ACCELERATED PUBLICATION

A mutant O-GlcNAcase as a probe to reveal global dynamics of protein O-GlcNAcylation during *Drosophila* embryonic development

Daniel Mariappa*[1], Nithya Selvan*†[1], Vladimir S. Borodkin*, Jana Alonso*, Andrew T. Ferenbach*, Claire Shepherd‡, Iva Hopkins Navratilova‡ and Daan M.F. van Aalten*‡[2]

*MRC Protein Phosphorylation and Ubiquitylation Unit, College of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, U.K.
†Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, U.K.
‡MRC-PPUU and Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, U.K.

O-GlcNAcylation is a reversible type of serine/threonine glycosylation on nucleocytoplasmic proteins in metazoa. Various genetic approaches in several animal models have revealed that O-GlcNAcylation is essential for embryogenesis. However, the dynamic changes in global O-GlcNAcylation and the underlying mechanistic biology linking them to embryonic development is not understood. One of the limiting factors towards characterizing changes in O-GlcNAcylation has been the limited specificity of currently available tools to detect this modification. In the present study, harnessing the unusual properties of an O-GlcNAcase (OGA) mutant that binds O-GlcNAc (O-*N*-acetylglucosamine) sites with nanomolar affinity, we uncover changes in protein O-GlcNAcylation as a function of *Drosophila* development.

Key words: O-GlcNAcase (OGA), Fluorescence polarization, *Drosophila*, embryonic development, O-GlcNAc probe.

INTRODUCTION

Protein O-GlcNAcylation is a dynamic nucleocytoplasmic post-translational modification (PTM) involving the transfer of *N*-acetylglucosamine (GlcNAc) to protein serine/threonine residues [1]. O-GlcNAc transfer is catalysed by O-GlcNAc transferase (OGT) and its removal is catalysed by O-GlcNAcase (OGA) [1]. OGT and OGA are both encoded by single genes in most animals, building an O-GlcNAc proteome of over 1000 proteins involved in metabolic pathways, cell signalling, transcriptional and translational control, epigenetic modification, intracellular trafficking and as components of nuclear pores [2]. Disruption of O-GlcNAc signalling is associated with pathogenesis, including diabetes, cancer and nervous system disorders [3–5]. Utilizing several genetic approaches, it has been demonstrated that O-GlcNAcylation is essential for embryonic development in mouse, *Drosophila* and zebrafish [6–8], although the underpinning mechanistic biology remains to be explored. *Drosophila* contains single copies of both *OGT* [also known as *super sex combs* (*sxc*)] and *OGA*. *sxc* mutants do not survive beyond the pupal stages, whereas a mutant lacking OGA protein is viable [7,9]. Although *sxc* is known to be essential for *Drosophila* development, changes in the O-GlcNAc proteome as a function of embryonic development have not yet been characterized.

Although the O-GlcNAc PTM was discovered three decades ago, its detection and quantification remains challenging. One of the pan-O-GlcNAc antibodies, CTD110.6, is a mouse monoclonal IgM raised against an O-GlcNAcylated peptide derived from the C-terminal domain of RNA polymerase II [10]. However, CTD110.6 has been shown to cross-react with N-glycosylated proteins [11,12], GlcNAc in the endoplasmic reticulum [13] and extracellular O-GlcNAc proteins [14,15]. The other commercially available O-GlcNAc antibody, RL2, shows selectivity towards glycosylated nucleoporins [16]. Other methods to detect nucleocytoplasmic O-GlcNAc include labelling with (UDP)-[^{3}H]galactose using galactosyltransferase [17], lectins [18], chemical labelling with modified galactose and recombinant galactosyltransferase [19–21] and ^{35}S-labelling using GlcNAc-specific sulfotransferases [22]. These techniques recognize only a subset of the O-GlcNAc proteome and/or are not compatible with more complex biological samples.

The bacterium *Clostridium perfringens* possesses an apparent OGA orthologue (*Cp*OGA) of unknown function, but with a remarkable ability to strip entire O-GlcNAc proteomes of O-GlcNAc [23,24]. Interestingly, when a critical catalytic residue (Asp^{298}, the catalytic acid protonating the glycosidic bond) is mutated to asparagine (*Cp*OGA^{D298N}), catalytic activity is lost, yet substrate binding is retained [23]. This has been exploited to determine crystal structures of Ser-O-GlcNAc peptides in complex with *Cp*OGA [25]. In the present study, we investigate the hypothesis that *Cp*OGA^{D298N} can also specifically bind to O-GlcNAcylated proteins and thus be used in a far-Western approach (Figure 1a) to reveal changes in protein O-GlcNAcylation as a function of *Drosophila* embryonic development.

MATERIALS AND METHODS

Protein production

pGEX6P1 plasmids containing N-terminally GST-tagged *Cp*OGA^{WT} (31–618) [25] and mutants were transformed into *Escherichia coli* BL21(DE3) pLysS cells. Cells were grown

Abbreviations: ACN, acetonitrile; ConA, concanavalin A; *Cp*OGA, *Clostridium perfringens* OGA orthologue; dHCF, *Drosophila* host cell factor; *Dm*OGT, *Drosophila* OGT; ETD, electron transfer dissociation; GalNAz, N-azidoacetylgalactosamine; GlcNAc, *N*-acetylglucosamine; Fmoc, 9-fluorenylmethyloxycarbonyl; HA, haemagglutinin; HCF, host cell factor; HEK, human embryonic kidney; hOGA, human OGA; LB-Amp, LB medium containing 50 μg/ml ampicillin; OGA, O-GlcNAcase; OGT, O-GlcNAc transferase; Ph, polyhomoeotic; PNGase F, peptide: N-glycosylase F; PTM, post-translational modification; SPR, surface plasmon resonance; *sxc*, *super sex combs*; TAB1, transforming growth factor-beta-activated protein kinase 1-binding protein 1; TFA, trifluoroacetic acid.

[1] These authors contributed equally to this work.

[2] To whom correspondence should be addressed (email dmfvanaalten@dundee.ac.uk).

