Recognition of a glycosylation substrate by the O-GlcNAc transferase TPR repeats

Karim Rafie1,1, Olawale Raimi1,1, Andrew T. Ferenbach1, Vladimir S. Borodkin1, Vaibhav Kapuria2 and Daan M. F. van Aalten1

1Centre for Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dundee, UK
2Center for Integrative Genomics, University of Lausanne 1015, Switzerland

O-linked N-acetylglucosamine (O-GlcNAc) is an essential and dynamic post-translational modification found on hundreds of nucleocytoplasmic proteins in metazoa. Although a single enzyme, O-GlcNAc transferase (OGT), generates the entire cytosolic O-GlcNAc proteome, it is not understood how it recognizes its protein substrates, targeting only a fraction of serines/threonines in the metazoan proteome for glycosylation. We describe a trapped complex of human OGT with the C-terminal domain of TAB1, a key innate immunity-signalling O-GlcNAc protein, revealing extensive interactions with the tetratricopeptide repeats of OGT. Confirmed by mutagenesis, this interaction suggests that glycosylation substrate specificity is achieved by recognition of a degenerate sequon in the active site combined with an extended conformation C-terminal of the O-GlcNAc target site.

1. Introduction

The attachment of a single β-N-acetylglucosamine (O-GlcNAc) sugar onto serine and threonine residues of nucleocytoplasmic proteins is a dynamic and abundant post-translational modification found in higher eukaryotes [1–3]. Remarkably, this modification is regulated by only two antagonistic enzymes: the O-GlcNAc transferase (OGT), which transfers the GlcNAc moiety onto acceptor residues from the donor sugar nucleotide UDP-GlcNAc, and the O-GlcNAc hydrolase (OGA), which removes it. To date more than 1000 O-GlcNAc proteins have been identified by mass spectrometry [4–10]. These proteins cover a wide range of cellular processes such as transcription and translation [11–13], trafficking and localization [14,15], as well as cell cycle progression [16–19]. However, it remains unclear how a single OGT enzyme is able to specifically recognize a limited number of serines/threonines on such a large number of substrates.

OGT is a multi-domain protein with a catalytic core at the C-terminus and 13 tetratricopeptide (TPR) repeats at the N-terminus, making up about half of the enzyme. Early experiments suggested that the TPR domain is involved in substrate recognition and/or protein–protein interactions [20–27]. The structure of the isolated OGT TPR domain revealed topological similarity to other helical repeat proteins and led to speculation that this domain might bind substrates in an extended conformation [20]. The first structural insights into the OGT catalytic domain came from an OGT orthologue in the bacterium Xanthomonas campestris [28,29]. This structure revealed that the sugar donor binding site is made up of the two lobes of the glycosyl transferase B (GT-B) fold, tightly fused to the superhelical TPR domain [28]. The subsequent structure of human OGT [30] revealed a very similar fold with the addition of an intervening domain of unknown function between the two catalytic lobes [30]. The structure suggested an ordered bi-bi mechanism of substrate binding, in which UDP-GlcNAc binds before the acceptor substrate [30]. Initial structural studies exploring Michaelis/substrate complexes with short acceptor peptides have

© 2017 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.
revealed limited substrate interactions with the enzyme [31,32]. More recently, Pathak et al. investigated the common binding modes of acceptor peptides to OGT [33]. Starting from a peptide library, they identified preference for certain acceptor peptide sequences, leading to definition of a degenerate sequon of OGT peptide substrates ([TS][PT][VT][S/T][RLV][ASY]). Crystal structures of complexes of OGT with some of these peptides revealed that OGT binds all the acceptor peptides studied so far in an extended conformation with similar conformation of the residues in the −3 to +2 position around the acceptor serine/threonine. Although the C-termini of these peptides point towards the TPR domain, these structural data do not explain how OGT recognizes larger protein substrates for glycosylation. The short sequence patterns alone are not sufficient to accurately predict the O-GlcNAc proteome, suggesting other mechanisms contribute to substrate recognition.

