



# Catalytic deficiency of O-GlcNAc transferase leads to X-linked intellectual disability

Veronica M. Pravata<sup>a,1</sup>, Villo Muha<sup>a,1</sup>, Mehmet Gundogdu<sup>a,1</sup>, Andrew T. Ferenbach<sup>a</sup>, Poonam S. Kakade<sup>b</sup>, Vasudha Vandadi<sup>a</sup>, Ariane C. Wilmes<sup>a</sup>, Vladimir S. Borodkin<sup>a</sup>, Shelagh Joss<sup>c</sup>, Marios P. Stavridis<sup>b</sup>, and Daan M. F. van Aalten<sup>a,2</sup>

<sup>a</sup>Division of Gene Regulation and Expression, School of Life Sciences, University of Dundee, DD1 5EH Dundee, United Kingdom; <sup>b</sup>Division of Cell and Developmental Biology, School of Life Sciences, University of Dundee, DD1 5EH Dundee, United Kingdom; and <sup>c</sup>West of Scotland Genetic Service, Queen Elizabeth University Hospital, G51 4TF Glasgow, United Kingdom

Edited by Carolyn R. Bertozzi, Stanford University, Stanford, CA, and approved June 17, 2019 (received for review January 29, 2019)

O-GlcNAc transferase (OGT) is an X-linked gene product that is essential for normal development of the vertebrate embryo. It catalyses the O-GlcNAc posttranslational modification of nucleocytoplasmic proteins and proteolytic maturation of the transcriptional coregulator Host cell factor 1 (HCF1). Recent studies have suggested that conservative missense mutations distal to the OGT catalytic domain lead to X-linked intellectual disability in boys, but it is not clear if this is through changes in the O-GlcNAc proteome, loss of protein–protein interactions, or misprocessing of HCF1. Here, we report an OGT catalytic domain missense mutation in monozygotic female twins (c. X:70779215 T > A, p. N567K) with intellectual disability that allows dissection of these effects. The patients show limited IQ with developmental delay and skewed X-inactivation. Molecular analyses revealed decreased OGT stability and disruption of the substrate binding site, resulting in loss of catalytic activity. Editing this mutation into the *Drosophila* genome results in global changes in the O-GlcNAc proteome, while in mouse embryonic stem cells it leads to loss of O-GlcNAcase and delayed differentiation down the neuronal lineage. These data imply that catalytic deficiency of OGT could contribute to X-linked intellectual disability.

intellectual disability | O-GlcNAc | neurodevelopment

The O-linked β-N-acetyl-D-glucosamine (O-GlcNAc) transferase (OGT) and the hydrolase O-GlcNAcase (OGA) together orchestrate protein O-GlcNAcylation, a dynamic co/posttranslational modification cycling on thousands of nucleocytoplasmic proteins (1, 2). OGT catalyses the covalent attachment of the monosaccharide O-GlcNAc to serine or threonine. This modification has been suggested to affect transcription (3–6), translation (1), protein stability (7, 8), and subcellular localization (9, 10). O-GlcNAcylation plays a key role in regulating stress response (11–13), differentiation (14, 15), nutrient sensing (16, 17), and autophagy (18).

Genetic studies have highlighted the importance of OGT and protein O-GlcNAcylation in development. OGT is essential for embryonic stem cell viability (19, 20) and mouse embryonic development (21). Zebrafish lacking OGT function exhibit shortened body axis and smaller brains (22). The *Drosophila melanogaster* OGT encoded by the Polycomb group gene *super sex combs (sxc)* plays a critical role in establishing *Drosophila* segment identity (23, 24).

The O-GlcNAc modification is particularly abundant in the brain (25, 26), where it controls memory formation (27–29), circadian rhythm (30, 31), and appetite (32, 33). Numerous neuron-specific proteins are O-GlcNAcylated, such as the Microtubule-associated proteins tau (34) and CRMP2 (35), synaptic vesicle proteins (36), the transcriptional factor cyclic-AMP response element binding protein (CREB) (28), and the AMPA receptor GluA2 subunit (27). Growing evidence suggests that O-GlcNAcylation is essential for normal development and function of the mammalian nervous system (21, 33, 37–39).

Although O-GlcNAcylation has been long implicated in chronic metabolic diseases, such as type II diabetes mellitus,

neurodegeneration, and cancer, its role in neurodevelopmental disorders has only recently become apparent. In the past year, a small number of missense mutations within the human OGT gene have been discovered in patients with X-linked intellectual disability (XLID) (40–43). Intellectual disability (ID) refers to a broad range of developmental disorders characterized by limited intellectual capacity, IQ below 70, and poor adaptive behavior with onset before the age of 18, affecting 1 to 3% of the population worldwide (44). Genetic factors are the major cause of this developmental condition involving over 700 ID genes (45). These genes encode for proteins that are required for neuronal development and activity, contributing to neuronal structure and function through several signaling pathways (46). Mutations in genes located on the X-chromosome account for ~5 to 10% of all ID causative genes affecting predominantly male individuals (44, 47).

To date, XLID-associated mutations in the OGT gene have only been reported for male patients, causing developmental delay and severe cognitive disability. Accompanying clinical phenotypes were dysmorphic features, such as clinodactyly, eye abnormalities, and microcephaly. A common molecular trait of the missense XLID OGT variants is that they appear to retain substantial OGT catalytic activity. This is in agreement with these mutations all being localized to the C-terminal noncatalytic tetratricopeptide repeat (TPR) domain of OGT that is responsible for substrate

## Significance

Protein O-GlcNAcylation is a posttranslational modification essential for development. Recently, mutations in the O-GlcNAc transferase (OGT) substrate binding domain have been described that lead to intellectual disability, but the mechanisms underpinning pathogenesis remain to be explored. This work describes the first point mutation in the OGT catalytic domain leading to effects on O-GlcNAcylation and Host cell factor 1 processing in vitro and in a stem cell/*Drosophila* model, resulting in delayed neuronal differentiation. This establishes a potential link between OGT activity and intellectual disability.

Author contributions: V.M.P., V.M., M.G., and D.M.F.v.A. designed research; V.M.P., V.M., and M.G. performed research; A.T.F., P.S.K., V.V., A.C.W., V.S.B., and S.J. contributed new reagents/analytic tools; V.M.P., V.M., M.G., M.P.S., and D.M.F.v.A. analyzed data; and V.M.P., V.M., M.G., and D.M.F.v.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, [www.pdb.org](https://www.pdb.org) (PDB ID code 6IBO).

<sup>1</sup>V.M.P., V.M., and M.G. contributed equally to this work.

<sup>2</sup>To whom correspondence may be addressed. Email: [dmfvanaalten@dundee.ac.uk](mailto:dmfvanaalten@dundee.ac.uk).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1900065116/-DCSupplemental](https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1900065116/-DCSupplemental).

Published online July 11, 2019.



or misprocessing of HCF1. Here, we characterize an OGT missense mutation in the catalytic domain as observed in female twins with ID and developmental delay. In vitro biochemical analysis of recombinant N567K OGT missense variant protein revealed that this mutation abrogates OGT and HCF1 proteolytic processing activity. Protein crystallography data explain the molecular basis of the observed loss of substrate binding. Introduction of this mutation into *D. melanogaster* by genome editing reveals global effects on the O-GlcNAc proteome. In mouse embryonic stem cells (mESCs) carrying this mutation, loss of OGA protein expression was observed, as well as defective HCF1 processing. The mutation caused defects in neurite outgrowth during neuronal differentiation. Taken together, our data provide evidence that loss of OGT activity could contribute to the XLID phenotype observed in the patients.

