

C-Terminal Tag Location Hampers *in Vitro* Profiling of OGT Peptide Substrates by mRNA Display

Jie Shi,^[a, b] Suhela Sharif,^[a] Cyril Balsollier,^[a] Rob Ruijtenbeek,^[a, c] Roland J. Pieters,^[a] and Seino A. K. Jongkees^{*[a]}

O-GlcNAc transferase (OGT) is the only enzyme that catalyzes the post-translational modification of proteins at Ser/Thr with a single β -*N*-acetylglucosamine (O-GlcNAcylation). Its activity has been associated with chronic diseases such as cancer, diabetes and neurodegenerative disease. Although numerous OGT substrates have been identified, its accepted substrate scope can still be refined. We report here an attempt to better define the peptide-recognition requirements of the OGT active site by using mRNA display, taking advantage of its extremely high throughput to assess the substrate potential of a library of all possible nonamer peptides. An antibody-based selection process is described here that is able to enrich an OGT substrate peptide from such a library, but with poor absolute recovery. Following four rounds of selection for O-GlcNAcylated peptides, sequencing revealed 14 peptides containing Ser/Thr, but these were shown by luminescence-coupled assays and peptide microarray not to be OGT substrates. By contrast, subsequent testing of an N-terminal tag approach showed exemplary recovery. Our approach demonstrates the power of genetically encoded libraries for selection of peptide substrates, even from a very low initial starting abundance and under suboptimal conditions, and emphasizes the need to consider the binding biases of antibodies and both C- and N-terminal tags in profiling peptide substrates by high-throughput display.

residues with a single β -*N*-acetylglucosamine sugar moiety, analogous to protein Ser/Thr phosphorylation.^[1,2] Like the phosphorylation cycle, O-GlcNAcylation is also reversibly controlled by “writer” and “eraser” enzymes. Remarkably, and unlike phosphorylation, O-GlcNAc transferase (OGT) is the only enzyme that catalyzes the attachment of O-GlcNAc directly to its substrates, using UDP-GlcNAc as the sugar donor. Similarly, O-GlcNAcase (OGA) is the only enzyme that removes this modification from proteins. O-GlcNAcylation has been found on a diverse range of proteins, which are involved in many fundamental cellular processes such as gene transcription,^[3] signal transduction,^[4] and the cell cycle.^[5] Aberrant regulation of O-GlcNAcylation has been linked to chronic disease such as cancer,^[6] diabetes,^[7] and neurodegenerative disease.^[8] Primary therapeutic effects have been seen in cases by modulating O-GlcNAcylation levels,^[9] so further investigation of OGT's catalytic properties is essential to understand disease development as well as for the design of inhibitors.

To date, about 1734 O-GlcNAcylation sites have been recorded (<https://www.phosphosite.org>) and recent proteomics studies estimated that this number should be even higher.^[10,11] One of the fundamental questions that remains unanswered is how only one enzyme can define its substrates and modify so many of them with any specificity. Attempts to define the OGT substrate specificity through a high-density peptide microarray showed that small peptide portions can be O-GlcNAcylated to varying degrees,^[12–14] suggesting that OGT has the ability to determine its substrate from hundreds of peptides by differential binding in its active site. Analysis of many OGT peptide substrates has led to the generation of a substrate preference pattern for the amino acids around O-GlcNAc as [TS] [PT] [VT] S/I [RLV] [ASY] (from –3 to +2 with respect to O-GlcNAcylation site underlined), some of which were also supported by crystal structures.^[12] Although this motif is not able to fully reflect the OGT substrate requirements, Pro at the –2 site was proposed to be necessary to access the OGT active site by inducing an extended conformation.^[15] These data strongly suggested that OGT is able to differentiate between peptide substrates based on primary sequence. In another approach, saturation mutagenesis of a PPVSR substrate peptide was investigated using quantitative mass spectrometry on self-assembled monolayers. In this study, variations in the motif made little difference, but addition of alanine in the +2 position or an aromatic or hydrophobic amino acid in the –4 position gave increased modification.^[16]

In addition to direct recognition of primary sequence, N-terminal domain tetratricopeptide repeats (TPR) in OGT are also

O-Linked GlcNAcylation (O-GlcNAcylation) is a dynamic and ubiquitous post-translational modification of proteins at Ser/Thr

[a] Dr. J. Shi, S. Sharif, C. Balsollier, Dr. R. Ruijtenbeek, Prof. R. J. Pieters, Dr. S. A. K. Jongkees
Department of Chemical Biology and Drug Discovery
Utrecht Institute for Pharmaceutical Sciences, Utrecht University
Universiteitsweg 99, Utrecht 3584 CG (The Netherlands)
E-mail: s.a.k.jongkees@uu.nl