A clue to how this might work came from the unusual OGT substrate host cell factor 1 (HCF1). HCF1 is a ubiquitously expressed chromatin-associated protein and a major transcriptional co-regulator involved in numerous cellular processes such as cell cycle progression (reviewed in [34]), which has also been shown to be heavily O-GlcNAcylated [35]. HCF1 is initially expressed as an approximately 210 kDa protein that is activated by limited proteolysis (pro tease maturation) within the proteolytic processing domain (PPD), consisting of multiple 20-residue repeats [36,37]. Strik ingly, in 2011 it was discovered that OGT not only glycosylates HCF1 but is also needed for its proteolytic maturation [13,35]. A depletion of OGT leads to an accumulation of full-length HCF1 protein and the PPD is proteolytically cleaved by OGT via an unusual glycosylated glutamate intermediate [24,35,38]. A recent structural study of a short PPD (HCF1PRO) repeat in complex with OGT revealed that part of the substrate bound in extended conformation in the TPR repeats [24]. OGT was shown to form an extensive array of polar interactions with the backbone of the HCF1PRO repeat peptide, as well as specific side chain interactions that were demonstrated to be essential for HCF1PRO binding [24]. However, it is as yet not clear if this binding mode also extends to OGT glycosylation substrates.

A well-characterized OGT glycosylation substrate is the TGFβ-activated kinase 1 (TAK1) binding protein 1 (TABI), a pseudophosphatase involved in the TGFβ-mediated inflammatory signalling pathway and found to be an essential activator of TAK1 [39,40]. The structure of the TAB1 N-terminal pseudophosphatase domain has been reported and revealed similarity to the PPM family of protein Ser/Thr protein phosphatases [41]. Previous studies have shown that phosphorylation at a C-terminal region of TAB1 regulates TAK1 activity [42–44]. We have recently discovered that TAB1 is dynamically O-GlcNAcylated at Ser395 in the C-terminal domain [45]. This glycosylation appears to be required for full activity of TAB1 and activation of downstream transcription and secretion of pro-inflammatory cytokines. Here, we exploit a novel approach to covalently trap OGT-substrate complexes to explore how OGT recognizes glycosylation substrates through its TPR domain. The structure of the OGT in complex with the TAB1 C-terminal domain combined with mutagenesis studies reveals that OGT recognizes the TAB1 substrate, and by extension a group of glycosylation substrates with similar disordered regions, through extensive essential interactions with the TPR repeats.

2. Results and discussion

2.1. The TAB1 O-GlcNAc site resides in a disordered region with similarity to other OGT targets

The O-GlcNAcylation sites on the OGT substrates TAB1 [45], colapsin response mediator 2 protein (CRMP2) [9] and casein kinase 2 (CK2) [46] are located in disordered regions close to the C-terminus (figure 1a). Although short peptides derived from these sites can be co-crystallized with OGT [31–33], we have been unsuccessful in using this approach with longer sequences/intact proteins to explore the role of the OGT TPR domain in substrate recognition. Aligning the sequences around the O-GlcNAc sites reveals similarities near the site of modification (figure 1b). Remarkably, this is also similar to the proteolytic cleavage site of a HCF1PRO repeat, with the major difference being a glutamate at the acceptor position (figure 1b).

To explore whether a fusion of the C-terminus of a peptide substrate to the N-terminus of a truncated OGT (312–1031) would generate physiological OGT-substrate complexes, we explored this approach first with HCF1PRO. A construct was designed where an 18-mer HCF1PRO repeat peptide (PPCETHETGTTNATTAT) was fused to the N-terminal Thr315 of OGT via a three glycine (3xGly) linker (figure 2). The fusion construct was overexpressed as a His6-tagged protein in E. coli, purified and crystallized. Well-diffracting protein crystals were obtained and synchrotron data were collected to 1.9 Å (electronic supplementary material, table S1). Molecular replacement and subsequent refinement revealed continuous unbiased ϕ[−ρ] = [0] density for both the HCF1PRO peptide and the 3xGly linker (figure 2b). Encouragingly, the conformation of the HCF1PRO peptide in the fusion protein was nearly identical to that observed in the previously published OGT-HCF1 peptide complex [24] (RMSD on Ca atoms = 0.2 Å). Thus, just as with the free peptide, the tethered HCF1PRO peptide backbone binds the OGT TPR domain in an extended conformation, interacting with residues lining the concave surface of the TPR superhelix (figure 2b). Therefore, a linear fusion of OGT and HCF1PRO reproduces the HCF1PRO binding mode.

