Carbohydrate Research 344 (2009) 1515–1525

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

A lectin from the Chinese bird-hunting spider binds sialic acids

Hans-Christian Siebert [a], Shan-Yun Lu [b], Rainer Wechselberger [c], Karin Born [d], Thomas Eckert [a], Songping Liang [b,e], Claus-Wilhelm von der Lieth [d,✠], Jesús Jiménez-Barbero [f], Roland Schauer [g], Johannes F. G. Vliegenthart [h], Thomas Lütteke [a], Tibor Kožár [i,*]

[a] Institut für Biochemie und Endokrinologie, Fachbereich für Veterinärmedizin, Justus-Liebig Universität Gießen, Frankfurter Straße 100, 35392 Gießen, Germany
[b] College of Life Sciences, Beijing University, Beijing 100871, China
[c] Department of NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands
[d] Zentrale Spektroskopie, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
[e] College of Life Sciences, Hunan Normal University, Changsha 410081, China
[f] Centro de Investigaciones Biólogicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain
[g] Biochemisches Institut, Universität Kiel, Olshausenstr. 40, D-24098 Kiel, Germany
[h] Department of Bio-Organic Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands
[i] Slovak Academy of Sciences, Institute of Experimental Physics, Department of Biophysics, 04001 Košice, Slovakia

ARTICLE INFO

Article history:
Received 16 March 2009
Received in revised form 30 May 2009
Accepted 2 June 2009
Available online 6 June 2009

Dedicated to our colleague Hans Kamerling on the occasion of his 65th birthday

Keywords:
Sialic acid
Lectin
Carbohydrate–protein interaction
Molecular modeling
NMR analysis

ABSTRACT

The affinity to sialic acid-containing oligosaccharides of the small-animal lectin SHL-I isolated from the venom of the Chinese bird-hunting spider *Selenocosmia huwena* is here described for the first time. By a strategic combination of NMR techniques, molecular modeling, and data mining tools it was possible to identify the crucial amino acid residues that are responsible for SHL-I's ability to bind sialic acid residues in a specific way. Furthermore, we are able to discuss the role of the functional groups of sialic acid when bound to SHL-I. Also the impact of Pro31 in its *cis*- or *trans*-form on SHL-I's ligand affinity is of special interest, since it answers the question if Trp32 is a crucial amino acid for stabilizing complexes between SHL-I and sialic acid. SHL-I can be considered as a proper model system that provides further insights into the binding mechanisms of small-animal lectins to sialic acid on a sub-molecular level.

1. Introduction

The *Selenocosmia huwena* lectin-1 (SHL-I), isolated from the venom of the Chinese bird-hunting spider *S. huwena*, agglutinates human and mice erythrocytes at a minimum concentration of 125 and 31 mg/mL, respectively. It was first described that mannosamine inhibits the agglutination activity.[1] The affinity to other saccharides was still unknown. The lectin SHL-I is one of the smallest occurring lectins and can therefore be referred to as a mini-lectin. It consists of 32 amino acid residues with three disulfide bridges. The three disulfide bridges of SHL-I correspond to the linkage form of HWTX-I, the 33-amino acid-containing neurotoxin from *S. huwena*.[2] In contrary to the neurotoxin HWTX-I, the lectin SHL-I shows a very low toxicity in mammals as well as in insects. Further toxic peptides from the venom of the Chinese bird-hunting spider *S. huwena* have been described.[3–7] Structural similarities and differences between these toxic spider peptides (especially HWTX-I) and SHL-I are discussed in this paper in order to learn which arrangement of amino acids is essential for SHL-I's lectin function. Structural similarities to plant lectins are found when comparing SHL-I with small hevein-like lectins. The structures of hevein and its isoform pseudo-hevein (both are *iso*-lectins from the *Hevea brasiliensis* latex) have been determined by NMR methods in the absence and in the presence of their specific binding ligands.[8–13] Three of the four disulfide bonds present in hevein and pseudo-hevein also show a 1-4, 2-5, 3-6 disulfide bridge pattern congruent to SHL-I and HWTX-I. The carbohydrate specificity of SHL-I was unclear until the binding affinity of *N*-acetylneuraminic acid to SHL-I could successfully be detected by NMR techniques and further confirmed by computational methods. Here, a strategic combination of laser photo CIDNP (Chemically Induced Nuclear Polarization) NMR methods, molecular modeling tools as well as ab initio calculations[14–17] provides a suitable way to successfully analyze the interactions between SHL-I and Neu5Ac,

* Corresponding author. Tel.: +421 55 720 4137; fax: +421 55 633 6292.
E-mail address: tibor@saske.sk (T. Kožár).
✠ Deceased, November 16, 2007.

0008-6215/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2009.06.002

Kdn, Neu5Acα2→3Galβ1→4Glc as well as Neu5Acα2→6-Galβ1→4Glc. A structural matter of interest is the *cis*- or *trans*-constitution of Pro31 since it must be considered as a factor that is potentially of great importance for the carbohydrate specificity and affinity of SHL-I.

2. Results and discussion

The differences between the amino acid sequences of the toxin (HWTX-I) and the lectin (SHL-I) in the venom of the Chinese bird-hunting spider *S. huwena* are shown in Figure 1.

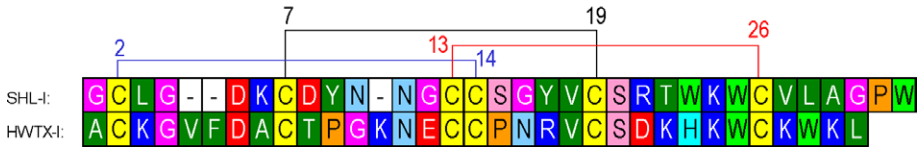

Figure 1. Amino acid sequences of the lectin (SHL-I) and the toxin (HWTX-I) in the venom of the Chinese bird-hunting spider *Selenocosmia huwena*, respectively.

