Recognition of Complex Core-Fucosylated N-Glycans by a Mini Lectin

Aurore Cabanettes, Lukas Perkams, Carolina Spies, Carlo Unverzagt, Annabelle Varrot

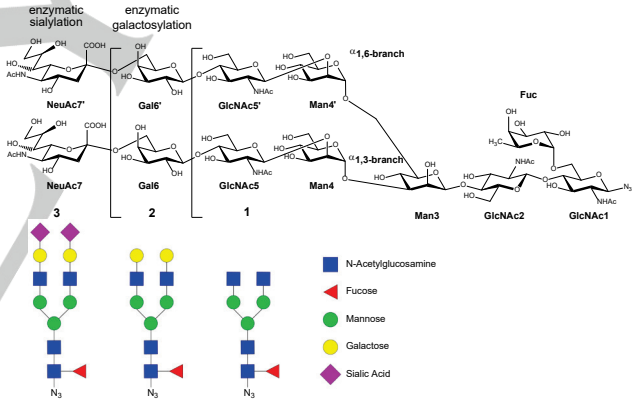
Recognition of complex core fucosylated N-glycans by a mini lectin

Aurore Cabanettes[a], Lukas Perkams[b], Carolina Spies[b], Carlo Unverzagt[b], Annabelle Varrot*[a]

Abstract: *We recombinantly produced and characterized the mini fungal lectin PhoSL. Despite a length of only 40 amino acids, PhoSL exclusively recognizes N-glycans with α1,6-linked fucose. Core fucosylation influences the intrinsic properties and bioactivities of mammalian N-glycoproteins and its level is linked to various cancers. Thus, PhoSL serves as a promising tool for glycoprofiling. Without structural precedence, we solved its crystal structure using the zinc anomalous signal. We revealed an interlaced trimer creating a novel protein fold termed β-prism III. Three biantennary core fucosylated N-glycan azides of 8 to 12 sugars were cocrystallized with PhoSL. The resulting highly resolved structures gave a detailed view on how the exclusive recognition of α1,6-fucosylated N-glycans by such a small protein occurs. This work also provided a protein consensus motif for the observed specificity as well as a glimpse on N-glycan flexibility upon binding.*

α1,6-Fucosylation (core fucosylation) consists of the transfer of an α1,6-linked fucose to the innermost N-acetyl-glucosamine (GlcNAc) of the N-glycan core of mammalian glycoproteins by the α1,6-fucosyltransferase (FUT8).[1] This post-translational modification regulates many biological processes such as cell adhesion or innate immunity. It is associated with pathologies like cancers, where altered glycosylation is a hallmark.[2] Enhanced core fucosylation is a preeminent factor in liver cancer and melanoma metastasis.[3] Several core fucosylated glycoproteins, such as alpha-fetoprotein (AFP-L3),[1-2] prostate-specific antigen (PSA)[4] or haptoglobin[5] are used as serological biomarkers for the detection of hepatocellular carcinoma, prostate and pancreatic cancers, respectively. Core fucosylation strongly affects the bioactivity of therapeutical N-glycoproteins. Its presence in the Fc region of monoclonal antibodies (mABs) impairs their recognition by FcγRIIIA receptors and thus, lowers their efficacy in antibody dependent cellular cytotoxicity (ADCC).[6] Hence, molecular tools for the rapid, easy and accurate analysis of core fucosylation in biomedical applications are needed.[7]

Lectins can be such tools since they are very versatile and can distinctly recognize particular glycan structures in situ.[8] Only five lectins with strict specificity for core fucosylation have been identified.[9] Based on primary protein sequence, they belong to, at least, two new protein families with unidentified fold where neither the 3D structure nor the molecular basis for selectivity is known. The first family, represented by BTL from the red alga *Bryothamnion triquetrum*, recognizes preferentially bi- and tri-antennary non-bisected core fucosylated N-glycans.[9a] The second family contains RSL from *Rhizopus stolonifer* and PhoSL from *Pholiota squarrosa*, the smallest fungal lectin to date with only 40 amino acids[9c, 9d] as well as SL2-1[9e] from the bacterium *Streptomyces rapamycinicus* presenting a triple repeat of PhoSL. PhoSL and SL2-1 recognize a variety of core fucosylated N-glycans: bi-, tri-, and tetraantennary as well as bisected.[9c, 9e] PhoSL detects core fucosylation on native glycoproteins, such as IgGs or thyroglobulin and core fucosylated biomarkers such as AFP-L3 and haptoglobin.[9c, 10] To get structural insights into its recognition of core fucosylation, we solved the structure of recombinant PhoSL (rPhoSL) in its apo form and in complex with three biantennary core fucosylated oligosaccharide azides of different length (Scheme 1).



Scheme 1. Core fucosylated biantennary oligosaccharide azides **1-3** used in this study.

First, we recombinantly produced PhoSL to allow better availability and reproducibility of the protein for future applications. Assuming disulfide formation and secretion for PhoSL, it was overexpressed into the periplasm of *Escherichia coli*, where the bacterial disulfide bond machineries are located. A cleavable and His6-tagged N-terminal fusion comprising the disulfide isomerase DsbA with its signal sequence was introduced to promote export into the periplasmic space, introduce disulfide bonds into the nascent protein and improve protein folding[11] (Figure 1a). After cleavage of the fusion with Tobacco Etch Virus (TEV) protease[12] the purified rPhoSL showed alike properties to the wild-type protein isolated from mushroom.[9c] Agglutination of rabbit red blood cells was found and strict specificity for α1,6-core fucosylated N-glycans was confirmed by glycan array analysis at the Consortium for Functional Glycomics. The latest glycan array

[a] A. Cabanettes, Dr. A. Varrot
Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France
E-mail: annabelle.varrot@cermav.cnrs.fr
[b] L. Perkams, C Spies, Prof. Dr Carlo Unverzagt
Bioorganische Chemie, Gebäude NW1, Universität Bayreuth, 95440 Bayreuth, Germany

Supporting information can be found via a link xxxx.

version demonstrated that rPhoSL, as SL2-1, also binds to core fucosylated N-glycans including those bearing polylactosamine motifs (Figure S1).[9e] In an enzyme-linked lectin assay (ELLA), biotinylated rPhoSL recognized immobilized core fucosylated glycoproteins such as thyroglobulin, fetuin and IgG-k but not conalbumin where core fucosylation is absent (Figure S2a). A submillimolar affinity was measured by isothermal micro-calorimetry with a Kd of 180 and 105 µM for core fucosylated chitobiose and N-glycan **1**, respectively (Figures S2b-c). Unexpectedly, this is 100 fold lower than for SL2-1.[9e] ITC measurements with L-fucose showed no binding. Size exclusion chromatography indicated a trimeric state of PhoSL in solution (data not shown). After initial screening, rPhoSL was crystallized in the presence of zinc and its apo structure was solved by single anomalous dispersion (SAD) at 1.7 Å resolution using the zinc anomalous signal (Table S1). rPhoSL was found as a trimer in the asymmetric unit with zinc atoms stabilizing surface loops (Figure 1b). The first 7 residues of chain A could not be modelled probably as a result of disorder in the crystal perhaps induced by a zinc atom close by. These residues were fully resolved in the other crystal forms obtained and described below (Figure 1c).

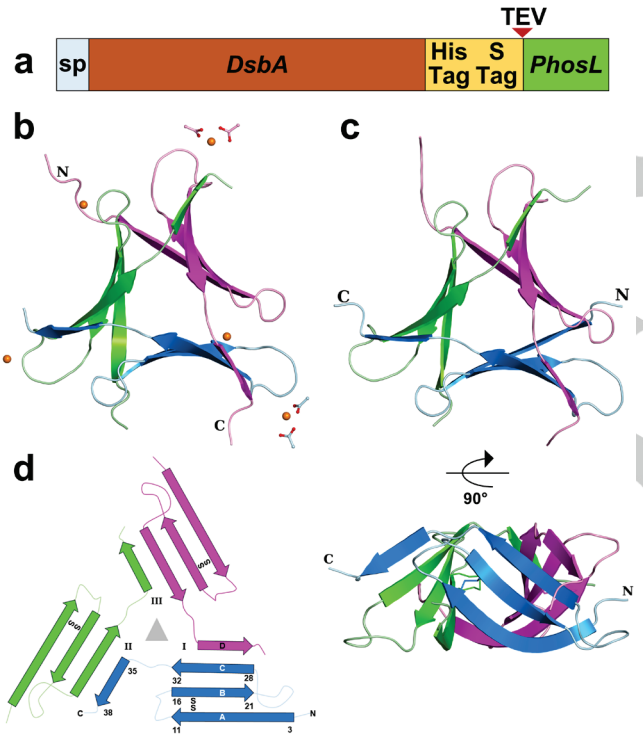


Figure 1. (a) rPhoSL construct for periplasmic export. (b) Carton representation of the β-prism III for the apo protein. rPhoSL trimer is colored by protein chain. Zinc atoms are shown as orange spheres and acetate ions as ball and sticks. (c) Top and side view of the β-prism III of rPhoSL in cocrystals with N-glycans. (d) Topology diagram of rPhoSL trimer.

rPhoSL forms a slightly curved four-stranded antiparallel β-sheet with β-strands termed A, B, C and D from *N*- to *C*-terminus and strands A and D oriented about 90º apart. Three protomers arrange around a three-fold symmetry axis to form a twelve-stranded β-prism. The β-sheets are not parallel to the prism axis but rotated by 45º and form a Velcro-like closure due to the swapping of β-strand D that stabilizes the trimer by interacting with β-strand C of a neighboring protein chain (Figures 1b-d). Additional contacts are observed between the extremities of strands A (Figure S3a). The β–prism presents an equilateral triangular shape with a 30º twist from its bottom, defined by the long strand A of each sheet to its top, formed by loop connecting strands C and D (Figures 1b-d). Its center is mainly filled by hydrophobic residues stacked in layers with V5 and L8 from strand A, W32 from strand C and the disulfide bridge C10-C17 between strand A and B (Figure S3b).

The four stranded β-meander adopted by rPhoSL is a common protein motif found in the greek key motif, β-propellers, β-prism II or β-barrels.[13] At the protomer level, PhoSL presents structural homology primarily with blades of many β-propeller proteins.[14] For the overall trimer, the search for structural homologs with PDBeFold[13b] revealed only very remote homology with the β-prism II fold (b.78 in SCOPe 2.06[13a]), first identified for the mannose specific *Galanthus nivalis* agglutinin (GNA) from snowdrop.[15] PhoSL and GNA can only be overlaid through one β-sheet since the assembly and the layout of their β-strands are very distinct leading to an opposite directionality of their β-sheets (Figures 2 and S4). Hence, PhoSL presents a novel structural fold for all β-proteins that we termed β-prism III.

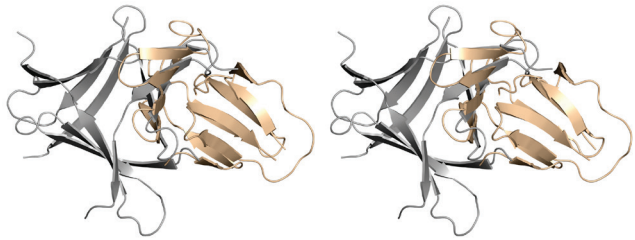


Figure 2. Stereo view of the overlay of β-prism III of PhoSL (grey) with β-prism II represented by the *Galenthus nivalis* agglutinin (GNA, PDB 1MSA, wheat).[15]

To obtain the molecular basis for rPhoSL specificity, it was cocrystallized with three core fucosylated N-glycans (**1**-**3**) of different length. The structures of the complexes were solved at 1.7 or 2.1 Å for two crystal forms where the asymmetric unit contains either four rPhoSL trimers (N-glycan **1**) or two protomers with the trimers formed by crystallographic symmetry (N-glycans **2** and **3**, Table S1). Those trimers present only small rigid body shifts for some surface loops compared to the apo structure (rmsd of 0.35-0.6 Å, Figure S3c). The three additional *N*-terminal amino acids remaining after TEV cleavage (GAM) present various orientations with no impact on the general fold or observed ligand interactions.

