

This is a repository copy of *Potent De Novo Macrocyclic Peptides That Inhibit O-GlcNAc Transferase through an Allosteric Mechanism*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/196718/>

Version: Accepted Version

---

**Article:**

Alteen, Matthew G., Peacock, Hayden, Meek, Richard W. et al. (5 more authors) (2023) Potent De Novo Macrocyclic Peptides That Inhibit O-GlcNAc Transferase through an Allosteric Mechanism. *Angewandte Chemie - International Edition*. e202215671. ISSN 1433-7851

<https://doi.org/10.1002/anie.202215671>

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Potent de novo macrocyclic peptides that inhibit O-GlcNAc transferase through an allosteric mechanism

Matthew G. Alteen<sup>†,[a]</sup> Hayden Peacock<sup>†,[b]</sup> Richard W. Meek,<sup>[c]</sup> Jil A. Busmann,<sup>[d]</sup> Sha Zhu,<sup>[a]</sup> Gideon J. Davies,<sup>[c]</sup> Hiroaki Suga,<sup>\*,[b]</sup> and David J. Vocadlo<sup>\*,[a,d]</sup>

<sup>[a]</sup> Department of Chemistry, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada

<sup>[b]</sup> Department of Chemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Bunkyo-ku, Tokyo 113-0033, Japan

<sup>[c]</sup> York Structural Biology Laboratory, Department of Chemistry, University of York, York, YO10 5DD, UK

<sup>[d]</sup> Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada

<sup>†</sup> Authors contributed equally

\* Corresponding authors: David J. Vocadlo ([dvocadlo@sfu.ca](mailto:dvocadlo@sfu.ca)), Hiroaki Suga ([hsuga@chem.s.u-tokyo.ac.jp](mailto:hsuga@chem.s.u-tokyo.ac.jp))

## Abstract

Glycosyltransferases are a notoriously difficult to inhibit superfamily of enzymes. Here we apply an mRNA display technology integrated with genetic code reprogramming, referred to as the RaPID (random non-standard peptides integrated discovery) system, to identify macrocyclic peptides with high binding affinities for O-GlcNAc transferase (OGT). These macrocycles inhibit OGT activity through an allosteric mechanism that is driven by their binding to the tetratricopeptide repeats of OGT. Saturation mutagenesis in a maturation screen using 39 amino acids, including 22 non-canonical residues, led to an improved unnatural macrocycle that is ~40 times more potent than the parent compound ( $K_i^{\text{app}} = 1.5$  nM). Subsequent derivatization delivered a biotinylated derivative that enabled one-step affinity purification of OGT from complex samples. The high potency and novel mechanism of action of these OGT ligands should enable new approaches to elucidate the specificity and regulation of OGT.

## **Introduction**

Modification of proteins with N-acetylglucosamine (GlcNAc) O-linked to the hydroxyl groups of serine and threonine residues (O-GlcNAc) is an abundant and widely conserved post-translational modification (PTM). Hundreds of nuclear and cytosolic proteins in mammalian cells are modified by O-GlcNAc. The presence of this PTM on diverse proteins is consistent with its emergence as a regulator of numerous cellular pathways ranging from transcription<sup>[1]</sup> and translation<sup>[2,3]</sup> through to participating in cell signaling<sup>[4,5]</sup>. Given these molecular roles, it is unsurprising that O-GlcNAc is implicated in a range of physiological processes ranging, for example, from stress response<sup>[6]</sup> to nutrient sensing<sup>[4,7]</sup>. Remarkably, the modification of proteins with O-GlcNAc is principally controlled by just two enzymes. The glycosyltransferase O-GlcNAc transferase (OGT) from CAZy family GT41<sup>[8]</sup> uses uridine diphosphate GlcNAc (UDP-GlcNAc) as a glycosyl donor substrate and catalyzes the transfer of GlcNAc residues to its target acceptor proteins<sup>[9]</sup>. Removal of O-GlcNAc is catalyzed by the GH84<sup>[8]</sup> glycoside hydrolase O-GlcNAcase (OGA), which cleaves the glycosidic bond to liberate the protein hydroxyl group<sup>[9,10]</sup>. How these enzymes are targeted to their substrates, and the means by which they are regulated, remains poorly understood<sup>[11]</sup>. Strikingly, modulating the action of these two enzymes has shown benefit in neurodegenerative diseases<sup>[12,13]</sup> and cancers<sup>[14,15]</sup>, stimulating great interest in antagonists of this pathway<sup>[16,17]</sup>. Accordingly, efforts are underway to better understand the mechanisms that govern their activity and, by extension, to identify molecules that regulate these enzymes in unique ways that could be exploited to drive advances in this growing field.

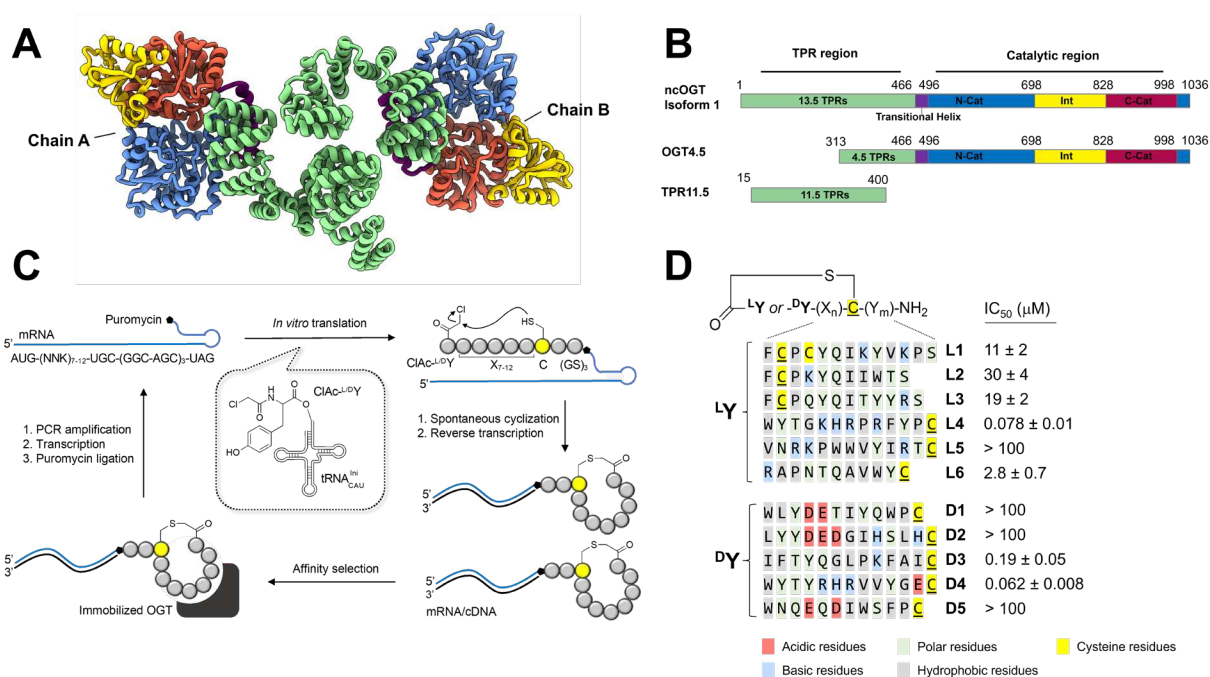
Over the past decade, significant progress has been made in generating chemical tools for OGA, resulting in the identification of several potent and selective inhibitors<sup>[18,19]</sup>. These compounds have enabled advances in medical imaging and translational pre-clinical studies<sup>[20,21]</sup>, even advancing into the clinic<sup>[22]</sup>. In contrast with OGA, for which potent and selective inhibitors have advanced through pre-clinical studies and into the clinic, the discovery of inhibitors of OGT has proven much more problematic. Rational approaches to inhibiting OGT in cells using metabolic precursors have yielded tool compounds that are helping to elucidate the role of OGT<sup>[23,24]</sup>, but these compounds can influence nucleotide sugar pools and may have limited selectivity<sup>[9]</sup>. High-throughput screening (HTS) has been used to discover small molecule OGT inhibitors with some success<sup>[25–27]</sup>, the most promising compounds being the OSMI series. These were identified and improved by Walker and co-workers using structure-guided optimization over several years to yield potent OGT inhibitors including the most advanced, OSMI-4 ( $K_d = 8$  nM), which is highly selective<sup>[28,29]</sup>. Notably, OSMI-4 and, indeed, all well-characterized inhibitors of

OGT, bind to the active site of the enzyme and are competitive inhibitors that consequently impair glycosylation of its entire set of protein substrates. We reasoned, given the limited chemical matter arising from HTS, that using alternative screening approaches could deliver promising new ligands. In addition, ligands binding outside the active site might serve as useful orthogonal chemical tools that could be used to help address questions about the biological roles of OGT.

OGT possesses a complex dimeric structure (Figure 1A and 1B) comprising a C-terminal catalytic section, appended to an N-terminal region of 13.5 tetratricopeptide repeats (TPRs)<sup>[30,31]</sup>. This TPR domain is known to play a role in dimerization<sup>[32]</sup> as well as binding to protein substrates within its superhelical groove<sup>[33]</sup>. Accordingly, we reasoned that ligands binding to this part of OGT could exhibit properties distinguishing them from more common competitive inhibitors. Given that OGT binds peptide substrates both within the active site and TPR domain, we speculated that a peptide display strategy might be an efficient route to discovering high affinity ligands for OGT, some of which might act as inhibitors. Among the various display methods that are known<sup>[34]</sup>, random non-standard peptides integrated discovery (RaPID) holds particular appeal<sup>[35]</sup>. This mRNA display system relies on the spontaneous reaction between an N-chloroacetylated initiator residue and a downstream cysteine residue of the nascent peptide chain to generate thioether-cyclized macrocycles (Figure 1C). Notably, RaPID also allows incorporation of non-natural amino acids, which are installed by adding the corresponding aminoacylated tRNAs to the *in vitro* translation reaction. These aminoacylated tRNAs are readily prepared using 'flexizymes', which are ribozyme catalysts<sup>[36]</sup>. This system permits the generation of diverse libraries of cyclic peptides, that can contain both canonical and non-canonical residues, which are covalently linked to mRNA/cDNA heteroduplexes that encode the peptide sequence. The speed and high sequence diversity ( $>10^{12}$ ) of the RaPID display approach, coupled with next-generation DNA sequencing, has led to the discovery of potent macrocyclic peptides binding to a variety of targets through affinity selection<sup>[37–39]</sup>. Though not having been previously applied to glycosyltransferases, we reasoned that RaPID could deliver high affinity ligands of OGT that might possess unique properties.

Here we report the discovery of potent macrocyclic peptides that bind the N-terminal TPR domain of OGT. We made the surprising observation that this series of compounds, which have low nanomolar affinity for OGT, inhibit the catalytic activity of this enzyme through an unprecedented allosteric mechanism. Follow-on saturation mutagenesis experiments that incorporated non-natural amino acids yielded macrocycles with even greater potency. Insertion of reactive functionalities into these high affinity macrocycles, followed by their subsequent

derivatization, afforded a biotinylated macrocyclic probe that allowed simple one-step affinity purification of OGT. These macrocyclic allosteric inhibitors represent unique chemical matter that can be exploited to create a range of chemical tools that will provide critical insight into the structure and function of OGT.



**Figure 1.** Discovery of macrocyclic peptides that bind O-GlcNAc transferase (OGT). **A**, Cryo-electron microscopy model of OGT, which forms a dimer and consists of an N-terminal superhelical TPR domain (green) and C-terminal catalytic sections (blue and red) with the intervening domain (yellow). **B**, Cartoon representation of the domain architecture of OGT with the same colour scheme as in **A**. **C**, Schematic outline of the random non-standard peptides integrated discovery (RaPID) mRNA display platform, permitting *in situ* generation of genetically encoded thioether-linked macrocycles. **D**, Sequences of enriched peptide binders of OGT identified through RaPID and corresponding  $IC_{50}$  values measured using recombinant ncOGT using a fluorescent activity assay.

## Results and Discussion

To identify potential peptide ligands for OGT, we used the RaPID system to perform an affinity selection against recombinant OGT bearing an N-terminal His<sub>6</sub> tag (Figure 1C). We designed the RaPID library so that between seven and twelve random natural amino acids were situated between a fixed C-terminal cysteine residue and a fixed N-terminal N-chloroacetyl-

tyrosine initiator residue, which react together and lead to cyclization of the translated peptide library. We performed two selection campaigns: one incorporating an initiator residue with L-chirality and one incorporating an initiator residue with D-chirality. Five rounds of selection were performed to enrich for sequences that displayed high binding affinity towards OGT. After these selections, we decoded the sequences of the enriched peptide binders using next generation sequencing of the attached cDNA barcodes and ranked the sequences according to their observed frequency within the sequencing data (Figure S1). The increasing level of library enrichment over successive rounds of selections and grouping of the final sequences into several distinct families (Figure S1) strongly suggested that the selection succeeded in delivering high affinity ligands of OGT.

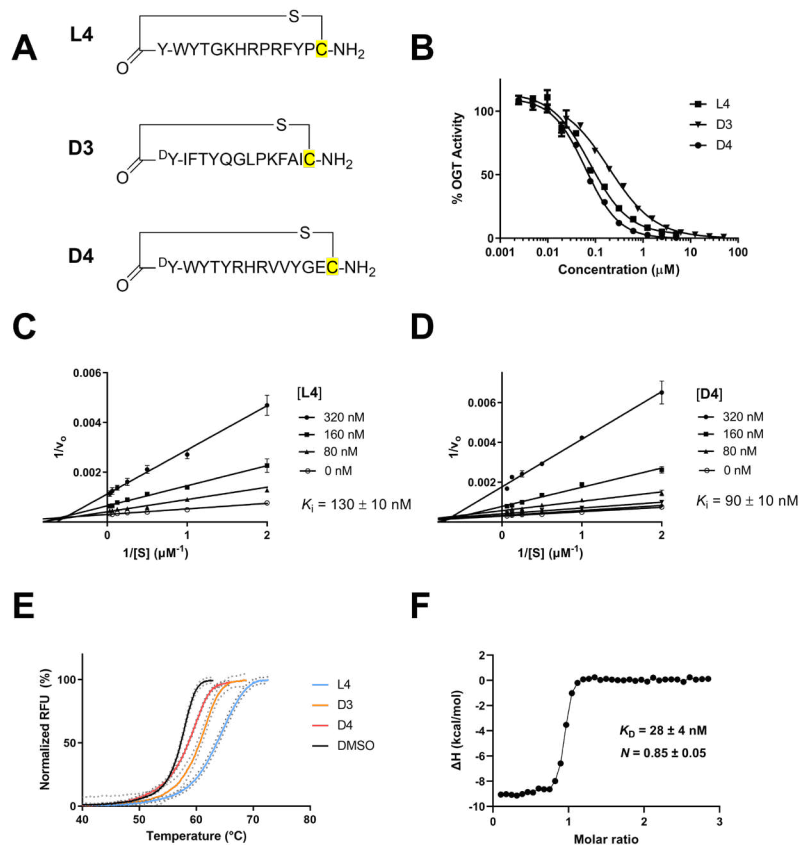
Based on these data, we synthesized and characterized 11 of the top ranked peptides (Figures S1, S2, and S3), encompassing both the L- and D-selections, for confirmation of binding using OGT bioassays. In choosing the sequences to synthesize, we endeavored to select at least one peptide from each distinct sequence family. Determination of the structure of peptide L1, which contains multiple cysteine residues, was performed using mass spectrometry (Figure S2). Because we were seeking ligands with inhibitory capacity, we assessed whether the synthesized candidate OGT ligands could inhibit the glycosyltransferase activity of OGT using a robust activity assay<sup>[25]</sup>. This assay measures the transfer of a fluorescently tagged GlcNAc residue to a peptide acceptor substrate, that interacts only with the catalytic domain. We found that seven of the macrocycles displayed concentration-dependent inhibition of OGT, while the remaining four macrocycles were inactive at concentrations below 100  $\mu$ M (Figure 1D). This lack of inhibition may stem from these cyclic peptides from non-inhibitory families binding to different sites. Among the seven inhibitory peptides, L4, D3, and D4 (Figure 2A) displayed  $IC_{50}$  values between 60-200 nM (Figure 2B), meriting further exploration of their properties.