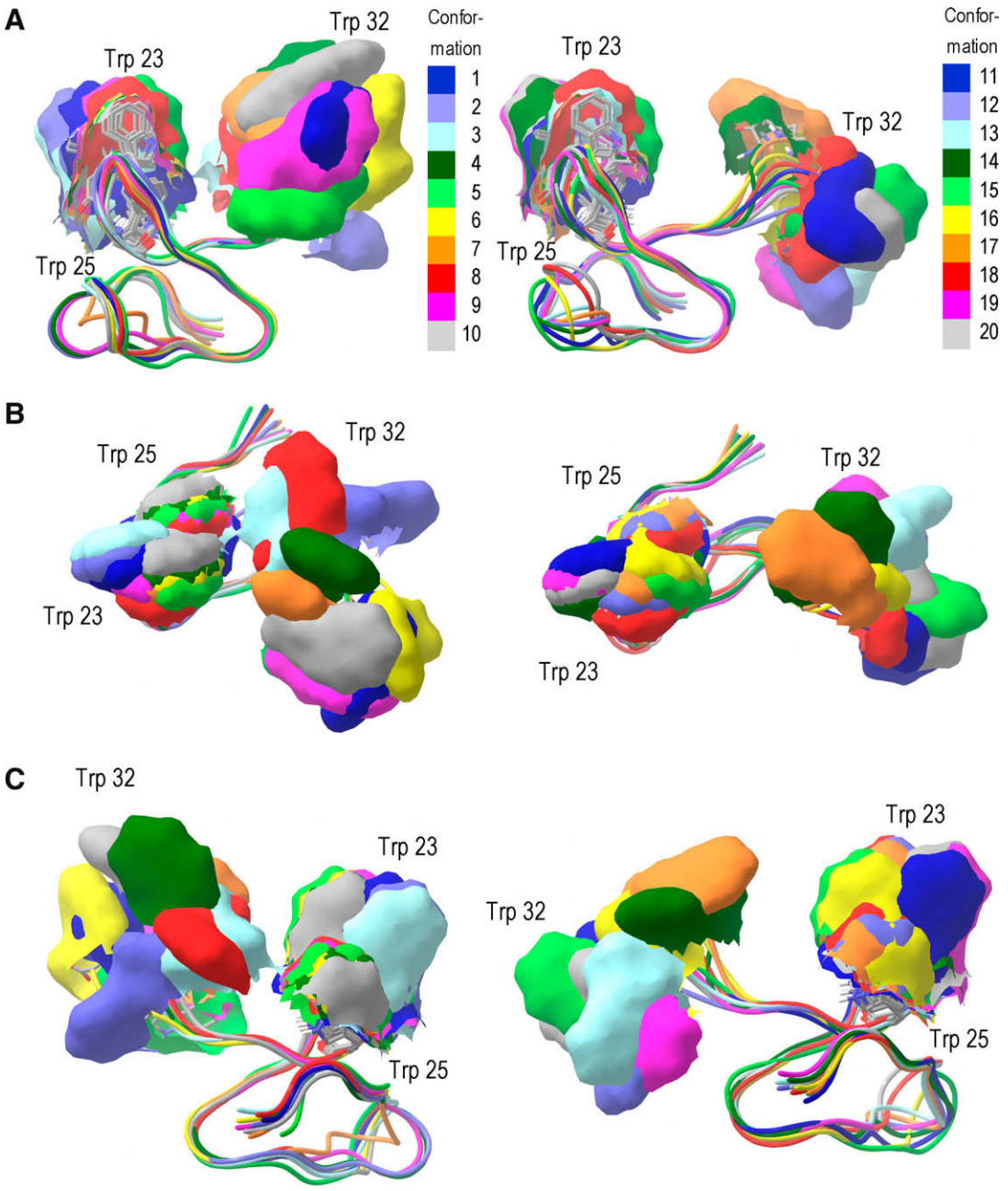

Figure 2. Conformational flexibility of SHL-I based on NMR data (PDB code 1QK7) illustrating influence of the trans (on the left; conformations 1–10) and cis (on the right; conformations 11–20) conformations of Pro31 on the orientation of Trp32. Only the surfaces of the Trp residues are visualized, illustrating the remarkable spatial distribution of Trp32. (A) front, (B) top, and (C) back view.

These two proteins with a high sequence homology differ remarkably in their biological function since HWTX-I is a neurotoxin and since the non-toxic SHL-I is defined by its carbohydrate affinity. SHL-I can be regarded as a suitable role-model that allows us to analyze how a protein becomes a lectin. However, such an analysis can only be carried out when the carbohydrate recognition domain (CRD) of SHL-I is known. The NMR structure of SHL-I in the ligand-free state is solved, and the set of twenty coordinates is stored in the protein data bank under 1QK7.pdb. Figure 2 illustrates the superposition of these coordinates grouped into two assemblies with *trans* and *cis* Pro31. The figure illustrates the conformational flexibility of SHL-I with the highlighted surface accessibility of the Trp residues.

First of all it was of interest which carbohydrates are specific binding partners of SHL-I. As a result of NMR-titration experiments at a pH value of 5.5 and 30 °C we determined the following association constants:

SHL-I + Neu5Acα2→6Galβ1→4Glc: K_{assoc}: 475 ± 126 M^{-1}
SHL-I + Neu5Acα2→3Galβ1→4Glc: K_{assoc}: 486 ± 128 M^{-1}
SHL-I + Neu5Ac: K_{assoc}: 23 ± 14 M^{-1}
SHL-I + mixture of O-acetylated sialic acids: K_{assoc}: 182 ± 79 M^{-1}

The CRD of SHL-I, as already mentioned has not yet been experimentally determined. We thus modeled the possible binding sites of SHL-I using the SITEMAP program.[18] Figure 3 illustrates that the *trans* and *cis* Pro31 CRDs result in different predictions of the binding sites for the different conformational sets of SHL-I.

Since SHL-I has structure similarities with the hevein lectins from the rubber tree that use Tyr and Trp residues for carbohydrate binding, the laser photo CIDNP method can be considered as a feasible experimental strategy to test whether these amino acids play a crucial role in SHL-I's carbohydrate affinity.[8,14,16] As can be seen in Figure 4, Tyr and Trp residues are surface exposed in the SHL-I structure. The degree of surface accessibility of the Trp and Tyr residues

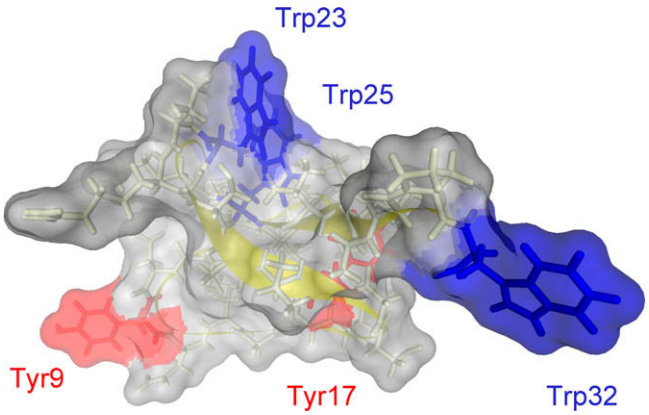

Figure 4. Surface presentation of SHL-I in the ligand-free state which is stored in the protein data bank under 1QK7 (only conformation 15 is visualized). The two Tyr and the three Trp residues are highlighted in red and blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

