Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 339 (2004) 259–267
Carbohydrate RESEARCH

Saturation transfer difference NMR and computational modeling of a sialoadhesin–sialyl lactose complex

Anirban Bhunia,[a] V. Jayalakshmi,[b] Andrew J. Benie,[a] Oliver Schuster,[a] Sørge Kelm,[c] N. Rama Krishna[b,*] and Thomas Peters[a,*]

[a]*Institute of Chemistry, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany*
[b]*Department of Biochemistry and Molecular Genetics, Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL 35294-2041, USA*
[c]*Institute for Physiological Biochemistry, University Bremen, 28334 Bremen, Germany*

Received 5 May 2003; revised 5 August 2003; accepted 16 September 2003

Abstract—The siglecs are a family of I-type lectins binding to sialic acids on the cell surface. Sialoadhesin (siglec-1) is expressed at much higher levels in inflammatory macrophages and specifically binds to α-2,3-sialylated *N*-acetyl lactosamine residues of glycan chains. The terminal disaccharide α-D-Neu5Ac-(2→3)-β-D-Gal is thought to be the main epitope recognized by sialoadhesin. To understand the basis of this biological recognition reaction we combined NMR experiments with a molecular modeling study. We employed saturation transfer difference (STD) NMR experiments to characterize the binding epitope of α-2,3-sialylated lactose, α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-D-Glc **1** to sialoadhesin at atomic resolution. The experimental results were compared to a computational docking model and to X-ray data of a complex of sialyl lactose and sialoadhesin. The data reveal that sialoadhesin mainly recognizes the *N*-acetyl neuraminic acid and a small part of the galactose moiety of **1**. The crystal structure of a complex of sialoadhesin with sialyl lactose **1** was used as a basis for a modeling study using the FlexiDock algorithm. The model generated was very similar to the original crystal structure. Therefore, the X-ray data were used to predict theoretical STD values utilizing the CORCEMA-STD protocol. The good agreement between experimental and theoretical STD values indicates that a combined modeling/STD NMR approach yields a reliable structural model for the complex of sialoadhesin with α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-D-Glc **1** in aqueous solution.

Keywords: Sialoadhesin; Sialyl lactose; STD NMR; FlexiDock; CORCEMA-STD

1. Introduction

Mammalian carbohydrate-binding proteins (animal lectins) have been classified into different groups on the basis of structural features of the lectins themselves, and the types of carbohydrate ligands that are recognized.[1] The siglec (*s*ialic acid binding *i*mmuno*g*lobulin like *lec*tins) family[2,3] constitutes a group of cellular recognition molecules that are characterized by sequence homology with members of the Ig super family. There are at least 11 different siglecs present in humans. The siglecs can be divided into two subgroups, one represented by sialoadhesin (siglec-1), CD22 (siglec-2), MAG (siglec-4a),

Abbreviations: pH*, pH uncorrected for deuterium isotope effect; STD, saturation transfer difference; GEM, group epitope mapping; H^{3ax}, H^3 axial; H^{3eq}, H^3 equatorial; ε, dielectric constant; DPFGSE, double pulsed field gradient spin echo; IgSF, immunoglobulin super family; Siglec, sialic acid binding Ig-like lectins; Neu5Ac, *N*-acetyl neuraminic acid; Gal, galactose; Glc, glucose; IL, interleukin; MAG, myelin-associated glycoprotein; CORCEMA, complete relaxation and conformational exchange matrix analysis; MCS, Multilevel Coordinate Search; SMP, Schwann cell myelin protein.

* Corresponding authors. Tel.: +49-451-500-4230; fax: +49-451-500-4241; e-mail addresses: nrk@uab.edu; thomas.peters@chemie.uni-luebeck.de

0008-6215/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2003.09.021

and SMP (siglec-4b), the other subgroup consisting of CD33 (siglec-3) related siglecs.[4] Several biological functions for siglecs have been proposed: MAG apparently plays a role in the inhibition of neurite outgrowth,[5,6] and CD22 is involved in the regulation of B-cell dependent immune responses.[7] For sialoadhesin, a role in the regulation of interactions between tissue macrophages and myeloid cells has been assumed, since sialoadhesin binds preferentially to myeloid cells from all stages of maturation,[8] and it is highly concentrated at the contact areas between macrophages and developing myeloid cells in the hemopoietic clusters of the bone marrow.[9] None of the cytokines was able to induce sialoadhesin expression, however, interleukin 4 (IL 4) prevented the induction of expression in presence of serum.[10] Each siglec exhibits a characteristic preference for both, the type of *N*-acyl neuraminic acid, and its linkage to the subterminal sugar. For instance, sialoadhesin, CD33, and MAG bind preferentially to *N*-acetyl neuraminic acid in α-2,3 linkages[2] whereas CD22 binds preferentially to Neu5Ac in α-2,6 linkages.[11] These differences in sugar-binding specificities are likely to be important in the cellular-recognition functions of these proteins.

To better understand the in vivo interaction of sialylated oligosaccharides with sialoadhesin we studied the interaction under physiological conditions, that is, in aqueous solution at neutral pH employing NMR experiments. For the complex of sialoadhesin with α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-D-Glc **1** a crystal structure is available,[12] which allows a comparison with the solution state data.

2. Results

2.1. STD NMR experiment

For the analysis of protein–ligand interactions a large number of NMR experiments exist. The use of NMR experiments to identify and characterize the binding of ligands to proteins has been reviewed recently.[13] For STD NMR[14] experiments a sample is required that contains both, the ligand and the receptor, with an excess of ligand (usually ca. 100-fold molar excess) over the receptor. A reference ^{1}H NMR spectrum is recorded with irradiation (hereafter denoted as on-resonance spectrum) of the protein signal envelope at a frequency where no ligand protons resonate. A second spectrum is recorded with the irradiation frequency (hereafter denoted as off-resonance spectrum) set to a value that is significantly apart from all resonance frequencies of either the protein or the ligand. Subtraction of the two spectra leads to a difference spectrum that contains only signals resulting from saturation transfer.

