Antibodies directed against GalNAc- and GlcNAc-*O*-Tyrosine posttranslational modifications – a new tool for glycoproteomic detection

Sandra Behren#,[a,b] Manuel Schorlemer#,[a,b] Gudula Schmidt,[c] Klaus Aktories,[c] and Ulrika Westerlind*[a,b]

[a] Dr. Sandra Behren, Dr. Manuel Schorlemer, Prof. Dr. Ulrika Westerlind
Department of Chemistry, Umeå University
90187 Umeå, Sweden
E-mail: ulrika.westerlind@umu.se

[b] Dr. Sandra Behren, Dr. Manuel Schorlemer, Prof. Ulrika Westerlind
Leibniz-Institut für Analytische Wissenschaften - ISAS - e.V.
44227, Dortmund, Germany

[c] Prof. Dr. Gudula Schmidt, Prof. Dr. Klaus Aktories
Institut für Experimentelle und Klinische Pharmakologie und Toxikologie,
Universität Freiburg, Otto-Krayer-Haus, Albertstraße 25,
79104 Freiburg, Germany

# These authors contributed equally to the manuscript.

Supporting information for this article is given via a link at the end of the document.

Abstract:

In the last decade, it was discovered that protein mucin-type *O*-glycosylation and *O*-GlcNAcylation modify Tyr residues besides the well explored Thr and Ser amino acids. Several glycoproteomic studies have identified α-GalNAc-*O*-Tyr modifications, and studies propose that β-GlcNAc-*O*-Tyr also exists as a new group of posttranslational modifications (PTMs). Specific bacterial toxins have further been identified to modify host GTPases with α-GlcNAc-*O*-Tyr top promote bacterial virulence. Despite being identified on numerous proteins, the biological roles, biosynthesis and expression of GalNAc- and GlcNAc-*O*-Tyr modifications are poorly understood. A major obstacle is the lack of tools to specifically detect and identify proteins containing these modifications. With this in mind, we prepared vaccine constructs and raised antibodies to enable selective detection of proteins carrying these new PTMs. The obtained polyclonal antibody sera were evaluated using ELISA and glycopeptide microarrays and were found to be highly selective for GlcNAc- and GalNAc-*O*-Tyr glycopeptides over the corresponding Ser- and Thr-modifications. For microarray analysis, synthetic GlcNAc- and GalNAc-*O*-Tyr Fmoc-amino acids were prepared and applied in Fmoc-SPPS to obtain an extensive *O*-glycopeptide library. After affinity purification, the antibodies were applied in western blot analysis and showed specific detection of α-GlcNAc-*O*-Tyr modified RhoA GTPase.

Introduction

Protein glycosylation is the most abundant and diverse form of all known post-translational modifications and plays key roles in many biological processes. As such, the specific glycosylation site is of great importance. Glycoproteomic analysis of *O*-linked glycosylation is highly challenging due to the complex and diverse glycan structures. Often, nearby and dense glycosylation sites make proteolytic cleavage and site specific analysis demanding, and in contrast to *N*-glycoproteomic workflows, endoglycosidases capable of efficient *O*-glycan release are not available. Because of their significant biological functions, "mucin-type" *O*-glycosylation (*O*-GalNAc-type) and *O*-GlcNAcylation (*O*-GlcNAc-type glycosylation) receive more and more attention from the glycoproteomic field. Whereas *O*-GlcNAcylation predominantly occurs on cytosolic, mitochondrial and nuclear proteins, and is usually not further elongated or modified to generate more complex glycan structures,[1] mucin-type *O*-glycosylation represents the most abundant and complex form of protein *O*-glycosylation and usually contains extension of the inner core GalNAc residue.[2] Since 2011, several studies reported that *N*-acetylhexosamine (HexNAc = α-GalNAc, α- or β-GlcNAc) does not only modify threonine (Thr) and serine (Ser) amino acid residues but also the hydroxyl group of the tyrosine (Tyr) side chain.[3] This new type of protein *O*-glycosylation was first discovered on Tyr-681 of the amyloid-β (Aβ) peptide the DAEFRHDSGYEVHHQK in the frame of a glycoproteomic study on CSF from Alzheimer's patients,[3a] and was in recent work found to influence the Aβ peptide aggregation properties.[3g] Due to glycan extension forming mono-, di- or tri-sialyl-core 1 ((α-Neu5Ac)$_{1\text{-}3}$-β-1,3-Gal-α-GalNAc-) structures, the identified peptides were considered to be modified with mucin-type *O*-glycosylation on the Tyr glycosylation site. Glycan analysis using "SimpleCell Strategy" and lectin weak affinity enrichment (LWAC) with the GalNAc specific lectin *Vicia villosa* agglutinin (VVA) enabled identification of additional GalNAc-*O*-Tyr modified peptides: Tyr-389 of the Nucleobindin-2 (NUCB2) peptide LEYHQVIQQMEQK, Tyr-50 of the Nucleobindin-1 (NUCB1) peptide GAPNKEETPATESPDTGLYYHR, Tyr-43 of the Extracellular matrix protein 1 (ECM1) peptide QLRPEHFQEVGYAAPPSPPLSR, Tyr-560 of the CD44 peptide RDPNHSEGSTTLLEGYTSHYPHTK and Tyr-113 of the proline-

rich acidic protein-1 (PRAP1) peptide VLSPEPDHDSLYHPPPEEDQGEERPR.[3b,3d] Glycoproteomic studies of isolated mitochondrial proteins led to the discovery of HexNAc-*O*-Tyr on Tyr-418 of the ATP synthase subunit beta (ATP5B) peptide IMDPNIVGNEHYDVAR, Tyr-186 of the voltage-dependent anion-selective channel protein 1 (VDAC1) peptide VTQSNFAVGYK and Tyr-96 of the aspartate aminotransferase (ASAT) peptide NLDKEYLPIGGLAEFCK.[3c] Due to glycopeptide affinity enrichment with the less selective lectin wheat germ agglutinin (WGA), it was not possible to conclude if the identified Tyr modified peptides contained a new PTM, β-GlcNAc-*O*-Tyr, or mucin-type *O*-glycosylation. However, the tissue localization and previously identified β-GlcNAc residues on nearby Ser and Thr sites would suggest a β-GlcNAc-*O*-Tyr modification.[4] Additionally, it has been reported that the *Photorhabdus asymbiotica* protein toxin (PaTox^{G}) modifies the switch II region of host GTPases, including the GTPases RhoA, Rac and Cdc42, with the PTM α-GlcNAc-*O*-Tyr, and thus modulates the GTPase activity.[3e,3f,5] In this case, NMR was used to confirm the α-anomeric configuration after protein modification using UDP-GlcNAc as donor. Another study also reported about bacterial toxins that modify GTPases with Glc or GlcNAc on Ser- and Thr-sites.[5]

Because Ser and Thr are the expected acceptor amino acids, Tyr residues were previously not considered as potential glycosylation sites, and the often dense *O*-glycosylation on glycoproteins makes site-specific characterization challenging. The fact that Tyr modifications exist besides Ser- and Thr-HexNAcylation raises questions about their biosynthesis, expression and unique biological functions, which are unexplored to a large extend. The lack of efficient methods to determine the exact structures and linkages of the attached carbohydrate moieties, and to specifically detect and enrich these modifications makes the study of *O*-GalNAc- and *O*-GlcNAc-Tyr extremely challenging. Proper assignment of the HexNAc isomers in previous studies was in many cases not possible, and the glycoforms were assigned based on general knowledge about carbohydrate biosynthesis or lectin binding preferences. In line with this problem, we previously elucidated the use of a MS-based methodology for HexNAc isomer differentiation which allows site-specific characterization of *O*-GlcNAc and *O*-GalNAc on glycopeptides.[6] A recent study showed that several plant lectins used in LWAC recognize GalNAc on Tyr.[7] However, these lectins also detect *O*-GalNAc on Ser and Thr sites as well. Consequently, new tools for specific detection, enrichment and identification of *O*-GalNAc- and *O*-GlcNAc-Tyr need to be developed to gain insights into the functions of this new group of PTM.

