



HAL
open science

A Bacterial Mannose Binding Lectin as a Tool for the Enrichment of C-and O-Mannosylated Peptides

Hermann Hütte, Birgit Tiemann, Aleksandra Shcherbakova, Valerian Grote, Marcus Hoffmann, Lorenzo Povolò, Mark Lommel, Sabine Strahl, Sergey Vakhrushev, Erdmann Rapp, et al.

► **To cite this version:**

Hermann Hütte, Birgit Tiemann, Aleksandra Shcherbakova, Valerian Grote, Marcus Hoffmann, et al.. A Bacterial Mannose Binding Lectin as a Tool for the Enrichment of C-and O-Mannosylated Peptides. Analytical Chemistry, 2022, 94 (20), pp.7329-7338. 10.1021/acs.analchem.2c00742 . hal-03679028

HAL Id: hal-03679028

<https://hal.science/hal-03679028>

Submitted on 25 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Bacterial Mannose Binding Lectin as a Tool for the Enrichment of C- and O- Mannosylated Peptides

Hermann J. Hütte¹, Birgit Tiemann¹, Aleksandra Shcherbakova¹, Valerian Grote², Marcus Hoffmann², Lorenzo Povolo³, Mark Lommel^{4#}, Sabine Strahl⁴, Sergey Y. Vakhrushev³, Erdmann Rapp^{2,5}, Falk F.R. Buettner¹, Adnan Halim³, Anne Imberty⁶, Hans Bakker^{1*}

¹Institute of Clinical Biochemistry, Hannover Medical School, Carl-Neuberg-Straße 1, 30625, Hannover, Germany

²Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany

³Copenhagen Center for Glycomics, Department of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3, Copenhagen, Denmark

⁴Centre for Organismal Studies (COS), Glycobiology, Heidelberg University, Im Neuenheimer Feld 360, 69120 Heidelberg, Germany

⁵glyXera GmbH, Leipziger Straße 44, 39120 Magdeburg, Germany

⁶Université Grenoble Alpes, CNRS, CERMAV, 601 rue de la chimie, 38000 Grenoble, France

Present address: Department of Microbiology, Faculty of Natural Sciences and Technology, Saarland University, Campus A1.5, 66123 Saarbrücken, Germany

Summary : *C-mannosylated peptides are easily detected by mass spectrometry (MS) from purified proteins but not from complex biological samples. Enrichment of specific glycopeptides by lectin affinity prior to MS analysis has been widely applied to support glycopeptide identification, but was up till now not available for C-mannosylated peptides. Here we used the α -mannose-specific Burkholderia cenocepacia lectin A (BC2L-A) and show that, in addition to its previously demonstrated high-mannose N-glycan binding capability, this lectin is able to retain C- and O-mannosylated peptides. Next to testing binding abilities to standard peptides, we applied BC2L-A affinity to enrich C-mannosylated peptides from complex samples of tryptic digests of HEK293 and MCF10A whole cell extracts, which led to the identification of novel C-mannosylation sites. In conclusion, BC2L-A enabled specific enrichment of C- and O-mannosylated peptides and might have superior properties over other lectins for binding to α -mannose-terminating glycans.*

Keywords : proteomics, lectin, C-mannosylation, O-mannosylation, N-linked glycosylation, glycosylation

Introduction

In MS-based proteomics, many types of peptides are underrepresented (Duncan et al., 2010), including C-mannosylated peptides. So far, these peptides could only be identified from purified proteins, either from natural source or upon recombinant expression (Buettner et al., 2013; Hofsteenge et al., 1999; Niwa et al., 2016; Verbij et al., 2016). Thus, discovery of new C-mannosylation sites is currently mainly based on an hypothesis-driven approach by MS analysis of specifically enriched proteins harboring already known WxxW/C consensus sequences (Furmanek and Hofsteenge, 2000). Enrichment of C-mannosylated peptides from complex biological samples would greatly enhance the possibility of discovering new C-mannosylation sites on proteins in an unbiased manner.

Very recently, an antibody was described that allowed enrichment of C-mannosylated peptides from a complex mixture of peptides (John et al., 2021). Several new C-mannosylation sites were discovered by this method. However, the used antibody requires one specific amino acid adjacent to the C-mannosylated tryptophan. In addition to antibodies, glycans can be specifically bound by lectins. These carbohydrate recognizing proteins are widely expressed in all forms of life (Sharon, 2008) and have diverse structures and recognition motifs (Bonnardel et al., 2019). The properties of lectins make them a useful tool to detect and purify glycoconjugates (Suttapitugsakul et al., 2020). A well-developed method is lectin affinity chromatography. In this procedure, biological samples are applied to an immobilized lectin to enrich specific glycoconjugates (Van Damme, 2011). A widely used lectin is Concanavalin A (ConA) (Freeze, 1995), a plant lectin recognizing mannose residues, notably present on N-glycans (Baenziger

and Fiete, 1979). Thus, the use of ConA affinity facilitated the disclosure of the N-glycoproteomes of diverse organisms (Bunkenborg et al., 2004; Kaji et al., 2007; Zielinska et al., 2010). Moreover, ConA has been used to identify new O-mannosylation sites on proteins (Neubert et al., 2016; Vester-Christensen et al., 2013; Winterhalter et al., 2013), resulting in the discovery of a new class of O-mannosyltransferases in mammals (Larsen et al., 2017a) and cytosolic O-mannosylation in yeast (Halim et al., 2015).

In addition to N-glycosylation and O-mannosylation, C-mannosylation is the third known type of protein glycosylation involving α -mannose. In contrast to all other glycoconjugates in which the mannose is linked via an oxygen to either another sugar residue (in N-glycans) or to serine/threonine (O-mannose), the C-mannose is attached to a tryptophan residue in a carbon-carbon linkage (Furmanek and Hofsteenge, 2000). The canonical consensus sequence for C-mannosylation is the WxxW motif, where only the first tryptophan is modified (Doucey et al., 1998; Furmanek et al., 2003). C-mannosylation has also been found on the second tryptophan residue or on tryptophan residues that are not part of the canonical consensus sequence (Ervin et al., 2005; John et al., 2021; Sasazawa et al., 2015), most likely due to the existence of at least two C-mannosyltransferases with distinct specificities (Shcherbakova et al., 2017). C-mannose is important for protein folding and stability (Shcherbakova et al., 2019). Lack of C-mannosylation results in compromised secretion of its target proteins, which, on organismal level, results in a pleiotropic phenotype in *C. elegans* (Honigberg and Kenyon, 2000) and fish (Cirksena et al., 2021).

Bioinformatics predicted that C-mannose is not rare – ~18 % of proteins passing the secretory pathway contain at least one C-mannosylation site (Julenius, 2007). Nevertheless, the number of so far identified C-mannosylated proteins is limited. These proteins mainly belong to two classes, cytokine type I receptors (Furmanek et al., 2003) and the thrombospondin type 1 repeat (TSR) superfamily (Hofsteenge et al., 2001). Previous attempts to use ConA lectin affinity to bind C-mannosyltryptophan were not successful (Nishikawa et al., 2004). Application of the mannose binding lectins *Galanthus nivalis* agglutinin (GNA) and *Lens culinaris* agglutinin (LCA) on the Cys subdomain 5 of MUC5A (Perez-Vilar, 2004) showed lack of binding

when the putative C-mannosylated tryptophan residue was mutated. So far, however, there is no direct evidence that the Cys subdomains of mucins are C-mannosylated (Ambort et al., 2011). Similar to ConA, both these lectins contain binding pockets that recognize more than just a single mannose (Kornfeld et al., 1981; Wright and Hester, 1996).

Looking for an alternative α -mannose-binding lectin we encountered the bacterial *Burkholderia cenocepacia* lectin A (BC2L-A) (Lameignere et al., 2008, 2010). BC2L-A belongs to the pseudomonas LecB family, characterized by the direct involvement of two calcium ions in the binding site (Mitchell et al., 2002). BC2L-A is the smallest member of the family with a single binding pocket, strict specificity, and high affinity for α -mannose, without the requirement for other epitopes. In this study, we show that BC2L-A is a C-mannose binding lectin and demonstrate that BC2L-A reveals advantages over ConA in the binding of O-mannosylated peptides. The strong binding capabilities of BC2L-A allowed us to enrich C-mannosylated peptides from complex mixtures by affinity chromatography, which revealed so far non-proven C-mannosylation sites of TSRs but also previously unknown C-mannosylation sites.

