JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau

Perspective

# Methods for Studying Site-Specific O-GlcNAc Modifications: Successes, Limitations, and Important Future Goals

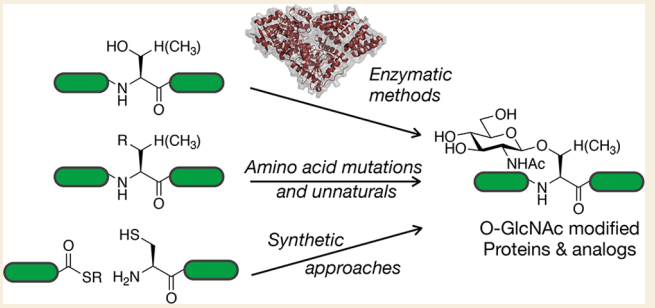
Stuart P. Moon, Afraah Javed,[#] Eldon R. Hard,[#] and Matthew R. Pratt*

Cite This: *JACS Au* 2022, 2, 74–83

Read Online

ACCESS | Metrics & More | Article Recommendations

**ABSTRACT:** O-GlcNAcylation is a dynamic post-translational modification which affects myriad proteins, cellular functions, and disease states. Its presence or absence modulates protein function via differential protein- and site-specific mechanisms, necessitating innovative techniques to probe the modification in highly selective manners. To this end, a variety of biological and chemical methods have been developed to study specific O-GlcNAc modification events both *in vitro* and *in vivo*, each with their own respective strengths and shortcomings. Together, they comprise a potent chemical biology toolbox for the analysis of O-GlcNAcylation (and, in theory, other post-translational modifications) while highlighting the need and space for more facile, generalizable, and biologically authentic techniques.




**KEYWORDS:** *Post-translational modifications, O-GlcNAc, chemical tools, protein modification, protein synthesis*

## INTRODUCTION

O-GlcNAc modification is an enzymatic post-translational modification (PTM) essential in mammals and insects[1−3] whereby intracellular proteins are functionalized by single monomers of *N*-acetylglucosamine at serine and threonine side chain hydroxyls (Figure 1). Unique from other forms of glycosylation, this GlcNAc monomer is not extended further into large polysaccharide chains. Instead, this PTM is highly analogous to phosphorylation in that it is reversible and dynamic via its "writer", O-GlcNAc transferase (OGT),[4] and its "eraser", O-GlcNAc hydrolase (OGA),[5] the activities of which are linked to the metabolic and disease state of the cell. The dynamic interplay of these enzymes lends O-GlcNAcylation important roles in cell signaling pathways,[6] cell fate determination,[7,8] transcription,[9] immunity,[10] and response to cellular stressors.[11] The modification has been shown to be dysregulated in many forms of cancer,[9,12] as well as in neurodegenerative,[13,14] metabolic,[15,16] and cardiovascular diseases,[17,18] stressing the importance of its homeostasis and of its study.

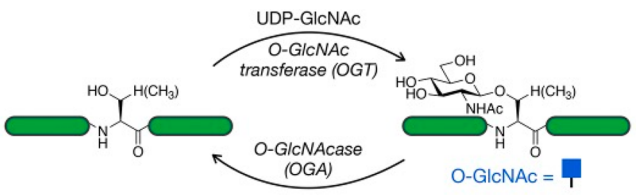



**Figure 1.** O-GlcNAc modification is the reversible addition of N-acetylglucosamine to serine/threonine residues of intracellular proteins. It is added by the enzyme O-GlcNAc transferase (OGT) and removed by O-GlcNAcase (OGA).

The mechanisms by which the modification impacts its substrates are highly multifaceted. Because OGT and kinases can compete for Ser and Thr residues, some effects of the modification have been attributed simply to the inhibition of phosphorylation,[19] while the PTM also affects other proteins by more direct means, primarily by dictating protein−protein interactions. The O-GlcNAc moiety has been dubbed a "grease and glue", either dampening or enhancing binding of its substrates to their interactors by its highly hydrophilic, uncharged steric bulk.[20] Its consequences are therefore differential both from substrate-to-substrate, but also from site-to-site within a given substrate, complicating their interrogation.

Because O-GlcNAcylation's effects are not generalizable across its substrates or substrate sites, several biological and chemical tools of varying specificity have been established for the study of the modification *in vitro* and *in vivo*. This review seeks to compile current methods used by our lab and others

Received: October 13, 2021
Published: December 15, 2021





https://doi.org/10.1021/jacsau.1c00455
*JACS Au* 2022, 2, 74–83

for the installation and examination of O-GlcNAc modification. We highlight their respective successes, limitations, and physiological relevance, as well as discuss the need for more simple, effective, and general chemical biology tools to more fully understand this PTM.

## O-GLCNAC MODULATION: BIOLOGICAL METHODS

### OGT Coexpression and *in Vitro* Substrate Modification

Studies involving O-GlcNAc-modified proteins often require copious amounts of O-GlcNAcylated material. To this end, coincubation of recombinant OGT with a protein of interest *in vitro* or *in vivo* coexpression of OGT and the desired substrate protein in *E. coli* allows for the convenient and efficient production of milligram-scale quantities of O-GlcNAc modified protein. As the gene for OGT is only found in eukaryotes, it must be introduced to the bacterial genome via a plasmid. Conveniently, the substrate for OGT's GlcNAc transferase activity, UDP-GlcNAc, is endogenously produced by bacteria for use as a cell wall building-block.[21]

Recombinant OGT expression was first explored in terms of substrate protein recognition in 2008.[22] After expressing three isoforms of OGT in *E. coli*, bacterial lysates were immunoblotted with RL2, an anti-O-GlcNAc antibody. Although OGT was expressed and catalytically active, no significant O-GlcNAcylation of bacterial proteins was detected. This is unsurprising because the lack of OGT in bacterial genomes suggests that O-GlcNAc modification is uncommon and unimportant for bacterial protein regulation. However, OGT successfully O-GlcNAc modified recombinantly coexpressed mammalian proteins known to be OGT substrates, establishing that bacteria produce enough UDP-GlcNAc for endogenous purposes as well as protein O-GlcNAcylation. The technique can be further optimized by coexpressing GlmM and GlmU, enzymes that promote UDP-GlcNAc synthesis to increase coexpressed OGT activity.[23]

This technique and *in vitro* glycosylation have been used to map modification sites of tau,[24] SIRT1,[25] emerin,[26] and lamin A,[27] and others. Coexpression of OGT and PKA kinases PKAc$\alpha$ and PKAc$\beta$ showed that the modification of these proteins enhances their kinase activity toward tau.[28] Further, OGT coexpression in yeast was used to study the modification's crosstalk with phosphorylation of SKN-1.[29] The major issue with this technique is that the resulting protein species often modified at multiple sites, preventing the straightforward examination of site-specific effects. Furthermore, this method does not yield 100% GlcNAc modified protein, but the O-GlcNAcylated protein can sometimes be isolated from the unmodified protein by high performance liquid chromatography (HPLC). Additionally, the *E. coli* protein NagZ, which endogenously hydrolyzes O-GlcNAc linkages to aid in peptidoglycan recycling, has also shown deglycosylation activity toward exogenous, O-GlcNAcylated proteins, decreasing the yields of modified POIs.[30] This issue can be avoided by inhibiting NagZ with small molecule OGA inhibitors such as PUGNAc or by genetic knockout of the problematic enzyme.

### Glycosite-to-Alanine/Valine Mutagenesis

While phosphomimetic (S-to-D) point mutations have been used to study phosphorylation, there are no direct amino acid substitutions that faithfully represent O-GlcNAcylated protein residues. Instead, loss-of-function serine-to-alanine/or threonine-to-valine mutations have been used to study the modification's absence site-specifically (Figure 2). Because

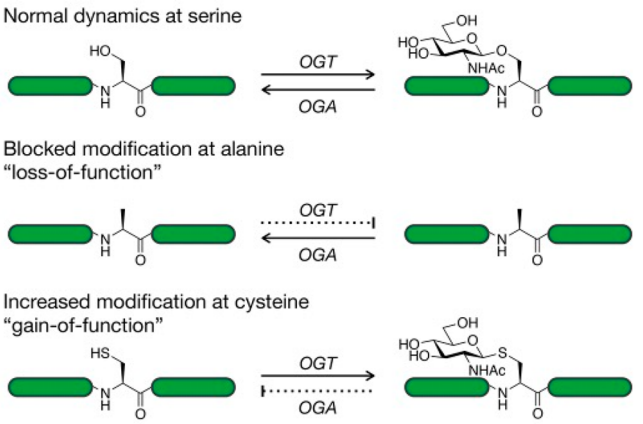



**Figure 2.** Standard mutagenesis methods to block or increase O-GlcNAcylation. The normal dynamics of O-GlcNAc modification can be blocked by mutation to an alanine resulting in a loss of function mutant. Alternatively, increased modification can be obtained by a cysteine mutant that inhibits OGA activity giving a gain of function mutant.

the PTM occurs on the hydroxyl groups of serine and threonine residues, mutations that replace the side chain hydroxyl group with a methyl group prevent endogenous O-GlcNAc modification while preserving the size of the original amino acids. This loss-of-function mutation enables one to observe the effects of "knocking out" a site-specific O-GlcNAc modification.