Herein we present the generation of GlcNAc- and GalNAc-*O*-Tyr specific antibodies to selectively detect, profile and enrich peptides and proteins containing these modifications. Therefore, antigen glycopeptide-CRM conjugates with high α-GalNAc- or β-GlcNAc-*O*-Tyr density were generated to raise specific polyclonal antibodies, which were subsequently immunologically analyzed and evaluated by enzyme-linked immunosorbent assay (ELISA) and microarray binding experiments (Fig. 1A). To determine the antibody binding specificities, a peptide library was prepared that included peptide sequences carrying α-GalNAc, α- or β-GlcNAc on Ser, Thr or Tyr, and immobilized on microarrays (Fig. 1B). Furthermore, the work aimed to explore if the obtained

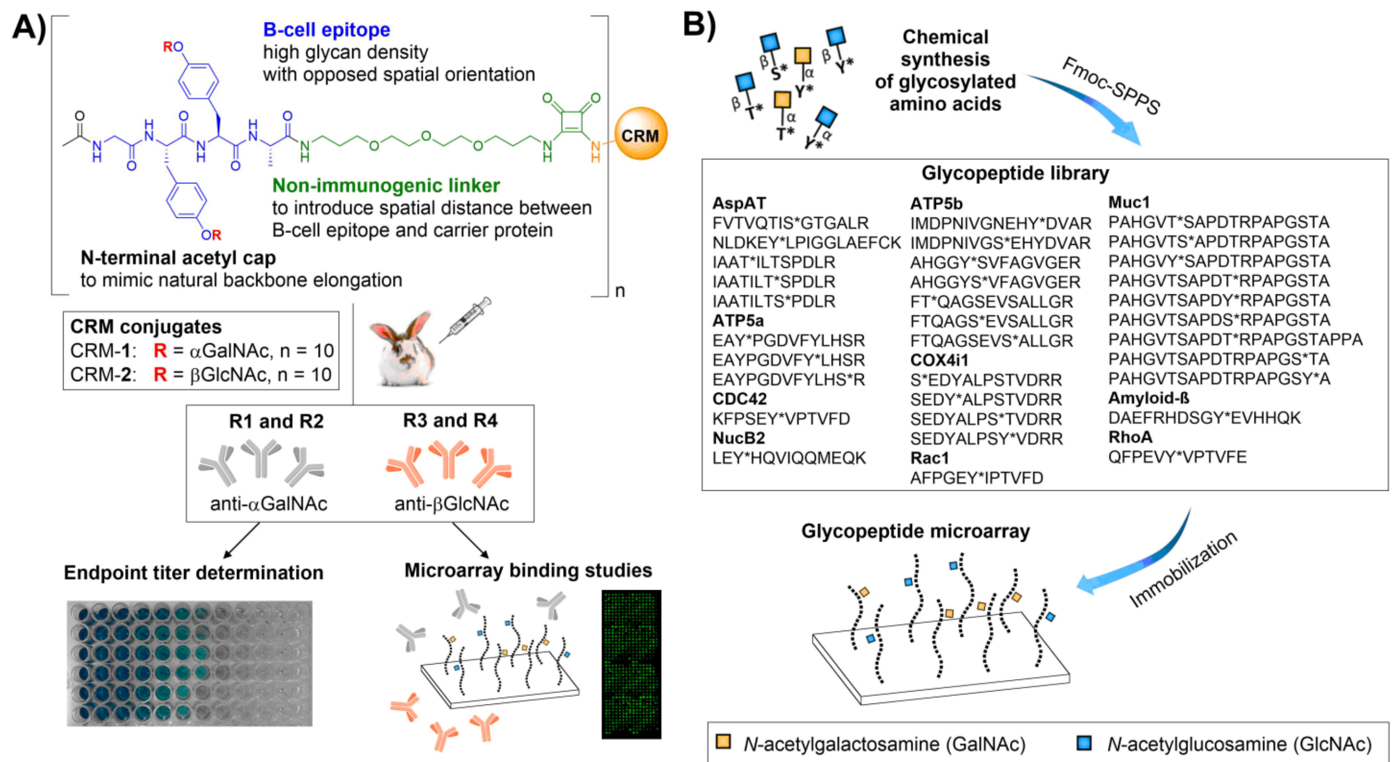
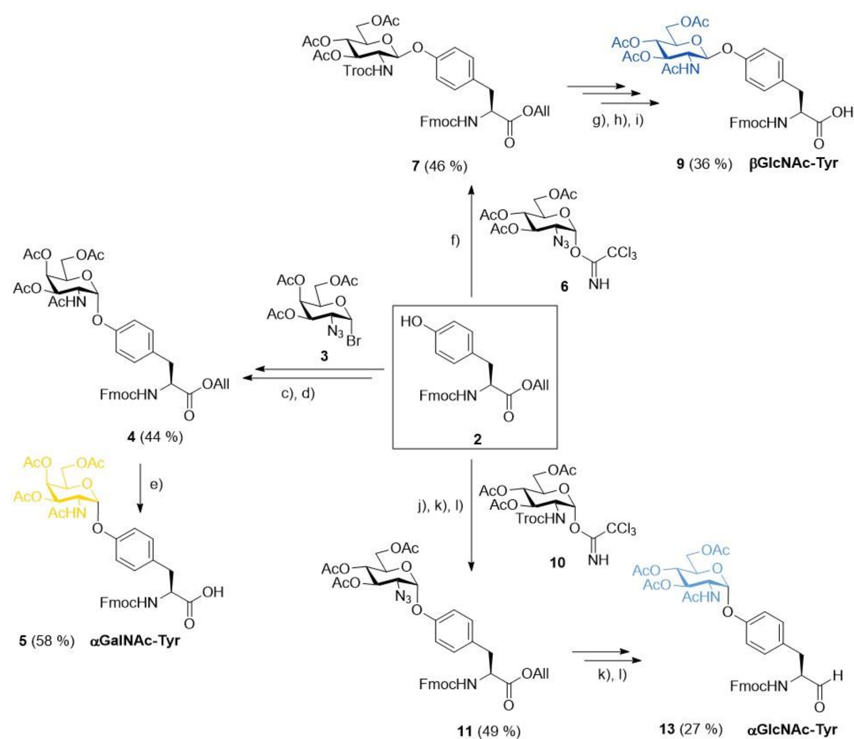


