Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

O-GlcNAcylated peptides and proteins for structural and functional studies

Aaron T Balana[1,3], Stuart P Moon[1,3] and Matthew R Pratt[1,2]

O-GlcNAcylation is an enzymatic post-translational modification occurring in hundreds of protein substrates. This modification occurs through the addition of the monosaccharide *N*-acetylglucosamine to serine and threonine residues on intracellular proteins in the cytosol, nucleus, and mitochondria. As a highly dynamic form of modification, changes in *O*-GlcNAc levels coincide with alterations in metabolic state, the presence of stressors, and cellular health. At the protein level, the consequences of the sugar modification can vary, thus necessitating biochemical investigations on protein-specific and site-specific effects. To this end, enzymatic and chemical methods to 'encode' the modification have been developed and the utilization of these synthetic glycopeptides and glycoproteins has since been instrumental in the discovery of the mechanisms by which O-GlcNAcylation can affect a diverse array of biological processes.

Addresses
[1] Departments of Chemistry, University of Southern California, Los Angeles, CA, 90089, United States
[2] Biological Sciences, University of Southern California, Los Angeles, CA, 90089, United States

Corresponding author: Pratt, Matthew R (matthew.pratt@usc.edu)
[3] These authors contributed equally.

Current Opinion in Structural Biology 2021, **68**:84–93

This review comes from a themed issue on **Protein-carbohydrate complexes and glycosylation**

Edited by **Jiaoyang Jiang** and **Spencer J Williams**

For a complete overview see the Issue and the Editorial

Available online 9th January 2021

https://doi.org/10.1016/j.sbi.2020.12.005

Introduction

O-GlcNAcylation is a post-translational modification (PTM) of intracellular proteins wherein the protein is attached to a monomer of β-*N*-acetylglucosamine (GlcNAc) through the side chains of serine and threonine residues (Figure 1a). Akin to protein phosphorylation, this modification is highly dynamic. Throughout a substrate's lifetime, *O*-GlcNAc can be added and removed cyclically by *O*-GlcNAc transferase (OGT) and *O*-GlcNAc hydrolase (OGA), respectively. OGT relies on its catalytic region, comprised of its N-Cat and C-Cat domains, for its transferase activity, while its interactions with substrates are mediated by a number of tetratricopeptide repeat (TPR) domains. OGA contains an N-terminal catalytic glycoside hydrolase domain fused by a stalk domain to a pseudo histone acetyltransferase (HAT) domain [1]. The levels of *O*-GlcNAc are highly responsive to the metabolic state of the cell because OGT uses the high-energy UDP-GlcNAc donor which is produced via the hexosamine biosynthetic pathway, intimately linking global *O*-GlcNAc levels to the metabolic state of the cell [2]. The disease state of the cell also impacts global *O*-GlcNAcylation as evidenced by the detection of perturbed modification levels in many cancers and neurodegenerative diseases [3,4] This modification imparts its functional effects on substrates in a variety of ways dependent on the protein's biochemical and biophysical properties. By competing with phosphorylation on the same Ser/Thr residues, *O*-GlcNAcylation can modulate protein function through PTM cross-talk [5]. The modification can also have profound effect on the protein-protein interactions of its substrates [6].

This review seeks to consolidate current literature concerning the site-specific effects of protein O-GlcNAcylation. We first present an overview of various techniques used to precisely encode the PTM onto polypeptides. Then, we review studies that use these modified polypeptides to examine the structure and function of OGT and OGA. We then summarize research by our lab and others into the biochemistry and biophysics of *O*-GlcNAc modified proteins. Finally, we highlight where the properties of O-GlcNAc have been exploited in non-native contexts.

Methods for the preparation of *O*-GlcNAc modified peptides and proteins

The earliest method described to produce highly homogeneous *O*-GlcNAcylated polypeptides involves the enzymatic modification of protein substrates using *in vitro* reactions with purified OGT [7]. Originally this technique involved the expression and purification of OGT and a putative substrate protein and their subsequent *in vitro* incubation. This approach proved useful on a variety of peptide and protein substrates and led to the characterization of the kinetic parameters for OGT's variable catalytic activities towards different substrates [8]; however, it requires the relatively difficult expression and refolding of OGT. As an alternative, an OGT and substrate co-expression approach was developed to work in culture in *Escherichia coli* [9] and eukaryotic systems [10] where both proteins are expressed prior to a single lysis

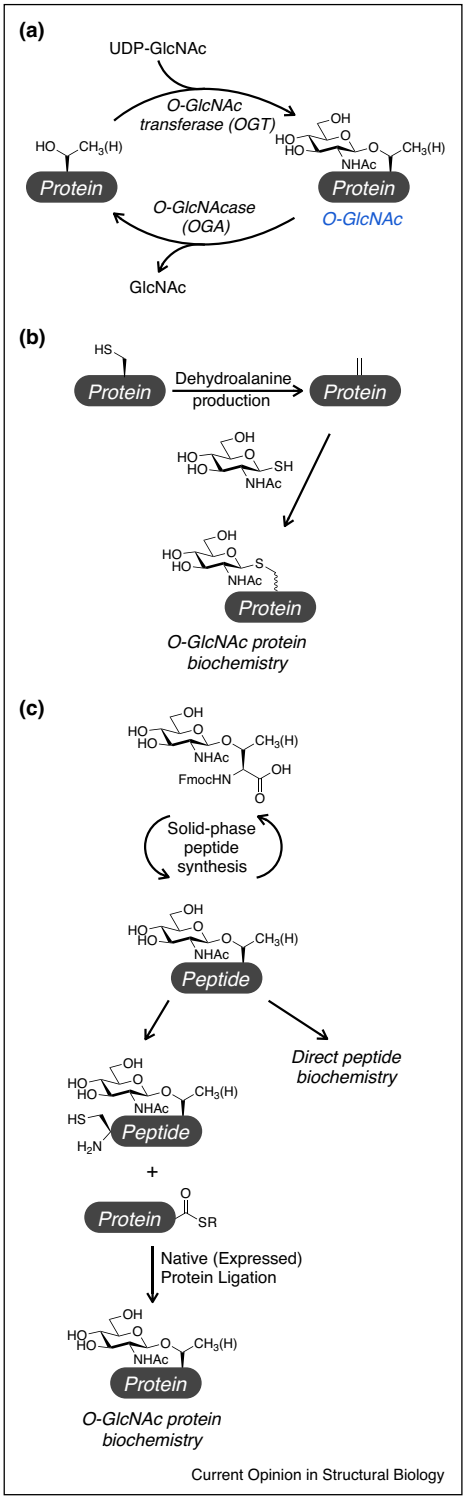
and isolation step. Further optimizations [11] led to improved efficiency and yield of protein expression and modification stoichiometry. More recently, the *E. coli* system also incorporated overexpression of UDP-GlcNAc biosynthetic pathway enzymes GlmM and GlmU for better donor sugar availability required during *in vivo* OGT modification [12•].

The major downsides to these enzymatic modification approaches is the resulting heterogeneity of the protein product and the presence of remaining unmodified substrate. The addition of *O*-GlcNAc by OGT is substoichiometric leading to varying levels of non-incorporation depending on protein substrate. While the completeness of the *in vitro* modification can be pushed with longer incubation, less stable proteins can degrade, precipitate, or lose activity with extended reaction times. In co-expression systems, the modification can also be removed by endogenous glycosidases [13] further contributing to lower modification efficiencies. Another source of heterogeneity stems from the fact that for a number of known substrates, OGT can modify multiple serine and threonine modification sites within the protein sequence. Enzymatic approaches hence result in a mixture of unmodified, singly-modified and multiply-modified proteins that are ultimately challenging to purify as the sugar modification does minimal alteration to the protein's size, polarity, or charge.