## Results and Discussion

**Monozygotic Twins with ID and Developmental Delay Possess a De Novo OGT Mutation.** Twin girls (twin 1 and twin 2) were born at 33-wk gestation by semielective caesarean section for maternal preeclampsia. Conception was unassisted. Twin 1 weighed 1,670 g at birth and required continuous positive airway pressure (CPAP) to aid breathing for 3 wk and supplemental oxygen for 58 d. Twin 2 weighed 2,150 g and needed CPAP for 2 d and supplemental oxygen for 20 d. Both babies were tube-fed initially but bottle feeding was established from 3 to 4 wk of age. The twins showed delay in reaching developmental milestones, especially in areas of speech and language development. By 4 y of age, twin 1 only used 3 to 4 single words; at 10 y she was able to put 2 to 3 words together. Language development of twin 2 was more advanced than her sister; at 4 y she was using short phrases and at 10 y she was speaking in sentences and able to recognize letters of the alphabet. Their movement and physical development were also affected: twin 1 by age 4 y and twin 2 by the age of 2 y were able to walk independently. However, twin 2 had significant gross motor difficulties and clumsiness at the age of 8 y. Both twins were diagnosed with cerebral visual impairment, twin 1 more severe than twin 2. They have attended an additional needs nursery and school.

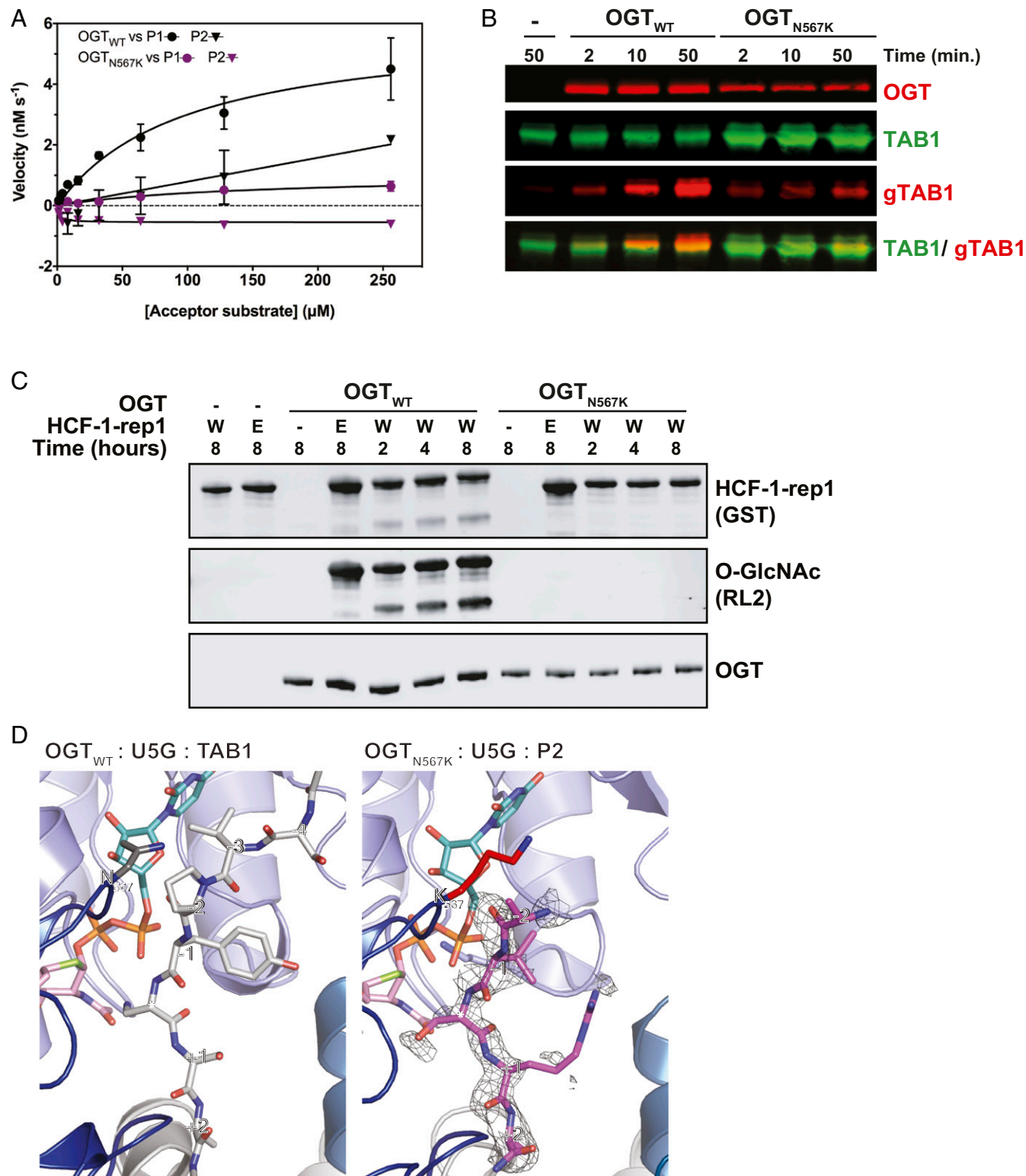
The twins (Fig. 1*A* and *B*) appeared to be affected by a similar developmental syndrome, including small nose, high arched palate, and fifth finger clinodactyly. Twin 1 has ataxic gait with knees bent, walking on toes with inverted feet. An MRI brain scan showed prominence of the lateral ventricles and an enlarged cistern magna in twin 1 at age 2 y, suggesting possible hypoplasia of the inferior cerebellar vermis. Although both siblings have kept generally good health, hypotonia was apparent in infancy and they required iron supplementation intermittently. Twin 1 had an adenoidectomy and treatment for gastro-esophageal reflux and constipation.

Sanger sequencing was performed as part of the Deciphering Developmental Disorders initiative, which currently lists genomic and clinical data from over 14,000 child patients with severe undiagnosed developmental disorders (55). This revealed a single pathogenic missense mutation in the OGT gene (X:70779215 T > A, N567K). The mutation was found absent in control population of 141,456 samples, comprising 125,748 exome sequences and 15,708 whole-genome sequences from unrelated individuals reported in the genome aggregation database, *gnomAD* (56). The mutation affected both monozygotic twins, whereas both of their parents were noncarriers. A 98:2 skew in X-inactivation was detected in both children using polymorphic markers that are differentially methylated on the active and inactive chromosome (*SI Appendix*, Fig. S1). However, being a de novo mutation, which 1 of the 2 chromosomes is predominantly inactivated remains unclear. Taking these data together, we have identified monozygotic twins with ID and developmental delay that possess a de novo OGT mutation.

**The N567K Mutation Abrogates OGT Activity In Vitro.** To understand the potential effects of the OGT N567K (OGT<sub>N567K</sub>) mutation, we next investigated changes in catalytic activity compared with full-length wild-type OGT (OGT<sub>WT</sub>) protein. Steady-state kinetics of OGT glycosyltransferase activity was measured against 2 acceptor peptide substrates, P1 (KKVPVSRA) and P2 (KKVAVSRA), that only varied at the -2 position adjacent to the O-GlcNAcylation site (57). OGT<sub>N567K</sub> activity was 12-fold reduced against P1 (Fig. 2*A* and Table 1) and not detectable against P2 (Fig. 2*A*). Next, OGT activity was measured against an intact protein substrate, TAK-1 binding protein (TAB1), where the reaction was followed with a TAB1 O-GlcNAc site-specific antibody (58). OGT<sub>N567K</sub> showed negligible glycosyltransferase activity (Fig. 2*B*).

OGT possesses a second catalytic activity in the form of promoting autocatalytic cleavage and activation of HCF1, which itself is an ID-associated gene (52). We next investigated the effects of the N567K mutation on OGT-mediated HCF1 processing. OGT proteolytic activity was measured using a GST-fusion construct of a minimal HCF1 fragment (HCF1-rep1) containing one of the PRO repeats that are known to be the targets of OGT-mediated proteolysis (49). Wild-type OGT not only promotes HCF1 proteolysis but was also able to hyperglycosylate this protein fragment (Fig. 2*C*). HCF1<sub>E1019Q</sub>, which lacks the key glutamate required for proteolytic processing, was not cleaved by wild-type OGT. Strikingly, even after 8 h of reaction time, OGT<sub>N567K</sub> was able to neither glycosylate nor cleave HCF1-rep1 (Fig. 2*C*). Taken together, the data show that the N567K mutation abrogates OGT activity.