Figure 1. A) Preparation of antigen vaccines CRM-1 and CRM-2, generation of polyclonal rabbit antibodies and evaluation of the obtained antibodies using ELISA and microarray binding studies; B) Synthesis of glycosylated amino acids, a (glyco-)peptide library and preparation of peptide microarrays

GalNAc- and GlcNAc-*O*-Tyr specific antibodies could be used to selectively detect this modification on protein level. Therefore, the antibodies were purified by affinity chromatography against GalNAc- and GlcNAc-*O*-Tyr antigen peptides, re-evaluated by ELISA and microarray assays and applied to specifically detect α-GlcNAc-*O*-Tyr modified RhoA, which was enzymatically modified using the *Photorhabdus asymbiotica* protein toxin (PaTox[G]), on a Western blot.

Results and Discussion

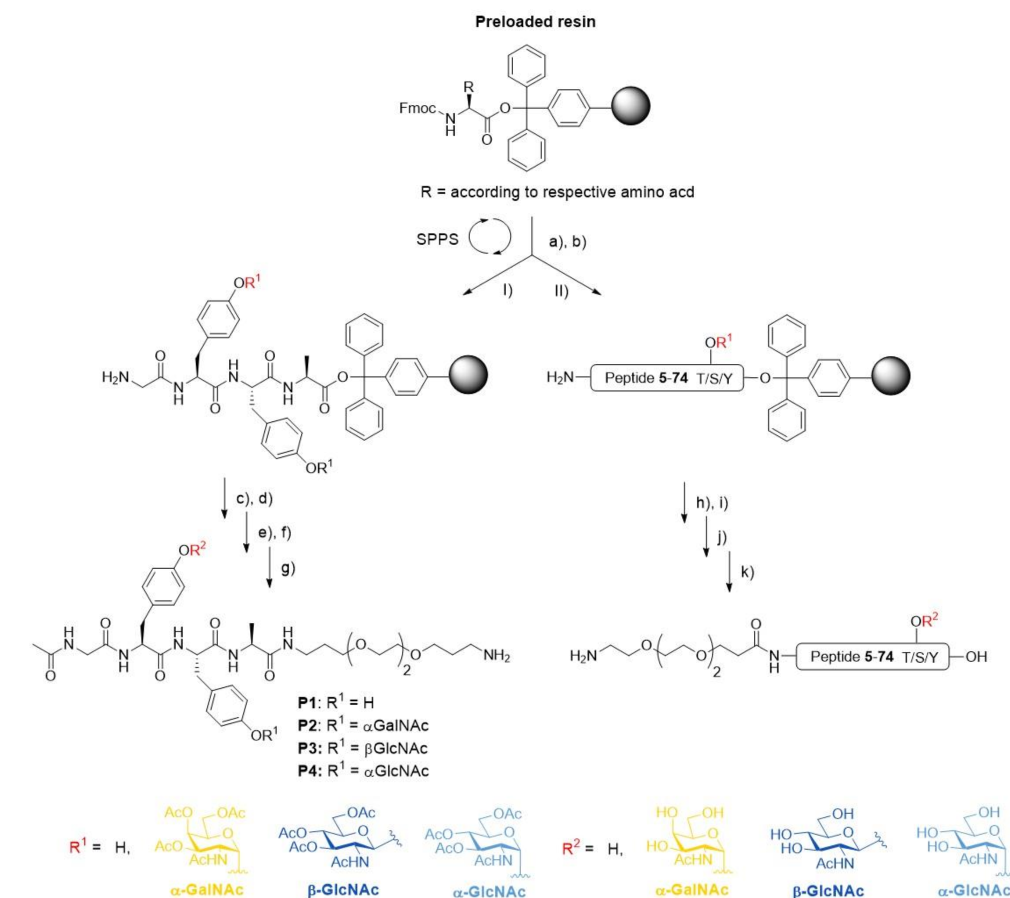
To prepare extensive GalNAc/GlcNAc-Tyr/Thr/Ser library containing 74 glycopeptides, glycosylated amino acid building blocks were assembled using common acceptor amino acids and corresponding glycosyl donors (Scheme 1), and subsequently incorporated into peptide sequences using solid phase peptide synthesis (SPPS) (Scheme 2).



Scheme 1. Synthesis of the αGalNAc-Tyr Fmoc amino acid building block 5, βGlcNAc-Tyr Fmoc amino acid building block 9 and αGlcNAc-Tyr Fmoc amino acid building block 13. a) Fmoc-OSu, $NaHCO_3$, dioxane/H_2O (3:1), 0 °C – RT, 4 d; b) allyl bromide, DIPA, THF, RT, overnight; c) $AgClO_4$, $AgCO_3$, MS 4 Å,DCM, 0 °C - RT, 2 d; d) AcSH, pyridine, RT, 72 h; e) $Pd(PPh_3)_4$, phenylsilane, THF, RT, 30 min; f) TMSOTf, MS 4 Å, DCM, 0 °C - RT, overnight; g) Zn, AcOH/THF (1:1), 50 °C, overnight; h) Ac_2O/pyridine (1:2), DMAP, RT, overnight; i) $Pd(PPh_3)_4$, phenylsilane, THF/DMF (1:1), RT, 1 h; j) TMSOTf, MS 4 Å, DCM/Et_2O (1:1), -20 °C, 2.5 h; k) AcSH, pyridine, RT, overnight; l) $Pd(PPh_3)_4$, phenylsilane, THF, RT, 30 min.