The three core fucosylated N-glycans could be located unambiguously at each interface between two protomers indicating that rPhoSL is trivalent (Figure 3). The core fucose and the N-acetyl group of GlcNAc2 are found in a deep groove formed essentially by surface loops (Figure 3a-b). One wall of the groove is composed of amino acids essentially involved in hydrogen bonding interactions: A1*, G12, D13 and Y15 whilst hydrophobic residues V3*, L21* and V36 form its bottom and F23* and W28* its other wall (* indicates the neighboring protomer, Figure 3b). The fucose methyl group is completely buried and the fucose ring makes CH-π interactions with the F23* phenyl group. The

conformation of the α1,6-glycosidic linkage is imposed by hydrophobic interactions with F23* and W28*. As a consequence, only an α1,6-linked fucose, which is attached via three flexible bonds, can be accommodated in this groove whereas secondary fucosides (e.g. α1,3) would result in steric clashes with the protein.

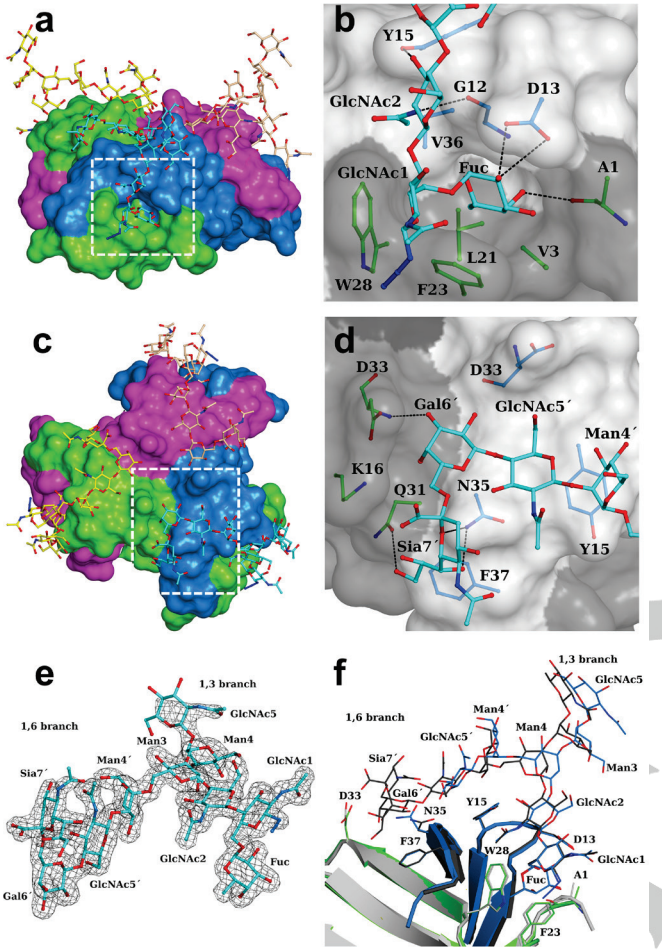


Figure 3. Binding of core fucosylated sialylated N-glycan **3** to rPhoSL. (a) Surface representation of rPhoSL and bound ligand (ball and stick) viewed from the side. (b) Interactions of rPhoSL with the core fucosylated chitobiose. (c) View from the top. (d) Interactions with the 1,6-branch of the N-glycan. (e) 2Fobs–Fcalc electron density for N-glycan **3** displayed at 1 σ (0.27 eA^3) (f) Overlay of the binding interface of N-glycan **1** (blue) and **3** (grey). Amino acids involved in sugar recognition are depicted and hydrogen bonds are represented by dotted lines.

In all complexed structures, the rigid α1,3-linked branch of the N-glycan does not interact with the protein and is totally exposed to the solvent (Figure 3a). The subsequent disorder resulted in no or poor electron density which did not permit its accurate building in many protein chains (Figures S4-5). The α1,3-arm was resolved only in the presence of stabilizing crystals contacts as for protein chain A in complex with N glycan **2** and **3** or protein chain D with N-glycan **1**. On the other hand, the flexible α1,6-branch is well ordered and completely resolved in the structures with N-glycans **2** or **3** and in eight of the twelve molecules of the complex with N-glycan **1** (Figures 3e and S5-6). The α1,6-arm wraps over the phenyl ring of Y15 at the top of β-prism. In N-glycan **1** the GlcNAc5' ring makes CH-π interaction with this aromatic ring and direct H-bonds with N35 side chain or water mediated contacts with Y15 hydroxyl and V36 main chain in N-glycan **1** (Figure S7). When the N-glycan is galactosylated (**2** or **3**) the α1,6-arm extends beyond β-strand D allowing the Gal6' moiety to bind to another groove (Figure 3c-e). This cavity is formed on one side by the loop connecting strands C and D and strand D of one protomer whilst the bottom and other side consists mainly of the *C*-terminal part of strand C of the other protomer and the side chain of K16*. The main chain nitrogen of D33* is hydrogen bonded to OH-3 of Gal6' whilst its side chain stacks against the OH-4 oriented toward the bottom of the groove (Figure 3d). The 2,6-linked sialic acid (Sia7') is turned towards the solvent and interacts with the protein mainly via its glycerol moiety via H-bonds with the side chain of Q31* and N35 and a hydrophobic interaction with the phenyl ring of F37. The ring oxygen of Sia7' is also hydrogen bonded to the side chain of N35 (Figure 3d, e).

The dihedral angles measured for all glycosidic linkages correspond to low energy conformations (Table S2).[16] The crystal packing in both crystal forms creates a slightly different environment for each binding site and thus gives an excellent set of snapshots of the flexibility of N-glycans in the bound state (Figures 3f and S8). Following the constraints imposed by the fucose binding groove, the mobility of the core fucose and N-acetyl group of GlcNAc2 is rather limited. Small changes in the dihedral angles of the α1,6-linkage generate small shifts in the chitobiose in particular for GlcNAc2. In order to accommodate those, a small rigid body shift occurs mainly for the protein loop connecting strands B and C allowing a good orientation of the aromatic ring of F23* and W28* and optimization of the hydrophobic interactions (Figure S8c). High mobility is observed for the 1,3-arm as it is totally exposed to the solvent (Figure S8a). The motion of the 1,6-arm becomes restricted whenever it contains a galactose moiety binding to the protein (Figures 3f and S8a-b). Overall, the rPhoSL is fairly static and sugar binding essentially influences the conformation of the BC loop and N35 side chain. The latter can adopt two conformations for the short glycan **1** but only one for the longer glycans **2** and **3** in order to avoid steric clashes with the galactose and sialic acid moieties (Figure 3f).

A homology search with the rPhoSL sequence gave homologs composed of one (RSL)[9d] or three rPhoSL repeats (SL2-1).[9e] Thus, the β-prism III fold may also be adopted from tandem repeats. Sequence alignment revealed the strict conservation of the two cysteines forming a disulfide bridge and most residues involved in core fucose binding as G12, L21, F23, W28 (Figure S9). A consensus sequence for PhoSL-like proteins can thus be proposed: $_{10}$CDGDXYXCXAXLXFGDGXW$_{28}$ (X being any amino acid). It defines this new lectin family and can be used to predict new members and core fucose binding. Variations observed for D13 and Y15 should decrease the affinity for ligands due to the loss of H-bonds whilst mutation of V36 to I36 should have little impact on ligand binding as fucose may still be accommodated. The sequence conservation is much lower for the secondary binding groove and may affect binding of the 1,6-arm (Figure S9). The nine residues at the *C*-terminus of the repeat seem responsible for the fine specificity of PhoSL-like proteins.

The mini lectin rPhoSL is able to specifically recognize core fucosylated N-glycans. rPhoSL self-assembles from 3 monomers into novel protein fold, the β-prism III, which is stabilized by stand exchange. One N-glycan (here up to 12 sugars, 2.4 kD) is bound per PhoSL protomer (40 aa, 4.5 kD). The binding sites located at the interfaces between monomers display a specific architecture for the exclusive recognition of α1,6 core fucose and provide simultaneous binding of the 1,6-arm even for sialylated N-glycans. PhoSL is the first structurally characterized representative of a novel lectin family.[17] Similarly to the *Psathyrella velutina* lectin (PVL), which is used for the detection of terminal GlcNAc in truncated glycans associated with cancers,[18] rPhoSL could become a precision probe for the detection of core fucosylation on glycoproteins.

Acknowledgements

The authors thank Emilie Gillon for help during ITC measurements. The authors are grateful to the EMBL-Grenoble staff for assistance in using the High Throughput Crystallisation (HTX) facility, to the European Synchrotron Radiation Facility (ESRF), Grenoble, France and SOLEIL synchrotron, Saint Aubin, France for access and technical support to beamlines BM14, ID23-1 and Proxima 1. The Consortium for Functional Glycomics (grant number GM62116) provided the glycan array resource. This work was supported by the Horizon 2020 program of the European Union (iNEXT grant, project no. 653706), the French National Research Agency (Glyco@Alps grant no. ANR-15-IDEX-02) and Deutsche Forschungsgemeinschaft (DFG: Un 63/4-2). A.U. is a recipient of a research grant from the Ministry of Research. C.U. thanks Roche for providing the sugar nucleotides.

Keywords: recombinant lectin • core fucosylation • β-prism fold • structural biology • glycosylation

[1] M. Schneider, E. Al-Shareffi, R. S. Haltiwanger, Glycobiology 2017, 27, 601-618.

[2] S. S. Pinho, C. A. Reis, Nat. Rev. Cancer 2015, 15, 540-555.

[3] P. Agrawal, B. Fontanals-Cirera, E. Sokolova, S. Jacob, C. A. Vaiana, D. Argibay, V. Davalos, M. McDermott, S. Nayak, F. Darvishian, M. Castillo, B. Ueberheide, I. Osman, D. Fenyö, L. K. Mahal, E. Hernando, Cancer Cell 2017, 31, 804-819.e807.

[4] E. Llop, M. Ferrer-Batalle, S. Barrabes, P. E. Guerrero, M. Ramirez, R. Saldova, P. M. Rudd, R. N. Aleixandre, J. Comet, R. de Llorens, R. Peracaula, Theranostics 2016, 6, 1190-1204.

[5] N. Okuyama, Y. Ide, M. Nakano, T. Nakagawa, K. Yamanaka, K. Moriwaki, K. Murata, H. Ohigashi, S. Yokoyama, H. Eguchi, O. Ishikawa, T. Ito, M. Kato, A. Kasahara, S. Kawano, J. Gu, N. Taniguchi, E. Miyoshi, Int. J. Cancer 2006, 118, 2803-2808.

[6] N. Yamane-Ohnuki, M. Satoh, MAbs 2009, 1, 230-236.

[7] P. Zhang, S. Woen, T. Wang, B. Liau, S. Zhao, C. Chen, Y. Yang, Z. Song, M. R. Wormald, C. Yu, P. M. Rudd, Drug Discovery Today 2016, 21, 740-765.

[8] a) O. H. Hashim, J. J. Jayapalan, C. S. Lee, PeerJ 2017, 5, e3784; b) J. C. Manning, A. Romero, F. A. Habermann, G. García Caballero, H. Kaltner, H.-J. Gabius, Histochemistry and Cell Biology 2017, 147, 199-222.

[9] a) A. S. do Nascimento, S. Serna, A. Beloqui, A. Arda, A. H. Sampaio, J. Walcher, D. Ott, C. Unverzagt, N. C. Reichardt, J. Jimenez-Barbero, K. S. Nascimento, A. Imberty, B. S. Cavada, A. Varrot, Glycobiology 2015, 25, 607-616; b) S. R. Inamdar, S. M. Eligar, S. Ballal, S. Belur, R. D. Kalraiya, B. M. Swamy, Glycoconj J. 2016, 33, 19-28; c) Y. Kobayashi, H. Tateno, H. Dohra, K. Moriwaki, E. Miyoshi, J. Hirabayashi, H. Kawagishi, J. Biol. Chem. 2012, 287, 33973-33982; d) Y. Oda, T. Senaha, Y. Matsuno, K. Nakajima, R. Naka, M. Kinoshita, E. Honda, I. Furuta, K. Kakehi, J. Biol. Chem. 2003, 278, 32439-32447; e) S. Vainauskas, R. M. Duke, J. McFarland, C. McClung, C. Ruse, C. H. Taron, Sci. Rep. 2016, 6, 34195.