**The N567K Mutation Disrupts the OGT Acceptor Binding Site.** To understand the molecular basis of the observed loss of OGT catalytic activity, we next initiated structural characterization of OGT<sub>N567K</sub>. Inspection of the wild-type OGT structure reveals that Asn567 maps to a loop in the N-terminal lobe of the catalytic domain that is completely conserved through evolution from *Caenorhabditis elegans* to man (Fig. 1*C*). The Asn567 side chain borders the -2 subsite as part of the OGT acceptor substrate binding cleft (Fig. 2*D*). Therefore, we hypothesized that the N567K mutation would affect OGT acceptor substrate binding, without affecting the overall OGT fold. Recombinant OGT<sub>N567K</sub> was obtained from *Escherichia coli* using construct boundaries previously employed to crystallize OGT<sub>WT</sub>-substrate complexes (57). OGT<sub>N567K</sub> was then cocrystallized with a donor analog, UDP-5S-GlcNAc, and 2 OGT acceptor peptide substrates, P1 (KKVPVSRA) and P2 (KKVAVSRA), previously used for activity measurements (Fig. 2*A*). Only the condition containing P2, which contains an alanine at the -2 subsite, yielded crystals that diffracted to 2.3 Å and allowed structure solution by molecular replacement and refinement (Table 2). The OGT<sub>N567K</sub> structure superposed onto that of OGT<sub>WT</sub> with an rmsd of 0.4 Å across 675 C<sub>α</sub> atoms, suggesting the mutation does not induce large conformational changes in the catalytic domain. Initial unbiased difference electron density maps revealed partial density corresponding to the acceptor peptide (Fig. 2*D*). Surprisingly, this only covered the P2 backbone in the +2 through to -2 subsites, suggesting that the N567K mutation may induce increased flexibility at the N-terminal tail of the acceptor peptide (Fig. 2*D* and *SI Appendix*, Fig. S2). In agreement with this, a comparative B-factor analysis suggested that the acceptor peptide bound to OGT<sub>N567K</sub> is substantially more flexible than that bound to OGT<sub>WT</sub> (*SI Appendix*, Fig. S2). Close inspection of the OGT<sub>N567K</sub> acceptor substrate binding cleft revealed that the protrusion caused by the N567K mutation partially occupies the position of the -2 proline frequently observed in OGT acceptor peptides (57, 59, 60). In summary, the N567K mutation appears to disrupt the OGT acceptor binding site.



**Fig. 2.** The N567K mutation abrogates OGT activity due to disrupted acceptor binding site. (A) Michaelis–Menten kinetics of OGT glycosyltransferase activity against P1 and P2 peptides. (B) Immunoblots showing OGT glycosyltransferase activity against TAB1. (C) Immunoblots showing OGT glycosyltransferase and proteolytic activities against HCF1-rep1. HCF1-rep1: GST-tagged host cell factor-1 fragment containing the first PRO repeat. W: wild type. E: E1019Q. (D) Crystal structures of OGT ternary complexes of OGT<sub>WT</sub> and OGT<sub>N567K</sub> active site in complex with UDP-5S-GlcNAc (U5G, turquoise, orange atoms) and acceptor peptide. OGT<sub>WT</sub> is shown in complex with RB2 (Retinoblastoma-like protein 2 peptide, light gray carbon atoms) [PDB ID code: 5C1D (57)]; OGT<sub>N567K</sub> is shown with P2 peptide (pink carbon atoms) (PDB ID code 6IBO). K567 is shown in red. An unbiased  $F_o - F_c$  difference map before inclusion of the peptide is shown as a mesh.

**Table 1. Michaelis–Menten kinetics of OGT activity against P1 peptide substrate**

Enzyme	$K_M$ (nM·s <sup>-1</sup> )	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_M$
OGT <sup>WT</sup>	99 ± 19	119 ± 11	1.2 ± 0.3
OGT <sup>N567K</sup>	144 ± 63	20 ± 4	0.1 ± 0.2

**The *sxc*<sup>N595K</sup> Mutation Leads to a Reduction of Protein O-GlcNAcylation in *D. melanogaster*.** Lack of functional OGT causes lethality and developmental arrest in most metazoan model organisms, including mice, frogs, zebrafish, and fruit flies (19, 22, 24, 61). To probe if the N567K mutation directly affects OGT catalytic activity in vivo, we exploited the genetically tractable system *D. melanogaster*, where phenotypes arisen due to null and hypomorphic alleles of *sxc* (the fly *ogt* ortholog) have been described (62, 63). *Drosophila sxc* null mutants (*sxc*<sup>1</sup> and *sxc*<sup>6</sup>) die at the pharate adult stage with homeotic transformation defects (62). A hypomorph allele, *sxc*<sup>H537A</sup>, having substantially reduced OGT catalytic activity (64), shows significantly decreased protein O-GlcNAcylation levels (63) yet produce fertile offspring. Crucially, the fruit fly provides a platform to dissect the different effects of the OGT N567K mutation. In the fly, feedback between proteome O-GlcNAcylation levels and OGA/OGT protein/mRNA levels has not been observed. Furthermore, although *Drosophila Hcf1* is heavily O-GlcNAc modified, proteolytic processing is catalyzed by the Taspase I protease instead of OGT (65).

We introduced the OGT N567K equivalent mutation (N595K) into *D. melanogaster* using a CRISPR/Cas9 gene-editing toolkit. *Vasa*:Cas9 embryos were injected in-house with a mixture of plasmids coding for the guide RNA and repair template DNA. In total, 100 potential candidate flies were screened exploiting restriction fragment-length polymorphism arising from the loss of a restriction digestion site introduced in parallel with the N595K mutation (*SI Appendix*, Fig. S3). We recovered 2 independent knockin lines (10.3 and 19.1) carrying the mutation, *sxc*<sup>N595K</sup>. Presence of the mutation was confirmed by sequencing the region ~250 base pairs upstream and downstream from the mutation. Additionally, the lines were further validated by sequencing the full-length *sxc* mRNA. Both lines developed to adulthood without apparent homeotic defects. Transheterozygous *sxc*<sup>1/sxc</sup><sup>H537A</sup>, *sxc*<sup>1/sxc</sup><sup>N595K-19.1</sup>, and *sxc*<sup>1/sxc</sup><sup>N595K-10.3</sup> flies with presumably further reduced OGT activity were viable (*SI Appendix*, Fig. S4).

To probe the effect of the N595K mutation on OGT enzymatic activity and stability in vivo, we subjected adult fly head and embryo lysates to Western blotting with an O-GlcNAc antibody, RL2 (Fig. 3A and *SI Appendix*, Fig. S5). O-GlcNAc levels were reduced to levels comparable to those in the hypomorph *sxc*<sup>H537A</sup> flies (Fig. 3A and *SI Appendix*, Fig. S5). OGT protein levels are expressed at wild-type level (Fig. 3A and *SI Appendix*, Fig. S5). Taken together, phenotypic and molecular characterization of *sxc*<sup>N595K</sup> *Drosophila* lines revealed that the mutation leads to reduction of protein O-GlcNAcylation.

**The N567K Mutation Leads to a mESC Neuronal Differentiation Phenotype.** The twins carrying the N567K OGT mutation, and the previously published male patients carrying OGT mutations in the TPR domain, all suffered from neurodevelopmental delay. We next explored the early events of neuronal development by studying effects on pluripotency and differentiation in mESCs, a cellular system amenable to genetic modification. We generated male mESC lines expressing C-terminal triple hemagglutinin (3HA)-tagged version of wild-type OGT (3HA-OGT<sup>WT</sup>) and carrying the N567K mutation (3HA-OGT<sup>N567K</sup>) using 2 rounds of CRISPR/Cas9-mediated gene editing at the endogenous locus (*SI Appen-*

*dix*, Figs. S6 and S7). At each step, at least 2 clones with the correct genotype were isolated and selected for further analysis.