Results

BC2L-A binds C-mannosylated peptides

The structure of the BC2L-A binding pocket and its affinities are well studied. Co-crystallization of BC2L-A revealed tight binding of a single methyl α -D-mannoside in 4C_1 conformation coordinating to calcium ions in each subunit of the dimer (Lameignere et al., 2008). However, a protein fragment comprising an intact TSR with two C-mannoses was not retained on a BC2L-A affinity column (Fig. 1A). Superimposition of C-mannosylated TSR on the binding site of BC2L-A demonstrated that the binding site is large enough to accommodate the tryptophan residue. However, steric hindrance is observed for the distal part of the protein, which rationalizes the absence of interactions (Fig. 1B). We then tested C-mannosylated peptides derived from the TSR, which are both shorter and more flexible, for binding to BC2L-A. The single TSR (TSR2 of *C. elegans* UNC-5) was produced in *Drosophila* S2 cells either with a maximum of two or without C-mannoses as described before (Shcherbakova et al.,

2019). The TSR was reduced and digested with trypsin, giving rise to two peptides of only two amino acids long (not detected by MS) and two larger peptides (peptide 1 and 2 in figure 1C) of which peptide 1 contained the two C-mannosylation sites. Separation of the digested TSR on the BC2L-A column and detection by UV at 280 nm revealed a single peak in the flowthrough fraction (designated unbound in Fig. 1D), but in contrast to the intact TSR (Fig. 1A), a second peak was observed after elution with EDTA (bound fraction). The binding capacity of the column could be restored by washing with calcium containing

buffer. Liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis of both peaks, presented as extracted ion chromatograms (XICs) in figure 1D, confirmed that only the C-mannosylated peptide 1 was bound by BC2L-A. Both mono-mannosylated (present as minor fraction) and di-mannosylated peptides were bound by the lectin and eluted by EDTA. In contrast, when the non-mannosylated TSR was digested and loaded on the BC2L-A column, both peptides 1 and 2 were identified within the unbound fraction (Fig. 1E). These experiments showed that BC2L-A specifically binds a C-mannosylated peptide.

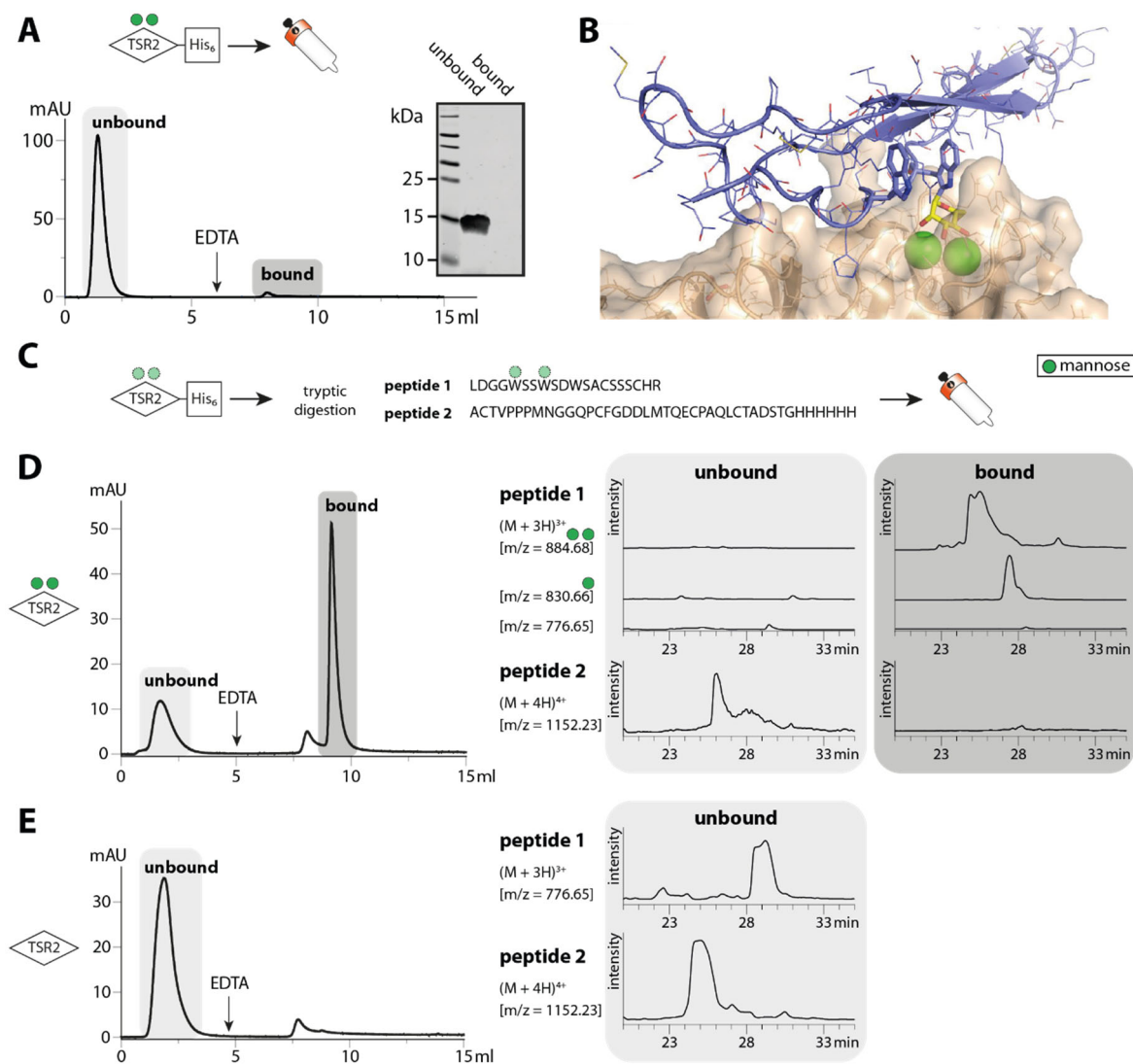


Figure 1: BC2L-A binds a C-mannosylated peptide. (A) BC2L-A lectin affinity chromatogram of the C-mannosylated *C. elegans* UNC-5 TSR2. Unbound (flowthrough) and bound fractions were analyzed by western blot using anti-His₅₆ antibody. (B) Superimposition of structure of TSR as blue ribbon (PDB code 1LSL) with modelled C-mannose on W426 on the structure of BC2L-A with transparent surface (PDB code 2VNV). Mannose is displayed as yellow sticks and calcium ions as green sphere. Steric hindrance is visible on the left of the panel (C) Tryptic peptides of His-tagged UNC-5 TSR2. (D) BC2L-A lectin affinity chromatogram of the digested C-mannosylated UNC-5 TSR2 (left) and XICs of peptides in the corresponding peak fractions (right). (E) BC2L-A lectin affinity chromatogram of the digested non-mannosylated UNC-5 TSR2 (left) and corresponding XICs (right).

BC2L-A selectively binds C-mannosylated peptides from a more complex peptide mixture

To test a larger protein with more irregular C-mannosylation, we expressed and purified a human thrombospondin 1 (THBS1) fragment, containing all three TSRs (TSR1-3) of the protein, in CHO cells. In mammalian cells, these TSRs show a variable degree of glycosylation with up to three C-mannoses on each TSR (Shcherbakova et al., 2017). Moreover, adjacent to the C-mannosylation sites, the TSRs exhibit O-fucosylation sites present on the same tryptic peptides (Gonzalez de Peredo et al., 2002).

Similar to TSR2 of *C. elegans* UNC-5, intact TSR1-3 was not bound by BC2L-A (Suppl. Fig. S1). Hence, we digested the purified protein with trypsin and

applied the digest to the BC2L-A column. The resulting chromatogram showed a bound and an unbound fraction (Fig. 2A). Analysis by LC-MS/MS demonstrated that C-mannosylated peptides were almost exclusively found in the bound fraction (Fig. 2B). Quantitative LC-MS/MS analysis of all C-mannosylation and O-fucosylation consensus sequence containing peptides confirmed the high C-mannose binding specificity of BC2L-A regardless of the number of attached C-mannoses per peptide (Fig. 2C). Furthermore, additional O-fucosylation of C-mannosylated peptides did not interfere with binding to BC2L-A. Non-mannosylated, but O-fucosylated peptides, regardless of the extension by glucose, were predominantly found in the unbound fraction. This confirmed the strong specificity of BC2L-A for mannose.