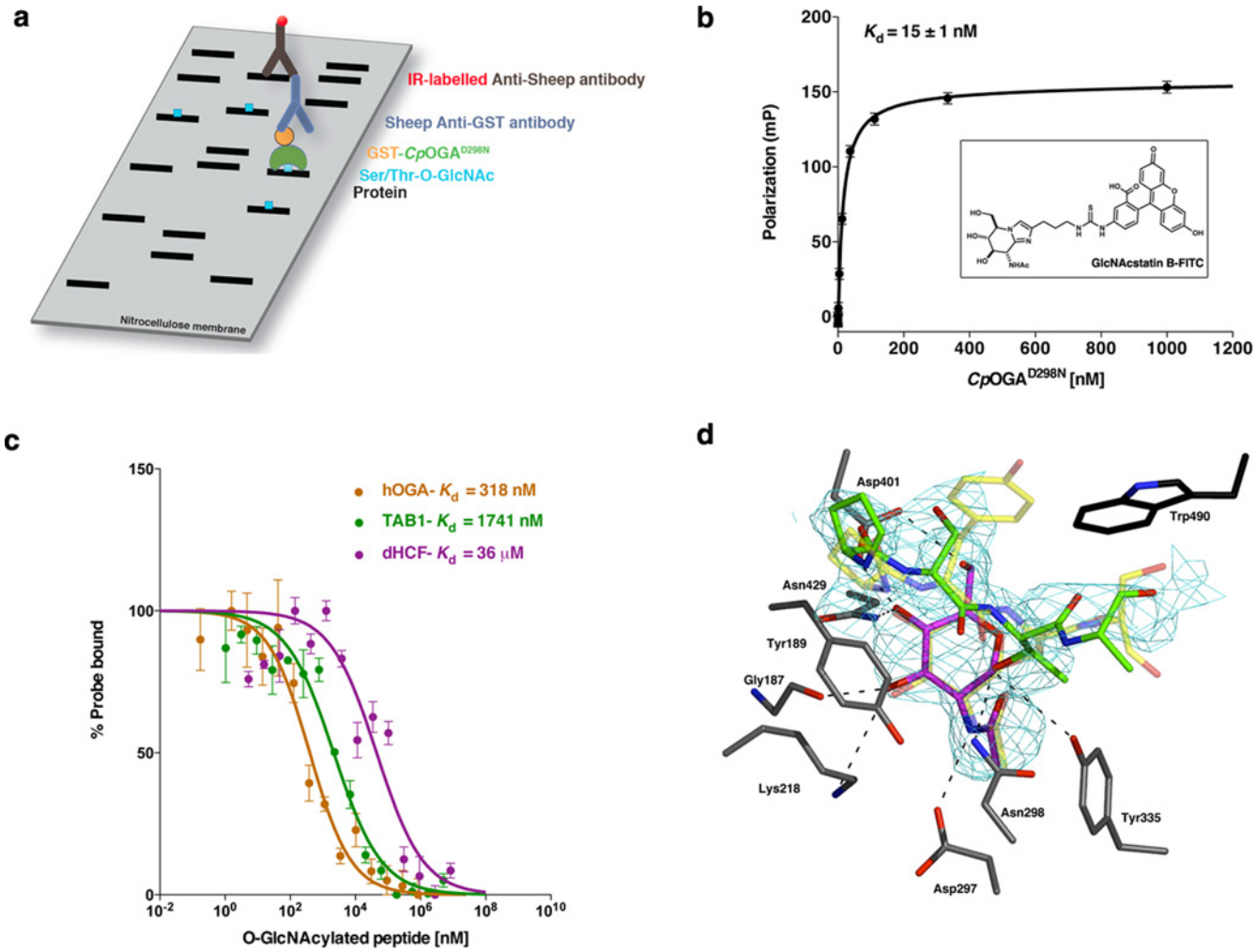

Figure 1 ***Cp*OGA^{D298N} binds to O-GlcNAcylated peptides**

(**a**) Cartoon outlining the far-Western blot approach using N-terminally GST-tagged (●) *Cp*OGA^{D298N} (◡). O-GlcNAc (■) on proteins (▬) transferred on to a nitrocellulose membrane is detected by incubating the membrane with GST–*Cp*OGA^{D298N} followed by anti-GST primary (⅄) and IR-labelled (●) secondary (⅄) antibodies. (**b**) Fluorescence polarization assay showing the binding of fluorescently labelled probe (structure in the inset; see the main text for more details) to *Cp*OGA^{D298N}. Binding was measured by incubating a fixed concentration of labelled probe with various concentrations of *Cp*OGA^{D298N}. Data points were fitted to a saturation binding equation using Prism (GraphPad). Experiments were performed in triplicate and error bars represent S.E.M. (**c**) Dose–response curves from fluorescence polarization assay showing the displacement from *Cp*OGA^{D298N} of a fixed concentration of fluorescent probe by increasing concentrations of O-GlcNAcylated peptides. The highest amount of probe bound to *Cp*OGA^{D298N} in the absence of competing O-GlcNAcylated peptides was arbitrarily set as 100 %. Data points were fitted to an equation for dose-dependent inhibition using Prism (GraphPad). Experiments were performed in triplicate and error bars represent S.E.M. (**d**) Structure of *Cp*OGA^{D298N} in complex with O-GlcNAcylated dHCF peptide (V^{614}PS**gT**MSAN^{621}). The carbons of the active site residues of *Cp*OGA^{D298N} are shown in grey sticks. The unbiased 2.44 Å (1 Å = 0.1 nm) $|F_o| - |F_c|$, ϕ_{calc} electron density map for the O-GlcNAcylated dHCF peptide is shown in cyan, contoured at 2.5σ. The dHCF peptide, PSgTA (as defined by the unbiased electron density), is shown as sticks with green carbons and the O-GlcNAc moiety with magenta carbons. Trp^{490}, which is not conserved in hOGA, is shown as sticks with black carbons. Superimposed for comparison is the structure of *Cp*OGA^{D298N} in complex with the human TAB1 peptide, VPYgSS (PDB code 2YDS [25]). The TAB1 peptide and O-GlcNAc moiety are shown in transparent sticks with yellow carbons. Hydrogen bonds are shown as dashed lines.

overnight at 37 °C in LB medium containing 50 μg/ml ampicillin (LB-Amp) and used at 10 ml/l to inoculate fresh LB-Amp. Cells were grown to an OD_{600} of 0.6–0.8, transferred to 18 °C and induced with 250 μM isopropyl beta-D-1-thiogalactopyranoside (IPTG) and harvested after 16 h by centrifugation for 30 min at 2000 ***g*** (4 °C). Cell pellets were resuspended in 10–20 ml/l of 50 mM Tris/HCl and 250 mM NaCl at pH 7.5 (lysis buffer) supplemented with protease inhibitors (1 mM benzamidine, 0.2 mM phenylmethylsulfonyl fluoride (PMSF) and 5 μM leupeptin), DNAse and lysozyme prior to lysis. Cells were lysed using a continuous flow cell disrupter (Avestin, three passes at 20 kpsi; 1 psi = 6.9 kPa) and the lysate was cleared by centrifugation (30 min, 48 000 ***g***, 4 °C). Supernatants were collected and loaded on to 2 ml of glutathione–Sepharose (GE Healthcare Life Sciences) pre-equilibrated with lysis buffer. The column was washed with 500 ml of lysis buffer. Proteins were eluted with 25 ml of lysis buffer supplemented with 50 mM reduced glutathione and dialysed into 1× tris buffered saline (TBS), pH 7.5. Proteins were further purified by size-exclusion chromatography using a Superdex 200, 26/60 column. transforming growth factor-beta-activated protein kinase 1- binding protein 1 (TAB1) was purified as described previously [26].

Synthesis of peptides and fluorescently labelled GlcNAcstatin

The glycosylated amino acid building blocks 3,4,6-triacetyl-*O*-GlcNAc-Fmoc-Ser-OH and 3,4,6-triacetyl-*O*-GlcNAc-Fmoc-Thr-OH were synthesized in-house according to the published procedures [25]. Microwave-assisted solid-phase peptide

synthesis was performed with a CEM Liberty automated peptide synthesizer on low-load (0.38 mmol/g) Rink amide 4-methylbenzhydrylamine (MBHA) resin 100–200 mesh (Novabiochem) using standard 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry protocols on a 0.05 mmol scale. After the removal of the N-terminal Fmoc group the peptidyl resin was acetylated (acetic anhydride, N,N-Diisopropylethylamine), carbohydrate residues were deprotected (20 % hydrazine, methanol) and peptides were globally deprotected and cleaved from the resin with trifluoroacetic acid (TFA)/triisopropylsilane (TIPS)/water (92.5:2.5:5, by vol.) cocktail for 2 h. The cleavage mixture was filtered off, the resin was washed with TFA twice and the combined filtrate was concentrated to one-tenth of the initial volume in vacuum. The oily residue was triturated with cold (0°C) methyl t-butyl ether and the precipitated peptide was collected by centrifugation. The crude peptides were purified to more than 95 % purity using Waters Peptide Separation Technology 19 mm × 100 mm C_{18} column at a flow rate 20 ml/min, on a preparative HPLC system consisting of Gilson 331/332 pumps, Gilson UV156 detector and Gilson 203 fraction collector controlled by Gilson Trilution LC software. The appropriate fractions were pooled and freeze-dried.