Figure 1A shows a ^{1}H NMR spectrum of the complex (1:1000 excess of ligand) using DPFGSE[15] for water suppression. Saturation was exclusively transferred to molecules that bind to sialoadhesin. Therefore, the difference spectrum (Fig. 1B) contained only signals of

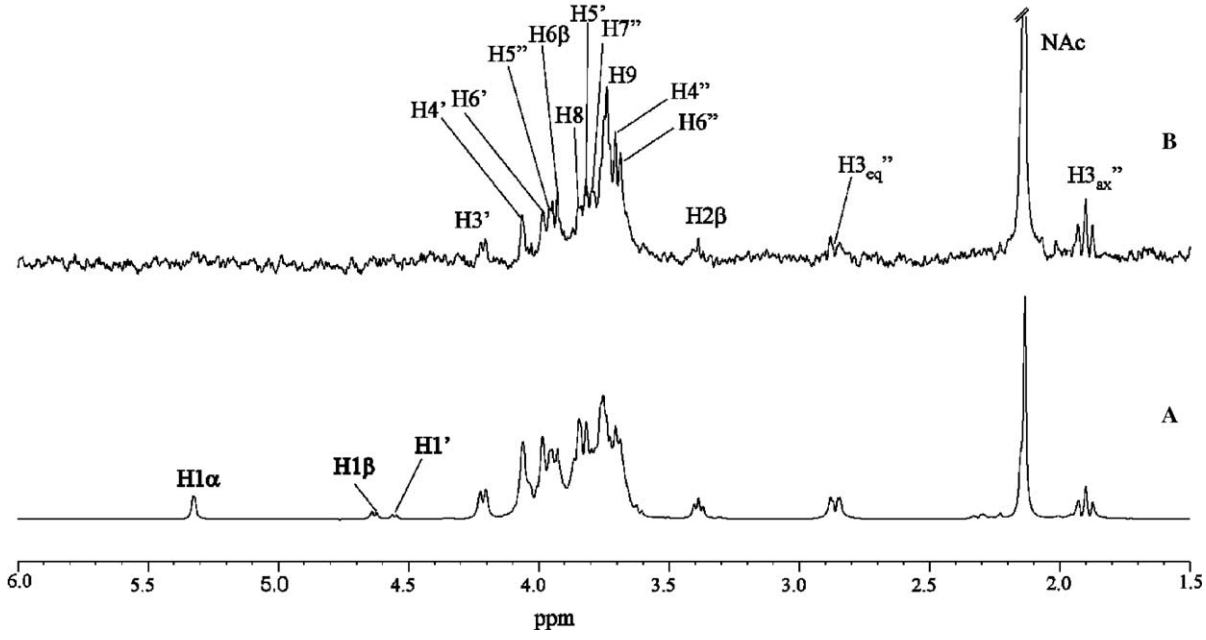


Figure 1. One dimensional STD spectra of sialyl lactose **1** in the presence of sialoadhesin at 500 MHz. (A) Reference ^{1}H NMR spectrum of sialyl lactose (5 mM) in the presence of sialoadhesin (5 μM) using DPFGSE for water suppression. (B) Corresponding STD NMR spectrum showing sialyl lactose in the presence of sialoadhesin. Only signals from sialyl lactose are visible indicating binding to sialoadhesin.

sialyl lactose **1**, with signals from low molecular weight impurities being removed.

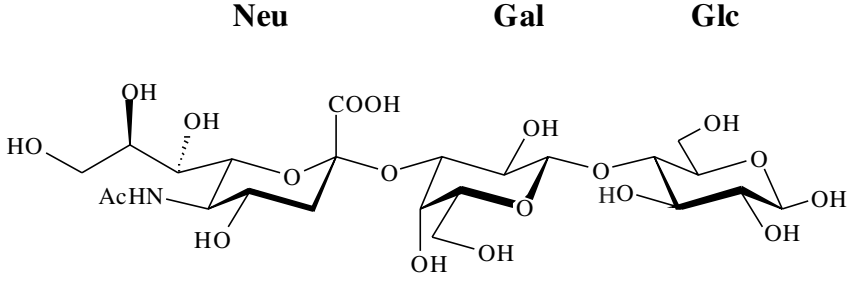
2.2. Group epitope mapping

Effective group epitope mapping (GEM)[16] is possible if the dissociation rate constant k_{off} is greater than, or of the same magnitude as the T_2 relaxation rate of the bound ligand because the ligand has to leave the binding site before all saturation has been equally distributed amongst all of the spins in the ligand via spin diffusion. Since sialyl lactose **1** is a weakly binding ligand with a dissociation constant (K_D) of 0.8 mM,[17] the dissociation rate constant k_{off} should be rather large, assuming a diffusion controlled on-rate of the ligand furnishing ideal conditions for STD NMR experiments.

The largest STD effect was observed for the *N*-acetyl methyl group of sialyl lactose **1**. Therefore, this signal was used as a reference, and was set to 100%. The relative degree of saturation for the individual protons is displayed in Figure 2, and immediately shows that the *N*-acetyl neuraminic acid residue is in most intimate contact with protons in the binding site of sialoadhesin. For protons not shown in Figure 2 no STD values were determined because of severe overlap with other resonances. For the anomeric protons of galactose and glucose no STD response was observed but since their signals were close to the HDO resonance this result was not used for the quantitative interpretation of the data. Saturation transfer to the protons of galactose was significantly less effective, and for the reducing glucose residue almost no STD response was detected.