[b] Dr. J. Shi
Key Laboratory of Carbohydrate Chemistry & Biotechnology
Ministry of Education, School of Biotechnology
Jiangnan University
214122 Wuxi (P. R. China)

[c] Dr. R. Ruijtenbeek
PamGene International BV
's-Hertogenbosch, 5211 HH (The Netherlands)

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202000624>

© 2020 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

believed to be responsible for recognition of physiological protein substrates by mediating a variety of protein-protein interactions.^[17–19] It is likely that OGT determines its physiological protein substrates following a synergistical mode of selectivity through a combination of TPR-mediated protein-protein interaction and active site-constrained amino acid preference.^[20] Despite this, short OGT peptide substrates, which cannot make contact with the OGT TPR domain, are still ideal tools to study its inherent catalytic properties.

The OGT substrate motif as currently understood can in some cases not accurately match or predict OGT substrates,^[13,14] because microarray-based peptide display is limited by sequence diversity and is unable to generate comprehensive sequence information at the required peptide length. Messenger RNA display is a very efficient *in vitro* technique that allows selection of proteins/peptides with functions of interest from $\geq 10^{12}$ unique variants in a single experiment. It has been widely used for applications such as the identification of ligands of a protein,^[21,22] finding new and selective glycosidase inhibitors,^[23] studying protein-protein^[24] and protein-DNA interactions,^[25] and enzyme-substrate interactions.^[26,27] We therefore felt that mRNA display would be an ideal method to survey short OGT peptide substrates and comprehensively study the substrate requirement for OGT directly at the peptide level. Here we detail some challenges encountered in trying to develop such a system.

To validate the utility of mRNA display to select short peptide substrates for OGT, we performed a proof of principle experiment following the selection scheme as depicted in Figure 1. However, instead of a DNA library, a positive DNA control that encodes a peptide known to be an OGT substrate

was prepared and employed to validate and optimize the selection process. This template DNA control includes a region containing a T7 promoter, a region encoding an OGT substrate peptide (RESSYDIYRVPSQS,^[13] O-GlcNAc attachment site underlined) and a region encoding a C-terminal flexible peptide spacer (GAGAGA) that connects to the mRNA tag. The product of this was verified by *in vitro* translation to yield a product with the desired molecular weight on LC-MS (Figure S2). The control DNA template was then transcribed into mRNA using T7 RNA polymerase, attached to a puromycin ("Puro") linker with a DNA spacer using T4 RNA ligase I, and this mRNA-Puro conjugate was added to an *in vitro* translation reaction with release factors omitted, allowing the encoding mRNA to efficiently capture its translated peptide covalently through puromycin. The mRNA portion was reverse-transcribed to form an mRNA/cDNA-peptide fusion to avoid the degradation of mRNA and prevent any unfavorable RNA folding that might interfere with a future pull-down step. Importantly, the resulting cDNA can be used to acquire peptide information after completion of enrichment.

Before the OGT reaction and subsequent pull-down steps, we assessed whether the presence of nucleotides and other components from the translation and reverse transcription reaction might affect OGT activity or antibody binding, as UDP has been reported as an inhibitor of OGT.^[28] The crude product after translation and reverse transcription was directly mixed with a standard microarray OGT assay, wherein OGT activity is reflected by the O-GlcNAcylation of an immobilized peptide substrate that is detected by antibody binding. Comparing with a parallel control OGT assay, the crude translation/reverse transcription product has no detectable effect on OGT activity (Figure S3), indicating that this reaction mixture is compatible with both OGT modification of the mRNA displayed library and binding of the product by RL2 antibody. The same crude reverse transcription product was then subjected to modification by OGT, while a negative control experiment was performed in identical conditions but with OGT omitted in case of non-enzymatic modification. Because OGT itself can be O-GlcNAcylated during the reaction and might interfere with following pull-down of O-GlcNAcylated peptide, ethanol precipitation of the mRNA/cDNA-peptide conjugate was performed after completion of the OGT reaction to remove proteins and small molecules. Pull-down of the O-GlcNAcylated peptide was achieved by RL2 O-GlcNAc antibody-charged protein G Magnetic beads (Figure S4), and the cDNA was eluted and collected by transferring supernatant at 95 °C. Analysis of the eluted cDNA was performed using quantitative PCR. As shown in Figure 2, qPCR analysis indicated a higher recovery of cDNA from reaction with OGT present, as compared with the control reaction wherein OGT was absent. However, the recovery rate was much lower than expected in a positive control, as this sequence has been shown to be an efficient substrate by a microarray assay.^[13] As the O-GlcNAcylation site in this peptide is close to the C terminus, we speculated that the bulky puromycin linker and nucleic acid might still hinder access of the final fusions to the catalytic site of OGT, even though a flexible C-terminal spacer (GAGAGA) was used.