Hence for investigations requiring highly homogeneous and site-specifically O-GlcNAc modified peptides or proteins, chemical methods have proven to be more useful. One proposed method is posttranslational mutagenesis which involves the chemoselective installation of the GlcNAc sugar onto engineered dehydroalanine residues on a target protein (Figure 1b) [14]. While this method is arguably the simplest and most straightforward way to install the modification, it does not perfectly recapitulate *O*-GlcNAc modification as the sugar is linked either through a cysteine thio-linkage [15] or a homohomoserine *O*-linkage [16] that is one carbon longer than found in nature. Most importantly, posttranslational mutagenesis causes racemization at the α-carbon generating an often inseparable mixture of diastereomer that may have different biochemistry.

The only method for site-specific and homogeneous *O*-GlcNAc modification of polypeptides, to date, is the installation of sugar-modified amino acid building blocks during solid phase peptide synthesis (Figure 1c). These *O*-GlcNAcylated Fmoc-serine or threonine monomers can be prepared using a variety of available synthetic routes [17], but are also conveniently available for purchase from commercial sources. To overcome the size limitations of peptide synthesis, chemical ligation techniques have also been employed to synthesize longer polypeptides or full-length proteins (Figure 1c). The gold

Figure 1



O-GlcNAc modification and methods to prepare site-specifically modified proteins. **(a)** *O*-GlcNAc modification is the dynamic addition of *N*-acetylglucosamine to serine/threonine side-chains of intracellular proteins. **(b)** Posttranslational mutagenesis can be used to transform a cysteine residue into an S-linked analog of *O*-GlcNAc. **(c)** Solid-phase peptide synthesis used alone or in combination with protein ligation techniques can be used to prepare *O*-GlcNAc modified peptides and proteins.

standard in the field is native chemical ligation (NCL), which involves the reaction of a peptide C-terminal thioester with an N-terminal cysteine containing peptide resulting in the formation of the native amide bond [18]. An extension of this technique, expressed protein ligation (EPL), enables the recombinant production of protein C-terminal thioesters, dramatically extending the scope of this approach [19]. As a robust and flexible techniques, NCL/EPL have been the most widely used method to prepare a number of site-specifically *O*-GlcNAc-modified full-length proteins.

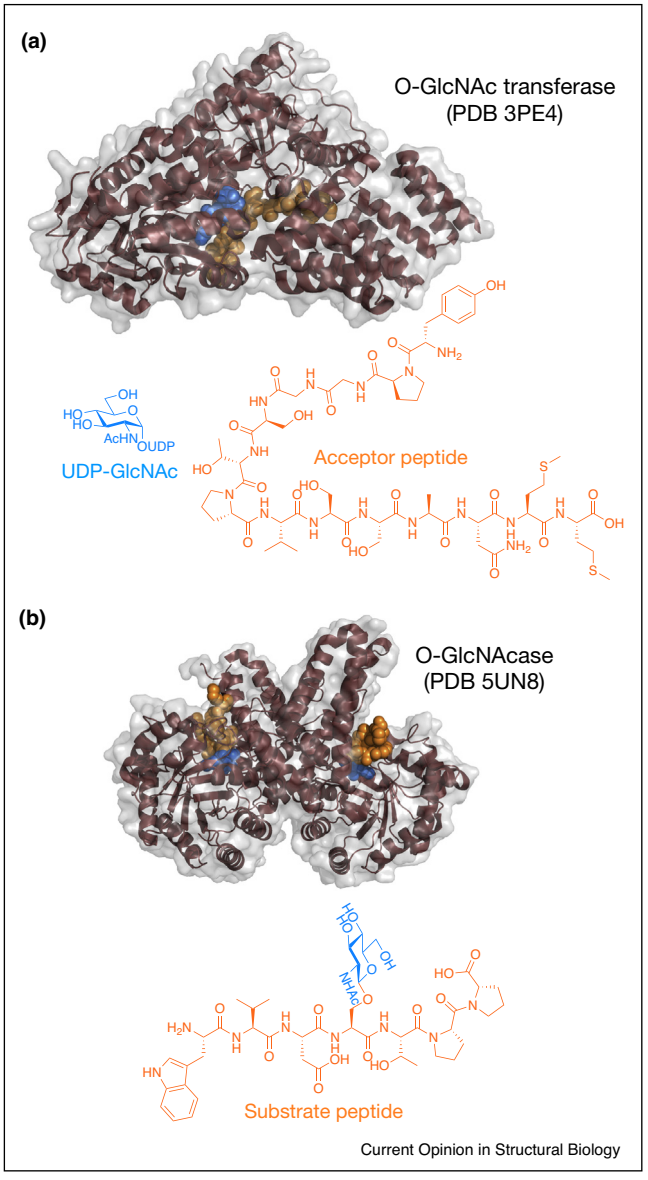
Synthetic polypeptides applied to the study of *O*-GlcNAc cycling enzymes

A composite model of the full-length human OGT was determined in 2011 both as a binary complex with UDP, and in a ternary complex with a casein kinase II-derived peptide substrate [20] (Figure 2a). For this structure, a truncated version of the enzyme with 4.5 of the 13 TPRs was used during the crystallization experiments and the TPR region was modeled in from a separate structure [21]. Although complex formation with the UDP-GlcNAc sugar donor was also attempted in this work, the hydrolysis of the sugar precluded the crystallization process. In order to include the sugar moiety in the structural characterization, an *O*-GlcNAcylated peptide from substrate TAB1 was used [22]. Alternatively, the use of a 5S-GlcNAc sugar analog previously shown to inhibit OGT activity [23] enabled successful crystallization of complexes of OGT with UDP-5SGlcNAc and various substrate peptides [22,24,25]. In these experiments, obtaining a crystal structure in the presence of synthetic peptide substrates was important in characterizing the distinct binding modes of OGT as it reveals the peptide-binding cleft not seen in the OGT-UDP or OGT-UDP-5SGlcNAc complexes.

While these structures definitively describe the UDP-GlcNAc binding pocket and the catalytic residues, the basis for how OGT recognizes its peptide substrates is less understood. Given that OGT does not appear to have a strict sequence preference for residues near the serine/threonine acceptor site, cataloging its substrates [26] has been a useful approach to determine potential contributing factors for recognition. Through *in vitro* modification of a 720-member synthetic peptide library [27], it was demonstrated that there is some degree of amino acid preference at the −3 to 2 sites of the peptide, suggesting that OGT's catalytic domain surrounding this substrate region imposes some selectivity constraints. This peptide library modification approach was also miniaturized to a microarray format [28] for high-throughput identification of novel OGT substrates. Also contributing to substrate recognition is the participation of the TPR region where the extended TPR interacts with the solvent-exposed region of protein substrates through a 'ladder' of asparagine residues [29,30]. This was corroborated in a microarray analysis of >6000 proteins showing that asparagine-to-alanine mutations in the TPR ladder results in retention OGT activity in shorter peptides but not full-length proteins [31]. Moreover, a similar protein microarray analysis was also utilized to demonstrate that additional aspartic acid residues in the TPR also contribute to substrate recognition [32].

