, 236
73. Li, H., Zhong, X., Chau, K. F., Williams, E. C., and Chang, Q. (2011) Loss of activity-induced phosphorylation of MeCP2 enhances synaptogenesis, LTP and spatial memory. *Nat. Neurosci.* **14**, 1001–1008
74. Tsuruzoe, S., Ishihara, K., Uchimura, Y., Watanabe, S., Sekita, Y., Aoto, T., *et al.* (2006) Inhibition of DNA binding of Sox2 by the SUMO conjugation. *Biochem. Biophys. Res. Commun.* **351**, 920–926

75. Myers, S. A., Peddada, S., Chatterjee, N., Friedrich, T., Tomoda, K., Krings, G., *et al.* (2016) SOX2 O-GlcNAcylation alters its protein-protein interactions and genomic occupancy to modulate gene expression in pluripotent cells. *Elife* **5**, e10647
76. Blackburn, P. R., Chacon-Camacho, O. F., Ortiz-González, X. R., Reyes, M., Lopez-Uriarte, G. A., Zarei, S., *et al.* (2018) Extension of the mutational and clinical spectrum of SOX2 related disorders: description of six new cases and a novel association with suprasellar teratoma. *Am. J. Med. Genet. A.* **176**, 2710–2719
77. Kauppinen, T. M., Chan, W. Y., Suh, S. W., Wiggins, A. K., Huang, E. J., and Swanson, R. A. (2006) Direct phosphorylation and regulation of poly(ADP-ribose) polymerase-1 by extracellular signal-regulated kinases 1/2. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7136–7141
78. Lai, Y. S., Chang, C. W., Pawlik, K. M., Zhou, D., Renfrow, M. B., and Townes, T. M. (2012) SRY (sex determining region Y)-box2 (Sox2)/poly ADP-ribose polymerase 1 (Parp1) complexes regulate pluripotency. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3772–3777
79. Gu, W., and Roeder, R. G. (1997) Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* **90**, 595–606
80. Negri, G., Magini, P., Milani, D., Colapietro, P., Rusconi, D., Scarano, E., *et al.* (2016) From whole gene deletion to point mutations of EP300-positive rubinstein-taybi patients: new insights into the mutational spectrum and peculiar clinical hallmarks. *Hum. Mutat.* **37**(2), 175–183
81. Miller Jenkins, L. M., Feng, H., Durell, S. R., Tagad, H. D., Mazur, S. J., Tropea, J. E., *et al.* (2015) Characterization of the p300 Taz2-p53 TAD2 complex and comparison with the p300 Taz2-p53 TAD1 complex. *Biochemistry* **54**, 2001–2010
82. Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., *et al.* (2018) ClinVar: Improving access to variant interpretations and supporting evidence. *Nucl. Acids Res.* **46**, D1062–D1067
83. Wulff-Fuentes, E., Berendt, R. R., Massman, L., Danner, L., Malard, F., Vora, J., *et al.* (2021) The human O-GlcNAc database and meta-analysis. *Sci. Data* **8**, 25
84. Wysocka, J., Myers, M. P., Laherty, C. D., Eisenman, R. N., and Herr, W. (2003) Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev.* **17**, 896–911
85. Zhu, Y., Willems, L. I., Salas, D., Cecioni, S., Wu, W. B., Foster, L. J., *et al.* (2020) Tandem bioorthogonal labeling uncovers endogenous cotranslationally O-GlcNAc modified nascent proteins. *J. Am. Chem. Soc.* **142**, 15729–15739
86. Wang, Z., Gucek, M., and Hart, G. W. (2008) Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13793–13798
87. Zhao, P., Viner, R., Teo, C. F., Boons, G. J., Horn, D., and Wells, L. (2011) Combining high-energy C-trap dissociation and electron transfer dissociation for protein O-GlcNAc modification site assignment. *J. Proteome Res.* **10**, 4088–4104
88. Dias, W. B., Cheung, W. D., and Hart, G. W. (2012) O-GlcNAcylation of kinases. *Biochem. Biophys. Res. Commun.* **422**, 224–228
89. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., *et al.* (2000) Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* **25**, 25–29
90. Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albu, L. P., Mushayama, T., *et al.* (2021) PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucl. Acids Res.* **49**, D394–D403
91. Van Rossum, G. (2020) *The Python Library Reference, release 3.8.2*. Python Software Foundation, Beaverton, OR
92. Chandra, R. V., and Varanasi, B. S. (2015) *Python Requests Essentials*, Packt Publishing Ltd, Birmingham, UK
93. Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., *et al.* (2010) A method and server for predicting damaging missense mutations. *Nat. Met.* **7**, 248–249
94. Kneussel, M., Haverkamp, S., Fuhrmann, J. C., Wang, H., Wässle, H., Olsen, R. W., *et al.* (2000) The gamma-aminobutyric acid type A receptor (GABAAR)-associated protein GABARAP interacts with gephyrin but is not involved in receptor anchoring at the synapse. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8594–8599
95. Yap, C. C., Murate, M., Kishigami, S., Muto, Y., Kishida, H., Hashikawa, T., *et al.* (2003) Adaptor protein complex-4 (AP-4) is expressed in the central nervous system neurons and interacts with glutamate receptor delta2. *Mol. Cell. Neurosci.* **24**, 283–295
96. Malakhova, M., Tereshko, V., Lee, S.-Y., Yao, K., Cho, Y.-Y., Bode, A., *et al.* (2008) Structural basis for activation of the autoinhibitory C-terminal kinase domain of p90 RSK2. *Nat. Struct. Mol. Biol.* **15**, 112–113
97. de Caestecker, M. P., Hemmati, P., Larisch-Bloch, S., Ajmera, R., Roberts, A. B., and Lechleider, R. J. (1997) Characterization of functional domains within Smad4/DPC4. *J. Biol. Chem.* **272**, 13690–13696
98. Rokudai, S., Laptenko, O., Arnal, S. M., Taya, Y., Kitabayashi, I., and Prives, C. (2013) MOZ increases p53 acetylation and premature senescence through its complex formation with PML. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 3895–3900
99. Jaber, E., Chitsazian, F., Shahidi, G. A., Rohani, M., Sina, F., Safari, I., *et al.* (2013) The novel mutation p.Asp251Asn in the β -subunit of succinate-CoA ligase causes encephalomyopathy and elevated succinylcarnitine. *J. Hum. Genet.* **58**, 526–530
100. Nishiyama, M., Oshikawa, K., Tsukada, Y., Nakagawa, T., Iemura, S., Natsume, T., *et al.* (2009) CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nat. Cell. Biol.* **11**, 172–182
101. Mandel, S., Rechavi, G., and Gozes, I. (2007) Activity-dependent neuroprotective protein (ADNP) differentially interacts with chromatin to regulate genes essential for embryogenesis. *Dev. Biol.* **303**, 814–824
102. Itoh, G., Kanno, S., Uchida, K. S. K., Chiba, S., Sugino, S., Watanabe, K., *et al.* (2011) CAMP (Cl3orf8, ZNF828) is a novel regulator of kinetochore-microtubule attachment. *EMBO J.* **30**, 130–144
103. Dallas, P. B., Pacchione, S., Wilsker, D., Bowrin, V., Kobayashi, R., and Moran, E. (2000) The human SWI-SNF complex protein p270 is an ARID family member with non-sequence-specific DNA binding activity. *Mol. Cell. Biol.* **20**, 3137–3146
104. Beinert, H., and Kennedy, M. C. (1993) Aconitase, a two-faced protein: enzyme and iron regulatory factor. *FASEB J.* **7**, 1442–1449
105. Hoang, H. T., Schlager, M. A., Carter, A. P., and Bullock, S. L. (2017) DYNC1H1 mutations associated with neurological diseases compromise processivity of dynein-dynactin-cargo adaptor complexes. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1597–E1606
106. Mou, T.-C., Gille, A., Fancy, D. A., Seifert, R., and Sprang, S. R. (2005) Structural basis for the inhibition of mammalian membrane adenylyl cyclase by 2'(3')-O-(N-Methylanthraniloyl)-guanosine 5'-triphosphate. *J. Biol. Chem.* **280**, 7253–7261
107. Kordeli, E., Lambert, S., and Bennett, V. (1995) AnkyrinG. A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier. *J. Biol. Chem.* **270**, 2352–2359
108. Harris, J. L., Jakob, B., Taucher-Scholz, G., Dianov, G. L., Becherel, O. J., and Lavin, M. F. (2009) Aprataxin, poly-ADP ribose polymerase 1 (PARP-1) and apurinic endonuclease 1 (APE1) function together to protect the genome against oxidative damage. *Hum. Mol. Genet.* **18**, 4102–4117
109. Xu, G., Shin, S. B., and Jaffrey, S. R. (2009) Global profiling of protease cleavage sites by chemoselective labeling of protein N-termini. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19310–19315
110. Takakubo, F., Cartwright, P., Hoogenraad, N., Thorburn, D. R., Collins, F., Lithgow, T., *et al.* (1995) An amino acid substitution in the pyruvate dehydrogenase E1 alpha gene, affecting mitochondrial import of the precursor protein. *Am. J. Hum. Genet.* **57**, 772–780