Chemical Biology

Increasing the Affinity of an O-Antigen Polysaccharide Binding Site in *Shigella flexneri* Bacteriophage Sf6 Tailspike Protein**

Sonja Kunstmann,[a, b, d] Olof Engström,[c] Marko Wehle,[b] Göran Widmalm,*[c] Mark Santer,*[b] and Stefanie Barbirz*[a]

Abstract: Broad and unspecific use of antibiotics accelerates spread of resistances. Sensitive and robust pathogen detection is thus important for a more targeted application. Bacteriophages contain a large repertoire of pathogen-binding proteins. These tailspike proteins (TSP) often bind surface glycans and represent a promising design platform for specific pathogen sensors. We analysed bacteriophage Sf6 TSP that recognizes the O-polysaccharide of dysentery-causing *Shigella flexneri* to develop variants with increased sensitivity for sensor applications. Ligand polyrhamnose backbone conformations were obtained from 2D ^{1}H,^{1}H-trNOESY NMR utilizing methine–methine and methine–methyl correlations. They agreed well with conformations obtained from molecular dynamics (MD), validating the method for further predictions. In a set of mutants, MD predicted ligand flexibilities that were in good correlation with binding strength as confirmed on immobilized *S. flexneri* O-polysaccharide (PS) with surface plasmon resonance. *In silico* approaches combined with rapid screening on PS surfaces hence provide valuable strategies for TSP-based pathogen sensor design.

Introduction

Glycan recognition is a ubiquitous theme in nature and intimately related to targeting, transport, adhesion and signalling processes both on an intracellular and extracellular level.[1] Most cellular envelopes are covered with glycan matrices and addressing these surfaces with tailor-made protein-binders has impact both in diagnostics and therapy.[2–5] This cell surface localization and involvement in reversible signalling events results in protein–carbohydrate complexes with relatively low affinities at single sites that gain functionality within a multivalent glycan–glycan binding environment.[6] For example, in lectins, typical carbohydrate binding proteins with high specificity in small sites, dissociation constants in the mM to µM range are found. This illustrates that the amphiphilic binding partners interact in an aqueous environment, often involving rather shallow protein surface grooves.[7] This is a clear difference to glycan binding sites found in antibodies that feature deep hydrophobic pockets and nM dissociation constants.[8]

Glycan binding proteins thus provide a rich pool of highly specific sites suitable for various applications, but their high-affinity engineering is difficult due to the various thermodynamic effects governing the complex formation.[9] In general, improvement of protein properties like stability, enzyme activity or binding affinity can be obtained by protein design and engineering techniques.[10] Scaffold design is used in de novo approaches to obtain proteins with defined geometries[11] or with altered binding specificities or enzymatic activities.[10] Experimental techniques rely on randomizing coding DNA sequences by error-prone PCR or directed evolution and subsequent selection procedures towards the desired protein property.[12] 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.[13–15] This strategy has been further extended by taking into account evolutionary profiles[16] or "re-epitoping" of antibodies and validation by crystal structure analysis.[17] For engineering of carbohydrate binding proteins, phage or plasmid display techniques have been employed.[18–20] Computational approaches also used ROSETTA,[21] but flexible ligands and water network distributions impede predictions via docking or Monte Carlo simulations with rotamer libraries. Algorithms like ROSETTALIGAND,[22] which are commonly used for other high-affinity design purposes, are in the process to be optimized for protein–carbohydrate interactions, that is, ROSETTADOCK.[23]

[a] Dr. S. Kunstmann, Dr. S. Barbirz
Physikalische Biochemie, Universität Potsdam
Karl-Liebknecht-Str. 24–25, 14476 Potsdam (Germany)
E-mail: barbirz@uni-potsdam.de

[b] Dr. S. Kunstmann, Dr. M. Wehle, Dr. M. Santer
Theory and Biosystems, Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14476 Potsdam (Germany)
E-mail: mark.santer@mpikg.mpg.de

[c] Dr. O. Engström, Prof. Dr. G. Widmalm
Department of Organic Chemistry, Arrhenius Laboratory
Stockholm University, 10691 Stockholm (Sweden)
E-mail: goran.widmalm@su.se

[d] Dr. S. Kunstmann
Current address: Department of Biotechnology and Biomedicine
Technical University of Denmark, Søltofts Plads
2800 Kgs. Lyngby (Denmark)

[**] A previous version of this manuscript has been deposited on a preprint server (https://doi.org/10.26434/chemrxiv.11410125.v1).

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/chem.202000495.

However, no common strategy exists for a rational selection of mutations for design of high-affinity carbohydrate-binding proteins. The inherent flexibility of the glycan ligand often has major impact on the energetics of protein–carbohydrate complex formation.[24–26] Flexibility in glycans can be functional to sample a large conformational space for binding sites and provide higher avidity in multivalent settings.[27] Although computationally more expensive, molecular dynamics simulations are thus well suited to model protein–glycan complexes because they can describe the flexibility of the glycan ligand.[28]

In this work, we analysed how flexibility of a glycan in a binding site links to its affinity of complex formation. As a proof-of-principle system, we chose a bacteriophage tailspike protein (TSP). TSPs act as recognition organelles for binding and orientation of bacteriophages during infection of the bacterial host.[29–31] TSPs have been also used as model systems to computationally link solvent network structures to experimental thermodynamic signatures of oligosaccharide ligand binding.[9] Bacteriophage Sf6 TSP specifically binds to the O-antigen of its host, *Shigella flexneri,* with the serogroup Y 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→]$_n$.[7,32] Sf6TSP is a highly stable trimeric protein with endorhamnosidase activity, producing oligosaccharides of mainly 2 RU, that is, octasaccharides, from the *S. flexneri* O-antigen polysaccharide.[7,32–34] Sf6TSP has three independent, elongated glycan binding sites for oligosaccharide O-antigen fragments produced by hydrolysis. Thorough glycan binding site description in Sf6TSP was achieved by a combination of MD simulations, X-ray crystallography and NMR spectroscopy (Figure 1).[7] Its N-terminal end anchors Sf6TSP to the phage tail and capsid, accordingly it can sample the O-antigen protruding perpendicularly from the bacterial cell surface.