Synthesis of the FITC-labelled fluorescence polarization probe was based on the synthetic procedure for GlcNAcstatin B [27] and details are available on request from D.M.F.v.A.

Protein crystallography

*Cp*OGA^{D298N} was purified and crystallized as described previously [25]. The glycopeptide complex with *Drosophila* host cell factor (dHCF) (VPS**gT**MSAN) peptide was achieved through soaking with 10 mM glycopeptide for 0.5 h prior to cryoprotection with 20 % glycerol in mother liquor. Diffraction data were collected at the Diamond Light Source (Didcot) I03 (Supplementary Table S1). Crystals belonged to space group $P6_1$ and contained one molecule per asymmetric unit, with 73.47 % solvent content. The structure was solved by molecular replacement, using PDB code 2YDS [25] as a search model, followed by iterative model building with COOT [28] and refinement with REFMAC5 [29] using 2 % of reflections as an R_{free} test set. Supplementary Table S1 gives details of the data collection, processing and refinement statistics. The model for the peptide was included when it became fully defined by the difference map upon refining the protein structure and adding hetero atoms. The occupancy for the peptide was set at 0.75.

Surface plasmon resonance

*Cp*OGA^{WT} (31–618) and mutants were purified as described previously [25]. Proteins were chemically biotinylated using the EZ-Link NHS-PEG4-Biotin kit (Thermo) according to the manufacturer's instructions, except that a 1:1 molar ratio of biotinylation reagent to protein was used. Proteins were captured on a neutravidin surface prepared on high-capacity amine sensor chip of a Mass-1 instrument (Sierra Sensors) at densities ~3600–3900 RU (relative units). All experiments were performed at 25 °C. Ligands were injected over captured proteins at flow rate of 30 μl/min in running buffer (50 mM HEPES, 250 mM NaCl, 10 mM EDTA, 5 mM tris (2-carboxyethyl) phosphine (TCEP) and, 0.05 % Tween 20, with each compound injected in duplicate in concentration series adjusted specifically around their affinities. Association was measured for 60 s and dissociation for 120 s. All data were double referenced for blank injections of buffer and biotin-blocked streptavidin surface. Analyser 2 (Sierra Sensors) and Scrubber 2 (BioLogic Software) were used to process and analyse the data.

Fluorescence polarization

Experiments were performed in PerkinElmer black 384-well plates and millipolarization units were measured using a Pherastar FS plate reader (BMG LABTECH) at excitation and emission wavelengths of 485 nm and 520 nm respectively. For determination of the equilibrium dissociation constant (K_d) of *Cp*OGA^{D298N} for the fluorescent probe, 5 nM probe was incubated with a range of concentrations of protein in 25 μl of total reaction volume containing 1× TBS buffer, pH 7.5, and a final concentration of 1 % dimethyl sulfoxide (DMSO). Reactions were allowed to stand at room temperature for 10 min and polarization was measured every 5 min for a period of 2.5 h (equilibrium was reached within 10 min). Readings were corrected for background emissions from reactions containing no *Cp*OGA^{D298N} and K_d was determined by fitting a non-linear regression curve with Prism (GraphPad) to readings obtained at 15 min. To avoid receptor depletion, reaction mixtures for competition binding experiments contained 5 nM fluorescent probe, 20 nM *Cp*OGA^{D298N} (receptor) and a range of concentrations of O-GlcNAcylated peptides under the aforementioned reaction conditions. The highest amount of fluorescent probe bound to *Cp*OGA^{D298N} in the absence of competing O-GlcNAcylated peptides was arbitrarily set as 100 %. EC_{50} values were determined by fitting non-linear regression curves with Prism (GraphPad) and converted into K_d as described elsewhere [30]. All experiments were performed in triplicate.

Drosophila stocks and embryo collections

w1118 wild-type flies were used for all experiments. Embryos were collected on apple juice agar plates at 25 °C for 30 min and aged for an additional 2, 4.5, 10.5 or 14.5 h. To make total embryo lysates, overnight (0–16 h) embryo collections were set up. The embryos thus collected and aged were dechorionated with bleach and snap frozen on dry ice. Other lines used were UAS::OGT^{WT}-HA (haemagglutinin) and tub::GAL4/TM3. To overexpress HA-tagged OGT^{WT}, UAS::OGT^{WT}–HA homozygous virgins were crossed with tub::GAL4/TM3 flies and 0–16 h embryos were collected on apple juice agar plates, dechorionated with bleach and used for immunoprecipitation experiments.

Lysates and immunoprecipitation

Drosophila embryo, human embryonic kidney (HEK)293 and S2 cell lysates were prepared identically for Western blot, far-Western blot and immunoprecipitation experiments. The frozen embryos were homogenized in lysis buffer (50 mM Tris/HCl, pH 8.0, 150 mM NaCl, 1 % Triton-X-100, 1 μM GlcNAcstatin C, 5 mM sodium fluoride, 2 mM sodium orthovanadate, 1 mM benzamidine, 0.2 mM PMSF, 5 μM leupeptin and 1 mM DTT). Confluent HEK293 or S2 cultures were washed with 1× PBS, pH 7.5, prior to lysis with lysis buffer. Lysates were then centrifuged at 16 000 ***g*** for 10 min, supernatants were collected and protein concentrations were estimated using the 660 nm protein assay (Thermo Scientific).

For immunoprecipitation experiments, lysates of OGT^{WT}–HA-overexpressing embryos were prepared with lysis buffer and 2.5 mg protein was incubated with Protein G-Dynabeads (Invitrogen) bound to 5 μg of anti-HA antibody (clone 12CA5). For immunoprecipitating O-GlcNAcylated proteins, 5 mg of S2

cell lysate was incubated with Protein G-Dynabeads bound to 10 μg of anti-O-GlcNAc antibody (RL2). After washing, the immunoprecipitated proteins were eluted by boiling with 1× SDS loading buffer. The proteins were then separated on 4–12 % NuPAGE gels, stained with Coomassie Blue, and gel pieces at the same molecular mass range in the control IgG and RL2 immunoprecipitates were excised and processed for mass spectrometry (MS). Where indicated, the immunoprecipitate, after washing away the lysate, was incubated with 10 μg of *Cp*OGA for 60 min at 30 °C.

Mass spectrometry (MS)

The excised gel pieces were subjected to enzymatic digestion as previously reported with minor modifications [31]. Briefly, excised bands were rinsed thrice with AmBic buffer [50 mM ammonium bicarbonate in 50 % methanol (HPLC grade, Merck)] followed by a reduction step with 10 mM DTT (Sigma–Aldrich). Subsequently, the gel pieces were rinsed twice with AmBic buffer and dried in a SpeedVac before alkylation with 55 mM iodoacetamide (Sigma–Aldrich) in 50 mM ammonium bicarbonate. Thereafter, the gel pieces were rinsed with AmBic buffer, dehydrated with acetonitrile (ACN; HPLC grade, Merck) and dried in a SpeedVac. The dry gel pieces were treated with trypsin (Promega; 20 ng/μl in 20 mM ammonium bicarbonate), incubating them at 37 °C for 16 h. Peptides were extracted thrice by incubating for 20 min with 40 μl of 60 % ACN in 0.5 % formic acid. The resulting peptide extracts were pooled, concentrated in a SpeedVac and stored at − 20 °C until MS analysis.