As noted above, we reasoned that the ligands from this display-based screening strategy might bind to either the catalytic region or the TPR array of OGT. To determine which region these ligands were binding, we examined whether L4, D3, and D4 retained inhibitory potency when tested against a truncated construct of OGT having a shortened N-terminal TPR array known as OGT4.5 (Figure 1B)<sup>[31]</sup>. To our surprise, we found that none of these inhibitory macrocycles showed significant inhibition of OGT4.5 at concentrations that led to inhibition of full-length OGT (Figure S4). These observations suggested that all of these cyclic peptides were binding to the N-terminal region of OGT comprising the first nine TPRs that are most distant from the active site. Accordingly, these ligands would need to inhibit the glycosyltransferase activity of full-length OGT

through an allosteric mechanism. To further test this hypothesis, we determined the mode of inhibition of the two best macrocycles (L4 and D4) toward full-length OGT (Figure 2C and 2D). As expected, we found a pattern of kinetic data consistent with non-competitive inhibition of OGT with respect to the UDP-GlcNAc donor substrate, supporting the idea that these macrocycles were not inhibiting OGT by binding to the active site.

We next tested whether these macrocyclic peptides were binding directly to OGT using a differential scanning fluorimetry (DSF)-based thermal shift assay<sup>[40]</sup>. To gain further insight into which domain these molecules were likely binding, we used the OGT4.5 construct as well as a construct lacking the catalytic domain and comprising just 11.5 TPRs (TPR<sub>11.5</sub> Figure 1B)<sup>[32]</sup>. We examined the thermal stability of these two OGT constructs in the presence and absence of macrocycles L4, D3, and D4. We found that each of these macrocyclic peptides, when incubated with the TPR domain at a concentration >10-fold above their apparent  $K_i$  values, resulted in a significant positive shift in the melting temperature ( $T_m$ ) of the TPR construct as compared to the  $T_m$  observed in the absence of ligand (Figure 2E). In contrast, we observed no thermal stabilization of the OGT4.5 construct upon incubation with these ligands (Figure S5), consistent with our not observing any inhibition of the enzymatic activity of this construct in our assays. We also performed isothermal titration calorimetry (ITC) as an orthogonal approach to measure direct binding and found that macrocycle L4 bound to the TPR<sub>11.5</sub> construct with a stoichiometry of ~1 and a dissociation constant ( $K_D$ ) of 28 nM (Figure 2F, Figure S6). We also wondered whether the mechanism of action was through disruption of the dimer arrangement of OGT. Size-exclusion chromatography analysis indicates that OGT in the presence of L4 is still dimeric (Figure S7). Taken together, the data from these assays confirm that the macrocycles bind to the TPR domain of OGT and not the catalytic section, indicated that the inhibition of full-length OGT we observed in our activity assay occurs through an allosteric mechanism. These macrocycles therefore represent the first series of ligands that allosterically inhibit OGT and, to our knowledge, the first allosteric antagonists of any glycosyltransferase.

Encouraged by the high affinity and novel mechanism of inhibition by which this series of macrocyclic peptides were acting on OGT, we judged that enhancing their potency would be worthwhile and could be achieved by performing deep mutational scanning using standard and non-proteinogenic amino acids. We therefore designed a saturation mutagenesis library targeted towards peptide D3. We selected D3 for modification because we perceived its physicochemical properties in having just one charged amino acid residue, coupled with ease of chemical synthesis, could make it a suitable starting point for a cell permeable derivative. Each residue of



**Figure 2.** *In vitro* characterization of top macrocyclic peptides show potent nanomolar inhibition towards OGT through binding of the TPR domain. **A**, Structures and sequences of prioritized macrocycles identified from RaPID screening. **B**, Concentration-response curves showing inhibition of OGT determined by an *in vitro* activity assay. Data were fitted to a normalized 4-parameter sigmoidal inhibition curve using Graphpad Prism 8. Error bars represent S.E.M. of triplicate measurements. **C**, **D**, Lineweaver-Burk plots of OGT activity in the presence of L4 or D4 macrocycles, revealing a non-competitive mode of inhibition. **E**, Thermal stability of the TPR<sub>11.5</sub> in the presence or absence of 20  $\mu$ M macrocycle assessed through differential scanning fluorimetry. Dotted grey lines indicate S.E.M of quadruplicate measurements. Melting temperatures ( $T_m$ ) were determined from the inflection point of curves obtained by fitting the data to the Boltzmann sigmoidal function. **F**, Integrated heat plot from isothermal titration calorimetry of purified TPR<sub>11.5</sub> binding peptide L4. Data were fitted using Malvern Origin software and plotted using Graphpad Prism 8.

the D3 peptide sequence, except for the initiator and the terminal cysteine, was mutated to 39 other residues: each of the natural amino acids, excluding methionine and cysteine, as well as a panel of 22 non-proteinogenic amino acids (Figure 3A). The non-proteinogenic amino acids were introduced using a methionine codon within the mRNA library and the identity of each non-proteinogenic amino acid was encoded, as previously reported, using unique barcoding primers in the reverse transcription step<sup>[41]</sup>. We incubated the resulting display library with immobilized OGT under equilibrium binding conditions. The immobilized OGT was then washed under



equilibrating conditions and the bound library members were eluted from the target. Next generation sequencing was used to evaluate the frequencies of each sequence in both the eluted and starting libraries. We calculated a relative enrichment score ( $E$ ) for each mutation, which indicated its enrichment in the eluted library relative to the starting D3 peptide sequence. We then converted these scores to  $\log_2(E)$  values and organized and presented them as a heatmap (Figure 3B, Table 1). A  $\log_2(E)$  value of zero indicates a mutation that is neither enriched nor depleted, therefore indicating it had no net effect on binding affinity. A positive  $\log_2(E)$  value indicates a favorable mutation (increased enrichment) and a negative value indicates an unfavorable mutation (decreased enrichment).

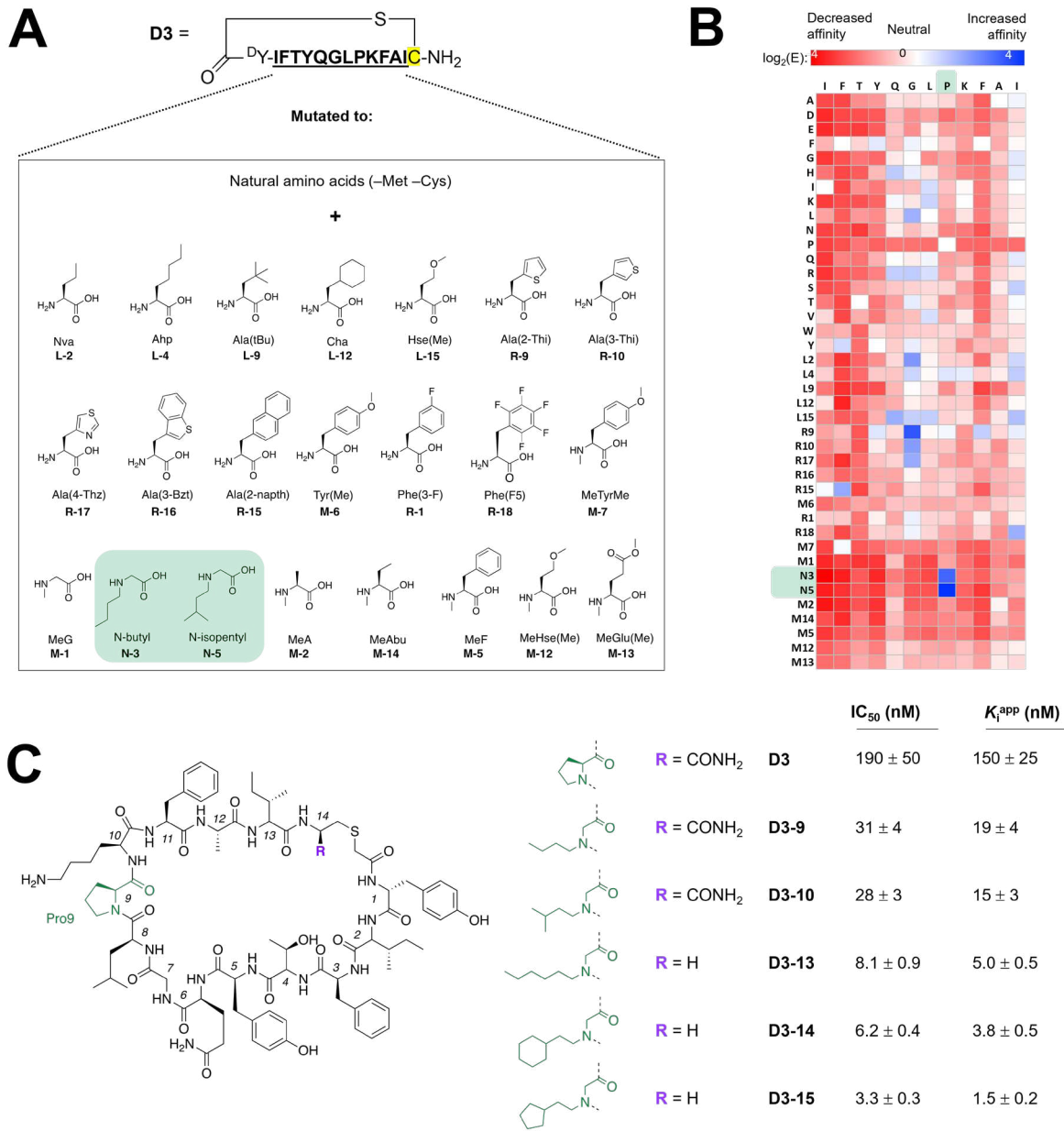
Analysis of this mutational scanning data indicated that the initial sequence of D3 peptide was already well optimized, as judged by the high proportion of negative enrichment scores of mutations at the majority of positions along the macrocyclic scaffold (Figure 3B). The first four residues of the N-terminal region were especially well conserved, with the exception of mutations that encoded similar aromatic residues in place of the phenylalanine at position 2. However, we noted that some variability was tolerated in the central region of the peptide, with some mutations resulting in  $\log_2(E)$  scores  $> 3$ . Based on these results, we prioritized the targeted synthesis of 12 additional macrocycles containing single mutations that resulted in increased enrichment relative to the D3 starting sequence. With these macrocycles in hand, we evaluated their inhibition of OGT using our activity assay as before (Figure S8). We observed a close correlation between the  $\log_2(E)$  scores of the mutant macrocycles and the corresponding  $IC_{50}$  values (Table 1), which supported the robustness of the saturation mutagenesis procedure. Notably, replacement of the proline at position 9 with non-proteinogenic N-butylglycine (D3-9) or N-isopentylglycine (D3-10) residues markedly increased the potency of inhibition from 230 nM to approximately 30 nM. Additionally, replacement of glycine at position 7 with amino acids bearing aliphatic chains (D3-3, D3-4) or 5-membered heterocycles (D3-5, D3-6, D3-7) resulted in modest 2- to 3-fold improvements in potency.

We were intrigued by the significant increase in potency of D3-9 and D3-10 that resulted from the presence of the backbone N-alkyl chains at position 7 (Figures 3A and 3B, green shading). We surmised that these alkyl groups might be interacting with a nearby hydrophobic pocket on the surface of OGT, affording increased binding affinity relative to the parental macrocycle. To capitalize on this observation, we synthesized an additional small panel of macrocycles with larger alkyl groups on the backbone nitrogen (Figure 3C). In addition, we removed the C-terminal amide (R group in Figure 3C), which we reasoned does not make any

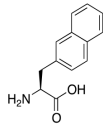






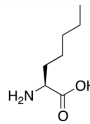
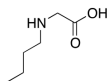
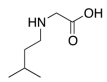
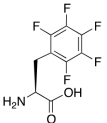
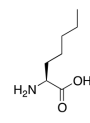
productive contacts with the protein because it serves as the site where the mRNA barcode is linked. We found that the incorporation of bulkier groups at position 9 increased the potency further to yield single digit nanomolar  $IC_{50}$  values (compounds D3-13, D3-14, and D3-15). The potency of the cyclopentyl derivative D3-15 is at least 40 times greater than that of the parent macrocycle, highlighting the utility of saturation mutagenesis in combination with structure-based design. To account for the tight-binding inhibition of these compounds relative to the concentration of enzyme present, we fitted data to the equation for tight-binding inhibition derived by Morrison<sup>[42]</sup>, which yielded similar single-digit  $K_i^{app}$  values (Figure S9). No binding of D3-15 to OGT4.5 was noted, consistent with these next generation macrocycles retaining the same mode of inhibition as their parent macrocycles. These data support the feasibility of further optimization of these ligands and define these macrocycles as among the most potent OGT ligands identified to date. Notably, the striking potency of these inhibitors is complemented by their unique allosteric mode of inhibition.

While D3-15 shows exceptional potency ( $K_i^{app} = 1.5$  nM), we found no sign of OGT inhibition upon incubation with cells. Nevertheless, spurred by the encouraging potency of these compounds, as well as their unique mechanism of inhibition, we sought to assess whether this macrocyclic peptide scaffold could be derivatized to incorporate a bioorthogonal functionality that could enable rapid derivatization. Given that the C-terminal cysteine is the site where the mRNA barcode is present within the RaPID library that is used in the selections, we hypothesized that the peptide near this site would tolerate modifications. We therefore synthesized an analogue of peptide D3-9 bearing a C-terminal propargylalanine residue (D3-Pra). This analogue showed no significant difference in  $IC_{50}$  towards OGT as compared D3 (data not shown), supporting this region of the molecule being tolerant to modification. We subsequently used copper catalyzed azide-alkyne click (CuAAC) chemistry to install a biotin moiety appended through a cleavable disulfide linker (D3-16, Figure 4A). This biotinylated analogue (D3-16) also retained high affinity for the enzyme (Figure 4B), supporting the utility of this position on the inhibitory scaffold to generate other potential tool compounds.

