Increasing the affinity of an O-antigen polysaccharide binding site in *Shigella flexneri* bacteriophage Sf6 tailspike protein

Sonja Kunstmann[1,2,†], Olof Engström[3], Marko Wehle[2], Göran Widmalm[3*], Mark Santer[2*], Stefanie Barbirz[1*]

1 Physikalische Biochemie, Universität Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany.

2 Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14476 Potsdam, Germany.

3 Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, 10691 Stockholm, Sweden.

† Present address: Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kgs. Lyngby, Denmark

* Corresponding authors: Tel: +49 331 567 9610, mark.santer@mpikg.mpg.de; Tel: +49 331 977 5322, barbirz@uni-potsdam.de; Tel: +46 8 16 37 42, goran.widmalm@su.se

Abstract

The unrestrained use of antibiotics against bacterial infections has accelerated the occurrence of resistances in many pathogens. In order to reduce use of antibiotics, sensitive and robust pathogen detection is necessary. Bacteriophage tailspike proteins (TSPs) are large, highly stable protein fibers with extended binding sites for microbial glycans which makes them a promising design platform for bacterial cell surface polysaccharide sensors.

We analysed TSP from bacteriophage Sf6 in complex with the O-polysaccharide of the pathogen *Shigella (S.) flexneri*. The conformational space populated by the polyrhamnose backbone of the *S. flexneri* O-polysaccharide as studied by an octasaccharide in complex with Sf6TSP could be well described with 2D ^{1}H,^{1}H-trNOESY NMR, utilizing a combination of methine-methine and methine-methyl correlations. The results are in good agreement with the conformations obtained from molecular dynamics (MD) simulations. To examine the impact of amino acid exchanges in the glycan binding site of Sf6TSP, MD simulations were used to predict increased O-polysaccharide binding affinities. We used surface plasmon resonance on *S. flexneri* O-polysaccharide surfaces to measure affinity increases in the obtained mutants. Combining *in silico* approaches with rapid screening on polysaccharide surfaces is thus a promising approach for the design of tailspike protein-based pathogen sensors.

Introduction

Improvement of protein properties like stability, enzyme activity or binding affinity can be obtained by protein design and engineering techniques.[1] Scaffold design is used in *de novo* approaches to obtain proteins with defined geometries[2] or with altered binding specificities or enzymatic activities.[1] Experimental techniques rely on randomizing coding DNA sequences by error-prone PCR or directed evolution and subsequent selection procedures towards the desired protein property.[3] Rational design often starts from computational approaches like ROSETTA that utilizes physical energy functions to sample the free energy space populated by a given amino acid sequence.[4–6] This strategy has been further extended by taking into account evolutionary profiles[7] or "re-epitoping" of antibodies and validation by crystal structure analysis.[8]

Proteins binding carbohydrate ligands are an especially important design target. Most cellular envelopes are covered with glycan matrices and addressing these surfaces with tailor-made protein-binders has impact both in diagnostics and therapy.[9–12] For engineering of carbohydrate binding proteins, phage-display has been employed to modify binding sites for charged carbohydrate structures in antibodies.[13,14] As another example carbohydrate-binding modules were used as robust protein scaffolds with high stability for computational design of high-affinity drug compound binders.[15] The rational design of protein-ligand high-affinity binders depicts a special challenge compared to improvement of enzyme specificities or affinities of protein-protein complex formation.[15] Moreover, the unique amphiphilic and conformational properties of carbohydrate-based ligands impose major challenges on modelling protein-glycan complexes.[16] Flexible ligands and water network distributions impede predictions via docking or Monte-Carlo simulations with rotamer libraries. Algorithms like ROSETTALIGAND,[17] which are commonly used for other high-affinity design purposes, are in the process to be optimized for protein-carbohydrate interactions, *i.e.* ROSETTADOCK.[18] Although computationally more expensive, techniques based on molecular dynamics simulations allow for higher precision in the description of protein-glycan complexes.[19] However, no common strategy exists for the automated selection of mutations for design of high-affinity carbohydrate-binding proteins.

MD simulations were shown to be well suited for sampling elongated carbohydrate ligands in the extended glycan binding sites found in bacteriophage cell surface receptor adhesion systems.[20–23] Tailspike proteins (TSP) act as recognition organelles for binding and orientation of bacteriophages during infection of the bacterial host.[24–26] TSPs have been also used as model systems to computationally link solvent network structures to experimental thermodynamic signatures of oligosaccharide ligand binding.[27] Sf6TSP from bacteriophage Sf6 is a specific receptor of the *Shigella (S.) flexneri* Y O-antigen with the repeat unit (RU) structure [→3)-α-L-Rha*p*-(1→3)-β-D-Glc*p*NAc-(1→2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→].[20,28] It is a highly stable trimeric protein with

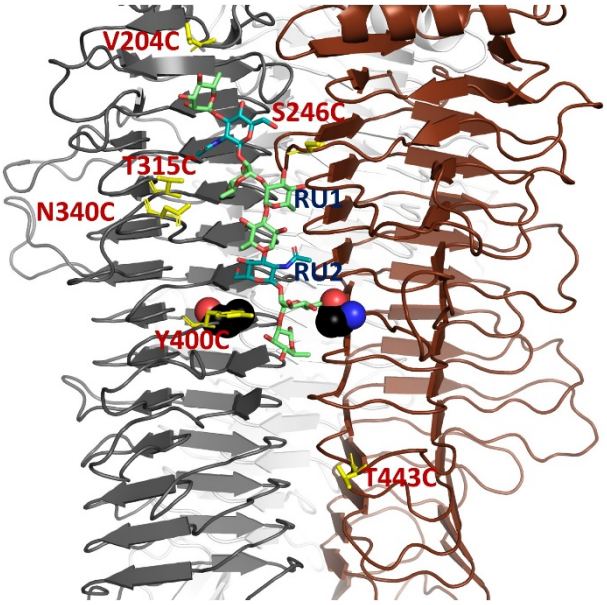


Figure 1. Sf6TSP glycan binding site in complex with octasaccharide and set of mutations
Three subunits of Sf6TSP E366A D366A (pdbID: 4URR, gray/brown/silver backbone, cartoon) are illustrated with octasaccharide ligand (sticks in green for Rhamnose and light blue for GlcNAc) as RU1 and RU2. All residues chosen for a cysteine mutation are shown as yellow sticks with the corresponding label in red. The catalytic active residue positions are depicted in black spheres of D399A (gray chain) and E366A (brown chain).

endorhamnosidase activity, producing oligosaccharides of mainly 2 RU, *i.e.* octasaccharides, from the *S. flexneri* O-antigen polysaccharide.[20,28–30] Sf6TSP has an elongated glycan binding site with low affinity towards oligosaccharide O-antigen fragments produced by hydrolysis. Nevertheless, combination of MD simulations, X-ray crystallography and NMR spectroscopy showed an octasaccharide binding state with a highly flexible reducing end pointing towards the enzyme's active site residues (Figure 1).[20] These suggested the protein as suitable for application of numerical tools for affinity prediction.

The goal of the present study was to characterize engineered Sf6TSP variants with increased glycan affinity. An enzymatically inactive Sf6TSP variant was used as a sensor to detect *S. flexneri.*[9] This pathogen causes dysentery in infants and is an important diagnostic target.[31–33]A high-affinity Sf6TSP would be valuable in improving detection limits for lipopolysaccharide (LPS) or O-polysaccharide (PS), the major glycan targets in TSP-based pathogen sensors and provide an important addition to antibody-based applications. TSP-based pathogen detection has been achieved on microtiter plates with an ELISA-like read-out.[9,34] In solution, fluorescence amplitude changes upon glycan ligand binding were used to detect O-polysaccharide via a cysteine coupled, environment-sensitive fluorescent label in the Sf6TSP binding site.[9] A diverse set of Sf6TSP mutants with single amino acid exchanges in the binding site was thus available for affinity studies in the present work (Figure 1).[9] Our study combines MD simulations with NMR data to confirm the conformational space sampled by a protein-bound oligosaccharide. Our test set shows that MD simulations can link mutations in the Sf6TSP binding site to glycan affinity as experimentally confirmed with surface plasmon resonance.

Results

MD simulations show varying glycan ligand flexibilities in different Sf6TSP mutants

Sf6TSP is a well-studied, weak-affinity carbohydrate binding protein as analysed with NMR, X-ray crystallography and MD simulations.[20] We used molecular dynamics simulations to analyze the conformational space occupied by the protein-bound oligosaccharide ligand in order to make assumptions on the affinity to Sf6TSP in different mutational backgrounds. The available Sf6TSP cysteine mutant set was based on the O-antigen hydrolysis deficient variant Sf6TSP E366A D399A ($Sf6TSP_{EADA}$). $Sf6TSP_{EADA}$ was crystallized with an octasaccharide ligand previously and is used as the reference binding system.[20] Moreover, for construction of an *S. flexneri* O-polysaccharide sensor, single cysteine exchanges were introduced into $Sf6TSP_{EADA}$.[9] Five of these mutations (V204C, S246C, T315C, N340C and Y400C) lie in close proximity to the octasaccharide binding site, whereas the mutation T433C is situated below the reducing end of the octasaccharide binding groove (Figure 1). MD simulations had previously shown that oligosaccharide fragments longer than octasaccharides could occupy this part of the protein surface, *i.e.* a modelled dodecasaccharide formed H-bond contacts with binding site residues at this proximal position.[20] Mutations V204C, T315C and T443C replace hydrophobic residues with cysteine as a potential hydrogen bond donor or acceptor whereas S246C, N340C and Y400C can potentially influence hydrogen-bonding patterns compared to the reference. All Sf6TSP cysteine variants were run in 100 ns MD simulations in TIP3P water in the presence of an octasaccharide ligand in the binding site.

To predict differences in binding affinity, octasaccharide conformations that occurred during the simulation were clustered to describe ligand flexibility in the different mutational backgrounds. For the $Sf6TSP_{EADA}$ reference octasaccharide complex containing no other mutations, five major ligand conformer clusters were defined (Figure 2). In contrast, simulations of V204C, S246C and T315C only yielded two major clusters, and these overlaid well with the crystal structure ligand conformer (Figure 2a-c). N340C, Y400C and T443C had a larger set of clusters than the reference structure, with tilted conformations at the oligosaccharide reducing end that bent away from the central binding groove (Figure 2d-f). In addition, the root mean square displacement (RMSD) of the ligand over the simulation time reflected different ligand behaviours (Supplementary Figure S1). For Y400C the mean RMSD and its standard deviation were notably increased compared to the $Sf6TSP_{EADA}$ reference (Table 1). During the 100 ns simulation, the octasaccharide gradually detached from the Y400C mutant binding site, suggesting a loss in oligosaccharide affinity. In contrast, ligand fluctuations were reduced with mutants V204C and S246C, indicating a less flexibly bound ligand, which may result in increased binding affinity. Although simulations with mutants T315C and N340C also showed a decreased mean ligand RMSD, here a higher standard deviation was found that might indicate changes in the H-bond donor-acceptor equilibrium with no clear indications for an effect on affinity. The T443C mutant

showed approximately the same mean RMSD and standard deviation as the $Sf6TSP_{EADA}$ reference. Thus, ligand flexibility analysis by MD simulations suggested that Sf6TSP mutations V204C and S246C create stronger octasaccharide binders, whereas the Y400C mutation results in loss of glycan affinity. Simulations with Sf6TSP N340C, T315C and T443C in contrast did not show ligand conformer deviations that were indicative for binding affinity changes in these mutants.