Identification of the O-GlcNAcylated proteins and O-GlcNAc sites was performed by ESI–IT (ion trap)–ETD (electron transfer dissociation) MS coupled to a nano-LC system (Ultimate 3000 RSLC, Dionex). Dried peptides were resuspended in 30 μl of 0.5 % formic acid and 10 μl of this was injected for mass spectrometric analysis. Tryptic peptides were concentrated on a trap column (2 cm×100 μm, Dionex) at 10 μl/min and separated on a 15 cm×75 μm Pepmap C_{18} reversed-phase column (Thermo Fisher Scientific). Peptides were eluted by a linear 60 min gradient of 95 % A/5 % B to 10 % A/90 % B (A: water, 0.1 % formic acid; B; 80 % ACN, 0.08 % formic acid) at 300 nl/min into a LTQ Velos ETD (Thermo Fisher Scientific). MS spectra were acquired in positive mode; first MS full scans were acquired followed by MS/MS in ETD mode. Up to ten most intense precursors were selected for ETD fragmentation with an activation time of 300 ms and non-dynamic exclusion. Proteome Discoverer version 1.4.0.288 software (Thermo Scientific) was used to process raw LC–MS/MS data, applying the Mascot (version 2.4, Matrix Science) search engine algorithm against SwissProt database with specified taxonomy (*Drosophila melanogaster*, number of sequences 5545) with the following Mascot parameters: 2+, 3+, 4+ and 5+ ions; precursor mass tolerance 100 ppm; fragment tolerance 0.6 Da and up to two missed cleavages. The variable modifications included were: oxidation (M; 15.99 Da), dioxidation (M; +31.98 Da) and HexNAc (ST; +203.0794 Da). All MS/MS data and database results were manually inspected in detail to verify the automatic assignment of fragment ions using the above software.

Western and far-Western blot analysis

Ten micrograms of the crude lysates was subjected to SDS/PAGE and transferred on to a nitrocellulose membrane before immunoblotting with 1 μg/ml RL2 (1:1000 dilution of 1 mg/ml, Abcam), 0.6 μg/ml CTD110.6 (1:500 dilution of 0.3 mg/ml, Cell Signaling Technologies), mouse anti-α-tubulin (1:10 000 dilution, Developmental Studies Hybridoma Bank), rabbit anti-tubulin (1:2500 dilution, Cell Signalling) and/or rabbit anti-actin (1:5000 dilution, Sigma). For the far-Western blot experiments, the blots were incubated with 10 μg/ml GST-*Cp*OGA^{D298N} for 30 min at room temperature followed by incubation with sheep anti-GST (1:5000 dilution) antibody. For each of the experiments, the respective IR dye-conjugated secondary antibodies or IR dye-conjugated streptavidin (Li-Cor or Life Technologies, 1:10 000 dilution) were used. Signal was detected with an Odyssey® Li-Cor IR imaging system. Labelling with UDP–N-azidoacetylgalactosamine (GalNAz) and biotinylation of the labelled proteins was performed using Click-IT O-GlcNAc enzymatic labelling system and the Click-IT Biotin protein analysis system (Life Technologies). To ascertain the specificity of the O-GlcNAc signal, embryo lysates (made without GlcNAcstatin C) were pre-treated with 10 μg of *Cp*OGA for 60 min at 30 °C before being processed. Control samples not treated with *Cp*OGA were also incubated at 30 °C for 60 min. To establish that *Cp*OGA^{D298N} does not cross react with N-glycosylated proteins, lysates were treated with peptide: N-glycosidase F (PNGase F) (NEB) as per the manufacturer's instructions. Removal of N-glycans was monitored by blotting with a specific lectin, concanavalin A (ConA), following a previously described protocol [32].

RESULTS AND DISCUSSION

Since the *Drosophila* O-GlcNAc proteome has not previously been studied, we initially immunoprecipitated O-GlcNAc proteins from *Drosophila* S2 cell lysates using the anti-O-GlcNAc RL2 antibody and analysed these with ETD MS/MS. Several peptides from dHCF were found to be O-GlcNAcylated with one site conserved with the human HCF1 orthologue (Supplementary Figure S1, V^{614}PS**gT**MSAN^{621}). Human HCF1 is one of the best-characterized OGT substrates that is heavily glycosylated and proteolytically cleaved by OGT, known to play a central role in epigenetic regulation of gene expression [33,34]. We next investigated whether the dHCF Thr-O-GlcNAc peptide V^{614}PS**gT**MSAN^{621}, containing an O-GlcNAcylated threonine residue, was able to bind *Cp*OGA^{D298N}. To facilitate this, we developed a fluorescence polarization assay based on a newly synthesized FITC-labelled OGA inhibitor that binds with 15 nM affinity to *Cp*OGA^{D298N} (Figure 1b). In competition experiments, the dHCF O-GlcNAc peptide was shown to bind *Cp*OGA^{D298N} with a K_d of 36 μM (Figure 1c). Encouraged by this, we also determined binding of Ser-O-GlcNAc peptides. Strikingly, V^{392}PY**gS**SA^{397} (derived from the O-GlcNAc site on the innate immunity signalling protein TAB1 [35], K_d = 1741 nM) and V^{402}AH**gS**GAK^{408} (derived from the O-GlcNAc site on human OGA (hOGA) [36], K_d = 318 nM) bind with nanomolar affinity. The unglycosylated peptide controls showed no detectable binding. These data are consistent with K_d values obtained from a surface plasmon resonance (SPR)-binding assay (Supplementary Figure S2). To explore the molecular basis of these interactions and possible differences in recognition of Ser- compared with Thr-O-GlcNAc proteins, we determined the crystal structure of *Cp*OGA^{D298N} in complex with the dHCF-derived peptide (Figure 1d). The structure reveals tight contacts between the GlcNAc moiety and the protein via hydrogen bonds to residues that are all conserved in hOGA and explains why binding is lost with the D401A mutant of the enzyme (*Cp*OGA^{D401A}), which abrogates the interactions with the GlcNAc O4 and O6 hydroxy groups [23]. Furthermore, Tyr^{189}, conserved in hOGA,