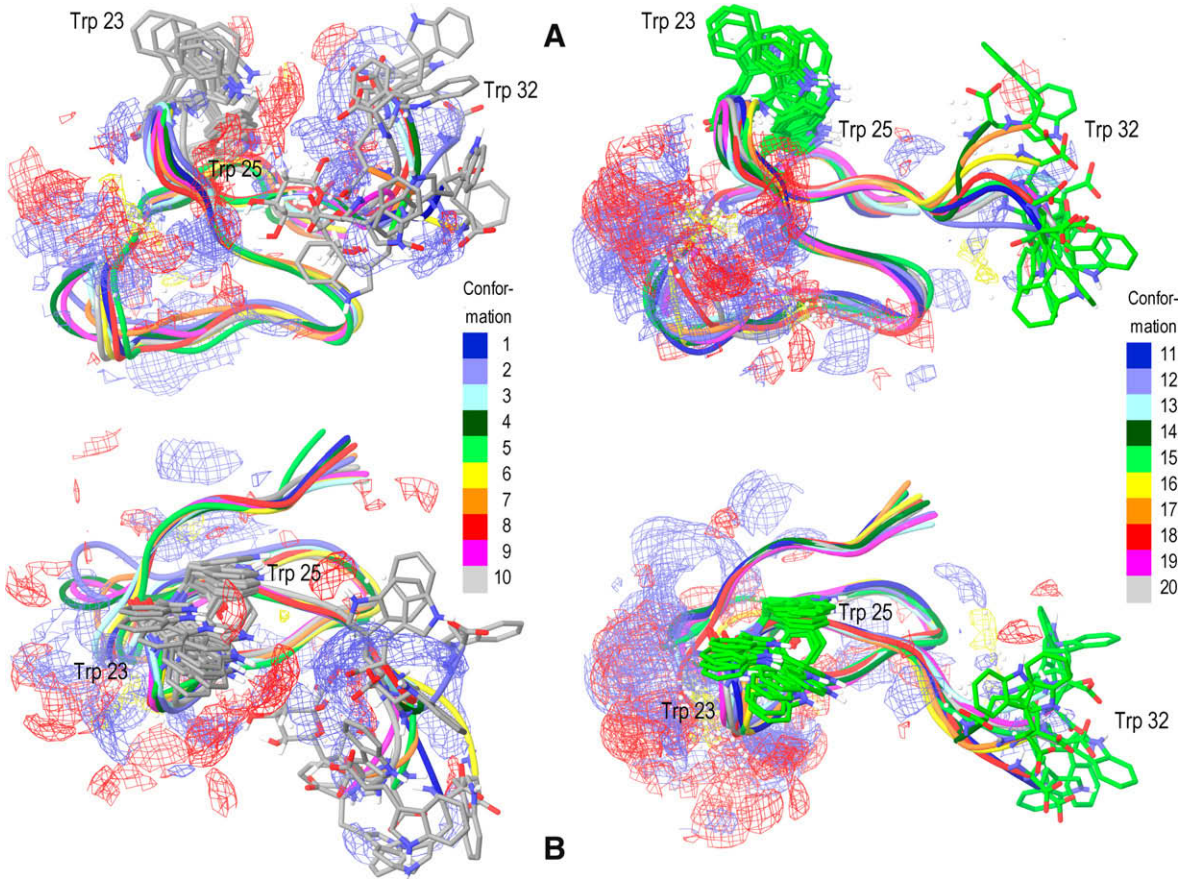

Figure 3. Superposition of SITEMAP (Schrödinger LLc.) predicted binding sites for the two conformational sets of SHL-I, that is, with *trans* Pro31 (on the left; gray carbons for Trp 23, 25, and 32) and for 10 *cis* Pro31 conformations on the right (green carbons). Hydrogen-bond donor maps are represented with blue mesh, hydrogen-bond acceptors with red mesh and the hydrophobic maps with yellow. (A) front view and (B) top view. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

does not change significantly after several semi-empirical minimizations on selected SHL-I conformations (Supplementary data) were carried out with the HYPERCHEM 8 [19] software package using AM1.

The CIDNP difference spectra, which are recorded in the absence and in the presence of sialyllactose (Neu5Acα2→3-Galβ1→4Glc), demonstrate that the intensities of the Trp23, Trp25, and Trp32 signals are suppressed after ligand addition, while the intensities of the Tyr signals are unaffected (Fig. 5A and B). Such results argue in favor of a significant decrease in the surface accessibility values of Trp23, Trp25, and Trp32 when SHL-I is complexed with Neu5Acα2→3Galβ1→4Glc. The corresponding parts of the two-dimensional NMR spectra (2D TOCSY and 2D NOESY) of SHL-I–Neu5Acα2→3Galβ1→4Glc complexes are shown in the Supplementary data.

A possible SHL-I binding position of the Neu5Acα2→3Galβ1 residue, which is in agreement with our CIDNP data, is shown in Figure 6.

In order to test whether the relevant amino acid residues for SHL-I's sialic acid binding affinity are also involved in similar protein–carbohydrate interaction processes, we performed a comprehensive analysis of the corresponding protein–glycan complexes deposited in the Protein Data Bank[20] (PDB at www.pdb.org). The GlyVicinity tool of the Glycosciences.de website (http://www.glycosciences.de/tools/glyvicinity/) was used for this purpose.[21]

A preferred role of Trp residues in the specific interaction process between sialic acid residues and proteins is underlined by the data shown in Table 1. When analyzing the 94 pdb entries of proteins bound to sialic acid or sialic acid-containing oligosaccharides, 112 Trp residues are determined to be involved in the binding of the 241 sialic acid residues. This corresponds to an occurrence of 5.5% of all amino acids in the vicinity to the sialic acids and an overrepresentation of 312.8% in relation to the normal abundance of Trp in proteins.