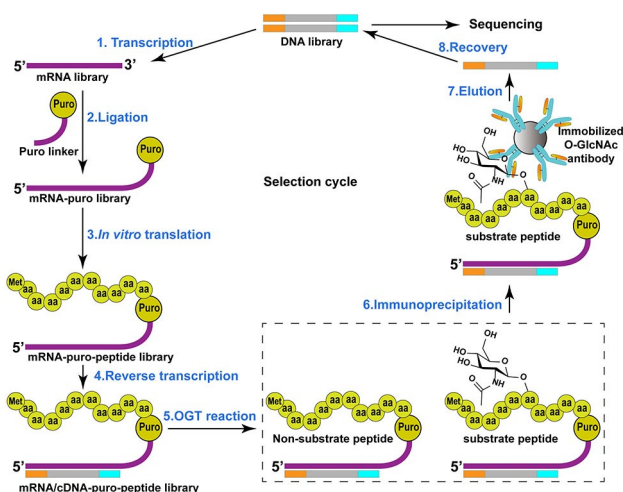


Figure 1. Schematic depiction of the selection cycle for OGT substrate peptides. 1. A DNA library or positive control template (5' and 3' UTR in orange and cyan) was transcribed into mRNA using T7 RNA polymerase. 2. Messenger RNA was ligated to a puromycin-containing linker by T4 RNA ligase I. 3. Peptide was translated and linked to its encoding RNA through the puromycin linker. 4. RNA was reverse transcribed to cDNA by reverse transcriptase. 5. mRNA/cDNA-peptide fusions were subjected to OGT reaction. 6. O-GlcNAcylated peptide-RNA conjugates were isolated by O-GlcNAc antibody pull-down assay. 7. DNA was heat eluted from the fusions. 8. Eluted DNA was amplified by PCR for sequencing or the next round of selection.

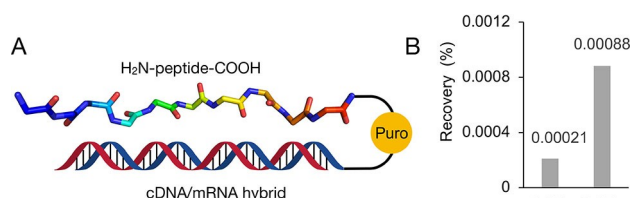


Figure 2. Recovery efficiency of the proof-of-principle selection using ZO-3 peptide. A) Illustration of the construct used, connecting the C terminus of the peptide (colored blue to red on carbons from N to C terminus) via puromycin to the mRNA strand of an mRNA/cDNA hybrid duplex. B) Pull-down of O-GlcNAcylated fraction of peptide-Puro-cDNA/mRNA conjugates with eluted cDNA quantified by qPCR. The recovery rate with and without OGT was calculated by dividing the absolute amount of DNA after pull-down by that before pull-down, using a standard curve for quantification.

Given that the OGT substrate sequence was close to the C terminus in our test substrate sequence, we decided to proceed with selection using a relatively long peptide library in the hope that a part of the library sequence would be sufficient to act as additional spacer where required. Since crystal structures of several peptide substrates in complex with OGT have suggested that the substrate specificity of OGT seems to be mainly confined to residues from -3 to $+2$ relative to the O-GlcNAcylation site,^[12] we decided to study OGT activity among a library of peptides with 9 randomized residues. To this end, the initial DNA library for this study contains a promoter region for transcription, a translation start site, a random region of 9 NNK codons (N is any nucleotide and K is guanine or thymine, giving a less biased coverage of all 20 amino acids vs NNN) encoding a 9-mer random peptide, a constant region encoding a flexible $(GA)_3$ peptide spacer then an ochre stop codon, and finally a puromycin linker annealing sequence. We did not introduce a Serine/Threonine at a fixed location in this library because both Ser and Thr appear to be common surrounding the OGT modification site, and analysis would thus not be simplified. The final DNA library was expected to express 5.12×10^9 unique peptides with sequences of $^fM-(X)_9-GAGAGA$, where X indicates any of the 20 natural amino acids and fM is formyl methionine. With $>10^{12}$ sequences examined in an mRNA display experiment, this library should be well covered (a feat that cannot be achieved by competing methodologies, including phage display). The assembly of this DNA library was achieved by two rounds of overlapping PCR, with the annealing temperature of each round of PCR optimized to avoid loss of library diversity through over-cycling (Figure S6). The DNA library was then transcribed *in vitro* to achieve the corresponding mRNA library, purified by preparative-scale urea PAGE, and further conjugated with a puromycin-containing oligonucleotide linker. *In vitro* translation of the Puro-mRNA library resulted in the mRNA-peptide conjugate library. As with the positive control, the mRNA was reverse transcribed into cDNA for stability and to minimize aptamer selection.