Sf6TSP binds octasaccharide O-antigen fragments with relatively low affinity due to a highly flexible reducing end pointing towards the enzyme's active site residues (Figure 1). Sf6TSP therefore is a valuable target for generating mutants with increased O-antigen affinity, as it can be used as a sensor for *S. flexneri*,[2] an important diagnostic target pathogen that causes dysentery in infants.[35–37] In solution, fluorescence amplitude changes upon glycan ligand binding were used to detect O-polysaccharide binding via a cysteine coupled, environment-sensitive fluorescent label in the Sf6TSP binding site.[2] Screening for the correct fluorescent probe positioning in Sf6TSP's glycan binding site had yielded a series of cysteine point mutants. They were thus available for affinity studies in the present work (Figure 1).[2] A high-affinity Sf6TSP would be valuable in further improving its 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. Our study combines MD simulations with NMR data to confirm the conformational space sampled by a protein-bound oligosaccharide. Experimentally approved with surface plasmon resonance analyses, our test set shows that MD simulations can link mutations in a protein binding site to varying affinities towards a given glycan ligand on the surface of a pathogen.

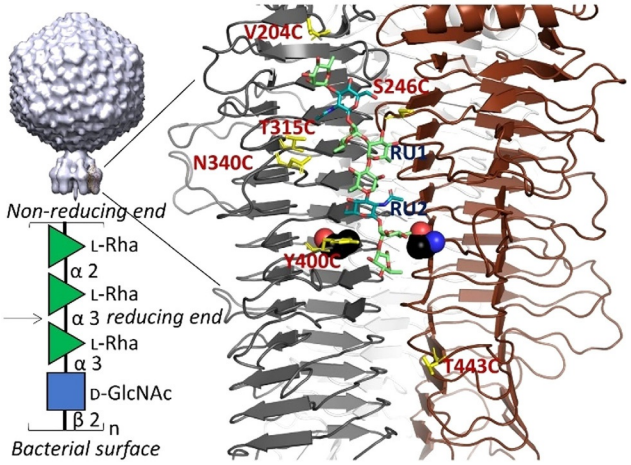

Figure 1. Sf6TSP glycan binding site in complex with octasaccharide and set of mutations. Bacteriophage Sf6 (left upper panel, EMDB: 1222) has six tailspike proteins (TSP) in its tail; a magnification of one inter-subunit binding site on a single, trimeric TSP is indicated. Repeat unit of the O-antigen polysaccharide ligand given in SNFG notation together with the Sf6TSP endorhamnosidase cleavage site and with respect to its orientation towards the bacterial surface. Three subunits of Sf6TSP E366A D399A (PDB ID: 4URR, grey/brown/silver backbone, cartoon, right panel) are illustrated with all residues chosen for a cysteine mutation shown as yellow sticks with the corresponding label in red. The catalytic residue positions are depicted in black spheres of D399A (grey chain) and E366A (brown chain). The protein is shown in complex with an octasaccharide (sticks in green for rhamnose and light blue for GlcNAc) as repeat unit 1 (RU1) and repeat unit 2 (RU2).

Results

Molecular dynamics simulations show varying glycan ligand flexibilities in different bacteriophage Sf6 tailspike protein mutants

Sf6TSP is a well-studied, weak-affinity carbohydrate binding protein as analysed with NMR spectroscopy, X-ray crystallography and computational analyses.[7] An Sf6TSP hydrolysis deficient variant with alanine exchanges for active site carboxylate residues (Sf6TSP E366A D399A, Sf6TSP$_{EADA}$) was used to construct a fluorescent probe sensitive for the O-antigen polysaccharide on *S. flexneri* bacteria.[2] A crystal structure of Sf6TSP$_{EADA}$ in complex with an O-antigen octasaccharide is available. Moreover, for construction of the *S. flexneri* sensor, single cysteine exchanges were introduced into Sf6TSP$_{EADA}$.[2] In this work, we used MD simulations to analyse the conformational space occupied by the protein-bound oligosaccharide ligand in order to make assumptions on the affinity to Sf6TSP in different mutational backgrounds. Five of the mutations (V204C, S246C, T315C, N340C and Y400C) lie in close proximity to the octasaccharide binding site, whereas the mutation T443C 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, that is, a modelled dodecasaccharide formed H-bond contacts with binding site residues at this proximal position.[7] 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. Details on the distinct protein–carbohydrate interactions of Sf6TSP with the octasaccharide ligand are provided in Figure S1a and Table S1.

To predict differences in binding affinity, octasaccharide conformations that occurred during the simulation were clustered to describe their ligand flexibility in the different mutational backgrounds. For the $Sf6TSP_{EADA}$ reference octasaccharide complex containing no other mutations, five major ligand conformer clusters were found (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 the different conformational behaviours in the protein binding site (Figure S2). 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 the mutations V204C and S246C, indicating increasing binding affinity. Simulations with mutants T315C and N340C also showed a decreased mean ligand RMSD, but with a higher standard deviation. This might be due to changes in the H-bond donor–acceptor equilibrium that is difficult to assign to an effect on affinity. The T443C mutant showed approximately the same mean RMSD and standard deviation as the $Sf6TSP_{EADA}$ reference.

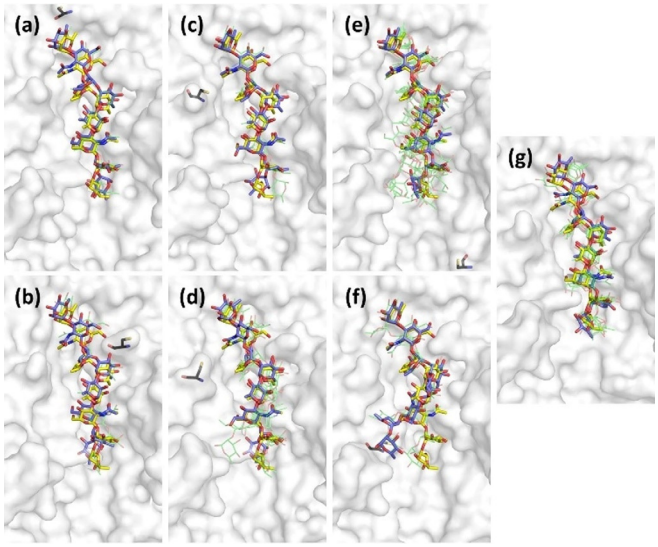

Figure 2. Octasaccharide ligands in the Sf6TSP binding site with flexibility analysis from 100 ns MD simulations. Typical conformational clusters obtained for oligosaccharide ligands are superimposed onto the ligand pose from crystal structure analysis (yellow) from a) V204C, b) S246C, c) T315C, d) N340C, 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.

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.

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 *S. flexneri* pathogens.[2] 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 (Figure S3). 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

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

Protein variant	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.