2.3. CORCEMA-STD calculations

The calculation of theoretical STD values was performed with the CORCEMA program[18,19] that has recently been extended for the prediction of STD effects.[20] A comparison of experimental STD values and the values calculated for the crystal structure of the complex is given in Figure 3. STD values are given in percent as relative values $[((I_{0(k)} - I(t)_{(k)})/I_{0(k)}) \times 100]$ as described in the experimental section. By fixing k_{on}, the other parameters were optimized to find the best fit between experimental and calculated STD values. This was performed using the Multilevel Coordinate Search (MCS) method to find the minimum *R*-factor by simultaneously optimizing free ligand correlation time, the correlation time for the protein and the complex, the dissociation constant K_D, and the order parameter S^2 for the dipolar interaction of methyl protons with other protons. Initial optimizations at each of the (1:100, 1:300, 1:500, and 1:1000) protein/ligand ratios were performed. The values of free ligand correlation time, K_D, and the order parameter S^2 are



Sialyl lactose **1**.

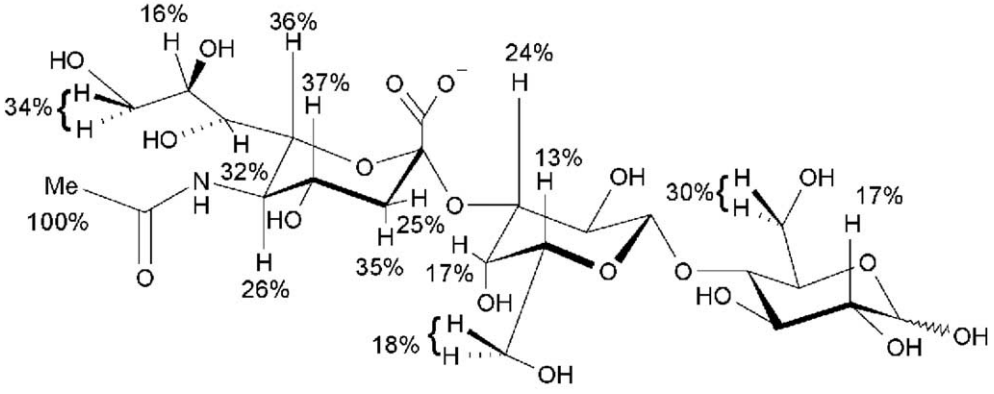


Figure 2. The relative STD effects of sialyl lactose bound to sialoadhesin. The values were calculated by determining individual signal intensities in the STD spectrum (I_{STD}), and in the reference ^{1}H NMR spectrum (I_0). The ratios of the intensities $(I_{STD} - I_0)/I_0$ were normalized using the largest STD effect (acetamido methyl group, 100%) as a reference. The number indicates the position of the proton experiencing an STD effect at a 1000-fold excess. The concentration of sialoadhesin was 5 μM and that of sialyl lactose was 5 mM.