This mutation strategy has been used in conjunction with mass spectrometry to map the specific sites at which proteins are O-GlcNAc modified. In a study by Kim et al., this technique was used to probe the PTM's effect on the proteasomal clearance of SMAD4.[31] Through mass spectrometry analysis, four residues of the protein were found to be O-GlcNAc modification sites, and these sites were mutated to alanines or valines using site-directed mutagenesis. Subsequent Western blotting analysis validated that each of the mutants were less O-GlcNAc-modified than the wild-type, and a quadruple mutation abolished almost all O-GlcNAcylation, validating that all four sites are O-GlcNAcylated. The technique was also used for comparison between multiple O-GlcNAc modification sites within the same protein. A study of ubiquitination levels of the four single-point mutants of SMAD4 elucidated that modification at Thr63 is the most important of the four for SMAD4 stabilization. Such site-specific differences highlight the value of molecular level investigations of O-GlcNAc modification. In addition, the use of mass spectrometry as a complementary technique can highlight differences that be otherwise overlooked. A study by Ma et al. showed that the loss of a highly conserved O-GlcNAc site (T305) on NF-$\kappa$B was compensated by increased O-GlcNAcylation at a different site (T352).[32]

The design of such loss-of-function mutants have highlighted important roles for O-GlcNAc in the induction of apoptosis,[33] modulation of transcription,[34] and inhibition of cancer cell growth.[31] Nevertheless, the structural and functional effects of replacing a polar side chain with a small, nonpolar side chain on protein structure must be considered. In addition, the mutagenesis results in the loss of a


https://doi.org/10.1021/jacsau.1c00455
JACS Au 2022, 2, 74−83

phosphorylation site, which is especially pertinent at certain sites considering the interplay between phosphorylation and O-GlcNAc modification. Others maintain that mutagenesis studies do not unambiguously probe site-specific deglycosylation, arguing that the consequences of the mutation may instead be artifacts of a reduction in OGT activity toward the substrate protein as a whole.[35]

## Glycosite-to-Cysteine Mutagenesis

In addition to the above loss-of-function mutation, O-GlcNAcylation has also been studied by using S-to-C gain-of-function mutations (Figure 2). It has been shown that OGT is capable of modifying cysteine residues, yielding S-GlcNAcylation that cannot be removed by OGA.[36] Although the biological relevance of this PTM is unknown, its hydrolytic stability and compatibility with OGT's promiscuity make it a promising analogue for studying the effects of site-specific, permanent O-GlcNAcylation.[37]

Our lab has synthetically incorporated S-GlcNAcylation onto $\alpha$-synuclein *in vitro* to show that the modification is resistant to OGA hydrolysis and that its effects are similar to O-GlcNAcylation at the same site.[37] Further, Withers and co-workers engineered a thioglycoligase that could directly attach GlcNAc moieties to cysteines through mutation of a catalytically active residue of OGT.[38] This enzymatic approach utilizes a commercially available glycosyl donor called *p*NP-GlcNAc, which makes the technique practical and convenient. In this work, the researchers generated S-GlcNAcylated tau protein as a proof-of-concept.

Recently, the van Aalten group developed a CRISPR-Cas9-based genetic method for the incorporation of S-GlcNAc in the place of O-GlcNAc on proteins in living cells.[39] This approach involves the mutation of known O-GlcNAc sites to cysteines that can be S-GlcNAcylated. S-GlcNAcylation can be detected by some site-specific and pan-selective O-GlcNAc antibodies. After confirming the S-GlcNAc transferase activity of OGT on various substrate sequences *in vitro*, they extended this approach to living cells. CRISPR-Cas9 was used to introduce a S405C mutation at the site of O-GlcNAc modification on OGA in mouse embryonic stem cells. O(S)-GlcNAcylated proteins in the lysate were chemoenzymatically labeled with UDP-GalNAz and mass shifted with DBCO-PEG-5K prior to SDS-PAGE, allowing for the discrimination of unmodified and S-GlcNAcylated OGA. Impressively, there was successful incorporation of S-GlcNAc with a stoichiometry of at least 70%. This method has the potential to allow for the study of O-GlcNAc at specific sites on proteins in living systems by artificially increasing their stoichiometry. However, in a fluorescence polarimetry competition assay, S-GlcNAcylated peptides were found to bind to OGA with affinities decreased by 2 orders of magnitude when compared to their O-GlcNAcylated counterparts. This loss of recognition by OGA suggests that the S-to-O mutation could impact the modification's recognition by other interactors, complicating conclusions drawn by using this technique.

## Substrate-Targeted OGT/OGA

While global O-GlcNAc modification levels can be altered using chemical inhibitors or modulation of OGT/OGA expression, it is more challenging to selectively install O-GlcNAc onto target proteins in cells. Such a method would be useful for investigating the biological effects of increased or decreased O-GlcNAcylation on specific glycoproteins of interest.

Research efforts in this domain have been spearheaded by the Woo group, who first reported an engineered OGT that was targeted to specific proteins through the use of nanobodies, small antigen-binding fragments of heavy chain only antibodies which are naturally produced by camelids and sharks (Figure 3a).[40,41] Truncated OGT was fused to a

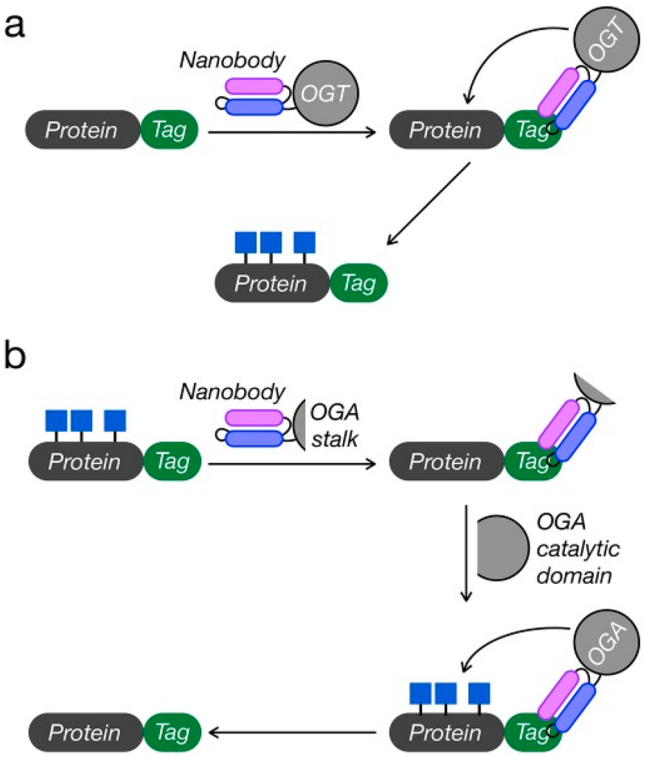



**Figure 3.** Substrate targeted OGT/OGA. (a) OGT catalytic domain can be targeted to proteins through a nanobody/tag pair, resulting in increased O-GlcNAc of the tagged, target protein. (b) Similarly, OGA can be nanobody targeted to remove O-GlcNAc. Splitting of the OGA stalk and catalytic domains prevents off-target activity.

nanobody (nanobody-OGT) with high affinity for a protein tag on target proteins JunB and Nup62 to induce O-GlcNAcylation. Levels of O-GlcNAcylated protein of interest were quantified by mass spectrometry in both wild-type OGT and nanobody-OGT overexpressing cells, and the nanobody-OGT fusions were found to have high selectivity. Nanobody-OGT also produced a similar glycosylation profile to that of wild-type OGT. In addition, this technique was demonstrated on endogenous $\alpha$-synuclein. Using a nanobody that was developed to target $\alpha$-synuclein, nanobody-OGT was expressed in HEK293T cells. O-GlcNAcylated proteins in the cell lysate then underwent mass-shifting to determine O-GlcNAc-modified protein stoichiometry.