The successful enrichment of O-GlcNAc modified peptides greatly relies on a specific and efficient pull-down method. We have previously observed that O-GlcNAc RL2 antibody could bind to some peptides that were not O-GlcNAcylated.^[14] To

remove peptides binding directly to either RL2 or protein G, the mRNA/cDNA-peptide fusion library was purified by ethanol-precipitation and three library pre-clear steps were performed by pull-down using a mixture of 50% uncharged protein G magnetic beads and 50% RL2-antibody charged protein G magnetic beads. While this pre-clear might lead to a loss of peptide diversity, it was seen as critical to prevent enrichment of peptides binding directly to these proteins without involvement of O-GlcNAc. After washing with phosphate buffered saline containing TWEEN-20 (PBS-T), DNA from all three sets of pre-clear beads was eluted by heat treatment and analyzed by qPCR. A small amount of recovery was observed from pre-clear 1, but no further recovery was found in pre-clear 2 and pre-clear 3, suggesting any direct antibody-binding peptides were completely removed (Figure 3). The supernatant after pre-clear was then subjected to OGT reaction at room temperature overnight. After completion of the OGT reaction, the mRNA/cDNA-peptide library was again ethanol precipitated to remove O-GlcNAcylated OGT and then incubated with RL2 antibody-saturated protein G magnetic beads for 1 hour at 4°C . The supernatant was removed and the beads were stringently washed three times with ice cold PBS-T. The DNA heat-eluted from these positive selection beads was analyzed by qPCR as for the pre-clears. As shown in Figure 3, a clear increase in the recovery of DNA was achieved from the selection following OGT reaction (Round 1), when it was compared with DNA from pre-clear beads. To further confirm this selection recovery of DNA, an additional pair of parallel control selections were performed using the same mRNA-peptide library but without OGT and without the UDP-GlcNAc sugar donor, respectively. The DNA recoveries (Cq value) from both control experiments were the same as that from pre-clear 3 (Figure S7), indicating the above recovery from the initial library was attributable to the O-GlcNAcylation of the peptides by OGT. The DNA from the first round of selection was subsequently amplified by PCR to allow further rounds of selection, of which four were performed. For reasons as yet unknown, the positive to negative recovery ratio of the selection did not increase substantially across these

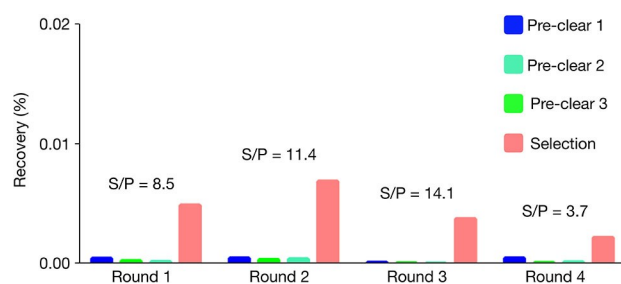


Figure 3. DNA recovery in each round of selection. The mRNA/cDNA-peptide library was pre-cleared three times using a 1:1 mixture of uncharged and RL2-antibody charged beads before OGT reaction. After OGT reaction, the mRNA/cDNA-peptide library was incubated with RL2-antibody charged beads to bind O-GlcNAcylated peptides. DNA eluted by heat from pre-clear beads ("Pre-clear 1", "Pre-clear 2", and "Pre-clear 3") and selection beads ("Selection") were analyzed by qPCR. The recovery ratio from each round was calculated by dividing the absolute amount of DNA after selection by that before selection (input), calibrated with a standard curve run in parallel. "S/P" values indicate the ratio of selection to pre-clear 3.

rounds, and in fact decreased from round 3 to round 4. For this reason the selection was stopped at this point.