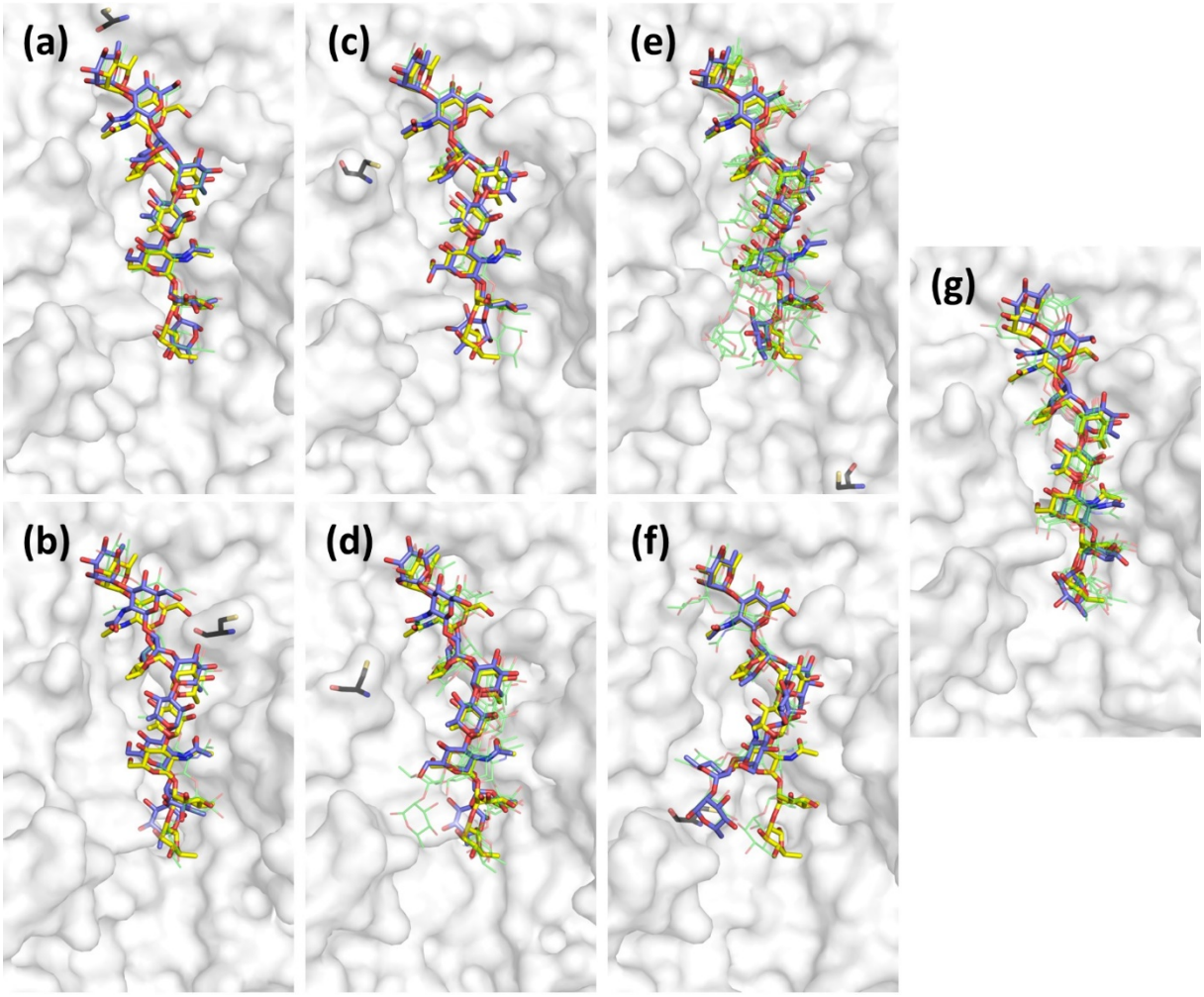


Figure 2. Octasaccharide ligands in the Sf6TSP binding site with flexibility analysis from 100 ns MD simulations.
Typical conformational clusters obtained for oligosaccharide ligands are superimposed onto the ligand pose from crystal structure analysis (yellow) from **(a)** V204C, **(b)** S246C, **(c)** T315C, **(d)** N340C, **(e)** T443C, **(f)** Y400C and **(g)** as the reference (E366A D399A). Most prevalent conformers are shown in blue, the full conformational space sampled is illustrated by conformers deviating from average (green). Black sticks indicate positions of the residue exchanged in each mutant.

Table 1. Oligosaccharide ligand properties in the Sf6TSP binding site during 100 ns MD simulation

	Reference	V204C	S246C	T315C	N340C	Y400C	T443C
RMSD / Å[a]	1.68±0.37	1.54±0.25	1.21±0.31	1.56±0.43	1.55±0.81	2.61±1.13	1.48±0.38
# clusters[b]	5	2	2	2	4(bended)	3(bended)	7(bended)
Relative H-bond occupancy[c]	1.00	1.34	1.56	1.03	1.08	1.42	1.01
# H-bonds[d]	14	10	11	13	12	14	13

[a]RMSD were calculated based on all atoms of the ligand. [b]Number of clusters were defined with a cut-off of 1.1 Å. [c]H-bond occupancies calculated by VMD1.9.1. with 3.5 Å distance and angles of 40°. [d]Only occupancies above 10 % were taken into account.

All Sf6TSP cysteine mutants had been probed earlier for thiol-based covalent attachment of fluorescent dyes to obtain environment-sensitive glycan-binding sensors for detection of *Shigella flexneri* pathogens.[9] Especially the N340C variant had shown high labelling efficiency with N-methyl-N-[2-[methyl(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]ethyl] (NBD). Moreover, Sf6TSP N340C-NBD ($Sf6TSP_{NBD}$) had the most notable fluorescence amplitude increase when exposed to *S. flexneri* O-polysaccharide Y, making it a promising candidate for a *S. flexneri* sensor. We therefore included $Sf6TSP_{NBD}$ in our MD simulations to study the glycan ligand pose in presence of the covalently attached fluorescent dye. For this, we created a parameter set for the NBD-modified cysteine at position 340, using either the AMBER or GAFF force field or a mixture of both. The three force field sets resulted in the same flexibility behaviour of label and ligand during the simulations (Supplementary Figure S2). In simulations with the ligand-free $Sf6TSP_{NBD}$, most of the time, the label was buried in the cleft between the two subunits forming the glycan binding site (Figure 3a). In presence of the octasaccharide, the label shows increased flexibility (Figure 3b) and is moved out of the cleft and the aromatic ring of the label is situated in parallel to NAG2 of the octasaccharide ligand. This results in a new CH,π-interaction, distorting the ligand from its original binding site position towards the label (Figure 3c). Additionally, the octasaccharide ligand had a notably increased RMSD in the binding site compared to the complex with the unlabelled protein (Figure 3d).

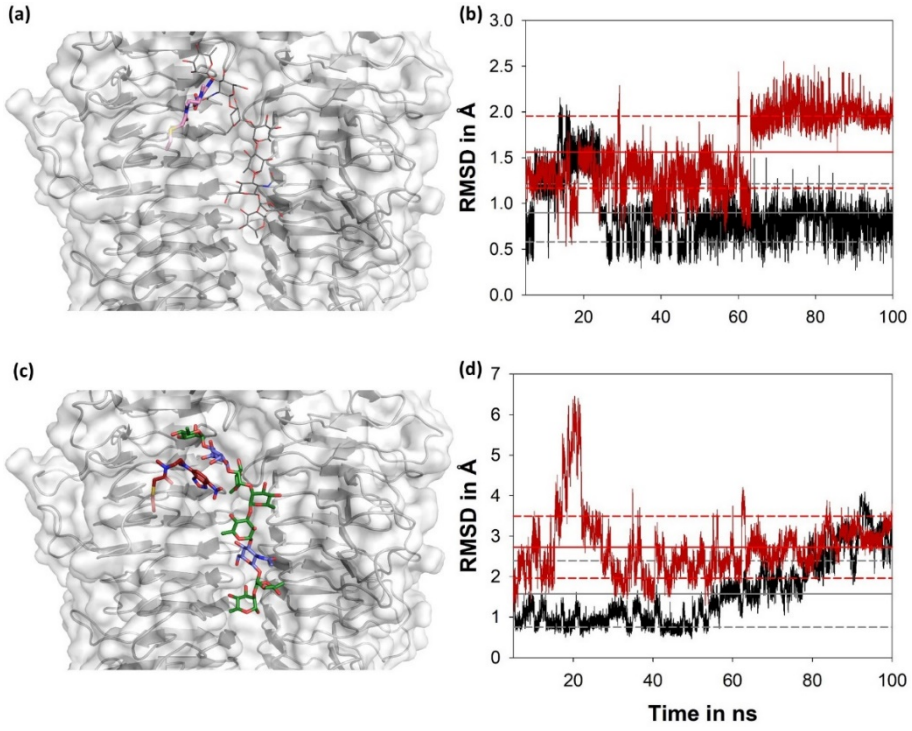


Figure 3. Sf6TSP N340C labelled with NBD simulated with octasaccharide ligand for 100 ns.
(a) Most abundant NBD label position (purple sticks) in a ligand free simulation. The label is buried between two protein subunits (white surface, grey cartoon) underneath the ligand binding site. An octasaccharide shown in grey thin sticks marks the ligand position in the crystal structure pdbID 4URR. **(b)** RMSD of the fluorescent label NBD covalently attached to Cys340 of Sf6TSP with (red) or without ligand (black). (Solid: average RMSD, Dashed: standard deviation). **(c)** Most abundant NBD label position (dark red sticks) in the presence of an octasaccharide (Rhamnoses: green, GlcNAc: blue). **(d)** Octasaccharide ligand RMSD in the Sf6TSP N340C binding site without label (black) or with label (red) depicted as in B. The figure shows the results with the AMBER force field parameters.

2D ^{1}H,^{1}H transfer NOESY NMR confirms torsion angles at oligosaccharide methyl and methine groups

The *S. flexneri* Y octasaccharide contains α-L-rhamnosyl residues to a large extent. Their methyl groups at C6 enable increased flexibility of glycans compared to contributions of other, fully hydroxylated monosaccharide building blocks.[21,35] In the 2D ^{1}H,^{1}H-trNOESY NMR analysis, the methyl groups at C6 facilitate a unique conformational landscape to be revealed, where correlation times between fast-spinning methyl groups and methine protons are significantly shorter than for the methine-methine interactions in the oligosaccharide;[36] this is the basis for using different correlation times in the analysis of the experimental NMR data. We analysed the Sf6TSP-octasaccharide complex and were able to confirm the observed ligand conformations obtained from MD simulations (Figure 4). Evaluation of the 2D ^{1}H,^{1}H-trNOESY derived proton-proton distances (Supplementary Table S1) resulted in two radial φ,ψ-distributions with two intersection points. An overlay with glycosidic linkage conformations previously determined for the *S. flexneri* O-serotype Y octaccharide with MD simulations and X-ray crystallography then could define the main populated ϕ,ψ torsional angle conformational space in the Sf6TSP-bound octasaccharide. Only linkage 7 had a less defined glycan geometry, which is in agreement with the increased flexibility of the reducing end as observed in simulation and crystallographic B-factors.[20] Additionally, effective proton-proton distances agree in each data set between simulation and NMR cross-relaxations (Supplementary Table S1).

We could hence extend the NMR methodological repertoire with this 2D ^{1}H,^{1}H-trNOESY technique and analyse conformational behaviour of an octasaccharide ligand in the Sf6TSP binding site. Methine-methyl correlations of rhamnose can be used to describe different glycan geometries in the ϕ,ψ-space. MD simulations further revealed the possible oligosaccharide conformers and thus reliably defined the glycan conformational space occupied when fixed in a defined protein environment.

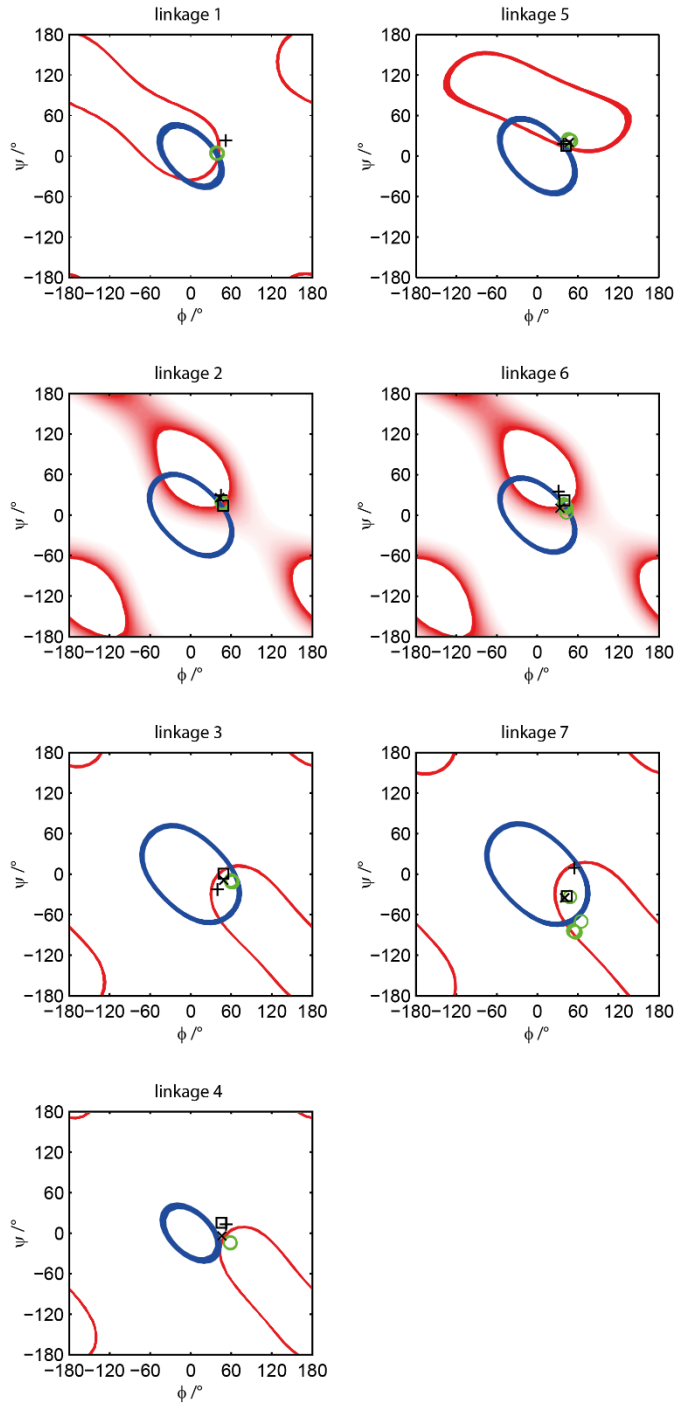


Figure 4. Distance curves for the octasaccharide describing the loci in terms of the glycosidic torsional angles ϕ and ψ obtained by 2D ^{1}H,^{1}H transfer NOESY NMR