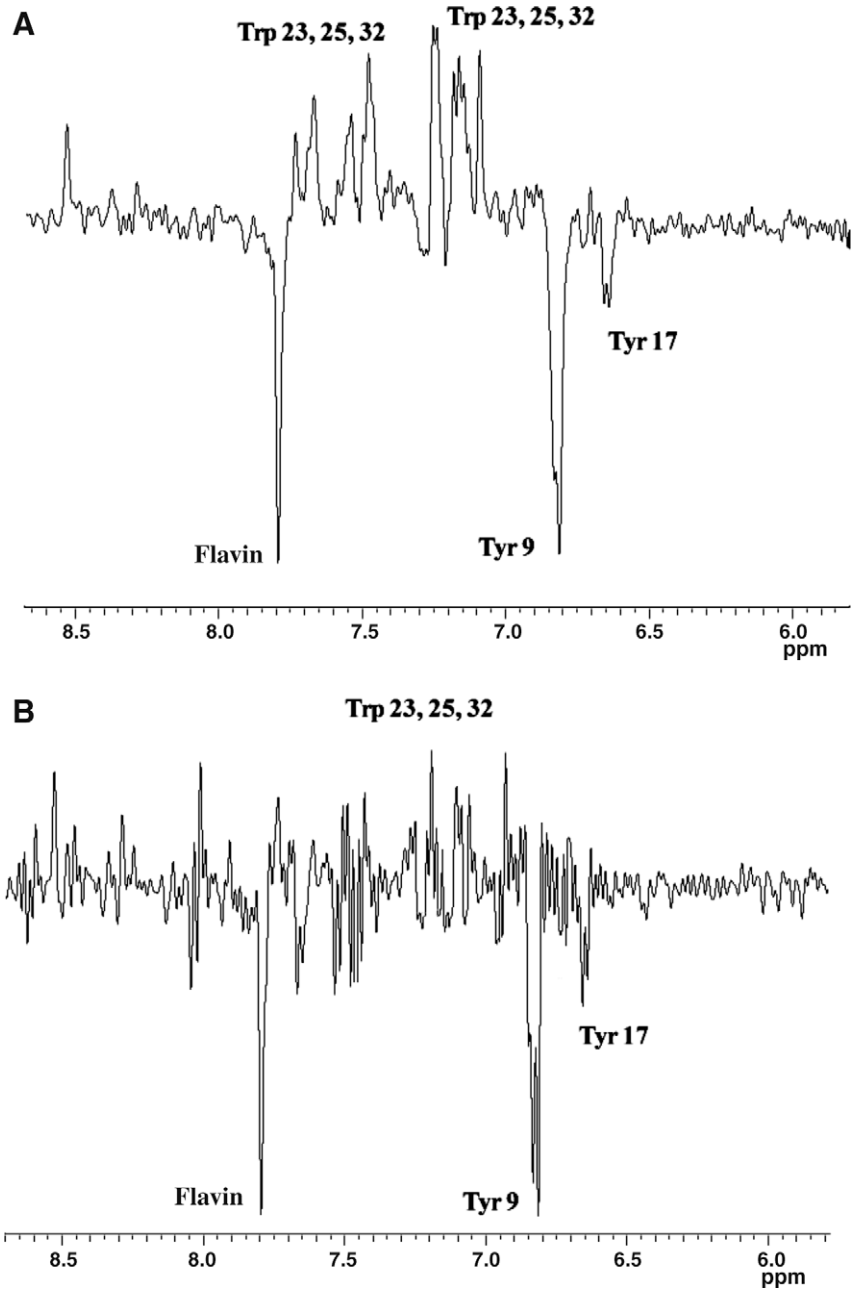

Figure 5. CIDNP difference spectra of SHL-I, recorded in (A) the absence of and (B) the presence of sialyllactose (Neu5Acα2→3Galβ1→4Glc).

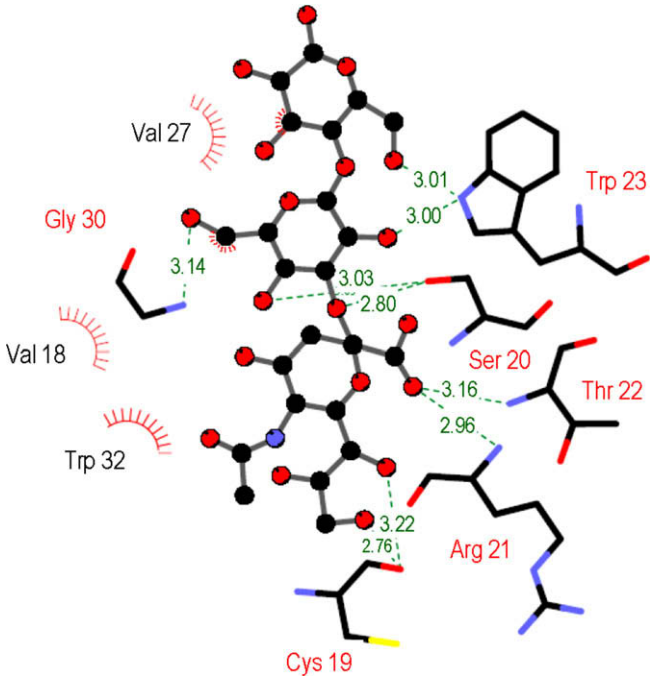

Figure 6. LIGPLOT visualization of the interaction pattern in one of the possible binding modes of the Neu5Acα2→3Galβ1→4Glc residues interacting with SHL-I as resulted from a GLIDE docking study. The amino acids involved in hydrogen bonds are indicated with red; those involved in steric interactions are marked with black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Arg, Ser, and Tyr each form more than 10% of all the amino acids found within a 4 Å radius of sialic acid residues in the PDB and can therefore be considered as amino acid with a potential binding relevance for sialic acid-containing oligosaccharides. These residues mainly form H-bonds with the Neu5Ac residues via their polar side-chain atoms. Tyr, however, can also form CH–π interactions with apolar regions of the sialic acid. This kind of interaction is also observed frequently between Neu5Ac and Trp. CH–π interactions between aromatic amino acids and carbohydrate residues such as glucose and galactose are mainly formed when apolar regions of the carbohydrate rings are stacked in parallel to the aromatic rings. Sialic acid, however, does not have any large apolar regions on its ring. Such regions are found at the methyl group of the *N*-acetyl substituent and at the aliphatic tail. These are the parts of the Neu5Ac residues that are predominantly interacting with the aromatic rings of Trp or Tyr residues.

After the CRD of SHL-I has been discussed in detail it is now of special interest to determine which functional groups of a sialic acid residue are crucial for a specific interaction with the spider lectin SHL-I. Therefore, we have chosen Kdn (2-keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid)[22] besides sialyllactose as a proper test ligand. Kdn is a derivative of Neu5Ac with an OH-group at position C-5 instead of an *N*-acetyl group. Kdn belongs to the sialic acid family, although it has no *N*-acetyl group, which is typical for this class of molecules. The use of Kdn in addition to a Neu5Ac-containing oligosaccharide allows us to clarify in general which contacts are essential for a specific interaction with the spider lectin and answers specially the question concerning the importance of the *N*-acetyl group for a specific interaction with SHL-I. The aromatic/NH parts of the one-dimensional NMR spectra of SHL-I recorded in the absence and in the presence of Kdn show already remarkable differences concerning the chemical shift alterations of the crucial Trp signals (Fig. 7). This result demonstrates that the *N*-acetyl group is not the essential contact point.