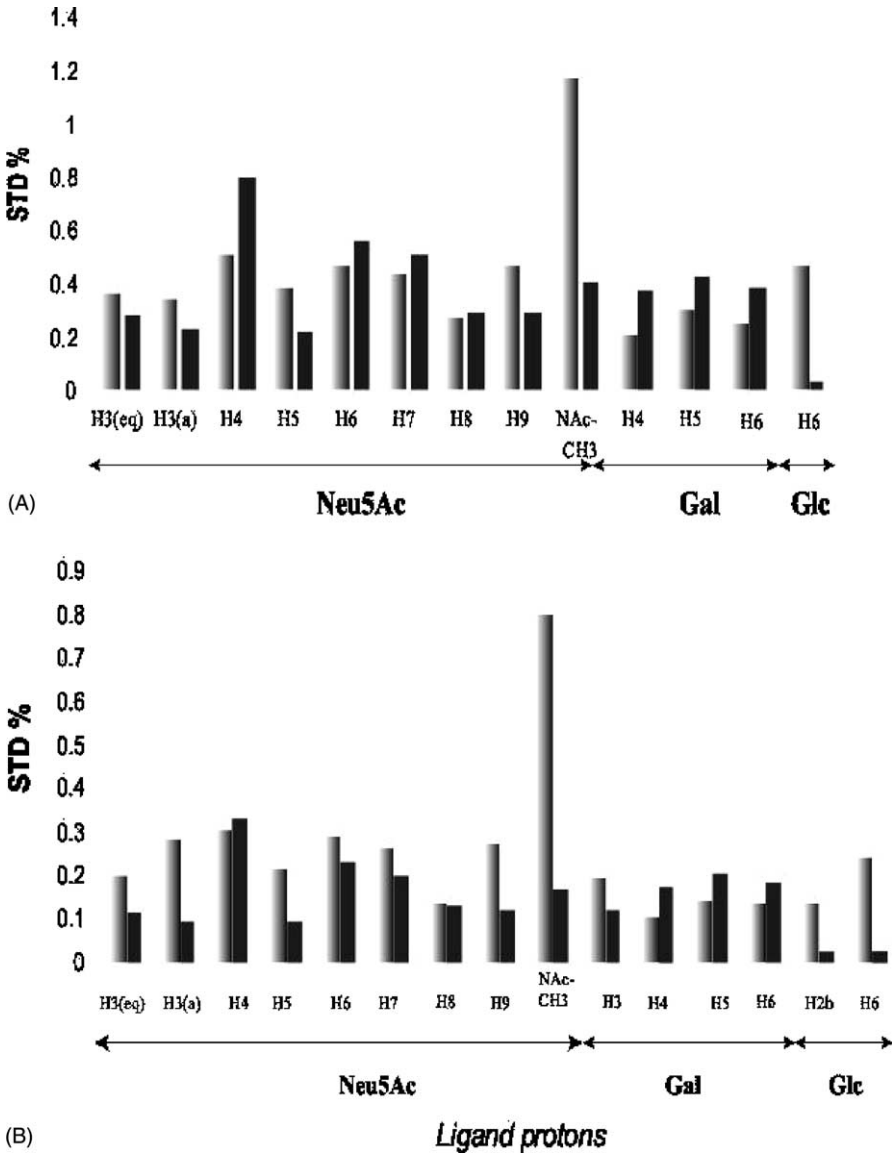


Figure 3. Comparison of experimental STD values (black bars) and predicted values from CORCEMA for the crystallographic structure (white bars). (A) Protein/ligand = 1:300 and (B) protein/ligand = 1:1000. STD values were calculated as $[((I_{0(k)} - I(t)_{(k)})/I_{0(k)}) \times 100]$, with $I_{0(k)}$ being the intensity of the signal of proton k without saturation transfer at time $t = 0$, and $I(t)_k$ being the intensity of proton k after a saturation transfer during the saturation time t.

found to be the same irrespective of the ligand concentration. In contrast, the value of correlation time of the protein and complex is close to 3×10^{-8} s when the ligand excess is 100, 300, and 500 but it is slightly different (4×10^{-8} s) at 1:1000 ratio of protein/ligand. The final set of calculations was performed by fixing the correlation time of sialoadhesin and the complex at 3×10^{-8} s for the four different protein/ligand ratios.

2.4. FlexiDock

Sialyl lactose **1** was first manually docked into the active site of sialoadhesin using the geometry of the crystal structure as a guide. The optimization of the binding geometry was then performed using the program FlexiDock (part of the SYBYL software package). The protein was kept rigid whereas the ligand was fully flexible and allowed to move in the binding site. The interaction energy was calculated using van der Waals, electrostatic, and torsional energy terms of the Tripos force field. The entire complex was energy minimized in 30 cycles using FlexiDock. The result of the docking is shown in Figure 4. The dihedral angles ϕ and ψ at the α-2,3-glycosidic linkages in **1** obtained from the FlexiDock model were almost identical to the values from the X-ray structure (Table 1). The comparison of the torsional angles in the *exo*-cyclic side chains of the pyranose rings of sialyl lactose **1** (Table 2) also corresponded well

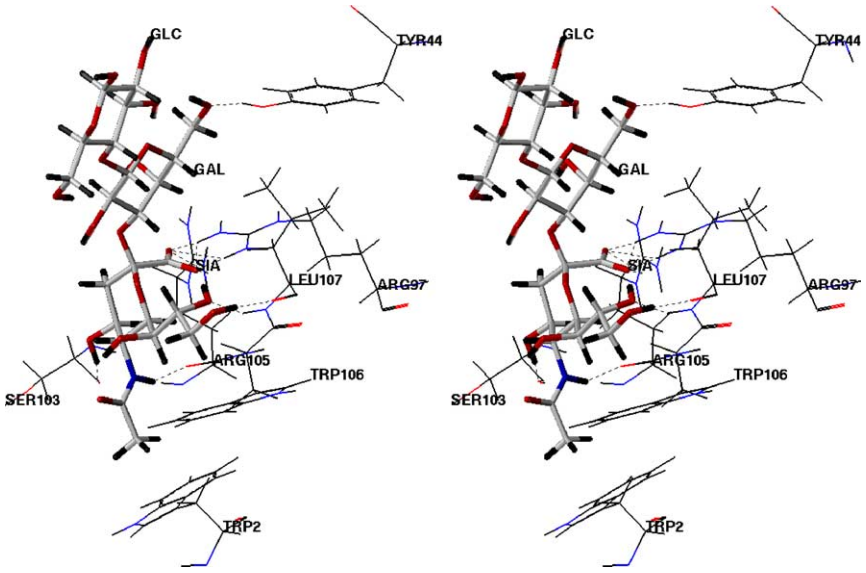


Figure 4. Stereo picture (relaxed view) of sialyl lactose **1** in the binding pocket of sialoadhesin showing the hydrogen bonds between protein and trisaccharide obtained from the FlexiDock calculations. The ligand is represented in stick mode whereas the amino acids are represented by lines. Dotted lines denote hydrogen bonds between the ligand and sialoadhesin.

Table 1. The dihedral angles (ϕ and ψ) at the glycosidic linkages were defined as ϕ, C1′–C2′–O2–C3 (Neu5Ac–Gal); ψ, C2′–O2′–C3–H3 (Neu5Ac–Gal) and ϕ, H1′–C1′–O4–C4 (Gal–Glc); ψ, C1′–O1′–C4–H4 (Gal–Glc)

	ϕ(Neu5Ac–Gal)	ψ(Neu5Ac–Gal)	ϕ(Gal–Glc)	ψ(Gal–Glc)
X-ray	−70	−18	34	−11
FlexiDock	−68	−19	49	−43

Table 2. The torsional angles (ω) for side chains were defined as ω(C7), C8–C7–C6–O6 (Neu5AC); ω(C8), C9–C8–C7–C6 (Neu5AC); ω(C9), O9–C9–C8–C7 (Neu5AC); ω(C6, Gal), O6–C6–C5–O5 (Gal); ω(C6, Glc), O6–C6–C5–O5 (Glc)

	$\omega(C7^{Neu5Ac})$	$\omega(C8^{Neu5Ac})$	$\omega(C9^{Neu5Ac})$	$\omega(C6^{Gal})$	$\omega(C6^{Glc})$
X-ray	−56	−173	−176	77	69
FlexiDock	−54	−164	179	60	174

with the values from the crystal structure. The global root mean square deviation (RMSD) of the FlexiDock model compared to the crystal structure was 0.75 Å.