We next assessed whether the biotinylated macrocycle D3-16 would be useful for affinity purification of OGT from various samples. Chemoproteomic approaches using high affinity ligands to enrich endogenous target proteins from tissue samples is of interest because it can allow identification of endogenous protein modifications and binding partners. Therefore, to investigate this, we added D3-16 to purified recombinant OGT and then subsequently added streptavidin coated magnetic beads. After gentle washing of the beads, OGT could be readily eluted from the

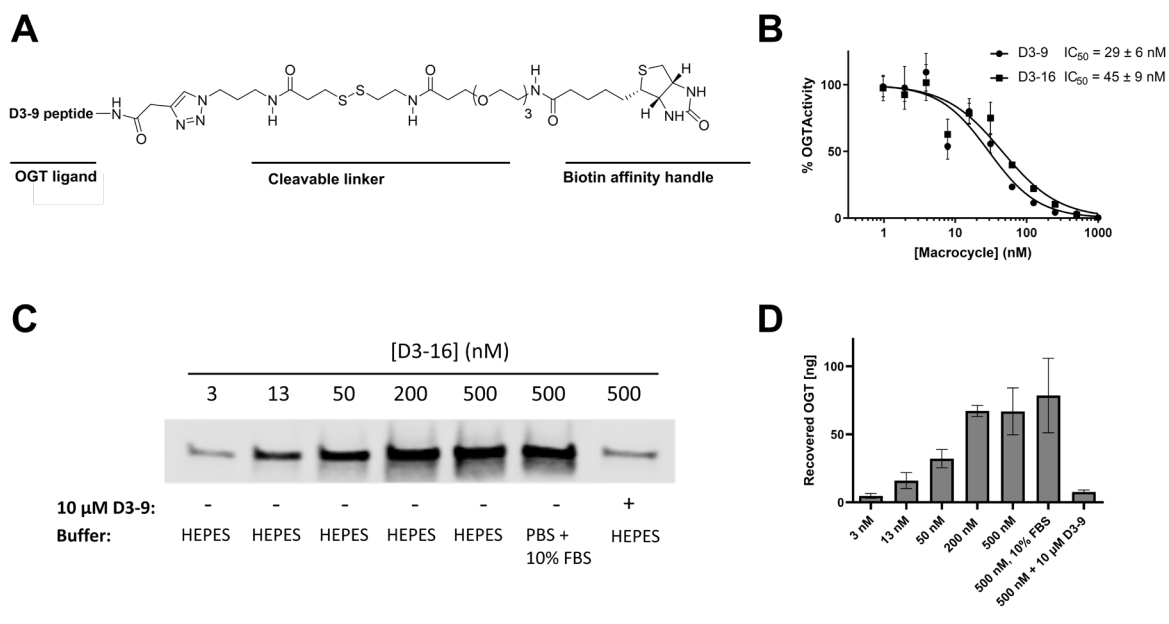


**Figure 3.** Site-saturation mutagenesis of peptide D3 for optimization of binding affinity. **A**, Mutagenesis strategy using a tailored mRNA library encoding the parent D3 peptide sequence permitting introduction of single mutations using non-natural amino acids and an expanded codon based *in vitro* translation system. The green shading highlights mutations of particular interest. **B**, Matrix outlining enrichment scores of mutant amino acid substitutions at each position along the core peptide sequence. Peptides with Log<sub>2</sub>(E) scores > 0 (blue) indicate increased enrichment relative to the parent sequence. Green shading highlights the N-alkyl glycine replacements of proline 9, which are highly enriched. **C**, Structure and resulting inhibitory potencies of chemically synthesized N-alkylglycine macrocycle analogues.

Compound	Site of mutation	Mutant residue	Structure	Log <sub>2</sub> (E)	IC <sub>50</sub> (nM)
D3	Parent	--	--	0	230 ± 22
D3-1	Phe 3	3-(2-Naphthyl)-alanine		1.32	86 ± 9
D3-2	Glu6	O-Methylhomoserine		1.24	170 ± 20
D3-3	Gly7	Norvaline		1.61	132 ± 14
D3-4	Gly 7	Leucine		1.21	115 ± 15
D3-5	Gly 7	3-(4-Thiazolyl)-alanine		1.30	108 ± 12
D3-6	Gly 7	3-(2-Thiophenyl)-alanine		2.59	67 ± 7
D3-7	Gly 7	3-(3-Thiophenyl)-alanine		1.52	126 ± 12
D3-8	Gly 7	Homonorleucine		0.65	635 ± 96
D3-9	Pro 9	N-Butylglycine		2.34	31 ± 4
D3-10	Pro 9	N-Isopentylglycine		3.14	28 ± 3
D3-11	Ile 13	3-(Pentafluorophenyl)-alanine		1.25	85 ± 10
D3-12	Ile 13	Homonorleucine		0.79	173 ± 10

**Table 1.** Log<sub>2</sub>(E) enrichment values for mutant D3 peptide series obtained from site-scanning mutagenesis and corresponding *in vitro* IC<sub>50</sub> values determined after synthesis of peptides.

beads through addition of 60 mM dithiothreitol (DTT), which acts to cleave the disulfide linker connecting the biotin handle with the OGT ligand. Subsequent analysis of eluted fractions by immunoblot indicated that OGT is effectively isolated and eluted during the pull-down experiment (Figure 4C and 4D). Additionally, we found that the amount of OGT isolated in these experiments was dependant on the concentration of biotinylated ligand added to the lysate and was blocked by the addition of non-biotinylated macrocycle D3-9, which acted as a competitive control. Together, these data indicated that the isolated OGT consisted of folded enzyme that interacted specifically with the streptavidin matrix only in the presence of biotinylated D3-16. This probe was similarly capable of binding and enriching OGT from a solution of 10% fetal bovine serum (FBS) in PBS, suggesting that it retained high affinity and specificity for OGT even when using complex protein samples. Based on these data, we expect that this tool compound will prove useful for future chemical proteomic analyses of OGT.



**Figure 4.** Biotinylated D3-16 macrocycle enables affinity purification of OGT from buffer and serum. **A**, Structure of biotinylated macrocycle D3-16, synthesized from D3-9, for affinity purification of OGT. **B**, D3-16 shows similar affinity for OGT compared to non-biotinylated analogue D3-9. Inhibition of OGT was assessed using a fluorescent activity assay as previously described. **C**, Representative western blot of recombinant OGT samples treated with various concentrations of D3-16. Samples were incubated with streptavidin-coated magnetic beads, washed, and then eluted with 60 mM DTT to cleave the linker and release bound protein from the beads. **D**, Quantification of amounts of OGT isolated from samples and analyzed by western blot. SDS-PAGE gels were loaded with a standard curve of recombinant OGT and quantified by fluorescence intensity after blotting. Data represent mean values from three independent experiments; error bars represent one standard deviation.

## Conclusion

The identification of inhibitors of glycosyltransferases has proved to be challenging and remains a topic of high interest<sup>[43]</sup>. Indeed, whilst few glycosyltransferase inhibitors are known,<sup>[25]</sup> several of these are peptidic in nature<sup>[44,45]</sup>. Most of these inhibitors, however, exhibit modest potencies. OGT is one of the few exceptions, for which the OSMI series of inhibitors have emerged as ligands with unusually high potency<sup>[28]</sup>. Nevertheless, all known glycosyltransferase inhibitors are, to our best knowledge, active site directed competitive inhibitors. Given that the RaPID system has been used to deliver ligands of various protein families, we reasoned that this selection-based display method could quickly deliver high potency inhibitors of glycosyltransferases. We also judged that the application of this peptide-based screen be particularly effective against OGT, especially because the enzyme binds peptide substrates within the superhelical groove of the TPR array that extends into the active site of the catalytic domain. Furthermore, given the ability of macrocyclic peptides to bind diverse protein topologies, we reasoned that some resulting ligands might be able to bind to the TPR array and, accordingly, inhibit OGT.

Using RaPID, we successfully discovered and optimized a series of high affinity thioether-linked macrocycles that bind to the TPR domain of OGT. Notably, we show that these compounds inhibit the catalytic activity of OGT despite not binding to the catalytic section and, as such, they operate through an allosteric mechanism. To our knowledge, this is an unprecedented mode of inhibition for glycosyltransferases. The dissociation and inhibition constants of the D3 series of macrocycles place these compounds (D3-15,  $K_i^{app} = 1.5$  nM) among the most potent inhibitors of OGT and, indeed, of any glycosyltransferase. Moreover, the allosteric mode of inhibition of these macrocycles confers on them attractive properties that differentiate them from competitive inhibitors<sup>[46,47]</sup>. Indeed, allosteric inhibitors offer benefits over competitive inhibitors in that inhibition is not mitigated by the accumulation of substrate, which has made them of great interest as pharmacological agents. Notably, the RaPID system has, in some cases, previously delivered allosteric enzyme inhibitors<sup>[48,49]</sup>, which suggests this platform has particular promise in furnishing antagonists having unusual and desirable inhibitory properties. While the current generation of macrocycles are not active within cells, enhancing the permeability of such compounds is a topic of intense current interest. Accordingly, we anticipate that application of emerging strategies, including promising new experimental<sup>[50-51]</sup> and computational methods,<sup>[52]</sup> to these series of macrocyclic OGT inhibitors will be a productive future line of research that will ultimately deliver selective cell active allosteric OGT inhibitors. Finally, the ability to append a biotin affinity tag to these OGT-targeted macrocycles, without perturbing the high affinity of these compounds,

enables the efficient isolation of OGT samples through a simple one-step affinity purification. We therefore expect that an important future application of these macrocycles will be to serve as valuable tool compounds that could be used to isolate endogenous OGT from various tissues and deliver new insights into its substrate specificity, post-translational modifications, protein interaction network, and regulation within cells.

### ***Funding Supports***

This work was supported by GlycoNet, the Canadian Glycomics Network (CD-1), the Canadian Cancer Society Research Institute (CCSRI-706825), and the Canadian Institutes of Health Research (PJT-156202) to D.J.V. and the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Specially Promoted Research (JP20H05618) to H.S. D.J.V. thanks the Canada Research Chairs program for support as a Tier I CRC in Chemical Biology.

### ***Author Contributions***

H.P. and M.G.A. contributed equally in conceiving of the study, designing experiments, analyzing results, and both writing and editing the manuscript. R.W.M. expressed and purified recombinant protein samples, performed biophysical assays and analyzed results, and contributed to writing the manuscript. J.A.B. performed and analyzed affinity purification and immunoblot experiments. S.Z. assisted with peptide synthesis. G.J.D. contributed to biophysical assay design. D.J.V. contributed to study conception and with H.S. designed research, oversaw research, edited the manuscript, and administered funding. All authors reviewed and provided input into the manuscript.

### ***References***

- [1] M. Resto, B.-H. Kim, A. G. Fernandez, B. J. Abraham, K. Zhao, B. A. Lewis, *J Biol Chem* **2016**, *291*, 22703-22713.
- [2] Y. Zhu, T. W. Liu, S. Cecioni, R. Eskandari, W. F. Zandberg, D. J. Vocadlo, *Nat Chem Biol* **2015**, *11*, 319–325.
- [3] X. Li, Q. Zhu, X. Shi, Y. Cheng, X. Li, H. Xu, X. Duan, L. C. Hsieh-Wilson, J. Chu, J. Pelletier, M. Ni, Z. Zheng, S. Li, W. Yi, *Proc Natl Acad Sci USA* **2022**, *119*, e2202821119.

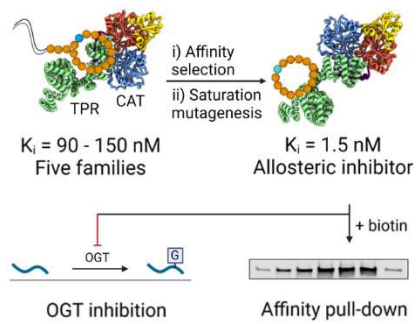
- [4] L. Chen, Y. Li, Z. Song, S. Xue, F. Liu, X. Chang, Y. Wu, X. Duan, H. Wu, *Proc Natl Acad Sci USA* **2022**, *116*, 7857-7866.
- [5] H. M. Itkonen, S. S. Gorad, D. Y. Duveau, S. E. S. Martin, A. Barkovskaya, T. F. Bathen, S. A. Moestue, I. G. Mills, *Oncotarget* **2016**, *7*, 12464–12476.
- [6] Z. Kazemi, H. Chang, S. Haserodt, C. McKen, N. E. Zachara, *J Biol Chem* **2010**, *285*, 39096–39107.
- [7] P. Chen, T. J. Smith, J. Wu, P. F. Siesser, B. J. Bisnett, F. Khan, M. Hogue, E. Soderblom, F. Tang, J. R. Marks, M. B. Major, B. M. Swarts, M. Boyce, J. Chi, *EMBO J* **2017**, *36*, 2233–2250.
- [8] E. Drula, M.-L. Garron, S. Dogan, V. Lombard, B. Henrissat, N. Terrapon, *Nucl Acids Res* **2022**, *50*, D571-D577.
- [9] M. G. Alteen, H. Y. Tan, D. J. Vocadlo, *Curr Opin Struct Biol* **2021**, *68*, 157–165.
- [10] C. Roth, S. Chan, W. A. Offen, G. R. Hemsworth, L. I. Willems, D. T. King, V. Varghese, R. Britton, D. J. Vocadlo, G. J. Davies, *Nat Chem Biol* **2017**, *13*, 610–612.
- [11] H. M. Stephen, T. M. Adams, L. Wells, *Glycobiology* **2021**, *31*, 724–733.
- [12] N. B. Hastings, X. Wang, L. Song, B. D. Butts, D. Grotz, R. Hargreaves, J. Fred Hess, K.-L. K. Hong, C. R.-R. Huang, L. Hyde, M. Laverty, J. Lee, D. Levitan, S. X. Lu, M. Maguire, V. Mahadomrongkul, E. J. McEachern, X. Ouyang, T. W. Rosahl, H. Selnick, M. Stanton, G. Terracina, D. J. Vocadlo, G. Wang, J. L. Duffy, E. M. Parker, L. Zhang, *Mol Neurodegener* **2017**, *12*, 39.
- [13] S. A. Yuzwa, X. Shan, M. S. Macauley, T. Clark, Y. Skorobogatko, K. Vosseller, D. J. Vocadlo, *Nat Chem Biol* **2012**, *8*, 393–399.
- [14] C. M. Ferrer, V. L. Sodi, M. J. Reginato, *J Mol Biol* **2016**, *428*, 3282–3294.
- [15] T. P. Lynch, C. M. Ferrer, S. R. E. Jackson, K. S. Shahriari, K. Vosseller, M. J. Reginato, *J Biol Chem* **2012**, *287*, 11070–11081.
- [16] M. Worth, C. W. Hu, H. Li, D. Fan, A. Estevez, D. Zhu, A. Wang, J. Jiang, *Chem Commun* **2019**, *55*, 13291–13294.
- [17] H. M. Itkonen, N. Poulouse, R. E. Steele, S. E. S. Martin, Z. G. Levine, D. Y. Duveau, R. Carelli, R. Singh, A. Urbanucci, M. Loda, C. J. Thomas, I. G. Mills, S. Walker, *Molec Cancer Res* **2020**, *18*, 1512–1521.
- [18] S. A. Yuzwa, M. S. Macauley, J. E. Heinonen, X. Shan, R. J. Dennis, Y. He, G. E. Whitworth, K. A. Stubbs, E. J. McEachern, G. J. Davies, D. J. Vocadlo, *Nat Chem Biol* **2008**, *4*, 483–490.
- [19] M. Bergeron-Brlek, J. Goodwin-Tindall, N. Cekic, C. Roth, W. F. Zandberg, X. Shan, V. Varghese, S. Chan, G. J. Davies, D. J. Vocadlo, R. Britton, *Angew Chem Int Ed* **2015**, *54*, 15429-15433.