However, two important questions concerning the mechanisms of the interactions between SHL-I and sialic acid moieties are still open. First, the influence of the *cis*- or *trans*-position of Pro31 on the lectin's affinity to sialic acids has to be answered in detail, and second, the possibility of different sialic acid binding modes should be discussed. Comparable problems have already been successfully addressed by an arsenal of sophisticated molecular methods so that it is promising to proceed this way.[23–26]

Docking studies were shown particularly important to gain insight into binding properties of SHL-I. When scrutinizing the role of Trp32 in carbohydrate binding, one has to consider that the positioning of Trp32 strongly depends on whether this residue is linked to Pro31 in its *trans* or in its *cis* form (see Fig. 3). Although the SITEMAP prediction for *cis* Pro31 does not correspond to the binding groove *trans* Pro31, we were still interested to find if the docking protocols provide equivalent results to the SITEMAP prediction of CRD. Figure 8 illustrates the results from AUTODOCK binding study for all 20 NMR conformations of SHL-I.

All ligands in the highest-ranked binding mode can bind according to AUTODOCK in relatively close positions for both SHL-I conformational sets, although there is probably a larger flexibility involved for binding with *cis* Pro31. This is remarkably shown in the part A of Figure 8 for Kdn.

In order to confirm the AUTODOCK results, we performed thorough docking runs using the newest version of the GLIDE program. We used the LIGPLOT program consequently in order to compare the resulting binding profiles from the AUTODOCK and GLIDE results. As mentioned within the methods section (Section 3.4), two basic runs were completed for GLIDE, based on (a) the high level ab initio optimized structures and (b) using the best geometries from the AUTODOCK runs. The highest-scoring poses selected from these runs were used for further analysis.

Figure 9 compares the results based on AUTODOCK and GLIDE docking protocols. The GLIDE refinement evidently improved the AUTODOCK results by providing more stabilizing interactions between SHL-I and the bound ligands.

Differences in the binding profiles are remarkably present for almost all the ligands studied. Despite the differences in the binding profiles, it is difficult to rationalize the influence of the *cis*/*trans* orientation of Pro31 on the binding affinities. For example, in the case of Kdn, the bound conformations are better clustered for Pro31 *trans* than in the case of Pro31 *cis*. When the ligand becomes larger as it is the case for Neu5Acα2→3Galβ1→4Glc and Neu5Acα2→6Galβ1→4Glc, the situation is inverted. The number of the lectin's amino acids involved in direct interactions with the trisaccharide ligands is higher.

Table 1
Amino acids present in the vicinity of sialic acid residues in the PDB

Residue	Asp	Glu	Lys	Asn	Arg	Gln	His	Gly	Ser	Ala	Pro	Thr	Tyr	Val	Cys	Met	Trp	Phe	Leu	Ile	Total
Count	68	156	117	150	211	74	68	136	240	57	60	127	213	46	2	8	112	31	103	46	2025
Perc. (%)	3.4	7.7	5.8	7.4	10.4	3.7	3.4	6.7	11.9	2.8	3.0	6.3	10.5	2.3	0.1	0.4	5.5	1.5	5.1	2.3	100
Dev. (%)	−34.8	24.7	−0.4	61.4	101.2	−9.5	49.2	−2.5	61.2	−61.8	−40.9	5.9	224.6	−64.8	−94.4	−83.0	312.8	−62.5	−45.0	−60.1	–

There are 94 PDB entries with a resolution of 3.0 Å or better that contain altogether 241 sialic acid residues in 229 non-covalently bound carbohydrate chains. A total of 2025 amino acids were found within a 4-Å radius of these residues.

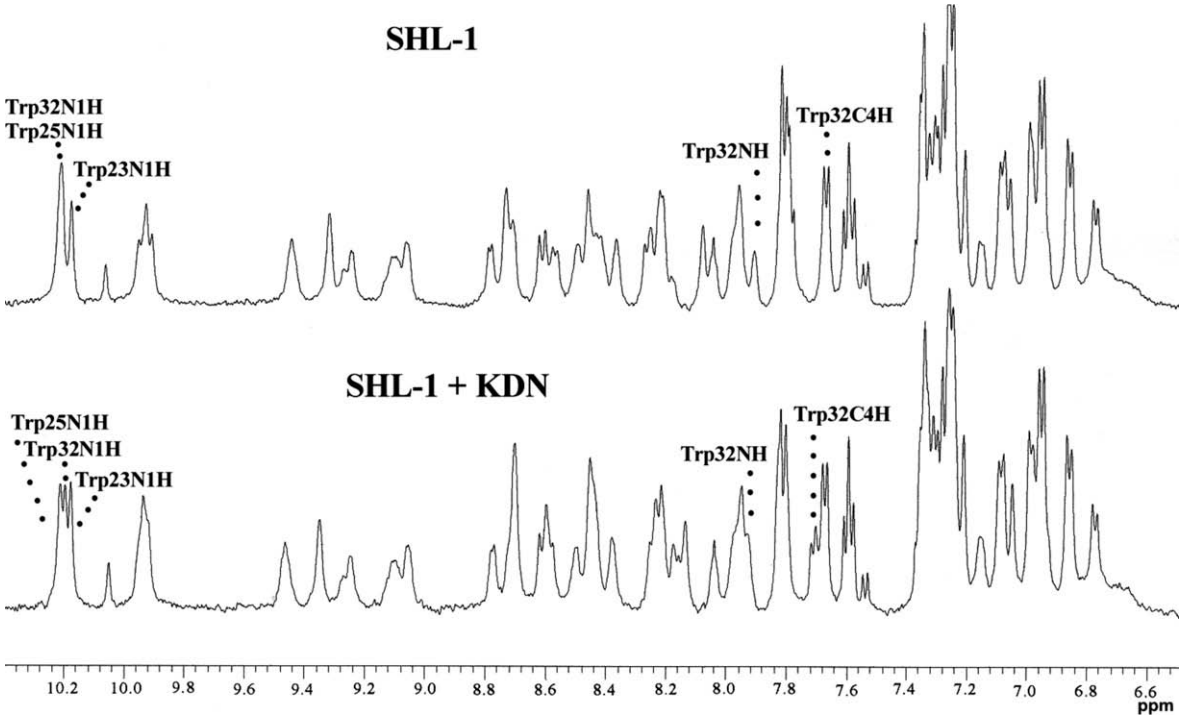

Figure 7. Aromatic and NH part of the one-dimensional NMR spectra of SHL-I, recorded in the absence (top) and in the presence (bottom) of Kdn. The docking-derived model structure shown in Figure 6A is in full agreement with the NMR results depicted in Figure 5.