The FlexiDock results (Fig. 4) for the complex of sialyl lactose with sialoadhesin lead to a model that was very similar to the crystal structure. The carboxy function of *N*-acetyl neuraminic acid forms a salt bridge with the guanidino group of Arg 97. The acetamido methyl group has van der Waal contacts with the indole ring of Trp 2, and the terminal carbon of the *N*-acetyl neuraminic acid side chain (C9) makes a hydrophobic contact to the aromatic side chain of Trp 106. The 8- and 9-hydroxyl groups of the *N*-acetyl neuraminic acid side chain formed hydrogen bonds with the main chain amide carbonyl of Leu 107 (~1.8 and ~1.5 Å, respectively). The amido nitrogen of the *N*-acetyl group of the *N*-acetyl neuraminic acid formed a hydrogen bond with the main chain carbonyl of Arg 105, and the 4-hydroxyl group interacted with the main chain carbonyl function of Ser 103. The galactose formed only one hydrogen bond with the protein between the 6-hydroxyl and the hydroxyl group of Tyr 44. The H5′ made van der Waal contacts with Leu 107. Therefore, the key interaction for the sialoadhesin was predicted to be a hydrophobic interaction between the aromatic ring of Trp 2 and with the *N*-acetyl methyl group of *N*-acetyl neuraminic acid on the basis of X-ray structure and FlexiDock modeling.

3. Discussion

In general, it is very important to know the size and the shape of the binding epitope of a ligand bound to a receptor protein. A quantitative analysis of STD effects allowed us to characterize the binding epitope of sialyl lactose **1** bound to sialoadhesin at atomic resolution. Here, we compared the experimental STD effects to theoretical data predicted from the crystal structure of the complex. A docking study performed on the basis of this crystal structure using the FlexiDock algorithm

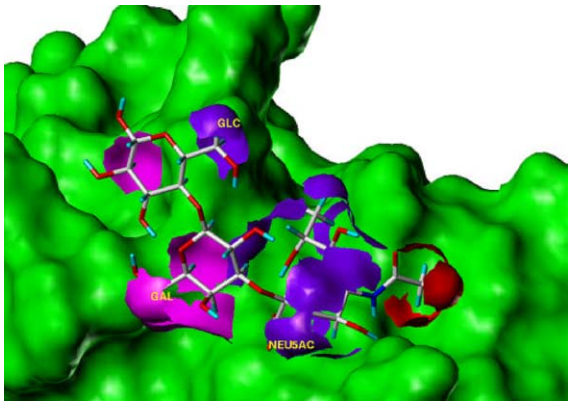


Figure 5. Experimental STD effects shown in the X-ray structure of sialyl lactose **1** bound to sialoadhesin. The green MOLCAD surface shows the surface of the protein. Differently colored MOLCAD surfaces are used to differentiate the size of STD effects. A red color indicates strong STD effects whereas purple and magenta show the medium and small STD effects, respectively.

substantiated the assumption that the overall conformation of the complex is very similar in the crystal and in aqueous solution.

The experimentally determined STD effects (cf. Figs. 1 and 2) have been used to illustrate in Figure 5 the dependence of the degree of saturation on the distance between ligand and protein protons. A color coding has been employed to illustrate the results of the group epitope mapping. It is immediately clear that those protons of sialyl lactose that are close to protons in the binding pocket of sialoadhesin show larger STD effects. The strongest STD effect was observed for the *N*-acetyl methyl group of sialyl lactose **1**. This is in excellent agreement with the X-ray structure of sialyl lactose **1** bound to sialoadhesin[12] where a strong hydrophobic interaction between Trp 2 of sialoadhesin and the *N*-acetyl group of sialyl lactose was observed.

The dramatic effect on the binding affinity of sialoadhesin for sialyl lactose upon mutation of Trp 2[12,17] already indicated that this interaction made a significant contribution to the binding energy. All of the protons of the *N*-acetyl neuraminic acid residue gave a response in the STD spectrum (Figs. 1 and 5), indicating that *N*-acetyl neuraminic acid is in close contact with the protein surface. From a comparison with the crystal structure it is seen that the H4″ and H9″ protons are oriented toward the protein interior with the corresponding hydroxyl functions OH4 and OH9 hydrogen bonded to Ser 103 and Leu 107. In particular, H9″ showed a hydrophobic interaction with Trp 106.

The protons of the galactose residue showed smaller STD effects (Figs. 1 and 5) indicating that this residue is in less intimate contact with the protein surface. This finding is also in good agreement with data from site-directed mutagenesis studies[12] of sialoadhesin. Mutation of Tyr 44 does not significantly affect adhesion[12] of sialoadhesin to sialyl lactose. In the crystal structure a hydrogen bond between the 6-hydroxyl group of galactose and Tyr 44 of sialoadhesin was observed. From the mutagenesis results it is concluded that this hydrogen bond is not critically important for the recognition reaction. All of the β-D-glucose protons, except for the protons H6 and H2 did not show any STD response. This suggests that the glucose moiety is not critical for the binding process.