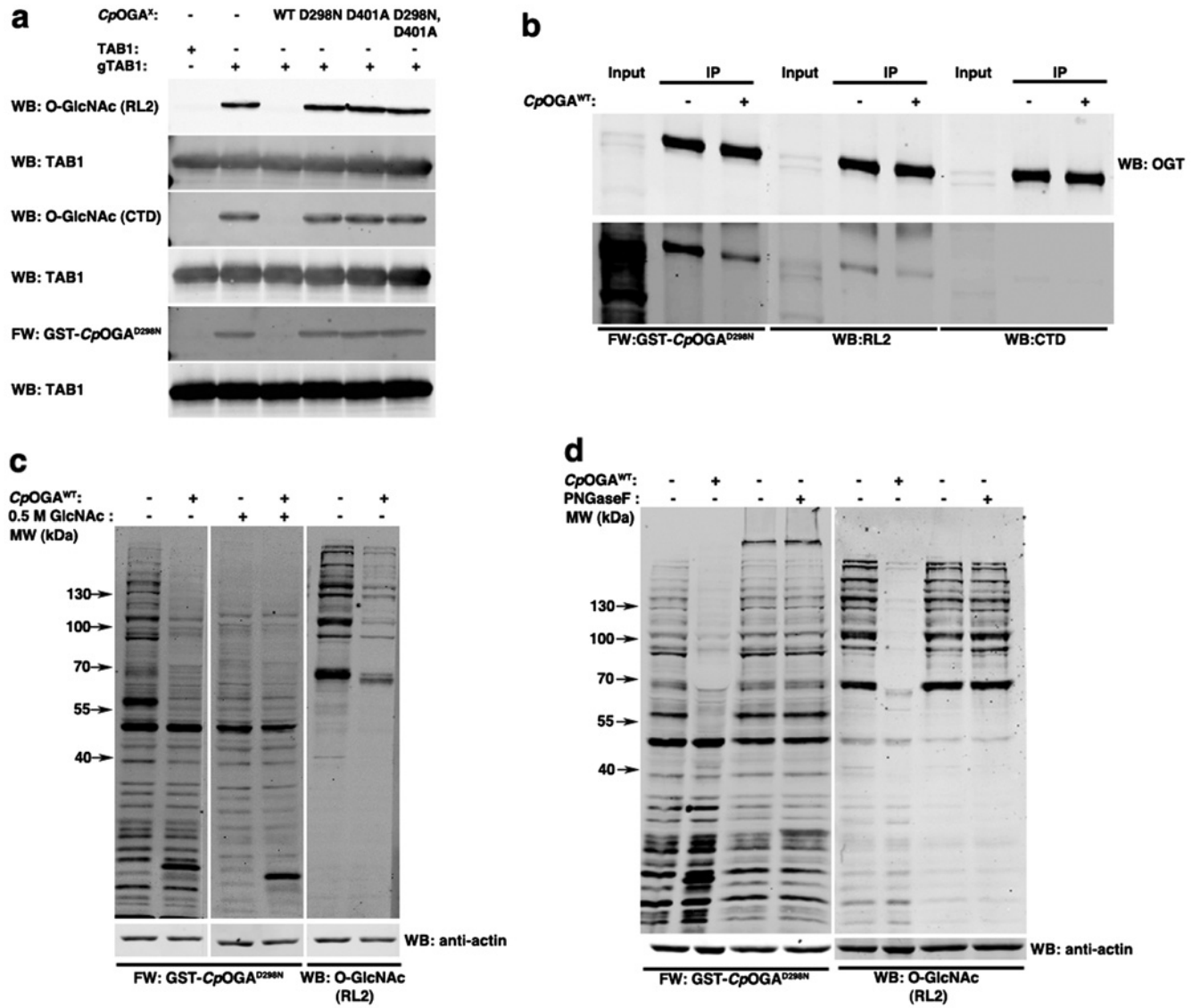

Figure 2 ***Cp*OGA^{D298N} can be utilized as a far-Western blotting tool**

(**a**) Immunoblots with anti-O-GlcNAc antibodies (RL2 and CTD110.6) or far-Western blotting with GST–*Cp*OGA^{D298N} was performed on either naked or *in vitro* O-GlcNAcylated TAB1 (labels adjacent to the blots). The specificity of the signal detected was confirmed by deglycosylating O-GlcNAcylated TAB1 by pre-treatment with *Cp*OGA^{WT}. Pre-treatment as indicated above the blots was performed to probe the activity of all of the *Cp*OGA constructs. (**b**) Embryonically overexpressed *Dm*OGT–HA was immunoprecipitated with anti-HA antibody and immunoblotted with anti-O-GlcNAc antibodies (RL2 and CTD110.6) or subjected to far-Western blotting with GST–*Cp*OGA^{D298N}. Treating the immunoprecipitate with *Cp*OGA^{WT} was used as a control. (**c**) HEK293 lysates without or with *Cp*OGA^{WT} pre-treatment were subjected to GST-*Cp*OGA^{D298N} far-Western blotting. In addition, the GST–*Cp*OGA^{D298N} far-Western blotting was performed in the presence of 0.5 M GlcNAc competition. The extreme right panel is an immunoblot of the same samples with anti-O-GlcNAc antibody RL2. (**d**) HEK293 lysates without or with PNGase F/*Cp*OGA^{WT} pre-treatment were subjected to GST–*Cp*OGA^{D298N} far-Western blotting. The right panel is an immunoblot of the same samples with anti-O-GlcNAc antibody RL2.

makes stacking interactions with the backbone of the O-GlcNAc peptide, explaining how this class of enzymes recognize their substrates in a sequence-independent, yet protein-dependent, manner. Interestingly, compared with previously determined structures of *Cp*OGA^{D298N} in complex with the TAB1-hOGA-derived Ser-O-GlcNAc peptides [25], the apolar extra γ-methyl group of the threonine residue appears to be unfavourably positioned near the Tyr^{335} hydroxy group, perhaps explaining the reduction in affinity compared with the Ser-O-GlcNAc peptides. Nevertheless, *Cp*OGA^{D298N} appears to be able to bind both Ser- and Thr-O-GlcNAc peptides, enabling its potential use as a general detector of O-GlcNAc proteins.

We next investigated the suitability of *Cp*OGA^{D298N} as a far-Western blotting probe as outlined in Figure 1a. Human TAB1 was O-GlcNAcylated *in vitro* and probed with GST-fused *Cp*OGA^{D298N} in a Western blot. The fusion protein was detected using sheep anti-GST antibody in combination with a Li-Cor (an infra red (IR) Western blot detection system) -compatible anti-sheep secondary antibody. Specific signal was detected only with GST–*Cp*OGA^{D298N} (Figure 2a). However, on performing the far-Western blotting with GST–*Cp*OGA^{WT}, (which hydrolyses the sugar, resulting in loss of binding), GST–*Cp*OGA^{D401A} or GST–*Cp*OGA$^{D298N, D401A}$ (both deficient in O-GlcNAc protein binding [23]; Supplementary Figure S3a), as expected, the specific signal