Next, we examined whether neuronal differentiation was altered in 3HA-OGT<sup>N567K</sup> cells. Expression of pluripotency markers Sox2 and Oct4 was unaffected during the first 6 d of differentiation (Fig. 3 C and D). Neurite outgrowth was assessed after 8 d. Neurite length was significantly reduced in 3HA-OGT<sup>N567K</sup> cells compared with 3HA-OGT<sup>WT</sup> control cells immune-labeled against the neuron-specific  $\beta$ -tubulin III protein (Fig. 3 E and F). Thus, the N567K mutation leads to an mESC neuronal differentiation phenotype.

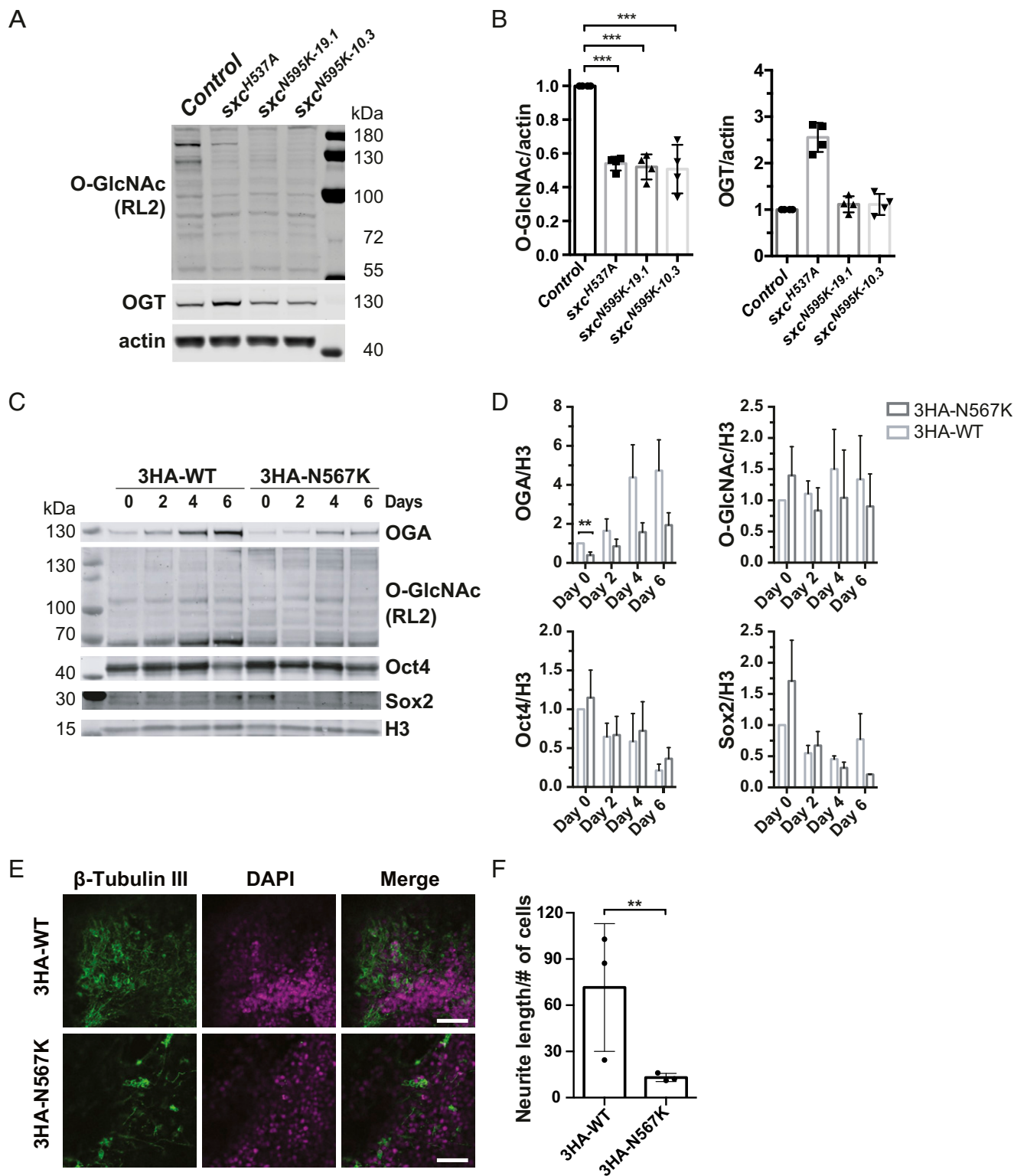
**The N567K Mutation Leads to Misprocessing of HCF1.** We next attempted to uncover the molecular mechanisms underpinning the observed differentiation phenotypes. We tested the levels of protein O-GlcNAcylation using Western blotting (Fig. 4A). Surprisingly, levels of protein O-GlcNAcylation in pluripotent mESCs appeared comparable in 3HA-OGT<sup>N567K</sup> and 3HA-OGT<sup>WT</sup> cells. Previous studies have shown that (as yet to be discovered) feedback mechanism regulates OGA and OGT expression levels in response to protein O-GlcNAcylation levels (42, 43, 66, 67). Indeed, while OGT levels appeared to be unaltered, OGA levels were abrogated in 3HA-OGT<sup>N567K</sup> pluripotent mESCs (Fig. 4 A and B), presumably to compensate for reduced OGT activity, but appeared normal during differentiation (Fig. 3 C and D).

Next, we investigated the effects of the N567K mutation on the posttranslational maturation and cellular localization of HCF1 in pluripotent mESCs. Western blot analysis on cytoplasmic and nuclear fractions revealed an elevated ratio of full-length versus total HCF1 in the nuclear fractions of 3HA-OGT<sup>N567K</sup> compared with 3HA-OGT<sup>WT</sup> mESCs, while the levels of HCF1 proteolytic products were reduced (Fig. 4 C and D). This is in agreement with

**Table 2. X-ray diffraction data processing and refinement statistics of the OGT<sup>N567K</sup> ternary complex**

	OGT <sup>N567K</sup>
Data collection	
Space group	F222
Cell dimensions	
$\alpha, \beta, \gamma$ (°)	137.7, 150.6, 200.0
A, B, C (Å)	90.0, 90.0, 90.0
Resolution (Å)	45.31–2.17 (2.17–2.23)
$R_{sym}$ or $R_{merge}$	0.07 (0.91)
$I/\sigma I$	11.3 (1.3)
Completeness (%)	100 (99)
Redundancy	5.2 (5.3)
Refinement	
Resolution (Å)	45.31–2.17 (2.17–2.23)
No. reflections	286,887 (23,302)
$R_{work}/R_{free}$	0.20/0.25
No. atoms	
Protein	5,530
Nucleotide sugar	39
Peptide	35
Water	328
$B$ -factors	
Protein	53.2
Nucleotide sugar	46.6
Peptide	91.9
Water	51.3
Rmsd	
Bond lengths (Å)	0.02
Bond angles (°)	1.8

Numbers in brackets show the highest resolution bin.



**Fig. 3.** The N567K OGT mutation leads to altered O-GlcNAc homeostasis in *D. melanogaster* and neurite outgrowth defect in mESCs. (A) Immunoblot on 3- to 5-d-old adult *Drosophila* head lysates. (B) Quantification of O-GlcNAc modification on proteins and OGT protein level, normalized to actin signal. Protein O-GlcNAc modification is decreased in *sxc*<sup>N595K 19.1</sup> ( $***P < 0.0001$ ) and *sxc*<sup>N595K 10.3</sup> ( $***P < 0.0001$ ) lines to a comparable level found in the hypomorphic *sxc*<sup>H537A</sup> flies ( $***P < 0.0001$ ).  $n = 4$ , mean  $\pm$  SD. ANOVA with Tukey test. (C) Immunoblots on full cell lysate of 3HA-OGT<sup>WT</sup> and 3HA-OGT<sup>N567K</sup> mESCs showing OGA, Oct4, Sox2, Histone 3 (H3), and protein O-GlcNAcylation (RL2) levels. mESCs were differentiated for 0, 2, 4, and 6 d in N2B27 medium. (D) Quantification of immunoblots of OGA, Oct4, Sox2, and global protein O-GlcNAcylation in differentiated 3HA-OGT<sup>WT</sup> and 3HA-OGT<sup>N567K</sup> mESCs normalized to H3 signal.  $n = 4$ , mean  $\pm$  SEM;  $**P = 0.012$ . Multiple *t* test using the Holm–Sidak method. (E) Immunofluorescence microscopy of 3HA-OGT<sup>WT</sup> and 3HA-OGT<sup>N567K</sup> mESCs after 8 d of neuronal differentiation in N2B27 medium. Cells were stained for tubulin III (green) and DAPI (magenta). (Scale bar, 50  $\mu$ m). (F) Quantification of neurite outgrowth. Total neurite length was measured per microscopic image by tracking  $\beta$ -tubulin III signal and then normalized to the number of nuclei (DAPI). (3HA-OGT<sup>WT</sup>, 108 to 800 nuclei per image; 3HA-OGT<sup>N567K</sup>, 260 to 740 nuclei per image)  $**P = 0.0079$ , Student *t* test, mean  $\pm$  SD,  $n = 3$ .