[10] K. Kusama, Y. Okamoto, K. Saito, T. Kasahara, T. Murata, Y. Ueno, Y. Kobayashi, Y. Kamada, E. Miyoshi, Glycoconjugate J. 2017, 34, 537-544.

[11] M. Berkmen, Protein Expr. Purif. 2012, 82, 240-251.

[12] R. B. Kapust, J. Tözsér, J. D. Fox, D. E. Anderson, S. Cherry, T. D. Copeland, D. S. Waugh, Protein Engineering, Design and Selection 2001, 14, 993-1000.

[13] a) N. K. Fox, S. E. Brenner, J. M. Chandonia, Nucleic Acids Res. 2014, 42, D304-309; b) E. Krissinel, K. Henrick, Acta Crystallogr. D Biol. Crystallogr. 2004, 60, 2256-2268.

[14] L. Holm, P. Rosenstrom, Nucleic Acids Res. 2010, 38, W545-549.

[15] G. Hester, H. Kaku, I. J. Goldstein, C. S. Wright, Nat. Struct. Mol. Biol. 1995, 2, 472-479.

[16] S. Pérez, A. Sarkar, A. Rivet, C. Breton, A. Imberty, in Glycoinformatics (Eds.: T. Lütteke, M. Frank), Springer New York, New York, NY, 2015, pp. 241-258.

[17] After submission of our manuscript, the RMN structure of PhoSL from a synthetic peptide was published in K. Yamasaki, T. Yamasaki, H. Tateno, *Sci Rep* **2018**, *8*, 7740.

[18] A. Audfray, M. Beldjoudi, A. Breiman, A. Hurbin, I. Boos, C. Unverzagt, M. Bouras, S. Lantuejoul, J. L. Coll, A. Varrot, J. Le Pendu, B. Busser, A. Imberty, PLoS One 2015, 10, e0128190.

Experimental Proceduces

Core fucosylated N-glycans synthesis. Compound **1** (scheme 1) was synthetized as previously described.[1] Compound **2** was synthetized as follow. Octasaccharide **1** (10.0 mg, 6.72 µmol) and UDP-Gal (17.1 mg, 30.20 µmol) were dissolved in 2.2 mL of sodium cacodylate (50 mM, pH 7.4) containing 2.2 mg of bovine serum albumin, 2.24 µmol of $MnCl_2$, 2.24 µmol of NaN_3 and 22.4 units of calf intestinal alkaline phosphatase (EC 3.1.3.1). GlcNAc β-1,4-galactosyltransferase (EC 2.4.1.22, 240 milliunits) was added and the reaction mixture was incubated for 48 h at 30°C. The pH was checked regularly and was maintained at 7.4 by addition of 1 M NaOH. After 48 h, the reaction mixture was purified by gel filtration (Superdex® 30, 16 x 600 mm, 0.1 M NH_4HCO_3, 1 mL min^{-1}, retention time 83 min) and lyophilized. Purification by HPLC (YMC® Hydro C18, 150 x 10 mm, 0-15 % acetonitrile, 2 mL min^{-1} 10 CV, retention time 40 min) furnished the galactosylated decasaccharide **2** (7.0 mg, 3.9 µmol, 58.0 %). R_f = 0.18 (*iso*-propanol/1 M ammonium acetate 2:1); $C_{68}H_{113}N_7O_{49}$ (1811.66); HR-MS (H_2O): $M_{calc.}$ = 906.8356 $(M+2H)^{2+}$, M_{found} = 906.8284 $(M+2H)^{2+}$.

^{1}H-NMR (500 MHz, D_2O): δ = 5.12 (s, 1H, H-1^4), 4.93 (s, 1H, H-$1^{4'}$), 4.91 (d, $J_{1,2}$ = 3.7 Hz, 1H, H-1^F), 4.77 (s, 1H, H-1^2), 4.75 (s, 1H, H-1^1), 4.69 (d, $J_{1,2}$ = 8.0 Hz, 1H, H-1^3), 4.58 (d, $J_{1,2}$ = 8.0 Hz, 2H, H-1^5, H-$1^{5'}$), 4.47 (d, $J_{1,2}$ = 8.0 Hz, 2H, H-1^6, H-$1^{6'}$), 4.25 (s, 1H, H-2^3), 4.19 (s, 1H, H-2^4), 4.14 (d, $J_{5,6}$ = 6.8 Hz, 1H, H-5^F), 4.11 (s, 1H, H-$2^{4'}$), 3.84-4.00 (m, 13H, H-$6a^5$, H-$6a^{5'}$, H-$6a^1$, H-4^6, H-$4^{6'}$, H-3^4, H-$3^{4'}$, H-3^F, H-$6a^6$, H-$6a^{6'}$, H-$6a^2$, H-$6a^4$, H-$6a^{4'}$), 3.45-3.84 (m, 41H, H-$6b^5$, H-$6b^{5'}$, H-4^1, H-2^F, H-4^F, H-2^1, H-2^2, H-2^5, H-$2^{5'}$, H-3^2, H-5^4, H-$6b^1$, H-3^5, H-$3^{5'}$, H-5^1, H-3^6, H-$3^{6'}$, H-5^3, H-$6b^6$, H-$6b^{6'}$, H-$5^{4'}$, H-5^5, H-$5^{5'}$, H-4^4, H-$4^{4'}$, H-2^6, H-$2^{6'}$, H-5^6, H-$5^{6'}$, H-4^2, H-5^2, H-$6b^2$, H-4^3, H-$6a^3$, H-$6b^{3'}$, H-$6b^4$, H-$6b^{4'}$, H-3^1, H-3^3, H-4^5, H-$4^{5'}$), 2.01-2.11 (m, 12H, Ac), 1.24 (d, $J_{5,6}$ = 7.0 Hz, 3H, CH_3^F).

^{13}C-NMR (125 MHz, D_2O): δ = 104.0 (C-1^6β, C-$1^{6'}$β, J = 166.3 Hz), 102.0 (C-1^3β, J = 166.3 Hz), 101.4 (C-1^2β, J = 161.4 Hz), 100.5 (C-1^4α, J = 171.1 Hz), 100.4 (C-1^5β, C-$1^{5'}$β, J = 160.6 Hz), 98.0 (C-1^Fα, J = 170.3 Hz), 97.8 (C-$1^{4'}$α, J = 171.2 Hz), 89.6 (C-1^1β, J = 162.8 Hz), 81.1 (C-4^2), 80.0 (C-3^3), 79.3 (C-4^5, C-$4^{5'}$), 79.3 (C-4^1), 77.2 (C-2^4), 77.1 (C-$2^{4'}$), 76.3 (C-5^1), 75.6 (C-5^3), 75.4 (C-5^5, C-$5^{5'}$), 75.3 (C-5^2), 75.2 (C-3^6, C-$3^{6'}$), 75.0 (C-5^4), 74.7 (C-5^6, C-$5^{6'}$), 74.5 (C-2^6, C-$2^{6'}$), 74.3 (C-$5^{4'}$), 73.8 (C-3^5, C-$3^{5'}$), 73.6 (C-3^1), 73.7 (C-5^3), 73.3 (C-4^F), 72.7 (C-4^3), 71.1 (C-2^6, C-$2^{6'}$), 71.0 (C-3^2), 70.9 (C-2^3), 70.8 (C-3^4, C-$3^{4'}$), 70.7 (C-3^F), 69.5 (C-4^6, C-$4^{6'}$), 69.1 (C-2^F), 69.0 (C-4^4, C-$4^{4'}$), 68.0 (C-5^F), 67.0 (C-6^1), 66.2 (C-6^3),

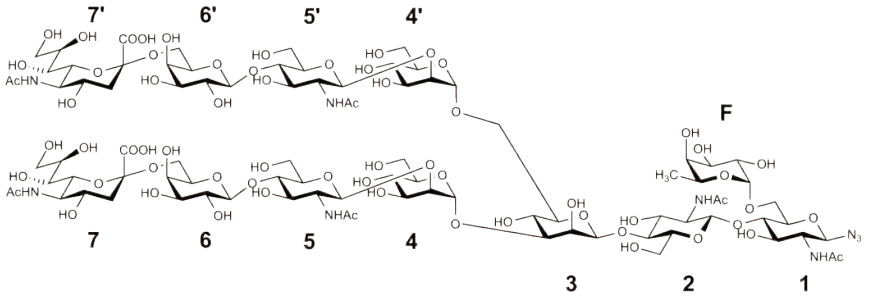
64.4 (C-6^6, C-$6^{6'}$), 62.1 (C-6^4, C-$6^{4'}$), 60.9 (C-6^5, C-$6^{5'}$), 55.7 (C-2^5, C-$2^{5'}$), 55.6 (C-2^2), 56.1 (C-2^1), 22.1 (CH_3^F), 16.4, 16.2, 16.0, 15.9, 15.9, 15.8 (NAc).

The following procedure was used for the synthesis of compound **3;** Decasaccharide **2** (3.0 mg, 1.66 µmol) and CMP-Neu5Ac (4.7 mg, 7.45 µmol) were dissolved in 282 µL of TRIS buffer (100 mM, pH 9). Gal α-2,6-sialyltransferase (from *Photobacterium damselae*, EC 2.4.99.1, 132.8 µg) was added and the reaction mixture was incubated at 30 °C. After 1 h another 2.4 mg CMP-Neu5Ac (3.81 µmol) were added and the mixture was incubated at 30 °C. After another 1.5 h the reaction mixture was purified by gel filtration (Sephadex® G25, 25 x 155 mm, 5 % ethanol, 2 mL min^{-1}, retention time 175 min) and lyophilized. Purification by HPLC (YMC® Hydro C18, 150 x 10 mm, 0-40 % acetonitrile, 2 mL min^{-1}, 15 CV, retention time 60 min) furnished the dodecasaccharide (2.3 mg, 0.98 µmol, 58.9 %). R_f = 0.10 (*iso*-propanol/1 M ammonium acetate 2:1); $C_{90}H_{147}N_9O_{65}$ (2393.85); HR-MS (H_2O): $M_{calc.}$ = 1195.9164 $(M-2H)^{2-}$, M_{found} = 1195.9095 $(M-2H)^{2-}$.

^{1}H-NMR (500 MHz, D_2O): δ = 5.13 (s, 1H, H-1^4), 4.94 (s, 1H, H-$1^{4'}$), 4.91 (d, $J_{1,2}$ = 3.6 Hz, 1H, H-1^F), 4.78 (d, $J_{1,2}$ = 3.5 Hz, 1H, H-1^2), 4.75 (s, 1H, H-1^1), 4.69 (d, $J_{1,2}$ = 7.4 Hz, 1H, H-1^3), 4.61 (d, $J_{1,2}$ = 7.4 Hz, 2H, H-1^5, H-$1^{5'}$), 4.45 (d, $J_{1,2}$ = 8.0 Hz, 2H, H-1^6, H-$1^{6'}$), 4.26 (s, 1H, H-2^3), 4.20 (s, 1H, H-2^4), 4.14 (d, $J_{5,6}$ = 7.0 Hz, 1H, H-5^F), 4.12 (s, 1H, H-$2^{4'}$), 3.88-4.04 (m, 15H, H-6a^5, H-6a$^{5'}$, H-6a^1, H-4^6, H-$4^{6'}$, H-3^4, H-$3^{4'}$, H-3^F, H-6a^6, H-6a$^{6'}$, H-6a^2, H-6a^4, H-6a$^{4'}$, H-$7^{7'}$, H-7^7), 3.47-3.88 (m, 53H, H-9a^7, H-9a$^{7'}$, H-6b^5, H-6b$^{5'}$, H-5^7, H-$5^{7'}$, H-4^1, H-2^F, H-4^F, H-2^1, H-2^2, H-2^5, H-$2^{5'}$, H-3^2, H-5^4, H-6b^1, H-9b^7, H-9b$^{7'}$, H-3^5, H-$3^{5'}$, H-5^1, H-3^6, H-$3^{6'}$, H-5^3, H-6b^6, H-6b$^{6'}$, H-6^7, H-$6^{7'}$, H-$5^{4'}$, H-5^5, H-$5^{5'}$, H-4^4, H-$4^{4'}$, H-4^7, H-$4^{7'}$, H-2^6, H-$2^{6'}$, H-5^6, H-$5^{6'}$, H-4^2, H-5^2, H-6b^2, H-4^3, H-6a^3, H-6b$^{3'}$, H-6b^4, H-6b$^{4'}$, H-3^1, H-3^3, H-4^5, H-$4^{5'}$, H-8^7, H-$8^{7'}$), 2.00-2.12 (m, 12H, Ac), 2.67 (d, $J_{7,8}$ = 11.5Hz, 2H, H-3a^7, H-3a$^{7'}$), 1.73 (t, $J_{3,3}$ = 12.3 Hz, 2H, H-3b^7, H-3b$^{7'}$), 1.24 (d, $J_{5,6}$ = 6.3 Hz, 3H, CH_3^F).