Knowledge of OGT's catalytic site was used to guide the development of small molecule OGT inhibitors through 5S-GlcNAc derivatization [33,34] or structure-guided

Figure 2



Structural characterization of *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA). Synthetic (glyco)peptides played key roles in helping determine the biochemistry of OGT and OGA using structural biology.

medicinal chemistry [35,36]. In addition, synthetic peptides have also been proposed as OGT inhibitors. Noting that the sugar moiety in UDP-GlcNAc does not contribute to its binding to OGT, a short peptide with an acceptor serine was linked directly to UDP with a carbon chain tether [37]. Termed goblins (**OG**T **b**isubstrate-**l**inked **in**hibitors), these bisubstrate-peptide conjugates were able to inhibit OGT activity, albeit having lower affinities compared to small molecule inhibitors. By replacing the serine with cysteine in the peptide portion to generate S-linked UDP-peptide conjugates, a newer generation of goblins was developed with improved affinities comparable to the best performing OGT inhibitors [38]; however, their bi-substrate nature yields poor cell permeability. Another modification on this approach removed the diphosphate of UDP as a strategy to improve cellular permeability leading to peptides conjugated to uridine-like scaffolds with some inhibitory activity but relatively weak affinity [39].

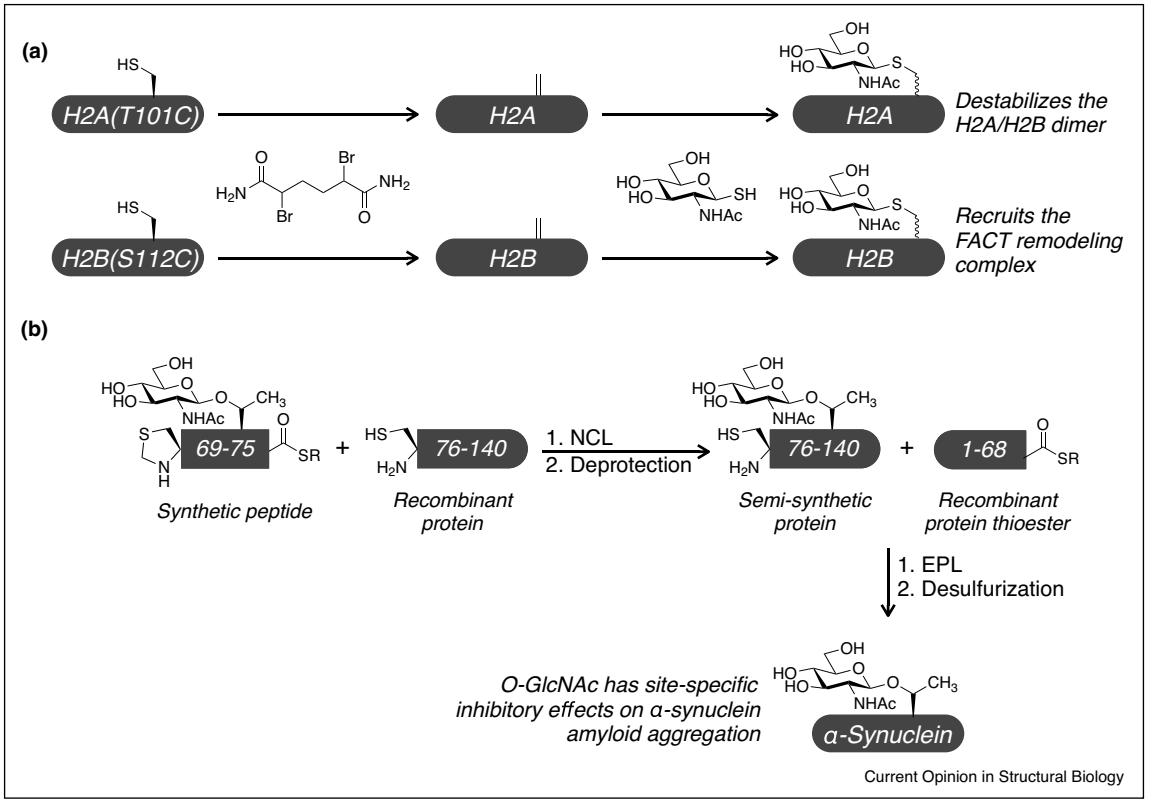
For structural studies on OGA, three separate groups concurrently published similar apo and inhibitor-bound crystal structures [40–42]. Both groups utilized truncated versions of OGA that lack the C-terminal HAT domain of yet unknown significance for its activity. These structures show a homodimeric configuration where the stalk domain of one OGA molecule is enclosed by the alpha helical barrel of the catalytic domain of the sister molecule. This interface appears to form the substrate-binding cleft as the co-crystal with the transition state analog inhibitor Thiamet-G [43] shows the pyranose ring of the inhibitor sitting at this interface. In addition to the inhibitor-bound structure, Li *et al.* also obtained a crystal structure of an *O*-GlcNAcylated p53 peptide bound to catalytically-inactive OGA (Figure 2b). This structure identified multiple key residues that strongly interacted with the GlcNAc moiety as well as a hydrophobic substrate cleft that contributes to the p53 peptide side chain recognition. Interestingly, later co-crystallization experiments of OGA with four other glycopeptides [44•] revealed that while the GlcNAc conformation generally remains the same, the different peptides can be bound in a variety of modalities, rationalizing OGA's adaptability to deglycosylate a diverse array of sequences.

A recent development in the field is the discovery that cysteine residues can also undergo enzymatic GlcNAc modification to form thio-linked S-GlcNAc sites. Notably, *S*-GlcNAcylation was previously proposed as an artificial strategy to prepare metabolically stable GlcNAc modified peptides and proteins [45,46]. Through a proteomics approach, *S*-GlcNAcylation was shown to occur in living systems while *in vitro* OGT modification of substrate peptides whose serine acceptors were replaced with cysteines confirmed that this process occurs enzymatically [47]. With the use of an *S*-GlcNAcylated synthetic peptide and a semi-synthetic protein, the stability of this modification towards human OGA removal was demonstrated in *in vitro* hydrolysis experiments [48]. Computational and biochemical analyses were also used to demonstrate that *S*-GlcNAc is a suitable structural and functional mimic for *O*-GlcNAc, at least in certain cases. This mimetic approach was later used to study the consequence of Ser-405 *O*-GlcNAc modification in OGA [49••]. An OGA peptide bearing a cysteine GlcNAc modification was similarly found to be more stable in *in vitro* hydrolysis experiments with bacterial CpOGA compared to its *O*-GlcNAc counterpart. Genetically converting OGA's serine-405 to cysteine in mammalian cells through CRISPR technology showed that *in vivo* S-GlcNAcylation results in higher modification stoichiometry as a consequence of its nonhydrolyzable nature, with a level of upregulation equivalent to OGA chemical inhibition. Importantly, this genetic substitution led to the discovery that GlcNAc modification at position 405 reduces OGA's stability and half-life compared to the unmodified enzyme.