Two-dimensional ϕ,ψ distance-plots are shown for all glycosidic linkages, denoted by superscripts in the *S. flexneri* O-serogroup Y octasaccharide with the structure α-L-Rha*p*-(1→3)$^{(1)}$-β-D-Glc*p*NAc-(1→2)$^{(2)}$-α-L-Rha*p*-(1→2)$^{(3)}$-α-L-Rha*p*-(1→3)$^{(4)}$-α-L-Rha*p*-(1→3)$^{(5)}$-β-D-Glc*p*NAc-(1→2)$^{(6)}$-α-L-Rha*p*-(1→2)$^{(7)}$-α-L-Rha*p*. Conformational range calculated from methine trNOEs (blue lines) and methyl trNOEs (red lines) is shown, as determined from the effective proton-proton distances r_{ij} presented in Supplementary Table S1. Overlaid are conformations obtained from six non-redundant X-ray models (green circles) or MD simulations with AMBER (+, black pluses) or CHARMM (□, black squares) force fields. In addition, conformers calculated from MD simulations with the Sf6TSP D399N are shown (x, black crosses, AMBER). These data points were obtained from Kang et al., 2016.[20] For the panels of linkages (2) and (6) the red color gradient corresponds to the sum of the overlapped NOE cross-peak volumes of the NAG2-acetyl to RAM3-H4 and NAG6-acetyl to RAM7-H4. Red depicts the total volume of the overlapped signals and white corresponds to NOE being absent or stronger than the sum of the overlapped cross-peak volumes (cf. Supplementary Table S1).

Multivalent Sf6TSP binding on surface-conjugated O-antigen polysaccharides assessed with surface plasmon resonance

To construct a platform for rapid screening of the different Sf6TSP binding site mutants, we set up a multivalent *S. flexneri* O-polysaccharide Y (SfY) surface plasmon resonance analysis platform. We validated binding properties to these surfaces with the polysaccharide hydrolysis deficient reference Sf6TSP double mutant E366A D399A prior to screening of the mutants.

The SfY O-polysaccharide was prepared from lipopolysaccharide by acidic hydrolysis of the lipid A part. This results in a polysaccharide with an inner core Kdo residue as the new reducing end.[37] However, trials to directly couple this O-polysaccharide via a hydrazide-modified carboxymethyl dextran surface were unsuccessful. To obtain sufficient amounts of surface immobilized SfY O-polysaccharide, mild oxidation with sodium periodate was thus required prior to coupling.[38] To rule out that the oxidation step had altered the SfY O-polysaccharide, we tested periodate-oxidized SfY O-polysaccharide as an enzyme substrate for the Sf6TSP wild type enzyme. We found comparable amounts of oligosaccharide products produced both from oxidized and non-oxidized SfY polysaccharide samples, in agreement with an unaltered carbohydrate substrate structure (Supplementary Figure S3).

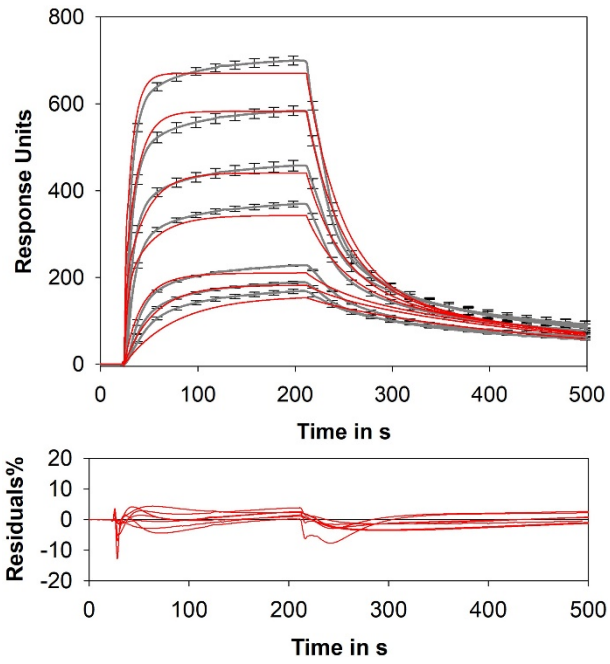


Figure 5. Surface plasmon resonance analysis of Sf6TSP binding to *S. flexneri* O-polysaccharide Y.
SPR response curves obtained by injections of serial dilutions of $Sf6TSP_{EADA}$ (0.08-18 µM subunit concentration, grey, average of triplicate measurements with standard deviations shown for every 10th data point). Curves were fitted with a 1:2 heterogeneous binding model (red).

We then assessed protein binding to the SfY O-polysaccharide modified surface with surface plasmon resonance (SPR) and found specific binding of the Sf6TSP reference protein (Figure 5). As Sf6TSP is a homotrimeric protein with three independent glycan binding sites between the subunits, we assume a multivalent interaction. A heterogeneous ligand parallel-binding model (1:2 model) provided the best description of the data (χ^2~380), whereas simple 1:1 binding could be excluded (Supplementary Figure S4a). Fitting the association and dissociation isotherms to the 1:2 model thus resulted in two equilibrium dissociation constants of K_{D1}=19.7±1.8 nM and K_{D2}=5.4±1.9 µM (Figure 5). Reasonable similar values were obtained when evaluating the equilibrium signals at 178 s of injection, *i.e.* K_{D1}=37.0±23.0 nM and K_{D2}=6.2±1.3 µM (Supplementary Figure S4a). Multivalent binding was additionally confirmed by concentration dependent data analysis, resulting in a curved, triphasic distribution in the Scatchard plot (Supplementary Figure S4b).

Furthermore, the kinetic constants obtained revealed a good agreement of the second dissociation rate, $k_{off,2}$≈0.024-0.027 s^{-1}, with the k_{off} of 0.0335 s^{-1} that was obtained previously from binding equilibrium relaxation analysis with fluorescence spectroscopy in the same system.[20] In all kinetic curve-fitting analyses, a systematic deviation of the fit from the data was observed, mainly due to a slight, but constant signal increase over the whole incubation time. However, this binding curve shape did not change with different incubation times and protein flow rates over the surface, excluding mass transport effects or unspecific binding (Supplementary Figure S4c).[39,40] The polysaccharide surface was highly stable and more than 300 experiments were performed on one single surface (Supplementary Figure S4d). All measurements were repeated on two chip surfaces (1 and 2) resulting in comparable kinetic constants (Table 2, Supplementary Table S2). However, equilibrium dissociation constants calculated from the SPR signals at the end of injection deviated from those calculated from kinetic data. This illustrates that batch-to-batch surface variations between the two chips apparently led to varying amounts of unspecific binding. This seemingly influenced the absolute signals evaluated in the equilibrium binding isotherm, whereas the kinetic constants remained unaffected. Nevertheless, all equilibrium dissociation constants from equilibrium and kinetic experiments were in the same order of magnitude. We conclude that the O-polysaccharide surfaces provided a robust and rapid SPR platform to screen for protein binding both with kinetic or equilibrium methods.

Table 2. Dissociation constants for Sf6TSP$_{EADA}$ derived from 1:2 heterogeneous ligand binding kinetics and equilibrium

Exp. #	$k_{on,1}$ $10^3 M^{-1} s^{-1}$	$k_{on,2}$ $10^3 M^{-1} s^{-1}$	$k_{off,1}$ $10^{-3} s^{-1}$	$k_{off,2}$ $10^{-3} s^{-1}$	$K_{D1,\ eq.}$ *nM*	$K_{D1,\ kin.}$ *nM*	$K_{D2,\ eq.}$ *µM*	$K_{D2,\ kin.}$ *µM*
1	154 ± 73.8	5.1 ± 5.8	1.50 ± 0.02	24.7 ± 0.018	85 ± 37	9.7 ± 6.2	47.1±100	4.88±18
2	157 ± 13.6	5.07 ± 1.59	3.09 ± 0.01	27.3 ± 0.004	37 ± 23	19.7 ± 1.8	6.15±1.33	5.38±1.87

Sf6TSP mutant binding screened on SfY polysaccharide surfaces with SPR

Different ligand flexibilities in the protein complexes were computer-generated from MD simulations. To experimentally access affinities of complex formation, we screened the Sf6TSP mutant set on the SfY polysaccharide surface described above. The automated set-up allowed to collect all SPR signals in triplicate at three different protein concentrations, *i.e.* nine individual measurements were carried out for each mutant to obtain a comparison of maximum response signals (Figure 6a and Supplementary Figure S5a). As described above, binding of the Sf6TSP reference to SfY O-polysaccharide surfaces showed multivalent binding with two equilibrium dissociation constants of ~10^{-8} M and ~10^{-6} M. For rapid and simple evaluation of the mutant binding strength, we thus chose to compare variations in the maximal response signals of each injection, assuming that all mutants exhibit the same non-specific signal of around 4 % on the reference channel on an individual chip surface (Supplementary Figure S5b).[41] The maximal responses for Sf6TSP binding scaled with the concentration in all mutants. Five out of six mutants showed similar or slightly increased binding signals compared to the reference. Only Y400C bound to the polysaccharide surface with evidently reduced response. We therefore

related the distribution of maximum SPR signals between the different mutants to ligand flexibilities calculated from MD simulations. Mean values of positional RMSD were similar for all mutants that showed similar or higher SPR responses compared to the reference (Figure 6b). However, individual ligand fluctuations were decreased for V204C and S246C that showed the highest SPR surface binding responses. In contrast, Y400C had a notably increased mean ligand RMSD with pronounced fluctuations, in agreement with this it showed a low SPR response, indicating weak binding to the SfY polysaccharide surface (Figure 6c).

To obtain a stronger differentiation between the mutants, we analysed hydrogen bond patterns from simulation in addition to ligand flexibility. These could only in part corroborate the estimate of binding strength (*cf.* Table 1 and Supplementary Table S3). For example, Y400C showed high hydrogen bond occupancies, although this mutant was a poor binder on the SfY polysaccharide surface. In contrast, mutants T315C, N340C and T443C, with a maximum SPR response similar to the reference, displayed a similar hydrogen bond occupancy. In addition, V204C and S246C had higher occupancies in agreement with higher SPR signals. This illustrates that analysis of the hydrogen bonding pattern alone was not sufficient to predict the affinity of complex formation. Rather, our simulations showed that single mutations on a protein surface can change completely or partly the observed hydrogen bonding patterns in number and occupancies, leading to individual energetic signatures of glycan complex formation in a given mutational background.

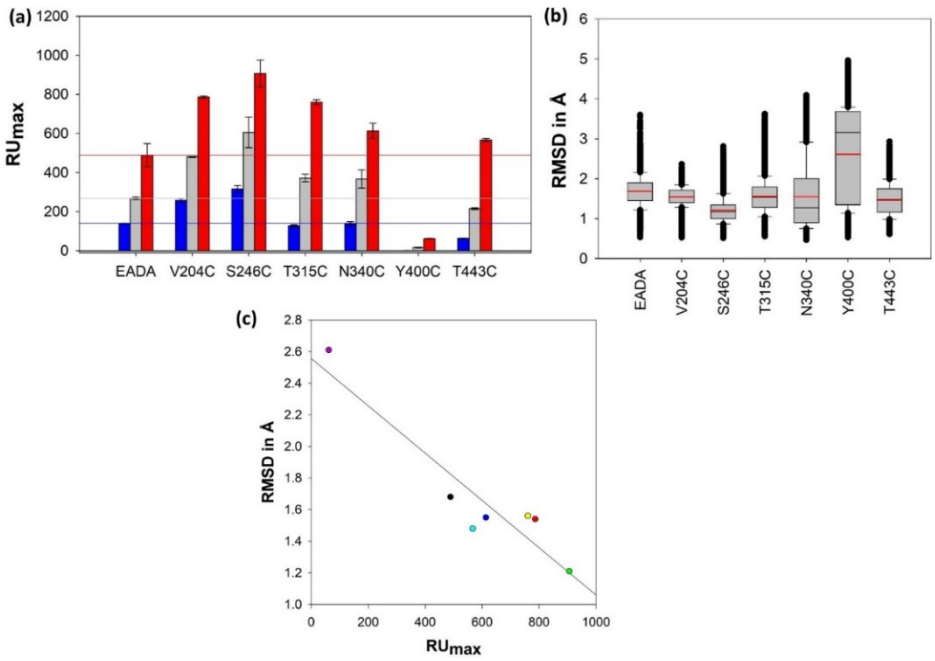


Figure 6. Interaction data derived by SPR of cysteine mutants in comparison with ligand flexibilities from MD simulations
(a) Maximal responses of injection with Sf6TSP$_{EADA}$ cysteine mutants measured as triplicates for protein concentrations of 0.08 (blue), 0.8 (grey) and 8 µM (red). Horizontal lines mark the maximal response of Sf6TSP$_{EADA}$ in the respective concentration. **(b)** Time dependent fluctuations of RMSD of the octasaccharide in the binding site of the corresponding cysteine mutant depicted as box plots. Red solid lines show the mean RMSD. Upper, middle and lower black lines indicate 75 %, median and 25 % of the fluctuation, respectively. Error bars represent one standard deviation and black circles (thick lines) all outliers of the fluctuation. **(c)** Correlation plot of the maximal responses of the mutants V204C (red), S246C (green), T315C (yellow), N340C (blue), Y400C (purple), T443C (cyan) and EADA (black) with 8 µM concentrations against the mean ligand RMSD.