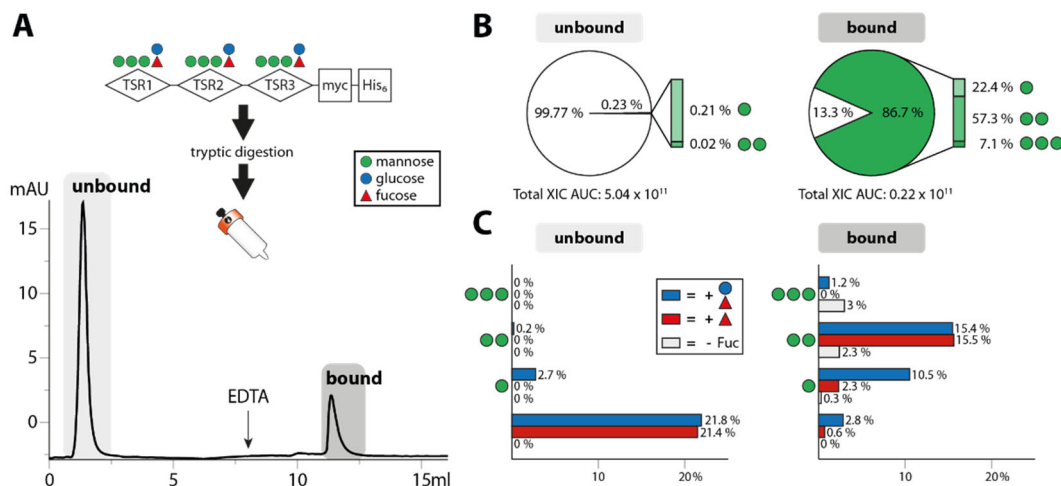


Fig. 2: BC2L-A retains C-mannosylated peptides of purified human THBS1. (A) A scheme of the thrombospondin construct (TSR1-3) expressed in CHO cells with three TSRs that can each carry up to three C-mannoses and a glucosylated O-fucose in the same tryptic peptide (upper panel) plus the corresponding BC2L-A lectin affinity chromatogram (lower panel). (B) LC-MS/MS quantification of all peptides from the unbound and bound lectin affinity fractions. Green areas represent C-mannosylated, white areas non-mannosylated peptides. The XIC-AUC (sum of the area under the curve) was set to 100% for unbound and bound fraction separately. The insets of the pie charts show the relative proportion of the number of mannoses present on single glycopeptides. (C) Distribution of only the peptides containing C-mannosylation and O-fucosylation motifs. The combined XIC-AUC of these peptides from unbound and bound fractions was set to 100%. Identified peptides and corresponding XICs, MS, and MS/MS spectra can be found in supplemental item 1.

BC2L-A binds both C- and O-mannosylated peptides

Currently, lectin-based enrichment of O-mannosylated proteins is achieved by the ConA lectin in a lectin weak affinity chromatography method, which requires a particularly long column (Halim et al., 2015; Neubert et al., 2016; Vester-Christensen et al., 2013; Winterhalter et al., 2013). So far, no other lectin with a higher affinity for O-mannoses has been described. Since BC2L-A is

capable of binding to C-mannoses and other mannose conjugates (Lameignere et al., 2008, 2010), we analyzed affinity of a purified C-mannosylated and a synthetic O-mannosylated peptide in comparison to ConA. We applied both peptides to a 1 ml ConA or BC2L-A column and compared their behavior to the non-mannosylated equivalent. As expected, the non-mannosylated peptides were bound neither by ConA nor by BC2L-A (Fig. 3A and B). Application of the mannosylated peptides revealed different affinities of both lectins.

The chromatogram of ConA revealed a weak affinity to the O-mannosylated as well as the C-mannosylated peptide. Both peptides showed a minor shift in retention time (Fig. 3A and B left). Retention of the O-mannosylated peptide seems slightly stronger, even though the C-mannosylated peptide contains two mannosylated sites. This

confirmed the weak affinity of ConA for peptide bound O-mannose and showed that ConA has an even lower affinity for the C-mannosylated peptide. In contrast, the interaction between both mannosylated peptides and BC2L-A was only be disrupted with EDTA (Fig. 3A and B right).



Figure 3: Binding affinities of BC2L-A and ConA to O-mannosylated peptides. (A) ConA (left) and BC2L-A (right) lectin affinity chromatograms of purified non- (dotted line) and C-mannosylated (solid line) peptide. LC-MS/MS analysis confirmed that no C-mannosylated peptide accumulated in the bound fraction of ConA. BC2L-A retained 98% of C-mannosylated peptide, which was eluted by EDTA (Suppl. Fig. S2). (B) ConA (left) and BC2L-A (right) lectin affinity chromatograms of non- (dotted line) and O-mannosylated (solid line) YATAV peptide. (C) ConA (left) and BC2L-A (right) lectin affinity chromatograms of a tryptic digest of N-cadherin with LC-MS/MS quantification of the detected O-mannosylated peptides in unbound and bound fractions. The pie charts show the relative proportion between O-mannosylated (in green) and non-O-mannosylated glycopeptides. The insets of the pie charts show the relative proportion of the number of mannoses present on single glycopeptides. Elution in ConA and BC2L-A chromatography was achieved by 0.1 M methyl α -D-mannoside or 1 mM EDTA, respectively. Identified peptides and corresponding XICs, MS, and MS/MS spectra can be found in supplemental item 2.

To analyze the binding compatibilities of O-mannosylated peptides from a more complex peptide mixture, a tryptic digest of the O-mannosylated mouse N-Cadherin expressed in CHO cells was applied to BC2L-A and ConA columns (Fig. 3C). Due to low absorbance at 280 nm of the O-mannosylated peptides, fractions of the bound and unbound peaks were collected in analogy to the lectin affinity chromatogram shown in figure 3B and analyzed by LC-MS/MS. For BC2L-A, a clear separation of O-mannosylated from non-mannosylated peptides was observed. Interestingly, although the YATAV peptide carrying a single mannose residue was only weakly bound by ConA (Fig. 3B), multiply O-mannosylated peptides were significantly enriched, consistent with the previously described ConA recognition capability for multiple mannoses (Naismith and Field, 1996). However, the binding was less efficient than to BC2L-A. Thus, BC2L-A is very well suited for the enrichment of O-mannosylated peptides.

BC2L-A binding of oligomannose-type N-glycans is comparable to ConA

N-cadherin is modified by O-mannoses and N-glycans on distinct peptides (Noor et al., 2021). N-

glycosylation analysis by LC-MS/MS of N-cadherin-derived tryptic peptides revealed only a few N-glycosylated peptides in bound and unbound fractions of the lectin affinity (Supplemental file 3). However, peptides with oligomannose-type N-glycans were only identified in the bound ConA and BC2L-A fractions. Hence, we further examined and compared the oligomannose-type N-glycan binding capability of BC2L-A and ConA using a more defined N-glycosylated protein. For this purpose, 10 µg of bovine Ribonuclease B (RNase B), a small glycoprotein that primarily carries oligomannose-type N-glycans in a single N-glycosylation site (Fu et al., 1994), was applied to the ConA or BC2L-A column. Elution of RNase B from the ConA column with 0.1 M methyl α-D-mannoside resulted in a peak fraction that was overlapped by the intrinsic absorbance of methyl α-D-mannoside (Fig. 4A). A similar elution peak of RNase B was observed for BC2L-A (Fig. 4B). However, no background absorbance was observed due to elution with EDTA. The retention of RNase B on both lectins was demonstrated by SDS-PAGE of the bound and unbound fractions (Fig. 4C), confirming high affinity of BC2L-A to oligomannose-type N-glycans, which was comparable to ConA.

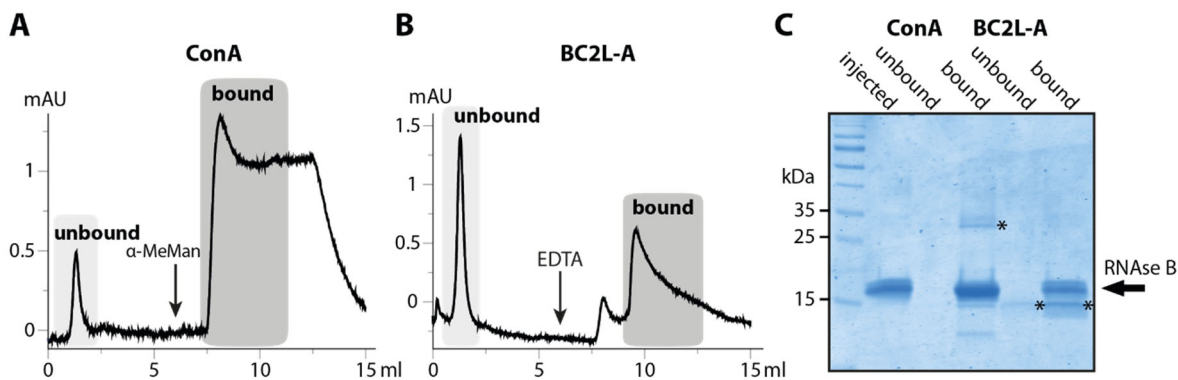


Figure 4: Oligomannose-type N-glycosylated RNase B binds to ConA and BC2L-A. (A) ConA and (B) BC2L-A lectin affinity chromatography of 10 µg bovine RNase B. Elution was achieved with 100 mM methyl α-D-mannoside or 1 mM EDTA, respectively. (C) Coomassie stained SDS-PAGE of the unbound and bound fractions revealing binding of RNase B. Additional bands (asterisks) correspond to loss of coupled lectins.