The peracetylated α-GalNAc-Fmoc-Tyr building block **5** was synthesized using a similar strategy to previously described protocols preparing α-GalNAc-Fmoc-Ser/Thr amino acids (Scheme 1, Supplementary Scheme S1B).[8] Here, the Koenigs-Knorr method was used with $AgCO_3$ and $AgClO_4$ activation to couple the Fmoc-Tyr allyl ester **2** with the glycosyl bromide donor **3** bearing a non-participating azide to yield exclusively the α-isomer. The glycan azide function was converted to a NHAc group by reductive amidation to give **4** in 44% yield from compound **2**.[9] In the final step, the allyl ester group was removed using phenylsilane and ($Pd(PPh_3)_4$ catalysis to give α-GalNAc-Fmoc-Tyr-OH **5** in 58% yield.[10] To generate the β- and α-GlcNAc-Fmoc-Tyr building blocks **9** and **13**, the tyrosine acceptor amino acid **2** was glycosylated with trichloroacetimidate glycosyl donors **6** and **10**, respectively (Scheme 2, Supplementary Schemes S1C and D).[11] The formation of the glycosidic linkage was controlled by either introducing a participating *N*-Troc at C-2 to promote β-glycosylation, or by installing a non-participating azide function to promote α-glycosylation, yielding **7** and **11** in 46% and 49%, respectively. The *N*-Troc group of **7** was then converted to an acetamide under reductive conditions using Zn/AcOH, followed by acetylation of the free amine to obtain **8** in 97% yield. The glycan azido group of **11** was reduced to the corresponding acetamide using thioacetic acid in pyridine to give **12** in 70% yield. The *C*-terminal allyl esters of **8** and **11** were removed to obtain the β- and α-GlcNAc-Fmoc-Tyr building blocks **9** and **13**, in 81% and 79% yield, respectively. The obtained peracetylated GalNAc/GlcNAc-Fmoc-Ser/Thr/Tyr building blocks were then used in Fmoc-SPPS to prepare α-GalNAc-*O*-Tyr and β-GlcNAc-*O*-Tyr antigen peptide-CRM vaccine conjugates as well as a library of glycopeptides carrying α-GalNAc, α- or β-GlcNAc on Ser, Thr or Tyr for microarray binding studies. Therefore, antigen glycopeptides **P1**-**P4** were generated containing short peptide sequences with high α-GalNAc- or β-GlcNAc-O-Tyr density, respectively, to serve as B-cell epitopes (Fig. 1A and Scheme 2). The antigen peptides were conjugated to CRM^{197}, which was used as carrier protein to elicit strong immune responses,[12] by using a non-immunogenic triethylene glycol linker[13] to obtain the corresponding vaccine constructs **CRM-1** and **-2**. The glycopeptide library was synthesized by automated Fmoc-SSPS. Standard Fmoc amino acids were coupled using HBTU/HOBt (Scheme 2). Glycosylated amino acids were pre-activated using the more reactive HATU/HOAt and manually added to the resin and the coupling time was extended to 8 h.[14] To prepare antigen glycopeptides **P1**-**P4**, the *N*-terminus was acetylated after complete peptide assembly. The peptides were released from the resin and a non-immunogenic linker was introduced to the free *C*-terminus for either subsequent conjugation to an immune carrier protein, or for immobilization on microarray slides.[15] The *tert*-butyloxycarbonyl (Boc) protecting group on the linker and the *O*-acetyl protecting groups on the glycans were removed and glycopeptides **P1**-**P4** were obtained after purification by preparative HPLC. To construct vaccine conjugates for immunization experiments and ELISA antibody endpoint titer determination, antigen peptides **P2** and **P3** were conjugated to CRM^{197} or BSA as carrier proteins, respectively. CRM^{197}, which is a non-toxic mutant of diphtheria toxin, was

chosen as immune carrier protein since it has successfully been used for immune stimulation with various glycoconjugates.[12] Glycopeptides **P2** (α-GalNAc-*O*-Tyr) and **P3** (β-GlcNAc-*O*-Tyr) were coupled to diethyl squarate, generating the corresponding squaric acid monoamides **14** and **15** (Supplementary Fig. S1). Finally, compounds **14** and **15** were conjugated to CRM^{197} or



Scheme 2. Synthesis of glycopeptides I) P1-P4 and II) P5-P74. a) Fmoc removal: 20 % piperidine in DMF; b) Amino acid coupling: Fmoc-Xaa-OH (8.0 equiv), HOBt (7.6 equiv), HBTU (7.6 equiv), DIPEA (16 equiv) in DMF, 40 min; Fmoc-Ser(βGlcNAc)-OH, Fmoc-Thr(βGlcNAc)-OH, Fmoc-Thr(αGalNAc)-OH, Fmoc-Tyr(βGlcNAc)-OH, Fmoc-Tyr(αGlcNAc)-OH or Fmoc-Tyr(αGalNAc)-OH (1.5 equiv), HATU (1.4 equiv), HOAt (1.4 equiv), DIPEA (3.0 equiv) in DMF, 8 h; c) Capping: acetic anhydride, HOBt, DIPEA, DMF, 2 h; d) Release from resin: TFA/TIPS/H_2O (95:5:5), 2 h; e) Spacer coupling: *N*-Boc-4,7,10-trioxa-1,13-tridecanediamine, HATU, HOAt, DIPEA, DMF, 4 h; f) Boc removal: dichloromethane/TFA (3:1), 4 h; g) Deprotection of the carbohydrates: 0.2 M NaOH/MeOH (pH 10.0), 18 h; h) Fmoc-triethylene glycol (3.0 equiv), HBTU (2.9 equiv), HOBt (2.9 equiv), DIPEA (6.0 equiv) in DMF, 2 h; i) Fmoc removal: 20 % piperidine in DMF; j) Release from resin: TFA/TIPS/H_2O (95:5:5); Deprotection of the carbohydrates: MeONa/MeOH (pH 8.5 – 9.5) or 0.2 M NaOH/MeOH (pH 10.0).

bovine serum albumin (BSA) to obtain the α-GalNAc-*O*-Tyr conjugates **CRM-1** and **BSA-1**, and β-GlcNAc-*O*-Tyr conjugates **CRM-2** and **BSA-2**. The glycopeptide-protein conjugates were purified by ultrafiltration and the loading rates were determined by MALDI-TOF mass spectrometry (Supplementary Fig. S2-S5). To induce antibodies against antigens α-GalNAc-Tyr and β-GlcNAc-Tyr, two rabbits were immunized with **CRM-1** (rabbit **R1** and **R2**) or **CRM-2** (rabbit **R3** and **R4**), respectively. The obtained antibody sera were analyzed by ELISA and microarray binding experiments. For ELISA experiments, 96-well microtiter plates were coated with BSA conjugates **BSA-1** or **BSA-2**, as well as with BSA only to determine the cutoff value. The rabbit sera were added to the wells at 2-fold dilutions and subsequently probed with a secondary antibody, and streptavidin-conjugated horseradish peroxidase (HRP) was used for colorimetric detection. The endpoint titers of the ELISAs were determined to be approximately 2 024 000 and 4 096 000 for rabbits **R1** and **R2**; and 4 096 000 and 2 048 000 for rabbits **R3** and **R4**, respectively (Supplementary Fig. S7). These findings indicate high concentrations of specific antibodies in the rabbit sera.

To evaluate the binding specificities of the obtained rabbit antisera, a peptide library consisting of 74 structurally well-defined synthetic glycopeptide sequences carrying α-GalNAc, α- or β-GlcNAc on Tyr, Ser or Thr was prepared (Scheme 2, Supplementary Table S1). This library contained different glycopeptide *O*-GalNAc and GlcNAc isomers, and included synthetic sequences of amyloid-β, Nuclobindin-2, previously identified by VVA LWAC enrichment and HCD-LC-MS as GalNAc-*O*-Tyr modified tryptic glycopeptide fragments, and isomeric glycopeptides of the human host RhoA, Rac1 and CDC42 GTPases, which were previously identified to be modified with α-GlcNAc-Tyr by bacterial toxins.[3a,b, e] Synthetic glycopeptide *O*-GalNAc and GlcNAc isomers were generated of previously identified tryptic glycopeptides of ATP5B and ASAT.[3c] Glycopeptides with identified nearby β-GlcNAc-Ser or -Thr sites of ATP5B and ASAT were synthesized including glycan isomer analogs modified at the specific site, or nearby potential glycosylation sites. Mitochondrial glycopeptide sequences of ATP synthase alpha (ATP5A) and Cytochrome c oxidase subunit 4 isoform 1 (Cox4i1) were synthesized, which were previously identified to contain β-GlcNAc-Ser or -Thr sites, and contained nearby Tyr sites.[4] Additionally, analogs of the MUC1 tandem repeat sequence were generated by replacing known and well explored GalNAc-Ser/Thr sites with HexNAc-Tyr to further explore the influence of Tyr modifications at different glycosylation sites on antibody recognition.[16] Peptides **P5**-**P74** were synthesized by Fmoc-SPPS according to the strategy described above. After complete peptide assembly, a triethylene glycol spacer was introduced to the *N*-terminus to later covalently bind the peptides to the microarray (Scheme 2). After assembly,