2.2. A linear fusion of OGT and HCF1PRO reproduces the HCF1PRO binding mode

To explore whether a fusion of the C-terminus of a peptide substrate to the N-terminus of a truncated OGT (312–1031) would generate physiological OGT-substrate complexes, we explored this approach first with HCF1PRO. A construct was designed where an 18-mer HCF1PRO repeat peptide (PPCETHETGTTNATTAT) was fused to the N-terminal Thr315 of OGT via a three glycine (3xGly) linker (figure 2a). The fusion construct was overexpressed as a His6-tagged protein in E. coli, purified and crystallized. Well-diffracting protein crystals were obtained and synchrotron data were collected to 1.9 Å (electronic supplementary material, table S1). Molecular replacement and subsequent refinement revealed continuous unbiased ϕ[−ρ] = [0] density for both the HCF1PRO peptide and the 3xGly linker (figure 2b). Encouragingly, the conformation of the HCF1PRO peptide in the fusion protein was nearly identical to that observed in the previously published OGT-HCF1 peptide complex [24] (RMSD on Ca atoms = 0.2 Å). Thus, just as with the free peptide, the tethered HCF1PRO peptide backbone binds the OGT TPR domain in an extended conformation, interacting with residues lining the concave surface of the TPR superhelix (figure 2b). Therefore, a linear fusion of OGT and HCF1PRO reproduces the HCF1PRO binding mode.

2.3. A linear TAB1:OGT fusion suggests that TAB1 makes extensive interactions with the OGT TPRs

We next explored the OGT-substrate fusion approach as a means of trapping complexes of OGT with TAB1. We
generated a TAB1:OGT fusion construct matching the HCF1PRO:OGT fusion, using an 18-mer TAB1 peptide derived from the S395 glycosylation site (VPYSSAQSTSKTSVTLSL; figure 2a). The chimaeric protein was overexpressed as a His\textsubscript{6}-fusion construct in \textit{E. coli} and purified as described for the HCF1PRO:OGT fusion protein (figure 2a). We were able to generate crystals of the TAB1:OGT fusion protein, solve the structure by molecular replacement and refine the complex against 2.5 Å\textsubscript{\textdegree} synchrotron diffraction data to \textit{R}_{\text{work}}/\textit{R}_{\text{free}} = 0.22/0.25 (electronic supplementary material, table S1). The unbiased \(|F_o| - |F_c|\) density allowed unambiguous building of the linker and peptide (figure 2b). The first eight amino acids of the TAB1 peptide (VPYSSAQST), covering the glycosylation site, were found in a similar conformation in the active site to the free TAB1 peptide in complex with OGT reported previously \cite{31} (figure 2b, RMSD on C\textsubscript{\text{\text{a}}} = 1.4 Å). The electron density revealed Ser395 to be glycosylated as a result of self-glycosylation during expression in \textit{E. coli}, which was confirmed by western blot analysis (figure 3a). The sugar occupies the same position as observed in a complex with a short synthetic TAB1 glycopeptide \cite{31} (maximum atomic shift = 0.1 Å). Intriguingly, there appears to be some extra electron density near Ser396 and Ser399 suggestive of additional glycosylation sites (electronic supplementary material, figure S1) that could be an artefact of the very high (local) concentrations of the fused substrate peptide, or glycosylation occurring in \textit{trans} as a result of the high protein concentrations (approx. 10 mg ml\textsuperscript{-1}) used in the crystallization experiments. In the TAB1:OGT fusion structure, the TAB1 peptide forms two side-chain-mediated interactions (Ser404/Thr406) with the TPR domain of OGT (Asp386/Asp420) (figure 2b). The sequence alignment shows the similarity between TAB1, CRMP2, HCF1 and HCF1\textsubscript{pro} repeat (figure 2b). These are remarkably similar to the interactions between the same OGT residues and Thr1090 and Thr1092 of the HCF1PRO repeat (figure 2b). Similarly, the interactions between the TAB1/HCF1 peptide backbones and the five regularly spaced OGT TPR asparagines are conserved (figure 2b). Furthermore, the overall conformations of the TAB1 and HCF1 peptides in the respective fusion constructs is similar (RMSD on C\textsubscript{\text{\text{a}}} = 1.3 Å). Thus, a linear TAB1:OGT fusion suggests that the TAB1 OGT substrate makes extensive interactions with the OGT TPRs.