Enrichment of C-mannosylated and O-mannosylated peptides from cell extracts

To test the ability of the BC2L-A lectin to enrich C-mannosylated peptides from a complex mixture, we used cell extracts of two human cell lines, HEK293 and breast epithelial cell line MCF10A. After digestion with trypsin of extracted proteins and

treatment with PNGase F to remove N-glycans, lysates were applied to a two meter long BC2L-A column according to the previously described method for the enrichment of O-mannosylated peptides by ConA (Vester-Christensen et al., 2013).

Using this approach, C-mannosylated peptides of 21 human (and one bovine, originating from the cell

culture medium) proteins were identified in the bound fractions (Fig. 5A). These peptides contained a total of 45 C-mannosylated sites, including many

which were so far not shown to be C-mannosylated (Fig. 5B). Additionally, we detected O-mannosylated peptides from 62 proteins.

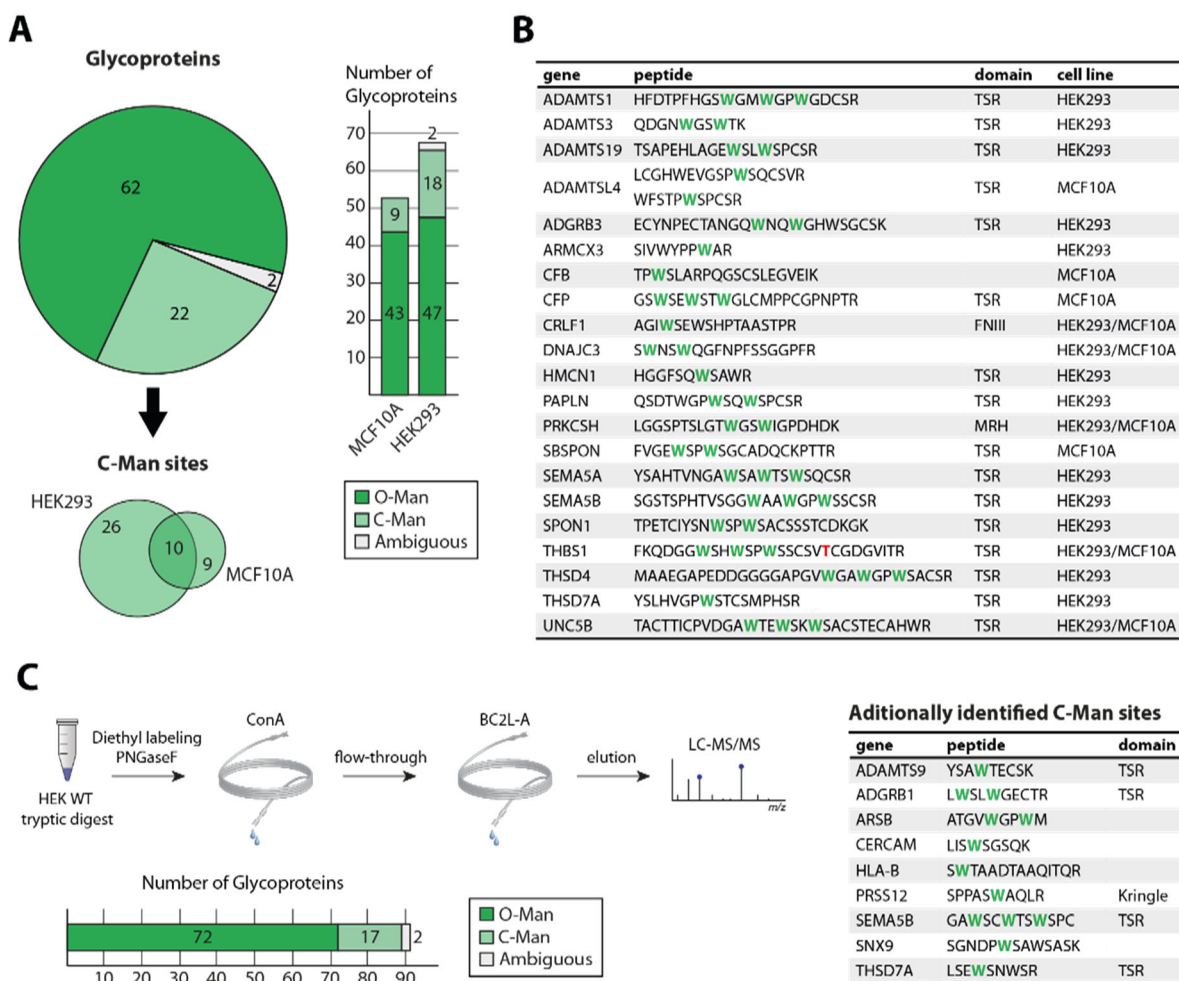


Figure 5: Mannosylated proteins identified from cell extracts after enrichment of corresponding C- and O-mannosylated glycopeptides by BC2L-A lectin affinity chromatography. (A) Identified C- and O-mannosylated proteins from HEK293 and MCF10A cell extracts. A total of 22 C-mannosylated unique proteins (21 human proteins and the bovine complement factor 9) were identified along with 62 O-mannosylated unique proteins by BC2L-A affinity chromatography. (B) List of the identified human C-mannosylated proteins by gene name and identified sequence. Tryptophans with an observed additional mass of 162 Da are depicted in green. O-fucosylated threonine is depicted in red. FNIII: Fibronectin type III domain of cytokine type I receptor family; MRH: Mannose-6-P receptor homology domain. (C) Additional C-mannosylated peptides identified by subsequent ConA and BC2L-A chromatography. The flowthrough of a ConA column was applied to a BC2L-A column. Details of the Identified peptides and corresponding MS/MS spectra can be found in supplemental item 4 and 5.

In the eluate of a ConA column, loaded with an identical HEK293 cell extract, we detected no C-mannosylated peptides. Moreover, the number of identified O-mannosylated peptides was lower than the number of peptides identified on BC2L-A (Suppl. Fig. S3). In addition, we analyzed a HEK293 cell extract by applying the flow through fraction of a ConA column to a BC2L-A column (Fig. 5C). The peptides of this sample were diethylated, which enhances ionization of peptides (Jung et al., 2019) and potentially leads to detection of a different set of peptides. We indeed identified eight peptides

(Fig. 5C) in the eluate of the BC2L-A column that were not identified in the sample presented in figure 5B. On the other hand, nine peptides were not identified in this sample (Supplemental item 4). Nineteen of all identified C-mannosylated peptides belong to TSR-containing proteins and revealed C-mannosylation on expected sites. Furthermore, the C-mannosylated peptide of the soluble cytokine receptor-like factor 1 (CRLF1) with a classical WSxWS motif and C-mannose on the first tryptophan was detected. Interestingly, two C-mannosylated peptides of the endoplasmic

reticulum resident proteins PRKCSH, the β -subunit of glucosidase II, and DNAJC3, a chaperone-like protein, were identified. Indeed, both peptides bear WxxW motifs; however, different from previously described sites, both tryptophan residues are C-mannosylated. Interestingly, five C-mannosylated peptides without WxxW/C motifs were detected. These belong to the extracellular complement factor B (CFB), CERCAM (a glycosyltransferase 25 family member with unknown activity), neurotrypsin (PRSS12), human leukocyte antigen (HLA-B or -C), and the putative mitochondrial armadillo repeat-containing X-linked protein 3 (ARMCX3) (Mou et al., 2009). Finally, we identified a C-mannosylated peptide of sorting nexin-9 (SNX9) with a classical WxxW consensus sequence. This protein is associated with clatrin coated vesicles, but lacks a secretion signal or predicted membrane domain (Lundmark and Carlsson, 2003). These data demonstrate that the lectin BC2L-A represents a highly functional tool for the enrichment and detection of C-mannosylated peptides from complex mixtures.

Discussion

C-mannose has been identified on a limited number of purified proteins, but there are no reports of the direct identification of C-mannosylated peptides from complex biological mixtures in proteomics studies without further enrichment. Here, we demonstrate that BC2L-A is a C-mannosyl peptide binding lectin which is compatible with lectin affinity chromatography enrichment of C-mannosylated peptides not only from purified proteins but also complex peptide mixtures to enhance their mass spectrometric detection. BC2L-A is able to retain peptides independent of the number of attached C-mannoses and presence of O-fucose.