Following the development of the proximity-directed OGT, the Woo group also generated nanobody-fused split O-GlcNAcase (Figure 3b).[42] Initial tests with the catalytic and stalk domains of OGA fused to a nanobody resulted in deglycosylation activity with poor target-protein selectivity. To optimize selectivity, pairs of truncated N- and C-terminal OGA fragments were screened for reduced enzymatic activity that could be restored by the addition of a nanobody. This approach can be customized to various target proteins and tags by various nanobodies without the perturbation of global O-GlcNAc levels.[43]


https://doi.org/10.1021/jacsau.1c00455
JACS Au 2022, 2, 74−83

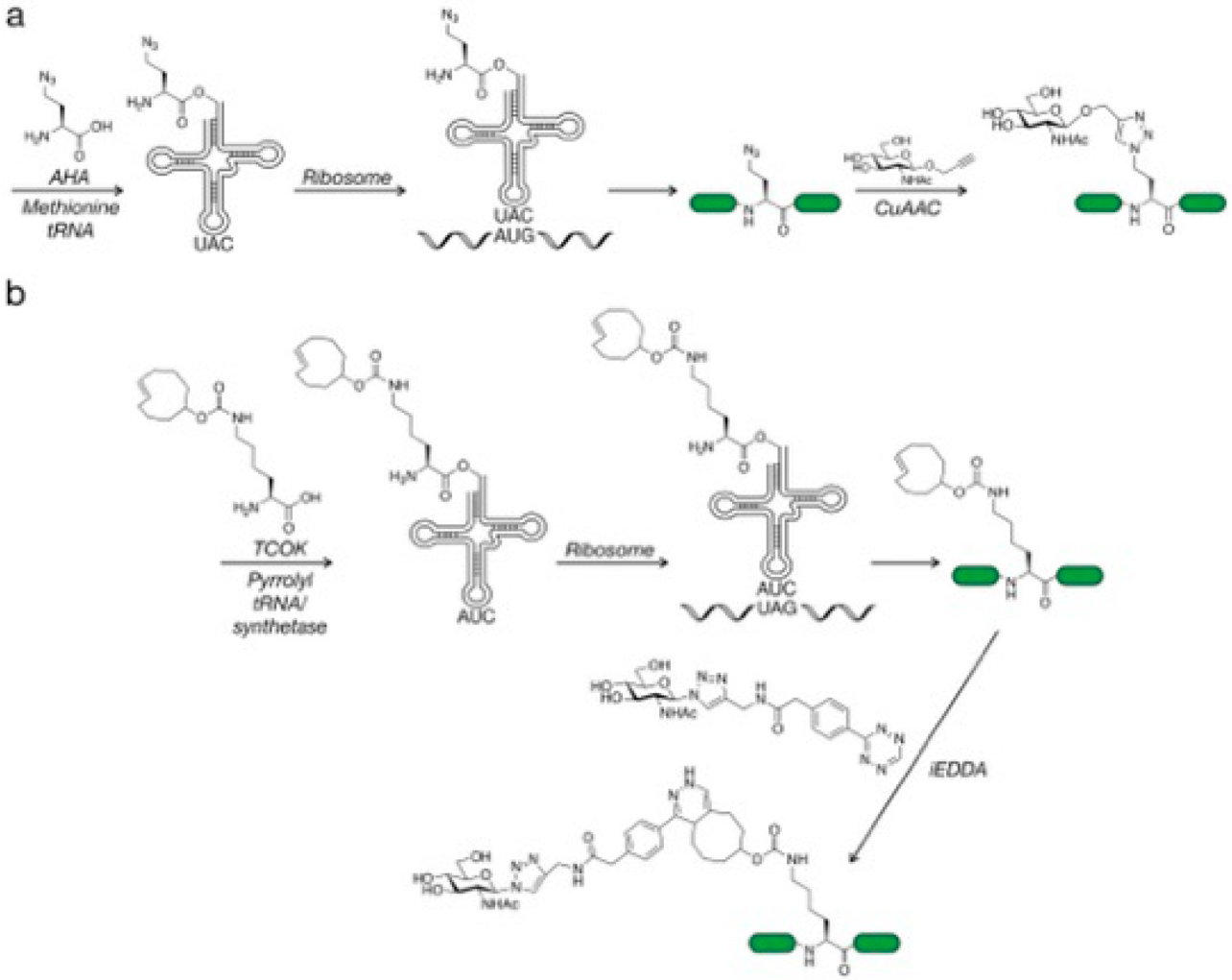



**Figure 4.** Incorporation of O-GlcNAc analogs using unnatural amino acids. (a) The promiscuity of natural methionine tRNA/synthetases can be exploited to incorporate azidohomoalanine (AHA) at methionine codons. Subsequent CuAAC chemistry can be used to bioorthogonally install an O-GlcNAc analog. (b) The pyrrolysine system can be used to site-specifically incorporate bioorthogonal amino acids at amber stop codons. Again, bioorthogonal chemistries can be used to install O-GlcNAc analogs.

Both substrate-targeted OGT and OGA enable characterization of the function of O-GlcNAc on specific proteins. Still, nanobody generation is a lengthy process, and it remains to be seen if this technique suffers from off-target effects. In the future, the discovery of more nanobodies that target endogenous proteins may facilitate the function of nanobody-fused split OGA without the use of tags that must be genetically encoded. Additionally, while this technique is protein-specific, it is incapable of controlling the sites (un)modified by the nanobody-fused enzymes.

## Genetic Code Expansion

Genetic code expansion (GCE) can be used to introduce site-specific PTMs both directly through the incorporation of premodified amino acids and indirectly by adding biorthogonal handles as scaffolds for the PTM.[44] These methods provide complete proteins that are stoichiometrically site-specifically functionalized and can be performed in living cells given the use of appropriate reagents. One approach to GCE involves the incorporation of unnatural amino acids through the natural promiscuity of endogenous tRNA/synthetase pairs. For example, the native methionine tRNA/synthetase in *E. coli* will accept azidohomoalanine (AHA) and homopropargyl glycine (HPG).[45] In methionine auxotrophic cells, the corresponding tRNAs will build up to sufficient concentrations that the ribosome will insert these unnatural amino acids into proteins at AUG codons (Figure 4a). Subsequent reaction of these azide- or alkyne-containing proteins under copper-catalyzed azide−alkyne cycloaddition (CuAAC) conditions has been used to install O-GlcNAc analogs.[46] Another common form of GCE takes advantage of orthogonal unnatural amino acid and tRNA synthetase pairs, as well as an amber stop codon at the site of interest within the POI's mRNA. This "amber suppression" mutagenesis enables the site-selective introduction of unnatural amino acids. In the case of O-GlcNAc, the pyrrolyl-tRNA/synthetase pair was used to incorporate bicyclo[6.1.0]non-4-yn-9-ylmethanol lysine (BCNK) or *trans*-cyclooctene-derivatized lysine (TCOK) into GFP.[47] These handles were then reacted with tetrazines through the inverse electron demand Diels−Alder (iEDDA) reaction to attach O-GlcNAc analogs (Figure 4b). The major disadvantage to these methods is the presence of the linker between the GlcNAc and the amino acid, which, under certain circumstances, can be larger than the O-GlcNAc itself.


https://doi.org/10.1021/jacsau.1c00455
JACS Au 2022, 2, 74−83

To date, there have been no successful incorporations of synthetic, O-GlcNAc-modified amino acids on proteins using genetic code expansion. An initial report demonstrating the incorporation of O-GlcNAc-Thr into myoglobin using amber suppression was later retracted.[48] O-GlcNAc-Ser is metabolized by *E. coli* for carbon, so it is not available in the cytoplasm in the time frame required for the amber stop codon suppression system. Acetylation of the sugar's free hydroxyls, a common mechanism to increase cellular uptake of polar molecules, is ineffective because *E. coli* do not endogenously express the deacetylases required to remove the protecting group.

## O-GLCNAC MODULATION: CHEMICAL METHODS

### SPPS and Protein Ligation

Site-specifically O(S)-GlcNAcylated peptides can be readily prepared using through solid-phase peptide synthesis (SPPS) with Fmoc-Ser($\beta$-Ac$_3$GlcNAc)-OH, Fmoc-Thr($\beta$-Ac$_3$GlcNAc)-OH, and Fmoc-Cys($\beta$-Ac$_3$GlcNAc)-OH building blocks. These modified amino acids are commercially available, yet costly, but there are several methods for producing them in-house, with the most common being the König-Knorr reaction. The König-Knorr reaction takes advantage of either halophilic activation using the heavy metal salt $HgBr_2$[49] or Lewis acid-activation using AgOTf,[50] TMSOTf,[51] or $InBr_3$.[52] An alternative strategy to glycosylation of amino acids is through thioglycoside activation.[51,53] These GlcNAcylated residues can be preactivated as pentafluorophenyl (PFP) esters for direct use in glycopeptide syntheses.[49] Both methods allow for the $\beta$-specific linkage of the O-GlcNAc sugar to the amino acid, and thus a peptide containing a native, site-specific O-GlcNAc moiety.