2.4. Interactions with the OGT TPRs contribute to TAB1 O-GlcNAcylation

Although the similarity to the HCF1 peptide binding mode and the presence of glycosylation on Ser395 suggests we have trapped a physiologically relevant TAB1:OGT complex, we further tested this model by structure-guided site-directed mutagenesis in the context of truncated OGT (312–1031) and TAB1 (7–409) as separate proteins. Two types of OGT mutants were generated: a single-point mutant in the active site (K842M), known to be essential for catalytic activity \cite{31} and a quintuple mutant where the five key asparagine...
residues that form the bulk of interactions in the TPR domain (Asn322, Asn325, Asn356, Asn390 and Asn424, figure 2b) were all mutated to alanines (from here on referred to as the 5N5A mutant). Based on the TAB1:OGT fusion protein complex, the 5N5A mutations would be expected to disrupt the binding of the C-terminal region of TAB1 to the TPR domain. Using western blot analysis, we probed OGT activity on TAB1 and blotted for O-GlcNAcylation using an O-GlcNAc Ser395 specific antibody [45]. As demonstrated previously, TAB1 is readily O-GlcNAcylated by WT OGT, whereas no glycosylation is observed with the catalytically inactive K842M mutant [31] (figure 3c). The 5N5A mutant shows significantly reduced activity on a free TAB1 peptide (KKPVSVPYSSAQSTSKTSVTLSL) matching the peptide used in the fusion construct (figure 3b), in agreement with the interactions formed by the key asparagines in the TPR domain of OGT observed in the structure (figure 2b).

However, when using a shorter synthetic TAB1 peptide (KKPVSVYSSAQSTKS) ending just before the start of the TPR repeats), the 5N5A mutant shows the same activity levels as WT OGT (electronic supplementary material, figure S2a). Intriguingly, the 5N5A mutant appears to show a more modest reduction of glycosyltransferase activity (approx. 50%), as calculated by quantifying fluorescent signal from the fluorophore conjugated secondary antibodies used in the western blot analysis (electronic supplementary material, figure S2b).
suggesting that while interactions of the TAB1 C-terminus with the OGT TPRs are important, further interactions with the globular pseudophosphatase domain of TAB1 may exist. Nevertheless, interactions with the OGT TPRs contribute to TAB1 O-GlcNAcylation.

3. Concluding remarks

The human O-GlcNAc transferase is a multi-domain protein and is essential in metazoa [12,47,48]. However, it is still unclear how a single OGT enzyme recognizes its multitude of substrates. Previous work has proposed sequence specificity targeting −2 to +3 relative to the acceptor residue [33,49]. Previous work has also suggested the involvement of the TPR domain in substrate recognition by incrementally removing repeats from the TPR domain, resulting in a loss of activity on substrates even on a peptide level, although the molecular basis of this was as yet unclear [20–23,25–27]. Using the proteolytic OGT substrate HCF1, Lazarus et al. [24] revealed the involvement of multiple OGT residues on the concave surface of the TPR domain in binding side chains and backbone of the HCF1PRO repeat proteolytic substrate. Here, we used a fusion approach to trap OGT-substrate complexes to investigate the role of the TPRs in recognition of glycosylation substrates. We first demonstrated that this fusion approach recapitulates the published HCF1 PRO peptide binding mode and then used that to reveal how the C-terminus of the OGT glycosylation substrate TAB1 is recognized by the enzyme. The TAB1 C-terminus binds in an extended conformation in the TPR domain, making extensive contacts with the concave surface through regularly spaced asparagines in OGT. An OGT mutant lacking these asparagines was deficient in glycosylation of TAB1. Interestingly, the data show a complete loss of O-GlcNAcylation of a free TAB1 C-terminal peptide, whereas activity on a TAB1 protein is more modestly reduced. These findings, coupled with recently published work on an OGT substrate sequence preference [33], suggest that OGT may bind its substrates through a combination of mechanisms. It is interesting to note that many other OGT substrates (e.g. Casein kinase II and CRMP2; figure 1a) also possess similarly disordered...
regions C-terminal of the O-GlcNAcylation site, suggesting that this may be a general mode of OGT substrate recognition. However, O-GlcNAc sites have also been reported to reside in/close to secondary structure motifs, as is the case for Histone H2B [50], p53 [51], the glucose-6-phosphate dehydrogenase G6PD [52] and SNAP-29 [53]. It is possible that a subset of substrates is O-GlcNAcylated in a co-translational fashion as proposed by recent work [54]. In this work, we have shown, using crystallography and site-directed mutagenesis, that the OGT substrate TAB1 binds the enzyme in the same way as the proteolytic substrate HCF1 [24] and that the five asparagine residues found on the concave surface of the TPR domain (Asn321, Asn322, Asn356, Asn390 and Asn424) are important for binding. Future studies could be directed at dissecting which other parts of OGT and/or substrate proteins contribute to substrate binding.