While far from optimal, some enrichment was observed and so DNA eluted from the third round of selection (highest selection to pre-clear recovery ratio) was prepared for sequencing by PCR amplification and subsequently cloning into a PCR2.1-TOPO vector. This was transformed into *Escherichia coli* DH5 α and plated, from which 35 colonies were found. Plasmids were extracted and verified by PCR using M13R and M13F primers. Among these 35 plasmids, 19 showed a band corresponding to the size of the DNA library (111 bp), indicating that DNA fragments from the *in vitro* selection were cloned into the vectors (Figure S8). Sequencing of these verified plasmids led to 16 DNA sequences of the correct library architecture, which were translated into the primary amino acid sequences as shown in Figure 4A. Notably, all sequences except peptides 2 and 11 bear at least one serine or threonine that could be O-GlcNAcylated by OGT, although these sequences do not match the currently accepted sequon.^[12] In addition, duplicate sequences were found in plasmids 8 through 10 and 13 through 14, suggesting that these two sequences have been more strongly enriched during the selection.

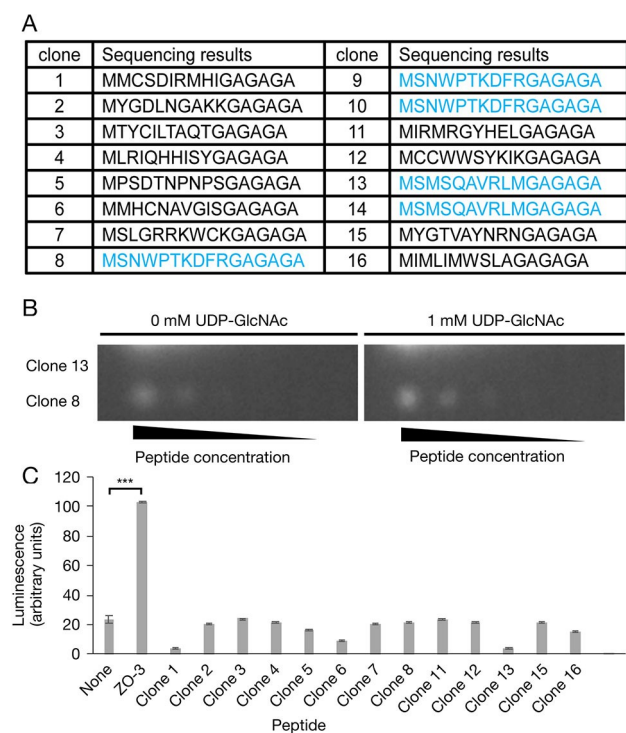


Figure 4. A) Peptide sequences from Sanger DNA sequencing results following cloning into the PCR2.1-TOPO vector (repeated sequences highlighted in blue). B) Peptides corresponding to clone 13 and 8 were synthesized and printed on a microarray at various concentrations. The OGT reaction was carried out in the presence of 1 mM UDP-GlcNAc, with reaction in the absence of UDP-GlcNAc used as control. C) All peptides were synthesized and tested as substrates for OGT by a UDP-Glo assay, with ZO-3 as a positive control. Bars represent the average of triplicate measurement with standard error. Significant activity above the no-peptide control is represented as *** ($p < 0.001$, two-tailed t-test).

Peptides 8 and 13 were thus synthesized and O-GlcNAcylated. They were tested by peptide microarray. Unfortunately, this revealed that peptide 13 is a false positive and peptide 8 showed binding of the antibody to the non-glycosylated form of the peptide, with only a slight increase in signal following OGT treatment (Figure 4B). This microarray approach is relatively labor intensive and relies on the same antibody as the selection, so rather than printing the remaining 11 sequences we sought further validation using a luminescence-coupled assay that detects liberated UDP following GlcNAc transfer (UDP-Glo). This removes the possibility of false positives based on antibody binding and allows a more convenient plate-based format. The results showed that none of these peptides gave luminescence above a no-peptide control, including sequence 8, and so we conclude that none are accepted as substrates by OGT (Figure 4C). Clone 8 can thus be concluded to have enriched through binding to the antibody, despite pre-clearing out such sequences. An appropriate control in future would be to carry out a direct enrichment for peptide binders to the antibody and also subject these to sequencing, allowing them to be excluded in subsequent analysis.