Beyond producing glycosylated peptides, it is possible to incorporate these peptides into full-length, native proteins with specifically modified glycosylated residues through native chemical ligation (NCL) and variants of the NCL method (Figure 5a). NCL is a chemoselective technique which involves the coupling of a C-terminal thioester to an N-terminal cysteine residue under mild aqueous conditions (pH 7−7.5) with high yields. This reaction generates a thioester-linked intermediate that rearranges spontaneously to form a peptide containing a native peptide bond to a cysteine residue through an S−N acyl shift. This method was first introduced in 1994 by Kent and colleagues.[54−56] NCL has numerous applications in the field of chemical biology, as it is used to synthesize native polypeptides with site-specifically modified residues. Complicating the applicability of NCL are limitations in the size of peptides produced by SPPS. Expressed Protein Ligation (EPL), an extension of NCL, can be used to overcome these issues by taking advantage of bacterial intein splicing mechanisms to express protein fragments that can be used in protein semisyntheses (Figure 5a).[57] To express protein thioesters, the fragment of interest can be genetically fused to a bacterial intein mutated such that the splicing mechanism is impeded at an intermediate stage. This results in a branched, thioester linkage between the fragment of interest and the intein which can be readily exchanged with exogenous thiols to yield stable, recombinant protein thioesters.[58,59] EPL enables the use of recombinant protein fragments in NCL, broadening the limits of total polypeptide size and the modifications introduced.[60,61]

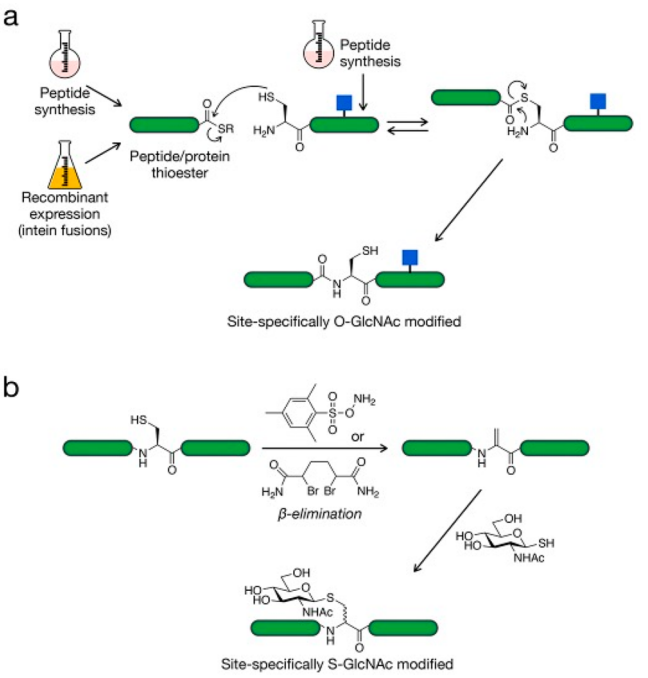



**Figure 5.** Site-specific incorporation of O-GlcNAc using protein ligation or post-translational mutagenesis. (a) Native/expressed protein ligation involves the selective reaction between protein thioesters and N-terminal cysteines to generate native amide bonds. (b) Cysteine residues can be forced to undergo $\beta$-elimination followed by addition of S-GlcNAc nucelophiles.

While it is relatively straightforward to design a synthetic scheme to access some protein targets, the sequences of many proteins and the positions of their modification sites must be amenable to the conditions of SPPS and NCL/EPL. The technique introduces relatively rare cysteine residues into completed sequences; however, for proteins that do not contain native cysteine residues, one can take advantage of metal- or nonmetal-based desulfurization methods to convert resulting cysteines to alanine residues.[62−64] To protect native cysteines that would be lost during broad deprotections, selenocysteine can be used as the NCL nucleophile and can be selectively deselenized in the presence of cysteine.[65,66] Additionally, the use of synthetic, thiolated/selenized amino acid analogs allows noncysteine or alanine ligation sites.[67] Further, hydrazines[68,69] and protected cysteine derivatives[70] can be used to mask ligation sites at N- and C-termini, respectively, before being activated to form reactive thioesters and free cysteines.

Using SPPS and EPL/NCL, one can produce native proteins with site-specifically O-GlcNAc-modified residues for study. SPPS was used to study the crosstalk between O-GlcNAcylation and Jak2 phosphorylation in using a synthetic peptide microarray.[71] The Pieters group found that a synthetically phosphorylated peptide substrate of OGT and Jak2 was highly resistant to O-GlcNAcylation; however, the same peptide could be phosphorylated by Jak2 when pre-O-GlcNAcylated synthetically. The same group later studied a peptide derived from ZO-3 which is phosphorylated at Tyr364 by Jak2 and O-GlcNAcylated on Ser369 by OGT.[72] It was found that phosphorylation at Tyr364 slightly reduced the removal of O-GlcNAc by OGA, while Ser369 glycosylation slightly enhanced the dephosphorylation of the nearby Tyr by phosphatases. One study used EPL and SPPS to investigate


https://doi.org/10.1021/jacsau.1c00455
JACS Au 2022, 2, 74−83

**Table 1. Comparison of Different Techniques**

| technique | protein selectivity | site selectivity | generalizability | ease of use | physiological relevance |
|---|---|---|---|---|---|
| OGT coexpression | - | - | ++ | +++ | ++ |
| Glycosite-to-A mutagenesis | +++ | +++ | +++ | +++ | + |
| Glycosite-to-C mutagenesis[a] | +++ | +++ | + | +++ | ++ |
| Targeted OGT/OGA | +++ | + | + | ++ | ++ |
| GCE (O-GlcNAc analogs) | +++ | ++ | ++ | + | - |
| Chemical ligation | +++ | +++ | ++ | + | +++ |
| post-translational mutagenesis[b] | +++ | ++ | ++ | + | + |

[a]Generalizability of glycosite-to-C mutagenesis is currently unclear as limited sites have been tested. [b]Site selectivity of post-translational mutagenesis is limited by the availability of a unique cysteine on the protein surface.

PTM crosstalk in kinase CK2, which is O-GlcNAc modified at Ser347 and phosphorylated at Thr334. By installing a metabolically stable S-linked GlcNAc at Ser347, the Cole group was able to determine that O-GlcNAcylation at this site blocks phosphorylation at the adjacent phosphorylation site.[73] The S-to-O mutation could alter the activity of the protein; however, as the atoms are similarly sized, the effect is subtle while the stabilization of the O-GlcNAc moiety is extremely useful. Native chemical ligation was also used in the first instance of semisynthetic tau protein by the Hackenberger group.[74] O-GlcNAc was site-specifically added onto Ser400 in the C-terminus of tau and the native protein was made using EPL and desulfurization. These methods allow for the study of O-GlcNAc's effect on tau protein for *in vitro* structural and functional studies.

Through NCL and EPL, O-GlcNAc's effect was also elucidated in studies of semisynthetic $\alpha$-synuclein by our laboratory.[75−77] It was found that O-GlcNAcylation results in site-specific differences in $\alpha$-synuclein aggregation in *in vitro* experiments and is generally inhibitory. We further showed that the PTM is protective against the protein's cleavage by calpain,[78] and that the anti-aggregation phenotypes imparted by the GlcNAc moiety are unique and not reproduced by other sugars.[79] We also used these techniques to study the effect of O-GlcNAcylation on semisynthetic HMGB1 modified at positions Ser100 and demonstrated the influence O-GlcNAc has on HMGB1-DNA interactions.[80] We showed that the PTM generally enhanced the interactions between the protein and DNA and resulted in error-prone repair of ICL-damaged plasmids in U2OS cell extracts. Finally, we have also studied the effect of O-GlcNAc modifications of small heat shock proteins (sHSPs) using NCL and EPL methods.[81] The results of this study found that O-GlcNAc modification near the IXI motif of semisynthetic sHSPs increases their anti-amyloid chaperone activity, and the O-GlcNAc modification of sHSPs is maintained even in globally reduced O-GlcNAc levels found in those with Alzheimer's disease.

## Post-Translational Mutagenesis

The production of unnatural amino acids through site-directed mutagenesis has been limited by the 20 natural amino acids and their $sp^3$-$sp^3$ C−C bonds. A study has shown that it is possible to form a range of alkyl halides on the side chains of amino acids through dehydroalanine (Dha) (readily formed using mild, carbon-centered free-radical chemistry at Cys residues) and Dha derivatives, allowing for the post-translational production of unnatural amino acids with high chemoselectivity and compatibility in biological systems (Figure 5b).[82,83] With this, the ability to insert side-chain alterations directly and selectively provides an easy route to natural and unnatural PTM incorporation, such as mimics of O-linked glycosides. A followup study used Dha and a thiolated GlcNAc to produce homogeneous histone protein H2A containing an O-GlcNAc mimic at Thr101.[84] The GlcNAcylated H2A was then used to form nucleosomes, and stability studies were carried out to show that O-GlcNAc at this position destabilizes the histone complex. Another study was performed using the same methods to probe the effects of O-GlcNAcylation at Ser112 on H2B in the nucleosome complex.[85] The Davis lab found that glycosylation at the 112 position recruits the FACT complex and aids in transcription elongation. Interestingly, a recent study by the Wang lab developed a novel method compatible *in cellulo*. Genetically Encoded Chemical Conversion (GECCO) takes advantage of a sulfur-fluoride exchange (SuFEx) reaction between a genetically encoded, unnatural fluorosulfate-L-tyrosine and a threonine or serine side chain to generate reactive dehydrobutyrine or dehydroalanine moieties *in situ* inter- or intramolecularly.[86] Thus, the Wang group was successful in attaching an S-GlcNAc monomer onto a Dha site at residue 184 of sfGFP. These reactions could prove to be the next step in post-translational mutagenesis methods, allowing the production of glycoprotein mimetics in living cells.