4. Material and methods

4.1. Construct design/ cloning

A codon-optimized version of hOGT 313–1031, based on the boundaries described in [31], was ordered from GenScript and cloned as a BamHI-NotI restriction fragment into a modified version of pGEX6p1 containing a 6His tag instead of a GST tag. PCR primers (6H_HCF1_GGG_hOGT_fwd GTATTCACTGACATCACACCACCATGACccgtagactgatgtaatgttgcg; 6H_HCF1_GGG_hOGT_rev CAGGGTGACTGGAGCGCAGCATGGGTaccgccaccaggtggcggtggtggcggtg; 6H_TAB1_GGG_hOGT_fwd GTATTCACTGACATCACACCACCATGACccgtagactgatgtaatgttgcg; 6H_TAB1_GGG_hOGT_rev CAGGGTGACTGGAGCGCAGCATGGGTaccgccaccaggtggcggtggtggcggtg) were then designed to introduce a TAB1 or HCF1 peptide followed by a GGG

4.2. Expression and purification of linear fusion constructs

Both HCF1PRO:OGT and TAB1:OGT fusion constructs were recombinantly expressed as His6-tagged proteins in E. coli BL21 (DE3) pLysS. Cultures were grown in LB media, supplemented with ampicillin, until an OD$_{600}$ of approximately 0.6 was reached. Expression was induced with 250 μM IPTG for 18 h at 16°C. Cells were harvested by centrifugation at 4800g in a J-6.1 centrifuge (Beckmann Coulter). The pellet was resuspended in lysis buffer (25 mM Tris, 150 mM NaCl, 0.5 mM TCEP, and 30 mM imidazole pH 8.5, containing approx. 0.1 mg ml$^{-1}$ lysozyme, 0.1 μg ml$^{-1}$ DNAse, 1 mM benzamidine, 0.2 mM PMSF, and 5 μM leupeptin). Cell lystate was spun down at 20,000g for 10 min in an Avanti J-25 centrifuge (Beckmann). The supernatant was incubated with Nickel-NTA resin for 2 h at 4°C. The beads were isolated by passing through a column and washed extensively with lysis buffer. Bound protein was eluted from the beads with lysis buffer containing 200 mM imidazole. The eluent was diluted to 25 mM NaCl in Tris–HCl pH 8.5 and purified further by anion-exchange chromatography using a HiTrap QFF 5 ml column (GE Healthcare). Fractions corresponding to the size of the fusion protein were pooled, concentrated to less than 2 ml and loaded onto a SuperDex 200 gel filtration column (GE Healthcare). Purity was checked by subjecting the fractions to SDS-PAGE analysis and pure fractions were pooled and buffered exchanged into a 50 mM Tris–HCl pH 8.5, 25 mM NaCl and 0.5 mM TCEP.

4.3. Crystallization and structure solution

The fusion constructs were crystallized at a protein concentration of approximately 10 mg ml$^{-1}$ in the presence of 5 mM UDP. Sitting drop vapour diffusion experiments were performed by combining 0.5 μl of protein with 0.5 μl mother liquor. Crystals grew in 2–3 days in 3.5 M sodium formate, 0.1 M Tris pH 8.5, containing approx. 0.1 mg ml$^{-1}$ lysozyme, 0.1 μg ml$^{-1}$ DNAse, 1 mM benzamidine, 0.2 mM PMSF, and 5 μM leupeptin.