Discussion

In this work, affinity enhancements in an elongated protein binding site for a complex bacterial cell surface oligosaccharide were probed with experimental and computational methods. The computational prediction was based on evaluation of ligand flexibilities during 100 ns simulations. MD results were further validated by the analysis of trNOE at methine and methyl signals. They are particularly useful for the conformational analysis of the deoxysugar rhamnose. Rhamnose is a main constituent in many microbial polysaccharides, for example rhamnose containing biofilms were shown to effectively block antimicrobial peptides from entering biofilms.[42]

Bacteriophage tailspike proteins recognize bacterial cell surface glycans with high specificity, which predestines them as sensor proteins for pathogens. Typical optical signal read outs might rely on peptide or fluorescent tags, the latter were coupled to the TSP of choice via genetically engineered cysteine residues.[9] Cysteine mutants of the bacteriophage Sf6TSP addressing *S. flexneri* serogroup Y therefore provided the test set for affinity screens of the Sf6TSP glycan binding groove. Typically, cysteines do not occur with high frequency in carbohydrate binding sites.[43,44] In our mutant set, all single amino acid exchanges were located in loops adjacent to the elongated binding site. Combining experimental and computational methods, we could identify the mutant Sf6TSP S246C as binder with an increased affinity due to decreased ligand flexibility. Serine 246 participates in the hydrogen bond contacts to the carbohydrate ligand via its backbone oxygen, consequently, exchange for a cysteine increased this hydrogen bond occupancy due to the reduced electronegativity of sulphur.[45]

Cysteines may form intermolecular disulphide bridges and are thus often excluded from rational design approaches.[46] Consequently, we cannot exclude that a cysteine at position 246 favours intermolecular disulphide bond formation. Sf6TSP is a native trimer with three glycan binding sites, oligomerisation would further increase this multivalence and have an avidity effect, even if single glycan binding sites might be buried in the higher protein oligomer assembly.[47,48] Nevertheless, the computational analysis of the Sf6TSP S246C binding site clearly shows that already on the level of a single oligosaccharide binding site the increased affinity is most probably linked to the reduced ligand flexibility.

In the case where a fluorescent label is present in the binding site, MD simulations moreover revealed the binding conformations that lead to fluorescence amplitude gain upon glycan binding in the Sf6TSP fluorescent sensor. Here, an additional CH,π-interaction between the fluorescent label and a GlcNAc residue in the octasaccharide occurred. Consequently, for the environment sensitive label NBD, an increased fluorescent signal can be observed, because glycan binding creates a more hydrophobic environment when shielding the aromatic ring from the solvent.[49,50]

MD simulations in this work stressed that the experimentally accessible affinity changes of Sf6TSP mutants can be well reproduced by a ligand flexibility analysis in the complex. This approach is successful because it implicitly contains all changes in binding site water distribution and hydrogen

bond formation without the need of their explicit analysis. A mobile water positional network on the protein surface is accessible with *MobyWat* analysis that is based on evaluation of short solvent MD simulations.[27,51,52] In Sf6TSP, however, this method could not distinguish water networks between the different mutants analysed, stressing that more sophisticated network analysis algorithms are needed in the future (Supplementary Figure S6).[53]

No general rule exists that assigns favourable or unfavourable contributions of water molecules to the driving forces for ligand binding.[54] In Sf6TSP, the glycan interaction site is formed by flexible protein loops that form the binding groove between two protein subunits.[28] The tailspike protein from bacteriophage HK620 has a very similar overall fold compared to Sf6TSP.[30] However, in HK620TSP the binding site lies not between subunits, but in a shallow surface depression formed by a rather hydrophobic ß-sheet, accommodating a ligand glucose branch in an occluded surface cavity.[28,55] As a consequence, in HK620TSP the redistribution of water molecules has a major impact on the enthalpy-entropy compensation during the ligand binding event.[27] In Sf6TSP, the binding site is less dominated by water molecules than in HK620TSP,[20] and the flexible loop regions can easily adjust to the ligand during complex formation without major solvent rearrangements being necessary. Oligosaccharide binding to Sf6TSP thus occurs as the most populated solution conformer,[21] and already subtle ligand conformational changes may favour dissociation, resulting in an overall low affinity.[56]

The mutant design approach in this work was rationally based on structural examination and computational analysis by conventional MD simulations and the RMSD measure to characterize ligand affinity in a “rule of thumb” manner. The use of accurate free energy methods for quantitative comparison across all mutants would imply a considerable and non-trivial extension of the computational part. Alchemical methods can be reliable in predicting the stability of polypeptides up to 20 amino acids.[57] In our case however, the frequent detachment of the glycan ligand to be expected during alchemical amino acid morphing may lead to significant undersampling. Either the increase in computational power is improved,[58] or the alchemical transformation are treated in a way that avoids undersampling from the beginning, by imposing suitable restraints on the ligand, which is clearly beyond the scope of the present work but will be reconsidered in a subsequent publication. Here, improved docking algorithms might be compared which described complex formation with long oligosaccharides, for example docking of a SfY pentasaccharide on a FAB fragment.[59] Also ROSETTA-based techniques have been extended to include carbohydrate moieties on glycoproteins and glycan ligand docking.[60] However, for all techniques applied it is important to emphasize that a reliable computational description of carbohydrate-protein complexes must be linked to the appropriate water scoring functions or include water explicitly.[61]

For the rapid comparison of proteins, differing in their affinity to a certain ligand a measurement set-up with options for screening is needed, in this work we probed mutants on surfaces functionalized

with bacterial polysaccharides. In their functional context, bacteriophage TSPs are indispensable tail parts make the mature phage an efficient, multivalent particle for adsorption to a bacterial surface to start infection.[23,62] Surface immobilized polysaccharides as in the SPR set-up in this work resemble this situation during phage infection. Similarly, with eukaryotic viruses, attachment studies were performed on glycosaminoglycan surfaces.[63] Compared to dissociation constants obtained at single ligand binding sites in a solution set-up, avidity effects occur upon multivalent protein binding to multivalent ligand surfaces. This results in notable decrease of dissociation constants, more than two orders of magnitude have been reported in surface plasmon resonance set-ups.[64] Hence, this prominent binding signal amplification is suitable for detecting small affinity changes at single sites within a multivalent system as analysed in this work.

For Sf6TSP, SPR signals obtained upon interactions with surface-immobilized O-antigen polysaccharide were best described by a bivalent heterogeneous binding model with two dissociation constants of ~20 nM and ~6 µM. Furthermore, a curved Scatchard plot was obtained as it has been typically described for multivalent binding interactions.[65,66] However, the rather simple, solution-based binding model we employed to fit our SPR data in this work neglects additional effects that occur during surface association, like geometrical and ligand surface density parameters. Models that are more sophisticated therefore should be used that take into account the reaction volume and a probability factor for the multivalent binding event.[67] Considering these parameters results in a set of differential equations, in which also heterogeneous ligands can be included.[68] For example, this model adjustments well described binding of the trivalent S-layer protein SbsB to the surface-immobilized *Geobacillus stearothermophilus* secondary cell wall polymer.[69]

In conclusion, we employed an interdisciplinary approach of NMR, MD simulations and SPR interaction measurements for the assessment of subtle affinity differences in an elongated carbohydrate-binding site. Importantly, multivalent binding observed for Sf6TSP on the activated SfY polysaccharide surface resembles the interactions of Sf6 bacteriophage with LPS covered *S. flexneri* surfaces. Here, the typical bimodal O-polysaccharide chain length distribution found in *S. flexneri* LPS results in a heterogeneous glycan ligand surface.[70] Similar SPR-based studies with carbohydrate-binding proteins have been described.[65,71–73] This emphasizes that multivalence effects are often necessary to provide the necessary amplification of specific protein-carbohydrate interactions in their functional context. Our work stresses that in a similar manner, these effects can be exploited to explore affinity fine tuning for protein-carbohydrate complexes.

Methods

Materials and Chemicals

All chemicals were of analytical grade, and ultrapure water was used throughout. Lipopolysaccharide of *S. flexneri* Y was a gift from Nils Carlin (Scandinavian Biopharma, Solna, Sweden). O-polysaccharide was obtained from LPS by acid hydrolysis as described.[74] Cloning and purification of Sf6TSP mutants have been described.[9,29]

Molecular dynamics simulations

Structures were parameterized with the AMBER03 force field for proteins[75] and the GLYCAM06 force field (v06j-1) for glycans.[76] Simulations were based on two subunits with one binding site cleft of the inactive mutant Sf6TSPΔN E366A D399A (residues 109-622, pdbID: 4URR) and on six cysteine mutants thereof in complex with octasaccharides of *Shigella flexneri* O-polysaccharide serogroup Y (Supplementary Table S4).[20] Parameters for N340C-NBD conjugates were generated with the ANTECHAMBER package using the AMBER and GAFF force field (v1.7) and a mixture for side chain residue (amber) and fluorescent label (gaff) (chimeric).[77,78] The simulation complex was placed in an orthogonal box with the dimensions ~140x100x120 $Å^3$. Simulations were run with the TIP3P water model[79] and charges were equalized with 18 sodium ions.

MD simulations were carried out using the GROMACS4.5.5 program package.[80–82] After energy minimization to a maximum force smaller than 1000 kJ/mol/nm (steep descent) the system was equilibrated in two simulations with 100 ps each. All simulations were run for 1 ns under isothermal-isobaric (NPT) ensemble conditions with Parrinello-Rahman barostat coupling (reference pressure of 1.0 bar and coupling time constant of 0.5 ps)[83] and a Langevin thermostat (reference temperature 298 K and coupling time constant 1.0 ps). Simulation time steps were 2 fs. Hydrogen bonds and protein backbone were constrained except for loops and turns using the LINCS algorithm in GROMACS with 1000 kJ mol^{-1} $Å^{-2}$.[84] All systems were simulated for 100 ns.

Binding site residues were defined as all amino acids in a 5 Å distance of octasaccharide ligand in Sf6TSP E366A D399A (pdbID: 4URR). The tool *g_rmsf* was used to calculate fluctuations of the carbohydrate ligand. The tool for clustering structures *g_cluster* by GROMACS4.6.4 was used to derive ligand clusters. Water positional analysis was performed using *MobyWat*[51,52] as described previously.[27] Hydrogen bond occupancies were scaled between individual complexes by summing up all hydrogen bond occupancies above a threshold of 10 % and by normalizing this sum to 1.0 for the Sf6TSP reference.

NMR spectroscopy

NMR experiments were performed at 56 °C on a 500 MHz Bruker Avance NMR spectrometer equipped with a TCI Z-Gradient cryoprobe using a sample containing Sf6TSP D399N (0.12 mM) and the

octasaccharide ligand (1.87 mM) in D_2O sodium phosphate buffer (100 mM, pD=7) as described earlier.[20] Besides the previously recorded 2D $^1H,^1H$ transfer NOESY spectrum with a mixing time of 120 ms, additional experiments were carried out with mixing times of 40, 60 and 100 ms. The cross-peak volumes were integrated and normalized by the average of calculated auto-peak volumes at t=0 of resonances RAM1-H1, RAM3-H4, NAG2-H3, RAM5-H1, NAG6-H1, NAG6-H3 and RAM7-H4.[20] The data were used to construct NOE build-up curves, from which proton-proton cross-relaxation rates were extracted as the slope of a second order polynomial fit at t=0.[85] Due to differences in effective correlation times of protons in the protein-bound oligosaccharide, namely, between the ones in fast spinning methyl groups[36] and those from methine protons, two different reference distances were used in the analysis relying on the isolated spin-pair approximation:[86] 2.59 Å of Me6-H5 of residue RAM1 and 2.49 Å H1-H2 of residue RAM1. Like this, the experimentally derived proton-proton distances in the bound oligosaccharide were obtained.