When considering the binding of the two larger trisaccharides (Neu5Acα2→3Galβ1→4Glc and Neu5Acα2-6Galβ1→4Glc) and comparing their binding behavior in group A (Pro31 *trans*) with that in group B (Pro31 *cis*), then the results argue in favor that there are fewer conformations predicted in the low-energy region for Pro31 *cis* than for Pro31 *trans*.

In order to maintain the ligand at a close-binding position for each conformation of the lectin, the lectin–ligand complex needs an enlargement of the binding energies. The program AUTODOCK, in addition to the binding modes, also predicts the numeric values of the binding affinities. In accordance with these values (visualized also in Figs. 8 and 9) none of the substituted monosaccharides seems to be in a stronger binding mode (AUTODOCK predicted affinities in the micromolar region) in contrary to the trisaccharides. A relatively small number of lectin–ligand interactions were predicted for Kdn and Neu5Ac. More lectin–ligand interactions were exhibited in the case of the larger trisaccharide ligands (as already illustrated in Figs. 8 and 9), resulting for certain lectin conformations into binding within a lower energy binding region.

It turned out from this systematic analysis that Arg21 is also an important binding partner for the tested ligands. The binding mode, in which Arg21 plays a crucial role, is illustrated on a detailed interaction profile in Figure 6.

In order to evaluate the flexibility of Trp32, which exists in the case of Pro31 *cis* as well as in the case of Pro31 *trans*, we performed semi-empirical calculations. These calculations support the assumption that Trp32 is a binding-relevant amino acid residue, since its arrangement in the low-energy conformations favors the induced fit of sialic acid residues into the carbohydrate recognition domain of SHL-I.

When considering the binding of the two larger trisaccharides (Neu5Acα2→3Galβ1→4Glc and Neu5Acα2→6Galβ1→4Glc) and comparing their binding behavior in group A with Pro31 in its *trans* position with that in group B (Pro31 *cis*), then we learn from our modeling data that there are more conformations predicted in the low-energy region for Pro31 *cis* than for Pro31 *trans*.

The GLIDE docking runs resulted in an equivalent finding for better ligand docking in the case of SHL-I conformations with *trans* Pro31 than with *cis* Pro31 conformations. The most advanced QPLD docking protocol resulted in the corresponding binding profiles, preferring more efficient binding with *trans*-oriented Pro31. Figure 10 illustrates the most probable SHL-I–ligand complexes resulting from QPLD docking.

Although the ligand molecules were placed in similar positions (Fig. 11) in relation to those chosen for the other docking programs, FLEXX docking did not demonstrate more efficient scoring for the trisaccharides in comparison to Kdn and Neu5Ac. While AUTODOCK and GLIDE did agree in prediction of higher binding affinity for the larger trisaccharides, in the case of FLEXX this was not confirmed because of the overestimation of the conformational energy contribution.

However, in both groups (A and B) we found several low-energy complexes. Therefore, we can conclude that there could be additional binding modes where Trp32 does not play a crucial role for the sialic acid interaction. In order to identify these binding modes, we systematically analyzed all interactions calculated for the lectin–ligand complexes in the two groups A and B resulting from different docking protocols.

More than one binding mode might be possible for SHL-I as illustrated in Figure 3. The overlay picture of SHL-I with the toxin (HWTX-I) from the venom of the spider (Fig. 12A and B) could thus be misleading without knowing all the other data. It is now possible to conclude that Trp32 has to be considered together with other amino acid residues (e.g., Trp23) as an important contact point responsible for SHL-I's affinity to sialic acids (Neu5Ac, Kdn) or sialic acid-containing oligosaccharides (Neu5Acα2→3Galβ1→4Glc, Neu5Acα2→3Galβ1→4Glc) (Fig. 12C).