the peptides were cleaved from the resin with simultaneous removal of acid-sensitive protecting groups on the amino acid side chains. The *O*-acetyl protecting groups on the glycans were removed by sodium methoxide treatment, and peptides **P1**-**P74** were obtained after purification by preparative HPLC.

The synthetic peptides **P1**-**P74** were immobilized on NHS-activated hydrogel slides (Nexterion® slide H, Schott) by microarray spotting and unreacted NHS groups were blocked with ethanolamine. The immobilized glycopeptides were then used to evaluate the specificity of the raised polyclonal antibodies for α-GalNAc, α- or β-GlcNAc-Tyr glycopeptides in comparison with the corresponding Ser and Thr analogs. To adjust for variations in antibody titers, the microarrays were incubated with a dilution series of the anti-α-GalNAc-*O*-Tyr rabbit antisera **R1**, **R2** and the anti-β-GlcNAc-*O*-Tyr rabbit antisera **R3**, **R4**. The rabbit antibodies were detected by a secondary biotinylated goat anti-rabbit IgG, followed by treatment with Cy5-labeled streptavidin, and the fluorescence was read out. Microarray analysis showed differences in binding of the rabbit sera towards the glycopeptides regarding different glycosylation sites and different GalNAc or GlcNAc isomers at the same glycosylation site. The microarray binding data for rabbit sera **R1**, **R2**, **R3** and **R4** are shown in Supplementary Fig. S8-S12 and surface dissociation constants (Surf. K_D) are given in Supplementary Table S2. The microarray data showed that all rabbit antisera exhibited a high affinity for GlcNAc- and GalNAc-*O*-Tyr glycopeptides. In contrast, the corresponding Ser and Thr peptides were not recognized. Only at high serum concentrations, weak binding to β-*O*-GlcNAc on Ser of rabbit antisera **R2** and **R3** was observed, but not for rabbit sera **R1** and **R4** (Fig. 2A). While weak affinity of rabbit sera **R3, R1**, and **R4** was observed for the unglycosylated antigen structure **P1**, none of the other unglycosylated peptides were recognized. Surprisingly, all rabbit antisera exhibited cross-reactivity between the different HexNAc-*O*-Tyr isoforms indicating that the raised antibodies tolerate minor structural conformational differences between closely related monosaccharide structures (Fig. 2B). Among the HexNAc-*O*-Tyr isomers, the antibodies showed binding preferences for α-GlcNAc-*O*-Tyr over the corresponding α-GalNAc and β-GlcNAc-*O*-Tyr peptides irrespective of whether or not the antibodies were directed against the α-GalNAc or β-GlcNAc antigen peptide. In order to obtain monospecific α-GalNAc-, α-GlcNAc- and β-GlcNAc-*O*-Tyr antibodies, rabbit serum **R2** was purified by affinity enrichment against the antigen glycopeptides α-GalNAc-*O*-Tyr **P2,** β-GlcNAc-*O*-Tyr **P3** and α-GlcNAc-*O*-Tyr **P4**. Rabbit serum **R2** was chosen for affinity purification because it exhibited the strongest binding towards α-GalNAc and β-GlcNAc on Tyr, and binding towards α-GlcNAc-*O*-Tyr peptides was comparable with rabbit serum **R3** (Supplementary Table S2). Additionally, the serum had a high endpoint titer and it was the only serum that did not recognize the unglycosylated antigen peptide **P1**. Polyclonal anti-α-GalNAc antibodies were purified from rabbit serum **R2** using affinity

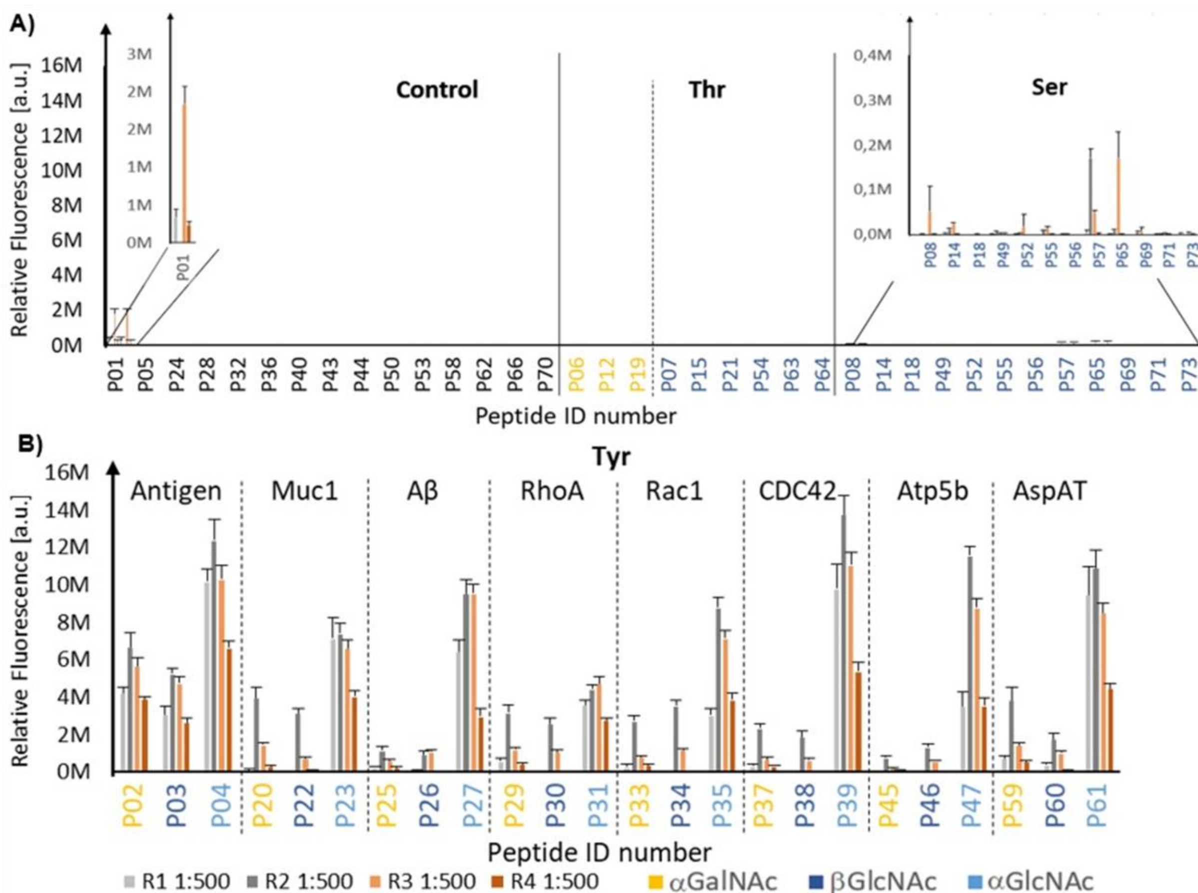


Figure 2. Binding of anti-α-GalNAc (R1 and R2) and anti-β-GlcNAc (R3 and R4) rabbit sera at dilution 1:500 to A) unglycosylated peptides and αGalNAc-Thr or βGlcNAc-Thr/Ser; B) αGalNAc-, βGlcNAc-, or αGlcNAc-Tyr glycopeptides.