From the CORCEMA-STD predictions based on the crystallographic structure, all protons of *N*-acetyl neuraminic acid display large STD values indicating that it is in intimate contact with the sialoadhesin. The H3′, H4′, H5′, and H6′ of galactose show smaller STD values. The predicted STD values for all the protons on glucose are negligibly small, illustrating that this residue is remote from the protein surface. All these predictions are in general agreement with the experimental data. If we compare the calculated and experimental STD values of individual protons, the calculated values are significantly smaller than the experimental values for the *N*-acetyl methyl group of *N*-acetyl neuraminic acid and for H6 of glucose. This is not surprising since in the crystal structure the *N*-acetyl methyl group is too far from the protein methyl groups to experience a direct saturation transfer. The crystal structure also predicts a somewhat larger STD value for H4″ than the experimental value. The *R*-factors (0.57–0.67) calculated at different protein/ligand ratios are in general indicative of a moderate overall fit between experimental and calculated values. Exclusion of STD values for the *N*-acetyl methyl group and H6 of glucose still resulted in *R*-factors ranging from 0.4 (at 1:300) to 0.49 (at 1:500) leading to a slightly improved fit. Taking into account an estimated experimental error of 10–20% for the STD data presented here the data suggest an overall correspondence between the crystal structure and the solution state. At the same time our data suggest that the crystallographic structure of the complex may not be entirely compatible with the structure in solution in the fine details. Therefore, our ongoing research focuses on the generation of a larger collection of experimental data, also including other protein–carbohydrate complexes, in order to better identify similarities and deviations between crystal structures and solution state conformations of protein–carbohydrate interactions.

4. Experimental

4.1. Nuclear magnetic resonance (NMR) spectroscopy

The expression and purification of N-terminal V-set immunoglobulin-like domain of sialoadhesin was done according to a published procedure.[9] Protein concentrations in the NMR samples were determined using UV

absorbance at 280 nm with extinction coefficient 1.43 M^{-1} cm^{-1}. The exchangeable protons of the protein were exchanged with a deuterated Tris buffer (pH* 8.0) by repeated washing using a micro-concentrator with a 10 kDa molecular weight exclusion limit (Sartorius, Germany). α-2,3-Sialyl lactose **1** (Dextra Laboratories, Germany)[21] was prepared by lyophilization from 99.9% D_2O. The lyophilization resuspension procedure was repeated three times. All NMR experiments were performed on a Bruker Avance DRX 500 spectrometer at 303 K using the DPFGSE method[15] for water suppression. For 1D STD NMR spectra, the spectral width was 10 ppm and all spectra were acquired using digital quadrature detection. Selective irradiation of the protein was achieved by a train of Gaussian shaped pulses with a 1% truncation and each of 50 ms in duration and separated by a 1 ms delay. The minimum irradiation time required to achieve the maximum signal to noise ratio in the STD spectra, in a given total experimental time, was determined by performing a series of difference experiments with the irradiation times 0.25, 0.5, 1, 2, 2.5, 3.75, and 5 s. The maximum achievable signal to noise ratio, in the given experimental time was found to occur with an irradiation time of 2 s, and this value was used for all subsequent experiments. The protein was irradiated at 0 ppm (on-resonance) and at 40 ppm (off-resonance). The total numbers of scans were 1600 for the reference spectra and 3200 for the difference spectra. Data processing was performed using the XWINNMR program suite (Bruker).

For STD NMR experiments, the following control experiments have been performed. Firstly, the frequencies were chosen such that the maximum irradiation of the protein signals occurs whilst at the same time avoiding interference with ligand signals. Secondly, the difference spectroscopy was performed via phase cycling and minimal subtraction artifacts were expected. This was experimentally demonstrated by performing the difference experiment with the on- and off-resonance frequencies set to the same values, in the case of this study 0 and 40 ppm. The resulting difference spectra contained no signals. Lastly, the sample containing only the ligand molecule was subjected to equivalent difference experiments as the samples containing the protein. The resulting difference spectra did not contain any signals for low concentration of the ligand. At higher concentrations of the ligand a correction is required because spurious effects were observed in the STD spectra of the ligand only. The correction was achieved by subtracting the effects from the final STD spectrum of the ligand with protein.

4.2. Computational methods

The FlexiDock algorithm was used to generate the computational docking model. All calculations were performed on a Silicon Graphics Octane (2xR12K) workstation using the SYBYL 6.7.2 program suite (Tripos, USA).[22] A crystal structure[12] of sialoadhesin co-crystallized with sialyl lactose (at 1.85 Å resolution) was obtained from Brookhaven Protein Data Bank (Acquisition code 1QFO). The crystal structure showed three asymmetric subunits A, B, and C and we used here subunit B for all subsequent computational modeling due to its high electron density and minimal temperature factors.

The genetic algorithm based FlexiDock provided a means of docking ligands into protein active sites keeping bond lengths and angles constant. FlexiDock was performed using a hydrogen van der Waal's radius of 1.0 Å, ε of 0.03, and cutoff distance for nonbonding interactions of 16 Å between atoms. The protein was fixed in space whilst the position and all torsion angles of the ligand were flexible. Water was removed during the calculation of FlexiDock. The ligand was initially positioned in the active site and then FlexiDock computation was performed. The entire complex of the protein and ligand was minimized in 30 steps among them 3000 population were run in 10 steps, 100,000 population in 15 steps, and 10^6 populations were run in five steps. The best structure of each run was set for the starting structure of the next run. The docking process was stopped when there was no further decrease of binding energy.