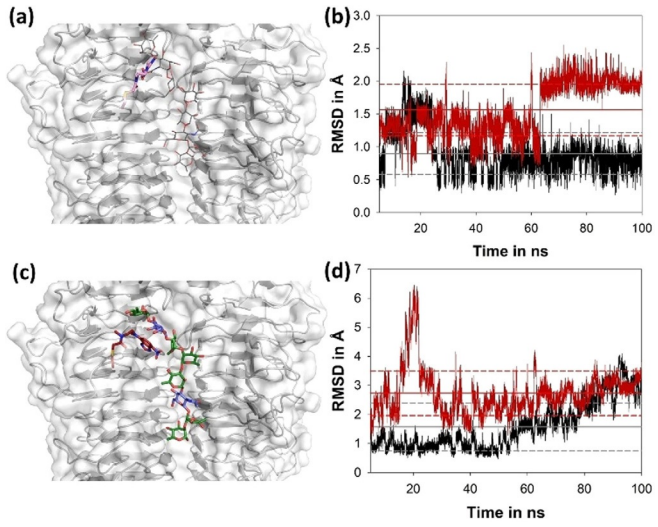

Figure 3. Sf6TSP N340C labelled with NBD simulated with octasaccharide ligand for 100 ns. a) Most abundant NBD label position (purple sticks) in a ligand free simulation. The label is buried between two protein subunits (white surface, grey cartoon) underneath the ligand binding site. An octasaccharide shown in grey thin sticks marks the ligand position in the crystal structure PDB ID 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.

(Figure 3 c, Figure S1b). Additionally, the octasaccharide ligand had a notably increased RMSD in the binding site compared to the complex with the unlabelled protein (Figure 3 d).

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

Interactions between oligo- or polysaccharides as ligands to proteins or antibodies can be investigated by trNOESY NMR experiments.[38] The *S. flexneri* Y octasaccharide contains α-L-rhamnosyl residues to a large extent, like for any of the O-antigens from various *S. flexneri* serotypes. The methyl groups at C6 of L-rhamnosyl residues thus enable an increased flexibility of the glycans compared to contributions of other, fully hydroxylated monosaccharide building blocks.[39,40] 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 (Figure S4);[41] 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 (Table S2) resulted in two radial ϕ,ψ-distributions with two intersection points. An overlay with glycosidic linkage conformations previously determined for the *S. flexneri* O-sero-

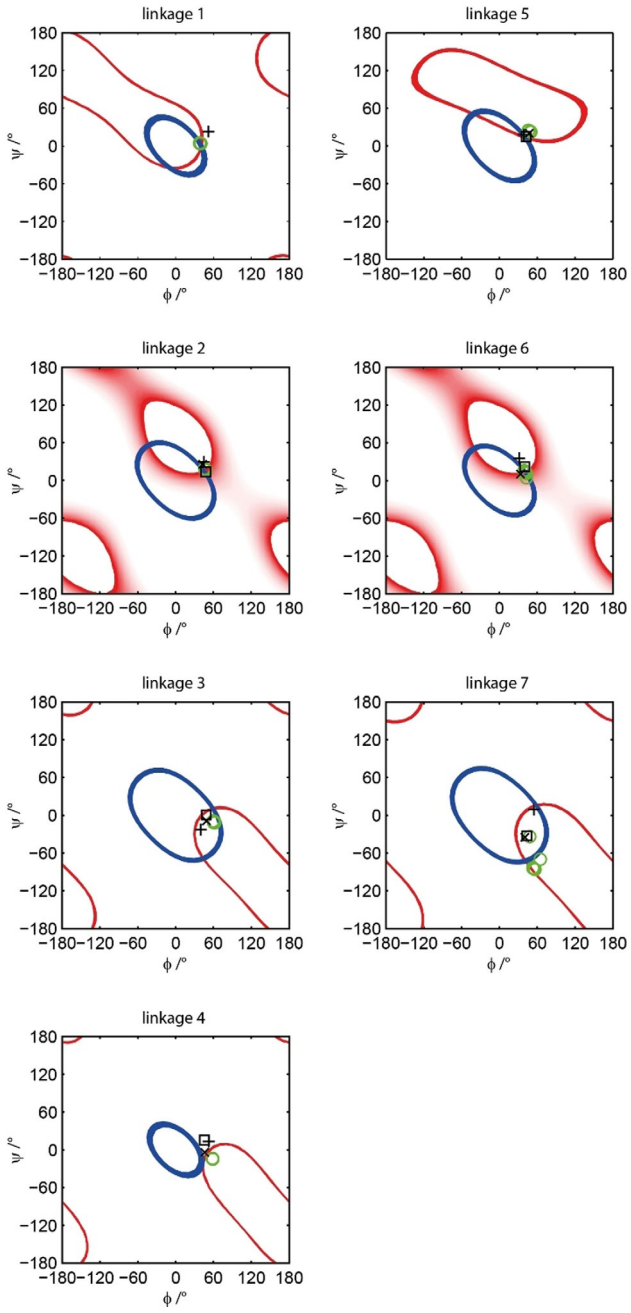

Figure 4. Distance curves for the octasaccharide describing the loci in terms of the glycosidic torsional angles ϕ and ψ obtained by 2D ^{1}H,^{1}H transfer NOESY NMR spectroscopy. Two-dimensional ϕ,ψ distance-plots are shown for all glycosidic linkages, denoted by superscripts in the *S. flexneri* O-serogroup 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 Table S2. Overlaid are conformations obtained from six non-redundant X-ray models (green circles) or MD simulations of Sf6TSP*wt* with AMBER/Glycam06 (+, black pluses) or CHARMM (□, black squares) force fields. In addition, conformers calculated from MD simulations with the Sf6TSP E366A D399A are shown (×, black crosses, AMBER/Glycam06). These data points were obtained from Kang et al..[7] For the panels of linkages (2) and (6) the red colour 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 (Table S2).

type Y octasaccharide 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.[7] Additionally, effective proton–proton distances agree in each data set between simulation, NMR cross-relaxations and crystallography (Table S2).

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.

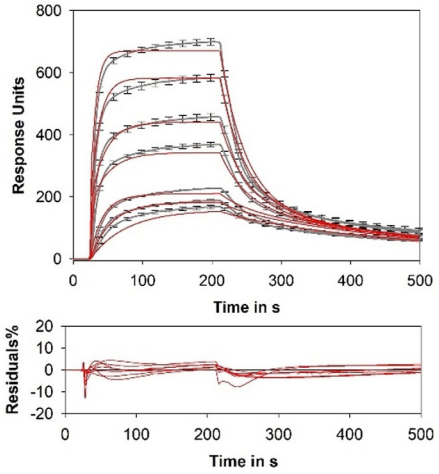

Figure 5. Surface plasmon resonance analysis of Sf6TSP binding to *S. flexneri* O-polysaccharide Y. SPR response curves obtained by injections of serial dilutions of $Sf6TSP_{EADA}$ (0.08-8 μM subunit concentration, grey, average of triplicate measurements with standard deviations shown for every 10th data point). Curves were fitted with a 1:2 heterogeneous binding model (red).