Structural and functional studies of *O*-GlcNAcylated substrates

O-GlcNAc can influence its substrates through its interplay with protein phosphorylation. The first semi-synthetic *O*-GlcNAcylated protein was generated using SPPS and EPL to study this cross-talk in the context of kinase CK2 [46]. In this case, the authors took advantage of the O-substitution to S-substitution mentioned above to produce a stable *O*-GlcNAc mimic for cell microinjection studies. They showed that phosphorylation at T344 improved the stability of the protein, and that *S*-GlcNAcylation at S347 block the endogenous phosphorylation at T344, resulting in CK2 degredation. Additionally, the presence of the glycan lead to an altered kinase substrate profile compared to wild-type and phosphorylated variants, presumably via contacts with CK2 substrates. To further study the cross-talk between these PTMs, the Pieters lab developed a synthetic peptide microarray to identify peptides that could be modified by both Jak2 and OGT [50,51]. They discovered that the phosphorylation (pY364) of a peptide corresponding to ZO-3 significantly impeded its *O*-GlcNAcylation (gS369) due to disruption of peptide-OGT contacts, whereas the inverse was not found to be true. A further study added to this finding by showing that phosphorylation of the ZO-3 peptide inhibited *O*-GlcNAc hydrolysis by OGA, while dephosphorylation was only slightly impacted by the presence of the glycan. With these findings and their own, Leney *et al.* used kinetic-based mass spectrometry assays to determine and validate a specific cross-talk motif: (pS/pT)P(V/A/T)(gS/gT) [52]. This motif is highly identified in the phosphoproteome, and it can be used to identify putative *O*-GlcNAcylation sites, particularly those with the potential for phosphorylation interplay.

Figure 3



Protein synthesis for studying *O*-GlcNAc biochemistry. **(a)** Posttranslational mutagenesis was used to install *O*-GlcNAc analogs onto the the histone proteins H2A and H2B, allowing the effects of site-specific *O*-GlcNAc on nucleosome biology. **(b)** Protein ligation was used to generate *O*-GlcNAc modified versions of α-synuclein, showing that these glycans have site-specific effects on amyloid formation. The synthesis of α-synuclein with O-GlcNAc at threonine 72 is shown as an example.

Additionally, *O*-GlcNAc modification can also impart its functional effects by positively or negatively influencing the abilities of its substrates to interact with their binding partners. The Davis lab studied these consequences on histone stability and function [15,53]. Semisynthetic *O*-GlcNAcylated histones were generated by expressing H2A or H2B with a cysteine mutations capable of conversion to dehydroalanine before introduction of an S-GlcNAc monomer via a Michael addition (Figure 3a). Glycosylation of H2A at T101 destabilizes its interaction with H2B, while modification of H2B's S112 recruits the FACT complex, both of which significantly aid in transcription elongation. The Boyce group, attempting to determine protein interactors with *O*-GlcNAc with synthetic, OGT-treated, biotinylated peptides, identified the 14-3-3 family of proteins [54]. Further structural studies showed that this interaction is mediated through extensive hydrogen bonding between the glycan and 14-3-3 isoform binding pockets. Two of our own recent studies also examine interactions impacted by *O*-GlcNAc modification. In our study of the *O*-GlcNAc modification of caspase-8, we generated *O*-GlcNAc-modified peptides derived from the sequences surrounding the protein's self-cleavage/activation sites [55]. *O*-GlcNAcylated peptides were significantly resistant to active caspase-8 cleavage, presumably via the obscuring of the cleavage sites by the glycans. We also used SPPS and EPL to generate α-synuclein variants modified at T72 and at T87 and studied their interactions with the protease calpain [56]. The presence of these glycans modified calpain binding as evidenced by changes in cleavage sites: both modifications obviated nearby cleavage sites, while glycosylation at T72 resulted in the appearance of a new cleavage site, implying that this modification both stabilizes and destabilizes different protein-protein interactions simultaneously.

A particularly important facet of these interaction effects is their implications in neurodegenerative disorders. A number of identified *O*-GlcNAcylation substrates participate in amyloidogenic pathways, often with the modification imparting anti-amyloid effects. One such protein is tau, the aggregation of which leads to Alzheimer's disease pathology. Frenkel-Pinter *et al.* used SPPS to construct *O*-

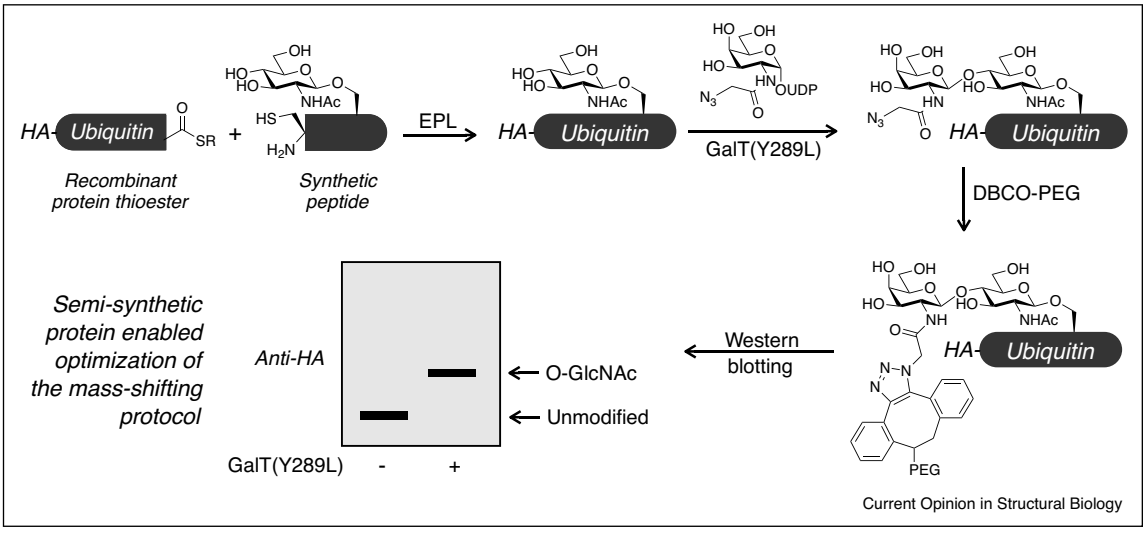
GlcNAc-modified peptides corresponding to PHF6, a hexamer required for tau oligomer formation [57]. These modified peptides aggregated to a far lesser degree than the unmodified control peptides, and, more importantly, inhibited the aggregation extent and kinetics of unmodified peptides in co-aggregation experiments. Further, the Hackenberger group has also used SPPS, NCL, and EPL to generate S400-glycosylated tau, which can be used to study the PTM's effect on aggregation using a full-length variant [58] (Figure 3b). Our group has also studied the aggregatory effects of *O*-GlcNAcylation of α-synuclein, the Parkinson's disease analog of tau. Recently, we used EPL and SPPS to construct full-length, *O*-GlcNAcylated variants of α-synuclein (gT72, gT75, gT81, gT87, and triply-modified gT72/gT75/gT81) and used these proteins to show that *O*-GlcNAc modification is extremely inhibitory to the aggregation process, but to site-dependent extents [59,60,61••]. Also, we showed that these PTMs can alter the architecture of the aggregates that form and can also impact their relative cytotoxicity in primary neurons. These effects are presumably due to the hydrophilic glycan's disruption of the hydrophobic interactions required for aggregation. Another player in the neurodegenerative aggregation scheme are small heat-shock proteins (sHsps), which protect against the formation of cytotoxic aggregates by acting as chaperones of unstructured proteins. In our recent work, we prepared *O*-GlcNAcylated sHsps Hsp27, αA-crystallin, and αB-crystallin via SPPS and EPL and determined that the modification is aggregation-protective because of the improvement of the chaperone activity of these sHsps [62••]. We show this effect is caused by the disruption the back-binding of the substrates' IXI domains to their chaperone clefts, circumventing their auto-regulation and imparting an 'always-on' effect [62••].