^{13}C-NMR (125 MHz, D_2O): δ = 104.6 (C-1^6β, C-$1^{6'}$β, J = 166.4 Hz), 102.0 (C-1^3β, J = 164.0 Hz), 101.4 (C-1^2β, J = 160.2 HZ), 100.3 (C-1^4α, J = 171.1 Hz), 100.0 (C-1^5β, C-$1^{5'}$β, J = 161.7 Hz), 99.8 (C-1^Fα, J = 171.0 Hz), 97.9 (C-$1^{4'}$α, J = 171.0 Hz), 89.6 (C-1^1β, J = 162.5 Hz), 81.6 (C-4^5, C-$4^{5'}$), 80.4 (C-4^2), 80.0 (C-3^3), 79.1 (C-4^1), 77.4 (C-2^4), 77.3 (C-$2^{4'}$), 76.4 (C-5^1), 75.6 (C-5^5, C-$5^{5'}$), 75.4 (C-5^2), 75.3 (C-3^6, C-$3^{6'}$), 74.8 (C-5^4), 74.7 (C-2^6, C-$2^{6'}$), 74.6 (C-5^6, C-$5^{6'}$), 73.9 (C-$5^{4'}$), 73.6 (C-6^7, C-$6^{7'}$), 73.4 (C-3^5, C-$3^{5'}$, C-5^3), 73.4 (C-3^1), 73.3 (C-3^2), 73.1 (C-4^F), 72.8 (C-7^7, C-$7^{7'}$), 72.7 (C-4^3), 71.8 (C-2^6, C-$2^{6'}$), 70.9 (C-2^3), 70.8 (C-3^4, C-$3^{4'}$), 70.7 (C-3^F), 69.6 (C-8^7, C-$8^{7'}$), 69.5 (C-4^7, C-$4^{7'}$), 69.4 (C-4^6, C-$4^{6'}$), 69.3 (C-2^F), 68.4 (C-4^4, C-$4^{4'}$), 68.3 (C-5^F), 67.4 (C-6^1), 66.3 (C-6^3), 64.4 (C-6^6, C-$6^{6'}$), 63.7 (C-9a^7,

C-9a$^{7'}$), 62.4 (C-6^{4}, C-6$^{4'}$), 61.3 (C-6^{5}, C-6$^{5'}$), 55.7 (C-2^{5}, C-2$^{5'}$), 55.6 (C-2^{2}), 55.5 (C-2^{1}), 52.9 (C-5^{7}, C-5S$^{7'}$), 41.1 (C-3^{7}, C-3$^{7'}$), 23.2, 23.1, 23.1, 23.0, 23.0, 22.9 (NAc), 22.2 (CH_3^F). The spectra were assigned according to the following convention:



Expression and purification of recombinant PhoSL. A nucleotide sequence, translated from the peptide sequence of PhoSL[2] was synthesized after codon optimization for expression in *Escherichia coli* (Genscript USA). The gene was introduced in the expression vector pET39b-TEV using *Nco*I and *Xho*I restriction sites. This homemade vector is derived from the pET39b plasmid (Novagen) where the enterokinase cleavage site was replaced for the cleavage site of the Tobacco Etch virus (TEV) protease. *E. coli* BL21 Star (DE3) cells harboring the plasmid pET39b-TEV-PhoSL were cultured in LB Broth medium with 30µg.mL^{-1} kanamycin at 37°C. When the culture reached an A_{600nm} of 0.4, the incubation temperature was changed to 16°C. When the culture reached an A_{600nm} of 0.8-0.9, protein expression was induced by the addition of 0.1 mM isopropyl β-D-thiogalactoside. Cells were harvested after one night at 16°C by centrifugation at 5000 x *g* for 10 min at 4°C, and the pellet from 1 liter culture was resuspended in 30 mL of buffer A (30 mM Tris-HCl, 20% sucrose, pH 8) and 1 mM EDTA for periplasmic extraction. The cells were gently stirred at room temperature for 10 min prior centrifugation at 10000 x *g* for 10 min at 4°C. The pellet obtained was resuspended in 30 mL of ice-cold 5 mM $MgSO_4$ and gently stirred for 10 min on ice. After centrifugation at 10000 x *g* for 10 min, the supernatant was subjected to affinity chromatography on 1 mL cOmplete^{TM} His-Tag column (Roche) equilibrated with 20 mM Tris-HCl pH 8, 300 mM NaCl. After washing with the equilibrium buffer, elution was performed with a 20 mL linear gradient of imidazole (0-500 mM). The fractions were analyzed on 15% SDS-PAGE gels prior pooling and desalting on PD10 column (Ge Healthcare) in 20 mM Tris-HCl pH 8. The His-tag was cleaved with the TEV protease[3] (1:50 w/w; enzyme:protein ratio) at 19°C overnight with addition of 0.5 mM EDTA. Then, the mixture was applied on the affinity column and the cleaved protein was collected in the flow-through. Protein contaminants were separated by anion-exchange chromatography on a 1 mL HiTrap DEAE Sepharose FF column (GE Healthcare). equilibrated with 20 mM Tris-HCl pH 8. Elution was performed

with a 30 mL linear gradient of NaCl (0-1M). After analysis of the fractions on 16.5% Tris-tricine gels, those containing pure protein were pooled and desalted on PD10 column (Ge Healthcare) in 25 mM MES pH 6, prior concentration by centrifugation using a Vivaspin (3KDa, Sartorius). Thereafter, recombinant PhoSL was used for physicochemical characterization and stored at 4°C.

Size exclusion chromatography. To study the oligomarisation state of rPhoSL, size exclusion chromatography (SEC) was performed on an ENrich™ SEC 70 10 x 300 column (Bio-Rad Ltd) at a flow rate of 1 ml min^{-1} in a buffer composed of 25 mM MES pH6. 100 µl of protein were injected and the eluate was monitored at 280 nm. A calibration curve was done with the following standards: Bovine serum albumine 66.5 kDa, Ovalbumin 43 kDa and RNAse A 13.7 kDa (GE Heathcare).

Glycan microarray analysis. Protein affinity towards a panel of 600 biologically important mammalian glycans was measured using the standard procedure of the protein–glycan interaction core (H) of the Consortium for Functional Glycomics (USA). Recombinant PhoSL was labeled with Alexa Fluor 488-TFP (carboxylic acid, 2,3,5,6 tetrafluorophenyl) (Fisher-Scientific) according to the manufacturer's instructions. The protein was analyzed at a concentration of 50 µM and 5 µM on Mammalian Printed Array Version 5.2. Experimental data were normalized to percentages of the highest RFU value for each analysis and an average binding was obtained for each glycan after averaging of the percentages at different lectin concentrations. The data were sorted according to their average binding and analyzed.

Enzyme linked lectin assay (ELLA) on glycoproteins. Microplates-96 wells (Greiner Bio-one #655061) were coated with 100 µl/well of glycoprotein. Ovalbumin, Conalbumin and Thyroglobulin from GE Healthcare or Fetuin and human IgG1-κ from Sigma-Aldrich were diluted in 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.76 mM KH_2PO_4, 0.1 mM $CaCl_2$ pH 7.4 (PBS-Ca) to reach 1000 ng/well and incubated 2 h at room temperature (RT). Unbound glycoproteins were removed by three washes with PBS-Ca. Then, wells were blocked with 200 µl of 2 % bovine serum albumin (BSA) in PBS-Ca for 1h at 37°C and then washed 5 times with PBS-T (PBS-Ca, 0.05% Tween 20). Biotinylated rPhoSL was prepared using Biotinamidohexanoyl-6-aminohexanoic acid N-hydroxysuccinimide ester (Sigma Aldrich #B3295) mixed with the protein resuspended in PBS-Ca, following manufacturer's instructions. For each assay, biotinylated rPhoSL was diluted with 0.1% BSA to reach 30 ng/well of protein and 100 µl were added before incubation for 2h at RT. The plate was washed as before with PBS-T. Biotinylated rPhoSL was detected using 100 µl/well of

streptavidin-peroxidase (from *Streptomyces avidinii*, Sigma-Aldrich cat. #S5512) diluted with 0.1% BSA to reach 10 ng/well and incubated 25 min at RT. After washing the plate 5 times with PBS-T, the color was developed using 200 µl per well of o-phenylenediamine dihydrochloride (SigmaFast OPD tablet, Sigma Aldrich #P9187) for a few minutes in the dark and the reaction was stopped by addition of 50 µl of 3 M HCl. The signal was read on a microplate reader (TECAN, Spark 10M) using a standard photometric method at 492 nm. All samples were analyzed in duplicate.

Isothermal titration calorimetry. rPhoSL in 25 mM MES pH 6 was centrifuged and its concentration was determined with NanoDrop 2000 spectrophotometer (Ozyme) at 280 nm using a theorical extinction coefficient of 14,000 $M^{-1}.cm^{-1}$. Titration was performed using an ITC200 microcalorimeter (Microcal Inc., Malvern Instruments) at 25°C with rPhoSL in the cell at 50 µM. 20 injections of 2 µl of carbohydrate ligand (10 mM N-glycan **1** or 5 mM core fucosylated chitobiose) in the same buffer every 3 min. Data were fitted with MicroCal Origin 7 software according to standard procedures. Fitted data yielded the stoichiometry (*n*), the association constant (K_a) and the enthalpy of binding (ΔH). Other thermodynamic parameters (the free energy ΔG and the entropy ΔS) were calculated from the equation $\Delta G = \Delta H - T\Delta S = -RT \ln K_a$, where T is the absolute temperature and R = 8.314 J/mol/K.

Crystallization and data collection. rPhoSL at 5.5- 6.3 mg mL^{-1} in 25 mM Mes pH 6 was incubated or not with 1 mM core fucosylated N-glycans (**1**, **2** or **3**) prior crystallization trials. The hanging-drop vapor-diffusion method was used with 2-µl drops containing a 50:50 (v/v) mix of protein and reservoir solution at 19 °C. First screening was performed using commercial kits (Clear Strategy Screen I and II and structure Screen 2; Molecular Dimensions Ltd) without success. The protein samples were then subjected to screening using a crystallization robot (HTXLab, EMBL) and 200 nl sitting drops testing 576 commercial conditions. Several hits were obtained and optimized manually in 2 µl hanging drops. Crystal with triangular shape were obtained for the apo rPhoSL using 18% 1,4 butanediol, 300 mM Zinc acetate and 0.1 M Imidazole-HCl pH 6-7 whilst diamond shaped crystals were obtained for rPhoSL-ligand in 3-3.4 M sodium malonate pH 5. Crystals were directly mounted in a cryoloop and flashed freezed in liquid nitrogen. The data for the apo structure were collected at the ESRF (European Synchrotron Radiation Facility), Grenoble, France on the beamline BM14 using a MARCCD detector. For the complexed structures, data were collected using a Pilatus 6M detector (Dectris Ltd) on beamline ID23-1 at ESRF for the N-glycan **1** and on beamline Proxima-1 at SOLEIL, Saint Aubin, France for N-glycans

2 and **3**. The data were processed using XDS.[4] All further computing was performed using the CCP4 suite.[5]

Structural determination and refinement. The apo structure of rPhoSL was solved by SAD method (zinc anomalous signal) using ShelXC/D.[6] A polyalanine model was first autobuilt in ShelXE and was improved and the sequence assigned in ArpWarp.[7] The molecular replacement technique was used to solve the structures of rPhoSL complexed with core fucosylated N-glycans using PHASER.[8] Four copies f of the apo trimer were searched for N-glycan **1** in $P2_1$ space group and 2 copies of the apo monomer were searched for N-glycans **2** and **3** in $I2_13$ space group. Structure refinement was performed in Refmac5.8[9] and manual model corrections and building in COOT.[10] For cross-validation analysis, 5% of the observations were set aside with the same set for N-glycans **2** and **3.** Hydrogen atoms were added in their riding positions and used for geometry and structure-factor calculations. TLS refinement in Refmac was used during the refinement of N-glycan **1** complexes structure.[11] The stereochemical quality of the refined models was validated on the wwPDB Validation server: http://wwpdb-validation.wwpdb.org and carbohydrates were checked in Privateer.[12] All figures were drawn with PyMOL Molecular Graphic System program (Version 2.0.4, Schrodinger, LLC).