the impaired OGT-mediated proteolytic cleavage of HCF1 observed *in vitro* (Fig. 2C). HCF1 contributes to the regulation of transcriptional programs associated with controlling cell cycle and pluripotency (68, 69). Interestingly, the expression of some of the HCF1 interactors are also regulated by HCF1 itself (51, 66, 69). As an initial attempt to uncover possible candidate genes affected by HCF1 misprocessing, we investigated the expression levels of known HCF1 partners by qRT-PCR (Fig. 4E). The Ets transcription factor GA-binding protein subunit  $\alpha$  (GABPA), a component of the nuclear respiratory factor-2 (NRF2) complex, showed a 2.3-fold increase in mRNA level in 3HA-OGT<sup>N567K</sup> cells compared with 3HA-OGT<sup>WT</sup> control samples ( $P = 0.002$ ). GABPA/NRF2 regulates mitochondrial biogenesis and cell cycle progression (70, 71). Taking these data together, the OGT N567K mutation leads to misprocessing of HCF1 in mESCs.

### Concluding Remarks

Previous studies have established a link between missense mutations in OGT and XLID in young males. These patients are hemizygous for the mutant OGT allele, with this genetic defect being present in all cells from early development to later life (41–43). However, residing exclusively in the OGT TPR domain, it was not clear whether these mutations led to the observed phenotypes through loss of glycosyltransferase activity, HCF1 misprocessing, or disruption of the OGT interactome (41–43). Here, we have presented female twins with XLID who are heterozygous for an N567K mutation in the OGT catalytic domain. Despite a wild-type *ogt* allele being present in all cells of these patients, its expression is arbitrarily silenced due to X chromosome inactivation at the late blastocyst stage of embryogenesis to compensate for gene dosage (72). Thus, these females represent an example of mosaic expression of an OGT XLID variant causing developmental abnormalities. Previous examples of female carriers of X-linked genetic disorders showed correlation between the X inactivation pattern and the severity of the condition (73–75). Hence, it is possible that the different etiology of ID in twins is the consequence of distinct spatial patterns and skew of X-inactivation.

The N567K variant of OGT retains negligible levels of catalytic activity. Thus, a possible mechanism underpinning the XLID phenotype is reduced protein O-GlcNAcylation on a subset of OGT target proteins at specific time points of development or differentiation. Previous genetic studies in mice have highlighted the vital role of OGT in both neuronal development and adult life (21, 33, 37–39). OGT is capable of transferring GlcNAc onto over 4,000 different substrate proteins; it possesses an acceptor-peptide specificity with a preference to a degenerated sequon motif (57, 76, 77). Our structural analysis has uncovered that the N567K mutation changes the surface of the substrate binding pocket, thereby altering substrate binding ability of the N567K OGT. Previously reported XLID associated mutations lie in the TPR domain of OGT (40–43, 78), so potentially influencing its scaffolding function or reducing O-GlcNAcylation on an as yet unknown, subset of substrates. The N567K mutation may affect the O-GlcNAc proteome globally yet leads to a similar clinical phenotype, suggesting it is the catalytic activity/dosage of OGT that is important.

We have identified one of the OGT substrates, HCF1, as a potential candidate that could at least partially explain some of the OGT N567K patient phenotypes. We showed that O-GlcNAcylation and proteolytic activation of HCF1 is significantly decreased, including in mESC nuclei. The HCF1 protein plays an important role in cell growth and cell cycle control through interacting with transcriptional complexes and epigenetic regulators (79–83). The HCF1 gene itself has been identified as an intellectual disability gene causing cobalamin type X deficiency, craniofacial abnormalities, and prenatal onset of microcephaly in the most extreme

cases (52, 53, 84). Less-severe HCF1 mutations have been associated with ID and autism spectrum disorder without cobalamin deficiency (54). Patients with N567K OGT variant did not suffer from cobalamin deficiency and expressed a much milder form of congenital anomalies than the severe HCF1 group, suggesting that a degree of HCF1 activity is retained. Interestingly, we observed an increase in full-length HCF1 protein level and decreased level of the mature proteolytic fragment in nuclei, showing that HCF1 processing is affected *in cellulo*, albeit with retention of some activity. Surprisingly, the expression of the transcriptional factor GABPA downstream of HCF1 was increased indicating that the N567K mutation in OGT could translate into moderate molecular changes downstream of HCF1. The opposite effect, decreased expression of GABPA, was detected upon overexpression of the HCF1 N-terminal fragment (HCF<sub>N</sub>) (68). GABPA belongs to the E-26 family of DNA binding factors and regulates expression of several proteins required for mitochondrial DNA transcription and replication (85). Moreover, GABPA modulates the expression of genes involved in cell cycle control, apoptosis, differentiation, and hormonal regulation. Its importance in early embryonic development was demonstrated in knockout mouse studies (86). However, there is no established link between GABPA and neurodevelopment. Furthermore, expression of several other transcriptional factors targeted by HCF1—such as *Creb1*, *E2F4*, *Krox-2*, *Sin3a*, *Thap2*, and *Zfp143*—were found unaltered, suggesting that partial abrogation of HCF1 cleavage may affect just a small subset of targets, whose expression is influenced by the altered ratio of proteolytically cleaved and uncleaved HCF1.

While OGT is critical for mESC maintenance of pluripotency and differentiation, we have detected no difference at the level of pluripotency markers, Sox2 and Oct4, between 3HA-OGT<sup>WT</sup> and 3HA-OGT<sup>N567K</sup> mESCs. The pluripotency factor Sox-2 is O-GlcNAc modified at Ser248 (87) and this posttranslational modification is indispensable for sustaining pluripotency. However, its O-GlcNAc modification is not required for differentiation, with cells expressing an O-GlcNAc-deficient mutant version of Sox2 exhibited enhanced reprogramming ability (87). Furthermore, mESCs expressing O-GlcNAc-deficient Sox2 displayed changes in their transcriptional network, specifically increasing the expression of genes vital to maintain pluripotency (87). It is feasible that a similar mechanism is responsible to compensate for pluripotency defects induced by the N567K OGT mutation.

A *Drosophila* model of the N567K OGT mutation showed reduced protein O-GlcNAcylation. However, the abundance of O-GlcNAc-modified proteins was similar in 3HA-OGT<sup>WT</sup> and 3HA-OGT<sup>N567K</sup> mESCs. Interestingly, we observed decreased OGA protein levels as a compensatory regulatory mechanism responsible for maintaining protein O-GlcNAc levels in undifferentiated mESCs. Our finding is in line with previous data reported on patient-derived cell lines carrying XLID mutations in the TPR domain of OGT (42, 43). Although there is no evidence for a link between OGA and XLID, given the role of OGA in regulating transcription (88–91), it cannot be excluded that changes in OGA levels directly contribute to the XLID phenotype.

In agreement with previous studies implying a requirement for OGT function in normal neuron differentiation, morphology, and health (27, 37, 92–94), abnormal OGT activity in 3HA-OGT<sup>N567K</sup> cells resulted in reduced neurite outgrowth. These data suggest a possible mechanistic link between the mutation and the microcephaly and neurodevelopmental cognitive impairments observed in the patients.