4.3. CORCEMA-STD calculations

A computer routine called CORCEMA-STD was developed for use in MATLAB to allow the more detailed analysis of STD data. The underlying theory of CORCEMA-STD has been recently described.[20] This program calculates the expected STD NMR effects for a reversibly forming ligand–receptor complex for any proposed molecular model and other parameters such as the identity of the saturated protein proton(s), correlation times, exchange rates, and spectrometer frequency. For the ideal case where there is infinite delay between each scan, the magnetizations in an STD experiment are given by

$$\mathbf{I}(t) = \mathbf{I}_0 + [1 - \exp\{-\mathbf{D}t\}]\mathbf{D}^{-1}\mathbf{Q}, \qquad (1)$$

where $\mathbf{I}(t)$ is a column matrix containing the magnetizations for the ligand and for those protein protons that do not experience a direct RF saturation. **Q** is a column matrix containing cross-relaxation terms between the protein protons that experience a direct RF saturation and the rest of the protons. The dynamic matrix **D** is a square matrix and is a sum of the relaxation rate matrix **R** and the exchange matrix **K**. These have been described in detail recently.[20] '*t*' is the time period for which

the protein proton(s) experience RF irradiation. The CORCEMA-STD program also has a provision for taking into account the effect of finite delays (t_d) between scans in calculating the STD effects,[20] and this finite delay was taken into account in the current analysis.

To analyze the data of the sialyl lactose–sialoadhesin complex, we used a two-state model involving free and bound states of the interacting species.[18,19] The read-in parameters are the number of ligand protons (N) and the protein protons (M) near the active site to be included in the CORCEMA-STD calculations, the number J of protein protons that experience direct RF irradiation and their identities, the rotational correlation times, kinetic parameters [equilibrium constant for the complex K_{eq} ($= 1/K_D$ where K_D is the dissociation constant), and k_{on} or k_{off} rates], the leakage relaxation rate of the individual protons to account for nonspecific relaxation, and the spectrometer frequency. To account for the effect of internal motions of the methyl groups, the corresponding spectral densities are calculated using the model-free formalism.[23] The order parameter S^2 and the internal correlation time for the methyl group are also read-in parameters. For intra-methyl relaxation, S^2 was set to 0.25[24] while for methyl–X relaxation S^2 is generally kept in the range of 0.6–0.9. For Tyr and Phe, a simple $\langle 1/r^6 \rangle$ average was used for the dipolar relaxation between the aromatic and other protons. Details of the CORCEMA-STD protocol have been described previously.[20]

Calculations were performed using the crystallographic structure of the sialyl lactose–sialoadhesin complex, PDB crystal structure entry 1QFO.[12] The crystal structure consisted of three copies (the A, B, and C chains) of the protein–ligand complex. However, the electron density for the ligand in the C-chain is rather poorly defined and hence its structure could not be used. After the addition of the hydrogens to the crystallographic structure using QUANTA, a bad contact between Arg 105 and Sia 201 protons in the A chain was observed, and hence this chain also was not used. Thus all the CORCEMA calculations were performed using only the B chain. In their uncomplexed states, for simplicity in calculations the protein and the ligand were assumed to retain the same conformation as in the complex. The program however allows the interacting molecules to have different conformations, in the free and bound states. To speed up the computation of the **R** matrix, spectral densities were calculated for only those proton pairs having a distance of 10 Å or less. In the calculations three sugars [Sia 201 (Neu5Ac), Gal 202, Glc 203] and the 20 amino acid residues within the binding pocket (Trp 2, Thr 37, Ala 38, Ile 39, Tyr 41, Tyr 44, Ser 45, Arg 48, Arg 97, Phe 98, Glu 99, Ile 100, Ser 101, Ser 103, Asn 104, Arg 105, Trp 106, Leu 107, Asp 108, and Val 109) were included.

Since the protein signals at 0 ppm were irradiated for the STD experiment, we made the reasonable assumption that the Leu, Val, and Ile methyl protons were instantaneously saturated. The saturation will take finite time to spread to other protein and ligand protons (bound and free) through dipolar networks and chemical exchange. From the intensity matrix $\mathbf{I}(t)$, the fractional intensity changes $[(I_{0(k)} - I(t)_{(k)})/I_{0(k)}]$ for different ligand protons k are calculated, and compared to the experimental STD values using an NOE R-factor defined as[25,26]

$$R\text{-factor} = \sqrt{\frac{\sum (S_{\text{expt},k} - S_{\text{calc},k})^2}{\sum (S_{\text{expt},k})^2}}. \qquad (2)$$

In these equations $S_{\text{expt},k}$ and $S_{\text{calc},k}$ refer to experimental and calculated STD values for proton k.

The Multilevel Coordinate Search (MCS) method was used to optimize parameters to get best fit between the experimental and predicted intensities. The NOE R-factor was used as the energy function to be minimized. A version of the MCS method was written based on the version presented by Huyer and Neumaier.[27] The algorithm performs the minimization by a standard coordinate search method. The method was carefully tested by the use of different starting points for the coordinate search. This alleviates local minima trapping by MCS, and identifies the global minimum within the parameter ranges used in the optimization.