In X-ray structures, C-mannose residues have been observed in both 4C_1 and 1C_4 conformations (Hamming et al., 2012; Klein et al., 2017). However, NMR studies show a preference for the unusual 1C_4 ring shape on RNase 2 and within TSRs (Frank et al., 2020; Jonker et al., 2020). The binding of mannose in BC2L-A is possible only in a 4C_1 conformation via the two calcium ions. To be able to bind to BC2L-A, the C-mannose thus needs to adopt a 4C_1 conformation. At least on peptides, the premise for this seems given. On intact proteins, inflexibility and

additional steric hindrance might impede binding to the lectin. We show here that BC2L-A does not bind intact thrombospondin repeats where the tryptophans intercalate with conserved arginine residues forming a so-called tryptophan-arginine ladder (Tan et al., 2002; Tossavainen et al., 2006). This structural feature partially shields the C-mannoses in the TSRs (Fig. 1A). It is, however, not excluded that BC2L-A can bind C-mannoses on intact proteins at more exposed regions.

For O-mannosylated peptides, BC2L-A could be an attractive alternative to ConA, especially for peptides with single O-mannoses. So far, separation of O-mannosylated peptides has been carried out by two meter long ConA columns to allow retaining of peptides with low binding affinity (Halim et al., 2015; Vester-Christensen et al., 2013; Winterhalter et al., 2013). We have used BC2L-A on both a short 1 ml column for the less complex samples and on a long column for the complex samples. In both cases, BC2L-A was superior for C-mannosylated peptides, but also showed better performance with O-mannosylated peptides than the ConA column. Next to the stronger binding of mannosylated peptides, BC2L-A also seems to be more selective against non-glycosylated peptides. The non-specific binding might be reduced due to the relative small size of BC2L-A compared to ConA. A further advantage of BC2L-A seems the possibility to elute bound glycopeptides with EDTA. Since calcium is directly involved in the interaction within the BC2L-A binding site but is loosely attached, a specific elution is already achieved by low EDTA concentrations. Although calcium acts as co-factor for ConA as well, elution with EDTA is not effective enough due to a stronger metal-lectin interaction (Stark and Sherry, 1979). Hence for elution, a high concentration of α -mannoside is required which in addition obscures the elution detection at 280 nm.

N-glycans are well retained by BC2L-A as well (Fig. 4). Co-crystal structure of BC2L-A and the trimannoside α Man1-3(α Man1-6) β Man revealed strong interactions of only one mannose residue (Lameignere et al., 2010), whereas the binding site of ConA interacts with all three mannoses of the same structure (Naismith and Field, 1996). BC2L-A might, therefore, retain a slightly different set of N-glycans, but both lectins quantitatively retained RNase B carrying oligomannose-type N-glycans. Crystal structures and kinetic data provide an explanation of the superior binding of BC2L-A to single α -mannose, as in the case of O- or C-

mannoses. Previous studies showed a ~30 fold higher affinity of BC2L-A to methyl- α -D-mannoside than ConA (Dam et al., 2005; Lameignere et al., 2010).

Whereas C-mannosylated peptides cannot be directly detected in total protein extracts of cells by mass spectrometric approaches, BC2L-A lectin affinity enrichment enables the identification of C-mannosylated peptides. Nevertheless, compared to the prediction of more than 2500 C-mannosylated proteins (Julenius, 2007), still relatively few C-mannosylated peptides have been detected in the present approach. This might result from diverse factors, like the presence of many LC-MS/MS incompatible peptide masses or the co-enrichment of N-glycosylated and O-mannosylated peptides. We exemplarily demonstrated that further C-mannosylated peptides could be identified by changes in the workflow. Diethylation of the peptides probably allowed detection of a partly different pool of peptides, but no overall increase of the number of identified peptides was observed by this or by a pre-extraction on a ConA column. Increase in the number of identifiable C-mannosylated peptides can, therefore, be expected from further diversification of protein digestion and peptide labeling methods.

The pool of newly identified C-mannosylation sites showed the existence of several C-mannosylated proteins that were never expected to be C-mannosylated. This showed that the previously used WxxW/C consensus sequence is of restricted value for the prediction of C-mannosylation sites. Taken together, we show that BC2L-A is a versatile tool to study C-mannosylation, with properties that might make it superior to other lectins for other types of glycosylation as well.

Materials and methods

Plasmid constructs

The BC2L-A gene from *Burkholderia cenocepacia* was cloned in the pET-22b plasmid (Novagen) with an N-terminal extension His tag (MGHHHHHH), in addition to the sequence in pdb:2WR9. *C. elegans* UNC-5 TSR2 in pMT/BIP (Invitrogen) was expressed and purified as described before (Shcherbakova et al., 2019). N-terminally His₆-tagged human thrombospondin-1 (THBS1) fragment containing the 3 TSRs (UniProt-P07996 amino acid 277-548)

was cloned via HindIII/XbaI in pSecTag-B (Invitrogen) and a V5-tag was introduced before the His-tag to allow detection by western blot (Fig. S1). Full-length *Mus musculus* N-cadherin (CDH1) was cloned in a modified pSecTag B vector containing a C-terminal Strep- and His₈-tag as described (Noor et al., 2021).

Preparation of lectin columns

For lectin purification, the BC2L-A plasmid was transformed in *E. coli* BL21-Gold and cultured in Power Broth medium with carbenicillin. At an OD₆₀₀ between 0.1 to 1, the expression was induced with 0.5 mM isopropyl- β -D-thiogalactopyranoside (IPTG). After 20 h incubation at 20 °C, the cells were harvested by centrifugation at 6000 g for 15 min. The pellet was resuspended in binding buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl and 20 mM imidazole) and cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche) followed by sonication (Branson Digital Sonifier). The lysate was centrifuged at 15,000 rpm and 4 °C for 30 min, filtered through a 0.8 μ m membrane filter (Merck) and applied to a 5 ml HisTrap™ HP (Cytiva). For the elution, a linear gradient of elution buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl and 500 mM imidazole) was applied over 8 ml. Fractions of the eluted protein were collected, concentrated to 50 mg/ml using a 3 kDa MWCO centrifugal tube (Amicon) and applied on a Superdex 75 10/300 GL column (Cytiva) equilibrated with coupling buffer (200 mM NaHCO₃, 500 mM NaCl, pH 8.3). Ten mg purified BC2L-A or ConA (Sigma) was coupled to a 1 ml HiTrap™ NHS-activated HP column (Cytiva) according to manufacturer's instruction.

Lectin affinity chromatography

The BC2L-A chromatography started with a 5 column volume (CV) wash in BC2L-A buffer A (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 100 μ M CaCl₂). After sample injection, the column was washed with 5-7 CV BC2L-A buffer A and fractions were collected. For the elution, the flow was switched to buffer B (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA) for 5-8 CV. Column reconditioning was performed by a 10 CV wash of BC2L-A buffer A. The same procedure was carried out for ConA, but with different buffers. ConA buffer A comprised 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM CaCl₂ and 1 mM MnCl₂ and ConA buffer B included additionally 0.1 M methyl α -D-mannoside. The ConA column was reconditioned by a 10 CV wash with 20 mM Tris-HCl pH 8.0, 150 mM NaCl.

Expression, Purification and tryptic digestion of *C.elegans* UNC5 TSR2

pMT-UNC5-TSR2 alone or together with C-mannosyltransferase DPY-19, were expressed in *Drosophila* S2 cells, as described before (Shcherbakova et al., 2019). The purified proteins from 10 ml culture were applied to C18 solid phase extraction (SPE, Empore C18-SD 7mm/3ml, 3M), eluted in 50% acetonitrile (ACN), dried and suspended in 100 μ l 100 mM ammonium bicarbonate (AMBIC) and 7 M urea. Disulfide bridges were reduced by the addition of 5 mM dithiothreitol (DTT), incubation for 45 min at 56 °C followed by alkylation with 14 mM iodoacetamide (IAA) for 30 min at room temperature, followed by 5 mM DTT for 15 min. After 5x dilution with 100 mM AMBIC, Trypsin Gold (Promega) was added and incubated at 37 °C over-night. The digestion was stopped with 1 mM phenylmethylsulfonyl fluoride and purified by C18 SPE, as described above (for Fig. 1) or purified by C18 chromatography (Xbridge TM Prep C18 5 μ m, 10 x 50 mm column, Waters) (for Fig. 3). In short, the sample was applied in 3% ACN/0.05 % TFA followed by elution with a 30 ml gradient to 50 % ACN/0.05 % TFA in order to obtain the di- or unmannosylated peptide. Elution fraction comprising peptides were evaporated. After evaporation, the purified peptides were resuspended in BC2L-A buffer A (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 100 μ M CaCl₂) and applied to lectin affinity chromatography.