In summary, we have shown that impaired catalytic activity of OGT leads to neurodevelopmental defects in humans, through combined downstream effects of impaired OGT activity and reduced HCF1 proteolytic processing. Further studies are required to define the mechanisms downstream of impaired OGT





catalytic activity that affect neurodevelopment resulting in intellectual disability.

## Materials and Methods

The study was approved by the University of Dundee Institutional Risk Assessment committee, and 1 of each participants' parents provided written, informed consent. The DDD study has UK Research Ethics Committee (REC) approval (10/H0305/83, granted by the Cambridge South Research Ethics Committee, and GEN/284/12 granted by the Republic of Ireland REC).

**X-Inactivation Analysis.** To determine if X chromosome inactivation occurred randomly or nonrandomly, the ratio of methylated and unmethylated maternal and paternal X chromosomes was determined. (SI Appendix, Fig. S1). Sites are methylated on the inactive X chromosome and unmethylated on the active X chromosome. The method utilizes the androgen receptor sequence that is encoded on the X chromosome (Xq11.2-q12) and contains a highly polymorphic polyglutamine (CAG) trinucleotide repeat in exon 1. There are 20 alleles of androgen receptor, with 11 to 31 repeats reported in the general population and 85 to 90% of patients are heterozygous for this locus, thus suitable for X-inactivation analysis. For experimental procedure, see SI Appendix, SI Materials and Methods.

**Structure Solution.** Crystallization of truncated OGT<sub>WT/N567K</sub> (residues 323–1044) was performed as described previously (57) (for protein expression, purification, and crystallization, see SI Appendix, SI Materials and Methods). The structure was solved by molecular replacement using the structure for OGT<sub>WT</sub> [PDB ID code 5C1D (57)] as the search model. The resulting model was manually refined using Coot (95) and REFMAC (96), respectively. The editing and refinement of the model was iterated until it was in complete agreement with the data. Scaling and model building statistics can be seen in Table 2.

**Enzyme Activity Assays.** Michaelis–Menten kinetics of OGT were measured using a fluorometric assay, as described previously (97), with the exception of reduced reaction volume of 25  $\mu$ L and usage of 384-well plate. As acceptor substrate, P1 (KKVPSVRA) and P2 (KKVAVSRA) were used. Additional O-GlcNAcylation assays were performed on TAB1 protein (residues 7–420). Proteolytic assays were performed as described previously (43), except a noncleavable HCF1-rep1 fragment was used instead of OGA treatment as the negative control. For further details, see SI Appendix, SI Materials and Methods.

**mESC Culture.** mESC AW2 line was acquired from the MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, University of Edinburgh (98). mESCs were cultured in an undifferentiated state in 0.1% gelatin (wt/vol)-coated plates in GMEM BHK-21 supplemented with 10% FBS (vol/vol), 0.1 mM MEM nonessential amino acids, 1 mM sodium pyruvate, 0.1 mM 2-mercaptoethanol, and 100 units/mL LIF at 37 °C in 5% CO<sub>2</sub>. For differentiation assays, cells were cultured following a previously published protocol (99). Briefly, cells were cultured in DMEM and F12 media mixed at 1:1 ratio (vol/vol), supplemented with modified N2 (25 mg/mL insulin, 100  $\mu$ g/mL apo-transferrin, 6 ng/mL progesterone, 16  $\mu$ g/mL putrescine, 30 nM sodium selenite, and 50  $\mu$ g/mL BSA fraction) and neurobasal supplement B27. Medium was renewed every 2 d. For generation of CRISPR/Cas9 lines, Immunocytochemistry, Western blotting, and qPCR analyses details are provided in SI Appendix, SI Materials and Methods.

**Drosophila Husbandry, Microinjection, and Genetics.** Fly stocks were maintained on Nutri-Fly Bloomington Formulation fly food at 25 °C. Vasa:Cas9 (#51323), *sxc*<sup>1</sup>, *bw*<sup>1</sup>, *sp*<sup>1</sup>/*ISM5* (#3058) and *Df(2R)BSC630/CyO* (#25705) stocks were obtained from the Bloomington *Drosophila* Stock Centre. The *sxc* hypomorphic allele, *sxc*<sup>H537A</sup>, was described in ref. 63. The *sxc*<sup>N595K</sup> *Drosophila* lines were generated by CRISPR/Cas9 mutagenesis, as described in SI Appendix, SI Materials and Methods.

**Western Blotting from Drosophila Samples.** Sample preparation of embryos and fly heads and Western blotting were performed as described in SI Appendix, SI Materials and Methods.

**ACKNOWLEDGMENTS.** This work was funded by a Wellcome Trust Investigator Award (110061) (to D.M.F.v.A.). V.M.P. was supported by the SPRINT/MND-MS Scholarship (University of Edinburgh). The X-ray diffraction experiments were performed on beamline ID23-2 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to Local Contact at the ESRF for providing assistance for using beamline ID23-2. We acknowledge Dr. Louise A. Young for the aid in the X-inactivation analysis and Dr. Andrew Smith for the AW2 mouse embryonic stem cell line. We acknowledge the Deciphering Developmental Disorders study, presenting independent research commissioned by the Health Innovation Challenge Fund (grant no. HICF-1009-003), a parallel funding partnership between Wellcome and the Department of Health, and the Wellcome Sanger Institute (grant no. WT098051). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. The authors would also like to thank the Genome Aggregation Database (gnomAD) and the groups that provided exome and genome variant data to this resource.

1. Y. Zhu *et al.*, O-GlcNAc occurs cotranslationally to stabilize nascent polypeptide chains. *Nat. Chem. Biol.* **11**, 319–325 (2015).
2. N. Zachara, Y. Akimoto, G. W. Hart, *The O-GlcNAc Modification* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2017).
3. S. Constable, J.-M. Lim, K. Vaidyanathan, L. Wells, O-GlcNAc transferase regulates transcriptional activity of human Oct4. *Glycobiology* **27**, 927–937 (2017).
4. N. Lamarre-Vincent, L. C. Hsieh-Wilson, Dynamic glycosylation of the transcription factor CREB: A potential role in gene regulation. *J. Am. Chem. Soc.* **125**, 6612–6613 (2003).
5. S. M. Ranuncolo, S. Ghosh, J. A. Hanover, G. W. Hart, B. A. Lewis, 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 (2012).
6. R. Deplus *et al.*, TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.* **32**, 645–655 (2013).
7. C. S. Chu *et al.*, O-GlcNAcylation regulates EZH2 protein stability and function. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1355–1360 (2014).
8. M. C. Gambetta, J. Müller, O-GlcNAcylation prevents aggregation of the Polycomb group repressor polyhomeotic. *Dev. Cell* **31**, 629–639 (2014).
9. Y. Skorobogatko *et al.*, O-linked  $\beta$ -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 (2014).
10. C. Peng *et al.*, Regulation of the hippo-YAP pathway by glucose sensor O-GlcNAcylation. *Mol. Cell* **68**, 591–604.e5 (2017).
11. N. E. Zachara, G. W. Hart, O-GlcNAc as a sensor of cellular state: The role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress. *Biochim. Biophys. Acta* **1673**, 13–28 (2004).
12. T. Ohn, N. Kedersha, T. Hickman, S. Tisdale, P. Anderson, A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nat. Cell Biol.* **10**, 1224–1231 (2008).
13. C. Han *et al.*, O-GlcNAcylation of SIRT1 enhances its deacetylase activity and promotes cytoprotection under stress. *Nat. Commun.* **8**, 1491 (2017).
14. L. M. Andres *et al.*, Chemical modulation of protein O-GlcNAcylation via OGT inhibition promotes human neural cell differentiation. *ACS Chem. Biol.* **12**, 2030–2039 (2017).
15. H. Jang *et al.*, O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell* **11**, 62–74 (2012).
16. S. P. Durning, H. Flanagan-Steet, N. Prasad, L. Wells, O-linked  $\beta$ -N-acetylglucosamine (O-GlcNAc) acts as a glucose sensor to epigenetically regulate the insulin gene in pancreatic beta cells. *J. Biol. Chem.* **291**, 2107–2118 (2016).
17. G. W. Hart, C. Slawson, G. Ramirez-Correa, O. Lagerlof, Cross talk between O-GlcNAcylation and phosphorylation: Roles in signaling, transcription, and chronic disease. *Annu. Rev. Biochem.* **80**, 825–858 (2011).
18. B. Guo *et al.*, O-GlcNAc-modification of SNAP-29 regulates autophagosome maturation. *Nat. Cell Biol.* **16**, 1215–1226 (2014).
19. R. Shafi *et al.*, 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. U.S.A.* **97**, 5735–5739 (2000).
20. T. Miura, S. Nishihara, O-GlcNAc is required for the survival of primed pluripotent stem cells and their reversion to the naïve state. *Biochem. Biophys. Res. Commun.* **480**, 655–661 (2016).
21. N. O'Donnell, N. E. Zachara, G. W. Hart, J. D. Marth, Ogt-dependent X-chromosome-linked protein glycosylation is a requisite modification in somatic cell function and embryo viability. *Mol. Cell Biol.* **24**, 1680–1690 (2004).
22. D. M. Webster *et al.*, O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development. *BMC Dev. Biol.* **9**, 28 (2009).
23. D. A. R. Sinclair *et al.*, *Drosophila* O-GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (*sxc*). *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13427–13432 (2009).
24. M. C. Gambetta, K. Oktaba, J. Muller, Essential role of the glycosyltransferase *Sxc/Ogt* in polycomb repression. *Science* **325**, 93–96 (2009).
25. L. K. Kreppel, M. A. Blomberg, G. W. Hart, Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *J. Biol. Chem.* **272**, 9308–9315 (1997).
26. W. A. Lubas, D. W. Frank, M. Krause, J. A. Hanover, O-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *J. Biol. Chem.* **272**, 9316–9324 (1997).
27. E. W. Taylor *et al.*, O-GlcNAcylation of AMPA receptor GluA2 is associated with a novel form of long-term depression at hippocampal synapses. *J. Neurosci.* **34**, 10–21 (2014).
28. J. E. Rexach *et al.*, Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nat. Chem. Biol.* **8**, 253–261 (2012).