The downside to post-translational mutagenesis methods is the resulting unnatural linkage to the O-GlcNAc modification. The cysteine thioether linkage is, again, a mimic of the native glycosidic bond, and the homoserine linkage is one carbon longer than would be found in nature. To compound issues, the $\alpha$-carbon is often racemized in these methods, which generally creates an inseparable mixture of diastereomers that can contain differing biochemical properties. To preserve the stereochemistry at the modification site, disulfide GlcNAc-S-Cys linkages can be desulfurized and converted into thioether-linked glycoconjugates through a method using polarized, electron-rich phosphines.[87]

## Conclusion, Limitations, and Future Directions

The catalogued list of O-GlcNAcylated proteins and sites of modification continues to grow with thousands of potential substrates in humans alone (https://www.oglcnac.mcw.edu).[88] However, any effects of most of these modifications are completely unknown. In this perspective, we have described varied approaches to install O-GlcNAc on certain proteins, sometimes in a site-specific manner. Techniques ranging from enzymatic modification of proteins by OGT to chemical synthesis of O-GlcNAc modified substrates have enabled important biochemical roles for O-GlcNAc to be elucidated. These studies have demonstrated that O-GlcNAc can alter protein structure and function in critical and multifaceted ways, and they will certainly be applied to make additional discoveries in the future. Each of these techniques has its own strengths and limitations (Table 1).


https://doi.org/10.1021/jacsau.1c00455
JACS Au 2022, 2, 74−83

In many aspects, the enzymatic modification of proteins by OGT is the technically most simple method for installing O-GlcNAc on proteins of interest. Specifically, many biochemists routinely perform recombinant protein expression in *E. coli*. Therefore, coexpression of a protein of interest with OGT in *E. coli* (or similar heterologous system) is an approach that many laboratories are well-placed to exploit. Unfortunately, as mentioned above, O-GlcNAc is often added to multiple residues on a protein of interest using this system, and these modification sites will typically have different levels of modification stoichiometry. This can result in complex mixtures of different glyco-proteoforms that can be very challenging to isolate from one another. Likewise, transient or even stable expression of proteins in mammalian cells is fairly routine; however, this same heterogeneity issue persists. The issue of site selectivity in both of these systems might be overcome through the mutation of serine/threonine residues to cysteine. As described above, mutation of serine to cysteine on OGA enabled a dramatic increase in O-GlcNAc stoichiometry at this site in mammalian cells.[39] One could envision using this strategy during recombinant expression by mutating a site of interest to cysteine, resulting in high levels of S-GlcNAc at this site, and then removing the remaining O-GlcNAcylation with OGA. However, the generality of this cysteine-mutation strategy for multiple proteins and sites needs to be explored further, and this is an important future area of investigation.

Targeted OGT/OGA approaches for increasing or decreasing O-GlcNAc on a protein of interest in living systems are a potentially powerful tool to complement genetic (e.g., RNAi) and small molecule inhibitor approaches.[89,90] A minor limitation of these systems are their relative complexity, which may limit, but not necessarily prevent, their application beyond easily transfectible cell systems. A more significant potential roadblock is the time and effort that might be needed for the development of nanobodies to target endogenous proteins of interest. Therefore, other targeting modalities should be explored. For example, a recent exciting publication demonstrated that RNA aptamers can be used to target OGT to $\beta$-catenin.[91]

In the case of chemical strategies, we have had significant success applying protein ligation methods (NCL/EPL) for the preparation of completely homogeneous O-GlcNAcylated proteins in biochemical studies.[37,75−81] While protein ligation is the only current method that is theoretically guaranteed to produce pure protein products, it is not without its limitations. Like all synthetic strategies using NCL/EPL, a fundamental issue is the relatively slow rate of the ligation reaction, requiring fairly high (millimolar) peptide/protein concentrations. Additionally, certain protein fragments can suffer from poor expression yields and unpredictable physical properties that make them difficult to purify in sufficient quantities. Therefore, methods that increase the rate of protein ligation, such as a lipid-facilitated protocol published by the Devaraj lab,[92] are still incredibly important for the field and should be explored. Additionally, robust and gentle ligation reactions that occur at noncysteine or -alanine junctions would greatly expand the potential viable synthetic strategies to any given protein target.

Post-translational mutagenesis overcomes some limitations of protein ligation because the full-length protein target can be expressed in full.[82,84,85] However, the selectivity of the chemistry for the installation of S-GlcNAc on one cysteine can be challenging, particularly in protein targets that have other crucial structural or catalytic cysteines. Additionally, while some proteins may not be affected by the racemization of the $\alpha$-carbon at the S-GlcNAcylated site, this will certainly not be true in every case, making the consequences of GlcNAc versus racemization difficult to distinguish. Therefore, any chemistry that could bias the stereochemical outcome of post-translational mutagenesis could be a key advance.

Current successes in genetic codon expansion suffer from even larger perturbations to the O-GlcNAc structure, which may make biochemical results difficult to interpret. For example, we recently demonstrated that even as subtle of a change as O-GlcNAc to O-GalNAc or O-glucose alters the aggregation behavior of the protein $\alpha$-synuclein.[79] While incorporation of bona fide O-GlcNAc or even S-GlcNAc through amber suppression in *E. coli* has been unsuccessful, this should still be a goal of the community. It is possible that different expression systems (e.g., Vmax X2[93]) or other tRNA/synthetase pairs may overcome the hurdles, and the direct incorporation of either serine or threonine O-GlcNAc by genetic codon expansion would be transformative.

In summary, the creation of various methods for the protein/site-selective installation of O-GlcNAc have allowed for a dramatic expansion in our understanding of this PTM. Researchers should continue to take advantage of the methods described above to complement other genetic, pharmacological, chemical biologic, and biochemical techniques. Additionally, we encourage "tool makers" to continue to expand the available approaches for studying O-GlcNAc and to take on major challenges in the field.

## AUTHOR INFORMATION


### Corresponding Author

**Matthew R. Pratt** − *Departments of Chemistry and Biological Sciences, University of Southern California, Los Angeles, California 90089, United States;* orcid.org/0000-0003-3205-5615; Email: matthew.pratt@usc.edu

### Authors

**Stuart P. Moon** − *Departments of Chemistry, University of Southern California, Los Angeles, California 90089, United States*

**Afraah Javed** − *Departments of Chemistry, University of Southern California, Los Angeles, California 90089, United States*

**Eldon R. Hard** − *Departments of Chemistry, University of Southern California, Los Angeles, California 90089, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.1c00455

### Author Contributions

#A.J. and E.R.H. contributed equally.


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health R01GM114537 and R01GM125939 and the National Science Foundation CHE-1905081 to M.R.P. S.P.M. was supported by NIGMS T32GM118289.



https://doi.org/10.1021/jacsau.1c00455
*JACS Au* 2022, 2, 74−83

## REFERENCES


(1) Yang, Y. R.; Song, M.; Lee, H.; Jeon, Y.; Choi, E.-J.; Jang, H.-J.; Moon, H. Y.; Byun, H.-Y.; Kim, E.-K.; Kim, D. H.; Lee, M. N.; Koh, A.; Ghim, J.; Choi, J. H.; Lee-Kwon, W.; Kim, K. T.; Ryu, S. H.; Suh, P.-G. O-GlcNAcase Is Essential for Embryonic Development and Maintenance of Genomic Stability. *Aging Cell* **2012**, *11* (3), 439−448.

(2) Shafi, R.; Iyer, S. P.; Ellies, L. G.; O'Donnell, N.; Marek, K. W.; Chui, D.; Hart, G. W.; Marth, J. D. 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.* **2000**, *97* (11), 5735−5739.

(3) Sinclair, D. A. R.; Syrzycka, M.; Macauley, M. S.; Rastgardani, T.; Komljenovic, I.; Vocadlo, D. J.; Brock, H. W.; Honda, B. M. Drosophila O-GlcNAc Transferase (OGT) Is Encoded by the Polycomb Group (PcG) Gene, Super Sex Combs (Sxc). *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (32), 13427−13432.

(4) Haltiwanger, R. S.; Holt, G. D.; Hart, G. W. Enzymatic Addition of O-GlcNAc to Nuclear and Cytoplasmic Proteins. Identification of a Uridine Diphospho-N-Acetylglucosamine:Peptide Beta-N-Acetylglucosaminyltransferase. *J. Biol. Chem.* **1990**, *265* (5), 2563−2568.

(5) Hanover, J. A.; Yu, S.; Lubas, W. B.; Shin, S. H.; Ragano-Caracciola, M.; Kochran, J.; Love, D. C. Mitochondrial and Nucleocytoplasmic Isoforms of O-Linked GlcNAc Transferase Encoded by a Single Mammalian Gene. *Arch. Biochem. Biophys.* **2003**, *409* (2), 287−297.

(6) Ong, Q.; Han, W.; Yang, X. O-GlcNAc as an Integrator of Signaling Pathways. *Front. Endocrinol. (Lausanne, Switz.)* **2018**, *9*, 599.