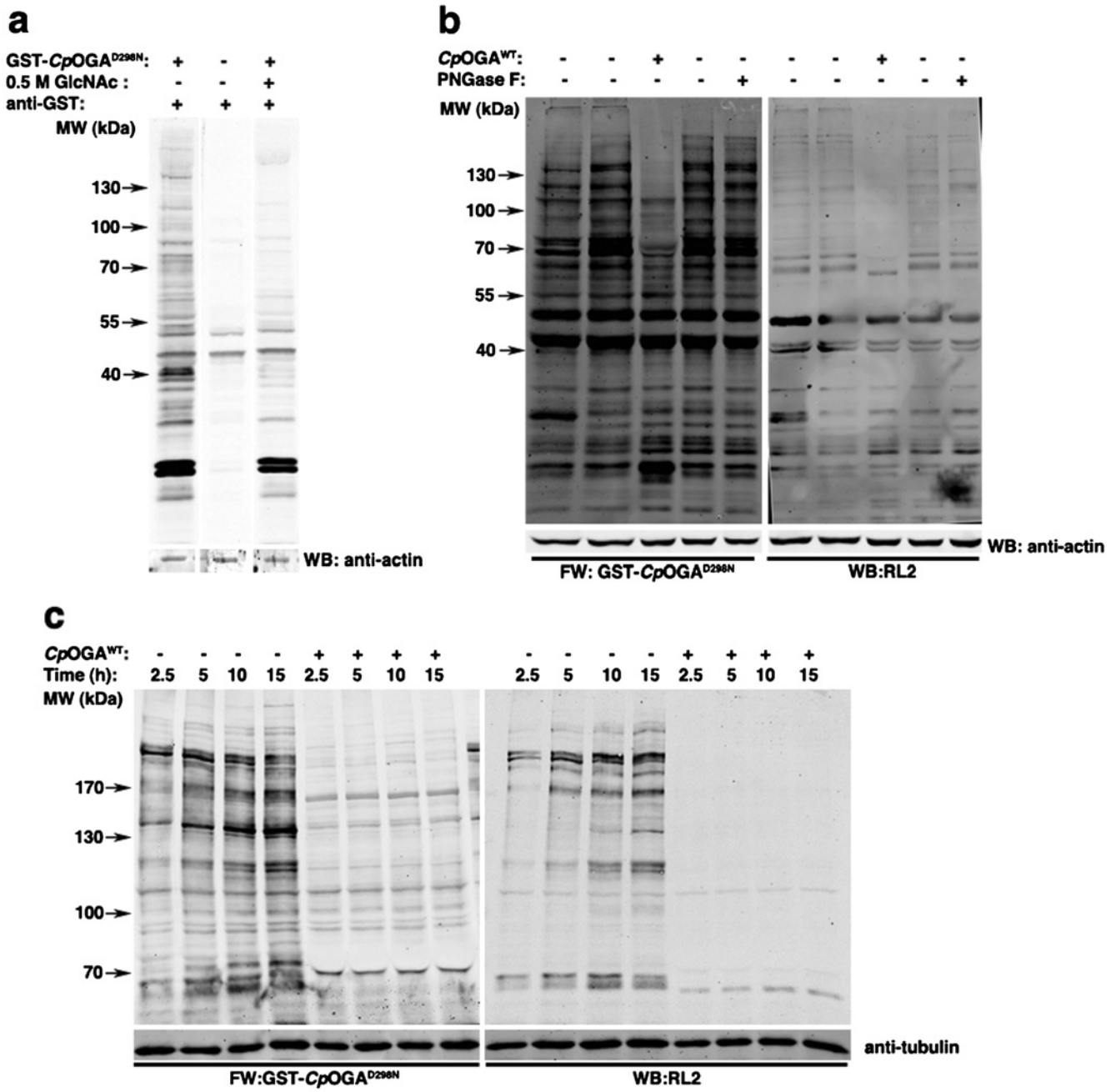

Figure 3 The O-GlcNAc proteome across *Drosophila* embryogenesis is dynamic

(**a**) *Drosophila* embryonic lysates from 0–16 h collections were subjected to GST or GST–*Cp*OGA^{D298N} far-Western blotting. In addition, the GST–*Cp*OGA^{D298N} far-Western blotting was performed in the presence of 0.5 M GlcNAc. (**b**) *Drosophila* embryonic lysates (0–16 h) without or with PNGase F/*Cp*OGA^{WT} pre-treatment were subjected to GST–*Cp*OGA^{D298N} far-Western blotting. The right panel is an immunoblot of the same samples with anti-O-GlcNAc antibody RL2. The lysates were separated by SDS/PAGE (10 % gel). (**c**) *Drosophila* embryonic lysates at 2.5, 5, 10 and 15 h without or with *Cp*OGA^{WT} pre-treatment were subjected to GST–*Cp*OGA^{D298N} far-Western blotting. The right panel is an immunoblot of the same samples with anti-O-GlcNAc antibody RL2. The lysates were separated by SDS/PAGE (6 % gel).

was not detected, establishing the suitability of *Cp*OGA^{D298N} for this purpose. When glycosylated TAB1 is pre-treated with *Cp*OGA^{WT}, thus deglycosylating it, the specific O-GlcNAc-dependent signal is no longer detectable on a far-Western blot with GST–*Cp*OGA^{D298N} or with the CTD110.6/RL2 anti-O-GlcNAc antibodies (Figure 2a). However, pre-treatment with *Cp*OGA^{D298N}, *Cp*OGA^{D401A} or *Cp*OGA^{D298N,D401A} did not lead to any change in the specific signal detected by GST–*Cp*OGA^{D298N} or with the CTD110.6/RL2 anti-O-GlcNAc antibodies, further confirming that these mutants are indeed inactive (Figure 2a).

Mammalian OGT is known to self-glycosylate [37] and hence we investigated auto-glycosylation of *Drosophila* OGT (*Dm*OGT) using *Cp*OGA^{D298N}. HA-tagged *Dm*OGT^{WT} was embryonically overexpressed and immunoprecipitated using anti-HA antibody. O-GlcNAcylation of immunoprecipitated OGT was readily detected using *Cp*OGA^{D298N} far-Western blotting (Figure 2b). The

specificity of this reactivity was also determined by treating the immunoprecipitate with *Cp*OGA^{WT}, which considerably reduced the signal (Figure 2b), thus demonstrating the utility of this far-Western blot approach in expanding the repertoire of detectable O-GlcNAc sites.

Next, we explored this approach using a more complex mammalian sample, HEK293 cell lysates. Specific O-GlcNAcylated proteins were detected when far-Western blotting was performed using GST–*Cp*OGA^{D298N}, although GST-dependent reactivity is visible in the <50 kDa range (Figure 2c). The specific signal of the >50 kDa bands diminished significantly when the lysates were pre-treated with *Cp*OGA^{WT} in order to remove the O-GlcNAc from proteins or on competing with 0.5 M GlcNAc (Figure 2c). Additional controls including treatment of blots with GST, anti-GST antibody or the Li-Cor anti-sheep secondary antibody alone further established the specificity of O-GlcNAc detection with GST–*Cp*OGA^{D298N} (Supplementary Figure S3b). The anti-O-GlcNAc antibody RL2 identified fewer bands (∼55 kDa, 100 kDa and > 130 kDa) as compared with GST–*Cp*OGA^{D298N} (Figure 2c). Since the secretory N-glycosylation pathway also utilizes an *N*-GlcNAc(β1,4)GlcNAc core, cross-reactivity of *Cp*OGA^{D298N} to N-glycosylated proteins was tested. Lysates were treated with PNGase F or *Cp*OGA^{WT} to remove N-glycosylation or O-GlcNAcylation on proteins prior to the far-Western/Western blot analyses. Treatment of lysates with PNGase F distinctly reduced reactivity to ConA, a lectin routinely used to detect N-glycans (Supplementary Figure S3c), but not the GST–*Cp*OGA^{D298N}-dependent signal (Figure 2d). As expected, pre-treatment with *Cp*OGA^{WT} significantly reduced detection by *Cp*OGA^{D298N}, establishing its specificity for nucleocytoplasmic O-GlcNAcylation (Figure 2d). As with the *in vitro* experiments, specific signal is not detected when GST–*Cp*OGA^{WT}, GST–*Cp*OGA^{D401A} or GST–*Cp*OGA^{D298N,D401A} are used as far-Western blotting probes on HEK293 lysates (Supplementary Figure S3d).

Finally, we used the far-Western blotting *Cp*OGA^{D298N} probe to investigate changes in the O-GlcNAc proteome in the context of *Drosophila* embryonic development. Specific O-GlcNAcylated proteins, spanning a large molecular mass range, were identified in 0–16 h *Drosophila* embryo lysates using this approach (Figure 3a). Specificity of the detection was established by either competing the *Cp*OGA^{D298N} binding with 0.5 M GlcNAc or the loss of specific signal on pre-treatment of the lysate with *Cp*OGA^{WT} (Figures 3a and 3b). The reactivity of *Cp*OGA^{D298N} was not altered when the embryo lysates were pre-treated with PNGase F, which did significantly reduce ConA binding (Supplementary Figure S3e), confirming that *Cp*OGA^{D298N} specificity is towards O-GlcNAc proteins and not N-glycosylated proteins in *Drosophila* embryo lysates. The utility of *Cp*OGA^{D298N} as a far-Western blotting probe was also compared with other O-GlcNAcylation detection methods. There is no specific signal detected at all with the CTD 110.6 antibody in *Drosophila* embryo lysates (Supplementary Figure S4a). Similarly, UDP–GalNAz labelling of embryo lysates resulted in only marginally improved specific signals (Supplementary Figure S4b).