Accession codes. Coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 6EKE, 6FX1, 6FX2 and 6FX3 for the structures of apo rPhoSL and in complex with N glycan **1**, **2** and **3** respectively.

Data availability. The full glycan array data are available from the corresponding author upon request.

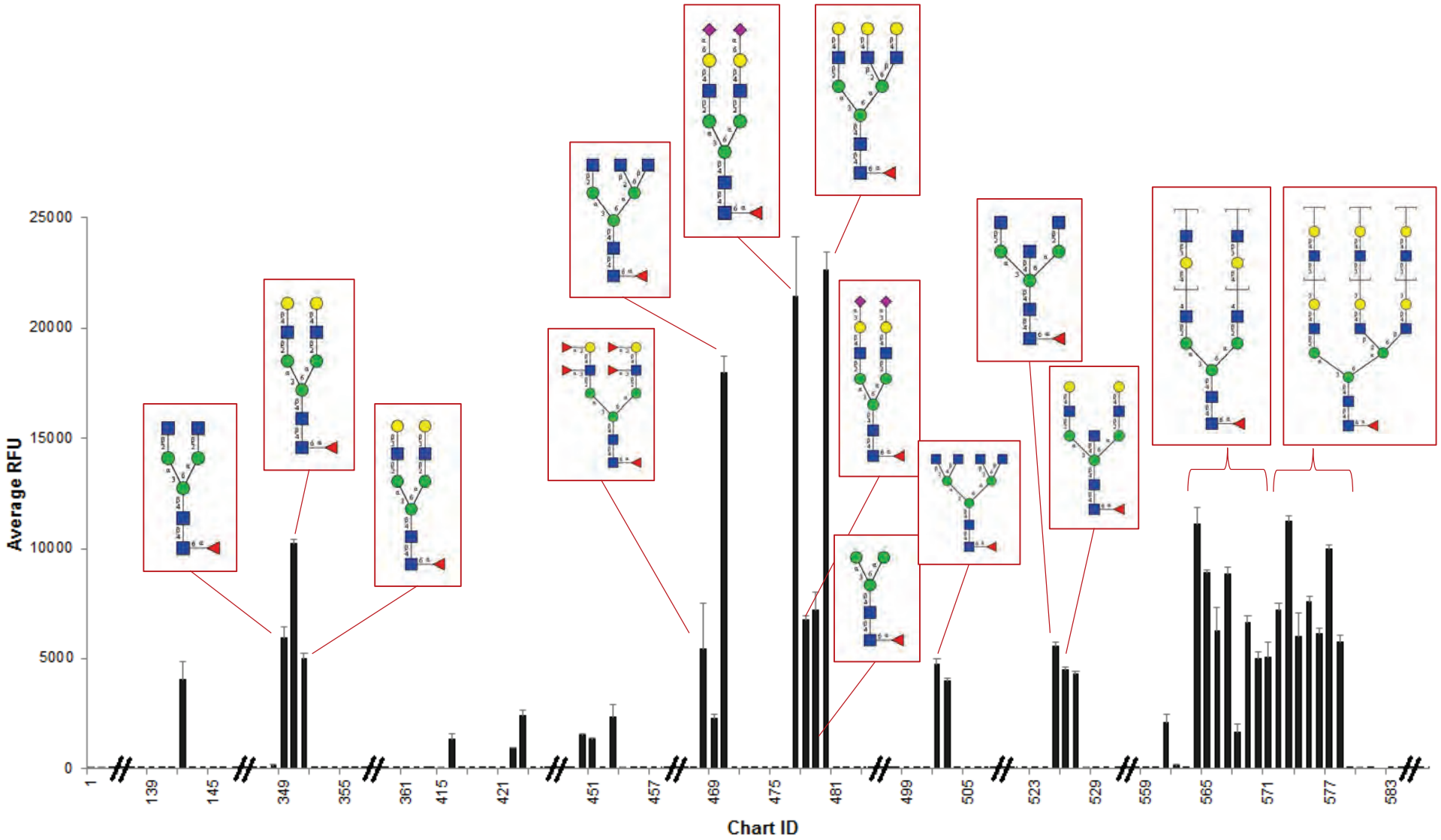


Figure S1 : Glycan arrays data for recombinant PhoSL. rPhoSL (5 µg mL^{-1}) labelled with Alexa488 on Mammalian Printed Array Ver 5.2. Only epitopes with signal superior to 400 RFU have been selected for the display and the highest affinity ligand are represented schematically. Complete data are given upon request to annabelle.varrot@cermav.cnrs.fr.

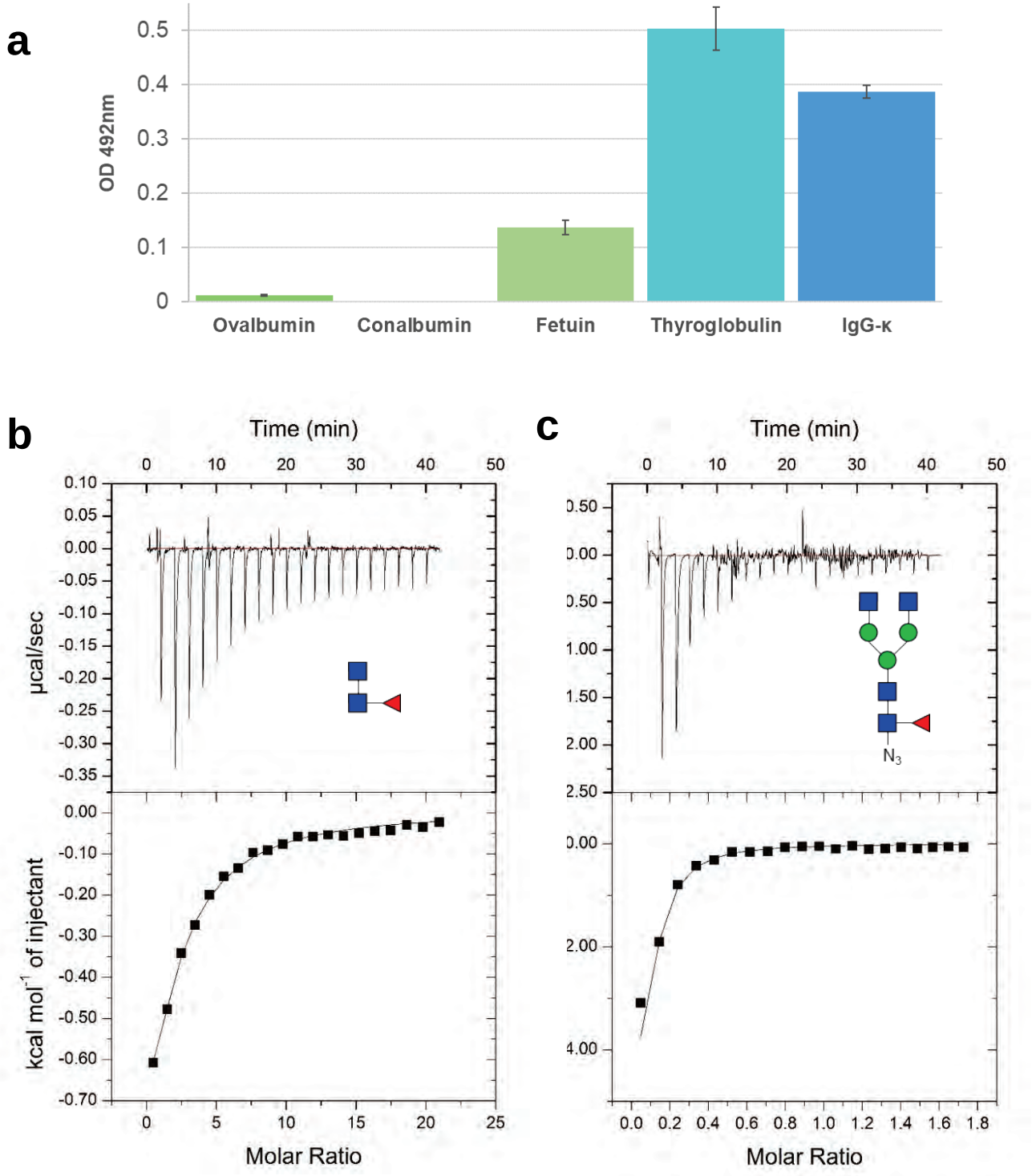


Figure S2. (**a**) Binding on biotinylated rPhoSL on immobilized glycoproteins by ELLA. (**b, c**) Isothermal microcalorimetry affinity measurements. Plots, measured by ITC200 (Microcal) were obtained from the titration of recombinant PhoSL (50 µM) with 5 mM core fucosylated chitobiose (**b**) and 10 mM core fucosylated N-glycan **1** (**c**) at 25 °C. Protein and glycans were prepared in 25 mM MES pH 6.0. The plots in the lower panels show the total heat released as a function of total ligand concentration for the titration shown in the upper panels. The solid lines represent the best least-square fit to experimental data using a one-site model. Schematic of the ligand are represented on top with red triangle for fucose, blue square for GlcNAc and green circle for mannose.