Transient expression in CHO cells

CHO (Chinese Hamster Ovary) cells were transiently transfected with either THSB1 TSR1-3 or CDH1 using polyethyleneimine (PEIMAX 40000, Polysciences). The medium was exchanged 6 h after transfection by DMEM/Ham's F-12 (Biochrom) with 5% FCS (Biochrom) and the cells were incubated for 3 d at 37 °C and 5 % CO₂. Secreted protein was purified via nickel affinity chromatography as previously described (Shcherbakova et al., 2017). Purified proteins were additionally applied to size exclusion chromatography via Superdex 75 or 100 10/300 GL (Cytiva). Fractions of hTSP1 TSR 1-3 and mCDH1 were precipitated with acetone and resuspended in 100 mM Tris-HCl (pH 8.3) and 7 M urea. Reduction and alkylation was performed as described above. For the dilution, 100 mM Tris-HCl (pH 8.0) was used followed by addition of Trypsin Gold (Promega). After incubation over night at 37 °C, the digestion was stopped by heating at 95 °C for 5 min and

applied to C18 chromatography (Xbridge TM Prep C18 5 μ m, 10 x 50 mm column, Waters). After sample injection the column was equilibrated with 3% ACN/0.05 % TFA for 5 ml. The elution was performed by a 2 ml gradient from 3 % ACN/0.05 % TFA to 50 % ACN/0.05 % TFA. The peptide fraction was collected, evaporated, resuspended in BC2L-A buffer A (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 100 μ M CaCl₂) and applied to lectin affinity chromatography.

Total cell tryptic digests

Adherent HEK293 AC2 cells were grown in DMEM medium (Sigma) supplemented with 10% FBS (Gibco), 1% GlutaMAX (Gibco), 1% Penicillin/Streptomycin (Gibco) while adherent MCF10A cells were grown in DMEM/F12 medium (Invitrogen) supplemented with 5% Horse Serum (Invitrogen), 20 ng/mL EGF (Sigma), 0.5 mg/mL Hydrocortisone (Sigma), 100 ng/mL Cholera Toxin (Sigma) and 10 μ g/mL Insulin (Sigma). Cells were washed in PBS twice and harvested at subconfluence by scraping. Total cell tryptic digests were prepared as previously described (Larsen et al., 2017b). Briefly, proteins were extracted by sonication in 50 mM ammonium bicarbonate, 0.1% Rapigest (Waters). Following reduction (5 mM DTT) and alkylation (10 mM IAA), the protein extract was digested with trypsin (Roche) and desalted on Sep-Pak C18 cartridges (Waters). Additional diethyl labelling of the tryptic peptides was performed as previously described (Jung et al., 2019). Finally, the tryptic digests were treated with PNGase F (Roche) at 37°C for 16 h.

Preparation and chromatography using a long lectin column

Purified BC2L-A lectin was conjugated to NHS-Agarose (Thermo Fisher Scientific) according to manufacturer's protocol and BC2L-A-agarose slurry was packed in PFA 1/16" (OD) x 0.040" (ID) tubing (IDEX Health & Science) as previously described (Steentoft et al., 2013) to a final column length of ~2 m. The 2 m long BC2L-A column was equilibrated in BC2L-A running buffer (25mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1 mM CaCl₂) at 4°C using a ÄKTA FPLC system (GE Healthcare Europe GmbH, Denmark). PNGase F digested tryptic peptides were injected onto the 2 m BC2L-A column and washed with BC2L-A running buffer until absorbance (214 nm) was <5 mAU. Finally, bound C- and O-mannosylated glycopeptides were eluted by injecting 3 mL 25 mM Tris-HCl pH 7.4, 150 mM NaCl,

0.1 mM CaCl₂, 20 mM EDTA and collected in six 500 µL fractions. Diethyl labelled and PNGase F treated peptides of HEK293 WT total cell tryptic digests were injected into a 3m ConA LWAC column prepared as previously described (Larsen et al., 2017a) and flow-through fractions were collected. The unbound peptides were desalted using a Sep-Pak C18 cartridge (Waters), injected into a 3.5 m BC2L-A column and separated similar to the 2 m BC2L-A column runs. Elution fractions were desalted by C18 Stage Tips (Empore), evaporated to dryness, and reconstituted in 20 µL 0.1% formic acid before mass spectrometric analyses.

Mass spectrometric analysis

LC-MS/MS analysis of BC2L-A bound and unbound UNC-5 TSR2 peptides (Fig. 1) was performed as described before (Buettner et al., 2013). Briefly, the evaporated samples were resuspended in 2 % ACN, 0.1 % formic acid, loaded and separated on an analytical C18 column (Waters, BEH130 C18, 100 µm x 100 mm, 1.7 µm particle size) using a nanoACQUITY UPLC System. The chromatographic system was online coupled to the ESI Q-TOF Ultima (Waters) operating in positive ion mode. Full MS scans were acquired between m/z 400–2,300. Fragment ions were generated in a data-dependent mode of the two most intensive precursors and MS/MS was acquired within a range between m/z 50–1,600. Spectra of non- and mannosylated peptides were manually analyzed using the MassLynx V4.1 Software (Waters) and extracted ion chromatograms were generated with a mass error of 0.3 Da.

LC-MS/MS analysis of bound and unbound thrombospondin, N-Cadherin peptides and UNC-5 TSR2 peptide 1 (Fig. 2, 3 and Suppl. Fig. S2) was carried out as described previously (Hoffmann et al., 2018). Briefly, samples were resuspended in 2% ACN, 0.05% TFA, injected onto a nano trap column (Acclaim PepMap100 C18, 5 µm, 100Å, 100 µm i.d. x 2 cm, Thermo Scientific) and subsequently separated on a C18 nano separation column (Acclaim PepMap RSLC C18, 2 µm, 100Å, 75 µm i.d. x 25 cm, Thermo Scientific) using an UltiMate 3000 RSLCnano system (Thermo Scientific). The chromatographic system was connected online via a nanoelectrospray source (Nanospray Flex™ ion source, NSI, Thermo Scientific) to the LTQ Orbitrap Elite hybrid mass spectrometer (Thermo Scientific) operated in positive ion mode. Full MS scans were acquired between m/z 300–2,000, at a resolution of 30,000 using the orbitrap mass analyzer. Fragment

ions were generated in a data-dependent mode where the top five highest intensity precursors (charge state z≥2) were selected for sequential higher-energy collisional dissociation (HCD) fragmentation. HCD spectra were acquired with a normalized collisional energy (NCE) of 20 on the orbitrap mass analyzer. For identification and relative quantification of (glyco)peptides, the Bionic software platform v3.10.4 (Protein Metrics Inc.) was used. The following parameters were used for the search with a FASTA sequence corresponding to the recombinantly expressed proteins: 5 ppm precursor ion mass tolerance; 20 ppm fragment ion mass tolerance; semispecific tryptic digestion with up to two missed cleavages; fixed modifications: carbamidomethylation of cysteine (+57.0214 Da); variable modifications: oxidation of methionine and tryptophan (+15.9949 Da), deamidation of asparagine and glutamine (+0.9840 Da), up to 5 hexose residues attached to serine and threonine (+162.0528 Da) for O-mannosylation, up to 3 hexose residues attached to tryptophan (+162.0528 Da) for C-mannosylation, either a fucose (+146.0579 Da) or fucose and glucose (+308.1107 Da) residue attached to serine and threonine for O-fucosylation. A score threshold of 100 was used for the initial filtering of results. MS-based relative peptide quantification was achieved by comparison of the area under the curve (AUC) derived from extracted-ion chromatograms of the respective precursor ions (XIC on MS-level). All annotated spectra and identified peptides are listed in the supplemental item 1–3.

Mass spectrometric analysis of C- and O-mannosylated peptides from total cell tryptic digests (Fig. 5) was carried out as previously described (Larsen et al., 2017b). Briefly, 10 µL of each BC2L-A elution fraction was injected using a EASY-nLC 1000 system (Thermo Fisher Scientific) interfaced via a nanoSpray Flex ion source to an Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). The EASY-nLC 1000 was operated using a single analytical column setup (PicoFrit Emitters, 75-µm inner diameter; New Objectives, Woburn, MA) packed in-house with Reprosil-Pure-AQ C18 phase (1.9-µm particle size; Dr. Maisch). Peptides were separated across 120 min at 200 nL/min; gradient elution was achieved by solvent A (0.1% formic acid) and solvent B (acetonitrile, 0.1% formic acid) which was set for 2–20% B for 95 min followed by 20–80% B for 10 min and finally 80% B for 15 min. Precursor MS1 scan (m/z 355–1700) was acquired

in the Orbitrap at a resolution setting of 120,000, followed by Orbitrap HCD-MS/MS and ETciD-MS/MS of multiply charged precursors ($z = 2-6$) in the MS1 spectrum; a minimum MS1 signal threshold of 10,000–50,000 ions was used for triggering data-dependent fragmentation events; MS2 spectra were acquired at a resolution of 60,000 (HCD and ETciD). Data analyses was carried out using Proteome Discoverer 1.4 software (Thermo Fisher Scientific). Data files (.raw and .mgf) were processed using the Sequest HT node and searched against the canonical human proteome downloaded (January 2013) from the UniProtKB database (<http://www.uniprot.org/>). Precursor mass tolerance was set to 10 ppm and fragment ion mass tolerance to 0.02 Da. Up to 2 missed trypsin (full- and semi-specific) cleavages were allowed. Carbamidomethylation (57.02146 Da) was set as a fixed modification for cysteine, and oxidation (15.9949 Da) was set as variable modification for methionine. In addition, Hex (162.0528 Da) was set as variable modification for serine, threonine, and tryptophan while dHex-Hex (308.1107 Da) was set as variable modifications for serine and threonine. Peptide confidence levels were calculated using the Target Decoy PSM Validator node and results were filtered for high-confidence ($p < 0.01$) identifications only. All C- and O-mannosylated peptide spectral matches were inspected manually to validate the accuracy of the identifications. The identified peptides and the annotated spectra list of C-mannosylated peptides can be found in the supplemental item 4 – 6.