To investigate changes in O-GlcNAcylation across *Drosophila* embryogenesis, embryos (growing at 25 °C) at the cellular blastoderm (2.5 h), germ band extension (5 h), germ band retraction (10 h) and dorsal closure (15 h) stages were collected and the far-Western blot experiment was performed. The number of O-GlcNAc proteins was the highest at the dorsal closure stage (Figure 3c). Although there was progressive accumulation of O-GlcNAc on several proteins with developmental time, many proteins are O-GlcNAcylated at the germ band retraction stage or later (Figure 3c). Whereas the overall O-GlcNAc profile detected using RL2 is comparable, numerous additional proteins are detected by the far-Western blot approach, (for instance, ∼70 kDa, 100 kDa, between 120 and 130 kDa and >170 kDa; Figure 3c). These data reveal that, like protein phosphorylation, different developmental stages are associated with different O-GlcNAc proteomes offering opportunities for future identification of the proteins involved and probing the mechanistic developmental biology underlying these changes from larger-scale tissue samples.

Conclusions

The approach in the present study was to investigate changes in the O-GlcNAc proteome during *Drosophila* embryonic development, harnessing the O-GlcNAc-binding properties of *Cp*OGA^{D298N}. During the course of developing this tool, we have established that *Cp*OGA^{D298N} has higher affinity for Ser-O-GlcNAc than for Thr-O-GlcNAc peptides. A crystal structure of a Thr-O-GlcNAc peptide, in complex with *Cp*OGA^{D298N}, has revealed that the position of the additional threonine γ-methyl is unfavourable, possibly leading to the lower affinity of *Cp*OGA^{D298N} for Thr-O-GlcNAc peptides. Whether this difference in affinities of *Cp*OGA^{D298N} towards Ser- compared with Thr-O-GlcNAc would lead to differential catalytic rates of the active enzyme or translates to physiological consequence(s) remains to be explored.

We have shown in the present study that the O-GlcNAc proteome during *Drosophila* embryonic development is dynamic. Changes in the O-GlcNAc proteome during *Xenopus* embryogenesis have been investigated using anti-O-GlcNAc RL2 antibody [38]. During *Xenopus* embryogenesis, the number of different that are O-GlcNAcylated proteins seems to be constant, with the most obvious changes being the extent of their O-GlcNAc levels, possibly because of the restricted specificity of the RL2 antibody. However, during *Drosophila* embryonic development, in addition to quantitative changes in O-GlcNAc levels on particular proteins, the number of proteins that are O-GlcNAcylated is also dynamic. A recent report identified polyhomoeotic (Ph), a polycomb group protein, as the major substrate for OGT during *Drosophila* development [39]. However, the present study demonstrates that there are multiple OGT substrates during *Drosophila* embryogenesis. Given that *sxc* mutants lacking maternal OGT are late embryonic lethals, it will be feasible to identify the O-GlcNAcylated proteins during *Drosophila* embryogenesis and investigate the significance of their O-GlcNAc status in this model. Reduced O-GlcNAcylation of proteins (other than Ph) could possibly lead to non-lethal subtler phenotypes that could be investigated during early embryogenesis. The impetus of this approach would be to understand the role of O-GlcNAcylation on sites conserved in mammalian orthologues using the strength of the *Drosophila* genetic model.

AUTHOR CONTRIBUTION

Daniel Mariappa and Daan van Aalten conceived the study. Daniel Mariappa, Nithya Selvan, Vladimir Borodkin and Daan van Aalten were involved in experimental design. Daniel Mariappa performed the cellular and *Drosophila* assays. Nithya Selvan performed protein expression, TAB1 *in vitro* assays, structural biology and fluorescence polarization (FP). Vladimir Borodkin performed peptide and chemical synthesis. Jana Alonso performed MS studies. Andrew Ferenbach performed molecular biology. Claire Shepherd and Iva Navratilova performed SPR. Daniel Mariappa, Nithya Selvan and Daan van Aalten interpreted the data and wrote the manuscript.

ACKNOWLEDGEMENTS

We thank the Diamond synchrotron for beam time on beamline I03. The refined crystal structure has been deposited in the PDB (code 4ZXL).

FUNDING

This work was supported by the Wellcome Trust Senior Research Fellowship [grant number WT087590MA] to D.M.F.v.A.