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

To obtain a rapid screening possibility 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 $Sf6TSP_{EADA}$ as a reference 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.[42] However, trials to directly couple this O-polysaccharide via Michael addition to a hydrazide-modified carboxymethyl dextran surface were unsuccessful. To obtain sufficient amounts of aldehyde groups and thus surface immobilized SfY O-polysaccharide, mild oxidation with sodium periodate was required prior to coupling.[43] To rule out that the oxidation step had altered the TSP binding capacity of the SfY O-polysaccharide, we tested periodate-oxidized 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 (Figure S5).

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 located between the subunits, we assume a multivalent interaction. A heterogeneous ligand parallel-binding model (1:2 model) provided the best description of the data ($\chi^2 \approx 380$), whereas simple 1:1 binding could be excluded (Figure S6a). Fitting the association and dissociation isotherms to the 1:2 model thus resulted in two equilibrium dissociation constants of $K_{D1} = 19.7(\pm 1.8)$ nM and $K_{D2} = 5.4(\pm 1.9)$ μM (Figure 5).

Reasonable similar values were obtained when evaluating the equilibrium signals at 178 s of injection, that is, $K_{D1} = 37.0(\pm 23.0)$ nM and $K_{D2} = 6.2(\pm 1.3)$ μM (Figure S6a). Multivalent binding was additionally confirmed by concentration dependent data analysis, resulting in a curved, triphasic distribution in the Scatchard plot (Figure S6b). Furthermore, the kinetic constants obtained revealed a good agreement of the second dissociation rate, $k_{off,2} \approx 0.024$–$0.027\ s^{-1}$, 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.[7] 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 (Figure S6c).[44,45] The polysaccharide surface was highly stable and more than 300 experiments were performed on one single surface (Figure S6d).

All measurements were repeated on two chip surfaces (1 and 2) resulting in comparable kinetic constants (Table 2, Table S3). 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 either 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, that is, nine individual measurements were carried out for each mutant to obtain a comparison of maximum response signals (Figure 6a and Figure S7a). As described above, binding of the Sf6TSP reference to SfY O-polysaccharide surfaces showed multivalent binding with two equilibrium dissociation constants of $\approx 10^{-8}$ M and $\approx 10^{-6}$ 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 had the same non-specific signal of around 9% on the reference channel of an individual chip surface (Figure S7b).[46] 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 related the distribution of maximum SPR signals between the different mutants to ligand flexibilities calculated as ligand RMSDs from MD simulations. Mean values of positional RMSDs were similar for all mutants that showed similar or higher SPR responses compared to the reference (Figure 6b). In V204C and S246C, individual ligand fluctuations were decreased; they 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 only had a low SPR response, indicating weak binding to the SfY polysaccharide surface (Figure 6c).

We also extracted hydrogen bonding patterns and occupancies from MD simulations of the ligand protein complexes to compare with binding strength estimates from SPR experiments (*cf.* Table 1 and Table S1). We found similar hydrogen bond occupancies in $Sf6TSP_{EADA}$ and mutants T315C, N340C

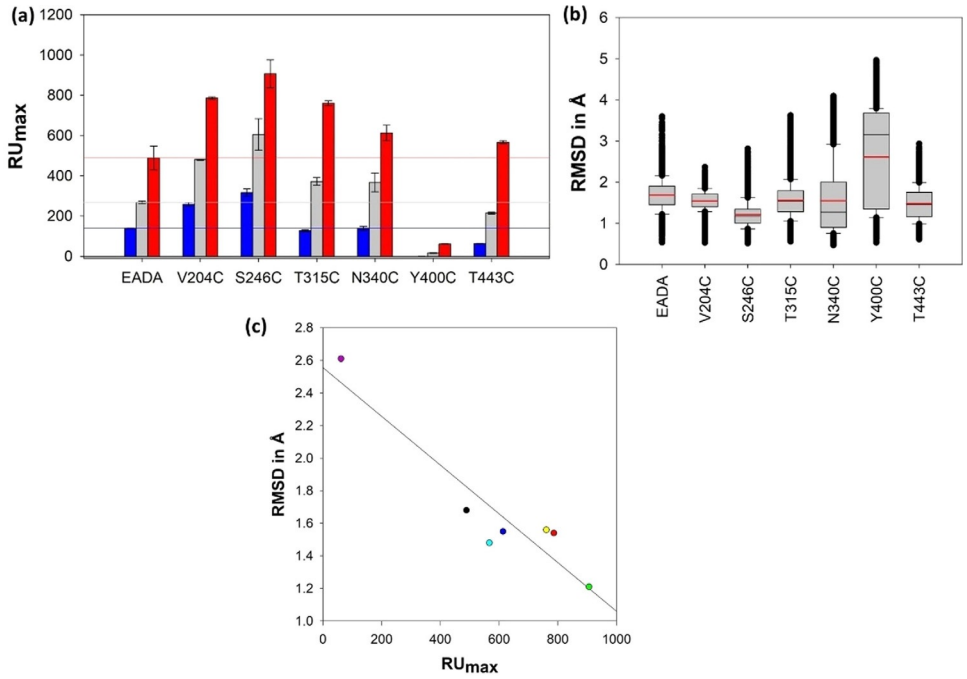

Figure 6. Polysaccharide surface binding of Sf6TSP mutants quantified with SPR and comparison with ligand flexibilities from MD simulations. a) Maximal SPR responses after 200 s of Sf6TSP mutant injections onto a *S. flexneri* O-serogroup Y polysaccharide surface. Each bar represents the mean value of three individual experiments at protein concentrations of 0.08 μM (blue), 0.8 μM (grey) or 8 μM (red). Horizontal lines mark responses for the $Sf6TSP_{EADA}$ reference without cysteine exchanges. b) Box-plotted time dependent fluctuations of octasaccharide RMSDs in cysteine mutants. 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, fluctuation outliers lie on the thick lines. c) Correlation of maximal SPR responses (protein concentration 8 μM) with mean ligand RMSDs (red: V204C, green: S246C, yellow: T315C, blue: N340C, purple: Y400C, cyan: T443C, black: reference EADA).

and T443C, that had similar SPR responses. V204C and S246C had higher occupancies, in agreement with higher SPR signals. In contrast, Y400C showed high hydrogen bond occupancies, although this mutant was a poor binder on the SfY polysaccharide surface. This illustrates that analysis of the hydrogen bonding pattern alone was not sufficient to fully predict affinities for polysaccharide surface binding. Rather, our simulations showed that single mutations on a protein surface may change completely or partly the observed hydrogen bonding patterns in number and occupancies, which may lead to individual glycan binding signatures in a given mutational background.