chromatography columns prepared by coupling of the immunogenic peptides **P2** (α-GalNAc-*O*-Tyr), **P3** (β-GlcNAc-*O*-Tyr) and **P4** (α-GlcNAc-*O*-Tyr) to HiTrap NHS-Activated HP

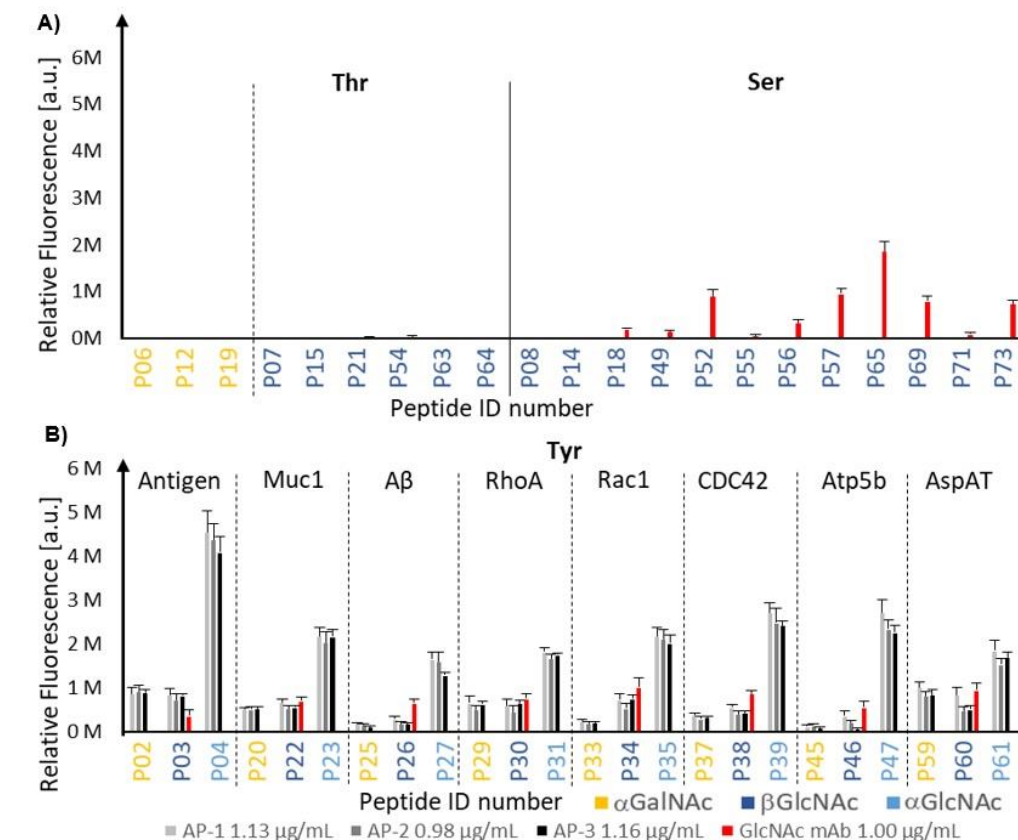


Figure 3. Binding of affinity purified rabbit antibodies AP-1, AP-2, AP-3 and anti-*O*-βGlcNAc mAb CTD 110.6 at concentrations 1.13 µg/mL, 0.98 µg/mL, 1.16 µg/mL and 1.00 µg/mL, respectively to A) αGalNAc-Thr or βGlcNAc-Thr/Ser glycopeptides; B) αGalNAc-, βGlcNAc-, or αGlcNAc-Tyr glycopeptides.

affinity columns, and purified antibodies **AP-1**, **AP-2** and **AP-3** were obtained, respectively. Next, the total protein concentrations of the purified antisera were quantified by 280 nm absorption measurements. Additionally, the enriched sera were analyzed for purity by SDS-gel electrophoresis. SDS-PAGE analysis of the purified rabbit **R2** antibodies showed a single band indicative of the desired purity (Supplementary Fig. S17). Then, the endpoint titers were determined by ELISA. To also obtain the endpoint titer for the anti-α-GlcNAc-*O*-Tyr antibodies **AP-3**, the α-GlcNAc-Tyr-BSA conjugate **BSA-3** was generated. Therefore, glycopeptide **P4** was coupled to diethyl squarate, generating the corresponding squaric acid monoamide **16**. In the next step, **16** was conjugated to BSA to obtain the desired conjugate **BSA-3**. The glycopeptide-protein conjugate was purified by ultrafiltration and the peptide to protein loading ratio was determined by MALDI-TOF mass spectrometry (Supplementary Fig. S6). ELISA analysis showed that all purified antibodies exhibited good immunoreactivity with the antigen conjugates α-GalNAc-*O*-Tyr **BSA-1**, β-GlcNAc-*O*-Tyr **BSA-2** and α-GlcNAc-*O*-Tyr **BSA-3**. End-point titers of purified antibodies for ELISA experiments were determined to be in a low ng/mL range (0.34-5.64 ng/mL) for anti-α-GalNAc-Tyr antibodies **AP-1**, anti-β-GlcNAc-Tyr antibodies **AP-2** and for anti-α-GlcNAc-Tyr antibodies **AP-3**, respectively (Supplementary Fig. S18). These titers indicate that high concentrations of specific high-affinity antibodies could be enriched during the affinity purifications. Finally, the binding specificities of the purified antibodies were evaluated using the glycopeptide microarray library. The microarray binding data for affinity purified antibodies **AP-1**, **AP-2**, and **AP-3** are shown in Fig. 3 and Supplementary Fig. S14-S16, and surface dissociation constants (Surf. K_D) are given in Supplementary Table S3. As desired, all purified antibodies showed no cross-reactivity towards the unglycosylated peptides, or Ser and Thr glycopeptides (Fig. 3A). The affinity enriched antisera **AP-1**, **AP-2**, and **AP-3** displayed similar binding patterns to the different GalNAc- and GlcNAc-*O*-Tyr glycopeptides (Fig. 3B) However, the affinity purifications did not result in a dramatic improvement of antibody selectivity towards the particular HexNAc-*O*-Tyr antigen immobilized on the affinity column. Instead, the higher affinity to the α-GlcNAc-*O*-Tyr over the α-GalNAc-*O*-Tyr and β-GlcNAc-*O*-Tyr glycopeptides remained (Fig. 3, Supplementary Table S3).