SDS-PAGE and Western Blotting

Lectin affinity peak fractions were collected and precipitated in 80 % acetone at $-20\text{ }^{\circ}\text{C}$ overnight. Precipitated proteins were resuspended in 1xLaemmli and separated by 15 % SDS-PAGE. Coomassie staining was performed by ROTI®Blue (Carl Roth) over night and western blots by semi-dry procedure on nitrocellulose membranes. For the detection, mouse anti-V5 (Acris Antibodies) or mouse anti-His₅ (Qiagen) antibody followed by IRDye 800 conjugated goat anti-mouse antibody (LI-COR) were used and scanned on a LI-COR Odyssey Infrared Scanner.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (FOR 2509 Projektnummer: 289991887) – Project-ID BA4091/6-2 (to H.B.),

BU2920/2-2 (to F.F.R.B.), RA2992/1-2 (to E.R.), STR 443/5-1 and STR 443/6-1 (to S.S.). A.I. was supported by PIA Glyco@Alps (ANR-15-IDEX-02) and Labex Arcane/CBH-EUR-GS (ANR-17-EURE-0003). L.P., S.Y.V. and A.H. were supported by the Danish National Research Foundation (DNRF107) and a research grant (00025438) from VILLUM FONDEN.

References

- Ambort, D., Van Der Post, S., Johansson, M.E.V., MacKenzie, J., Thomsson, E., Kregel, U., and Hansson, G.C. (2011). Function of the CysD domain of the gel-forming MUC2 mucin. *Biochem. J.* *436*, 61–70.
- Baenziger, J.U., and Fiete, D. (1979). Structural determinants of concanavalin A specificity for oligosaccharides. *J. Biol. Chem.* *254*, 2400–2407.
- Bonnardel, F., Kumar, A., Wimmerova, M., Lahmann, M., Perez, S., Varrot, A., Lisacek, F., and Imberty, A. (2019). Architecture and Evolution of Blade Assembly in β -propeller Lectins. *Structure* *27*, 764-775.e3.
- Buettner, F.F.R., Ashikov, A., Tiemann, B., Lehle, L., and Bakker, H. (2013). *C. elegans* DPY-19 Is a C-Mannosyltransferase Glycosylating Thrombospondin Repeats. *Mol. Cell* *50*, 295–302.
- Bunkenborg, J., Pilch, B.J., Podtelejnikov, A. V., and Wiśniewski, J.R. (2004). Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics* *4*, 454–465.
- Cirkseña, K., Hütte, H.J., Shcherbakova, A., Thumberger, T., Sakson, R., Weiss, S., Jensen, L.R., Friedrich, A., Todt, D., Kuss, A.W., et al. (2021). The C-Mannosylome of Human Induced Pluripotent Stem Cells Implies a Role for ADAMTS16 C-Mannosylation in Eye Development. *Mol. Cell. Proteomics* *20*, 100092.
- Dam, T.K., Oscarson, S., Roy, R., Das, S.K., Pagé, D., Macaluso, F., and Brewer, C.F. (2005). Thermodynamic, kinetic, and electron microscopy studies of concanavalin A and Dioclea grandiflora lectin cross-linked with synthetic divalent carbohydrates. *J. Biol. Chem.* *280*, 8640–8646.
- Van Damme, E.J.M. (2011). Lectins as Tools to Select for Glycosylated Proteins. In *Methods in Molecular Biology*, pp. 289–297.
- Doucey, M. a, Hess, D., Cacan, R., and Hofsteenge, J. (1998). Protein C-mannosylation is enzyme-catalysed and uses dolichyl-phosphate-mannose as a precursor. *Mol. Biol. Cell* *9*, 291–300.
- Duncan, M.W., Aebersold, R., and Caprioli, R.M. (2010). The pros and cons of peptide-centric proteomics. *Nat. Biotechnol.* *28*, 659–664.
- Ervin, L.A., Ball, L.E., Crouch, R.K., and Schey, K.L. (2005). Phosphorylation and Glycosylation of Bovine Lens MP20. *Investig. Ophthalmology Vis. Sci.* *46*, 627.
- Frank, M., Beccati, D., Leeflang, B.R., and Vliegthart, J.F.G.