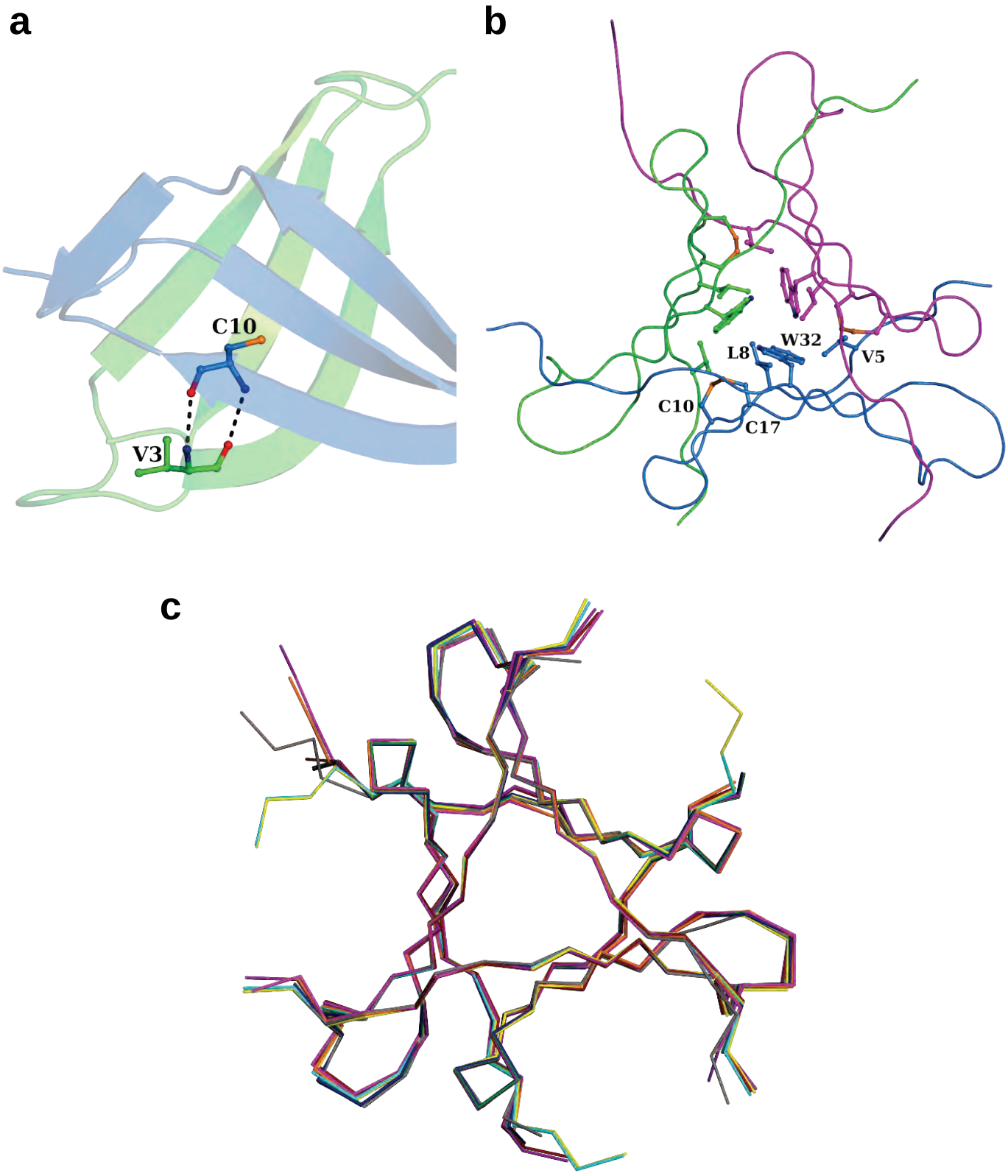


Figure S3. (**a**) Zoom on the interactions stabilizing strand A between two neighboring protein chains. (**b**) Residues filling the center of the β-prism. (**c**) Overlay of rPhoSL trimer in the apo structure (grey) with the trimers of the structures in complex with N-glycan **1** (red, purple, magenta, orange), with N-glycan **2** (blue, cyan) and with N-glycan **3** (green, yellow). Trimers for N-glycan **2** and **3** were created from crystallographic symmetry.

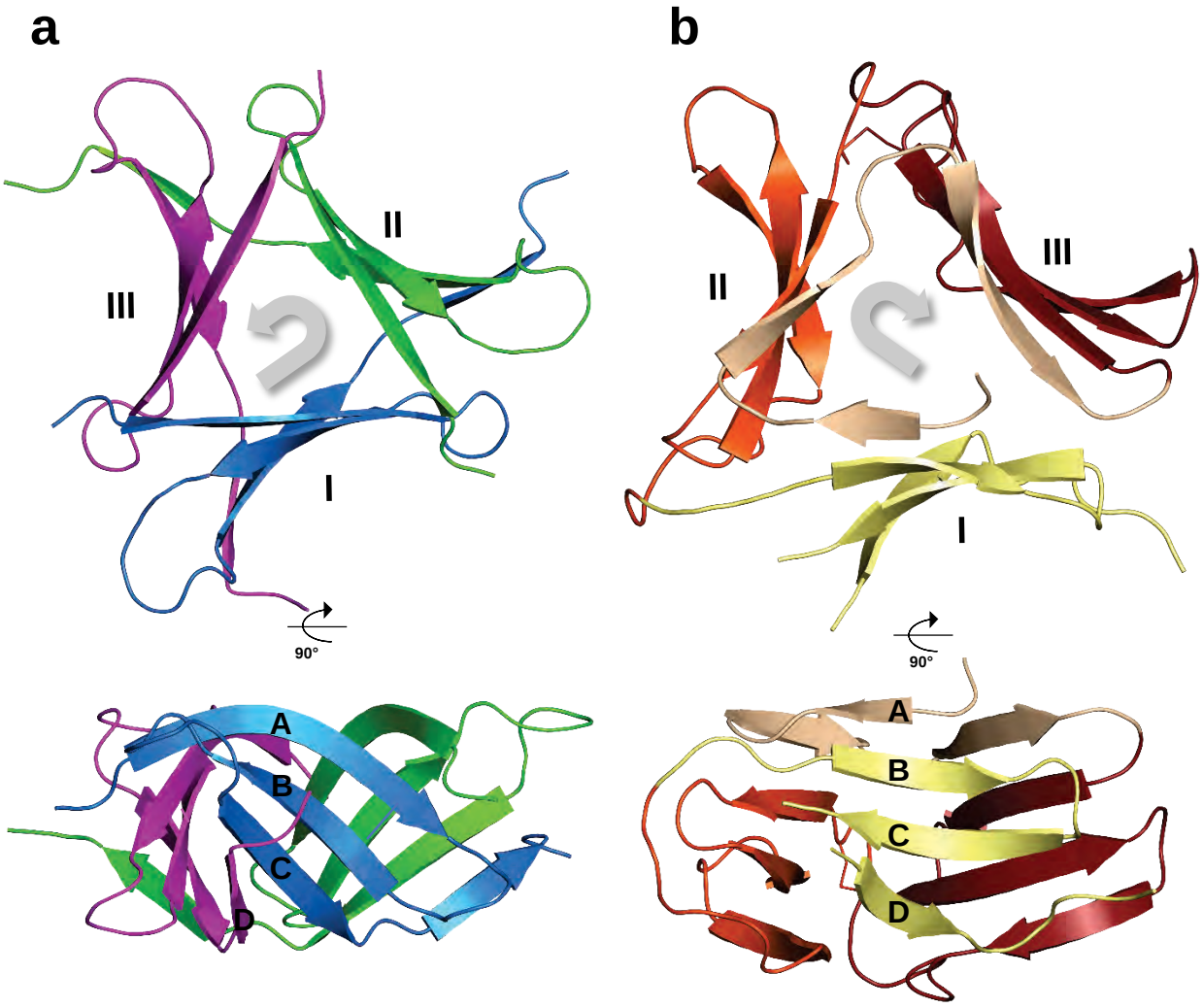


Figure S4. Differences in assembly, layout and directionality of β-strands and β-sheets between β-prisms II and III. Cartoon representation viewed from strands A faces and the side for rPhoSL (**a**) and GNA from *Galanthus nivalis* (**b**).[13] The strand composition from *N*- to *C*-terminus is A_I-B_I-C_I-D_{III}-A_{II}-B_{II}-C_{II}-D_{II}-A_{III}-B_{III}-C_{III}-D_{II} for PhoSL and A_I-A_{II}-A_{III}-B_{III}-C_{III}-D_{III}-B_{II}-C_{II}-D_{II}-B_I-C_I-D_I for GNA with I, II, II representing the three β-sheets. Disulfide bridges are represented by lines. rPhoSL is colored by protomer. β-strands A are colored in beige for GNA whilst the others B,C,D are colored per β-sheets.

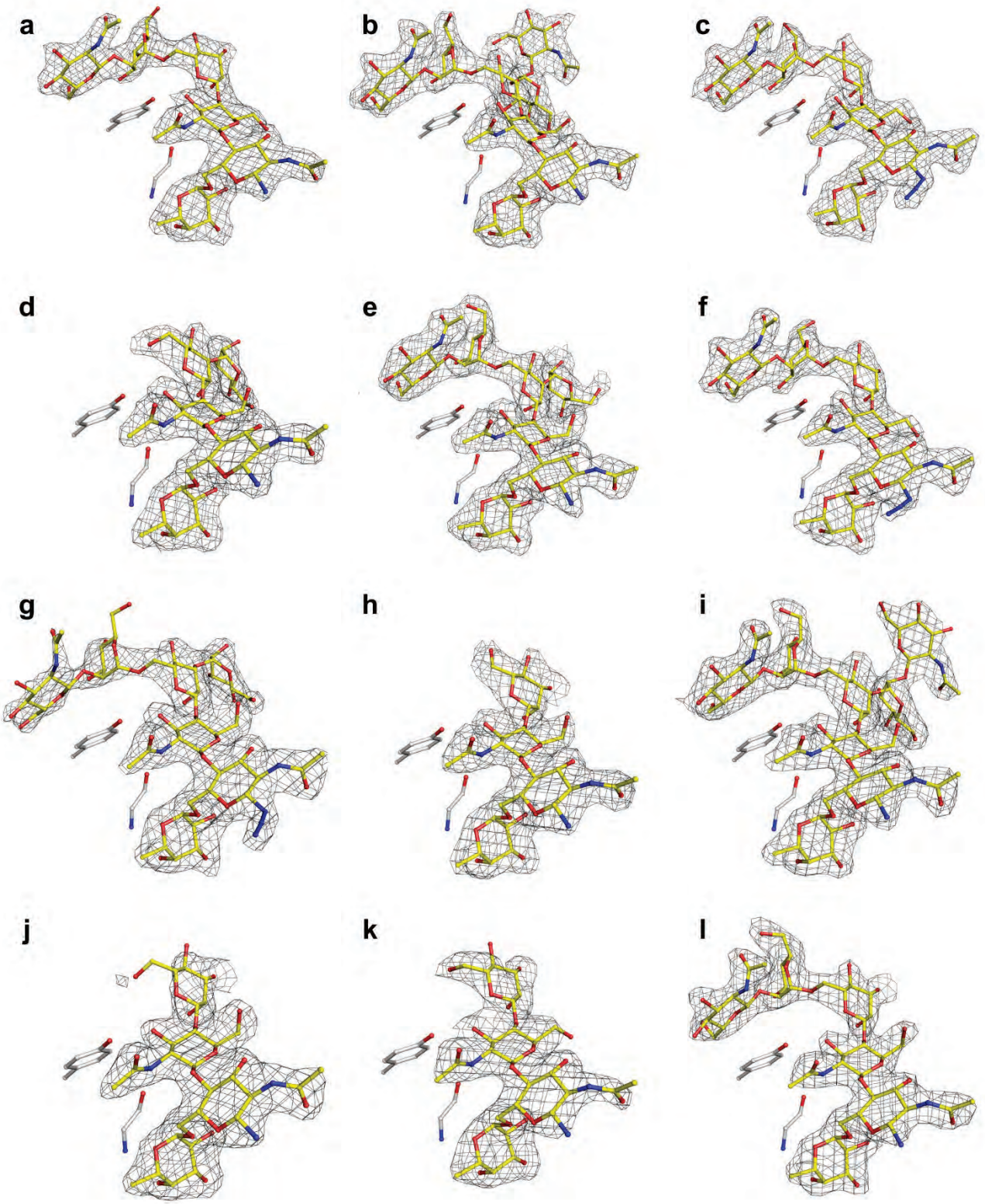


Figure S5. (**a-l**) $2F_{obs}–F_{calc}$ electron density for the modelled N-glycan **1** for each protein chain of rPhoSL displayed at 1 σ (0.2 eA^3). Tyr15 and Gly12 are represented in ball and sticks and each panel is name according to the protein chain it is related to in the coordinates of PDB 6FX1.