Additionally, we evaluated a commercial *O*-β-GlcNAc specific monoclonal mouse antibody (CTD 110.6), which is commonly used to detect Ser- and Thr-*O*-GlcNAcylation. The determined mAb binding specificities towards *O*-β-GlcNAc-Ser, -Thr and -Tyr were compared with our raised Tyr specific rabbit antisera (Fig. 3). While the *O*-GlcNAc mAb was highly selective for peptides modified with *O*-β-GlcNAc, surprisingly not all peptides *O*-β-GlcNAcylated on Ser and Thr peptides were recognized. However, the monoclonal antibody bound to all β-GlcNAc-*O*-Tyr peptides. Peptides glycosylated on Tyr were stronger recognized than the respective Ser and Thr glycopeptides indicating that hydrophobic interactions with the Tyr residue increases the binding drastically. Additionally, we could observe that our affinity purified antibodies **AP-1-3** show binding intensities comparable with the monoclonal anti-*O*-β-GlcNAc antibody at similar antibody concentrations. These findings suggest that our anti-HexNAc-*O*-Tyr antibodies can be used for bioanalytical applications using similar antibody concentrations as for the commercial monoclonal mouse antibody. The affinity enriched HexNAc-*O*-Tyr antibody sera **AP-1, 2** or **3,** which are selective towards Tyr

modified glycopeptides, and the β-GlcNAc selective monoclonal anti-*O*-β-GlcNAc antibody CTD 110.6 exhibit affinities that complement each other. In a combined format, these antibodies would enable selective detection of β-GlcNAc-*O*-Tyr modified glycopeptides.

Finally, we applied our purified antibodies to specifically detect HexNAc-*O*-Tyr modified proteins by western blot analysis. Therefore, we enzymatically modified RhoA with α-GlcNAc using UDP-GlcNAc and PaToxA^{G}, which is an enzyme that specifically glycosylates a specific Tyr site of host GTPases with α-GlcNAc. After enzyme treatment, SDS gel electrophoresis was performed with α-GlcNAcylated RhoA and unmodified RhoA (Supplementary Fig. S19). The proteins were transferred to a polyvinylidene fluoride (PVDF) membrane and probed with purified rabbit antibodies **AP-1**, **AP-2** and **AP-3**. The resulting protein-antibody complexes were detected with a secondary labeled antibody (donkey anti-rabbit IgG Alexa Fluor 488 conjugate) and scanned with a fluorescence imaging system (Fig. 4 and Supplementary Fig. S20). Whereas the unmodified RhoA was not recognized, the α-GlcNAcylated RhoA was successfully detected with our HexNAc-*O*-Tyr specific rabbit antibodies. These data suggest that we generated HexNAc-*O*-Tyr specific polyclonal rabbit antibodies **AP-1**, **AP-2** and **AP-3** that can detect α-GlcNAc-*O*-Tyr modified proteins.

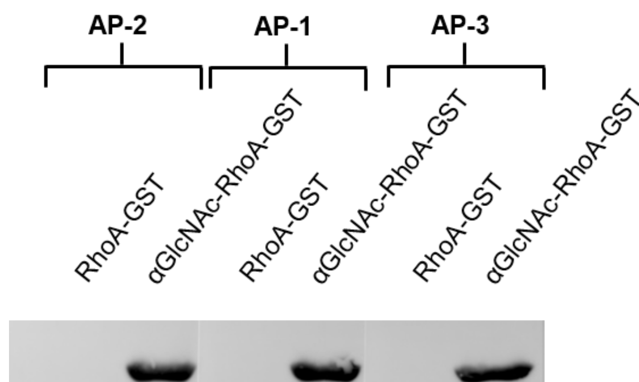


Figure 4. Western blot of RhoA-GST (lanes 1, 3 and 5) α-GlcNAc-RhoA-GST (lanes 2, 4 and 6) probed with HexNAc Tyr specific antibodies AP-1 (lanes 3 and 4), AP-2 (lanes 1 and 2) and AP-3 (lanes 5 and 6). Whereas the primary antibodies did not detect the unmodified RhoA-GST, they bound to α-GlcNAcylated RhoA. The lower band indicates partial breakdown of RhoA.

Conclusions

In summary, we generated polyclonal rabbit antibodies that are highly specific for α-GalNAc-, β-GlcNAc- and α-GlcNAc-*O*-Tyr glycopeptides, and can be applied to detect this relatively new group of protein modifications on protein level. Furthermore, a methodology was developed to efficiently synthesize GalNAc- and GlcNAc-*O*-Tyr amino acid building blocks, which were subsequently applied in Fmoc-SPPS to generate α-GalNAc- and β-GlcNAc-Tyr antigen glycopeptides. These antigen peptides were then conjugated to CRM^{197} and the obtained vaccine constructs were used to immunize rabbits to produce HexNAc-*O*-Tyr specific antibody sera. The sera were evaluated using ELISA and microarray binding assays. Consequently, we generated GalNAc- and GlcNAc-*O*-Tyr antigen peptide BSA conjugates for ELISA assays as well as a library of 74 synthetic unglycosylated and α-GalNAc-, β-GlcNAc- and α-GlcNAc-*O*-Tyr/Ser/Thr glycopeptides, which were subsequently printed on microarray slides. Microarray binding studies with the polyclonal rabbit sera showed that the raised antibodies exhibited strong binding to all HexNAc isoforms with a higher affinity for α-GlcNAc-*O*-Tyr glycopeptides. In contrast, low or no binding affinities were observed for unglycosylated and glycosylated Ser or Thr peptides. To evaluate abilities of the obtained antibodies to selectively detect this new protein modification in more complex samples, we affinity purified a rabbit serum with a high titer and high affinities for all HexNAc *O*-Tyr glycopeptide isomers, and applied the antigen-enriched antibodies in western blot analysis. Therefore, we enzymatically modified the host GTPase RhoA with α-GlcNAc-*O*-Tyr using the bacterial toxin PaToxA^{G} and blotted the obtained conjugate as well as the unmodified RhoA. Whereas the unmodified protein was not recognized by our purified rabbit antibodies, α-GlcNAc-*O*-Tyr-RhoA was successfully detected. In conclusion, we generated HexNAc-*O*-Tyr specific polyclonal rabbit antibodies that can detect α-GlcNAc-*O*-Tyr modified proteins. These findings support our hypothesis that the affinity purified antibodies can be applied to detect HexNAc-*O*-Tyr glycosylation in biological samples. Additionally, the *O*-β-GlcNAc specific monoclonal mouse antibody (CTD 110.6), which is commonly used for detection of Ser- and Thr *O*-GlcNAcylation by the glycobiology community, was evaluated on glycopeptide microarrays and found to also recognize β-GlcNAc-*O*-Tyr modifications. In conclusion, the affinity enriched HexNAc-*O*-Tyr antibodies **AP-1**, **AP-2** and **AP-3** and the *O*-β-GlcNAc mAb CTD 110.6 are new tools available to explore the glycobiology behind α-GalNAc-, β-GlcNAc- and α-GlcNAc-*O*-Tyr protein glycosylations.