The fusion constructs were crystallized at a protein concentration of approximately 10 mg ml$^{-1}$ in the presence of 5 mM UDP. Sitting drop vapour diffusion experiments were performed by combining 0.5 μl of protein with 0.5 μl mother liquor. Crystals grew in 2–3 days in 3.5 M sodium formate, 0.1 M Tris pH 8.5, containing approx. 0.1 mg ml$^{-1}$ lysozyme, 0.1 μg ml$^{-1}$ DNAse, 1 mM benzamidine, 0.2 mM PMSF, and 5 μM leupeptin. Crystals were cryoprotected using 10% glycerol in mother liquor and 2.5 M lithium sulfate for TAB1:OGT and HCF1PRO:OGT, respectively. Data were collected by subjecting the fractions to SDS-PAGE analysis and pure fractions were pooled and buffered exchanged into a 50 mM Tris–HCl pH 8.5, 25 mM NaCl and 0.5 mM TCEP.

The fusion constructs were crystallized at a protein concentration of approximately 10 mg ml$^{-1}$ in the presence of 5 mM UDP. Sitting drop vapour diffusion experiments were performed by combining 0.5 μl of protein with 0.5 μl mother liquor. Crystals grew in 2–3 days in 3.5 M sodium formate, 0.1 M Tris pH 8.5, containing approx. 0.1 mg ml$^{-1}$ lysozyme, 0.1 μg ml$^{-1}$ DNAse, 1 mM benzamidine, 0.2 mM PMSF, and 5 μM leupeptin. Crystals were cryoprotected using 10% glycerol in mother liquor and 2.5 M lithium sulfate for TAB1:OGT and HCF1PRO:OGT, respectively. Data were collected by subjecting the fractions to SDS-PAGE analysis and pure fractions were pooled and buffered exchanged into a 50 mM Tris–HCl pH 8.5, 25 mM NaCl and 0.5 mM TCEP.

4.4. In vitro TAB1 glycosylation assay

For glycosylation assays, TAB1 and hOGT (WT and mutants) were expressed and purified as described previously [30,31,33,41]. Ten micromolar of TAB1 was incubated with 50 nM hOGT (WT or mutants) in TBS reaction buffer (0.1 M Tris–HCl pH 7.4, 150 mM NaCl) supplemented with 0.5 mM TCEP and 0.1 mg ml$^{-1}$ BSA. The reaction was started.
by adding UDP-GlcNAc to a final concentration of 100 µM and incubating the reaction mixtures at 25°C. Ten microlitres of sample mixtures was taken at indicated times and mixed with 4x LDS sample loading buffer to a final volume of 50 µl and boiled at 95°C for 5 min. Proteins were resolved using precast SDS-PAGE gels (NuPAGE 4–12% Bis-Tris gels, Invitrogen) and blotted onto nitrocellulose membranes (GE Healthcare). The primary antibodies were used at the following concentrations: Anti-TAB1-O-GlcNAc (1:1000 [45]), anti-TAB1 (1:1000, Division of Signal Transduction and Translation, University of Dundee) and anti-OGT (1:2000, DM17, Sigma-Aldrich, Cat#: O6264). Li-Cor secondary antibodies (IRDye 680 Donkey anti-rabbit and IRDye 800 Donkey anti-rabbit, anti-sheep) were used at dilutions of 1:10 000. Blots were imaged using the Li-Cor Odyssey infrared imaging system (Li-Cor, Lincoln, NE). Quantification of the O-GlcNAc specific signal (gTAB1) was performed using ImageStudioLite (Li-Cor) and normalized to total OGT (tOGT) and total TAB1 (tTAB1) signal. Data were plotted with GraphPad Prism 7.