For construction of trNOE-derived distance maps, an octasaccharide molecular model retrieved from the Sf6TSP wild type MD simulations with AMBER force field was used as a template for calculating atom-atom distances.[20] Molecular models covering the full glycosidic conformational space were generated in Vega ZZ (release 2.3.1.2)[87] by scanning the torsional angles ϕ and ψ in 10° intervals. Additional models were generated for relevant methyl groups for which, in addition to ϕ and ψ, also the torsion angle related to the methyl group, centred at the C5-C6 bond in RAM or the CO-CH_3 bond in NAG, was rotated for a total of 120° in 10° increments. Relevant atom-atom distances for each conformation in the trajectories were extracted using VMD 1.9.1. The trNOE-derived maps were then generated in MATLAB (R2012a, Mathworks) using the atom-atom distances as input. For methyl groups, the internuclear distances of the three methyl protons were averaged according to r^{-6} with respect to each other and to the methyl bond rotation, thus giving a single effective distance at a given ϕ and ψ torsion angle. The atom-atom matrices were used to calculate theoretical trNOE, by assuming ISPA[86] and using a RAM1-H1 – RAM1-H2 distance of 2.49 Å and a RAM1-H5 – RAM1-Me distance of 2.59 Å, for methine-methine and methyl-methine atom pairs, respectively. The distance maps were generated by plotting the contours for which calculated trNOE±10% is equal to its experimental counterpart on a 2D grid.

Surface plasmon resonance

Prior to surface immobilization, O-polysaccharide preparations were oxidized with sodium periodate. *Shigella flexneri* Y polysaccharide was prepared as stock solutions of 10 mg ml^{-1} in water and diluted in 1:10 10 mM sodium phosphate pH 6.2 and 10 mM sodium periodate for oxidation for 30 min at 25 °C.[38,88] Ethylene glycol was added to 20 % (v/v) final concentration to stop the oxidation process and

oxidized polysaccharide was purified by a desalting column (PD10, GE Healthcare, Freiburg, Germany), concentrated by ultrafiltration (Amicon 4k, Millipore, Darmstadt, Germany) and stored at -20 °C.

All SPR experiments were run in a Reichert SPR 7500 DC (Reichert, Buffalo, NY, USA) at 25 °C with all solutions filtered (0.45 µm) and degassed. Hydrazide activation of carboxymethyl dextran surfaces (CMD200D, Xantec, Düsseldorf, Germany) was performed as described previously.[38] Oxidized polysaccharide (10 mg ml^{-1} in water) was injected on the hydrazide surface for 20 min at 4 µl min^{-1}, washed with 10 mM sodium phosphate pH 6.0 for 7 min and reduced with 50 mM sodium cyanoborhydride (0.1 M in acetate buffer pH 4.0) for 20 min.

Interaction experiments were performed at 20 µl min^{-1} in 50 mM sodium phosphate pH 7.0. Protein samples were injected for 3 min and dissociation was monitored for 5 min. The surface was regenerated by a 4 min injection of 100 mM sodium acetate pH 4.0. Data were processed with the program TraceDrawer1.7 (Reichert) and association and dissociation rate constants were fitted based on models for bivalent equilibrium and heterogeneous binding models (Supplementary Figure S7).

References

1. Pokala, N. & Handel, T. M. Review: Protein design—where we were, where we are, where we're going. *J. Struct. Biol.* **134**, 269–281 (2001).
2. Höcker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, A. & Sterner, R. Dissection of a $(\beta\alpha)_8$-barrel enzyme into two folded halves. *Nat. Struct. Mol. Biol.* **8**, 32–36 (2001).
3. *Directed Evolution Library Creation: Methods and Protocols*. (Humana Press, 2003).
4. Marshall, S. A., Lazar, G. A., Chirino, A. J. & Desjarlais, J. R. Rational design and engineering of therapeutic proteins. *Drug Discov. Today* **8**, 212–221 (2003).
5. Rohl, C. A., Strauss, C. E. M., Misura, K. M. S. & Baker, D. Protein structure prediction using ROSETTA. in *Methods in Enzymology* vol. 383 66–93 (Academic Press, 2004).
6. Kuhlman, B. Designing protein structures and complexes with the molecular modelling program Rosetta. *J. Biol. Chem.* jbc.AW119.008144 (2019) doi:10.1074/jbc.AW119.008144.
7. Pearce, R., Huang, X., Setiawan, D. & Zhang, Y. EvoDesign: designing protein–protein binding interactions using evolutionary interface profiles in conjunction with an optimized physical energy function. *J. Mol. Biol.* **431**, 2467–2476 (2019).
8. Nimrod, G. *et al.* Computational design of epitope-specific functional antibodies. *Cell Rep.* **25**, 2121-2131.e5 (2018).
9. Kunstmann, S. *et al.* Bacteriophage Sf6 tailspike protein for detection of *Shigella flexneri* pathogens. *Viruses* **10**, 431 (2018).
10. Marchetti, R. *et al. Burkholderia pseudomallei* capsular polysaccharide recognition by a monoclonal antibody reveals key details toward a biodefense vaccine and diagnostics against melioidosis. *ACS Chemical Biology* **10**, 2295–2302 (2015).
11. Fiebig, T. *et al.* Efficient solid-phase synthesis of meningococcal capsular oligosaccharides enables simple and fast chemoenzymatic vaccine production. *J. Biol. Chem.* **293**, 953–962 (2018).
12. Keogh, D. *et al.* Generating novel recombinant prokaryotic lectins with altered carbohydrate binding properties through mutagenesis of the PA-IL protein from *Pseudomonas aeruginosa*. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1840**, 2091–2104 (2014).

13. Mao, S. *et al.* Phage-display library selection of high-affinity human single-chain antibodies to tumor-associated carbohydrate antigens sialyl LewisX and LewisX. *Proc. Natl. Acad. Sci.* **96**, 6953–6958 (1999).
14. Schoonbroodt, S. *et al.* Engineering antibody heavy chain CDR3 to create a phage display Fab library rich in antibodies that bind charged carbohydrates. *J. Immunol.* **181**, 6213–6221 (2008).
15. Morin, A. *et al.* Computational design of an endo-1,4-β-xylanase ligand binding site. *Protein Eng. Des. Sel.* **24**, 503–516 (2011).
16. Woods, R. J. Predicting the structures of glycans, glycoproteins, and their complexes. *Chem. Rev.* **118**, 8005–8024 (2018).
17. Meiler, J. & Baker, D. ROSETTALIGAND: protein–small molecule docking with full side-chain flexibility. *Proteins Struct. Funct. Bioinforma.* **65**, 538–548 (2006).
18. Kilambi, K. P. *et al.* Extending RosettaDock with water, sugar, and pH for prediction of complex structures and affinities for CAPRI rounds 20-27: RosettaDock for CAPRI rounds 20-27. *Proteins Struct. Funct. Bioinforma.* **81**, 2201–2209 (2013).
19. *Computational Drug Discovery and Design*. vol. 819 (Springer New York, 2012).
20. Kang, Y. *et al.* Bacteriophage tailspikes and bacterial O-antigens as a model system to study weak-affinity protein–polysaccharide interactions. *J. Am. Chem. Soc.* **138**, 9109–9118 (2016).
21. Kang, Y., Barbirz, S., Lipowsky, R. & Santer, M. Conformational diversity of O-antigen polysaccharides of the Gram-negative bacterium *Shigella flexneri* serotype Y. *The Journal of Physical Chemistry B* **118**, 2523–2534 (2014).
22. Broeker, N. K. *et al.* Complex carbohydrate recognition by proteins: fundamental insights from bacteriophage cell adhesion systems. *Perspect. Sci.* **11**, 45–52 (2017).
23. Broeker, N. K. *et al.* Time-resolved DNA release from an O-antigen–specific *Salmonella* bacteriophage with a contractile tail. *J. Biol. Chem.* **294**, 11751–11761 (2019).
24. Andres, D., Baxa, U., Hanke, C., Seckler, R. & Barbirz, S. Carbohydrate binding of *Salmonella* phage P22 tailspike protein and its role during host cell infection. *Biochemical Society Transactions* **38**, 1386–1389 (2010).
25. Lindberg, A. A. Bacteriophage receptors. *Annu. Rev. Microbiol.* **27**, 205–241 (1973).
26. Lindberg, A. A., Wollin, R., Gemski, P. & Wohlhieter, J. A. Interaction between bacteriophage Sf6 and *Shigella flexneri*. *J. Virol.* **1**, 38 (1978).
27. Kunstmann, S. *et al.* Solvent networks tune thermodynamics of oligosaccharide complex formation in an extended protein binding site. *J. Am. Chem. Soc.* **140**, 10447–10455 (2018).
28. Müller, J. J. *et al.* An intersubunit active site between supercoiled parallel β-helices in the trimeric tailspike endorhamnosidase of *Shigella flexneri* phage Sf6. *Structure* **16**, 766–775 (2008).
29. Freiberg, A. *et al.* The tailspike protein of *Shigella* phage Sf6: A structural homolog of *Salmonella* phage P22 tailspike protein without sequence similarity in the β-helix domain. *Journal of Biological Chemistry* **278**, 1542–1548 (2003).
30. Barbirz, S., Becker, M., Freiberg, A. & Seckler, R. Phage tailspike proteins with β-solenoid fold as thermostable carbohydrate binding materials. *Macromol. Biosci.* **9**, 169–173 (2009).
31. Chang, Z. *et al.* The changing epidemiology of bacillary dysentery and characteristics of antimicrobial resistance of *Shigella* isolated in China from 2004–2014. *BMC Infectious Diseases* **16**, (2016).
32. Kotloff, K. L. *et al.* Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *The Lancet* **382**, 209–222 (2013).
33. Chen, W. H. & Kotloff, K. L. *Shigella* vaccine development: finding the path of least resistance. *Clinical and Vaccine Immunology* **23**, 904–907 (2016).
34. Schmidt, A., Rabsch, W., Broeker, N. K. & Barbirz, S. Bacteriophage tailspike protein based assay to monitor phase variable glucosylations in *Salmonella* O-antigens. *BMC Microbiology* **16**, (2016).
35. Galochkina, T. *et al.* Conformational dynamics of the single lipopolysaccharide O-antigen in solution. *ChemPhysChem* **17**, 2839–2853 (2016).

36. Widmalm, G., Pastor, R. W. & Bull, T. E. Molecular dynamics simulation of methyl group relaxation in water. *J. Chem. Phys.* **94**, 4097–4098 (1991).
37. Lüderitz, O., Westphal, O., Staub, A. M. & Nikaido, H. Chapter 4 - Isolation and chemical and immunological characterization of bacterial lipopolysaccharides. in *Bacterial Endotoxins* 145–233 (Academic Press, 1971).
38. Gedig, E. T. Surface chemistry in SPR technology. *Handb. Surf. Plasmon Reson.* 173–220 (2008).
39. Copeland, R. A. Protein-ligand binding equilibria. *Enzym. Pract. Introd. Struct. Mech. Data Anal.* 76–108 (2000).
40. Schuck, P. & Minton, A. P. Kinetic analysis of biosensor data: elementary tests for self-consistency. *TIBS* **21**, 458–460 (1996).
41. Jung, M. *et al.* Affinity map of bromodomain protein 4 (BRD4) interactions with the histone H4 tail and the small molecule inhibitor JQ1. *J. Biol. Chem.* **289**, 9304–9319 (2014).
42. Bellich, B. *et al.* Influence of bacterial biofilm polysaccharide structure on interactions with antimicrobial peptides: a study on *Klebsiella pneumoniae*. *Int. J. Mol. Sci.* **19**, (2018).
43. Kulharia, M., Bridgett, S. J., Goody, R. S. & Jackson, R. M. InCa-SiteFinder: a method for structure-based prediction of inositol and carbohydrate binding sites on proteins. *J. Mol. Graph. Model.* **28**, 297–303 (2009).
44. Taroni, C., Jones, S. & Thornton, J. M. Analysis and prediction of carbohydrate binding sites. *Protein Eng.* **13**, 89–98 (2000).
45. Pal, D. & Chakrabarti, P. Different types of interactions involving cysteine sulfhydryl group in proteins. *J. Biomol. Struct. Dyn.* **15**, 1059–1072 (1998).
46. Sharabi, O., Erijman, A. & Shifman, J. M. Computational methods for controlling binding specificity. in *Methods in Enzymology* vol. 523 41–59 (Elsevier, 2013).
47. Lis, H. & Sharon, N. Lectins: carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.* **98**, 637–674 (1998).
48. Shinohara, Y., Hasegawa, Y., Kaku, H. & Shibuya, N. Elucidation of the mechanism enhancing the avidity of lectin with oligosaccharides on the solid phase surface. *Glycobiology* **7**, 1201–1208 (1997).
49. Gether, U., Lin, S. & Kobilka, B. K. Fluorescent labelling of purified β2 adrenergic receptor evidence for ligand-specific conformational changes. *J. Biol. Chem.* **270**, 28268–28275 (1995).
50. Ferrero, V. E. V., Di Nardo, G., Catucci, G., Sadeghi, S. J. & Gilardi, G. Fluorescence detection of ligand binding to labelled cytochrome P450BM3. *Dalton Trans* **41**, 2018–2025 (2012).
51. Jeszenői, N., Bálint, M., Horváth, I., van der Spoel, D. & Hetényi, C. Exploration of interfacial hydration networks of target–ligand complexes. *J. Chem. Inf. Model.* **56**, 148–158 (2016).
52. Jeszenői, N., Horváth, I., Bálint, M., van der Spoel, D. & Hetényi, C. Mobility-based prediction of hydration structures of protein surfaces. *Bioinformatics* **31**, 1959–1965 (2015).
53. Taylor, N. R. Small world network strategies for studying protein structures and binding. *Comput. Struct. Biotechnol. J.* **5**, 1–7 (2013).
54. Ladburry, J. E. Just add water! The effect of water on the specificity of protein ligand binding sites and its potential application to drug design. *Chem. Biol.* 973–980 (1996).
55. Barbirz, S. *et al.* Crystal structure of *Escherichia coli* phage HK620 tailspike: podoviral tailspike endoglycosidase modules are evolutionarily related. *Molecular Microbiology* **69**, 303–316 (2008).
56. Ahmad, M., Helms, V., Kalinina, O. V. & Lengauer, T. The role of conformational changes in molecular recognition. *J. Phys. Chem. B* **120**, 2138–2144 (2016).
57. Gapsys, V. & de Groot, B. L. pmx Webserver: a user friendly interface for alchemistry. *J. Chem. Inf. Model.* **57**, 109–114 (2017).