Acknowledgements

This study was supported by a exploration grant to Ulrika Westerlind from the Boehringer Ingelheim Stiftung and by funding from the Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen, the Regierende Bürgermeister von Berlin - inkl. Wissenschaft und Forschung, and the Bundesministerium für Bildung und Forschung.

Keywords: Posttranslational modifications • Glycosylation • Glycopeptides • Antibodies • Microarrays

References

[1] a) G. W. Hart, *Annual Rev. Biochem.* **1997**, *66*, 315-335; b) F. I. Comer, G. W. Hart, *J. Biol. Chem.* **2000**, *275*, 29179-29182.

[2] a) D. J. Gill, H. Clausen, F. Bard, *Trends in cell biology* **2011**, *21*, 149-158; b) E. P. Bennett, U. Mandel, H. Clausen, T. A. Gerken, T. A. Fritz, L. A. Tabak, *Glycobiology* **2011**, *22*, 736-756.

[3] a) A. Halim, G. Brinkmalm, U. Ruetschi, A. Westman-Brinkmalm, E. Portelius, H. Zetterberg, K. Blennow, G. Larson, J. Nilsson, *Proc. Nat. Acad. Sci.* **2011**, *108*, 11848-11853; b) C. Steentoft, S. Y. Vakhrushev, M. B. Vester-Christensen, K. T. B. G. Schjoldager, Y. Kong, E. P. Bennett, U. Mandel, H. Wandall, S. B. Levery, H. Clausen, *Nature Methods* **2011**, *8*, 977-982; c) J. C. Trinidad, R. Schoepfer, A. L. Burlingame, K. F. Medzihradszky, *Mol. Cell. Proteomics.* **2013**, *12*, 3474-3488; d) S. Y. Vakhrushev, C. Steentoft, M. B. Vester-Christensen, E. P. Bennett, H. Clausen, S. B. Levery, *Mol. Cell. Proteomics.* **2013**, *12*, 932-944; e) T. Jank, X. Bogdanovic, C. Wirth, E. Haaf, M. Spoerner, K. E. Bohmer, M. Steinemann, J. H. C. Orth, H. R. Kalbitzer, B. Warscheid, C. Hunte,

K. Aktories, *Nature Struct. Mol. Biol.* **2013**, *20*, 1273-1280; f) T. Jank, S. Eckerle, M. Steinemann, C. Trillhaase, M. Schimpl, S. Wiese, D. M. Van Aalten, W. Driever, K. Aktories, *Nature commun.* **2015**, *6*, 7807; g) D. Liu, Q. Wei, W. Xia, C. He, Q. Zhang, L. Huang, X. Wang, Y. Sun, Y. Ma, X. Zhang, Y. Wang, X. Shi, C. Liu, S. Dong, *J.Am. Chem. Soc.* **2021**, *143*, 20216-20223.

[4] J. Ma, T. Liu, A. C. Wei, P. Banerjee, B. O'Rourke, G. W. Hart, *J. Biol. Chem.* **2015**, *290*, 29141-29153.

[5] K. Aktories, *Nat. Rev. Microbiol.* **2011,** *9,* 487–498.

[6] a) J. Yu, M. Schorlemer, A. Gomez Toledo, C. Pett, C. Sihlbom, G. Larson, U. Westerlind, J. Nilsson, *Chem. Eur. J.* **2016**, *22*, 1114-1124; b) A. Halim, U. Westerlind, C. Pett, M. Schorlemer, U. Ruetschi, G. Brinkmalm, C. Sihlbom, J. Lengqvist, G. Larson, J. Nilsson, *J. Proteome Res.* **2014**, *13*, 6024-6032.

[7] R. Gibadullin, D. W. Farnsworth, J. J. Barchi Jr, J. C. Gildersleeve, *ACS Chem. Biol.* **2017**, *12*, 2172-2182.

[8] a) W. Koenigs, E. Knorr, Chem. *Berichte* **1901**, *34*, 957-981; b) H. Paulsen, J. P. Holck, *Carbohydr. Res.* **1982**, *109*, 89-107; c) U. Westerlind, A. Hobel, N. Gaidzik, E. Schmitt, H. Kunz, *Angewandte Chem. Int. Ed.* **2008**, *47*, 7551-7556.

[9] T. Rosen, I. M. Lico, D. T. W. Chu, *J. Org. Chem.* **1988**, *53*, 1580-1582.

[10] P. Wang, J. Nilsson, G. Brinkmalm, G. Larson, X. F. Huang, *Chem. Commun.* **2014**, *50*, 15067-15070.

[11] a) R. R. Schmidt, J. Michel, *Tet. Lett.* **1984**, *25*, 821-824; b) R. R. Schmidt, J. Michel, *Angew. Chem. Int. Ed.* **1980**, *19*, 731-732; c) R. R. Schmidt, *Angew. Chem. Int. Ed.* **1986**, *25*, 212-235.

[12] a) G. Giannini, R. Rappuoli, G. Ratti, *Nucleic Acids Res* **1984,** *12*, 4063–4069; b) J. Yu, O. C. Grant, C. Pett, S. Strahl, R. J. Woods, U. Westerlind, *Chem. Eur. J.* **2017,** *23*, 3466-3473; c) U. Möginger, A. Resemann, C. E. Martin, S. Parameswarappa, S. Govindan, E.-C. Wamhoff, F. Broecker, D. Suckau, C. Lebev Pereira, C. Anish, P. H. Seeberger, D. Kolarich, *Sci. Rep.* **2016,** *6,* 20488-20502.

[13] a) S. Keil, C. Claus, W. Dippold, H. Kunz, *Angew. Chem., Int. Ed.* **2001**, *40*, 366-369; b) S. Dziadek, A. Hobel, E. Schmitt, H. Kunz, *Angew. Chem., Int. Ed.* **2005**, *44*, 7630-7635; c) U. Westerlind, H. Schroder, A. Hobel, N. Gaidzik, A. Kaiser, M. Niemeyer Christof, E. Schmitt, H. Waldmann, H. Kunz, *Angew. Chem. Int. Ed.* **2009**, *48*, 8263-8267.

[14] a) L. A. Carpino, A. El-Faham, C. A. Minor, F. Albericio, *Chem. Commun.* **1994**, 201-203; b) L. A. Carpino, *J. Am. Chem. Soc.* **1993**, *115*, 4397-4398; c) V. Dourtoglou, B. Gross, V. Lambropoulou, C. Zioudrou, *Synthesis* **1984**, *1984*, 572-574.

[15] M. Trester-Zedlitz, K. Kamada, S. K. Burley, D. Fenyö, B. T. Chait, T. W. Muir, *J. Am. Chem. Soc.* **2003**, *125*, 2416-2425.

[16] F. G. Hanisch *Biol. Chem.* **2001** *382*, 143–149.