(7) Zhang, Z.; Parker, M. P.; Graw, S.; Novikova, L. V.; Fedosyuk, H.; Fontes, J. D.; Koestler, D. C.; Peterson, K. R.; Slawson, C. O-GlcNAc Homeostasis Contributes to Cell Fate Decisions during Hematopoiesis. *J. Biol. Chem.* **2019**, *294* (4), 1363−1379.

(8) Sheikh, M. A.; Emerald, B. S.; Ansari, S. A. Stem Cell Fate Determination through Protein O-GlcNAcylation. *J. Biol. Chem.* **2021**, *296*, 100035.

(9) Parker, M. P.; Peterson, K. R.; Slawson, C. O-GlcNAcylation and O-GlcNAc Cycling Regulate Gene Transcription: Emerging Roles in Cancer. *Cancers* **2021**, *13* (7), 1666.

(10) Chang, Y.-H.; Weng, C.-L.; Lin, K.-I. O-GlcNAcylation and Its Role in the Immune System. *J. Biomed. Sci.* **2020**, *27* (1), 57.

(11) Liu, Y.; Yao, R.-Z.; Lian, S.; Liu, P.; Hu, Y.-J.; Shi, H.-Z.; Lv, H.-M.; Yang, Y.-Y.; Xu, B.; Li, S.-Z. O-GlcNAcylation: The "Stress and Nutrition Receptor" in Cell Stress Response. *Cell Stress Chaperones* **2021**, *26* (2), 297−309.

(12) Ma, Z.; Vocadlo, D. J.; Vosseller, K. Hyper-O-GlcNAcylation Is Anti-Apoptotic and Maintains Constitutive NF-κB Activity in Pancreatic Cancer Cells. *J. Biol. Chem.* **2013**, *288* (21), 15121−15130.

(13) Pinho, T. S.; Correia, S. C.; Perry, G.; Ambrósio, A. F.; Moreira, P. I. Diminished O-GlcNAcylation in Alzheimer's Disease Is Strongly Correlated with Mitochondrial Anomalies. *Biochim. Biophys. Acta, Mol. Basis Dis.* **2019**, *1865* (8), 2048−2059.

(14) Lee, B. E.; Kim, H. Y.; Kim, H.-J.; Jeong, H.; Kim, B.-G.; Lee, H.-E.; Lee, J.; Kim, H. B.; Lee, S. E.; Yang, Y. R.; Yi, E. C.; Hanover, J. A.; Myung, K.; Suh, P.-G.; Kwon, T.; Kim, J.-I. O-GlcNAcylation Regulates Dopamine Neuron Function, Survival and Degeneration in Parkinson Disease. *Brain* **2020**, *143* (12), 3699−3716.

(15) Ma, J.; Hart, G. W. Protein O-GlcNAcylation in Diabetes and Diabetic Complications. *Expert Rev. Proteomics* **2013**, *10* (4), 365−380.

(16) Ducheix, S.; Magré, J.; Cariou, B.; Prieur, X. Chronic O-GlcNAcylation and Diabetic Cardiomyopathy: The Bitterness of Glucose. *Front. Endocrinol.* **2018**, *9*, 642.

(17) Bolanle, I. O.; Riches-Suman, K.; Williamson, R.; Palmer, T. M. Emerging Roles of Protein O-GlcNAcylation in Cardiovascular Diseases: Insights and Novel Therapeutic Targets. *Pharmacol. Res.* **2021**, *165*, 105467.

(18) Umapathi, P.; Mesubi, O. O.; Banerjee, P. S.; Abrol, N.; Wang, Q.; Luczak, E. D.; Wu, Y.; Granger, J. M.; Wei, A.-C.; Reyes Gaido, O. E.; Florea, L.; Talbot, C. C.; Hart, G. W.; Zachara, N. E.; Anderson, M. E. Excessive O-GlcNAcylation Causes Heart Failure and Sudden Death. *Circulation* **2021**, *143* (17), 1687−1703.

(19) Hart, G. W.; 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* (1), 825−858.

(20) Yang, X.; Qian, K. Protein O-GlcNAcylation: Emerging Mechanisms and Functions. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (7), 452−465.

(21) Konopka, J. B. N-Acetylglucosamine Functions in Cell Signaling. *Scientifica* **2012**, *2012*, 1−15.

(22) Riu, I. H.; Shin, I. S.; Do, S. Il. Sp1Modulates NcOGT Activity to Alter Target Recognition and Enhanced Thermotolerance in E. Coli. *Biochem. Biophys. Res. Commun.* **2008**, *372* (1), 203−209.

(23) 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* (12), 949−957.

(24) Yuzwa, S. A.; Yadav, A. K.; Skorobogatko, Y.; Clark, T.; Vosseller, K.; Vocadlo, D. J. Mapping O-GlcNAc Modification Sites on Tau and Generation of a Site-Specific O-GlcNAc Tau Antibody. *Amino Acids* **2011**, *40* (3), 857−868.

(25) Han, C.; Gu, Y.; Shan, H.; Mi, W.; Sun, J.; Shi, M.; Zhang, X.; Lu, X.; Han, F.; Gong, Q.; Yu, W. O-GlcNAcylation of SIRT1 Enhances Its Deacetylase Activity and Promotes Cytoprotection under Stress. *Nat. Commun.* **2017**, *8* (1), 1491.

(26) Berk, J. M.; Maitra, S.; Dawdy, A. W.; Shabanowitz, J.; Hunt, D. F.; Wilson, K. L. O-Linked β-N-Acetylglucosamine (O-GlcNAc) Regulates Emerin Binding to Barrier to Autointegration Factor (BAF) in a Chromatin- and Lamin B-Enriched "Niche"*. *J. Biol. Chem.* **2013**, *288* (42), 30192−30209.

(27) Simon, D. N.; Wriston, A.; Fan, Q.; Shabanowitz, J.; Florwick, A.; Dharmaraj, T.; Peterson, S. B.; Gruenbaum, Y.; Carlson, C. R.; Grønning-Wang, L. M.; Hunt, D. F.; Wilson, K. L. OGT (O-GlcNAc Transferase) Selectively Modifies Multiple Residues Unique to Lamin A. *Cells* **2018**, *7* (5), 44.

(28) Xie, S.; Jin, N.; Gu, J.; Shi, J.; Sun, J.; Chu, D.; Zhang, L.; Dai, C.; Gu, J.; Gong, C.-X.; Iqbal, K.; Liu, F. O-GlcNAcylation of Protein Kinase A Catalytic Subunits Enhances Its Activity: A Mechanism Linked to Learning and Memory Deficits in Alzheimer's Disease. *Aging Cell* **2016**, *15* (3), 455−464.

(29) Li, H.; Liu, X.; Wang, D.; Su, L.; Zhao, T.; Li, Z.; Lin, C.; Zhang, Y.; Huang, B.; Lu, J.; Li, X. O-GlcNAcylation of SKN-1 Modulates the Lifespan and Oxidative Stress Resistance in Caenorhabditis Elegans. *Sci. Rep.* **2017**, *7* (1), 43601.

(30) Goodwin, O. Y.; Thomasson, M. S.; Lin, A. J.; Sweeney, M. M.; Macnaughtan, M. A. E. Coli Sabotages the in Vivo Production of O-Linked β-N-Acetylglucosamine-Modified Proteins. *J. Biotechnol.* **2013**, *168* (4), 315−323.

(31) Kim, Y. J.; Kang, M. J.; Kim, E.; Kweon, T. H.; Park, Y. S.; Ji, S.; Yang, W. H.; Yi, E. C.; Cho, J. W. O-GlcNAc Stabilizes SMAD4 by Inhibiting GSK-3β-Mediated Proteasomal Degradation. *Sci. Rep.* **2020**, *10* (1), 1−13.

(32) Ma, Z.; Chalkley, R. J.; Vosseller, K. Hyper-O-GlcNAcylation Activates Nuclear Factor κ-Light-Chain-Enhancer of Activated B Cells (NF-κB) Signaling through Interplay with Phosphorylation and Acetylation. *J. Biol. Chem.* **2017**, *292* (22), 9150−9163.

(33) Wu, J. L.; Chen, Y. J.; Wu, H. Y.; Tsai, D. Y.; Chiang, M. F.; Chen, Y. J.; Gao, S.; Lin, C. C.; Lin, C. H.; Khoo, K. H.; Lin, K. I. Temporal Regulation of Lsp1 O-GlcNAcylation and Phosphorylation during Apoptosis of Activated B Cells. *Nat. Commun.* **2016**, *7*, 12526.

(34) Housley, M. P.; Rodgers, J. T.; Udeshi, N. D.; Kelly, T. J.; Shabanowitz, J.; Hunt, D. F.; Puigserver, P.; Hart, G. W. O-GlcNAc Regulates FoxO Activation in Response to Glucose. *J. Biol. Chem.* **2008**, *283* (24), 16283−16292.