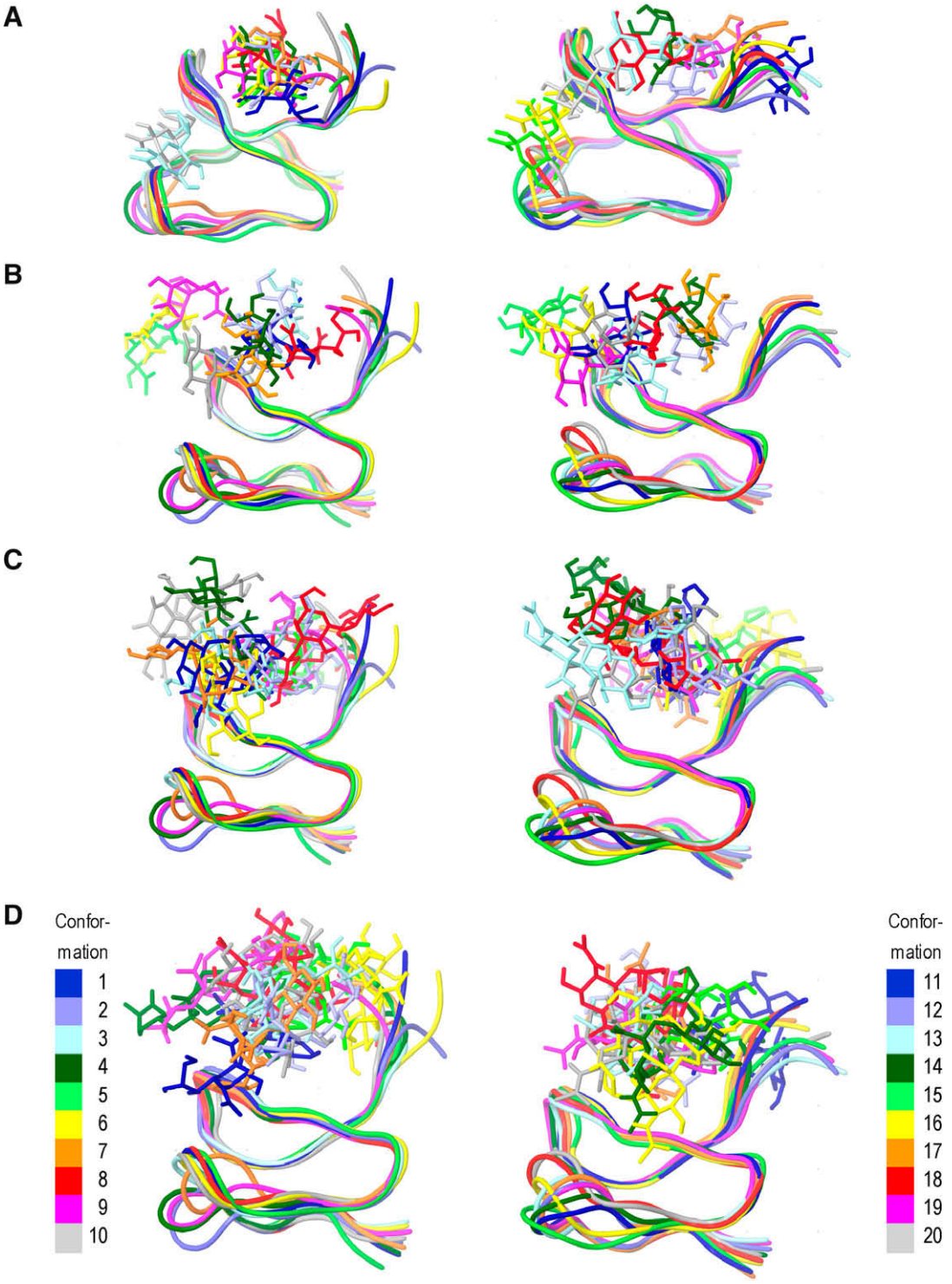

Figure 8. Superposition of AUTODOCK docking results (the highest score only) for the two conformational sets of SHL-I. Ligand assignment: (A) Kdn, (B) Neu5Ac, (C) Neu5Acα2→3Galβ1→4Glc, and (D) Neu5Acα2→6Galβ1→4Glc. The ligand color coding is equivalent to the SHL-I conformation shown in Figure 2.

Not only is Trp32 important for a proper lectin design, but also the existence of an Arg residue instead of Asp (comparing the lectin SHL-I with the toxin HWTX-I) is essential for sialic acid binding as illustrated on the interactions frequency plot (Fig. 9). In addition, filling the possible binding pocket with amino acids such as Leu30 in the case of the toxin HWTX-I might be another factor limiting the carbohydrate ligand's recognition by the toxin.

3. Materials and methods

3.1. Materials

SHL-I was isolated and purified according to a procedure described in the literature.[1] The purity of the peptide was controlled by electrophoresis, mass spectroscopy, amino acid sequence analyzis, and NMR spectroscopy. The binding of carbohydrates to

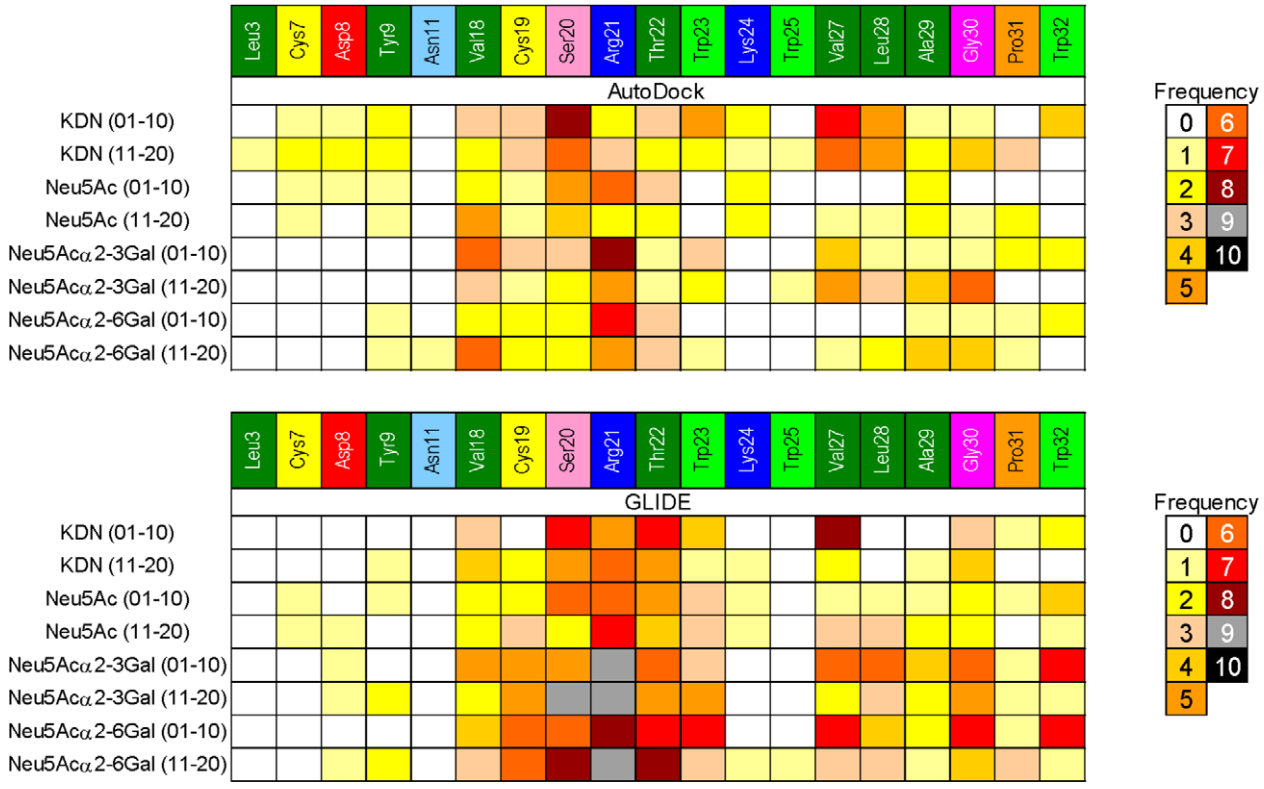

Figure 9. Interaction frequency plots (summing up both hydrogen-bond and steric interactions) for carbohydrate ligands and SHL-I amino acids based on AUTODOCK and GLIDE docking calculations in summary for all 20 conformations of the lectin. Only the best docked geometries were incorporated into the plots.