4.5. Steady-state kinetics

hOGT activity was determined in reactions containing 50 nM of either WT or 5N5A His6-hOGT (312–1031), 50 mM Tris–HCl pH 7.4, 0.1 mg ml\(^{-1}\) BSA, 10 µM sodium dithioate and varying concentrations of the TAB1 peptide KKPVSVPYSSAQSTS, in a 200 µM of the TAB1 peptide KKPVSVPYSSAQSTSL or at a fixed concentration and varying concentrations of the TAB1 peptide KKPVSVPYSSAQSTS, in a total volume of 100 µl. Reaction mixtures were preincubated for 15 min before initiating the reaction by adding UDP-GlcNAc to a final concentration of 50 µM. Reactions were incubated for 30 min at 21°C before addition of 200 µl of 75 µM pyrocatechol violet/15 µM fluorophore, DP-sensitive xanthene-based Zn(II) compound [33,62,63], in 25 mM HEPES pH 7.4, 10 mM NaCl, 50% (v/v) MeOH. UDP formation was detected on a Gemini EM fluorescent Microplate reader (Molecular Devices) using excitation and emission wavelengths of 485 nm and 530 nm, respectively. Turnover did not exceed 10% for either substrate. Data are presented as average of three measurements, with error bars showing s.e.m. Data were analysed with GraphPad Prism 7.

4.6. Western blot analysis of purified TAB1:OGT

Samples of purified TAB1:OGT fusion protein were incubated for 30 min at 37°C in the presence and absence of approximately 10 µg ml\(^{-1}\) CpOGA, a promiscuous bacterial O-GlcNAc hydrolase [64]. Samples were supplemented with 4x LDS-loading buffer and boiled for 5 min at 95°C. A total of 0.5 µg of each untreated and treated TAB1:OGT fusion protein were subjected to SDS-PAGE analysis and transferred onto a nitrocellulose membrane (GE Healthcare), using a wet-transfer system (Invitrogen). The membrane was blocked in 5% BSA for 30 min at 21°C before incubating with anti-O-GlcNAc AB (RL2, 1:1000, Abcam, catalogue no. ab2739) and anti-OGT AB (1:2000, Abcam, catalogue no. 177941). Li-Cor secondary antibodies IRDye 680 Donkey anti-mouse (anti-O-GlcNAc) and IRDye 800 Donkey anti-rabbit (anti-OGT) were used at dilutions of 1:10 000. Blots were imaged using the Li-Cor Odyssey infrared imaging system (Li-Cor, Lincoln, NE).

4.7. Peptide synthesis

Microwave-assisted solid-phase peptide synthesis was performed with CEM Liberty automated peptide synthesizer on Rink amide MBHA resin (Novabiochem) using standard Fmoc chemistry protocols. The peptide was cleaved from the resin and deprotected with i-Pr\(_3\)SiH-H\(_2\)O-TFA 2.5:5:92.5 mixture for 2 h. The crude peptide was obtained after dilution of the cleavage mixture with diethyl ether and centrifugation. It was finally purified by reverse-phase HPLC at C18 Waters Xbridge OBD 5 µm 19×100 column in a linear gradient of buffer B (acetonitrile−0.1% trifluoroacetic acid) in buffer A (water−0.1% trifluoroacetic acid) 5−40% in 5 min, flow rate 20 ml min\(^{-1}\). Appropriate fractions were pooled and freeze-dried to provide the target compound as fluffy solid.

5. Accession codes

X-ray diffraction data and refined structures have been deposited in the Protein Data Bank under accession codes 5LWV (HCF1PRO:OGT) and 5LVV (TAB1:OGT).

Data accessibility. This article has no additional data.

Authors’ contributions. D.M.F.v.A. conceived the study; O.R., K.R., A.T.F. and V.K. performed experiments; V.S.B. performed peptide synthesis; K.R. and D.M.F.v.A. analysed and interpreted the data; K.R. and D.M.F.v.A. wrote the manuscript with input from all authors.

Competing interests. The authors declare no conflict of interest.

Funding. K.R. is funded by a BBBSRC Studentship (1416998). This work was funded by a Wellcome Investigator Award (110061) to D.M.F.v.A.

Acknowledgements. We thank the European Synchrotron Radiation Facility (ESRF) for beam time on beam line ID29 and assistance. We thank Winship Herr for useful discussions.

References


sites in TAK1-binding protein (TAB) 1. PLoS ONE 6, e29256. (doi:10.1371/journal.pone.0029256)