Although none of these sequences were confirmed to be substrates, we carried out high-throughput sequencing (HTS) of the input, negative and positive selection libraries in rounds 1 through 3 by Illumina Miseq. Such a high-throughput analysis of the input and selected libraries would yield more statistically robust information that could be used to compare the relative abundance of each sequence in the selected population,^[29] and which we hoped would help in the identification of new OGT substrates. The results of this sequencing showed that, even after selection, the libraries remained at extremely high diversity, with few repeated sequences and no clear pattern of preferred amino acids emerged (Figure S9). In addition, none of the sequences found from Sanger sequencing were found in the HTS results. The high diversity of this dataset, coupled with no way to define the site of modification for sequence alignment, prevented clear analysis of the results. Unexpectedly, the HTS data did show that the test sequence used for method development (fMRESSYDIYRVPSSQS) was enriching during the selection more than any other sequence. This sequence is not of the correct length to derive from the nine-amino-acid library, and was not deliberately spiked into the selection. We presume that this sequence entered the selection through cross-contamination from use of the same pipettes, despite the use of filter tips and other precautions. While not the result that we sought, this is nonetheless an indication that our approach is sufficiently powerful to enrich substrate sequences from a very low background starting point, even with the poor recovery seen in Figure 2B.

As the crystal structure of substrate-bound OGT shows a tunnel at the C terminus, we hypothesized that this might be interfering with the tag in our mRNA display experiment. Thus, to attempt to clarify the reason for our low enrichment, an additional pull-down experiment was performed using a synthetic DNA-peptide fusion bearing the same peptide sequence as in Figure 2, but in which the DNA was instead attached to the N terminus of the peptide. As described in

Figure S5, this DNA-peptide fusion was achieved using an succinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) crosslinker to attach a 3' amine-modified oligo to a synthetic peptide with an N-terminal cysteine of the OGT substrate sequence to that used in the puromycin-based method validation experiments.^[30] After OGT reaction with this peptide-oligo fusion, the same selection process was carried out as in that proof-of-principle experiment. A dramatically improved recovery of cDNA was observed in this case (Figure 5), corresponding to roughly ~2400-fold higher recovery with OGT compared to without. While not suitable for direct application in high-throughput substrate profiling, this construct nonetheless indicates that the reason for our low recovery was indeed the placement of the oligonucleotide acid tag. This could perhaps be solved by linker optimization in this case, while combining C- and N-terminal tags would likely present a more general strategy (particularly with enzymes for which no structural information is available). The location of the mRNA tag at the C terminus of the peptide is an inherent limitation of mRNA display as currently applied, but phage display does offer this possibility.

Although stringent pre-clear steps were performed against direct binding of the initial library, we still found a peptide in the enriched pool that binds directly to the antibody. This suggests that the use of an antibody may not be the best approach for specific *O*-GlcNAc peptide selection, even with a higher recovery from an optimized linker or N-terminal tag. Antibodies or lectins with higher specificity may be able to overcome the limitations of the current RL2 antibody-based selection approach, but any binding protein with sufficient affinity for an *O*-GlcNAc modified peptide is likely to bring its own inherent biases. A better alternative may be found in

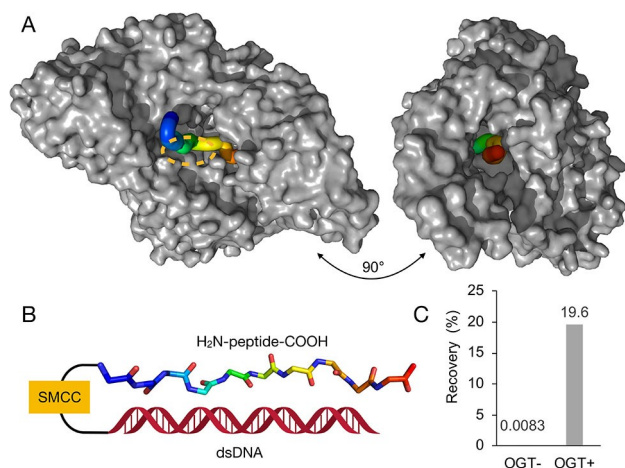


Figure 5. A) Illustration of the C-terminal occlusion of peptide substrates bound to OGT. Gray: OGT; rainbow cartoon: peptide, with blue at the N terminus to red at the C terminus; dotted yellow circle: OGT catalytic site (using PDB ID: 4GYY). B) Illustration of the construct used to test pull-down by N-terminal display, connecting a synthetic peptide (colored blue to red on carbons from N to C terminus) at the side chain of an N-terminal cysteine via SMCC to an amine-modified strand of DNA (together with an amine-free complementary strand). C) Pull-down of *O*-GlcNAcylated fraction of dsDNA-SMCC-peptide conjugates, with eluted DNA quantified by qPCR as for Figure 1.