- (2020). C-Mannosylation Enhances the Structural Stability of Human RNase 2. *IScience* 23, 101371.
- Freeze, H.H. (1995). Lectin Affinity Chromatography. *Curr. Protoc. Protein Sci.* 00.
- Fu, D., Chen, L., and O'Neill, R.A. (1994). A detailed structural characterization of ribonuclease B oligosaccharides by ¹H NMR spectroscopy and mass spectrometry. *Carbohydr. Res.* 261, 173–186.
- Furmanek, A., and Hofsteenge, J. (2000). Protein C-mannosylation: facts and questions. *Acta Biochim. Pol.* 47, 781–789.
- Furmanek, A., Hess, D., Rogniaux, H., and Hofsteenge, J. (2003). The WSAWS motif is C-hexosylated in a soluble form of the erythropoietin receptor. *Biochemistry* 42, 8452–8458.
- Gonzalez de Peredo, A., Klein, D., Macek, B., Hess, D., Peter-Katalinic, J., and Hofsteenge, J. (2002). C-Mannosylation and O-Fucosylation of Thrombospondin Type 1 Repeats. *Mol. Cell. Proteomics* 1, 11–18.
- Halim, A., Larsen, I.S.B., Neubert, P., Joshi, H.J., Petersen, B.L., Vakhrushev, S.Y., Strahl, S., and Clausen, H. (2015). Discovery of a nucleocytoplasmic O-mannose glycoproteome in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15648–15653.
- Hamming, O.J., Kang, L., Svensson, A., Karlsen, J.L., Rahbek-Nielsen, H., Paludan, S.R., Hjorth, S.A., Bondensgaard, K., and Hartmann, R. (2012). Crystal Structure of Interleukin-21 Receptor (IL-21R) Bound to IL-21 Reveals That Sugar Chain Interacting with WSXWS Motif Is Integral Part of IL-21R. *J. Biol. Chem.* 287, 9454–9460.
- Hoffmann, M., Pioch, M., Pralow, A., Hennig, R., Kottler, R., Reichl, U., and Rapp, E. (2018). The Fine Art of Destruction: A Guide to In-Depth Glycoproteomic Analyses-Exploiting the Diagnostic Potential of Fragment Ions. *Proteomics* 18, 1800282.
- Hofsteenge, J., Blommers, M., Hess, D., Furmanek, A., and Miroshnichenko, O. (1999). The Four Terminal Components of the Complement System Are C-Mannosylated on Multiple Tryptophan Residues. *J. Biol. Chem.* 274, 32786–32794.
- Hofsteenge, J., Huwiler, K.G., Macek, B., Hess, D., Lawler, J., Mosher, D.F., and Peter-Katalinic, J. (2001). C-Mannosylation and O-Fucosylation of the Thrombospondin Type 1 Module. *J. Biol. Chem.* 276, 6485–6498.
- Honigberg, L., and Kenyon, C. (2000). Establishment of left/right asymmetry in neuroblast migration by UNC-40/DCC, UNC-73/Trio and DPY-19 proteins in *C. elegans*. *Development* 127, 4655–4668.
- John, A., Järvå, M.A., Shah, S., Mao, R., Chappaz, S., Birkinshaw, R.W., Czabotar, P.E., Lo, A.W., Scott, N.E., and Goddard-Borger, E.D. (2021). Yeast- and antibody-based tools for studying tryptophan C-mannosylation. *Nat. Chem. Biol.* 17, 428–437.
- Jonker, H.R.A., Saxena, K., Shcherbakova, A., Tiemann, B., Bakker, H., and Schwalbe, H. (2020). NMR Spectroscopic Characterization of the C-Mannose Conformation in a Thrombospondin Repeat Using a Selective Labeling Approach. *Angew. Chemie Int. Ed.* 59, 20659–20665.
- Julenius, K. (2007). NetCGlyc 1.0: prediction of mammalian C-mannosylation sites. *Glycobiology* 17, 868–876.
- Jung, J., Jeong, K., Choi, Y., Kim, S.A., Kim, H., Lee, J.W., Kim, V.N., Kim, K.P., and Kim, J.-S. (2019). Deuterium-Free, Three-Plexed Peptide Diethylation for Highly Accurate Quantitative Proteomics. *J. Proteome Res.* 18, 1078–1087.
- Kaji, H., Kamiie, J.I., Kawakami, H., Kido, K., Yamauchi, Y., Shinkawa, T., Taoka, M., Takahashi, N., and Isobe, T. (2007). Proteomics reveals n-linked glycoprotein diversity in *Caenorhabditis elegans* and suggests an atypical translocation mechanism for integral membrane proteins. *Mol. Cell. Proteomics* 6, 2100–2109.
- Klein, C., Waldhauer, I., Nicolini, V.G., Freimoser-Grundschober, A., Nayak, T., Vugts, D.J., Dunn, C., Bolijn, M., Benz, J., Stihle, M., et al. (2017). Cergutuzumab amunaleukin (CEA-IL2v), a CEA-targeted IL-2 variant-based immunocytokine for combination cancer immunotherapy: Overcoming limitations of aldesleukin and conventional IL-2-based immunocytokines. *Oncoimmunology* 6, 1–15.
- Kornfeld, K., Reitman, M.L., and Kornfeld, R. (1981). The carbohydrate-binding specificity of pea and lentil lectins. Fucose is an important determinant. *J. Biol. Chem.* 256, 6633–6640.
- Lameignere, E., Malinová, L., Slávková, M., Duchaud, E., Mitchell, E.P., Varrot, A., Šedo, O., Imberty, A., and Wimmerová, M. (2008). Structural basis for mannose recognition by a lectin from opportunistic bacteria *Burkholderia cenocepacia*. *Biochem. J.* 411, 307–318.
- Lameignere, E., Shiao, T.C., Roy, R., Wimmerova, M., Dubreuil, F., Varrot, A., and Imberty, A. (2010). Structural basis of the affinity for oligomannosides and analogs displayed by BC2L-A, a *Burkholderia cenocepacia* soluble lectin. *Glycobiology* 20, 87–98.
- Larsen, I.S.B., Narimatsu, Y., Joshi, H.J., Yang, Z., Harrison, O.J., Brasch, J., Shapiro, L., Honig, B., Vakhrushev, S.Y., Clausen, H., et al. (2017a). Mammalian O-mannosylation of cadherins and plexins is independent of protein O-mannosyltransferases 1 and 2. *J. Biol. Chem.* 292, 11586–11598.
- Larsen, I.S.B., Narimatsu, Y., Joshi, H.J., Siukstaite, L., Harrison, O.J., Brasch, J., Goodman, K.M., Hansen, L., Shapiro, L., Honig, B., et al. (2017b). Discovery of an O-mannosylation pathway selectively serving cadherins and protocadherins. *Proc. Natl. Acad. Sci.* 114, 11163–11168.
- Lundmark, R., and Carlsson, S.R. (2003). Sorting Nexin 9 Participates in Clathrin-mediated Endocytosis through Interactions with the Core Components. *J. Biol. Chem.* 278, 46772–46781.
- Mitchell, E., Houles, C., Sudakevitz, D., Wimmerova, M., Gautier, C., Pérez, S., Wu, A.M., Gilboa-Garber, N., and Imberty, A. (2002). Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients. *Nat. Struct. Biol.* 9, 918–921.
- Mou, Z., Tapper, A.R., and Gardner, P.D. (2009). The Armadillo repeat-containing protein, ARM CX3, physically and functionally interacts with the developmental regulatory factor Sox10. *J. Biol. Chem.* 284, 13629–13640.
- Naismith, J.H., and Field, R.A. (1996). Structural basis of trimannoside recognition by concanavalin A. *J. Biol. Chem.* 271, 972–976.
- Neubert, P., Halim, A., Zauser, M., Essig, A., Joshi, H.J.,

- Zatorska, E., Larsen, I.S.B., Loibl, M., Castells-Ballester, J., Aebi, M., et al. (2016). Mapping the O-Mannose glycoproteome in *saccharomyces cerevisiae*. *Mol. Cell. Proteomics* *15*, 1323–1337.
- Nishikawa, T., Kajii, S., Sato, C., Yasukawa, Z., Kitajima, K., and Isobe, M. (2004). α -C-Mannosyltryptophan is not recognized by conventional mannose-binding lectins. *Bioorganic Med. Chem.* *12*, 2343–2348.
- Niwa, Y., Suzuki, T., Dohmae, N., and Simizu, S. (2016). Identification of DPY19L3 as the C-mannosyltransferase of R-spondin1 in human cells. *Mol. Biol. Cell* *27*, 744–756.
- Noor, S.I., Hoffmann, M., Rinis, N., Bartels, M.F., Winterhalter, P.R., Hoelscher, C., Hennig, R., Himmelreich, N., Thiel, C., Ruppert, T., et al. (2021). Glycosyltransferase POMGNT1 deficiency strengthens N-cadherin-mediated cell–cell adhesion. *J. Biol. Chem.* *296*, 100433.
- Perez-Vilar, J. (2004). C-Mannosylation of MUC5AC and MUC5B Cys subdomains. *Glycobiology* *14*, 325–337.
- Sasazawa, Y., Sato, N., Suzuki, T., Dohmae, N., and Simizu, S. (2015). C-mannosylation of thrombopoietin receptor (c-Mpl) regulates thrombopoietin-dependent JAK-STAT signaling. *Biochem. Biophys. Res. Commun.* *468*, 262–268.
- Sharon, N. (2008). Lectins: Past, present and future. *Biochem. Soc. Trans.* *36*, 1457–1460.
- Shcherbakova, A., Tiemann, B., Buettner, F.F.R., and Bakker, H. (2017). Distinct C-mannosylation of netrin receptor thrombospondin type 1 repeats by mammalian DPY19L1 and DPY19L3. *Proc. Natl. Acad. Sci.* *114*, 2574–2579.
- Shcherbakova, A., Preller, M., Taft, M.H., Pujols, J., Ventura, S., Tiemann, B., Buettner, F.F.R., and Bakker, H. (2019). C-mannosylation supports folding and enhances stability of thrombospondin repeats. *Elife* *8*, 1–16.
- Stark, C.A., and Sherry, A.D. (1979). Metal ion induced conformational changes in concanavalin A: Evidence for saccharide binding to one metal free structure. *Biochem. Biophys. Res. Commun.* *87*, 598–604.
- Steentoft, C., Bennett, E.P., and Clausen, H. (2013). Glycoengineering of Human Cell Lines Using Zinc Finger Nuclease Gene Targeting: SimpleCells with Homogeneous GalNAc O-glycosylation Allow Isolation of the O-glycoproteome by One-Step Lectin Affinity Chromatography. pp. 387–402.
- Suttapitugsakul, S., Sun, F., and Wu, R. (2020). Recent Advances in Glycoproteomic Analysis by Mass Spectrometry. *Anal. Chem.* *92*, 267–291.
- Tan, K., Duquette, M., Liu, J.H., Dong, Y., Zhang, R., Joachimiak, A., Lawler, J., and Wang, J.H. (2002). Crystal structure of the TSP-1 type 1 repeats: A novel layered fold and its biological implication. *J. Cell Biol.* *159*, 373–382.
- Tossavainen, H., Pihlajamaa, T., Huttunen, T.K., Raulo, E., Rauvala, H., Permi, P., and Kilpeläinen, I. (2006). The layered fold of the TSR domain of *P. falciparum* TRAP contains a heparin binding site. *Protein Sci.* *15*, 1760–1768.
- Verbij, F.C., Stokhuijzen, E., Kaijen, P.H.P., Van Alphen, F., Meijer, A.B., and Voorberg, J. (2016). Identification of glycans on plasma-derived ADAMTS13. *Blood* *128*, e51–e58.
- Vester-Christensen, M.B., Halim, A., Joshi, H.J., Steentoft, C., Bennett, E.P., Lavery, S.B., Vakhrushev, S.Y., and Clausen, H. (2013). Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 21018–21023.
- Winterhalter, P.R., Lommel, M., Ruppert, T., and Strahl, S. (2013). O-glycosylation of the non-canonical T-cadherin from rabbit skeletal muscle by single mannose residues. *FEBS Lett.* *587*, 3715–3721.
- Wright, C.S., and Hester, G. (1996). The 2.0 Å structure of a cross-linked complex between snowdrop lectin and a branched mannopentaose: Evidence for two unique binding modes. *Structure* *4*, 1339–1352.
- Zielinska, D.F., Gnad, F., Wiśniewski, J.R., and Mann, M. (2010). Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* *141*, 897–907.