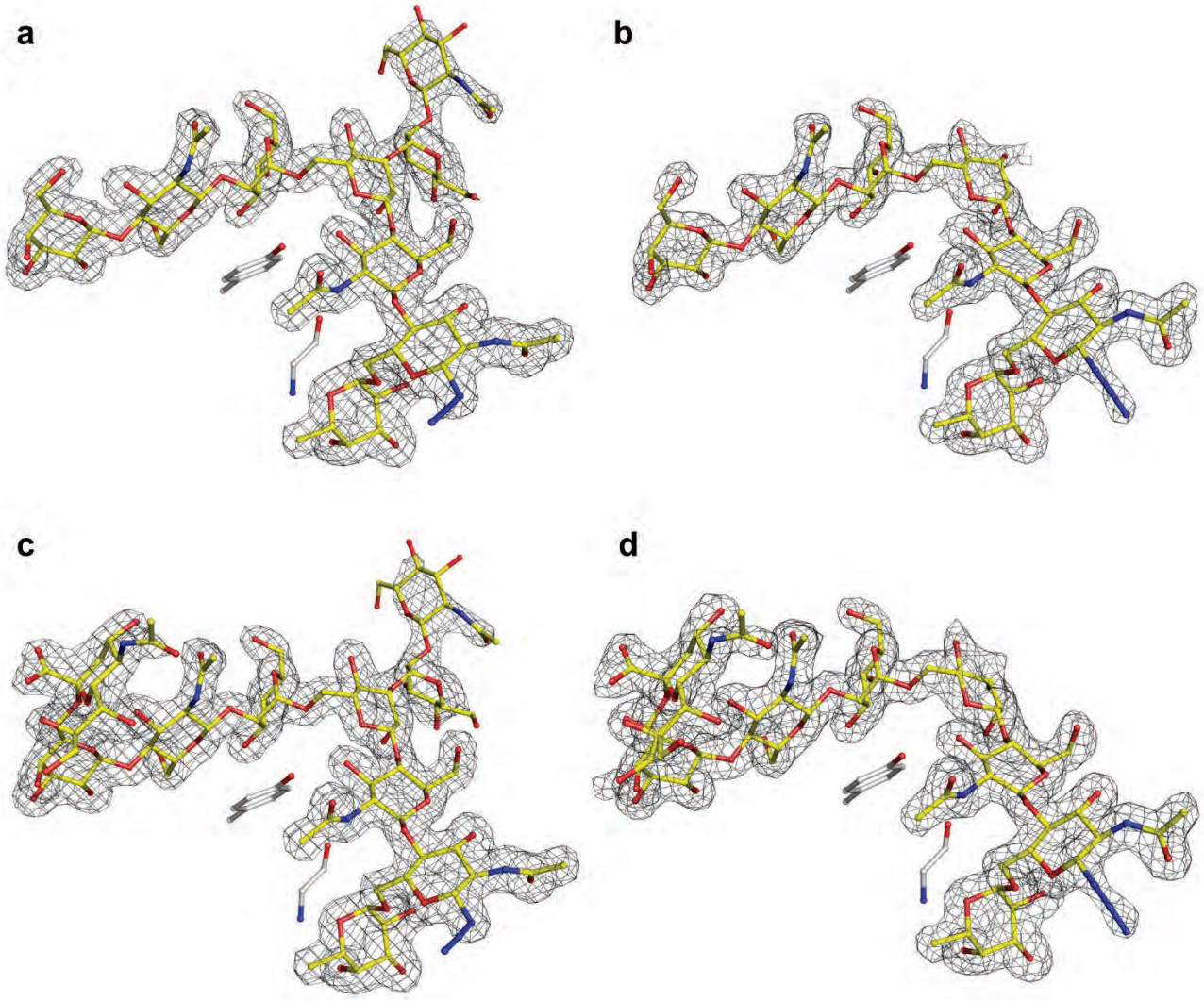


Figure S6. (**a-b**) $2F_{obs}–F_{calc}$ electron density for the modelled N-glycan **2** in protein chain A (a) and B (b) of rPhoSL displayed at 1 σ (0.25 eA^3). $2F_{obs}–F_{calc}$ electron density for the modelled N-glycan **3** in protein chain A (c) and B (d) of rPhoSL displayed at 1 σ (0.27 eA^3). Tyr15 and Gly12 are represented in ball and sticks.

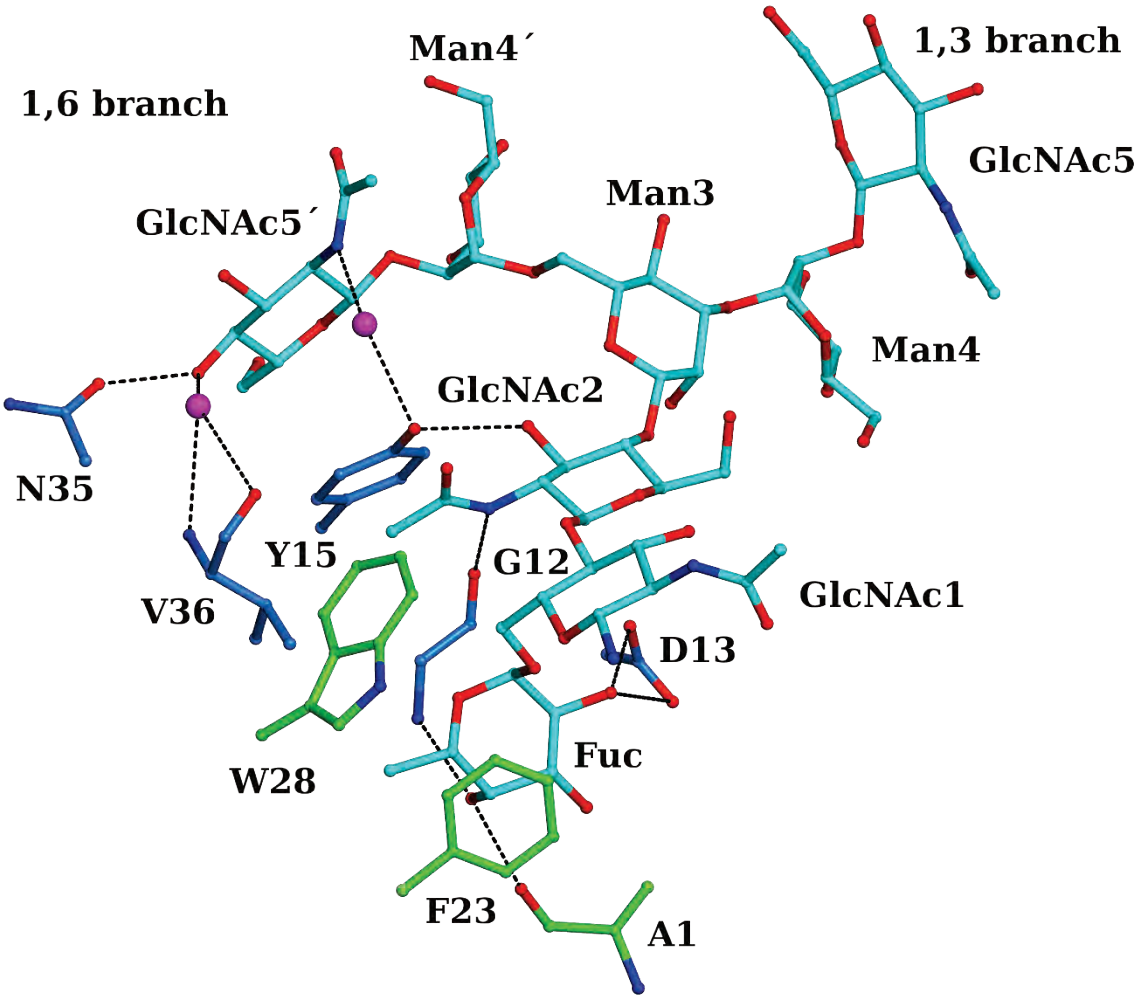


Figure S7. Zoom in the interactions of rPhoSL with N-glycan **1**

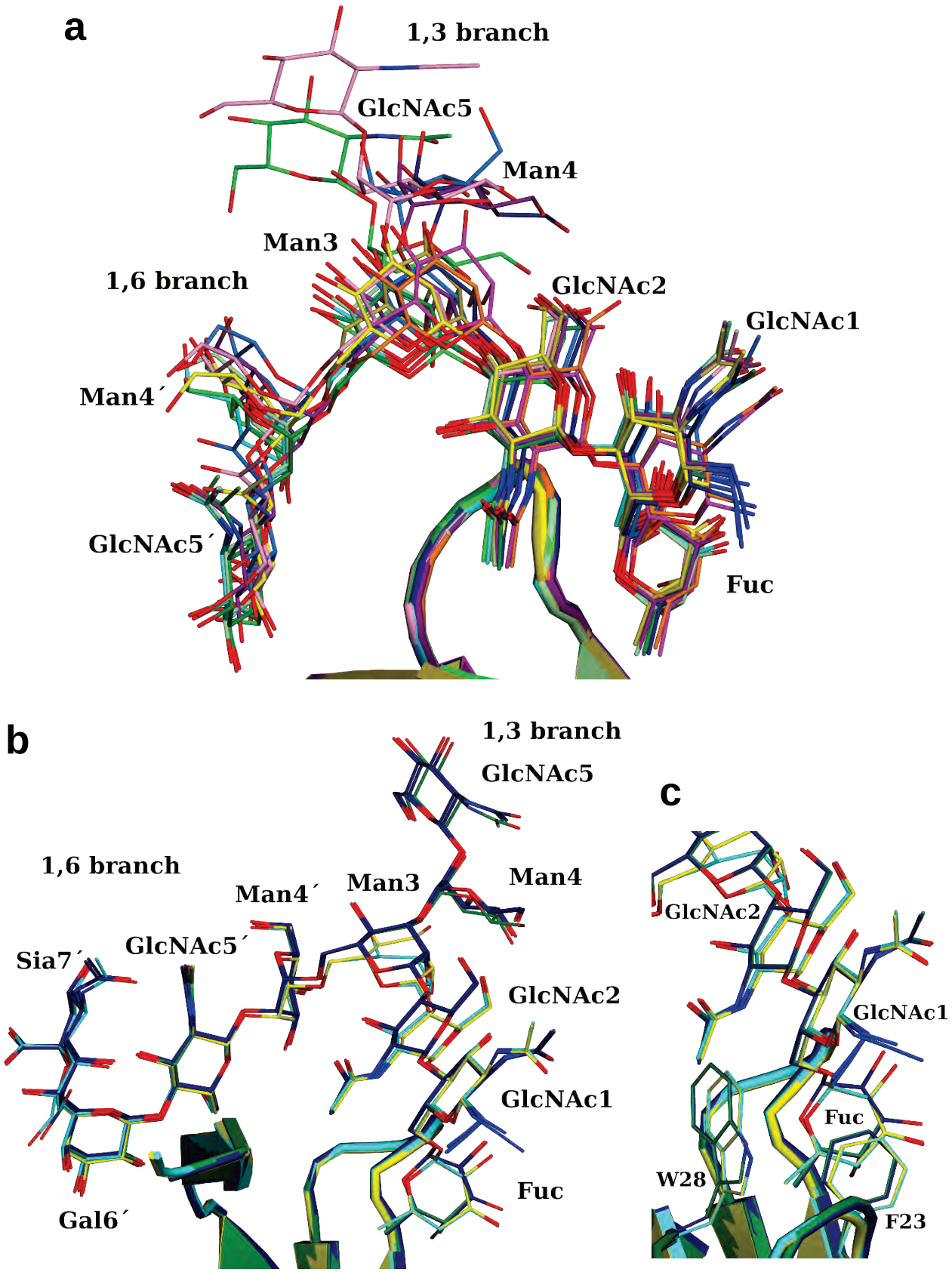


Figure S8. N-glycans flexibility. (a) Overlay of N-glycan **1** in its twelve binding sites colored by protein chain with A: dark green, B: green, C: pale green, D: dark blue, E: blue, F: cyan, G: purple, H: magenta, I: pink, J: brown, K: orange, L: yellow. (b-c) Overlay of the binding of N-glycan **2** (chain A: green, chain B: yellow) and N-glycan **3** (chain A, blue; chain B: cyan) to rPhoSL binding site with a zoom on the fucose environment (c).