Other application for synthetic *O*-GlcNAc peptides

While various chemical and biological methods have been developed for the detection, enrichment, and identification of *O*-GlcNAc substrates, the most convenient and highly used technique is through the use of antibodies during Western blotting, immunoprecipitation, ELISAs, or tissue staining [63]. A number of these antibodies are pan-selective and are able to recognize GlcNAc regardless of the protein sequence or identity [64]. For the production of such antibodies, various antigens have been used during immunization including nuclear pore complex fractions [65] or synthetic glycopeptides [66,67]. Unfortunately, some of these antibodies suffer from selectivity by cross-reacting with terminal residues in complex glycans [63]. Furthermore, these antibodies possibly do not exhibit true 'pan' selectivity based on non-reactivity to *bona fide* *O*-GlcNAc-modified synthetic proteins [61••]. In order to improve sensitivity, as well as to allow site-specific detection, protein-specific antibodies have also been developed for a number of substrates again with the use of synthetic glycopeptides [68].

The extent to which *O*-GlcNAc modification affects protein function is closely tied to its stoichiometry, which varies widely from substrate to substrate. Traditional mass spectrometry techniques can precisely identify sites of modification, but are unable to determine modification number on a particular protein under a given set of conditions. To this end, our lab (in collaboration with

Figure 4



Determining *O*-GlcNAc stoichiometry using mass shifting. Synthetic, homogeneously *O*-GlcNAc modified ubiquitin was used to optimize conditions for chemoenzymatic mass-shifting of *O*-GlcNAc modified proteins.

the Hsieh-Wilson lab) optimized a Western blotting assay capable of determining modification stoichiometry [10,69•]. Briefly, in this technique, endogenously *O*-GlcNAcylated proteins in cell lysates are appended with azide-labeled galactose monomers via a selectively-mutated galactosyltransferase. These azide handles can then be linked to DBCO-functionalized PEG, resulting in gel-shifting mass tags proportional to the modification's stoichiometry which are visualizable via Western blotting. To optimize this protocol, we used EPL to generate an HA-tagged ubiquitin thioester which we reacted with either an unmodified or *O*-GlcNAcylated small peptide prepared via SPPS. This enabled us to have a 100% modified protein control to test different conditions and optimize the overall protocol (Figure 4).

Finally, *O*-GlcNAc modifications have also been introduced to pharmacologically-active peptides as a strategy to modulate their therapeutic potential. Notably, these peptides are non-native substrates for *O*-GlcNAcylation by OGT. Inspired by the observation that *O*-GlcNAc modifications improve the stability of certain substrates against proteolytic cleavage [56], the addition of the sugar to GPCR peptide agonists GLP-1 (glucagon-like peptide) and PTHRP (parathyroid hormone receptor peptide) was proposed as a potentially generalizable peptide engineering strategy to improve serum half-lives [70•] Indeed, the strategy worked in certain peptide variants without adversely affecting potency or affinity for the receptor. In addition, the production of the bacteriostatic *Lactobacillus plantarum* di-GlcNAcylated 43-mer peptide glycocin F has also been described [71]. Glycocin F is naturally GlcNAc modified by the bacterial glycosyltransferases at Ser19 and Cys43. Using a sophisticated semi-synthesis strategy, variants of glycocin F were prepared and tested for differences in antibacterial activities. Interestingly, a glycocin F variant that had *O*-GlcNAc to *S*-GlcNAc substitution showed improved biological activity likely as a consequence of enhanced stability of *S*-GlcNAc against hydrolysis by bacterial glycosidases.

Conclusions

O-GlcNAc modifications play myriad roles in a number of different cellular processes by imposing varied biochemical and biophysical characteristics on its substrates. By leveraging a chemical protein synthesis toolbox, researchers can generate homogeneously modified proteins for the interrogation of modification consequences. These techniques have facilitated studies into the readers, writers, and erasers of *O*-GlcNAc modification, as well as the direct impacts on its substrates themselves. Further, semisynthesis of *O*-GlcNAcylated proteins has enabled the optimization of techniques to study the endogenous modification and has allowed for the modulation of polypeptide-based therapeutics. Together, these works highlight the utility of protein semi-synthesis to probe the biological implications of protein PTMs and establish that *O*-GlcNAcylation is highly multifaceted in terms of its substrates and its effects.

Conflict of interest statement

Nothing declared.

Acknowledgements

Our research in this area is supported by the National Institutes of Health (R01GM114537), the National Science Foundation (CHE1905081), and the Anton Burg Foundation. S.p.m. is supported by NIGMST32GM118289, and A.T.B is supported as a Dornsife Chemistry-Biology Interface Trainee. We apologize to those researchers whose work we have not included due to space limitations.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as

- • of special interest
- •• of outstanding interest

1. Joiner CM, Li H, Jiang J, Walker S: **Structural characterization of the O-GlcNAc cycling enzymes: insights into substrate recognition and catalytic mechanisms**. *Curr Opin Struct Biol* 2019, **56**:97-106 http://dx.doi.org/10.1016/j.sbi.2018.12.003.

2. Walgren JLEE, Vincent TS, Schey KL, Buse MG: **High glucose and insulin promote O-GlcNAc modification of proteins, including α-tubulin**. *Am J Physiol Metab* 2003, **284**:E424-E434 http://dx.doi.org/10.1152/ajpendo.00382.2002.

3. Ma Z, Vocadlo DJ, Vosseller K: **Hyper-O-GlcNAcylation is anti-apoptotic and maintains constitutive NF-KB activity in pancreatic cancer cells**. *J Biol Chem* 2013, **288**:15121-15130 http://dx.doi.org/10.1074/jbc.M113.470047.

4. Pinho TS, Correia SC, Perry G, Ambrósio AF, Moreira PI: **Diminished O-GlcNAcylation in Alzheimer's disease is strongly correlated with mitochondrial anomalies**. *Biochim Biophys Acta - Mol Basis Dis* 2019, **1865**:2048-2059 http://dx.doi.org/10.1016/j.bbadis.2018.10.037.

5. Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O: **Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease**. *Annu Rev Biochem* 2011, **80**:825-858 http://dx.doi.org/10.1146/annurev-biochem-060608-102511.

6. Tarbet HJ, Toleman CA, Boyce M: **A sweet embrace: control of protein–protein interactions by O-linked β-N-acetylglucosamine**. *Biochemistry* 2017, **57**:13-21 http://dx.doi.org/10.1021/acs.biochem.7b00871.

7. Lubas WA, Hanover JA: **Functional expression of O-linked GlcNAc transferase**. *J Biol Chem* 2000, **275**:10983-10988 http://dx.doi.org/10.1074/jbc.275.15.10983.

8. Shen DL, Gloster TM, Yuzwa SA, Vocadlo DJ: **Insights into O-linked N-acetylglucosamine (O-GlcNAc) processing and dynamics through kinetic analysis of O-GlcNAc transferase and O-GlcNAcase activity on protein substrates**. *J Biol Chem* 2012, **287**:15395-15408 http://dx.doi.org/10.1074/jbc.M111.310664.

9. Lim, Hong K, Ha CH, Chang HI: **Production of O-GlcNAc modified recombinant proteins in *Escherichia coli***. *J Microbiol Biotechnol* 2002, **12**:306-311 http://dx.doi.org/10.1590/S1415-47572004000300022.