58. Kutzner, C. *et al.* Best bang for your buck: GPU nodes for GROMACS biomolecular simulations. *J. Comput. Chem.* **36**, 1990–2008 (2015).
59. Nivedha, A. K., Thieker, D. F., Makeneni, S., Hu, H. & Woods, R. J. Vina-Carb: improving glycosidic angles during carbohydrate docking. *J. Chem. Theory Comput.* **12**, 892–901 (2016).
60. Labonte, J. W., Adolf-Bryfogle, J., Schief, W. R. & Gray, J. J. Residue-centric modelling and design of saccharide and glycoconjugate structures. *J. Comput. Chem.* **38**, 276–287 (2017).
61. Li, L., Xu, W. & Lü, Q. Improving protein-ligand docking with flexible interfacial water molecules using SWRosettaLigand. *J. Mol. Model.* **21**, (2015).
62. Broeker, N. K. & Barbirz, S. Not a barrier but a key: How bacteriophages exploit host's O-antigen as an essential receptor to initiate infection. *Mol. Microbiol.* **105**, 353–357 (2017).
63. Peerboom, N. *et al.* Binding kinetics and lateral mobility of HSV-1 on end-grafted sulfated glycosaminoglycans. *Biophys. J.* **113**, 1223–1234 (2017).
64. Lundquist, J. J. & Toone, E. J. The cluster glycoside effect. *Chem. Rev.* **102**, 555–578 (2002).
65. Häyrinen, J. *et al.* High affinity binding of long-chain polysialic acid to antibody, and modulation by divalent cations and polyamines. *Mol. Immunol.* **39**, 399–411 (2002).
66. MacKenzie, C. R. *et al.* Analysis by surface plasmon resonance of the influence of valence on the ligand binding affinity and kinetics of an anti-carbohydrate antibody. *J. Biol. Chem.* **271**, 1527–1533 (1996).
67. Müller, K. M., Arndt, K. M. & Plückthun, A. model and simulation of multivalent binding to fixed ligands. *Anal. Biochem.* **261**, 149–158 (1998).
68. van Steeg, T. J., Bergmann, K. R., Dimasi, N., Sachsenmeier, K. F. & Agoram, B. The application of mathematical modelling to the design of bispecific monoclonal antibodies. *mAbs* **8**, 585–592 (2016).
69. Mader, C., Huber, C., Moll, D., Sleytr, U. B. & Sara, M. Interaction of the crystalline bacterial cell surface layer protein SbsB and the secondary cell wall polymer of *Geobacillus stearothermophilus* PV72 assessed by real-time surface plasmon resonance biosensor technology. *Journal of Bacteriology* **186**, 1758–1768 (2004).
70. Morona, R. Genetic modulation of *Shigella flexneri* 2a lipopolysaccharide O-antigen modal chain length reveals that it has been optimized for virulence. *Microbiology* **149**, 925–939 (2003).
71. Brogioni, B. & Berti, F. Surface plasmon resonance for the characterization of bacterial polysaccharide antigens: a review. *MedChemComm* **5**, 1058 (2014).
72. Houser, J. *et al.* Structural insights into *Aspergillus fumigatus* lectin specificity: AFL binding sites are functionally non-equivalent. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 442–453 (2015).
73. Nakajima, H. Kinetic analysis of binding between Shiga toxin and receptor glycolipid Gb3Cer by surface plasmon resonance. *J. Biol. Chem.* **276**, 42915–42922 (2001).
74. Haishima, Y., Holst, O. & Brade, H. Structural investigation on the lipopolysaccharide of *Escherichia coli* rough mutant F653 representing the R3 core type. *European Journal of Biochemistry* **203**, 127–134 (1992).
75. Duan, Y. *et al.* A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012 (2003).
76. Kirschner, K. N. *et al.* GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **29**, 622–655 (2008).
77. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260 (2006).
78. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

79. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1983).
80. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
81. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
82. Wehle, M. *et al.* Mechanical compressibility of the glycosylphosphatidylinositol (GPI) anchor backbone governed by independent glycosidic linkages. *J. Am. Chem. Soc.* **134**, 18964–18972 (2012).
83. Parrinello, M. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182 (1981).
84. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: a linear constraint solver for molecular simulations. *J Comp Chem* **18**, 1463–1472 (1997).
85. Widmalm, G., Byrd, R. A. & Egan, W. A conformation study of α-L-Rha*p*-(1→ 2)-α-L-Rha*p*-(1→ OMe) by NMR nuclear Overhauser effect spectroscopy (NOESY) and molecular dynamics calculations. *Carbohydrate research* **229**, 195–211 (1992).
86. Thomas, P. D., Basus, V. J. & James, T. L. Protein solution structure determination using distances from two-dimensional nuclear Overhauser effect experiments: effect of approximations on the accuracy of derived structures. *Proc. Natl. Acad. Sci.* **88**, 1237–1241 (1991).
87. Pedretti, A., Villa, L. & Vistoli, G. VEGA: a versatile program to convert, handle and visualize molecular structure on Windows-based PCs. *J. Mol. Graph. Model.* **21**, 47–49 (2002).
88. Wolfe, C. A. & Hage, D. S. Studies on the rate and control of antibody oxidation by periodate. *Anal. Biochem.* **231**, 123–130 (1995).

Acknowledgements

The authors acknowledge funding from the International Max Planck Research School on Multiscale Biosystems (S.K.), the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 713683 (COFUNDfellowsDTU) (S.K.), the Swedish Research Council (no. 2017-03703) (G.W.), the Knut and Alice Wallenberg Foundation (G.W.), and the Deutsche Forschungsgemeinschaft (BA4046/1-2) (S.B.). We thank Nils Carlin (Scandinavian Biopharma) for providing us with *S. flexneri* LPS preparations. We thank Mandy Schietke for excellent technical assistance.

Author contributions

S.B., M.S. and G.W. designed and planned the study. O.E. and G.W. performed and analysed the NMR experiments. S.K. performed and analysed SPR and MD experiments. M.W. produced the parameters for the labelled ligand simulations. All authors wrote and reviewed the manuscript.

Competing Interests

The authors declare no competing interests.

Supplementary information

Increasing the affinity of an O-antigen polysaccharide binding site in *Shigella flexneri* bacteriophage Sf6 tailspike protein

Sonja Kunstmann[1,2,†], Olof Engström[3], Marko Wehle[2], Göran Widmalm[3*], Mark Santer[2*], Stefanie Barbirz[1*]

1 Physikalische Biochemie, Universität Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany.

2 Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14476 Potsdam, Germany.

3 Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, 10691 Stockholm, Sweden.

† Present address: Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kgs. Lyngby, Denmark

* Corresponding authors: Tel: +49 331 567 9610, mark.santer@mpikg.mpg.de; Tel: +49 331 977 5322, barbirz@uni-potsdam.de; Tel: +46 8 16 37 42, goran.widmalm@su.se

Acknowledgements

The authors acknowledge funding from the International Max Planck Research School on Multiscale Biosystems (S.K.), the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 713683 (COFUNDfellowsDTU) (S.K.), the Swedish Research Council (no. 2017-03703) (G.W.), the Knut and Alice Wallenberg Foundation (G.W.), and the Deutsche Forschungsgemeinschaft (BA4046/1-2) (S.B.). We thank Nils Carlin (Scandinavian Biopharma) for providing us with *S. flexneri* LPS preparations. We thank Mandy Schietke for excellent technical assistance.

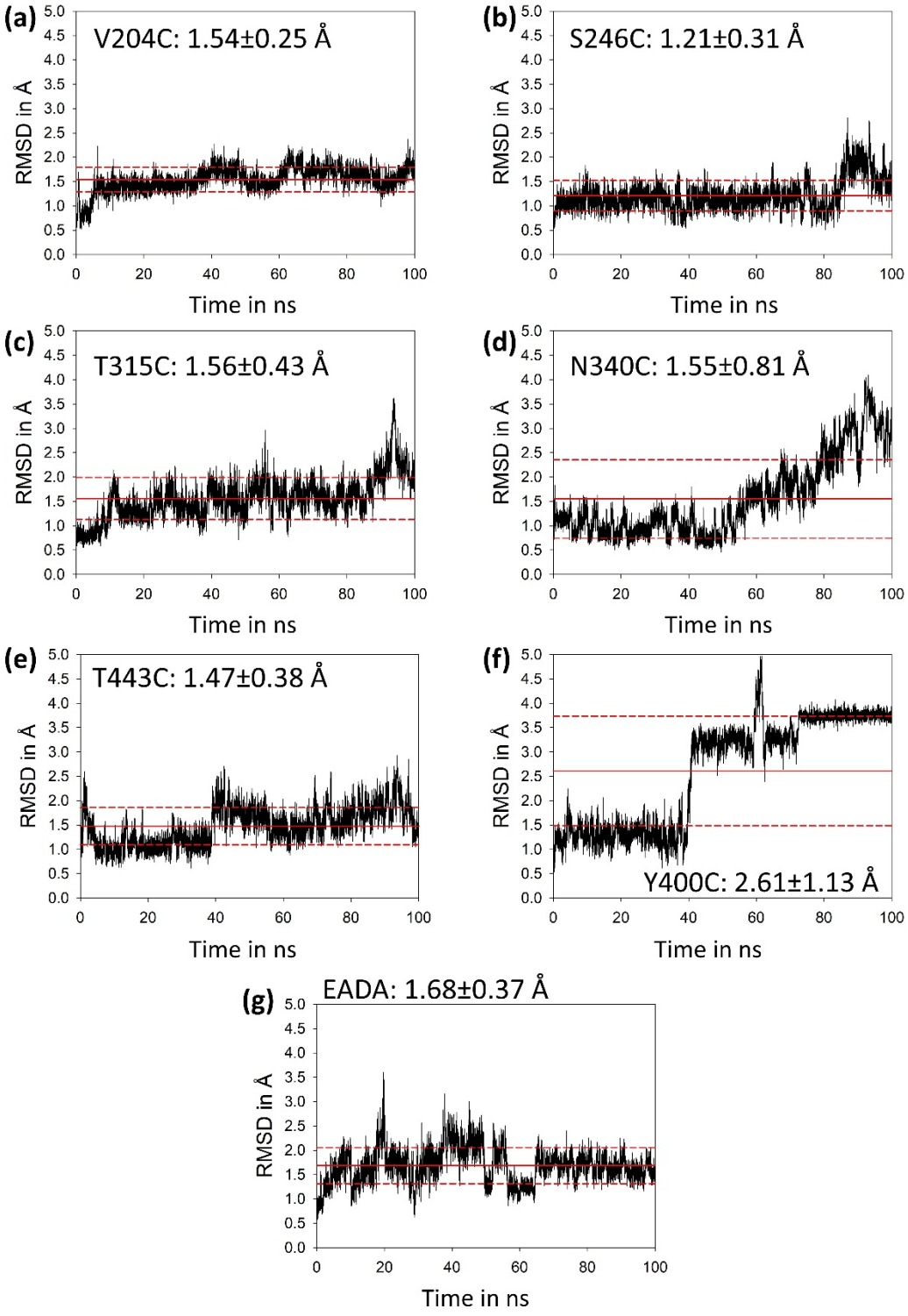
Author contributions

S.B., M.S. and G.W. designed and planned the study. O.E. and G.W. performed and analysed the NMR experiments. S.K. performed and analysed SPR and MD experiments. M.W. produced the parameters for the labelled ligand simulations. All authors wrote and reviewed the manuscript.