azido-functionalized sugar donors such as UDP-GlcNAz and UDP-GlcNAc₆N₃, which have been shown to be accepted as substrates for OGT *in vitro*,^[31,32] and can subsequently be further covalently modified to attach biotin. However, these azido modifications exhibit a reduced rate for conversion of substrate as compared with the natural donor UDP-GlcNAc^[33] and, more importantly, may also induce a bias in the selected peptide sequences. A further alternative tool to selectively introduce an azide only on *O*-GlcNAc modified peptides or proteins is by means of a mutant β -1,4-galactosyltransferase that specifically transfers azido-modified galactose (GalNAz) from UDP-GalNAz to *O*-GlcNAc residues on glycosylated peptides or proteins.^[34,35] This tandem enzyme modification approach would avoid any bias or low turnover resulting from changing the OGT donor sugar while also overcoming the limitation of the antibody approach employed here. A final enrichment strategy that could be applicable here is through the use of GlcNAc electrophilic probes that react with a cysteine in a mutant OGT. On glycosyl transfer this gives a covalent link from peptide substrate to OGT and so allows substrate co-enrichment by pull-down of the enzyme.^[36] There are thus several reported approaches that offer solutions to this problem.

We have demonstrated here the feasibility of using mRNA display to enrich for an OGT substrate from a low starting ratio by use of antibody pull-down. Our data suggest that the described enrichment process works for a known substrate even at an extremely low starting ratio (derived in this case from cross-contamination), but with serious inherent issues from the antibody-based enrichment process and library architecture when applied to an unbiased library. Because of these limitations, with the current version of the selection protocol it was not possible to obtain any new insights into OGT's substrate requirement. Despite this, we report promising first steps towards an approach for high-throughput profiling of glycosyltransferase peptide substrates, as well as providing detail of problems encountered and proposed solutions.

Experimental Section

All the experimental details can be found in the Supporting Information.

Acknowledgements

This work was supported by a scholarship to J.S. from the China Scholarship Council (<http://www.csc.edu.cn/>).

Conflict of Interests

The authors declare no conflict of interests.

Keywords: mRNA display · *O*-GlcNAc transferase · *O*-GlcNAcylation · peptide selection · substrate specificity