29. E. L. Ardiel *et al.*, CaMK (CMK-1) and O-GlcNAc transferase (OGT-1) modulate mechanosensory responding and habituation in an interstimulus interval-dependent manner in *Caenorhabditis elegans*. *bioRxiv*:10.1101/115972 (10 March 2017).
30. E. Y. Kim *et al.*, A role for O-GlcNAcylation in setting circadian clock speed. *Genes Dev.* **26**, 490–502 (2012).
31. L. Sui *et al.*, Clustered DNA damage induced by protons radiation in plasmid. *Chin. Sci. Bull.* **58**, 3217–3223 (2013).
32. C. L. Dai, J. H. Gu, F. Liu, K. Iqbal, C. X. Gong, Neuronal O-GlcNAc transferase regulates appetite, body weight, and peripheral insulin resistance. *Neurobiol. Aging* **70**, 40–50 (2018).
33. O. Lagerlöf *et al.*, The nutrient sensor OGT in PVN neurons regulates feeding. *Science* **351**, 1293–1296 (2016).
34. C. S. Arnold *et al.*, The microtubule-associated protein tau is extensively modified with O-linked N-acetylglucosamine. *J. Biol. Chem.* **271**, 28741–28744 (1996).
35. R. N. Cole, G. W. Hart, Cytosolic O-glycosylation is abundant in nerve terminals. *J. Neurochem.* **79**, 1080–1089 (2001).
36. K. Vosseller *et al.*, O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **5**, 923–934 (2006).
37. H. Francisco *et al.*, O-GlcNAc post-translational modifications regulate the entry of neurons into an axon branching program. *Dev. Neurobiol.* **69**, 162–173 (2009).
38. O. Lagerlöf, G. W. Hart, R. L. Haganir, O-GlcNAc transferase regulates excitatory synapse maturity. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 1684–1689 (2017).
39. A. C. Wang, E. H. Jensen, J. E. Rexach, H. V. Vinters, L. C. Hsieh-Wilson, Loss of O-GlcNAc glycosylation in forebrain excitatory neurons induces neurodegeneration. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 15120–15125 (2016).
40. H. Bouazzi, G. Lesca, C. Trujillo, M. K. Alwasiyah, A. Munnich, Nonsyndromic X-linked intellectual deficiency in three brothers with a novel MED12 missense mutation [c.5922G>T (p.Glu1974His)]. *Clin. Case Rep.* **3**, 604–609 (2015).
41. N. Selvan *et al.*, 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 (2018).
42. K. Vaidyanathan *et al.*, Identification and characterization of a missense mutation in the O-linked  $\beta$ -N-acetylglucosamine (O-GlcNAc) transferase gene that segregates with X-linked intellectual disability. *J. Biol. Chem.* **292**, 8948–8963 (2017).
43. A. P. Willems *et al.*, Mutations in N-acetylglucosamine (O-GlcNAc) transferase in patients with X-linked intellectual disability. *J. Biol. Chem.* **292**, 12621–12631 (2017).
44. S. Bassani *et al.*, The neurobiology of X-linked intellectual disability. *Neuroscientist* **19**, 541–552 (2013).
45. G. Neri, C. E. Schwartz, H. A. Lubs, R. E. Stevenson, X-linked intellectual disability update 2017. *Am. J. Med. Genet. A* **176**, 1375–1388 (2018).
46. B. Boda, C. Mas, D. Müller, Activity-dependent regulation of genes implicated in X-linked non-specific mental retardation. *Neuroscience* **114**, 13–17 (2002).
47. R. E. Stevenson, C. E. Schwartz, R. C. Rogers, Malformations among the X-linked intellectual disability syndromes. *Am. J. Med. Genet. A* **161A**, 2741–2749 (2013).
48. V. Kapuria *et al.*, Proteolysis of HCF-1 by Ser/Thr glycosylation-incompetent O-GlcNAc transferase:UDP-GlcNAc complexes. *Genes Dev.* **30**, 960–972 (2016).
49. M. B. Lazarus *et al.*, HCF-1 is cleaved in the active site of O-GlcNAc transferase. *Science* **342**, 1235–1239 (2013).
50. T. Bhuiyan, P. Waridel, V. Kapuria, V. Zoete, W. Herr, Distinct OGT-binding sites promote HCF-1 cleavage. *PLoS One* **10**, e0136636 (2015).
51. B. Khurana, T. M. Kristie, A protein sequestering system reveals control of cellular programs by the transcriptional coactivator HCF-1. *J. Biol. Chem.* **279**, 33673–33683 (2004).
52. L. Huang *et al.*, A noncoding, regulatory mutation implicates HCFC1 in nonsyndromic intellectual disability. *Am. J. Hum. Genet.* **91**, 694–702 (2012).
53. L. A. Jolly *et al.*, HCFC1 loss-of-function mutations disrupt neuronal and neural progenitor cells of the developing brain. *Hum. Mol. Genet.* **24**, 3335–3347 (2015).
54. C. Koufaris, A. Alexandrou, G. A. Tanteles, V. Anastasiadou, C. Sismani, A novel HCFC1 variant in male siblings with intellectual disability and microcephaly in the absence of cobalamin disorder. *Biomed. Res. Rep.* **4**, 215–218 (2016).
55. H. V. Firth, C. F. Wright; DDD Study, The deciphering developmental disorders (DDD) study. *Dev. Med. Child Neurol.* **53**, 702–703 (2011).
56. K. J. Karczewski *et al.*, Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*:10.1101/531210 (30 January 2019).
57. S. Pathak *et al.*, The active site of O-GlcNAc transferase imposes constraints on substrate sequence. *Nat. Struct. Mol. Biol.* **22**, 744–750 (2015).
58. S. Pathak *et al.*, O-GlcNAcylation of TAB1 modulates TAK1-mediated cytokine release. *EMBO J.* **31**, 1394–1404 (2012).
59. M. B. Lazarus *et al.*, Structural snapshots of the reaction coordinate for O-GlcNAc transferase. *Nat. Chem. Biol.* **8**, 966–968 (2012).
60. M. Schimpl *et al.*, O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis. *Nat. Chem. Biol.* **8**, 969–974 (2012).
61. S. Kenwick, E. Amaya, N. Papalopolu, Pilot morpholino screen in *Xenopus tropicalis* identifies a novel gene involved in head development. *Dev. Dyn.* **229**, 289–299 (2004).
62. P. W. Ingham, A gene that regulates the bithorax complex differentially in larval and adult cells of *Drosophila*. *Cell* **37**, 815–823 (1984).
63. D. Mariappa, A. T. Ferenbach, D. M. F. van Aalten, Effects of hypo-O-GlcNAcylation on *Drosophila* development. *J. Biol. Chem.* **293**, 7209–7221 (2018).
64. D. Mariappa *et al.*, Dual functionality of O-GlcNAc transferase is required for *Drosophila* development. *Open Biol.* **5**, 150234 (2015).