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 in *Streptococcus*.[47] Rhamnose containing biofilms were shown to effectively block antimicrobial peptides from entering biofilms.[48]

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 engineered cysteine residues.[2] 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. Cysteine mutants were introduced on the inactive mutant E366A D399A, as this mutant showed an increased affinity in comparison to the wild type and interaction measurement methods could be established with this mutant.[7] Typically, cysteines do not occur with high frequency in carbohydrate binding sites, and future studies should extend the data set to residues with higher propensity for glycan interactions.[49,50] 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, exchange for a cysteine increased the fluctuations of this hydrogen bond (Figure S8a,b). In hydrogen bonds, thiol groups mostly act as hydrogen donors to carbonyl groups, and, less frequently, also to carboxyl groups.[51] The C246 thiol group thus facilitated contacts with the neighbouring carboxyl group of D245, leading to stiffening of a ligand adjacent loop (residues 243–250) in comparison to the reference mutant E366A D399A (Figure S8c).

Cysteines may form intermolecular disulphide bridges and are thus often excluded from rational design approaches.[52] 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.[53,54] During our experiments we did not observe formation of higher oligomers from disulphide formation with the Sf6TSP cysteine mutants (Table S5). Additionally, 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 next to the binding site, MD simulations showed label interactions with the ligand that may explain the 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.[55,56]

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. It was shown that an analysis with the tool *MobyWat*, based on evaluation of solvent MD simulation on short time scales, provides a relatively simple access to the mobile water positional network on the protein surface.[9,57,58] In Sf6TSP, however, this method could not distinguish water networks between the different mutants (Figure S9), emphasizing the need for further mathematical algorithms to solvent network analysis.[59]

No general rule exists that assigns favourable or unfavourable contributions of water molecules to the driving forces for ligand binding.[60] In Sf6TSP, the glycan interaction site is formed by flexible protein loops that form the binding groove between two protein subunits.[32] The tailspike protein from bacteriophage HK620 has a very similar overall fold compared to Sf6TSP.[34] However, in HK620TSP the binding site lies not between subunits, but in a shallow surface depression formed by a rather hydrophobic β-sheet, accommodating a ligand glucose branch in an occluded surface cavity.[32,61] As a consequence, in HK620TSP the redistribution of water molecules has a major impact on the enthalpy-entropy compensation during the ligand binding event.[9] In Sf6TSP, the binding site is less dominated by water molecules than in HK620TSP,[7] 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,[39] and already subtle ligand conformational changes may favour dissociation, resulting in an overall low affinity.[62]

The mutant design approach in this work was rationally based on structural examination and computational analysis

by conventional MD simulations 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 nontrivial extension of the computational part. Alchemical methods can be reliable in predicting the stability of polypeptides up to 20 amino acids.[63] 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,[64] or the alchemical transformations 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.[65] Also ROSETTA-based techniques have been extended to include carbohydrate moieties on glycoproteins and glycan ligand docking.[66] 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.[67]

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 $\approx$20 nM and $\approx$6 μM. Furthermore, a curved Scatchard plot was obtained as it has been typically described for multivalent binding interactions.[68,69] 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. Binding the multivalent TSP to the multivalent sensor surface mimics the situation during phage infection, where the TSP as part of the phage tail has to bind perpendicular to the bacterial cell surface. Although the also perpendicularly protruding O-antigen chains would suggest a mainly parallel orientation of TSP and polysaccharide chains, cryoelectron tomography analyses showed that whole phage particles bound obliquely to the cell surface.[70] Therefore, an SPR surface-immobilized polysaccharide that might contain also surface-parallel regions mimics this situation, as periodate treatment of the polysaccharide might interfere with its fully perpendicular attachment. Here, interaction with more than one glycan binding site per TSP would then require a 60° rotation around the symmetry axis and result in an energetic penalty. Indeed, the negative cooperativity found in the Scatchard plot might point to this unfavourable surface binding where two sites on one TSP must bind simultaneously. Models that are more sophisticated should be therefore used that take into account the reaction volume and a probability factor for the multivalent binding event.[71] Considering these parameters results in a set of differential equations, in which also heterogeneous ligands can be included.[72] For example, these types of model adjustments well described binding of the trivalent S-layer protein SbsB to the surface-immobilized *Geobacillus stearothermophilus* secondary cell wall polymer.[73]

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 rendering the mature phage into an efficient, multivalent particle for adsorption to a bacterial surface to start infection.[74,75] 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.[76] A similar approach was chosen with eukaryotic viruses, where attachment studies were performed on glycosaminoglycan surfaces.[77] 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.[6] Similar SPR-based studies with carbohydrate-binding proteins have also been described.[68,78–80] Also in this work, a substantial binding signal amplification was obtained compared to the low affinity of a single octasaccharide binding site on Sf6TSP. Polysaccharide surface immobilization was thus the key for detecting small affinity changes at single protein sites within the multivalent TSP-polysaccharide binding system.

Conclusions

We employed an interdisciplinary approach of NMR spectroscopy, MD simulations and SPR interaction measurements for the assessment of subtle affinity differences in an elongated carbohydrate-binding site. Our work stresses that in a similar manner, these effects can be exploited to explore affinity fine tuning for protein–carbohydrate complexes.

Experimental Section

Materials and chemicals

All chemicals were of analytical grade and purchased from Sigma Aldrich (Taufkirchen, Germany) unless stated otherwise, and ultrapure water (PURELAB® flex, ELGA Veolia Water Technologies, Celle, Germany) 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 acidic hydrolysis as described.[81] Cloning and purification of Sf6TSP mutants have been described.[2,33]

Molecular dynamics simulations

Structures were parameterized with the AMBER03 force field for proteins[82] and the GLYCAM06 force field (v06j-1) for glycans.[83] Simulations were based on two subunits with one binding site cleft of the inactive mutant Sf6TSPΔN E366A D399A (residues 109–622, PDB ID: 4URR) and on six cysteine mutants thereof in complex with octasaccharides of *S. flexneri* O-polysaccharide serogroup Y (Table S4).[7] 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).[84,85] The simulation complex was placed in an orthogonal box with the dimensions $\approx 140\times100\times120$ Å^3. Simulations were run with the TIP3P water model[86] and charges were equalized with 18 sodium ions.