10. Rexach JE, Rogers CJ, Yu S-HH, Tao J, Sun YE, Hsieh-Wilson LC: **Quantification of O-glycosylation stoichiometry and dynamics using resolvable mass tags**. *Nat Chem Biol* 2010, **6**:645-651 http://dx.doi.org/10.1038/nchembio.412.

11. Han C, Shan H, Bi C, Zhang X, Qi J, Zhang B, Gu Y, Yu W: **A highly effective and adjustable dual plasmid system for O-GlcNAcylated recombinant protein production in *E. coli***. *J Biochem* 2015, **157**:477-484 http://dx.doi.org/10.1093/jb/mvv006.

12. • Gao H, Shi M, Wang R, Wang C, Shao C, Gu Y, Yu W: **A widely compatible expression system for the production of highly O-GlcNAcylated recombinant protein in *Escherichia coli***. *Glycobiology* 2018, **28**:949-957 http://dx.doi.org/10.1093/glycob/cwy077.
This work demonstrated that overexpression of enzymes required in bacterial UDP-GlcNAc biosynthesis during co-expression in *E. coli* results in improved sugar donor bioavailability and greater modification stoichiometries.

13. Goodwin OY, Thomasson MS, Lin AJ, Sweeney MM, Macnaughtan MA: ***E. coli* sabotages the in vivo production of O-linked β-N-acetylglucosamine-modified proteins**. *J Biotechnol* 2013, **168**:315-323 http://dx.doi.org/10.1016/j.jbiotec.2013.10.008.

14. Chalker JM, Lercher L, Rose NR, Schofield CJ, Davis BG: **Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications**. *Angew Chem Int Ed* 2012, **51**:1835-1839 http://dx.doi.org/10.1002/anie.201106432.

15. Lercher L, Raj R, Patel NA, Price J, Mohammed S, Robinson CV, Schofield CJ, Davis BG: **Generation of a synthetic GlcNAcylated nucleosome reveals regulation of stability by H2A-Thr101 GlcNAcylation**. *Nat Commun* 2015, **6**:7978 http://dx.doi.org/10.1038/ncomms8978.

16. Wright TH, Bower BJ, Chalker JM, Bernardes GJLL, Wiewiora R, Ng WL, Raj R, Faulkner S, Vallee MRJ, Phanumartwiwath A *et al.*: **Posttranslational mutagenesis: a chemical strategy for exploring protein side-chain diversity**. *Science (80-.)* 2016, **354** http://dx.doi.org/10.1126/science.aag1465 aag1465–aag1465.

17. De Leon CA, Lang G, Saavedra MI, Pratt MR: **Simple and efficient preparation of O- and S-GlcNAcylated amino acids through InBr3-catalyzed synthesis of β-N-acetylglycosides from commercially available reagents**. *Org Lett* 2018, **20**:5032-5035 http://dx.doi.org/10.1021/acs.orglett.8b02182.

18. Dawson PE, Muir TW, Clark-Lewis I, Kent SB: **Synthesis of proteins by native chemical ligation**. *Science (80-)* 1994, **266**:776-779 http://dx.doi.org/10.1126/science.7973629.

19. Muir TW, Sondhi D, Cole PA: **Expressed protein ligation: a general method for protein engineering**. *Proc Natl Acad Sci U S A* 1998, **95**:6705-6710 http://dx.doi.org/10.1073/pnas.95.12.6705.

20. Lazarus MB, Nam Y, Jiang J, Sliz P, Walker S: **Structure of human O-GlcNAc transferase and its complex with a peptide substrate**. *Nature* 2011, **469**:564-567 http://dx.doi.org/10.1038/nature09638.

21. Jínek M, Rehwinkel J, Lazarus BD, Izaurralde E, Hanover JA, Conti E: **The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin α**. *Nat Struct Mol Biol* 2004, **11**:1001-1007 http://dx.doi.org/10.1038/nsmb833.

22. Schimpl M, Zheng X, Borodkin VS, Blair DE, Ferenbach AT, Schüttelkopf AW, Navratilova I, Aristotelous T, Albarbarawi O, Robinson DA *et al.*: **O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis**. *Nat Chem Biol* 2012, **8**:969-974 http://dx.doi.org/10.1038/nchembio.1108.

23. Gloster TM, Zandberg WF, Heinonen JE, Shen DL, Deng L, Vocadlo DJ: **Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells**. *Nat Chem Biol* 2011, **7**:174-181 http://dx.doi.org/10.1038/nchembio.520.

24. Lazarus MB, Jiang J, Gloster TM, Zandberg WF, Whitworth GE, Vocadlo DJ, Walker S: **Structural snapshots of the reaction coordinate for O-GlcNAc transferase**. *Nat Chem Biol* 2012, **8**:966-968 http://dx.doi.org/10.1038/nchembio.1109.

25. Lazarus MB, Jiang J, Kapuria V, Bhuiyan T, Janetzko J, Zandberg WF, Vocadlo DJ, Herr W, Walker S: **HCF-1 is cleaved in the active site of O-GlcNAc transferase**. *Science (80-)* 2013, **342**:1235-1239 http://dx.doi.org/10.1126/science.1243990.

26. Chalkley RJ, Thalhammer A, Schoepfer R, Burlingame AL: **Identification of protein O-GlcNAcylation sites using electron transfer dissociation mass spectrometry on native peptides**. *Proc Natl Acad Sci U S A* 2009, **106**:8894-8899 http://dx.doi.org/10.1073/pnas.0900288106.

27. Pathak S, Alonso J, Schimpl M, Rafie K, Blair DE, Borodkin VS, Schüttelkopf AW, Albarbarawi O, Van Aalten DMFF: **The active site of O-GlcNAc transferase imposes constraints on substrate sequence**. *Nat Struct Mol Biol* 2015, **22**:744-749 http://dx.doi.org/10.2210/pdb5bnw/pdb.

28. Shi J, Sharif S, Ruijtenbeek R, Pieters RJ: **Activity Based high-throughput screening for novel O-GlcNAc transferase substrates using a dynamic peptide microarray**. *PLoS One* 2016, **11** http://dx.doi.org/10.1371/journal.pone.0151085 e0151085.

29. Clarke AJ, Hurtado-Guerrero R, Pathak S, Schüttelkopf AW, Borodkin V, Shepherd SM, Ibrahim AFMM, van Aalten DMFF: **Structural insights into mechanism and specificity of O-GlcNAc transferase**. *EMBO J* 2008, **27**:2780-2788 http://dx.doi.org/10.1038/emboj.2008.186.

30. Rafie K, Raimi O, Ferenbach AT, Borodkin VS, Kapuria V, van Aalten DMF: **Recognition of a glycosylation substrate by the O-GlcNAc transferase TPR repeats**. *Open Biol* 2017, **7** http://dx.doi.org/10.1098/rsob.170078 170078.

31. Levine ZG, Fan C, Melicher MS, Orman M, Benjamin T, Walker S: **O-GlcNAc transferase recognizes protein substrates using an asparagine ladder in the tetratricopeptide repeat (TPR) superhelix**. *J Am Chem Soc* 2018, **140**:3510-3513 http://dx.doi.org/10.1021/jacs.7b13546.

32. Joiner CM, Levine ZG, Aonbangkhen C, Woo CM, Walker S: **Aspartate residues far from the active site drive O-GlcNAc transferase substrate selection**. *J Am Chem Soc* 2019, **141**:12974-12978 http://dx.doi.org/10.1021/jacs.9b06061.