65. F. Capotosti, J. J.-D. Hsieh, W. Herr, Species selectivity of mixed-lineage leukemia/trithorax and HCF proteolytic maturation pathways. *Mol. Cell. Biol.* **27**, 7063–7072 (2007).
66. K. Qian *et al.*, Transcriptional regulation of O-GlcNAc homeostasis is disrupted in pancreatic cancer. *J. Biol. Chem.*, 10.1074/jbc.RA118.004709 (2018).
67. Z. Zhang, E. P. Tan, N. J. VandenHull, K. R. Peterson, C. Slawson, O-GlcNAc expression is sensitive to changes in O-GlcNAc homeostasis. *Front. Endocrinol.* **5**, 206 (2014).
68. J. L. Vogel, T. M. Kristie, The novel coactivator C1 (HCF) coordinates multiprotein enhancer formation and mediates transcription activation by GABP. *EMBO J.* **19**, 683–690 (2000).
69. S. Deléhouzée *et al.*, GABP, HCF-1 and YY1 are involved in Rb gene expression during myogenesis. *Genes Cells* **10**, 717–731 (2005).
70. A. Jaworski, C. L. Smith, S. J. Burden, GA-binding protein is dispensable for neuromuscular synapse formation and synapse-specific gene expression. *Mol. Cell. Biol.* **27**, 5040–5046 (2007).
71. A. Perdomo-Sabogal *et al.*, Human lineage-specific transcriptional regulation through GA-binding protein transcription factor alpha (GABPA). *Mol. Biol. Evol.* **33**, 1231–1244 (2016).
72. I. Okamoto *et al.*, Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* **472**, 370–374 (2011).
73. R. C. Allen, R. G. Nachtman, H. M. Rosenblatt, J. W. Belmont, Application of carrier testing to genetic counseling for X-linked agammaglobulinemia. *Am. J. Hum. Genet.* **54**, 25–35 (1994).
74. K. H. Ørstavik *et al.*, X chromosome inactivation in carriers of Barth syndrome. *Am. J. Hum. Genet.* **63**, 1457–1463 (1998).
75. G. P. S. Knudsen *et al.*, Increased skewing of X chromosome inactivation in Rett syndrome patients and their mothers. *Eur. J. Hum. Genet.* **14**, 1189–1194 (2006).
76. J. C. Trinidad *et al.*, Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Mol. Cell. Proteomics* **11**, 215–229 (2012).
77. R. Jochmann, P. Holz, H. Sticht, M. Stürzl, Validation of the reliability of computational O-GlcNAc prediction. *Biochim. Biophys. Acta* **1844**, 416–421 (2014).
78. M. Gundogdu *et al.*, The O-GlcNAc transferase intellectual disability mutation L254F distorts the TPR helix. *Cell Chem. Biol.* **25**, 513–518.e4 (2018).
79. Y. J. Machida, Y. Machida, A. A. Vashisht, J. A. Wohlschlegel, A. Dutta, The deubiquitinating enzyme BAP1 regulates cell growth via interaction with HCF-1. *J. Biol. Chem.* **284**, 34179–34188 (2009).
80. H. Yu *et al.*, The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression. *Mol. Cell. Biol.* **30**, 5071–5085 (2010).
81. M. Dejosez *et al.*, Ronin/Hcf-1 binds to a hyperconserved enhancer element and regulates genes involved in the growth of embryonic stem cells. *Genes Dev.* **24**, 1479–1484 (2010).
82. E. Julien, W. Herr, Proteolytic processing is necessary to separate and ensure proper cell growth and cytokinesis functions of HCF-1. *EMBO J.* **22**, 2360–2369 (2003).
83. S. Rodriguez-Jato, A. Busturia, W. Herr, *Drosophila melanogaster* dHCF interacts with both PcG and TrxG epigenetic regulators. *PLoS One* **6**, e27479 (2011).
84. H.-C. C. Yu *et al.*, An X-linked cobalamin disorder caused by mutations in transcriptional coregulator HCFC1. *Am. J. Hum. Genet.* **93**, 506–514 (2013).
85. F. Bruni, P. L. Polosa, M. N. Gadaleta, P. Cantatore, M. Roberti, Nuclear respiratory factor 2 induces the expression of many but not all human proteins acting in mitochondrial DNA transcription and replication. *J. Biol. Chem.* **285**, 3939–3948 (2010).
86. S. Risteovski *et al.*, The ETS transcription factor GABPalpha is essential for early embryogenesis. *Mol. Cell. Biol.* **24**, 5844–5849 (2004).
87. S. A. Myers *et al.*, SOX2 O-GlcNAcylation alters its protein-protein interactions and genomic occupancy to modulate gene expression in pluripotent cells. *Elife* **5**, e10647 (2016).
88. I. Akan, D. C. Love, K. R. Harwood, M. R. Bond, J. A. Hanover, *Drosophila* O-GlcNAcase deletion globally perturbs chromatin O-GlcNAcylation. *J. Biol. Chem.* **291**, 9906–9919 (2016).
89. Z. Zhang *et al.*, O-linked N-acetylglucosamine (O-GlcNAc) transferase and O-GlcNAcase interact with Mi2 $\beta$  protein at the  $\alpha\gamma$ -globin promoter. *J. Biol. Chem.* **291**, 15628–15640 (2016).
90. M. Resto *et al.*, O-GlcNAcase is an RNA polymerase II elongation factor coupled to pausing factors SPT5 and TIF1 $\beta$ . *J. Biol. Chem.* **291**, 22703–22713 (2016). Correction in: *J. Biol. Chem.* **292**, 16524–16525 (2017).
91. S. Das *et al.*, O-GlcNAcylation of GLI transcription factors in hyperglycemic conditions augments Hedgehog activity. *Lab. Invest.* **99**, 260–270 (2019).
92. H. Hwang, H. Rhim, Functional significance of O-GlcNAc modification in regulating neuronal properties. *Pharmacol. Res.* **129**, 295–307 (2018).
93. J. Fourneau, M. H. Canu, C. Cieniewski-Bernard, B. Bastide, E. Dupont, Synaptic protein changes after a chronic period of sensorimotor perturbation in adult rats: A potential role of phosphorylation/O-GlcNAcylation interplay. *J. Neurochem.* **147**, 240–255 (2018).
94. M. K. Tallent *et al.*, In vivo modulation of O-GlcNAc levels regulates hippocampal synaptic plasticity through interplay with phosphorylation. *J. Biol. Chem.* **284**, 174–181 (2009).
95. M. D. Winn *et al.*, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
96. A. A. Vagin *et al.*, REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2184–2195 (2004).
97. V. S. Borodkin *et al.*, Bisubstrate UDP-peptide conjugates as human O-GlcNAc transferase inhibitors. *Biochem. J.* **457**, 497–502 (2014).
98. X. Zhou *et al.*, Hes1 desynchronizes differentiation of pluripotent cells by modulating STAT3 activity. *Stem Cells* **31**, 1511–1522 (2013).
99. W. Wongpaiboonwattana, M. P. Stavridis, Neural differentiation of mouse embryonic stem cells in serum-free monolayer culture. *J. Vis. Exp.*, e52823 (2015).