MD simulations were carried out using the GROMACS4.5.5 program package.[87–89] After energy minimization to a maximum force smaller than 1000 kJ mol^{-1} nm^{-1} (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)[90] 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}.[91] 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 (PDB ID: 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*[57,58] as described previously.[9] 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 E366A D399A (0.12 mM) and the octasaccharide ligand (1.87 mM) in D_2O sodium phosphate buffer (100 mM, pD = 7) as described earlier.[7] Besides the previously recorded 2D ^{1}H,^{1}H transfer NOESY spectrum with a mixing time of 120 ms, additional experiments were carried out with mixing times of 40, 60 and 100 ms employing a zero-quantum suppression filter with a pulsed-field-gradient strength of 9% of the maximum (53.0 G cm^{-1}). The experiments were recorded with 256×16 k data points in the F_1 and F_2 dimensions, respectively, using 32 scans per increment and 8 dummy scans, an acquisition time of 1.6 s and an interscan delay of 3 s. Prior to Fourier transformation, forward linear prediction with 120 coefficients was applied in the F_1 dimension using a total of 1k × 16k points in F_1 and F_2, respectively; 90° shifted squared sine-bell window functions were used in both dimensions. 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.[7] 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 $t = 0$.[92] Due to differences in effective correlation times of protons in the protein-bound oligosaccharide, namely, between the ones in fast spinning methyl groups[41] and those from methine protons, two different reference distances were used in the analysis relying on the isolated spin-pair approximation:[93] 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/Glycam06 force field was used as a template for calculating atom–atom distances.[7] Molecular models covering the full glycosidic conformational space were generated in Vega ZZ (release 2.3.1.2)[94] 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[92] 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. *S. 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.[43,95] 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.[43] 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. Cysteine containing Sf6TSP mutants were checked for the absence of disulphide mediated aggregate formation by UV light scattering (Table S5), to ensure that prolonged storage in the autosampler of the instrument at < 10 °C during the three days' measurement time had not formed higher oligomers. 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 (Figure S10).

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 pro-

gramme 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.

Conflict of interest

The authors declare no conflict of interest.

Keywords: carbohydrates · molecular dynamics simulations · NMR spectroscopy · protein–carbohydrate interactions · surface plasmon resonance