(35) Tan, E. P.; Caro, S.; Potnis, A.; Lanza, C.; Slawson, C. O-Linked N-Acetylglucosamine Cycling Regulates Mitotic Spindle Organization. *J. Biol. Chem.* **2013**, *288* (38), 27085−27099.




https://doi.org/10.1021/jacsau.1c00455
*JACS Au* 2022, 2, 74−83

(36) Maynard, J. C.; Burlingame, A. L.; Medzihradszky, K. F. Cysteine S-Linked N-Acetylglucosamine (S-GlcNAcylation), A New Post-Translational Modification in Mammals. *Mol. Cell. Proteomics* **2016**, *15* (11), 3405−3411.
(37) De Leon, C. A.; Levine, P. M.; Craven, T. W.; Pratt, M. R. 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* (27), 3507−3517.
(38) Tegl, G.; Hanson, J.; Chen, H.; Kwan, D. H.; Santana, A. G.; Withers, S. G. Facile Formation of β-thioGlcNAc Linkages to Thiol-Containing Sugars, Peptides, and Proteins Using a Mutant GH20 Hexosaminidase. *Angew. Chem.* **2019**, *131* (6), 1646−1651.
(39) Gorelik, A.; Bartual, S. G.; Borodkin, V. S.; Varghese, J.; Ferenbach, A. T.; van Aalten, D. M. F. Genetic Recoding to Dissect the Roles of Site-Specific Protein O-GlcNAcylation. *Nat. Struct. Mol. Biol.* **2019**, *26* (11), 1071−1077.
(40) Ramirez, D. H.; Aonbangkhen, C.; Wu, H. Y.; Naftaly, J. A.; Tang, S.; O'Meara, T. R.; Woo, C. M. Engineering a Proximity-Directed O-GlcNAc Transferase for Selective Protein O-GlcNAcylation in Cells. *ACS Chem. Biol.* **2020**, *15* (4), 1059−1066.
(41) Bever, C. S.; Dong, J.-X.; Vasylieva, N.; Barnych, B.; Cui, Y.; Xu, Z.-L.; Hammock, B. D.; Gee, S. J. VHH Antibodies: Emerging Reagents for the Analysis of Environmental Chemicals. *Anal. Bioanal. Chem.* **2016**, *408* (22), 5985−6002.
(42) Ge, Y.; Ramirez, D. H.; Yang, B.; D'Souza, A. K.; Aonbangkhen, C.; Wong, S.; Woo, C. M. Target Protein Deglycosylation in Living Cells by a Nanobody-Fused Split O-GlcNAcase. *Nat. Chem. Biol.* **2021**, *17* (5), 593−600.
(43) Ramirez, D. H.; Ge, Y.; Woo, C. M. O-GlcNAc Engineering on a Target Protein in Cells with Nanobody-OGT and Nanobody-SplitOGA. *Curr. Protoc.* **2021**, *1* (5), 1 DOI: 10.1002/cpz1.117.
(44) Lee, K. J.; Kang, D.; Park, H.-S. Site-Specific Labeling of Proteins Using Unnatural Amino Acids. *Mol. Cells* **2019**, *42* (5), 386−396.
(45) Dieterich, D. C.; Link, A. J.; Graumann, J.; Tirrell, D. A.; Schuman, E. M. Selective Identification of Newly Synthesized Proteins in Mammalian Cells Using Bioorthogonal Noncanonical Amino Acid Tagging (BONCAT). *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (25), 9482−9487.
(46) Van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G. Expanding the Diversity of Chemical Protein Modification Allows Post-Translational Mimicry. *Nature* **2007**, *446* (7139), 1105−1109.
(47) Machida, T.; Lang, K.; Xue, L.; Chin, J. W.; Winssinger, N. Site-Specific Glycoconjugation of Protein via Bioorthogonal Tetrazine Cycloaddition with a Genetically Encoded Trans-Cyclooctene or Bicyclononyne. *Bioconjugate Chem.* **2015**, *26* (5), 802−806.
(48) Xu, R.; Hanson, S. R.; Zhang, Z.; Yang, Y.-Y.; Schultz, P. G.; Wong, C.-H. Site-Specific Incorporation of the Mucin-Type N-Acetylgalactosamine-α-O-Threonine into Protein in Escherichia Coli. *J. Am. Chem. Soc.* **2009**, *131* (38), 13883.
(49) Carvalho, I.; Scheuerl, S. L.; Kartha, K. P. R.; Field, R. A. Practical Synthesis of the 2-Acetamido-3,4,6-Tri-O-Acetyl-2-Deoxy-β-D-Glucosides of Fmoc-Serine and Fmoc-Threonine and Their Benzyl Esters. *Carbohydr. Res.* **2003**, *338* (10), 1039−1043.
(50) Meinjohanns, E.; Meldal, M.; Bock, K. Efficient Synthesis of O-(2-Acetamido-2-Deoxy-β-d-Glucopyranosyl)-Ser/Thr Building Blocks for SPPS of O-GlcNAc Glycopeptides. *Tetrahedron Lett.* **1995**, *36* (50), 9205−9208.
(51) Saha, U. K.; Schmidt, R. R. Efficient Synthesis of O-(2-Acetamido-2-Deoxy-β-D-Glucopyranosyl)-Serine and -Threonine Building Blocks for Glycopeptide Formation. *J. Chem. Soc., Perkin Trans. 1* **1997**, *1* (12), 1855−1860.
(52) De Leon, C. A.; Lang, G.; Saavedra, M. I.; Pratt, M. R. 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* (16), 5032−5035.
(53) Marotta, N. P.; Cherwien, C. A.; Abeywardana, T.; Pratt, M. R. O-GlcNAc Modification Prevents Peptide-Dependent Acceleration of α-Synuclein Aggregation. *ChemBioChem* **2012**, *13* (18), 2665−2670.
(54) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J. C. M.; Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev.* **2019**, *119* (12), 7328−7443.
(55) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. Synthesis of Proteins by Native Chemical Ligation. *Science (Washington, DC, U. S.)* **1994**, *266* (5186), 776−779.
(56) Kent, S. B. H. Total Chemical Synthesis of Proteins. *Chem. Soc. Rev.* **2009**, *38* (2), 338−351.
(57) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed Protein Ligation: A General Method for Protein Engineering. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (12), 6705−6710.
(58) Xu, M. Q.; Perler, F. B. The Mechanism of Protein Splicing and Its Modulation by Mutation. *EMBO J.* **1996**, *15* (19), 5146−5153.
(59) Chong, S.; Mersha, F. B.; Comb, D. G.; Scott, M. E.; Landry, D.; Vence, L. M.; Perler, F. B.; Benner, J.; Kucera, R. B.; Hirvonen, C. A.; Pelletier, J. J.; Paulus, H.; Xu, M.-Q. Single-Column Purification of Free Recombinant Proteins Using a Self-Cleavable Affinity Tag Derived from a Protein Splicing Element. *Gene* **1997**, *192* (2), 271−281.
(60) Vila-Perelló, M. *Expressed Protein Ligation Methods and Protocols*; Springer Science+Business Media, 2020.
(61) Muir, T. W. Semisynthesis of Proteins by Expressed Protein Ligation. *Annu. Rev. Biochem.* **2003**, *72*, 249−289.
(62) Jin, K.; Li, X. Advances in Native Chemical Ligation−Desulfurization: A Powerful Strategy for Peptide and Protein. *Chem. - Eur. J.* **2018**, *24* (66), 17397−17404.
(63) Wan, Q.; Danishefsky, S. J. Free-Radical-Based, Specific Desulfurization of Cysteine: A Powerful Advance in the Synthesis of Polypeptides and Glycopolypeptides. *Angew. Chem., Int. Ed.* **2007**, *46* (48), 9248−9252.
(64) Yan, L. Z.; Dawson, P. E. Synthesis of Peptides and Proteins without Cysteine Residues by Native Chemical Ligation Combined with Desulfurization. *J. Am. Chem. Soc.* **2001**, *123* (4), 526−533.
(65) Malins, L. R.; Mitchell, N. J.; McGowan, S.; Payne, R. J. Oxidative Deselenization of Selenocysteine: Applications for Programmed Ligation at Serine. *Angew. Chem.* **2015**, *127* (43), 12907−12912.
(66) Metanis, N.; Keinan, E.; Dawson, P. E. Traceless Ligation of Cysteine Peptides Using Selective Deselenization. *Angew. Chem.* **2010**, *122* (39), 7203−7207.
(67) Kulkarni, S. S.; Sayers, J.; Premdjee, B.; Payne, R. J. Rapid and Efficient Protein Synthesis through Expansion of the Native Chemical Ligation Concept. *Nat. Rev. Chem.* **2018**, *2* (4), 122.
(68) Fang, G. M.; Li, Y. M.; Shen, F.; Huang, Y. C.; Li, J. B.; Lin, Y.; Cui, H. K.; Liu, L. Protein Chemical Synthesis by Ligation of Peptide Hydrazides. *Angew. Chem., Int. Ed.* **2011**, *50* (33), 7645−7649.
(69) Flood, D. T.; Hintzen, J. C. J.; Bird, M. J.; Cistrone, P. A.; Chen, J. S.; Dawson, P. E. Leveraging the Knorr Pyrazole Synthesis for the Facile Generation of Thioester Surrogates for Use in Native Chemical Ligation. *Angew. Chem., Int. Ed.* **2018**, *57* (36), 11634−11639.
(70) Naruse, N.; Kobayashi, D.; Ohkawachi, K.; Shigenaga, A.; Otaka, A. Copper-Mediated Deprotection of Thiazolidine and Selenazolidine Derivatives Applied to Native Chemical Ligation. *J. Org. Chem.* **2020**, *85* (3), 1425−1433.
(71) Shi, J.; Tomašič, T.; Sharif, S.; Brouwer, A. J.; Anderluh, M.; Ruijtenbeek, R.; Pieters, R. J. Peptide Microarray Analysis of the Cross-Talk between O-GlcNAcylation and Tyrosine Phosphorylation. *FEBS Lett.* **2017**, *591* (13), 1872−1883.
(72) Sharif, S.; Shi, J.; Ruijtenbeek, R.; Pieters, R. J. Study of Cross Talk between Phosphatases and OGA on a ZO-3-Derived Peptide. *Amino Acids* **2019**, *51* (4), 739−743.
(73) Tarrant, M. K.; Rho, H. S.; Xie, Z.; Jiang, Y. L.; Gross, C.; Culhane, J. C.; Yan, G.; Qian, J.; Ichikawa, Y.; Matsuoka, T.; Zachara, N.; Etzkorn, F. A.; Hart, G. W.; Jeong, J. S.; Blackshaw, S.; Zhu, H.;