- [20] H. G. Selnick, J. F. Hess, C. Tang, K. Liu, J. B. Schachter, J. E. Ballard, J. Marcus, D. J. Klein, X. Wang, M. Pearson, M. J. Savage, R. Kaul, T. S. Li, D. J. Vocadlo, Y. Zhou, Y. Zhu, C. Mu, Y. Wang, Z. Wei, C. Bai, J. L. Duffy, E. J. McEachern, *Med Chem* **2019**, *62*, 10062–10097.
- [21] S. Lu, M. B. Haskali, K. M. Ruley, N. J. F. Dreyfus, S. L. DuBois, S. Paul, J. S. Liow, C. L. Morse, A. Kowalski, R. L. Gladding, J. Gilmore, A. J. Mogg, S. Michelle Morin, P. J. Lindsay-Scott, J. Craig Ruble, N. A. Kant, S. Shcherbinin, V. N. Barth, M. P. Johnson, M. Cuadrado, E. Jambrina, A. J. Mannes, H. N. Nuthall, S. S. Zoghbi, C. D. Jesudason, R. B. Innis, V. W. Pike, *Sci Transl Med* **2020**, *12*, 517-530.
- [22] S. M. Smith, A. Struyk, D. Jonathan, R. Declercq, J. Marcus, D. Toolan, X. Wang, J. B. Schachter, M. Cosden, M. Pearson, F. Hess, H. Selnick, C. Salinas, W. Li, J. Duffy, E. McEachern, D. Vocadlo, J. J. Renger, H. D. Eric, M. Forman, D. Schoepp, *Alzheimers Dement* **2016**, *12*, P261–P261.
- [23] T. M. Gloster, W. F. Zandberg, J. E. Heinonen, D. L. Shen, L. Deng, D. J. Vocadlo, *Nat Chem Biol* **2011**, *7*, 174–181.
- [24] T. W. Liu, W. F. Zandberg, T. M. Gloster, L. Deng, K. D. Murray, X. Shan, D. J. Vocadlo, *Angew Chem Int Ed* **2018**, *57*, 7644–7648.
- [25] M. G. Alteen, C. Gros, R. W. Meek, D. A. Cardoso, J. A. Busmann, G. Sangouard, M. C. Deen, H.-Y. Tan, D. L. Shen, C. C. Russell, G. J. Davies, P. J. Robinson, A. McCluskey, D. J. Vocadlo, *Angew Chem Int Ed* **2020**, *132*, 9688–9696.
- [26] B. J. Gross, B. C. Kraybill, S. Walker, *J Am Chem Soc* **2005**, *127*, 14588–14589. [27] B. J. Gross, J. G. Swoboda, S. Walker, *J Am Chem Soc* **2008**, *130*, 440-441
- [28] S. E. S. Martin, Z. W. Tan, H. M. Itkonen, D. Y. Duveau, J. A. Paulo, J. Janetzko, P. L. Boutz, L. Törk, F. A. Moss, C. J. Thomas, S. P. Gygi, M. B. Lazarus, S. Walker, *J Am Chem Soc* **2018**, *140*, 13542–13545.
- [29] R. F. Ortiz-Meoz, J. Jiang, M. B. Lazarus, M. Orman, J. Janetzko, C. Fan, D. Y. Duveau, Z. W. Tan, C. J. Thomas, S. Walker, *ACS Chemical Biology* **2015**, *10*, 1392–1397.
- [30] R. W. Meek, J. N. Blaza, J. A. Busmann, M. G. Alteen, D. J. Vocadlo, G. J. Davies, *Nat Commun* **2021**, *12*, 6508.
- [31] M. B. Lazarus, Y. Nam, J. Jiang, P. Sliz, S. Walker, *Nature* **2011**, *469*, 564–569.
- [32] M. Jínek, J. Rehwinkel, B. D. Lazarus, E. Izaurrealde, J. A. Hanover, E. Conti, *Nat Struct Mol Biol* **2004**, *11*, 1001–1007.
- [33] Z. G. Levine, C. Fan, M. S. Melicher, M. Orman, T. Benjamin, S. Walker, *J Am Chem Soc* **2018**, *140*, 3510–3513.
- [34] W. Jaroszewicz, J. Morcinek-Orłowska, K. Pierzynowska, L. Gaffke, G. Węgrzyn, *FEMS Microbiol Rev* **2021**, *46*, fuab052.

- [35] T. Passioura, H. Suga, *Chem Commun* **2017**, *53*, 1931–1940.
- [36] Y. Goto, T. Katoh, H. Suga, *Nat Protoc* **2011**, *6*, 779–790.
- [37] M. Nawatha, J. M. Rogers, S. M. Bonn, I. Livneh, B. Lemma, S. M. Mali, G. B. Vamisetti, H. Sun, B. Bercovich, Y. Huang, A. Ciechanover, D. Fushman, H. Suga, A. Brik, *Nat Chem* **2019**, *11*, 644–652.
- [38] J. Morimoto, Y. Hayashi, H. Suga, *Angew Chem Int Ed* **2012**, *51*, 3423–3427.
- [39] E. Stefan, R. Obexer, S. Hofmann, K. Vu Huu, Y. Huang, N. Morgner, H. Suga, R. Tampé, *Elife* **2021**, *10*, e67732.
- [40] F. H. Niesen, H. Berglund, M. Vedadi, *Nat Protoc* **2007**, *2*, 2212–2221.
- [41] J. M. Rogers, T. Passioura, H. Suga, *Proc Natl Acad Sci USA* **2018**, *115*, 10959–10964.
- [42] J. F. Morrison, *Biochim Biophys Acta, Enzymol* **1969**, *185*, 269–286.
- [43] D. J. Vocadlo, T. L. Lowary, C. R. Bertozzi, R. L. Schnaar, J. D. Esko, in *Essentials of Glycobiology, 4<sup>th</sup> Edition*, Cold Spring Harbor Laboratory Press, **2022**.
- [44] K. Rafie, A. Gorelik, R. Trapannone, V. S. Borodkin, D. M. F. van Aalten, *Bioconjugate Chem* **2018**, *29*, 1834–1840. [45] K. Y. Lee, H. G. Kim, M. R. Hwang, J. il Chae, J. M. Yang, Y. C. Lee, Y. K. Choo, Y. Ik Lee, S. S. Lee, S. il Do, *J Biol Chem* **2002**, *277*, 49341–49351.
- [46] X. Lu, J. B. Smaill, K. Ding, *Angew Chem Int Ed* **2020**, *59*, 13764–13776.
- [47] C. J. Wenthur, P. R. Gentry, T. P. Mathews, C. W. Lindsley, *Annu Rev Pharmacol Toxicol* **2014**, *54*, 165–184.
- [48] H. Yu, P. Dranchak, Z. Li, R. Macarthur, M. S. Munson, N. Mehzabeen, N. J. Baird, K. P. Battalie, D. Ross, S. Lovell, C. K. S. Carlow, H. Suga, J. Inglese, *Nat Commun* **2017** *8*, 14932.
- [49] Y. Matsunaga, N. K. Bashiruddin, Y. Kitago, J. Takagi, H. Suga, *Cell Chem Biol* **2016**, *23*, 1341–1350.
- [50] G. J. Saunders, A. K. Yudin, *Angew Chem Int Ed* **2022**, *61*, e202206866.
- [51] S.A. Dai, Q. Hu, R. Gao, E. E. Blythe, K. K. Touhara, H. Peacock, Z. Zhang, M. von Zastrow, H. Suga, K.M. Shokat, *Cell* **2022**, *185*, 3950–3965
- [52] G. Bhardwaj et al, *Cell* **2022**, *185*, 3520–3532.

## TOC Graphic:



Using the random non-standard peptides integrated discovery (RaPID) system led to macrocyclic peptide ligands of *O*-GlcNAc transferase (OGT). The most potent of these ligands bind to the tetratricopeptide repeat region of OGT and inhibit this enzyme in an allosteric manner, making them promising tool compounds for studying OGT.

## Supporting Information

### **Potent de novo macrocyclic peptides that inhibit O-GlcNAc transferase through an allosteric mechanism**

*Matthew G. Alteen<sup>\*,[a]</sup> Hayden Peacock<sup>\*,[b]</sup> Richard W. Meek,<sup>[c]</sup> Jil A. Busmann,<sup>[d]</sup> Sha Zhu,<sup>[a]</sup> Gideon J. Davies,<sup>[c]</sup> Hiroaki Suga,<sup>† [b]</sup> and David J. Vocadlo<sup>† [a,d]</sup>*

*<sup>[a]</sup> Department of Chemistry, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada*

*<sup>[b]</sup> Department of Chemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Bunkyo-ku, Tokyo 113-0033, Japan*

*<sup>[c]</sup> York Structural Biology Laboratory, Department of Chemistry, University of York, York, YO10 5DD, UK*

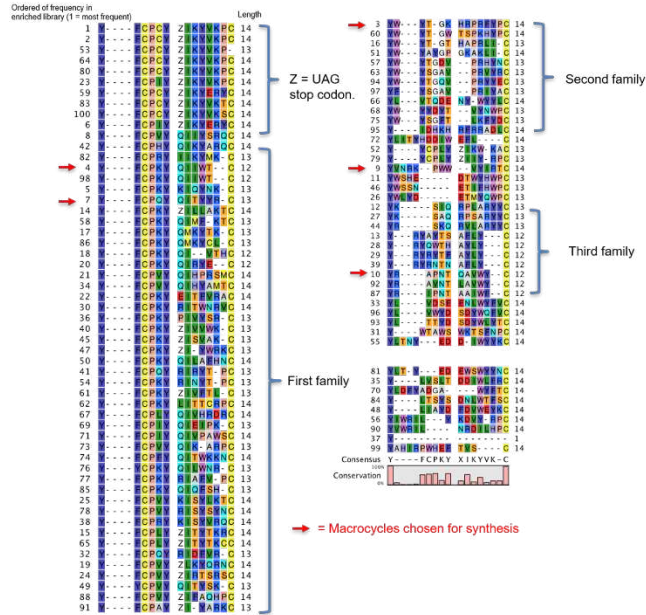
*<sup>[d]</sup> Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada*

*† Corresponding authors: David J. Vocadlo ([dvocadlo@sfu.ca](mailto:dvocadlo@sfu.ca)), Hiroaki Suga ([hsuga@chem.s.u-tokyo.ac.jp](mailto:hsuga@chem.s.u-tokyo.ac.jp))*

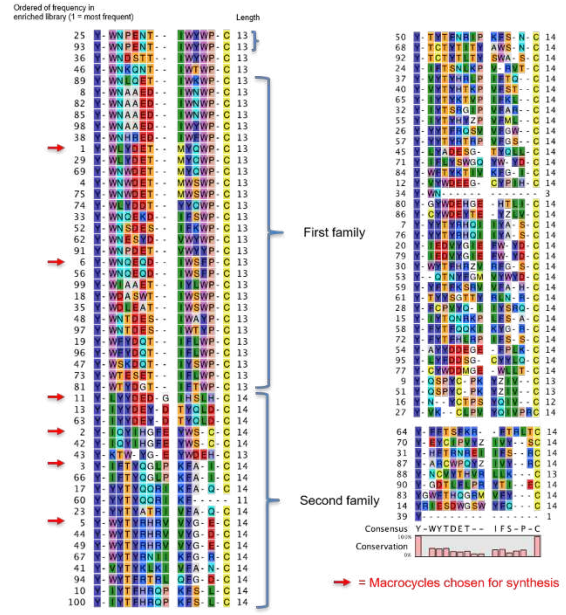
*\*Authors contributed equally*

# SUPPLEMENTAL FIGURES

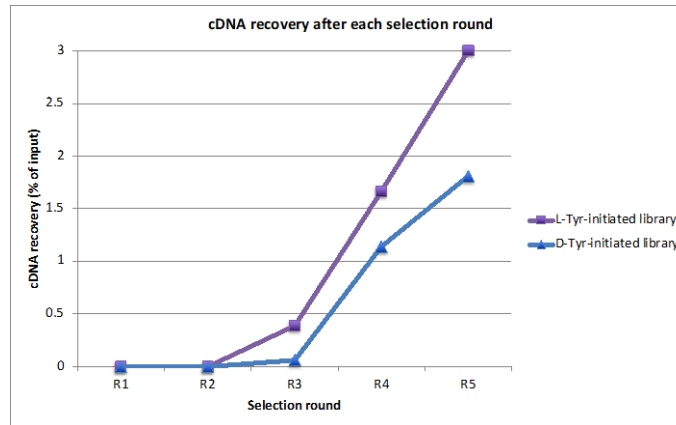
**A**



**B**



**C**



**Figure S1:** Sequencing results of initial RaPID screen against ncOGT. **A**, Enriched sequences obtained after five rounds of selection using an N-chloroacetyl-L-tyrosine initiator library. **B**, Enriched sequences obtained after five rounds of selection using an N-chloroacetyl-D-tyrosine initiator library. The 100 most frequent sequences were ranked according to relative enrichment and binned into families based on sequence similarity. Red arrows indicate sequences which were selected for synthesis and confirmation using *in vitro* assays. **C**, Recovery of cDNA as a percentage of input over each round of selection indicates significant enrichment is achieved after five round of selection.