[1] A. Varki, *Glycobiology* **2017**, *27*, 3 – 49.
[2] S. Kunstmann, T. Scheidt, S. Buchwald, A. Helm, L. Mulard, A. Fruth, S. Barbirz, *Viruses* **2018**, *10*, 431.
[3] R. Marchetti, M. J. Dillon, M. N. Burtnick, M. A. Hubbard, M. T. Kenfack, Y. Blériot, C. Gauthier, P. J. Brett, D. P. AuCoin, R. Lanzetta, A. Silipo, A. Molinaro, *ACS Chem. Biol.* **2015**, *10*, 2295 – 2302.
[4] T. Fiebig, C. Litschko, F. Freiberger, A. Bethe, M. Berger, R. Gerardy-Schahn, *J. Biol. Chem.* **2018**, *293*, 953 – 962.
[5] D. Keogh, R. Thompson, R. Larragy, K. McMahon, M. O'Connell, B. O'Connor, P. Clarke, *Biochim. Biophys. Acta* **2014**, *1840*, 2091 – 2104.
[6] J. J. Lundquist, E. J. Toone, *Chem. Rev.* **2002**, *102*, 555 – 578.
[7] Y. Kang, U. Gohlke, O. Engström, C. Hamark, T. Scheidt, S. Kunstmann, U. Heinemann, G. Widmalm, M. Santer, S. Barbirz, *J. Am. Chem. Soc.* **2016**, *138*, 9109 – 9118.
[8] N. K. Vyas, M. N. Vyas, M. C. Chervenak, D. R. Bundle, B. M. Pinto, F. A. Quiocho, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15023 – 15028.
[9] S. Kunstmann, U. Gohlke, N. K. Broeker, Y. Roske, U. Heinemann, M. Santer, S. Barbirz, *J. Am. Chem. Soc.* **2018**, *140*, 10447 – 10455.
[10] N. Pokala, T. M. Handel, *J. Struct. Biol.* **2001**, *134*, 269 – 281.
[11] B. Höcker, S. Beismann-Driemeyer, S. Hettwer, A. Lustig, R. Sterner, *Nat. Struct. Mol. Biol.* **2001**, *8*, 32 – 36.
[12] *Directed Evolution Library Creation: Methods and Protocols* (Eds.: F. H. Arnold, G. Georgiou) Humana, **2003**.
[13] S. A. Marshall, G. A. Lazar, A. J. Chirino, J. R. Desjarlais, *Drug Discovery Today* **2003**, *8*, 212 – 221.
[14] C. A. Rohl, C. E. M. Strauss, K. M. S. Misura, D. Baker, *Methods in Enzymology*, Vol. 383, 66 – 93, Academic Press, **2004**.
[15] B. Kuhlman, *J. Biol. Chem.* **2019**, *294*, 19436 – 19443.
[16] R. Pearce, X. Huang, D. Setiawan, Y. Zhang, *J. Mol. Biol.* **2019**, *431*, 2467 – 2476.
[17] G. Nimrod, S. Fischman, M. Austin, A. Herman, F. Keyes, O. Leiderman, D. Hargreaves, M. Strajbl, J. Breed, S. Klompus, K. Minton, J. Spooner, A. Buchanan, T. J. Vaughan, Y. Ofran, *Cell Rep.* **2018**, *25*, 2121 – 2131.e5.
[18] M. Chen, X. Shi, R. M. Duke, C. I. Ruse, N. Dai, C. H. Taron, J. C. Samuelson, *Nat. Commun.* **2017**, *8*, 15487.
[19] S. Mao, C. Gao, C.-H. L. Lo, P. Wirsching, C.-H. Wong, K. D. Janda, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6953 – 6958.
[20] S. Schoonbroodt, M. Steukers, M. Viswanathan, N. Frans, M. Timmermans, A. Wehnert, M. Nguyen, R. C. Ladner, R. M. Hoet, *J. Immunol.* **2008**, *181*, 6213 – 6221.
[21] A. Morin, K. W. Kaufmann, C. Fortenberry, J. M. Harp, L. S. Mizoue, J. Meiler, *Protein Eng. Des. Sel.* **2011**, *24*, 503 – 516.
[22] J. Meiler, D. Baker, *Proteins* **2006**, *65*, 538 – 548.
[23] K. P. Kilambi, M. S. Pacella, J. Xu, J. W. Labonte, J. R. Porter, P. Muthu, K. Drew, D. Kuroda, O. Schueler-Furman, R. Bonneau, J. J. Gray, *Proteins Struct. Funct. Bioinf.* **2013**, *81*, 2201 – 2209.
[24] A. Gimeno, S. Delgado, P. Valverde, S. Bertuzzi, M. A. Berbís, J. Echavarren, A. Lacetera, S. Martín-Santamaría, A. Surolia, F. J. Cañada, J. Jiménez-Barbero, A. Ardá, *Angew. Chem. Int. Ed.* **2019**, *58*, 7268 – 7272; *Angew. Chem.* **2019**, *131*, 7346 – 7350.
[25] M. Lepsik, R. Sommer, S. Kuhaudomlarp, M. Lelimousin, E. Paci, A. Varrot, A. Titz, A. Imberty, *Eur. J. Medic. Chem.* **2019**, *177*, 212 – 220.
[26] D. Andres, U. Gohlke, N. K. Broeker, S. Schulze, W. Rabsch, U. Heinemann, S. Barbirz, R. Seckler, *Glycobiology* **2013**, *23*, 486 – 494.
[27] W. Peng, R. P. de Vries, O. C. Grant, A. J. Thompson, R. McBride, B. Tsogtbaatar, P. S. Lee, N. Razi, I. A. Wilson, R. J. Woods, J. C. Paulsen, *Cell Host Microbe* **2017**, *21*, 23 – 34.
[28] *Computational Drug Discovery and Design* (Ed.: R. Baron), Springer, New York, **2012**.
[29] D. Andres, U. Baxa, C. Hanke, R. Seckler, S. Barbirz, *Biochem. Soc. Trans.* **2010**, *38*, 1386 – 1389.
[30] A. A. Lindberg, *Annu. Rev. Microbiol.* **1973**, *27*, 205 – 241.
[31] A. A. Lindberg, R. Wollin, P. Gemski, J. A. Wohlhieter, *J. Virol.* **1978**, *27*, 38.
[32] J. J. Müller, S. Barbirz, K. Heinle, A. Freiberg, R. Seckler, U. Heinemann, *Structure* **2008**, *16*, 766 – 775.
[33] A. Freiberg, R. Morona, L. Van Den Bosch, C. Jung, J. Behlke, N. Carlin, R. Seckler, U. Baxa, *J. Biol. Chem.* **2003**, *278*, 1542 – 1548.
[34] S. Barbirz, M. Becker, A. Freiberg, R. Seckler, *Macromol. Biosci.* **2009**, *9*, 169 – 173.
[35] Z. Chang, J. Zhang, L. Ran, J. Sun, F. Liu, L. Luo, L. Zeng, L. Wang, Z. Li, H. Yu, Q. Liao, *BMC Infectious Diseases* **2016**, *16*, 685.
[36] K. L. Kotloff, J. P. Nataro, W. C. Blackwelder, D. Nasrin, T. H. Farag, S. Panchalingam, Y. Wu, S. O. Sow, D. Sur, R. F. Breiman, A. S. Faruque, A. K. Zaidi, D. Saha, P. L. Alonso, B. Tamboura, D. Sanogo, U. Onwuchekwa, B. Manna, T. Ramamurthy, S. Kanungo, J. B. Ochieng, R. Omore, J. O. Oundo, A. Hossain, S. K. Das, S. Ahmed, S. Qureshi, F. Quadri, R. A. Adegbola, M. Antonio, M. J. Hossain, A. Akinsola, I. Mandomando, T. Nhampossa, S. Acácio, K. Biswas, C. E. O'Reilly, E. D. Mintz, L. Y. Berkeley, K. Muhsen, H. Sommerfelt, R. M. Robins-Browne, M. M. Levine, *The Lancet* **2013**, *382*, 209 – 222.
[37] W. H. Chen, K. L. Kotloff, *Clin. Vaccine Immunol.* **2016**, *23*, 904 – 907.
[38] J. Boutet, P. Blasco, C. Guerreiro, F. Thouron, S. Dartevelle, F. Nato, F. J. Cañada, A. Ardá, A. Phalipon, J. Jiménez-Barbero, L. A. Mulard, *Chem. Eur. J.* **2016**, *22*, 10892 – 10911.
[39] Y. Kang, S. Barbirz, R. Lipowsky, M. Santer, *J. Phys. Chem. B* **2014**, *118*, 2523 – 2534.
[40] T. Galochkina, D. Zlenko, A. Nesterenko, I. Kovalenko, M. Strakhovskaya, A. Averyanov, A. Rubin, *ChemPhysChem* **2016**, *17*, 2839 – 2853.
[41] G. Widmalm, R. W. Pastor, T. E. Bull, *J. Chem. Phys.* **1991**, *94*, 4097 – 4098.
[42] O. Lüderitz, O. Westphal, A. M. Staub, H. Nikaido, in *Bacterial Endotoxins*, Academic Press, Amsterdam, **1971**, pp. 145 – 233.
[43] E. T. Gedig, *Handb. Surf. Plasmon Reson.* **2008**, 173 – 220.
[44] R. A. Copeland, *Enzym. Pract. Introd. Struct. Mech. Data Anal.* **2000**, 76 – 108.
[45] P. Schuck, A. P. Minton, *Trends Biochem. Sci.* **1996**, *21*, 458 – 460.
[46] M. Jung, M. Philpott, S. Muller, J. Schulze, V. Badock, U. Eberspacher, D. Moosmayer, B. Bader, N. Schmees, A. Fernandez-Montalvan, B. Händler, *J. Biol. Chem.* **2014**, *289*, 9304 – 9319.
[47] R. J. Edgar, V. P. van Hensbergen, A. Ruda, A. G. Turner, P. Deng, Y. L. Breton, N. M. El-Sayed, A. T. Belew, K. S. McIver, A. G. McEwan, A. J. Morris, G. Lambeau, M. J. Walker, J. S. Rush, K. V. Korotkov, G. Widmalm, N. M. van Sorge, N. Korotkova, *Nat. Chem. Biol.* **2019**, *15*, 463 – 471.
[48] B. Bellich, C. Lagatolla, A. Tossi, M. Benincasa, P. Cescutti, R. Rizzo, *Int. J. Mol. Sci.* **2018**, *19*, 1685.
[49] M. Kulharia, S. J. Bridgett, R. S. Goody, R. M. Jackson, *J. Mol. Graph. Model.* **2009**, *28*, 297 – 303.
[50] C. Taroni, S. Jones, J. M. Thornton, *Protein Eng.* **2000**, *13*, 89 – 98.
[51] D. Pal, P. Chakrabarti, *J. Biomol. Struct. Dyn.* **1998**, *15*, 1059 – 1072.
[52] O. Sharabi, A. Erijman, J. M. Shifman, *Methods in Enzymology*, Elsevier, **2013**, pp. 41 – 59.
[53] H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637 – 674.
[54] Y. Shinohara, Y. Hasegawa, H. Kaku, N. Shibuya, *Glycobiology* **1997**, *7*, 1201 – 1208.
[55] U. Gether, S. Lin, B. K. Kobilka, *J. Biol. Chem.* **1995**, *270*, 28268 – 28275.
[56] V. E. V. Ferrero, G. Di Nardo, G. Catucci, S. J. Sadeghi, G. Gilardi, *Dalt. Trans.* **2012**, *41*, 2018 – 2025.
[57] N. Jeszenői, M. Bálint, I. Horváth, D. van der Spoel, C. Hetényi, *J. Chem. Inf. Comput. Sci.* **2016**, *56*, 148 – 158.
[58] N. Jeszenői, I. Horváth, M. Bálint, D. van der Spoel, C. Hetényi, *Bioinformatics* **2015**, *31*, 1959 – 1965.