REFERENCES

1 Hart, G., Slawson, C., Ramirez-Correa, G. and Lagerlof, O. (2011) Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu. Rev. Biochem. **80**, 825–883 CrossRef PubMed
2 Hart, G.W. (2014) Three decades of research on O-GlcNAcylation - a major nutrient sensor that regulates signaling, transcription and cellular metabolism. Front. Endocrinol. **5**, 183 CrossRef PubMed
3 Vaidyanathan, K. and Wells, L. (2014) Multiple tissue-specific roles for the O-GlcNAc post-translational modification in the induction of and complications arising from type II diabetes. J. Biol. Chem. **289**, 34466–34471 CrossRef PubMed
4 Ma, Z. and Vosseller, K. (2014) Cancer metabolism and elevated O-GlcNAc in oncogenic signaling. J. Biol. Chem. **289**, 34457–34465 CrossRef PubMed
5 Zhu, Y., Shan, X., Yuzwa, S.A. and Vocadlo, D.J. (2014) The emerging link between O-GlcNAc and Alzheimer disease. J. Biol. Chem. **289**, 34472–34481 CrossRef PubMed
6 Shafi, R., Iyer, S.P., Ellies, L.G., O'Donnell, N., Marek, K.W., Chui, D., Hart, G.W. and Marth, J.D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. Proc. Natl. Acad. Sci. U.S.A. **97**, 5735–5739 CrossRef PubMed
7 Gambetta, M.C., Oktaba, K. and Muller, J. (2009) Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. Science **325**, 93–96 CrossRef PubMed
8 Webster, D.M., Teo, C.F., Sun, Y., Wloga, D., Gay, S., Klonowski, K.D., Wells, L. and Dougan, S.T. (2009) O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development. BMC Dev. Biol. **9**, 28 CrossRef PubMed
9 Radermacher, P.T., Myachina, F., Bosshardt, F., Pandey, R., Mariappa, D., Müller, H.A. and Lehner, C.F. (2014) O-GlcNAc reports ambient temperature and confers heat resistance on ectotherm development. Proc. Natl. Acad. Sci. U.S.A. **111**, 5592–5597 CrossRef PubMed
10 Comer, F.I. and Hart, G.W. (2001) Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II. Biochemistry **40**, 7845–7852 CrossRef PubMed
11 Isono, T. (2011) O-GlcNAc-specific antibody CTD110.6 cross-reacts with N-GlcNAc2-modified proteins induced under glucose deprivation. PLoS One **6**, e18959 CrossRef PubMed
12 Reeves, R.A., Lee, A., Henry, R. and Zachara, N.E. (2014) Characterization of the specificity of O-GlcNAc reactive antibodies under conditions of starvation and stress. Anal. Biochem. **457**, 8–18 CrossRef PubMed
13 Ogawa, M., Sakakibara, Y. and Kamemura, K. (2013) Requirement of decreased O-GlcNAc glycosylation of Mef2D for its recruitment to the myogenin promoter. Biochem. Biophys. Res. Commun. **433**, 558–562 CrossRef PubMed
14 Tashima, Y. and Stanley, P. (2014) Antibodies that detect O-linked beta-D-N-acetylglucosamine on the extracellular domain of cell surface glycoproteins. J. Biol. Chem. **289**, 11132–11142 CrossRef PubMed
15 Ogawa, M., Sawaguchi, S., Kawai, T., Nadano, D., Matsuda, T., Yagi, H., Kato, K., Furukawa, K. and Okajima, T. (2015) Impaired O-linked N-acetylglucosaminylation in the endoplasmic reticulum by mutated epidermal growth factor (EGF) domain-specific O-linked N-acetylglucosamine transferase found in adams-oliver syndrome. J. Biol. Chem. **290**, 2137–2149 CrossRef PubMed
16 Snow, C.M., Senior, A. and Gerace, L. (1987) Monoclonal antibodies identify a group of nuclear pore complex glycoproteins. J. Cell Biol. **104**, 1143–1156 CrossRef PubMed
17 Torres, C.R. and Hart, G.W. (1984) Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. J. Biol. Chem. **259**, 3308–3317 PubMed
18 Roquemore, E.P., Chou, T.Y. and Hart, G.W. (1994) Detection of O-linked N-acetylglucosamine (O-GlcNAc) on cytoplasmic and nuclear proteins. Methods Enzymol. **230**, 443–460 CrossRef PubMed
19 Khidekel, N., Arndt, S., Lamarre-Vincent, N., Lippert, A., Poulin-Kerstien, K.G., Ramakrishnan, B., Qasba, P.K. and Hsieh-Wilson, L.C. (2003) A chemoenzymatic approach toward the rapid and sensitive detection of O-GlcNAc posttranslational modifications. J. Am. Chem. Soc. **125**, 16162–16163 CrossRef PubMed
20 Clark, P.M., Dweck, J.F., Mason, D.E., Hart, C.R., Buck, S.B., Peters, E.C., Agnew, B.J. and Hsieh-Wilson, L.C. (2008) Direct in-gel fluorescence detection and cellular imaging of O-GlcNAc-modified proteins. J. Am. Chem. Soc. **130**, 11576–11577 CrossRef PubMed
21 Ramakrishnan, B. and Qasba, P.K. (2002) Structure-based design of beta 1,4-galactosyltransferase I (beta 4Gal-T1) with equally efficient N-acetylgalactosaminyltransferase activity: point mutation broadens beta 4Gal-T1 donor specificity. J. Biol. Chem. **277**, 20833–20839 CrossRef PubMed
22 Wu, Z.L., Robey, M.T., Tatge, T., Lin, C., Leymarie, N., Zou, Y. and Zaia, J. (2014) Detecting O-GlcNAc using in vitro sulfation. Glycobiology **24**, 740–747 CrossRef PubMed
23 Rao, F.V., Dorfmueller, H.C., Villa, F., Allwood, M., Eggleston, I.M. and van Aalten, D.M. (2006) Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis. EMBO J. **25**, 1569–1578 CrossRef PubMed
24 Dorfmueller, H.C., Borodkin, V.S., Schimpl, M., Shepherd, S.M., Shpiro, N.A. and van Aalten, D.M. (2006) GlcNAcstatin: a picomolar, selective O-GlcNAcase inhibitor that modulates intracellular O-glcNAcylation levels. J. Am. Chem. Soc. **128**, 16484–16485 CrossRef PubMed
25 Schimpl, M., Borodkin, V.S., Gray, L.J. and van Aalten, D.M. (2012) Synergy of peptide and sugar in O-GlcNAcase substrate recognition. Chem. Biol. **19**, 173–178 CrossRef PubMed
26 Conner, S.H., Kular, G., Peggie, M., Shepherd, S., Schüttelkopf, A.W., Cohen, P. and Van Aalten, D.M. (2006) TAK1-binding protein 1 is a pseudophosphatase. Biochem. J. **399**, 427–434 CrossRef PubMed
27 Dorfmueller, H.C., Borodkin, V.S., Schimpl, M. and van Aalten, D.M. (2009) GlcNAcstatins are nanomolar inhibitors of human O-GlcNAcase inducing cellular hyper-O-GlcNAcylation. Biochem. J. **420**, 221–227 CrossRef PubMed
28 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta. Crystallogr. D Biol. Crystallogr. **60**, 2126–2132 CrossRef PubMed
29 Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta. Crystallogr. D Biol. Crystallogr. **53**, 240–255 CrossRef PubMed
30 Nikolovska-Coleska, Z., Wang, R., Fang, X., Pan, H., Tomita, Y., Li, P., Roller, P.P., Krajewski, K., Saito, N.G., Stuckey, J.A. and Wang, S. (2004) Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. Anal. Biochem. **332**, 261–273 CrossRef PubMed
31 Alonso, J. and Santaren, J.F. (2005) Proteomic analysis of the wing imaginal discs of *Drosophila melanogaster*. Proteomics **5**, 474–489 CrossRef PubMed
32 Mariappa, D., Sauert, K., Mariño, K., Turnock, D., Webster, R., van Aalten, D.M., Ferguson, M.A. and Müller, H.A. (2011) Protein O-GlcNAcylation is required for fibroblast growth factor signaling in *Drosophila*. Sci. Signal. **4**, ra89 CrossRef PubMed
33 Zargar, Z. and Tyagi, S. (2012) Role of host cell factor-1 in cell cycle regulation. Transcription **3**, 187–192 CrossRef PubMed
34 Capotosti, F., Guernier, S., Lammers, F., Waridel, P., Cai, Y., Jin, J., Conaway, J.W., Conaway, R.C. and Herr, W. (2011) O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1. Cell **144**, 376–388 CrossRef PubMed
35 Pathak, S., Borodkin, V.S., Albarbarawi, O., Campbell, D.G., Ibrahim, A. and van Aalten, D.M. (2012) O-GlcNAcylation of TAB1 modulates TAK1-mediated cytokine release. EMBO J. **31**, 1394–1798 CrossRef PubMed
36 Khidekel, N., Ficarro, S.B., Clark, P.M., Bryan, M.C., Swaney, D.L., Rexach, J.E., Sun, Y.E., Coon, J.J., Peters, E.C. and Hsieh-Wilson, L.C. (2007) Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics. Nat. Chem. Biol. **3**, 339–348 CrossRef PubMed
37 Whelan, S.A., Lane, M.D. and Hart, G.W. (2008) Regulation of the O-linked beta-N-acetylglucosamine transferase by insulin signaling. J. Biol. Chem. **283**, 21411–21417 CrossRef PubMed
38 Dehennaut, V., Lefebvre, T., Leroy, Y., Vilain, J.P., Michalski, J.C. and Bodart, J.F. (2009) Survey of O-GlcNAc level variations in *Xenopus* laevis from oogenesis to early development. Glycoconj. J. **26**, 301–311 CrossRef PubMed
39 Gambetta, M.C. and Muller, J. (2014) O-GlcNAcylation prevents aggregation of the polycomb group repressor polyhomeotic. Dev. Cell **31**, 629–639 CrossRef PubMed

Received 26 May 2015/30 June 2015; accepted 14 July 2015
Accepted Manuscript online 14 July 2015, doi:10.1042/BJ20150610