Cole, P. A. Regulation of CK2 by Phosphorylation and O-GlcNAcylation Revealed by Semisynthesis. *Nat. Chem. Biol.* **2012**, *8* (3), 262−269.
(74) Schwagerus, S.; Reimann, O.; Despres, C.; Smet-Nocca, C.; Hackenberger, C. P. R. Semi-Synthesis of a Tag-Free O-GlcNAcylated Tau Protein by Sequential Chemoselective Ligation. *J. Pept. Sci.* **2016**, *22* (5), 327−333.
(75) Marotta, N. P.; Lin, Y. H.; Lewis, Y. E.; Ambroso, M. R.; Zaro, B. W.; Roth, M. T.; Arnold, D. B.; Langen, R.; Pratt, M. R. O-GlcNAc Modification Blocks the Aggregation and Toxicity of the Protein α-Synuclein Associated with Parkinson's Disease. *Nat. Chem.* **2015**, *7* (11), 913−920.
(76) Lewis, Y. E.; Galesic, A.; Levine, P. M.; De Leon, C. A.; Lamiri, N.; Brennan, C. K.; Pratt, M. R. O-GlcNAcylation of α-Synuclein at Serine 87 Reduces Aggregation without Affecting Membrane Binding. *ACS Chem. Biol.* **2017**, *12* (4), 1020−1027.
(77) Levine, P. M.; Galesic, A.; Balana, A. T.; Mahul-Mellier, A.-L. L.; Navarro, M. X.; De Leon, C. A.; Lashuel, H. A.; Pratt, M. R. 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* (5), 1511−1519.
(78) Levine, P. M.; De Leon, C. A.; Galesic, A.; Balana, A.; Marotta, N. P.; Lewis, Y. E.; Pratt, M. R. O-GlcNAc Modification Inhibits the Calpain-Mediated Cleavage of α-Synuclein. *Bioorg. Med. Chem.* **2017**, *25* (18), 4977−4982.
(79) Galesic, A.; Rakshit, A.; Cutolo, G.; Pacheco, R. P.; Balana, A. T.; Moon, S. P.; Pratt, M. R. Comparison of N-Acetyl-Glucosamine to Other Monosaccharides Reveals Structural Differences for the Inhibition of α-Synuclein Aggregation. *ACS Chem. Biol.* **2021**, *16* (1), 14−19.
(80) Balana, A. T.; Mukherjee, A.; Nagpal, H.; Moon, S. P.; Fierz, B.; Vasquez, K. M.; Pratt, M. R. O-GlcNAcylation of High Mobility Group Box 1 (HMGB1) Alters Its DNA Binding and DNA Damage Processing Activities. *J. Am. Chem. Soc.* **2021**, *143* (39), 16030−16040.
(81) Balana, A. T.; Levine, P. M.; Craven, T. W.; Mukherjee, S.; Pedowitz, N. J.; Moon, S. P.; Takahashi, T. T.; Becker, C. F. W.; Baker, D.; Pratt, M. R. O-GlcNAc Modification of Small Heat Shock Proteins Enhances Their Anti-Amyloid Chaperone Activity. *Nat. Chem.* **2021**, *13* (5), 441−450.
(82) Wright, T. H.; Bower, B. J.; Chalker, J. M.; Bernardes, G. J. L.; Wiewiora, R.; Ng, W. L.; Raj, R.; Faulkner, S.; Vallée, M. R. J.; Phanumartwiwath, A.; Coleman, O. D.; Thézénas, M. L.; Khan, M.; Galan, S. R. G.; Lercher, L.; Schombs, M. W.; Gerstberger, S.; Palm-Espling, M. E.; Baldwin, A. J.; Kessler, B. M.; Claridge, T. D. W.; Mohammed, S.; Davis, B. G. Posttranslational Mutagenesis: A Chemical Strategy for Exploring Protein Side-Chain Diversity. *Science (Washington, DC, U. S.)* **2016**, *354* (6312), 1 DOI: 10.1126/science.aag1465.
(83) Yang, A.; Ha, S.; Ahn, J.; Kim, R.; Kim, S.; Lee, Y.; Kim, J.; Soll, D.; Lee, H.-Y.; Park, H.-S. A Chemical Biology Route to Site-Specific Authentic Protein Modifications. *Science (Washington, DC, U. S.)* **2016**, *354* (6312), 623−626.
(84) Lercher, L.; Raj, R.; Patel, N. A.; Price, J.; Mohammed, S.; Robinson, C. V.; Schofield, C. J.; Davis, B. G. Generation of a Synthetic GlcNAcylated Nucleosome Reveals Regulation of Stability by H2A-Thr101 GlcNAcylation. *Nat. Commun.* **2015**, *6*, 7978.
(85) Raj, R.; Lercher, L.; Mohammed, S.; Davis, B. G. Synthetic Nucleosomes Reveal That GlcNAcylation Modulates Direct Interaction with the FACT Complex. *Angew. Chem., Int. Ed.* **2016**, *55* (31), 8918−8922.
(86) Yang, B.; Wang, N.; Schnier, P. D.; Zheng, F.; Zhu, H.; Polizzi, N. F.; Ittuveetil, A.; Saikam, V.; DeGrado, W. F.; Wang, Q.; Wang, P. G.; Wang, L. Genetically Introducing Biochemically Reactive Amino Acids Dehydroalanine and Dehydrobutyrine in Proteins. *J. Am. Chem. Soc.* **2019**, *141* (19), 7698−7703.
(87) Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. From Disulfide- to Thioether-Linked Glycoproteins. *Angew. Chem., Int. Ed.* **2008**, *47* (12), 2244−2247.
(88) Wulff-Fuentes, E.; Berendt, R. R.; Massman, L.; Danner, L.; Malard, F.; Vora, J.; Kahsay, R.; Olivier-Van Stichelen, S. The Human O-GlcNAcome Database and Meta-Analysis. *Sci. Data* **2021**, *8* (1), 25.
(89) Lin, C.-H.; Liao, C.-C.; Chen, M.-Y.; Chou, T.-Y. Feedback Regulation of O-GlcNAc Transferase through Translation Control to Maintain Intracellular O-GlcNAc Homeostasis. *Int. J. Mol. Sci.* **2021**, *22* (7), 3463.
(90) Mueller, T.; Ouyang, X.; Johnson, M. S.; Qian, W.-J.; Chatham, J. C.; Darley-Usmar, V.; Zhang, J. New Insights Into the Biology of Protein O-GlcNAcylation: Approaches and Observations. *Frontiers in Aging* **2021**, *1*, 620382.
(91) Zhu, Y.; Hart, G. Targeting the O-GlcNAc Transferase to Specific Proteins Using RNA Aptamers. *FASEB J.* **2021**, *35* (S1), 1 DOI: 10.1096/fasebj.2021.35.S1.01899.
(92) Jin, S.; Brea, R. J.; Rudd, A. K.; Moon, S. P.; Pratt, M. R.; Devaraj, N. K. Traceless Native Chemical Ligation of Lipid-Modified Peptide Surfactants by Mixed Micelle Formation. *Nat. Commun.* **2020**, *11* (1), 2793.
(93) González, S. S.; Ad, O.; Shah, B.; Zhang, Z.; Zhang, X.; Chatterjee, A.; Schepartz, A. Genetic Code Expansion in the Engineered Organism Vmax X2: High Yield and Exceptional Fidelity. *ACS Cent. Sci.* **2021**, *7* (9), 1500−1507.


https://doi.org/10.1021/jacsau.1c00455
*JACS Au* 2022, 2, 74−83