Competing Interests

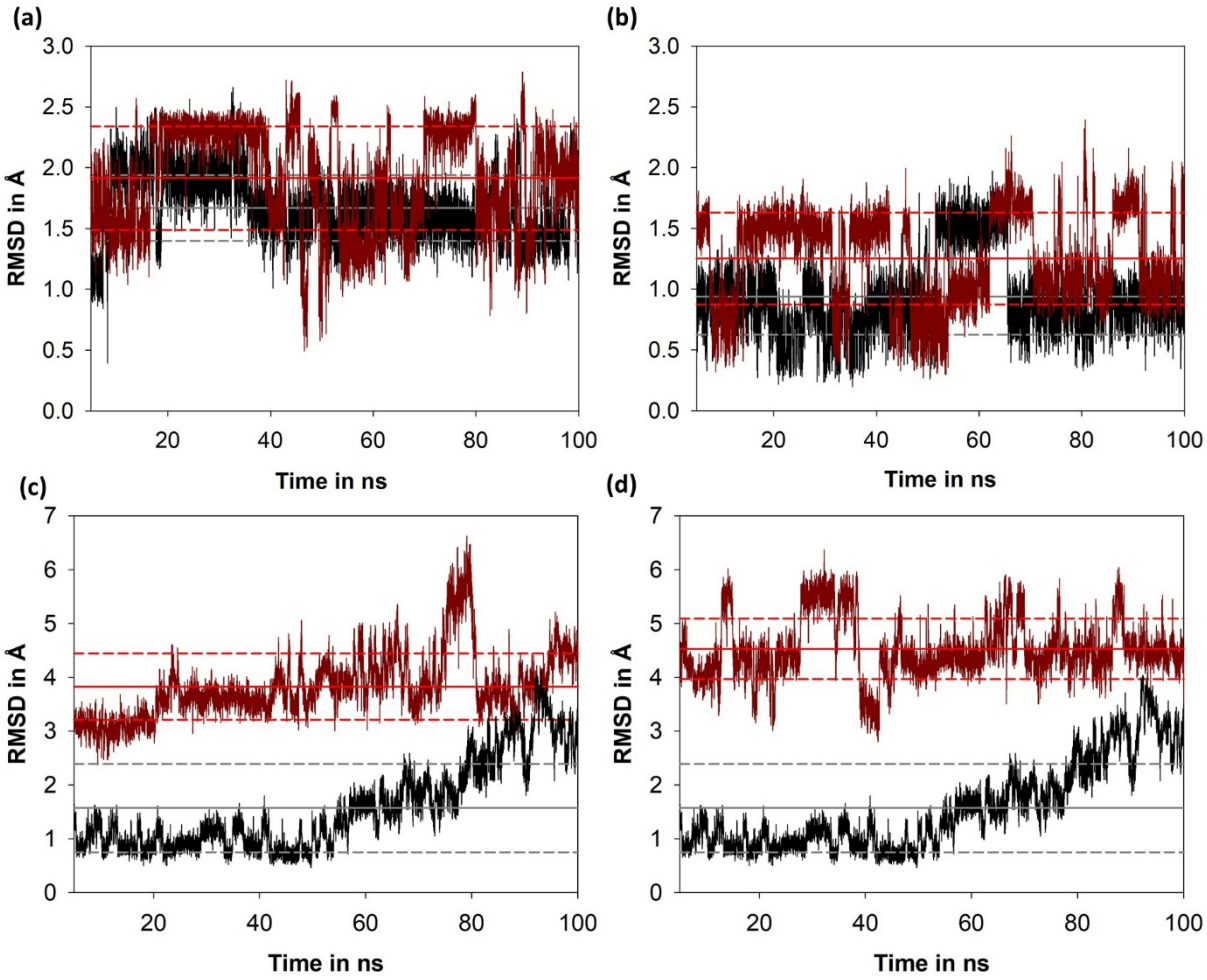
The authors declare no competing interests.

Figure S1: RMSD fluctuations over 100 ns MD simulations with octasaccharide ligand



Root mean square deviation of the octasaccharide ligand for cysteine mutants of Sf6TSP$_{EADA}$ with V204C **(a)**, S246C **(b)**, T315C **(c)**, N340C **(d)**, T443C **(e)**, Y400C **(f)** and EADA as reference **(g)**. The red solid line depicts the mean RMSD with the red dashed lines corresponding to one standard deviation. Numbers depict the mean of the RMSD with one standard deviation.

Figure S2: MD simulation fluctuations of NBD labels in different force fields



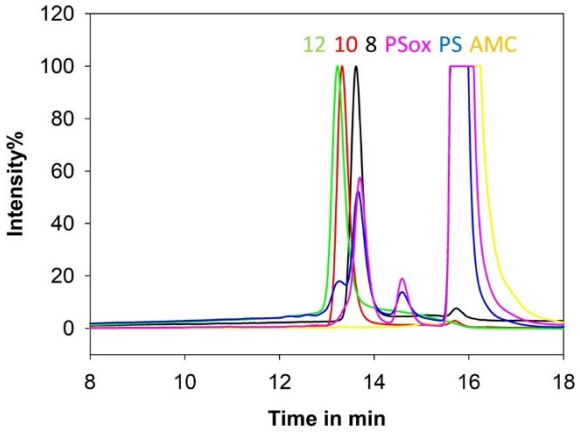
Root mean square displacement of the NBD label connected to Sf6TSP N340C were calculated from 100 ns simulation without ligand (black) and with ligand (red) with solid lines depicting the averaged RMSD and dashed lines as standard deviations with GAFF parameters **(a)** and with a chimeric version of AMBER and GAFF **(b)**. Root mean square displacement of the octasaccharide ligand in the Sf6TSP N340C binding site without label (black) or with label (red) calculated from 100 ns simulation with GAFF parameters **(c)** and with a chimeric version of AMBER and GAFF **(d)**.

Table S1: NMR cross-relaxation rates σ and effective proton-proton distances r_{ij} (Å) of TSP-octasaccharide complex Sf6TSP D399N.

Proton pair	σ	r_{ij}	MD Sf6TSP*wt* AMBER	MD Sf6TSP D399N AMBER	MD Sf6TSP*wt* CHARMM
RAM1-H1 - NAG2-H3	0.583	2.36	2.27	2.44	2.20
RAM1-H1 - RAM1-H2[a]	0.418	2.49	2.54	2.49	2.53
RAM1-H2 - NAG2-Ac	0.110	3.27	3.92	3.85	3.48
RAM1-H4 - RAM1-Me	0.359	2.69	2.93	2.70	2.92
RAM1-H5 - RAM1-Me[a]	0.451	2.59	2.58	2.59	2.63
NAG2-H1 - NAG2-H2	0.315	2.62	3.05	3.08	2.96
NAG2-H1 - NAG2-H3	0.313	2.62	2.93	2.83	2.52
NAG2-H1 - RAM3-H2	0.552	2.38	2.63	2.49	2.56
RAM3-H1 - RAM4-H2	0.375	2.54	2.35	2.41	2.26
RAM3-Me - RAM4-H1	0.159	3.08	3.87	3.66	3.24
RAM4-H1 - RAM4-Me	0.346	2.70	2.59	2.71	2.60
RAM4-H1 - RAM5-H3	0.864	2.21	2.38	2.21	2.56
RAM4-Me - RAM5-H2	0.302	2.77	3.20	3.41	4.08
RAM5-H1 - NAG6-Ac	0.097	3.34	3.78	3.44	3.78
RAM5-H1 - NAG6-H3	0.459	2.46	2.39	2.68	2.41
RAM5-H1 - RAM5-H2	0.450	2.46	2.69	2.66	2.46
RAM5-H4 - RAM5-Me	0.315	2.75	2.88	2.94	2.98
RAM5-H5 - RAM5-Me	0.412	2.63	2.54	2.60	2.65
NAG6-H1 - RAM7-H2	0.645	2.32	2.41	2.20	2.29
RAM7-H1 - RAM8-H2	0.287[b]	2.58	2.66	2.32	2.36
RAM7-Me - RAM8-H1	0.090[b]	3.28	4.37	3.70	4.19

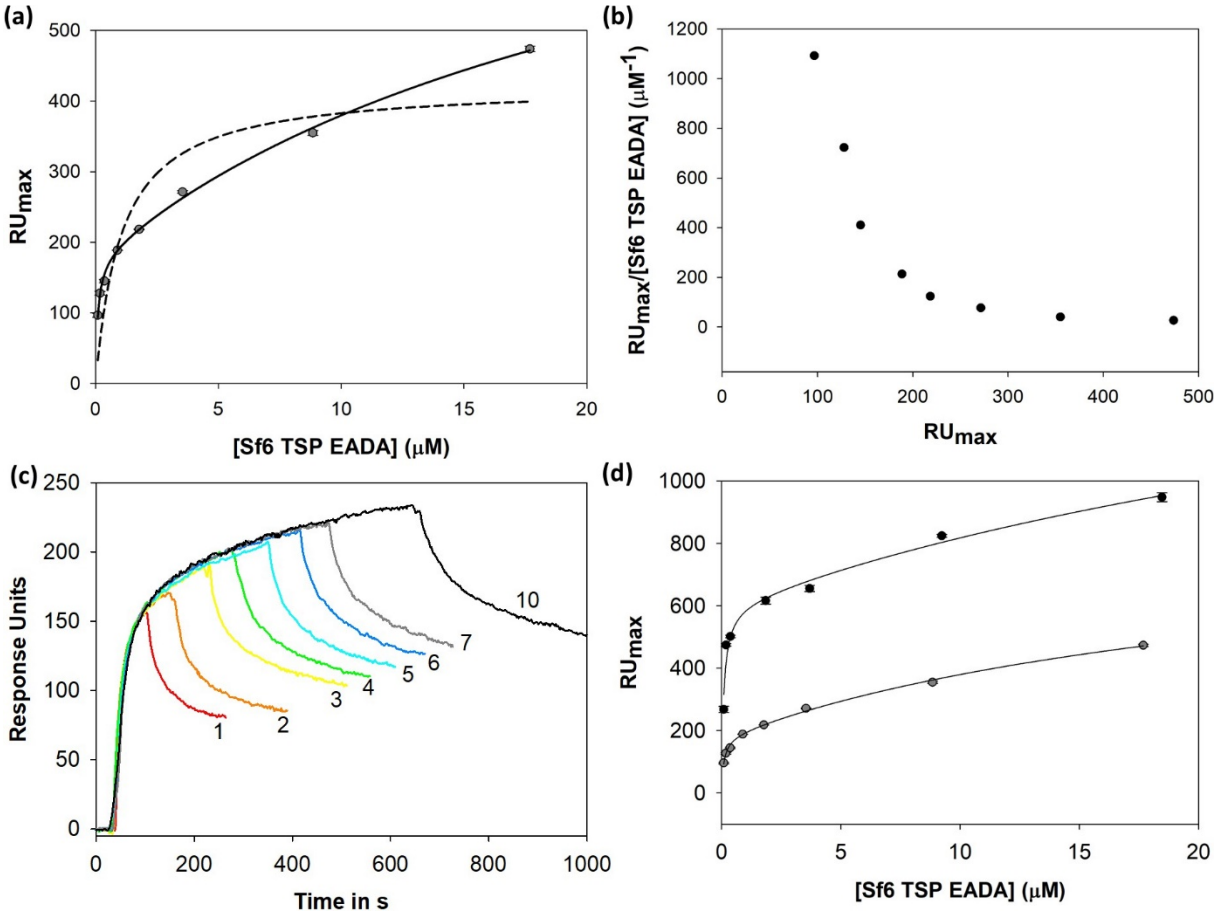
[a]Reference distances; [b]Measured from resonances of the α-anomeric form of RAM8 present to 83% in solution.

Figure S3: Analysis of oxidized polysaccharide cleavage products by reverse phase HPLC



In the reaction vial 1 mg PS was solved in 50 mM sodium phosphate buffer, pH 7 and 20 µg ml^{-1} Sf6TSP*wt* were added. The reaction was incubated overnight at room temperature. The samples were precipitated with 80 % ice-cold ethanol for 2 h at -40 °C and the supernatant was dried in a Speed Vac at 40 °C. Obtained oligosaccharides were labelled in 96 % methanol and 4 % acetic acid with a final concentration of 6 mM AMC (7-Amino-4-methylcumarin) for 1 h at 37 °C. For reduction 10 mM $NaCNBH_3$ in methanol was added and the reaction was incubated overnight. The samples were dried in a Speed Vac and oligosaccharides were extracted two times with 50 µL water and injected on the reverse phase column as 1:1000 dilutions.
Chromatographies were run on a Ultrasphere 5ODS (250 mm x 4.6 mm, Hi-Chrom) analytical column with a Ultrasphere 5ODS (45 mm x 4.6 mm, Hi-Chrom) guard column with a flow rate of 1 ml min^{-1}. Buffer A consisted of 0.1 % TFA (v/v) and buffer B of 0.1 % TFA (v/v) with 70 % acetonitrile (v/v). A total program of 35 min was run with a gradient from 10-40 % B over 20 min followed by a 5 min gradient from 40-100 % B and a reconditioning for 10 min from 100-10 % B.
Chromatograms were monitored with a RF-10A XL fluorescence detector with an excitation wavelength at 360 nm and emission wavelength at 450 nm (Response: 0.5 sec; Gain: x4; Sensitivity: Medium-High; Recorder Range: 1).