[59] N. R. Taylor, *Comput. Struct. Biotechnol. J.* **2013**, *5*, e201302006.
[60] J. E. Ladbury, *Chem. Biol.* **1996**, *3*, 973 – 980.
[61] S. Barbirz, J. J. Müller, C. Uetrecht, A. J. Clark, U. Heinemann, R. Seckler, *Mol. Microbiol.* **2008**, *69*, 303 – 316.
[62] M. Ahmad, V. Helms, O. V. Kalinina, T. Lengauer, *J. Phys. Chem. B* **2016**, *120*, 2138 – 2144.
[63] V. Gapsys, S. Michielssens, D. Seeliger, B. L. de Groot, *J. Comput. Chem.* **2015**, *36*, 348 – 354.
[64] C. Kutzner, S. Páll, M. Fechner, A. Esztermann, B. L. de Groot, H. Grubmüller, *J. Comput. Chem.* **2015**, *36*, 1990 – 2008.
[65] A. K. Nivedha, D. F. Thieker, S. Makeneni, H. Hu, R. J. Woods, *J. Chem. Theory Comput.* **2016**, *12*, 892 – 901.
[66] J. W. Labonte, J. Adolf-Bryfogle, W. R. Schief, J. J. Gray, *J. Comput. Chem.* **2017**, *38*, 276 – 287.
[67] L. Li, W. Xu, Q. Lü, *J. Mol. Model.* **2015**, *21*, 294.
[68] J. Häyrinen, S. Haseley, P. Talaga, M. Mühlenhoff, J. Finne, J. F. Vliegenthart, *Mol. Immunol.* **2002**, *39*, 399 – 411.
[69] C. R. MacKenzie, T. Hirama, S. Deng, D. R. Bundle, S. A. Narang, N. M. Young, *J. Biol. Chem.* **1996**, *271*, 1527 – 1533.
[70] C. Wang, J. Tu, J. Liu, I. J. Molineux, *Nat. Microbiol.* **2019**, *4*, 1049 – 1056.
[71] K. M. Müller, K. M. Arndt, A. Plückthun, *Anal. Biochem.* **1998**, *261*, 149 – 158.
[72] T. J. van Steeg, K. R. Bergmann, N. Dimasi, K. F. Sachsenmeier, B. Agoram, *mAbs* **2016**, *8*, 585 – 592.
[73] C. Mader, C. Huber, D. Moll, U. B. Sleytr, M. Sara, *J. Bacteriol.* **2004**, *186*, 1758 – 1768.
[74] N. K. Broeker, Y. Roske, A. Valleriani, M. S. Stephan, D. Andres, J. Koetz, U. Heinemann, S. Barbirz, *J. Biol. Chem.* **2019**, *294*, 11751 – 11761.
[75] N. K. Broeker, S. Barbirz, *Mol. Microbiol.* **2017**, *105*, 353 – 357.
[76] R. Morona, *Microbiology* **2003**, *149*, 925 – 939.
[77] N. Peerboom, S. Block, N. Altgärde, O. Wahlsten, S. Möller, M. Schnabelrauch, E. Trybala, T. Bergström, M. Bally, *Biophys. J.* **2017**, *113*, 1223 – 1234.
[78] B. Brogioni, F. Berti, *MedChemComm* **2014**, *5*, 1058.
[79] J. Houser, J. Komarek, G. Cioci, A. Varrot, A. Imberty, M. Wimmerova, *Acta Crystallogr. Sect. D* **2015**, *71*, 442 – 453.
[80] H. Nakajima, *J. Biol. Chem.* **2001**, *276*, 42915 – 42922.
[81] Y. Haishima, O. Holst, H. Brade, *Eur. J. Biochem.* **1992**, *203*, 127 – 134.
[82] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, *J. Comput. Chem.* **2003**, *24*, 1999 – 2012.
[83] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.* **2008**, *29*, 622 – 655.
[84] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* **2006**, *25*, 247 – 260.
[85] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157 – 1174.
[86] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
[87] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435 – 447.
[88] S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, *Bioinformatics* **2013**, *29*, 845 – 854.
[89] M. Wehle, I. Vilotijevic, R. Lipowsky, P. H. Seeberger, D. Varon Silva, M. Santer, *J. Am. Chem. Soc.* **2012**, *134*, 18964 – 18972.
[90] M. Parrinello, *J. Appl. Phys.* **1981**, *52*, 7182.
[91] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J Comp Chem.* **1997**, *18*, 1463 – 1472.
[92] G. Widmalm, R. A. Byrd, W. Egan, *Carbohydr. Res.* **1992**, *229*, 195 – 211.
[93] P. D. Thomas, V. J. Basus, T. L. James, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 1237 – 1241.
[94] A. Pedretti, L. Villa, G. Vistoli, *J. Mol. Graph. Mod.* **2002**, *21*, 47 – 49.
[95] C. A. Wolfe, D. S. Hage, *Anal. Biochem.* **1995**, *231*, 123 – 130.

Manuscript received: January 29, 2020
Revised manuscript received: March 9, 2020
Accepted manuscript online: March 19, 2020
Version of record online: May 19, 2020