**A) Sequences:**



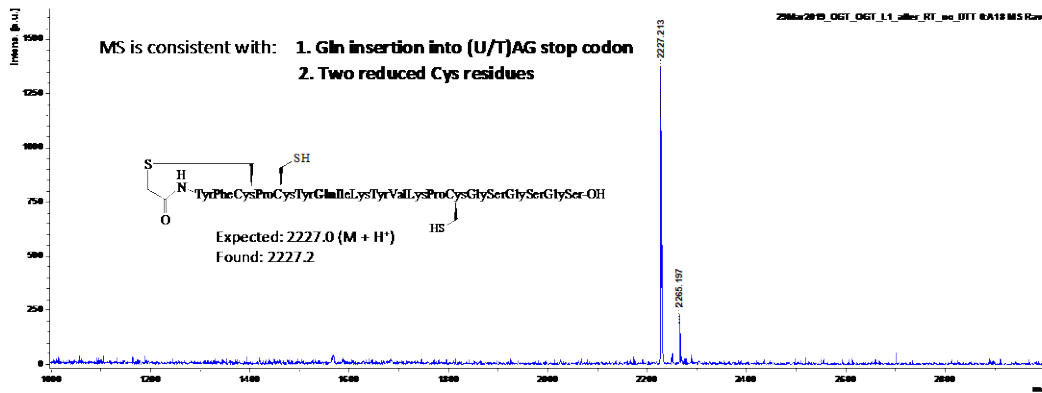
**L1 DNA for translation:**

TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAGATATACATATGTTTGTCCGTGTATATAGATTAAGTATGTGAAGCCTTGC TAAGACGGGGGGCGGAAA

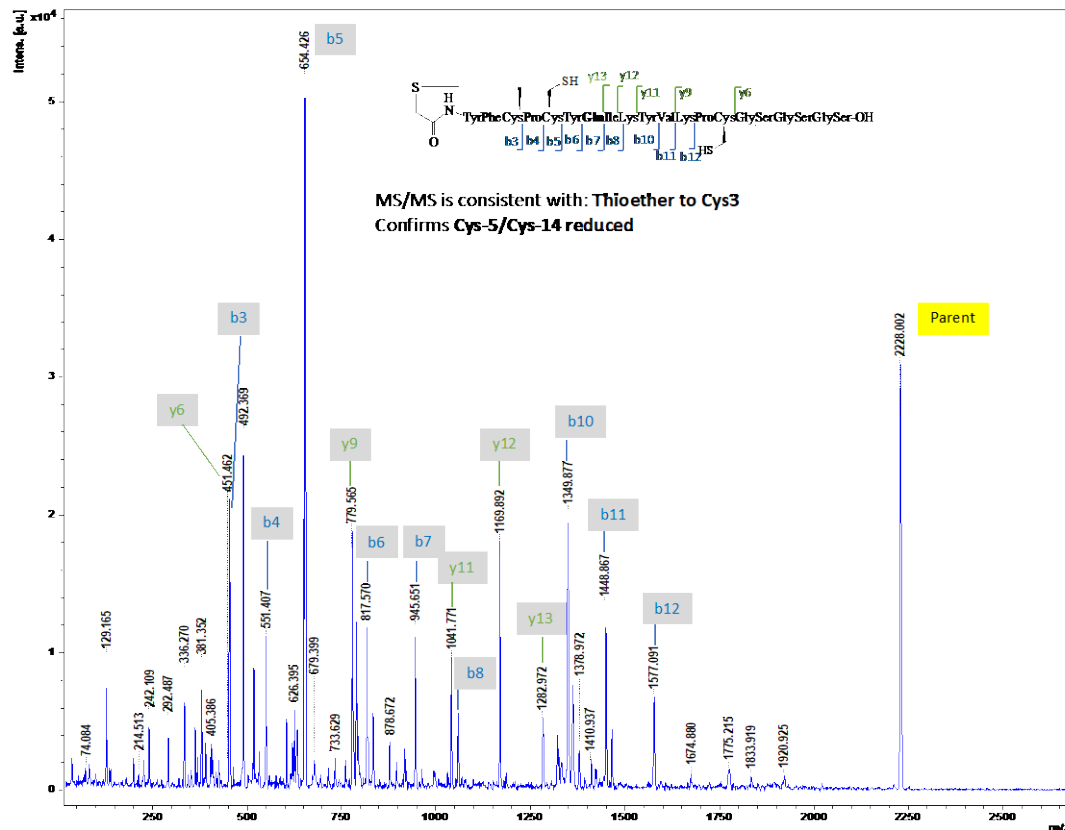
**B) Work-flow:**

1. *In vitro* translation (30 min, 37 °C, then min, 25 °C)
2. Cyclisation (10 mM EDTA, 30 min, 37 °C)
3. Submit to reverse transcription conditions (no enzyme): 1 h, 42 °C
4. Desalting, MALDI MS and MALDI MS/MS

**C) MALDI MS:**

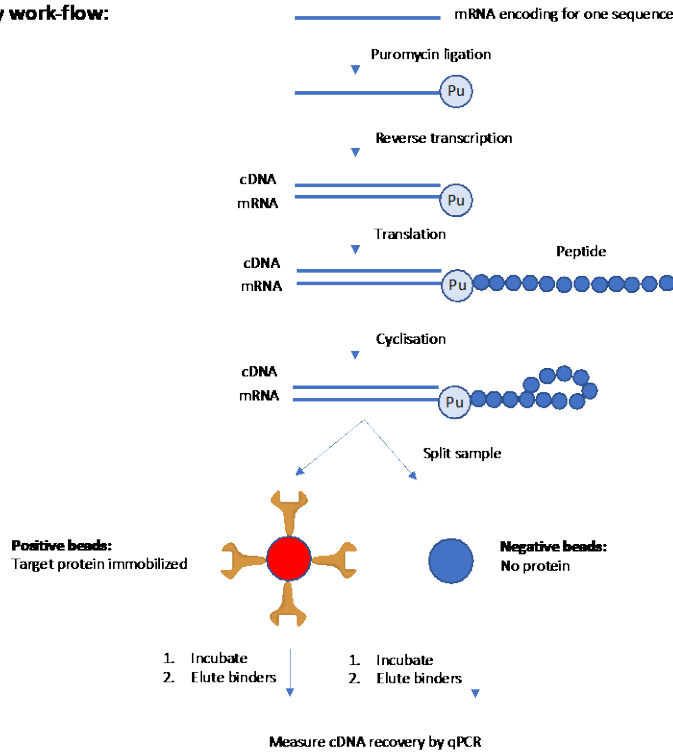


**D) MALDI MS/MS:**



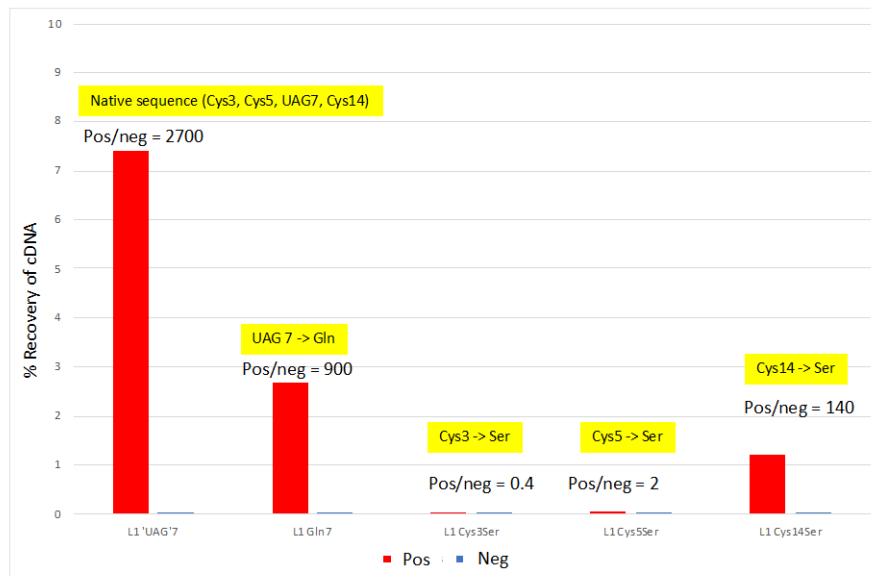
**Figure S2:** Elucidation of L1 structure using MALDI mass spectrometry. A) Amino acid sequence of L1 peptide before cyclization, and its corresponding DNA for *in vitro* translation with the peptide coding region underlined (ClAc = N-chloroacetyl). B) Workflow for *in vitro* translation of L1 peptide and analysis via MALDI MS. C) MALDI spectrum obtained after *in vitro* translation of L1 peptide and application of reverse transcription conditions (to replicate the state of the peptide in the finalized display library). No enzyme was added in the reverse transcription step, as this is unlikely to affect the state of the peptide. The observed mass is consistent with the L1 sequence in which a glutamine residue is inserted at position 7, and two reduced cysteine residues are present. D) Result of MALDI MS/MS on the same sample as above. The fragmentation pattern is consistent with a thioether ring to cysteine 3 (i.e. lariat geometry). The fragmentation pattern also confirms the presence of two reduced cysteine residues.

**A) Clone assay work-flow:**



**B) Clone assay results:**

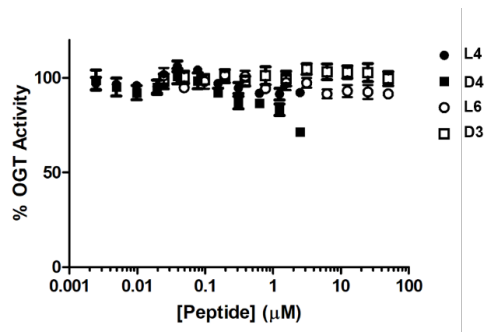
L1 = CIAc-YFC<sup>3</sup>PC<sup>5</sup>Y<sup>7</sup>IKYVKPC<sup>14</sup> Z = (U/T)AG stop codon



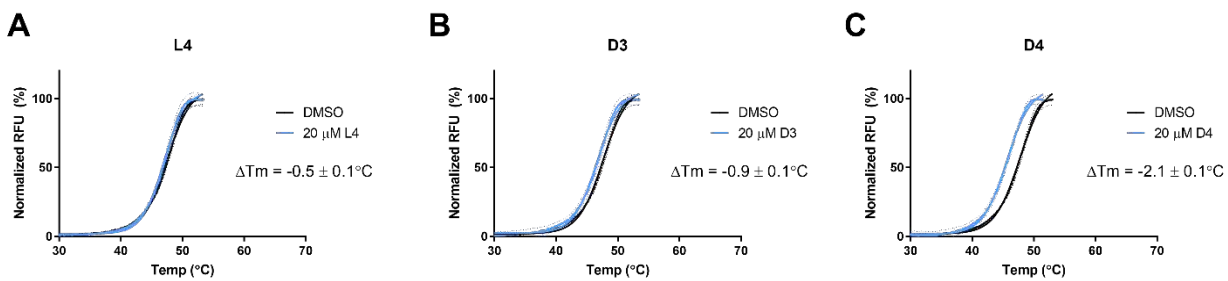
- Conclusion:**
- \* Explicit replacement of UAG stop codon with Gln (position 7) generates a target-binding clone: Confirms read-through insertion of glutamine
  - \* Mutation of Cys3 generates a non-binding clone: Cys3 is required (consistent with Cys3 lariat geometry indicated by MS)
  - \* Mutation of Cys5 generates a non-binding clone: Cys-5 is required
  - \* Mutation of Cys14 generates a target-binding clone: Cys14 is not required



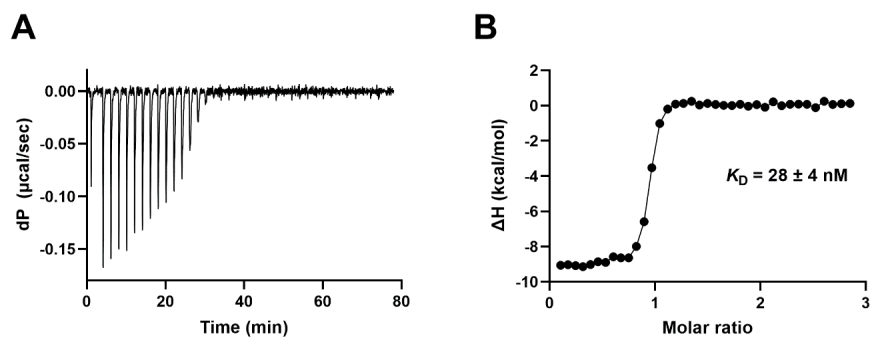
**Figure S3:** Confirmation of the sequence requirements of peptide L1 using a clone assay. A) Workflow of the clone assay. The assay is similar to a single round of RaPID selection. However, after preparation of the final display library, the sample is split equally between “positive” beads (containing immobilized target protein) and “negative” beads (naked). After eluting the binders, the amount of cDNA recovered from each condition is measured by qPCR. The positive/negative recovery ratio is used as an indication of the target binding capability of each clone. The assay is not fully quantitative; for the purposes of this experiment, a positive/negative ratio of  $> 100$  is interpreted as an indication of a target-binding clone. B) Chart of positive and negative recoveries of clones relating to the L1 sequence, and a conclusion based on the results.



**Figure S4:** 1<sup>st</sup> generation macrocycles that show the greatest potency of inhibition towards full-length OGT do not inhibit the truncated 4.5OGT construct. Assays were performed under identical conditions and concentrations as performed for full-length enzyme, except that a final concentration 20 nM 4.5OGT was used in place of 20 nM ncOGT. Error bars represent S.E.M. values of triplicate measurements.

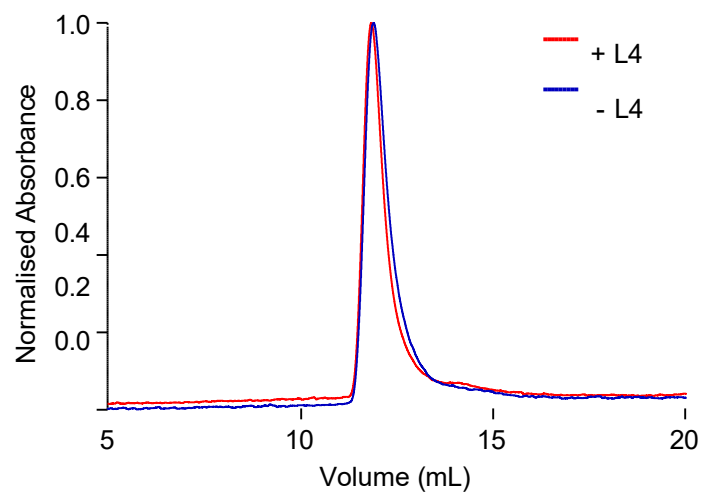


**Figure S5:** Inhibitory macrocyclic peptides do not show evidence of binding to the OGT4.5 catalytic domain construct. Melt curves were performed by DSF using identical conditions as shown in figure S3. Dashed lines indicate S.E.M. (n=3).