The dissociation constant (K_D) of the sialyl lactose–sialoadhesin complex was reported as 0.8 mM from the changes in chemical shift of sugar resonances and 1.4 ± 0.4 mM from the changes in chemical shift of protein resonances.[17] Thus the calculations were performed by restricting the K_D values to the range of 0.8–1.4 mM. The methyl group internal correlation times (τ_m) was set at a reasonable value of 5 ps.[28] The k_{on} was set to $10^8\ s^{-1}\ M^{-1}$ and the leakage factor was set at $0.1\ s^{-1}$. For optimization of the parameters the following ranges were employed initially on the 300:1 ratio data set: $\tau_{\text{r-protein}}$ (10^{-8}–10^{-7} s), $\tau_{\text{r-ligand}}$ (0.1–2 ns), external methyl S^2 (0.6–0.85), and K_D (0.8–1.4 mM) at 1:300 ratio of protein (5 μM)/ligand. The calculations were repeated with the narrow ranges $\tau_{\text{r-protein}}$ (2×10^{-8}–6×10^{-8} s), $\tau_{\text{r-ligand}}$ (1–2 ns), external methyl S^2 (0.7–0.85), and K_D (0.8–1.4 mM) for the same ratio of protein to ligand and found that it converges to the same global minimum. To speed up the computational time, we restricted further calculations with the narrow ranges of parameters for the remaining protein/ligand ratios. In all the calculations the protein concentration was kept fixed at 5 μM. The predicted STDs in Figure 4 were calculated using the following parameters corresponding to global minimum: $\tau_{\text{r-ligand}} = 2$ ns, $\tau_{\text{r-protein}} = 3 \times 10^{-8}$ s, $S_1^2 = 0.85$, and $K_D = 1.4$ mM.

Acknowledgements

We wish to thank the VW foundation for generous support (VW Project Program Grant 'Conformational Control of Biomolecular Function'). We would also like to thank the University of Luebeck for funding extra equipment for the 500 MHz NMR spectrometer. In addition, partial support of this work through NCI grants CA-13148 (NMR core facility), and CA-84177 is also acknowledged.

References

1. Drickamer, K. *Curr. Opin. Struct. Biol.* **1995**, *5*, 612–616.
2. Kelm, S.; Pelz, A.; Schauer, R.; Filbin, M.; Tang, S.; Bellard, M.-E. D.; Schnarr, R. I.; Mahoney, J. A.; Hartnell, A.; Bradfield, P.; Crocker, P. R. *Curr. Biol.* **1994**, *4*, 965–972.
3. Crocker, P. R.; Clark, E. A.; Filbin, M.; Gordon, S.; Jones, E. Y.; Kehrl, J. H.; Kelm, S.; Le Douarin, N.; Powel, L.; Roder, J. *Glycobiology* **1998**, *8*, v–vi.
4. Crocker, P. R. *Curr. Opin. Struct. Biol.* **2002**, *12*, 609–615.
5. McKerracher, I.; David, S.; Jackson, D. L.; Kottis, V.; Dunn, R. J.; Braun, P. E. *Neuron* **1994**, *13*, 805–811.
6. Li, M.; Shibata, A.; Li, C. M.; Braun, P. E.; Mckerracher, L.; Roder, J.; Kater, S. B.; David, S. *J. Neurosci. Res.* **1996**, *46*, 404–414.
7. Nitschke, L.; Carsetti, R.; Ocker, B.; Köhler, G.; Lamers, M. C. *Curr. Biol.* **1997**, *7*, 133–143.
8. Crocker, P. R.; Freeman, S.; Gordon, S.; Kelm, S. *J. Clin. Invest.* **1995**, *95*, 635–645.
9. Crocker, P. R.; Werb, Z.; Gordon, S.; Bainton, D. F. *Blood* **1991**, *76*, 1131–1138.
10. McWilliam, A. S.; Tree, P.; Gordon, S. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10522–10526.
11. Kelm, S.; Schauer, R.; Manuguerra, J. C.; Gross, H. J.; Crocker, P. R. *Glycoconj. J.* **1994**, *11*, 576–585.
12. May, A. P.; Robinson, R. C.; Vinson, M.; Crocker, P. R.; Jones, E. Y. *Mol. Cell* **1998**, *1*, 719–728.
13. Meyer, B.; Peters, T. *Angew. Chem., Int. Ed.* **2003**, *42*, 864–890; *Angew. Chem., Int. Ed.* **2003**, *115* 890–918.
14. Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784–1788.
15. Stott, K.; Stonehouse, J.; Keeler, J.; Hwang, T.-S.; Shaka, A. J. *J. Am. Chem. Soc.* **1995**, *117*, 4199–4200.
16. Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108–6117.
17. Crocker, P. R.; Vinson, M.; Kelm, S.; Drickamer, K. *Biochem. J.* **1999**, *341*, 355–361.
18. Moseley, H. N. B.; Curto, E. V.; Krishna, N. R. *J. Magn. Reson. B* **1995**, *108*, 243–261.
19. Krishna, N. R.; Moseley, H. N. B. Complete Relaxation and Conformational Exchange Matrix (CORCEMA) Analysis of NOESY Spectra of Reversibly Forming Ligand–Receptor Complexes. In: *Biological Magnetic Resonance; Structure Computation and Dynamics in Protein NMR*; Kluwer Academic/Plenum: New York, 1999; Vol. 17; pp 223–310.
20. Jayalakshmi, V.; Krishna, N. R. *J. Magn. Reson.* **2002**, *155*, 106–118.
21. Dorland, L.; van Halbeek, H.; Vliegenthart, J. F. G.; Schauer, R.; Wiegandt, H. *Carbohydr. Res.* **1986**, *151*, 233–245.
22. Sybyl molecular modeling software, version 6.7, Tripos Associates, 1699 South Hanely Road, St. Louis, MO 63144-2917.
23. Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.
24. Dellwo, M. J.; Wand, A. J. *J. Am. Chem. Soc.* **1993**, *115*, 1886–1893.
25. Krishna, N. R.; Agresti, D. G.; Glickson, J. D.; Water, R. *Biophys. J.* **1978**, *24*, 791–814.
26. Xu, Y.; Sugar, I. P.; Krishna, N. R. *J. Biomol. NMR* **1995**, *3*, 361–365.
27. Huyer, W.; Neumaier, A. *J. Global Optim.* **1999**, *14*, 331–335.
28. Palmer, A. G.; Case, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 9059–9067.