- [1] X. Yang, K. Qian, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 452–465.
- [2] Q. Zeidan, G. W. Hart, *J. Cell Sci.* **2010**, *123*, 13–22.
- [3] S. Özcan, S. S. Andrali, J. E. L. Cantrell, *Biochim. Biophys. Acta Gene Regul. Mech.* **2010**, *1799*, 353–364.
- [4] J. Zhong, M. Martinez, S. Sengupta, A. Lee, X. Wu, R. Chaerkady, A. Chatterjee, R. N. O'Meally, R. N. Cole, A. Pandey, et al., *Proteomics* **2015**, *15*, 591–607.
- [5] S. Zhang, K. Roche, H. P. Nasheuer, N. F. Lowndes, *J. Biol. Chem.* **2011**, *286*, 37483–37495.
- [6] J. P. Singh, K. Zhang, J. Wu, X. Yang, *Cancer Lett.* **2015**, *356*, 244–250.
- [7] P. S. Banerjee, J. Ma, G. W. Hart, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6050–6055.
- [8] F. Liu, K. Iqbal, I. Grundke-Iqbal, G. W. Hart, C. X. Gong, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10804–10809.
- [9] S. A. Caldwell, S. R. Jackson, K. S. Shahriari, T. P. Lynch, G. Sethi, S. Walker, K. Vosseller, M. J. Reginato, *Oncogene* **2010**, *29*, 2831–2842.
- [10] P. V. Hornbeck, B. Zhang, B. Murray, J. M. Kornhauser, V. Latham, E. Skrzypek, *Nucleic Acids Res.* **2015**, *43*, D512–D520.
- [11] J. Ma, G. W. Hart, *Clin. Proteomics* **2014**, *11*, 8.
- [12] S. Pathak, J. Alonso, M. Schimpl, K. Rafie, D. E. Blair, V. S. Borodkin, A. W. Schüttelkopf, O. Albarbarawi, D. M. F. van Aalten, *Nat. Struct. Mol. Biol.* **2015**, *22*, 744–750.
- [13] J. Shi, T. Tomašič, S. Sharif, A. J. Brouwer, M. Anderluh, R. Ruijtenbeek, R. J. Pieters, *FEBS Lett.* **2017**, *591*, 1872–1883.
- [14] J. Shi, S. Sharif, R. Ruijtenbeek, R. J. Pieters, *PLoS One* **2016**, *11*, e0151085.
- [15] M. B. Lazarus, Y. Nam, J. Jiang, P. Sliz, S. Walker, *Nature* **2011**, *469*, 564–567.
- [16] W. Kightlinger, L. Lin, M. Rosztoczy, W. Li, M. P. Delisa, M. Mrksich, M. C. Jewett, *Nat. Chem. Biol.* **2018**, *14*, 627–635.
- [17] K. Rafie, O. Raimi, A. T. Ferenbach, V. S. Borodkin, V. Kapuria, D. M. F. van Aalten, *Open Biol.* **2017**, *7*, DOI: 10.1098/rsob.170078.
- [18] S. P. N. Iyer, G. W. Hart, *J. Biol. Chem.* **2003**, *278*, 24608–24616.
- [19] M. Jínek, J. Rehwinkel, B. D. Lazarus, E. Izaurrealde, J. A. Hanover, E. Conti, *Nat. Struct. Mol. Biol.* **2004**, *11*, 1001–1007.
- [20] J. Shi, R. Ruijtenbeek, R. J. Pieters, *Glycobiology* **2018**, *28*, 814–824.
- [21] D. S. Wilson, A. D. Keefe, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3750–3755.
- [22] W. W. Ja, O. Wiser, R. J. Austin, L. Y. Jan, R. W. Roberts, *ACS Chem. Biol.* **2006**, *1*, 570–4.
- [23] S. A. K. Jongkees, S. Caner, C. Tysoe, G. D. Brayer, S. G. Withers, H. Suga, *Cell Chem. Biol.* **2017**, *24*, 381–390.
- [24] H. Wang, R. Liu, *Expert Rev. Proteomics* **2011**, *8*, 335–346.
- [25] S. Tateyama, K. Horisawa, H. Takashima, E. Miyamoto-Sato, N. Doi, H. Yanagawa, *Nucleic Acids Res.* **2006**, *34*, e27.
- [26] T. P. Cujec, P. F. Medeiros, P. Hammond, C. Rise, B. L. Kreider, *Chem. Biol.* **2002**, *9*, 253–264.
- [27] S. R. Fleming, P. M. Himes, S. V. Ghodge, Y. Goto, H. Suga, A. A. Bowers, *J. Am. Chem. Soc.* **2020**, *142*, 5024–5028.
- [28] H. C. Dorfmueller, V. S. Borodkin, D. E. Blair, S. Pathak, I. Navratilova, D. M. F. Van Aalten, *Amino Acids* **2011**, *40*, 781–792.
- [29] J. Glanville, S. D'Angelo, T. A. Khan, S. T. Reddy, L. Naranjo, F. Ferrara, A. R. M. Bradbury, *Curr. Opin. Struct. Biol.* **2015**, *33*, 146–160.
- [30] B. A. R. Williams, J. C. Chaput, *Curr. Protoc. Nucleic Acid Chem.* **2010**, 1–20.
- [31] D. J. Voadlo, H. C. Hang, E. Kim, J. A. Hanover, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9116–9121.
- [32] A. Mayer, T. M. Gloster, W. K. Chou, D. J. Voadlo, M. E. Tanner, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1199–1201.
- [33] X. Ma, P. Liu, H. Yan, H. Sun, X. Liu, F. Zhou, L. Li, Y. Chen, M. M. Muthana, X. Chen, et al., *PLoS One* **2013**, *8*, 2–9.
- [34] P. M. Clark, J. F. Dweck, D. E. Mason, C. R. Hart, S. B. Buck, E. C. Peters, B. J. Agnew, L. C. Hsieh-Wilson, *J. Am. Chem. Soc.* **2008**, *130*, 11576–11577.
- [35] V. Dehennaut, M. C. Slomianny, A. Page, A. S. Vercoutter-Edouart, C. Jessus, J. C. Michalski, J. P. Vilain, J. F. Bodart, T. Lefebvre, *Mol. Cell. Proteomics* **2008**, *7*, 2229–2245.
- [36] C. W. Hu, M. Worth, D. Fan, B. Li, H. Li, L. Lu, X. Zhong, Z. Lin, L. Wei, Y. Ge, et al., *Nat. Chem. Biol.* **2017**, *13*, 1267–1273.

Manuscript received: September 2, 2020

Revised manuscript received: October 6, 2020

Accepted manuscript online: October 6, 2020

Version of record online: November 6, 2020