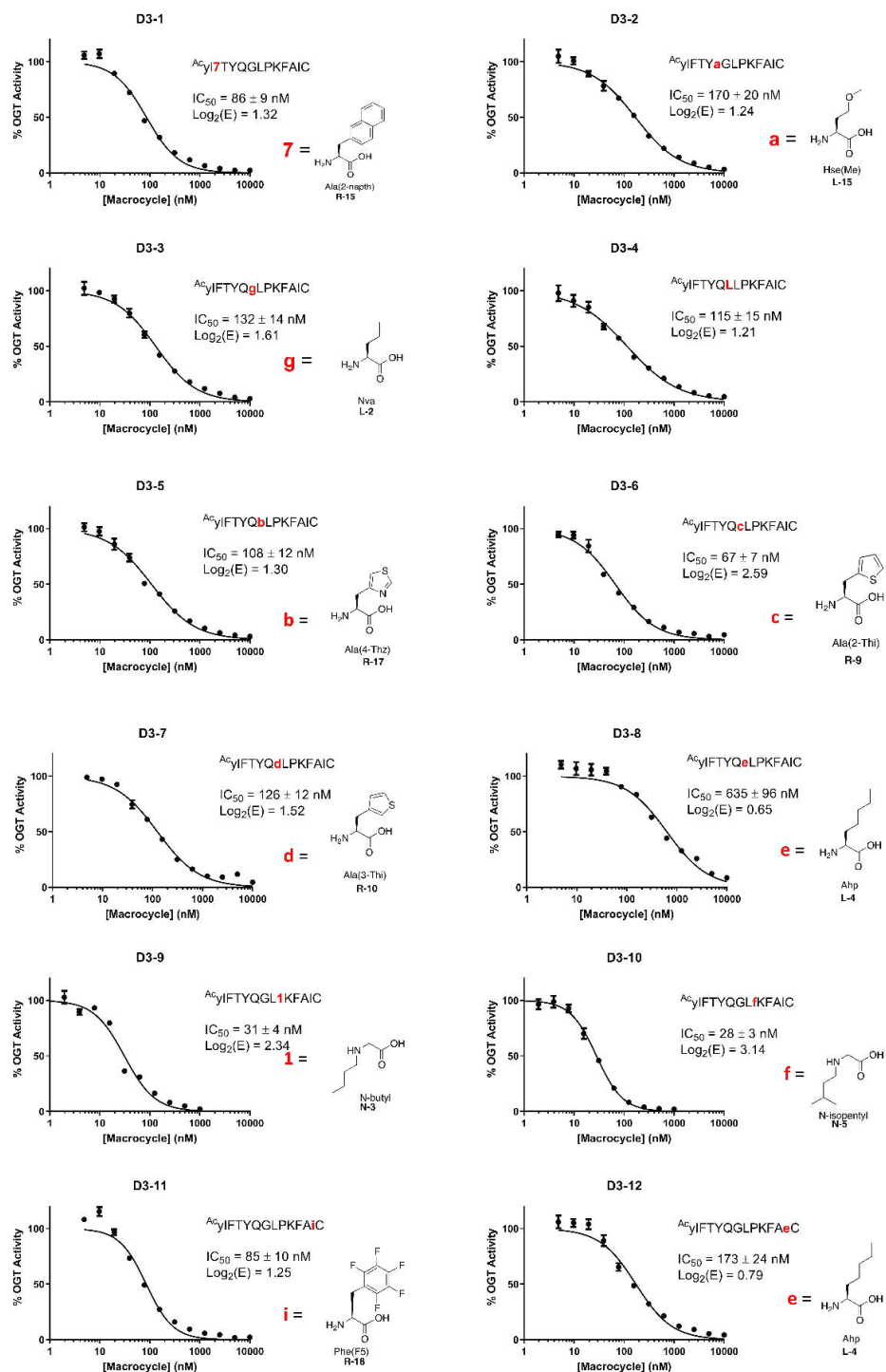


Compound	N (sites)	$K_D$ (nM)	$\Delta H$ (kcal mol <sup>-1</sup> )	$\Delta G$ (kcal mol <sup>-1</sup> )	$-T\Delta S$ (kcal mol <sup>-1</sup> )
L4 peptide	$0.85 \pm 0.05$	$28.4 \pm 3.6$	$-9.19 \pm 0.4$	$-10.3 \pm 0.08$	$-1.11 \pm 0.47$

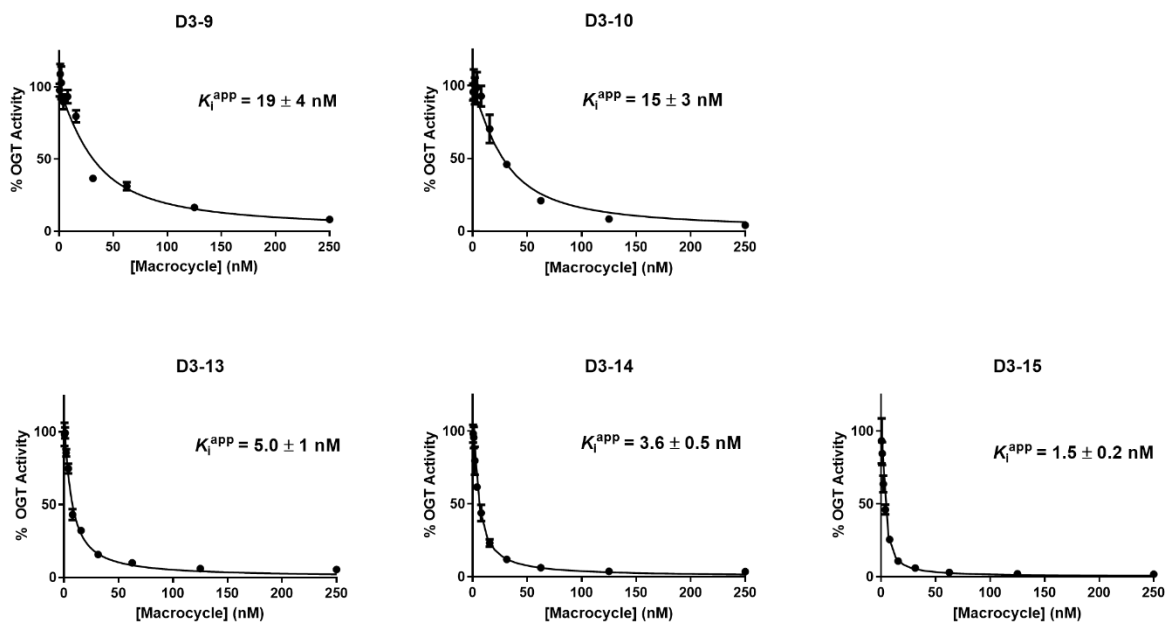
**Figure S6:** ITC analysis of the L4 peptide binding to the TPR<sub>11.5</sub> construct of OGT A, Raw injection profile. B, Titration curve with each injection used for analysis marked with a point. Titrations were carried out at 25°C using 200 µM peptide in the injection syringe and 14.3 µM protein in the sample cell. A series of 38, 1 µL injections were performed with 120 seconds between injections and a stirring speed of 750 RPM. Thermodynamic parameters were calculated using a one-site fitting model in Origin software. C, Calculated results from four independent replicates with cell concentration varying from 14.3 µM to 21 µM with fixed L4 peptide concentration. Values ± standard deviation.



**Figure S7:** A size exclusion assay performed using full-length OGT with and without L4 peptide shows that the L4 macrocyclic peptide does not affect OGT dimerization. The experiment was run in duplicate.

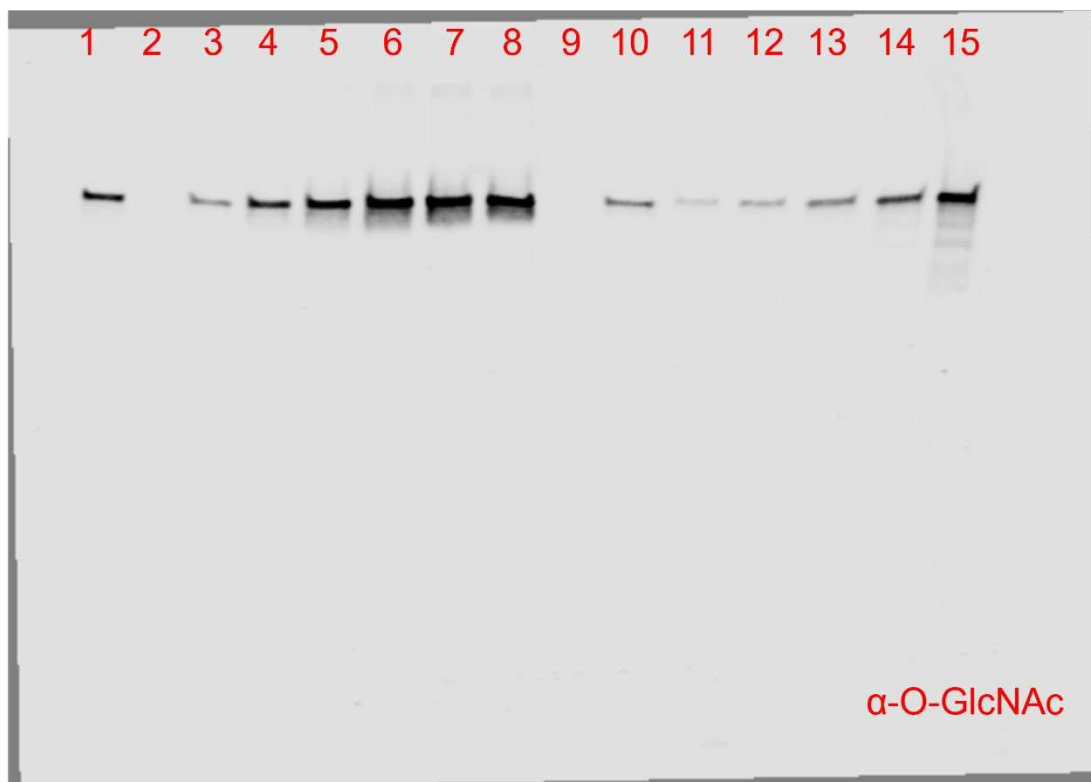


**Figure S8:** concentration-response curves and calculated  $\text{IC}_{50}$  values for 2<sup>nd</sup>-generation D3 macrocycles derived from saturation mutagenesis. Assays were performed using an *in vitro* OGT activity assay as described in the general method using 2  $\mu\text{M}$  glycosyl donor, 5  $\mu\text{M}$  peptide acceptor and 20 nM ncOGT. Curves were fitted to a log(inhibitor) vs normalized response four parameter inhibition model using GraphPad Prism 8. Data points represent mean values of four replicates per concentration; error bars indicate S.E.M.



**Figure S9:** Concentration response curves and calculated  $K_i^{app}$  values for the most potent series of N-alkyl D3 macrocycle mutants. Inhibition data was obtained as previously described, except that a final concentration of 5 nM ncOGT was used. Data points represent mean values of four replicates per concentration; error bars indicate one standard deviation.  $K_i^{app}$  values were obtained by fitting the data to Equation 1 for tight-binding inhibition derived by Morrison and Williams<sup>[1,2]</sup>:

$$\frac{v_i}{v_o} = 1 - \frac{([E] + [I] + K_i^{app}) - \sqrt{([E] + [I] + K_i^{app})^2 - 4[E][I]}}{2[E]} \quad (1)$$



**Figure S10:** Complete western blot image from OGT pull-down experiment. 10  $\mu$ L of sample was loaded in each lane. Assay was performed according to the procedure described in the Supplemental Methods section. Lane 1: Supernatant after addition of streptavidin beads to OGT and D3-16. Lane 2: empty. Lane 3: Elution fraction after incubation with 3 nM D3-16. Lane 4: Elution fraction after incubation with 13 nM D3-16. Lane 5: Elution fraction after incubation with 50 nM D3-16. Lane 6: Elution fraction after incubation with 200 nM D3-16. Lane 7: Elution fraction after incubation with 500 nM D3-16. Lane 8: Elution fraction after incubation with 500 nM D3-16 in PBS with 10% FBS (v/v). Lane 9: empty. Lane 10: Elution fraction after competition of 500 nM D3-16 with 10  $\mu$ M D3-9. Lane 11: 2.5 ng rOGT standard. Lane 12: 5 ng rOGT standard. Lane 13: 10 ng rOGT standard. Lane 14: 20 ng rOGT standard. Lane 15: 40 ng rOGT standard.



## SUPPLEMENTAL METHODS

### RaPID macrocyclic library design

A thioether-macrocyclic peptide library was constructed using *N*-(2-chloroacetyl)-L- and *N*-(2-chloroacetyl)-D-tyrosine (ClAcY or ClAcDY) as initiators in a flexible *in vitro* translation (FIT) system reaction<sup>[3]</sup>. The underlying mRNA library was designed to have an AUG (ClAcY/ClAcDY) initiator codon followed by 7–12 NNK codons (N = G, C, A or U; K = G or U), that code for random proteinogenic amino acid residues, followed by a fixed UGC codon that codes for Cys. After *in vitro* translation, a thioether bond forms spontaneously between the *N*-terminal ClAc group of the initiator residue and the sulfhydryl group of the downstream Cys residue to generate the macrocyclic peptide backbone. The theoretical diversity of the peptide library is  $> 10^{12}$ .

### Selection of macrocyclic peptides binding to OGT

Affinity selection was performed using the both the ClAcY- and ClAcDY-initiated libraries against His<sub>6</sub>-tagged full-length OGT by employing the random non-standard peptide integrated discovery (RaPID) approach<sup>[4]</sup>. The mRNA library and ClAcY-tRNA<sup>fMet</sup><sub>CAU</sub> / ClAcDY-tRNA<sup>fMet</sup><sub>CAU</sub> were prepared as previously reported<sup>[5,6]</sup>, and the RaPID selection was conducted as previously reported. Briefly, 1  $\mu$ M mRNA library was ligated to a puromycin-linked oligonucleotide (1.5  $\mu$ M) using T4 RNA ligase at 25 °C for 30 min. After purification by phenol–chloroform extraction and ethanol precipitation, 1.2  $\mu$ M mRNA-puromycin conjugate was translated at 37 °C for 30 min in a methionine-deficient FIT reaction containing 25  $\mu$ M ClAc-Y-tRNA<sup>fMet</sup><sub>CAU</sub> or 25  $\mu$ M ClAc-DY-tRNA<sup>fMet</sup><sub>CAU</sub> to generate the peptide libraries. Following translation, the libraries were incubated at 25 °C for 12 min to facilitate the formation of mRNA-peptide conjugates. Next, EDTA was added (to a final concentration of 20 mM) and the mixture was incubated at 37 °C for 30 min to remove the mRNA-peptide conjugates from the ribosomes. The product was subsequently reverse-transcribed using RNase H- reverse transcriptase (Promega) at 42 °C for 1 h to yield peptide–mRNA/cDNA conjugates. The peptide–mRNA/cDNA conjugates were then buffer-exchanged into selection buffer (phosphate buffered saline, pH 7.4, containing 0.05 % (v/v) tween-20 and 1 mM dithiothreitol) using a short column of Sephadex G25 (GE Healthcare). Acetylated bovine serum albumin was added (to a final concentration of 1 mg/mL) and the solution was incubated in the presence of Dynabeads His-Tag Isolation to remove His-tagged recombinant proteins and bead-binding macrocyclic peptides. This yielded the final library used for affinity selection.

For affinity selection to OGT, the library solution was incubated with 200 nM His<sub>6</sub>-tagged full-length OGT (immobilized on Dynabeads His-Tag Isolation) for 30 min at 4 °C to isolate OGT binders. The beads were washed with ice-cold selection buffer and the peptide–mRNA/cDNA binders were isolated from the beads by incubating in 1 × *Taq* buffer (10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, pH 8.3) for 5 min at 95 °C. The amount of eluted cDNAs was measured by quantitative PCR. The remaining cDNAs were amplified by PCR and then purified and transcribed to produce an enriched mRNA library for the next round of selection. The selection was repeated for five rounds, and after the fifth round the recovered cDNAs were sequenced by MiSeq (Illumina). Single read 150 bp was sufficient to cover the NNK region. DNA reads with average Phred33+ score > 25 in the NNK region were translated and unique peptide sequences were tallied using custom python scripts.

### **Mass spectrometry analysis of translated peptides**

In vitro translation was performed in a methionine-deficient FIT reaction supplemented with initiator tRNA, as described above. Following translation, EDTA was added (to a final concentration of 10 mM), and the reaction was submitted to the conditions of cyclization and reverse transcription, as described above. However, no reverse transcriptase enzyme was added in the reverse transcription step. The mixture was desalted using C18 C-tips (washing with 4% ACN + 0.5% acetic acid; eluting with a 50% saturated solution of  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) matrix in 80% ACN + 0.5% acetic acid) and submitted to MALDI MS and MALDI MS/MS analysis (Bruker Autoflex II).