33. Liu T-WW, Zandberg WF, Gloster TM, Deng L, Murray KD, Shan X, Vocadlo DJ: **Metabolic inhibitors of O-GlcNAc transferase that act in vivo implicate decreased O-GlcNAc levels in leptin-mediated nutrient sensing**. *Angew Chem Int Ed* 2018, **57**:7644-7648 http://dx.doi.org/10.1002/anie.201803254.

34. Worth M, Hu CW, Li H, Fan D, Estevez A, Zhu D, Wang A, Jiang J: **Targeted covalent inhibition of: O-GlcNAc transferase in cells**. *Chem Commun* 2019, **55**:13291-13294 http://dx.doi.org/10.1039/c9cc04560k.

35. Liu Y, Ren Y, Cao Y, Huang H, Wu Q, Li W, Wu S, Zhang J: **Discovery of a low toxicity O-GlcNAc transferase (OGT) inhibitor by structure-based virtual screening of natural products**. *Sci Rep* 2017, **7**:1-11 http://dx.doi.org/10.1038/s41598-017-12522-0.

36. Martin SESS, Tan Z-WW, Itkonen HM, Duveau DY, Paulo JA, Janetzko J, Boutz PL, Törk L, Moss FA, Thomas CJ *et al.*: **Structure-based evolution of low nanomolar O-GlcNAc transferase inhibitors**. *J Am Chem Soc* 2018, **140**:13542-13545 http://dx.doi.org/10.1021/jacs.8b07328.

37. Borodkin VS, Schimpl M, Gundogdu M, Rafie K, Dorfmueller HC, Robinson DA, Van Aalten DMF: **Bisubstrate UDP-peptide conjugates as human O-GlcNAc transferase inhibitors**. *Biochem J* 2014, **457**:497-502 http://dx.doi.org/10.1042/BJ20131272.

38. Rafie K, Gorelik A, Trapannone R, Borodkin VS, Van Aalten DMF: **Thio-linked UDP-peptide conjugates as O-GlcNAc transferase inhibitors**. *Bioconjug Chem* 2018, **29**:1834-1840 http://dx.doi.org/10.1021/acs.bioconjchem.8b00194.

39. Zhang H, Tomašič T, Shi J, Weiss M, Ruijtenbeek R, Anderluh M, Pieters RJ: **Inhibition of: O-GlcNAc transferase (OGT) by peptidic hybrids**. *Medchemcomm* 2018, **9**:883-887 http://dx.doi.org/10.1039/c8md00115d.

40. Elsen NL, Patel SB, Ford RE, Hall DL, Hess F, Kandula H, Kornienko M, Reid J, Selnick H, Shipman JM *et al.*: **Insights into activity and inhibition from the crystal structure of human O-GlcNAcase**. *Nat Chem Biol* 2017, **13**:613-615 http://dx.doi.org/10.1038/nchembio.2357.

41. Li B, Li H, Lu L, Jiang J: **Structures of human O-GlcNAcase and its complexes reveal a new substrate recognition mode**. *Nat Struct Mol Biol* 2017, **24**:362-369 http://dx.doi.org/10.1038/nsmb.3390.

42. Roth C, Chan S, Offen WA, Hemsworth GR, Willems LI, King DT, Varghese V, Britton R, Vocadlo DJ, Davies GJ: **Structural and**

functional insight into human O-GlcNAcase. *Nat Chem Biol* 2017, **13**:610-612 http://dx.doi.org/10.1038/nchembio.2358.

43. Yuzwa SA, Macauley MS, Heinonen JE, Shan X, Dennis RJ, He Y, Whitworth GE, Stubbs KA, McEachern EJ, Davies GJ *et al.*: **A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo**. *Nat Chem Biol* 2008, **4**:483-490 http://dx.doi.org/10.1038/nchembio.96.

44. • Li B, Li H, Hu C-WW, Jiang J: **Structural insights into the substrate binding adaptability and specificity of human O-GlcNAcase**. *Nat Commun* 2017, **8**:1-5 http://dx.doi.org/10.1038/s41467-017-00865-1.
By using four different O-GlcNAcylated substrate peptides during OGA crystallization, the authors showed that substrates for deglycosylation can be bound bi-directionally and while the backbones remain in a conserved conformation, the side chains residues can be stabilized in multiple modes.

45. Ohnishi Y, Ichikawa M, Ichikawa Y: **Facile synthesis of N-Fmoc-serine-S-GlcNAc: a potential molecular probe for the functional study of O-GlcNAc**. *Bioorg Med Chem Lett* 2000, **10**:1289-1291 http://dx.doi.org/10.1016/s0960-894x(00)00223-7.

46. Tarrant MK, Rho H-SS, Xie Z, Jiang YL, Gross C, Culhane JC, Yan G, Qian J, Ichikawa Y, Matsuoka T *et al.*: **Regulation of CK2 by phosphorylation and O-GlcNAcylation revealed by semisynthesis**. *Nat Chem Biol* 2012, **8**:262-269 http://dx.doi.org/10.1038/nchembio.771.

47. Maynard JC, Burlingame AL, Medzihradszky KF: **Cysteine S-linked N-acetylglucosamine (S-GlcNAcylation), a new post-translational modification in mammals**. *Mol Cell Proteomics* 2016, **15**:3405-3411 http://dx.doi.org/10.1074/mcp.M116.061549.

48. De Leon CA, Levine PM, Craven TW, Pratt MR: **The sulfur-linked analogue of O-GlcNAc (S-GlcNAc) is an enzymatically stable and reasonable structural surrogate for O-GlcNAc at the peptide and protein levels**. *Biochemistry* 2017, **56**:3507-3517 http://dx.doi.org/10.1021/acs.biochem.7b00268.

49. •• Gorelik A, Bartual SG, Borodkin VS, Varghese J, Ferenbach AT, van Aalten DMF: **Genetic recoding to dissect the roles of site-specific protein O-GlcNAcylation**. *Nat Struct Mol Biol* 2019, **26**:1071-1077 http://dx.doi.org/10.1038/s41594-019-0325-8.
This manuscript demonstrated that the genetic mutation of endogenous serine GlcNAc acceptor to cysteine can result in hydrolytic stabilization of the modification *in vivo* in the context of an OGA modification site. It further utilized this genetic introduction of S-GlcNAcylation to enable site-specific studies of the consequences of the modification within a living system.

50. Shi J, Tomašič T, Sharif S, Brouwer AJ, Anderluh M, Ruijtenbeek R, Pieters RJ: **Peptide microarray analysis of the cross-talk between O-GlcNAcylation and tyrosine phosphorylation**. *FEBS Lett* 2017, **591**:1872-1883 http://dx.doi.org/10.1002/1873-3468.12708.

51. Sharif S, Shi J, Ruijtenbeek R, Pieters RJ: **Study of cross talk between phosphatases and OGA on a ZO-3-derived peptide**. *Amino Acids* 2019, **51**:739-743 http://dx.doi.org/10.1007/s00726-019-02699-1.

52. Leney AC, El Atmioui D, Wu W, Ovaa H, Heck AJR: **Elucidating crosstalk mechanisms between phosphorylation and O-GlcNAcylation**. *Proc Natl Acad Sci U S A* 2017, **114**:E7255-E7261 http://dx.doi.org/10.1073/pnas.1620529114.

53. Raj R, Lercher L, Mohammed S, Davis BG: **Synthetic nucleosomes reveal that GlcNAcylation modulates direct interaction with the FACT complex**. *Angew Chem Int Ed Engl* 2016, **55**:8918-8922 http://dx.doi.org/10.1002/anie.201603106.