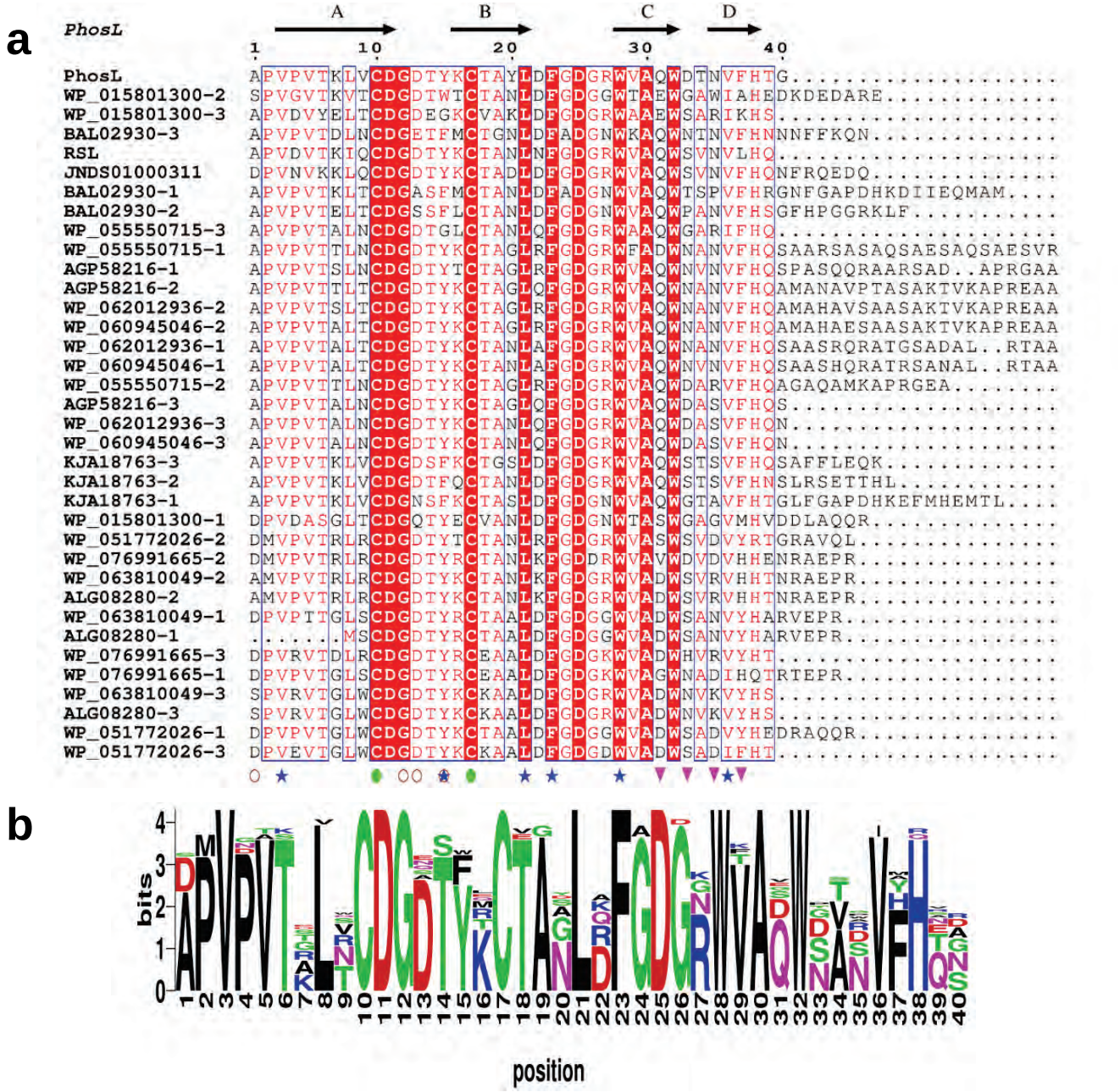


Figure S9. (a) Multiple sequence alignment of rPhoSL with homologous domain in hypothetical proteins. Aligned protein sequences are from *Pholiota squarrosa* (PhoSL); *Actinosynnema mirum* (WP_015801300); *Pholiota nameko* (BAL02930); *Rhizopus delemar* Type II NRRL 21477 (RSL, NODE_1316_length_1457_cov_6.197667_1), *Rhizopus stolonifer* B9770 (JNDS01000311, shotgun sequence ctg7180000282406_1); *Streptomyces sp.* NBRC 110028 (WP_055550715); *Streptomyces sporocinereus* (WP_062012936); *Streptomyces hygroscopicus* (WP_060945046) *Streptomyces rapamycinicus* NRRL 5491 (AGP58216); *Hypholoma sublateritium* FD-334 SS-4 (KJA18763); *Saccharothrix sp.* NRRL B-16314 (WP_051772026); *Actinosynnema sp.* ALI-1.44 (WP_076991665); *Kibdelosporangium phytohabitans* (WP_063810049 and ALG08280). Signal peptide has been removed for clarity. PhoSL numbering, secondary structure and the name of its □-strands are displayed on top. Cysteines involved in the conserved disulphide bridge are marked with a full green circle, those involved in hydrophobic interactions or hydrogen binding with the core fucosylated chitobiose are represented with a blue star and a brown circle, respectively and the residues interacting with the 1.6 branch with a triangle. The figure was drawn with ESPript 3.0 server.[14] (b) Sequence conservation using BlockLogo is depicted in the bottom panel.[15]

Table S1. Data Collection and Refinement Statistics for rPhoSL crystal structures.

	rPhoSL apo			rPhoSL-glycan 1	rPhoSL-N-glycan 2		rPhoSL-N-glycan 3	
Crystallisation conditions	18% 1,4-butanediol 300 mM ZnAc 100mM imidazole ph 7.0			3.2 M Na malonate pH 5.0	3.2 M Na malonate pH 5.0		3.4 M Na malonate pH 5.0	
Data collection								
Beamline ESRF	ESRF BM14			ESRF ID23-1	SOLEIL Proxima 1		SOLEIL Proxima 1	
Space group	$P2_1$			$P2_122_1$	$I2_13$		$I2_13$	
Cell dimensions *a b c (Å)* α β γ (°)	28.2 67.7 31.0 90 97.25 90			91.0 95.6 110.6 90 90 90	101.7 101.7 101.7 90 90 90		101.9 101.9 101.9 90 90 90	
Wavelength	0.984			0.976	0.979		0.979	
Resolution (Å)	33.84-1.70 (1.73-1.70)			47.88-2.10 (2.16-2.10)	36.01-1.7 (1.73-1.70)		41.54-1.7 (1.73-1.7)	
Measured reflections	50099			296329	628467		335255	
Unique reflections	12666			56993	19497		19409	
R_{merge}[a]	0.045 (0.349)			0.045 (0.466)	0.048 (0.550)		0.050 (0.663)	
Mean *I* / s*I*[a]	14.8 (3.0)			18.6 (3.2)	45.9 (7.3)		32.8 (4.6)	
Completeness (%)[a]	99.5 (97.6)			99.9 (100)	100.0 (100.0)		100.0 (100.0)	
Redundancy[a]	4.0 (3.6)			5.2 (5.4)	32.2 (29.7)		17.3 (16.9)	
$CC_{1/2}$[a]	99.7(89.2)			0.999 (0.907)	1.000 (0.968)		1.000 (0.917)	
Estimated mean FOM / Pseudo free CC (%)	0.639 / 68.2							
Refinement								
Resolution (Å)	33.84-1.70			47.88-2.10	36.01-1.70		41.54-1.70	
No. of work/ free Reflections	12732 / 665			57008 / 2785	19500 / 976		19412 /973	
R_{work} / R_{free}[b] (%)	17.1 / 21.1			18.4 / 21.9	16.4 / 18.8		15.7 / 18.7	
Rmsd bond lengths (Å)	0.0151			0.015	0.0153		0.0145	
Rmsd bond angles (°)	1.62			1.84	1.89		2.06	
No. of atoms	Chain A	Chain B	Chain C	For the 12 chains	Chain A	Chain B	Chain A	Chain B
Protein	277	338	339	3801	315	331	317	331
Ligand				889	113	88	133	112
Others heteroatoms	12	12	9	102				
Water	41	39	46	320	145		151	
Average *B*-factors	Chain A	Chain B	Chain C					
Protein	25	24.7	22.8	40.3	26.4	25.7	24.2	22.9
Ligand				50	40.6	33.7	38	31
Others heteroatoms	36.9	39.5	34.2	55.3				
Water	36.8	37	36.8	45.6	40.9		40.4	
Ramachandran plot (%) Allowed	100			100	100		100	
Favored	99.2			97.4	98.7		98.7	
Outliers	0			0	0		0	
PDB code	6EKE			6FX1	6FX2		6FX3	

[a]Values in parentheses are for highest-resolution shell. [b] R_{free} was calculated with 5% of the reflections.

Table S2. Torsion angles observed per protein chain for the glycosidic linkages of the core fucosylated N-glycans **1**, **2**, and **3** bound to rPhoSL.

Complex	N-glycan 1												N-glycan 2		N-glycan 3	
Chain	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**A**	**B**	**A**	**B**
Fuc(α1,6)GlcNAc																
Φ*	-69	-66	-61	-62	-68	-64	-62	-67	-65	-59	-53	-60	-67	-58	-67	-65
Ψ*	-142	-143	-145	-153	-140	-150	-141	-152	-145	-144	-151	-156	-145	-146	-144	-142
Ω*	-59	-56	-55	-62	-65	-51	-61	-63	-66	-61	-70	-55	-60	-59	-59	-54
GlcNAc(β1,4)GlcNAc																
Φ*	-71	-66	-75	-75	-73	-74	-72	-79	-76	-71	-72	-80	-85	-81	-83	-78
Ψ*	-132	-130	-131	-124	-136	-130	-136	-128	-134	-135	-128	-129	-141	-129	-143	-131
GlcNAc(β1,4)Man																
Φ*	-77	-73	-74	-76	-69	-78	-71	-78	-74	-74	-66	-71	-74	-86	-73	-77
Ψ*	-141	-148	-130	-128	-125	-135	-127	-119	-124	-117	-126	-120	-139	-146	-141	-144
Man(α1,3)Man																
Φ*		65		83	73		54		-83				72		69	
Ψ*		132		131	103		143		131				133		129	
Man(β1,2)GlcNAc on 1,3 branch																
Φ*		-92							-78				-90		-89	
Ψ*		153							152				146		152	
Man(α1,6)Man																
Φ*	81	71	76		56	89	53		61			59	79	75	81	72
Ψ*	-165	177	-164		172	-164	-178		168			171	-162	-162	-160	-159
Man(β1,2)GlcNAc on 1,6 branch																
Φ*	-91	-84	-92		-94	-83	-72		-104			-105	-93	-87	-89	-84
Ψ*	111	135	100		151	91	110		154			143	104	103	101	102
GlcNAc (β1,4)Gal on 1,6 branch																
Φ*													-87	-87	-90	-90
Ψ*													-146	-148	-149	-148
Sia(α2,6)Gal on 1,6 branch																
Φ*															68	56
Ψ*															-179	-175

***φ** (O_5-C_1-O_X-C_X) and **ψ** (C_1-O_X-C_X-C_X+1) with x, the number of the carbon atom of the second monosaccharide with which the 1→ x glycosidic bond is formed and **Ω** (O_6-C_6-C_5-O_5) for two monosaccharides linked by a 1 → 6 linkage.

References

[1] A. S. do Nascimento, S. Serna, A. Beloqui, A. Arda, A. H. Sampaio, J. Walcher, D. Ott, C. Unverzagt, N. C. Reichardt, J. Jimenez-Barbero, K. S. Nascimento, A. Imberty, B. S. Cavada, A. Varrot, *Glycobiology* **2015**, *25*, 607-616.

[2] Y. Kobayashi, H. Tateno, H. Dohra, K. Moriwaki, E. Miyoshi, J. Hirabayashi, H. Kawagishi, *J. Biol. Chem.* **2012**, *287*, 33973-33982.

[3] R. B. Kapust, J. Tözsér, J. D. Fox, D. E. Anderson, S. Cherry, T. D. Copeland, D. S. Waugh, *Protein Engineering, Design and Selection* **2001**, *14*, 993-1000.

[4] W. Kabsch, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125-132.

[5] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235-242.

[6] G. M. Sheldrick, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 479-485.

[7] G. Langer, S. X. Cohen, V. S. Lamzin, A. Perrakis, *Nat. Protoc.* **2008**, *3*, 1171-1179.

[8] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658-674.

[9] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355-367.

[10] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486-501.

[11] M. D. Winn, M. N. Isupov, G. N. Murshudov, *Acta Crystallogr D Biol Crystallogr* **2001**, *57*, 122-133.

[12] J. Agirre, J. Iglesias-Fernandez, C. Rovira, G. J. Davies, K. S. Wilson, K. D. Cowtan, *Nat. Struct. Mol. Biol.* **2015**, *22*, 833-834.

[13] G. Hester, H. Kaku, I. J. Goldstein, C. S. Wright, *Nat. Struct. Mol. Biol.* **1995**, *2*, 472-479.

[14] X. Robert, P. Gouet, *Nucleic Acids Res.* **2014**, *42*, W320-324.

[15] L. R. Olsen, U. J. Kudahl, C. Simon, J. Sun, C. Schönbach, E. L. Reinherz, G. L. Zhang, V. Brusic, *J. Immunol. Methods* **2013**, *400-401*, 37-44.