### **Determination of L1 peptide structure**

Sequence L1 originally contained three cysteine residues and an internal stop codon. Using MALDI mass spectrometry analysis, we determined that the thioether macrocycle was exclusively forming with the first cysteine residue (position 3) (Figure S2). We also observed that the internal stop codon (originally at position 7) was being read through with the insertion of a glutamine residue. These results were confirmed with a clone assay analysis (Figure S3). Furthermore, we determined that one of the remaining cysteine residues (originally at position 14) could be substituted with serine without destroying target binding activity, which improved the

ease of synthesis (Figure S3). These observations were combined to yield the final L1 sequence, shown in Figure 1D (main text).

### **Saturation mutagenesis**

The saturation mutagenesis experiment was conducted as previously reported<sup>[7]</sup>. Briefly, the site saturation DNA library was prepared so that each codon of the D3 sequence was changed to a single NNK codon (to allow for the sampling of all natural amino acids) and a single ATG codon (to allow for the introduction of non-natural amino acids). The non-natural amino acids were charged on to tRNA<sup>EnGlu</sup><sub>CAU</sub> under flexizyme catalysis, using conditions and activated precursors that have been reported previously<sup>[8]</sup>. 22 small-scale translation reactions were conducted, each incorporating a different non-natural amino acid at the AUG codon. After translation, the identity of the non-natural amino acids was encoded using unique barcoding reverse primers (detailed in reference<sup>[7]</sup>). The resulting peptide–mRNA/cDNA conjugates were combined in to a single sample, and exchanged in to selection buffer, as detailed in the ‘**Selection of macrocyclic peptides**’ section above. The library was incubated with 200 nM His<sub>6</sub>-tagged full-length OGT (immobilized on Dynabeads His-Tag Isolation) for 2 h at 25 °C. The beads were washed twice with selection buffer, each wash was performed for 2 h at 25 °C. A ‘negative’ condition was performed to confirm that the data obtained from the experiment was genuine. In the ‘negative’ condition no OGT was immobilized on the beads, however the experiment was identical in every other way. The peptide–mRNA/cDNA binders were isolated from the beads, as detailed above, and the recovered cDNAs were sequenced by MiSeq (Illumina). The calculation of the relative enrichment score (*E*) of each mutant sequence was performed as previously described<sup>[7]</sup>.

### **General method for synthesis of macrocyclic peptides**

Macrocyclic peptides were chemically synthesized using a Syro automated peptide synthesizer (Biotage) and standard Fmoc solid-phase peptide synthesis techniques at 25 μmol scale on NovaPEG rink amide resin (0.44 mmol/g substitution), as previously described<sup>[5,6]</sup>. Standard couplings were performed using HBTU/HOBt chemistry, as previously described. Couplings occurring directly after an N-alkylated residue were performed using the following method: (1) Oxyma (28.3 mg, 8 eq), DIC (31.2 μL, 8 eq) and Fmoc-AA (8 eq) were combined in 1 mL DMF and allowed to incubate at room temperature for 10 min.

- (2) This activated mixture was transferred to the waiting peptide synthesis resin.
- (3) The coupling reaction was conducted at 65 °C for 25 min, with shaking.
- (4) The mixture was drained from the resin and the coupling was repeated again from step (1). Post-coupling washing and deprotection steps were performed in the same manner as for the standard couplings<sup>[6]</sup>.

Following synthesis, the chloroacetyl group was installed on the *N*-terminal amine group of the peptides by treating the resin with 0.15 M chloroacetic acid, 0.15 M 1-hydroxybenzotriazole hydrate and 0.15 M *N,N'*-diisopropylcarbodiimide (final concentrations) in *N,N*-dimethylformamide for 1 h. The peptides were deprotected and cleaved from the resin using a solution of 92.5% (v/v) trifluoroacetic acid, 2.5% (v/v) water, 2.5% (v/v) triisopropylsilane and 2.5% (v/v) 2,2-(ethylenedioxy)diethanethiol and then precipitated by the addition of ice-cold diethyl ether. To conduct the cyclization reaction, the peptide pellet was dissolved in 10 mL 4 : 1 DMSO/H<sub>2</sub>O, adjusted to pH > 8 by addition of triethylamine (20 – 40 μL) and incubated for 1 h at 42 °C. The progress of the cyclization reaction was monitored by MALDI-TOF mass spectrometry ( $\alpha$ -cyano-4-hydroxycinnamic acid matrix). After completion of cyclization, the reactions were quenched by acidification with trifluoroacetic acid (equal volume to that of triethylamine). The peptide solutions were directly purified by reverse-phase HPLC (Shimadzu Prominence LC-20AP with a Merck Chromolith Prep column, 200 × 25 mm i.d.) using a gradient of acetonitrile in water supplemented with 0.1% (v/v) trifluoroacetic acid. The fractions containing product were identified using MALDI-TOF mass spectrometry and lyophilized. The purity of the peptides was confirmed by analytical ultra high-pressure liquid chromatography using a Nexera X2 system (Shimadzu) fitted with a C18 reversed phase column. Approximately 1 nmol of peptide was separated using a 20 – 70% (v/v) aqueous acetonitrile gradient supplemented with 0.1% (v/v) trifluoroacetic acid and monitored by absorbance at 280 nm.

### **Synthesis of peptide L1**

Di-cysteine peptide L1 was synthesized as above, however low loading resin (CEM Rink Amide ProTide 0.18 mmol/g) was employed. Monomethoxytrityl (Mmt)-protected Cys was used in position 3 and diphenylmethyl (Dpm)-protected Cys was used in position 5. Chloroacetylation was performed as above. The resin was washed with DMF and DCM, then the Mmt protecting group was selectively cleaved using 5% (v/v) trifluoroacetic acid, 2.5% (v/v) triisopropylsilane in DCM (4

x 1 mL, 15 min each). The resin was washed with DCM, then DMF. The cyclization reaction was conducted on-resin by treating the resin with 5% v/v *N,N*-diisopropylethylamine, 0.5 M LiCl in DMF for 3 h at 37 °C. The peptide was cleaved from the resin, fully deprotected, and purified as above.

### **Installation of N-alkyl glycine residues**

The N-alkyl glycine amino acids (in peptides **D3-8**, **D3-9**, **D3-13**, **D3-14**, **D3-15**) were installed using the sub-monomer approach<sup>[9]</sup> as follows:

- (1) 1 mL of 0.6 M bromoacetic acid in DMF, and 86  $\mu$ L of DIC were added to the waiting peptide synthesis resin.
- (2) The resin was incubated at 30 °C for 30 min with shaking, then drained.
- (3) The coupling was repeated from step (1)
- (4) The resin was washed 5x with DMF
- (5) 1 mL of 1M alkylamine (see below) was added to the resin.
- (6) The resin was incubated at 30 °C for 1 h with shaking, then drained.
- (7) The coupling was repeated from step (5)
- (8) The resin was washed 5x with DMF, and submitted to the next step of the peptide synthesis cycle.

#### *Alkylamine solutions:*

N-butylamine, N-isopentylamine and N-hexylamine were used as solutions in N-methyl-2-pyrrolidone. N-Cyclohexylethanamine and N-cyclopentylethanamine were derived from their corresponding HCl salts<sup>[10]</sup> directly before use, as follows: 0.95 eq of 11 M aqueous KOH was added to a 1 M suspension of the alkylamine HCl salt in DMSO. The mixture was thoroughly combined by vortex, then the supernatant (solution of free alkylamine) was isolated from the precipitate by centrifugation.

### **Preparation of cysteamine-terminal peptides**

The cysteamine-terminal peptides **D3-13**, **D3-14** and **D3-15** were synthesized as above, however cysteamine 2-chlorotrityl resin (Sigma Aldrich) was used in place of rink amide resin.

## Characterization data of synthesized macrocyclic peptides

- L1:** Purity 94% (UPLC), retention time 12.5 min, calculated mass 1777.8 [M+H]<sup>+</sup>, found 1777.5.
- L2:** Purity 60% (UPLC), retention time 13.5 min, calculated mass 1587.7 [M+H]<sup>+</sup>, found 1587.6.
- L3:** Purity 97% (UPLC), retention time 12.8 min, calculated mass 1770.8 [M+H]<sup>+</sup>, found 1771.0.
- L4:** Purity 95% (UPLC), retention time 10.1 min, calculated 1913.2 [M+H]<sup>+</sup>, found 1913.1.
- L5:** Purity 95% (UPLC), retention time 12.5 min, calculated mass 1923.0 [M+H]<sup>+</sup>, found 1923.2.
- L6:** Purity 89% (UPLC), retention time 10.8 min, calculated mass 1510.7 [M+H]<sup>+</sup>, found 1511.2.
- D1:** Purity 95% (UPLC), retention time 14.4 min, calculated mass 1818.8 [M+H]<sup>+</sup>, found 1819.4.
- D2:** Purity 75% (UPLC), retention time 11.1 min, calculated mass 1766.9 [M+H]<sup>+</sup>, found 1767.3.
- D3:** Purity 97% (UPLC), retention time 15.0 min, calculated mass 1703.0 [M+H]<sup>+</sup>, found 1703.3.
- D4:** Purity 94% (UPLC), retention time 10.7 min, calculated mass 1934.1 [M+H]<sup>+</sup>, found 1933.5.
- D5:** Purity 92% (UPLC), retention time 13.9 min, calculated mass 1754.9 [M+H]<sup>+</sup>, found 1755.4.
- D3-1:** Purity 93% (UPLC), retention time 10.2 min, calculated mass 1752.9 [M+H]<sup>+</sup>, found 1752.6.
- D3-2:** Purity 91% (UPLC), retention time 10.3 min, calculated mass 1689.9 [M+H]<sup>+</sup>, found 1689.3.
- D3-3:** Purity 86% (UPLC), retention time 10.4 min, calculated mass 1744.9 [M+H]<sup>+</sup>, found 1744.4.
- D3-4:** Purity 82% (UPLC), retention time 10.7 min, calculated mass 1758.9 [M+H]<sup>+</sup>, found 1758.3.
- D3-5:** Purity 85% (UPLC), retention time 9.8 min, calculated mass 1799.9 [M+H]<sup>+</sup>, found 1799.3.
- D3-6:** Purity 94% (UPLC), retention time 10.4 min, calculated mass 1798.9 [M+H]<sup>+</sup>, found 1798.3.
- D3-7:** Purity 91% (UPLC), retention time 10.6 min, calculated mass 1798.9 [M+H]<sup>+</sup>, found 1798.2.
- D3-8:** Purity 91% (UPLC), retention time 11.3 min, calculated mass 1772.9 [M+H]<sup>+</sup>, found 1772.4.
- D3-9:** Purity 82% (UPLC), retention time 10.6 min, calculated mass 1718.9 [M+H]<sup>+</sup>, found 1718.5.
- D3-10:** Purity 85% (UPLC), retention time 11.1 min, calculated mass 1732.9 [M+H]<sup>+</sup>, found 1732.4.
- D3-11:** Purity 90% (UPLC), retention time 10.5 min, calculated mass 1826.8 [M+H]<sup>+</sup>, found 1826.3.

**D3-12:** Purity 91% (UPLC), retention time 10.1 min, calculated mass 1716.9 [M+H]<sup>+</sup>, found 1716.3.

**D3-13:** Purity 92% (UPLC), retention time 12.4 min, calculated mass 1703.9 [M+H]<sup>+</sup>, found 1704.3.

**D3-14:** Purity 91% (UPLC), retention time 12.9 min, calculated mass 1729.9 [M+H]<sup>+</sup>, found 1730.3.

**D3-15:** Purity 93% (UPLC), retention time 12.4 min, calculated mass 1715.9 [M+H]<sup>+</sup>, found 1716.4.

### **Full-length OGT expression and purification**

Preparation of ncOGT was performed as previously described<sup>[11]</sup>. Briefly, the pET28a vector (isoform 1, UNIPROT identifier O15294-3) encoding the human OGT gene was transformed into competent *Escherichia coli* BL21 (DE3) cells (Invitrogen). Successful transformants were cultured in Terrific Broth (TB) supplemented with 50 µg/mL kanamycin at 37°C until an optical density of 1.8 absorbance units was reached. Protein expression was induced with 0.2 mM isopropyl β-D-thiogalactoside (IPTG) at 16°C for 18 h. Cells were harvested and resuspended in 2 mL BugBuster protein extraction reagent (EMD Millipore) per gram of cell pellet in the presence of 1 mg/mL lysozyme, 0.2 mg/mL DNase and an EDTA-free protease inhibitor tablet (Roche). The mixture was gently rocked at 4°C for 30 min and then clarified by centrifugation for 2 x 30 min at 18,000 x *g*. The cell lysate was then applied to a 1 mL HisTrap nickel column (GE Healthcare) which was pre-equilibrated in 50 mM HEPES, 300 mM NaCl, 20 mM imidazole and 1 mM DTT at pH 7.5. The protein was purified using an imidazole gradient of 20-500 mM over 20 column volumes on an AKTA FPLC (GE Healthcare). Fractions judged as pure were pooled and dialyzed at 4 °C in 50 mM HEPES buffer containing 500 mM NaCl and 1 mM DTT at pH 7.4. Aliquots of OGT were flash frozen and stored at -80°C for future use.

### **Expression and purification of truncated OGT constructs**

A modified pET-YSBLIC-3C plasmid encoding TPR<sub>11.5</sub> (residues 15-400, beginning A<sup>14</sup>ELAH, UNIPROT entry: O15294-3)<sup>[12]</sup> with a 3C protease cleavable N-terminal 6xHis-tag was transformed into *E. coli* BL21(DE3) cells for expression<sup>[13]</sup>. Cells were grown at 37°C in 4 litres of

TB media supplemented with 35 µg/ml kanamycin until an OD<sub>600</sub> of 1.6-2.0 was reached. Expression was induced with IPTG to a final concentration of 0.4 mM and cultures were shaken overnight at 16 °C. Cells were pelleted at 5500 g for 20 mins and stored at -20 °C. Pellets were resuspended in buffer A (20 mM Tris, 500 mM NaCl and 40 mM imidazole, pH 8.0), lysozyme, DNase I and 1 mM AEBSF before being lysed by cell disruptor. Cell debris was removed by centrifugation at 41700 g for 1 hour at 4 °C. Supernatant was loaded onto a 5 ml HisTrap FF column (GE Healthcare) pre-equilibrated in buffer A. Column was washed with 10 column volumes of buffer A before protein was eluted through a linear gradient of buffer B (20 mM Tris, 500 mM NaCl and 500 mM imidazole, pH 8.0) in buffer A. Fractions containing TPR<sub>11.5</sub> were pooled and supplemented with DTT to a final concentration of 2 mM. The 6xHis-tag was removed through incubation with HRV 3C protease at a ratio of 1 µg protease:100 µg protein overnight at 4°C. The protein sample was diluted with 20 mM Tris pH 8.0 to a final imidazole concentration of 50 mM before being passed through a 5 ml HisTrap FF column pre-equilibrated in buffer A. Flowthrough was collected and concentrated to 2 ml using 10K MWCO Vivaspin® centrifugal concentrator (Sartorius). Protein was further purified by size-exclusion using a 16/600 Superdex 200 column (GE Healthcare) in 20 mM Na-HEPES, 200 mM NaCl and 2 mM DTT, pH7.5. Fractions containing purified protein (as assessed by SDS-PAGE) were pooled, concentrated as before, snapfrozen in liquid nitrogen, and stored at -70 °C.