54. Toleman CA, Schumacher MA, Yu S-H, Zeng W, Cox NJ, Smith TJ, Soderblom EJ, Wands AM, Kohler JJ, Boyce M: **Structural basis of O-GlcNAc recognition by mammalian 14-3-3 proteins**. *Proc Natl Acad Sci U S A* 2018, **115**:5956-5961 http://dx.doi.org/10.1073/pnas.1722437115.

55. Chuh KN, Batt AR, Zaro BW, Darabedian N, Marotta NP, Brennan CK, Amirhekmat A, Pratt MR: **The new chemical reporter 6-Alkynyl-6-Deoxy-GlcNAc reveals O-GlcNAc modification of the apoptotic caspases that can block the cleavage/activation of caspase-8**. *J Am Chem Soc* 2017, **139**:7872-7885 http://dx.doi.org/10.1021/jacs.7b02213.

56. Levine PM, De Leon CA, Galesic A, Balana A, Marotta NP, Lewis YE, Pratt MR: **O-GlcNAc modification inhibits the calpain-mediated cleavage of α-synuclein**. *Bioorg Med Chem* 2017, **25**:4977-4982 http://dx.doi.org/10.1016/j.bmc.2017.04.038.

57. Frenkel-Pinter M, Richman M, Belostozky A, Abu-Mokh A, Gazit E, Rahimipour S, Segal D: **Distinct effects of O-GlcNAcylation and phosphorylation of a tau-derived amyloid peptide on aggregation of the native peptide**. *Chem – A Eur J* 2018, **24**:14039-14043 http://dx.doi.org/10.1002/chem.201802209.

58. Schwagerus S, Reimann O, Despres C, Smet-Nocca C, Hackenberger CPR: **Semi-Synthesis of a Tag-FreeO-GlcNAcylated tau protein by sequential chemoselective ligation**. *J Pept Sci* 2016, **22**:327-333 http://dx.doi.org/10.1002/psc.2870.

59. Marotta NP, Lin YH, Lewis YE, Ambroso MR, Zaro BW, Roth MT, Arnold DB, Langen R, Pratt MR: **O-GlcNAc modification blocks the aggregation and toxicity of the protein α-Synuclein associated with Parkinson's disease**. *Nat Chem* 2015, **7**:913-920 http://dx.doi.org/10.1038/nchem.2361.

60. Lewis YE, Galesic A, Levine PM, De Leon CA, Lamiri N, Brennan CK, Pratt MR: **O-GlcNAcylation of α-Synuclein at Serine 87 reduces aggregation without affecting membrane binding**. *ACS Chem Biol* 2017, **12**:1020-1027 http://dx.doi.org/10.1021/acschembio.7b00113.

61. •• Levine PM, Galesic A, Balana AT, Mahul-Mellier A-LL, Navarro MX, De Leon CA, Lashuel HA, Pratt MR: **α-Synuclein O-GlcNAcylation alters aggregation and toxicity, revealing certain residues as potential inhibitors of Parkinson's disease**. *Proc Natl Acad Sci U S A* 2019, **116**:1511-1519 http://dx.doi.org/10.1073/pnas.1808845116.
The authors employed an EPL/SPPS/desulfurization strategy to generate native, site-specifically glycosylated α-synuclein (gT72, gT75, gT81, gT87, and 3gT72/75/81). These semisynthetic proteins demonstrated the modification's multi-faceted, site-dependent, anti-neurodegenerative phenotypes in the context of full-length proteins.

62. •• Balana AT, Levine PM, Mukherjee S, Pedowitz NJ, Moon SP, Takahashi TT, Becker CFW, Pratt MR: **O-GlcNAcylation of small heat shock proteins enhances their anti-amyloid chaperone activity**. *bioRxiv* 2019 http://dx.doi.org/10.1101/869909. 869909.
This work utilized EPL and SPPS to generate a number of site-specifically O-GlcNAcylated sHsps to demonstrate that, in this context, the modification is anti-aggregatory via an improvement in chaperone activity and that this increased chaperone activity is due to a glycan-induced disruption of IXI-chaperone cleft binding, a critical auto-regulatory interaction. Notably, they demonstrate that the effect of O-GlcNAc towards altering HSP27's activity is substoichiometric.

63. Ma J, Hart GW: **O-GlcNAc profiling: from proteins to proteomes**. *Clin Proteomics* 2014, **11**:1-16 http://dx.doi.org/10.1186/1559-0275-11-8.

64. Tashima Y, Stanley P: **Antibodies that detect O-Linked β-D-N-acetylglucosamine on the extracellular domain of cell surface glycoproteins**. *J Biol Chem* 2014, **289**:11132-11142 http://dx.doi.org/10.1074/jbc.M113.492512.

65. Snow CM, Senior A, Gerace L: **Monoclonal antibodies identify a group of nuclear pore complex glycoproteins**. *J Cell Biol* 1987, **104**:1143-1156 http://dx.doi.org/10.1083/jcb.104.5.1143.

66. Comer FI, Vosseller K, Wells L, Accavitti MA, Hart GW: **Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine**. *Anal Biochem* 2001, **293**:169-177 http://dx.doi.org/10.1006/abio.2001.5132.

67. Teo CF, Ingale S, Wolfert MA, Elsayed GA, Nöt LG, Chatham JC, Wells L, Boons G-JJ: **Glycopeptide-specific monoclonal antibodies suggest new roles for O-GlcNAc**. *Nat Chem Biol* 2010, **6**:338-343 http://dx.doi.org/10.1038/nchembio.338.

68. Gorelik A, van Aalten DMF: **Tools for functional dissection of site-specific O-GlcNAcylation**. *RSC Chem Biol* 2020, **1**:98-109 http://dx.doi.org/10.1039/d0cb00052c.

69. • Darabedian N, Thompson JW, Chuh KN, Hsieh-Wilson LC, Pratt MR: **Optimization of chemoenzymatic mass tagging by**

strain-promoted cycloaddition (SPAAC) for the determination of O-GlcNAc stoichiometry by western blotting. *Biochemistry* 2018, **57**:5769-5774 http://dx.doi.org/10.1021/acs.biochem.8b00648.

The authors prepared a ubiquitin protein C-terminally tagged with an O-GlcNAcylated peptide through expressed protein ligation. This semi-synthetic protein was then used as a 'standard' for benchmarking and optimization of the mass-tagging protocol. The work is a good demonstration of using synthetic glycopeptides for method/technique development.

70. • Levine PM, Balana AT, Sturchler E, Koole C, Noda H, Zarzycka B, Daley EJ, Truong TT, Katritch V, Gardella TJ *et al.*: **O-GlcNAc engineering of GPCR peptide-agonists improves their stability and in vivo activity**. *J Am Chem Soc* 2019, **141**:14210-14219 http://dx.doi.org/10.1021/jacs.9b05365.

This work proposed peptide O-GlcNAcylation as an artificial strategy to potentially improve stability and therapeutic properties of peptide hormones GLP-1 and PTH.

71. Amso Z, Bisset SW, Yang S-HH, Harris PWRR, Wright TH, Navo CD, Patchett ML, Norris GE, Brimble MA: **Total chemical synthesis of Glycocin F and analogues: S-glycosylation confers improved antimicrobial activity**. *Chem Sci* 2018, **9**:1686-1691 http://dx.doi.org/10.1039/c7sc04383j.