Figure S4: Surface plasmon resonance data and controls

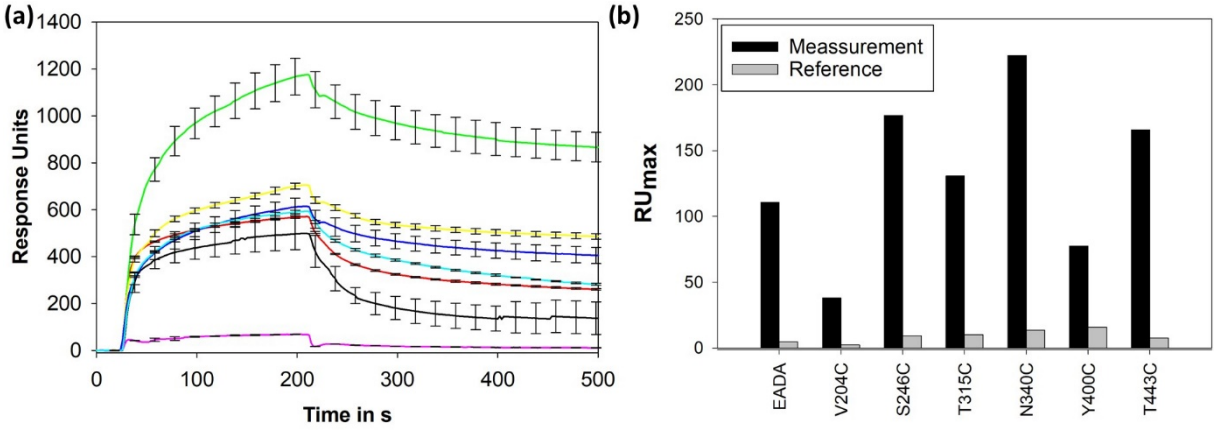


(a) Maximal response (RU_{max}) of the curves at different protein concentrations (grey dots) shown as mean of three measurements with error bars as one standard deviation. The maxima are fitted with 1:1 (dashed line) and 1:2 Langmuir binding isotherms (solid line). **(b)** Scatchard plot analysis of the data resulted in a curved, triphasic distribution. **(c)** Sensograms of 0.8 µM $Sf6TSP_{EADA}$ injected on immobilized SfY PS on carboxymethyldextran at different interaction times and flow rates: 1 min→50 µl/min; 2 min→35 µl/min; 3 min→20 µl/min; 4 min→17 µl/min; 5 min→15 µl/min; 6 min→11 µl/min; 7 min→10 µl/min and 10 min→7 µl/min. **(d)** Maximal responses as mean of triplicate measurements with standard deviations as error bars against the protein concentration from the first measurements (black circles, upper) and the last measurements (grey circles, lower) of $Sf6TSP_{EADA}$ on Chip1.

Table S2: Fitted maximal responses of surface plasmon resonance measurements

Exp. #	$RU_{max,1\ kin}$	$RU_{max,2\ kin}$	χ^2_{kin}	$RU_{max,1\ eq.}$	$RU_{max,2\ eq.}$	R^2_{eq}
1	303 ± 1	713 ± 2	2155	603±79	1252±2491	0.9782
2	176 ± 1	630 ± 0	382	211±28	632±30	0.9981

Figure S5: Surface Plasmon Resonance data from cysteine mutants



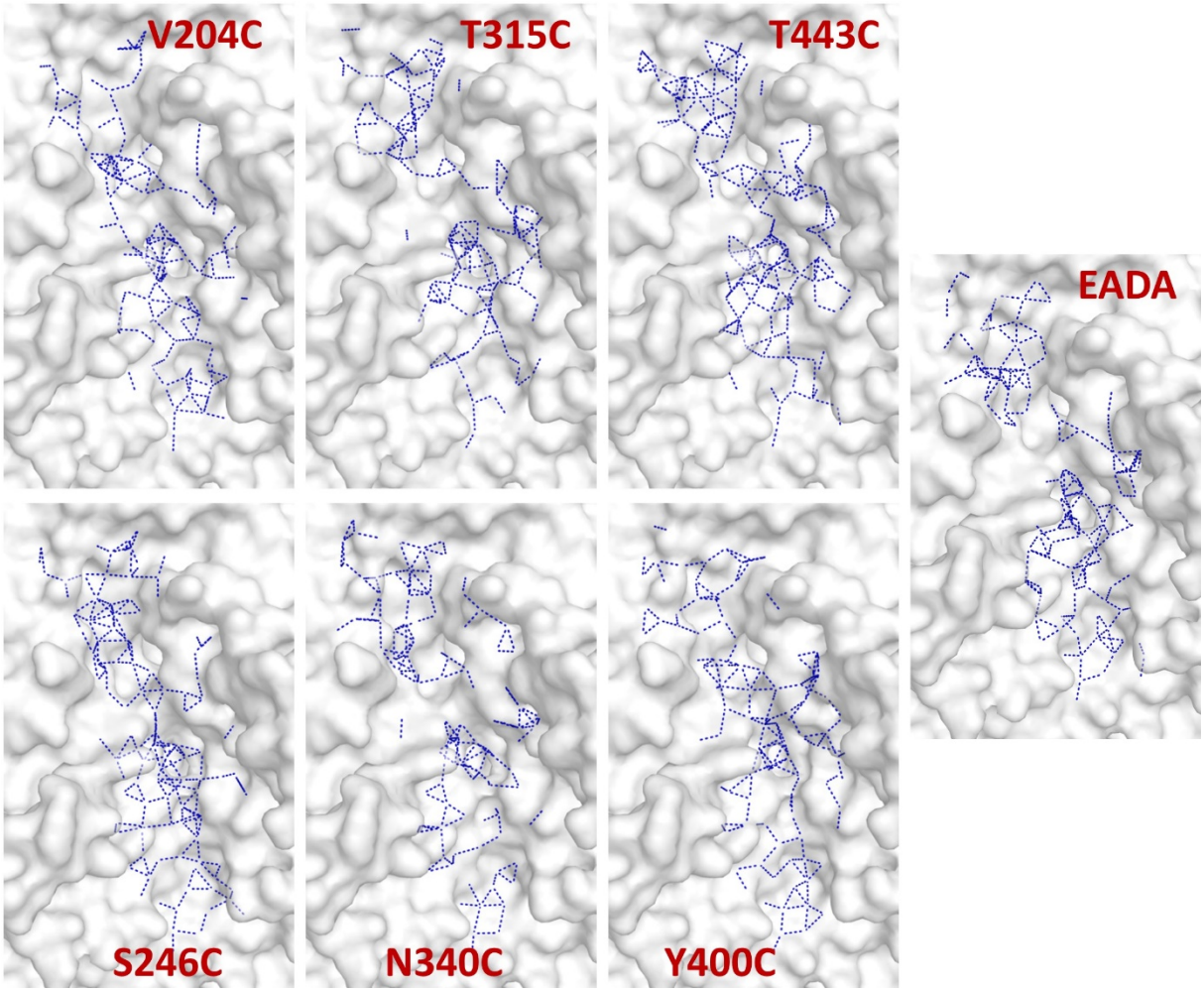
(a) Sensograms of 8 µM protein injections over the surface with immobilized SfY polysaccharide (Chip1) of Sf6TSP$_{EADA}$ (black) and each of the cysteine mutants: V204C (red), S246C (green), T315C (yellow), N340C (blue), Y400C (purple) and T443C (cyan). **(b)** The maximal responses measured at the end of 3 min. injections on the second chip are depicted for all Sf6TSP$_{EADA}$ cysteine mutants on the measurement channel (black) and the reference channel (grey). The unspecific signal on the reference channel ranges around 4 % for all protein variants except for Y400C, where the reference is increased to 20 %. However, the total signal on the measurement channel is lower for Y400C compared to most of the other cysteine mutants.

Table S3: Occupancies of hydrogens bonds between Sf6TSP$_{EADA}$ cysteine mutants and octasaccharide ligand

Protein	Ligand	EADA-octa	V204C-octa	S246C-octa[a]	T315C-octa	N340C-octa	Y400C-octa	T443C-octa
ALA203-O	RAM1-O2	15.27 %						15.87 %
ARG230-NH1	RAM1-O2	25.65 %	56.49 %	56.89 %	38.32 %	23.15 %	43.01 %	33.03 %
ARG230-NH2	RAM1-O2	16.67 %	10.18 %	16.37 %	22.46 %	35.13 %	17.56 %	13.97 %
ARG257-NH1	RAM1-O2	12.87 %			12.48 %			
GLU293-OE1	RAM1-O2		28.44 %	31.24 %			22.85 %	10.58 %
GLU293-OE2	RAM1-O2	28.04 %	67.56 %	23.25 %			67.07 %	
ARG230-NH1	RAM1-O3					20.26 %		
ARG230-NH2	RAM1-O3	14.07 %						
ARG257-NH2	RAM1-O3	13.27 %						11.48 %
ARG230-NH2	RAM1-O5	10.08 %	26.05 %		10.98 %		18.16 %	11.88 %
SER246-O	NAG2-O6				10.48 %		18.36 %	30.44 %
ASP247-OD1	RAM3-O3	29.74 %	89.32 %	39.42 %	31.04 %	42.61 %		22.36 %
ASP247-OD2	RAM3-O3	31.64 %		32.73 %	58.58 %	37.03 %	91.72 %	50.20 %
THR248-OG1	RAM3-O4	11.48 %	11.28 %	31.94 %	21.66 %	11.48 %		
SER246-O	RAM4-O3	73.65 %	69.06 %	71.46 %	45.71 %	49.00 %	46.21 %	65.47 %
ASP245-O	RAM4-O4						29.44 %	
GLN280-NE2	RAM4-O4					16.57 %		
GLN280-NE2	NAG6-O7	17.07 %	61.58 %	68.76 %	44.61 %	32.44 %	18.46 %	32.34 %
TYR282-OH	NAG6-O7				10.48 %			
SER337-OG	RAM7-O4						13.07 %	
GLY339-N	RAM7-O4						18.66 %	
GLN325-OE1	RAM7-O4	26.35 %						19.06 %
TRP421-NE1	RAM8-O4				19.06 %			13.57 %
ASP397-OD2	RAM8-O4			75.95 %		41.52 %		
ASP397-OD1	RAM8-O3			61.58 %		32.83 %		
TYR400-OH	RAM8-O3				11.38 %			
TYR400-OH	RAM8-O2		18.16 %					
ASP425-OD1	RAM8-O1						25.75 %	
ASP425-OD2	RAM8-O1					10.68 %	32.24 %	
Relative Occupancy		**1.00**	**1.34**	**1.56**	**1.03**	**1.08**	**1.42**	**1.01**
Total number of H-bonds		**14**	**10**	**11**	**13**	**12**	**14**	**13**

[a] Hydrogen bonds in this mutant which are listed as Ser246 are formed by Cys246.

Figure S6: Water networks of Sf6TSP$_{EADA}$ cysteine mutants



Water networks were derived from 100 ns simulation with *MobyWat* and the IDa list selection restricted to the top 20 % in the prediction mode. Resulting structures were superimposed with the crystal structure of Sf6TSP$_{EADA}$ and all water within 5 Å distance to the octasaccharide were selected. Networks are presented as dashed lines showing hydrogen bonds with maximal 3.5 Å distance.

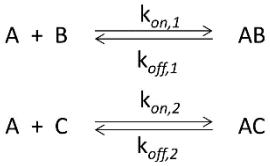
Table S4: Molecular dynamics simulations of Sf6TSP-oligossaccharide complexes described in this study

	EADA	***V204C***	***S246C***	***T315C***	***N340C***	***Y400C***	***T443C***
Ligand	Octa	Octa	Octa	Octa	Octa	Octa	Octa
No. atoms system	176007	176015	176004	176001	176004	175999	176001
No. atoms protein	15102	15092	15102	15096	15096	15082	15096
No. atoms ligand	177	177	177	177	177	177	177
No. water	53570	53576	53569	53570	53571	53574	53570

Octa: [⟶3)-α-L-Rha*p*-(1⟶3)-β-D-Glc*p*NAc-(1⟶2)-α-L-Rha*p*-(1⟶2)-α-L-Rha*p*(1⟶]$_2$

Figure S7: Binding models for SPR data evaluation

Heterogeneous binding:



$$\frac{d[AB]}{dt} = \left(k_{on,1}[A] - k_{off,1}\right)[AB]$$

$$\frac{d[AC]}{dt} = \left(k_{on,2}[A] - k_{off,2}\right)[AC]$$

$$Y = [AB] + [AC] \text{ (Recorded signal)}$$

Formulas are reproduced from TraceDrawer Handbook Vers1_6 (Reichert Technologies Life Sciences Buffalo, NY, USA).

Equilibrium binding:

$$RU_{max} = \frac{RU_{max,1}[L]}{K_{D1} + [L]} + \frac{RU_{max,2}[L]}{K_{D2} + [L]}$$