OGT4.5 was purified as previously described<sup>[11]</sup>. Briefly, OGT4.5 with a 3C-protease cleavable N-terminal HisTag was expressed in *E. coli* BL21(DE3) cells. Cells were grown at 37 °C in 6 litres of TB media supplemented with 35 µg/ml kanamycin until an OD<sub>600</sub> of 1.0-1.2 was reached. Cultures were cooled at 16 °C for 30 mins before expression was induced with a final concentration of 0.2 mM IPTG and shaken overnight at 16 °C. Cells were pelleted at 5500 g for 20 mins and stored at -20 °C. Pellets were resuspended in buffer C (20 mM Tris pH 7.4, 150 mM NaCl, and 40 mM imidazole), lysozyme, DNase I, and 1 mM AEBSF. Resuspended cells were lysed by cell disruptor and insoluble material was removed by centrifugation at 41700 g for 1 hour at 4 °C. Supernatant was loaded onto a 5 ml HisTrap FF column (GE Healthcare) pre-equilibrated in buffer C, Column was washed with 10 column volumes of buffer C before protein was eluted by a stepwise gradient of buffer D (20 mM Tris 7.4, 150 mM NaCl, and 400 mM imidazole) in buffer C. Fractions containing OGT4.5 were pooled and supplemented with THP to a final concentration of 1 mM. His-tag was removed by incubation of protein with HRV 3C protease at a ratio of 1 µg protease:100 µg protein overnight at 4°C. The protein solution was diluted to a final NaCl concentration of 15 mM with 20 mM Tris pH 7.4 and applied to a 5 ml HisTrap FF column (GE Healthcare) followed by a 5 ml HiTrap Q HP column (GE Healthcare) pre-equilibrated in 20



mM Tris and 40 mM imidazole, pH 8.0. OGT4.5 was eluted from the 5 ml HiTrap Q HP column with a linear gradient of 20 mM Tris pH 8.0 through to 20mM Tris pH 8.0 and 1M NaCl. Fractions containing purified protein were pooled, concentrated with 10K MWCO Vivaspin® spin-concentrators (Sartorius) to 2 ml. Concentrated protein was size excluded on a 16/600 Superdex 200 column (GE Healthcare) in 20 mM Tris pH 8.0 and 150 mM NaCl. Purified OGT4.5 was pooled, supplemented with THP to 1 mM, concentrated as before and snap frozen. Sample purity was assessed by SDS-PAGE analysis. OGT4.5 was stored at -70 °C until use.

### **General method for *in vitro* OGT activity assays**

Kinetic assays to assess inhibition of macrocycles towards recombinant OGT were performed as previously described with minor modifications<sup>[11,14]</sup>. All assays were carried out in black 384-well microplates (Nunc 262260, Thermo Fisher Scientific, Mississauga, ON) at a 25 µL well volume using PBS pH 7.4 supplemented with 12.5 mM MgCl<sub>2</sub> and 1 mM DTT as the reaction buffer. In a typical experiment, stocks of macrocyclic peptide in DMSO were diluted to various concentrations and were pre-incubated with 20 nM OGT on ice for 15 min. The reaction was commenced by addition of a master mix containing both glycosyl donor (UDP-GlcN-BODIPY) and acceptor substrate (biotinylated HCF-serine peptide). The plate was then incubated at ambient temperature for up to 60 min, during which time the reaction rate was shown to be linear. The reaction is terminated by the addition of 25 µL stop mix containing 2 mM UDP and either 0.2 mg/mL streptavidin-coated magnetic beads (Trilink Technologies M-1002). The plate was incubated at room temperature for at least 30 min to allow for binding to resin, and is then subjected to an automated washing procedure using a BioTek EL406 plate washer containing a magnetic adapter and a BioTek 384F magnet (6 wash cycles, 100 µL PBS dispensed per well, 6 mm aspiration height offset, 4 min initial resting time on magnet, 1 min rest between washes). At the end of the wash cycle, a final dispense of 50 µL PBS per well is performed using the syringe dispenser of the BioTek washer. The fluorescence signal is read using a Biotek Neo2 multimode plate reader using 490nm excitation and 525 nm emission wavelengths. Each well is read using an area scan function which records the mean signal of 9 individual scans in a grid pattern with 1.0 micron spacing between each point. Data were analyzed and plotted using GraphPad Prism 8.

### **Determination of inhibition constants**

All  $IC_{50}$  experiments were performed under balanced conditions using donor and acceptor substrate concentrations equal to their determined  $K_m$  values (2  $\mu$ M UDP-GlcN-BODIPY, 5  $\mu$ M biotin-HCF1 peptide). Reactions were performed at room temperature for 60 min and were then stopped, washed, and read as described above. Titrations of at least 10 different inhibitor concentrations, as well as a DMSO control sample, were pre-incubated with enzyme for 15 min prior to addition of substrate. Concentrations of DMSO were kept constant throughout all wells. The percent normalized activity of each sample was calculated as a ratio of the fluorescence values in the positive controls (DMSO, 100% activity) versus the negative controls (2 mM UDP, 0% activity). The data were fitted using a sigmoidal 4-parameter log(inhibitor) vs response function in GraphPad Prism 8;  $IC_{50}$  values were calculated from the inflection point of the sigmoidal curves.  $K_i$  values for L4, D3 and D4 were assessed by performing experiments in a similar manner, except that the concentration of donor substrate was varied using concentrations at least 5-fold above and 5-fold below the  $K_{M(app)}$  in the presence of varying concentrations of inhibitor.  $K_i$  values were calculated using the non-linear regression function of GraphPad Prism 8 by fitting the data to a mixed-model inhibition function. For all inhibitors with  $IC_{50}$  values < 5-fold above the concentration of enzyme,  $K_i$  values were determined by fitting the  $IC_{50}$  data to the tight-binding inhibition model for a non-competitive inhibitor derived by Morrison<sup>[1]</sup>.

### **Size Exclusion Assay**

A 40  $\mu$ l solution of 26  $\mu$ M full-length OGT  $\pm$  260  $\mu$ M L4 peptide was incubated at ambient temperature for 20 mins. Solutions were applied to a Superdex 200 Increase 10/300 GL column pre-equilibrated in 25 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT using an ÄKTA Pure system. Sample was run at 0.5 mL/min and protein elution was monitored at A280. Experiment was run in duplicate.

### **Differential Scanning Fluorimetry**

Differential Scanning Fluorimetry was performed using an Applied Biosystems QuantStudio 3 RT-PCR machine equipped with a 96-well heating block. Reactions were performed in MicroAmp 96-well optical PCR plates (Applied Biosystems) at a final volume of 20  $\mu$ L per well. Reactions contained 0.06 mg/mL protein, 1X Protein Thermal Shift Dye (Thermo Fisher Scientific) and 20  $\mu$ M final concentration of ligand in 1X PBS. Reaction components were made up on ice; thermal shift assays were carried out at temperatures ranging from 25 to 95°C at a ramp rate of 0.05°C/sec. Fluorescence measurements were taken using a 580 nm excitation / 623 nm emission filter set. Reactions were performed in triplicate and data were analyzed using

QuantStudio Design and Analysis software and GraphPad Prism 8. Melt curves were fitted to a Boltzmann sigmoidal curve and apparent  $T_m$  values were determined by calculating the inflection point of the fitted curves.

### **Isothermal Titration Calorimetry**

ITC experiments were carried out using a Malvern MicroCal Auto-ITC<sub>200</sub> (Malvern, UK). Prior to measurement, recombinant TPR<sub>11.5</sub> was exchanged into ITC buffer (20 mM Na-HEPES pH 7.4, 200 mM NaCl and 1% DMSO) using 7K MWCO Zeba™ spin Desalting Columns as per manufacturers instructions. The L4 peptide was dissolved in DMSO and diluted with ITC buffer lacking DMSO to a final concentration of 200  $\mu$ M (yielding a sample with a final concentration of 1% DMSO). Concentration of TPR<sub>11.5</sub> and the L4 peptide after dilution was verified by  $A_{280}$  absorbance using extinction coefficients of 45310  $M^{-1} cm^{-1}$  and 9970  $M^{-1} cm^{-1}$ , respectively. To generate isotherms, the L4 peptide was titrated into varying concentrations of TPR<sub>11.5</sub> (14.3-21  $\mu$ M) over 38, 1  $\mu$ L injections with 120 second delays between each injection (first injection was 0.5  $\mu$ L with a 180 second spacing). Experiments were run at a temperature of 25°C with stirring at 750 rpm. Binding affinity was calculated using a one-site model in the MicroCal PEAQ-ITC analysis software. Experiment was run in quadruplicate.

### **OGT pull-down and western blotting**

Recombinant full-length human OGT (5  $\mu$ g) was added to 1 ml aliquots of pre-cooled buffer (50 mM HEPES, 300 mM NaCl, 5 mM  $MgCl_2$ , 5 mM EDTA, 10% Glycerol, 0.02% Triton X-100, pH 7.4) or 10% FBS (Life Technologies) in buffer. OGT was then incubated with various concentrations of pull-down probe D3-16 for 1 hour at 4°C on a rotator. For control samples, OGT was pre-incubated with 500 nM D3-9 peptide for 15 min at 4°C prior to addition D3-16. Following probe incubation, 40  $\mu$ L streptavidin-coated magnetic beads (NanoLINK, 10 mg/mL) were added to the mixture and incubated for 1 hour at 4°C on a rotator to enable pull-down of probe and OGT complexes. Prior to incubation, the beads were blocked for 30 min at 4°C with SuperBlock (PBS) blocking buffer (Thermo Scientific) in a ratio of 1:9 (v/v) and washed 3 times with buffer using a magnet (Stemcell, EasySep). The beads were then isolated using a magnet, the supernatant (SN) collected, and the beads were washed 5 times using buffer. The beads were then resuspended in 150  $\mu$ L buffer containing 60 mM DTT and incubated at 850 rpm for 10 minutes at RT on a thermomixer (Eppendorf) to specifically elute OGT off the beads. The elution was collected and mixed with SDS-PAGE loading buffer and heated to 95°C on a heating block (VWR) for 5 min. 5

$\mu\text{l}$  of each sample was then loaded onto and separated by SDS-PAGE (Mini-PROTEAN TGX Precast Gels, Bio-Rad) and transferred to a 0.45  $\mu\text{m}$  nitrocellulose membrane (Bio-Rad). For immunoblotting of OGT, rabbit OGT antibody (1:2000 dilution, clone D1D8Q, Cell Signaling) and IRDye 680LT goat anti-rabbit secondary antibody (1:20000 dilution, LiCor) were used for visualization of OGT bands using an Odyssey imaging system (Li-Cor).

### **Synthesis of biotinylated probe D3-16**

2 mg of D3-9-Pra and 1.2 mg disulfide-biotin-azide (Click Chemistry Tools, Scottsdale, AZ, Cat# 1168-5) were dissolved in 2 mL DMSO in a 5 mL round bottom flask under an argon atmosphere. Separately, 1.2 eq  $\text{CuSO}_4$  and 1.5 eq BTTAA (Click Chemistry Tools, Scottsdale, AZ, Cat# 1236-100) were dissolved in 2 mL water and then added to the DMSO mixture. The reaction was initiated by addition of 1.2 eq sodium ascorbate in 0.5 mL water, and the reaction was stirred for 2 hours at room temperature with constant sparging of the solvent with argon gas. The reaction mixture was then concentrated and purified on an Agilent 1200 series HPLC equipped with an Agilent XDB-C18 Eclipse reversed-phase column (9.4 x 250 mm, 5 $\mu$  particle size). The product was eluted in mobile phase containing water and acetonitrile with 0.02% formic acid, using a gradient of 35% to 50% acetonitrile over 20 minutes at a flow rate of 2 mL/min, yielding 1.2 mL D3-16. 92% purity (HPLC), retention time 15.2 min, HRMS m/z calculated 1254.1034,  $[\text{M}+2\text{H}]^{2+}$ , found 1254.1123.

## References

- [1] J. F. Morrison, *Biochimica et Biophysica Acta (BBA) - Enzymology* **1969**, *185*, 269–286.
- [2] J. W. Williams, J. F. Morrison, *Methods Enzymol* **1979**, *63*, 437–467.
- [3] Y. Goto, T. Katoh, H. Suga, *Nature Protocols* **2011**, *6*, 779–790.
- [4] T. Passioura, H. Suga, *Chemistry - A European Journal* **2013**, *19*, 6530–6536.
- [5] K. Ito, K. Sakai, Y. Suzuki, N. Ozawa, T. Hatta, T. Natsume, K. Matsumoto, H. Suga, *Nature Communications* **2015**, *6*, 1–12.
- [6] H. Yu, P. Dranchak, Z. Li, R. Macarthur, M. S. Munson, N. Mehzabeen, N. J. Baird, K. P. Battalie, D. Ross, S. Lovell, C. K. S. Carlow, H. Suga, J. Inglese, *Nature Communications* **2017**, *8*, 1–13.
- [7] J. M. Rogers, T. Passioura, H. Suga, *Proc Natl Acad Sci U S A* **2018**, *115*, 10959–10964.
- [8] T. Passioura, W. Liu, D. Dunkelmann, T. Higuchi, H. Suga, *J Am Chem Soc* **2018**, *140*, 11551–11555.
- [9] H. Tran, S. L. Gael, M. D. Connolly, R. N. Zuckermann, *J Vis Exp* **2011**, DOI 10.3791/3373.
- [10] G. M. Figliozzi, R. Goldsmith, S. C. Ng, S. C. Banville, R. N. Zuckermann, *Methods in Enzymology* **1996**, *267*, 437–447.
- [11] M. G. Alteen, C. Gros, R. W. Meek, D. A. Cardoso, J. A. Busmann, G. Sangouard, M. C. Deen, H.-Y. Tan, D. L. Shen, C. C. Russell, G. J. Davies, P. J. Robinson, A. McCluskey, D. J. Vocadlo, *Angewandte Chemie International Edition* **2020**, *132*, 9688–9696.
- [12] M. Jínek, J. Rehwinkel, B. D. Lazarus, E. Izaurralde, J. A. Hanover, E. Conti, *Nature Structural and Molecular Biology* **2004**, *11*, 1001–1007.
- [13] D. Bonsor, S. F. Butz, J. Solomons, S. Grant, I. J. S. Fairlamb, M. J. Fogg, G. Grogan, *Organic & Biomolecular Chemistry* **2006**, *4*, 1252–1260.
- [14] R. W. Meek, J. N. Blaza, J. A. Busmann, M. G. Alteen, D. J. Vocadlo, G. J. Davies, *Nature Communications* **2021**, *12*:1 **2021**, *12*, 1–10.