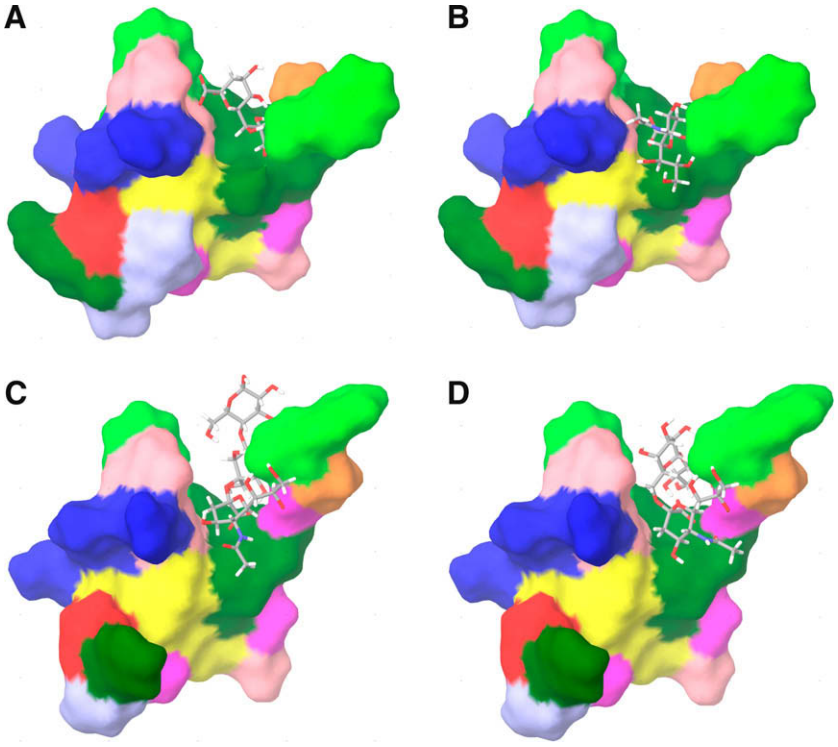

Figure 10. QPLD results illustrating the best binding modes. The SHL-I surface is colored according to AA properties, grouped as follows: DARK GREEN: Ala, Ile, Leu, Met, Val; BLUE: Lys, Arg; LIGHT BLUE: Gln, Asn; RED: Asp; YELLOW: Cys; PURPLE: Gly; ORANGE: Pro; PINK: Ser, Thr; GREEN: Trp. (A)Kdn, SHL-I Conf 09; (B) Neu5Ac, SHL-I Conf 09; (C) Neu5Acα2→3Galβ1→4Glc, SHL-I Conf 10; and (D) Neu5Acα2→6Galβ1→4Glc, SHL-I Conf 10. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

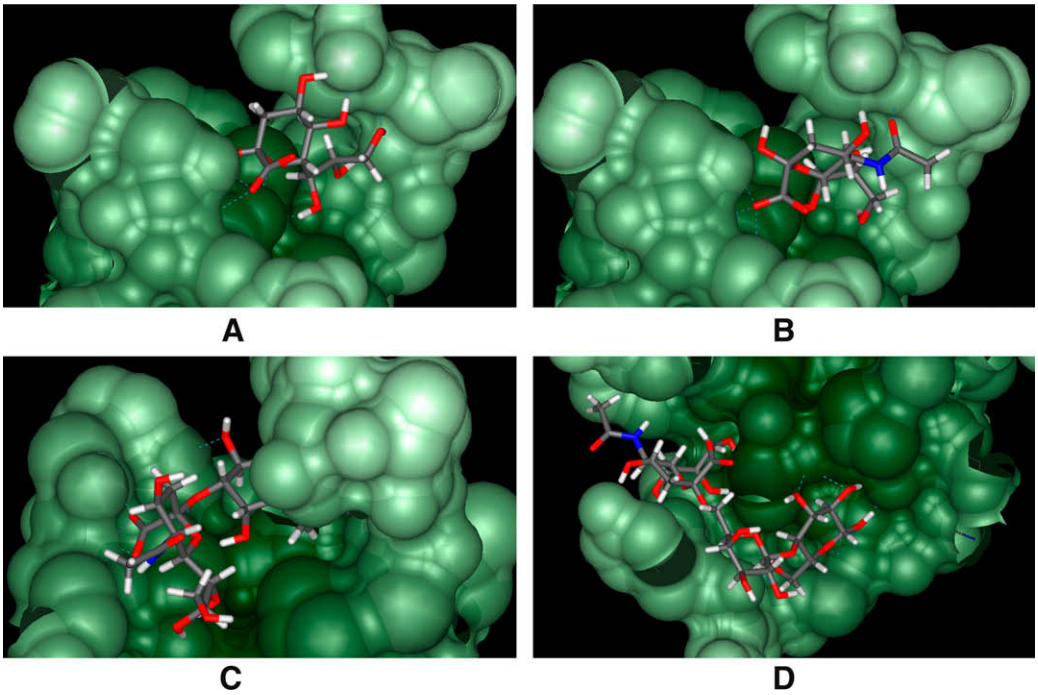

Figure 11. The best binding modes for SHL-I according to FLEXX docking. Only the surface of the binding site is visualized. (A) Kdn, SHL-I Conf 09; (B) Neu5Ac, SHL-I Conf 09; (C) Neu5Acα2→3Galβ1→4Glc, SHL-I Conf 07; and (D) Neu5Acα2→6Galβ1→4Glc, SHL-I Conf 16.

SHL-I was monitored by NMR titration experiments and laser photo CIDNP techniques.

3.2. NMR spectroscopy

3.2.1. 1D experiments

NMR titration experiments with various ligands were carried out by recording 1D 500 MHz ^{1}H NMR spectra of a series of SHL-I samples with variable carbohydrate concentrations (up to 10 different concentrations). The concentration of SHL-I during the NMR experiments was kept constant: 3 mM. First, a 0.5 mL sample of SHL-I in PBS buffer was used to obtain the ^{1}H NMR chemical shifts of the sugar-free protein sample. The 1D NMR spectrum for the sample with the highest ligand:protein ratio was recorded by dissolving the corresponding glycan (at a concentration of 30 mM) in 0.5 mL of the SHL-I-containing solution.

3.2.2. 2D experiments

Samples were prepared by dissolving the lyophilized powder in 0.5 mL of 20 mM phosphate buffer (90% H_2O, 10% D_2O). ^{1}H NMR spectra were recorded on Bruker AMX-600 and AMX-750 MHz spectrometers. The 2D TOCSY and 2D NOESY experiments were carried out in the phase-sensitive mode using the time proportional phase incrementation (TPPI) method for quadrature detection in F1. Typically, a data matrix of 512×2000 points was used in order to digitize a spectral width of 9000 Hz. Sixty scans were accumulated per increment with a relaxation delay of 1 s. Prior to Fourier transformation, we expanded the data to 1000×2000 by zero filling in the F1 direction. Baseline correction was applied in both dimensions. The corresponding shift was optimized for the different spectra. The 2D TOCSY spectra were acquired using MLEV-17 during the 60 ms of isotropic mixing period. The 2D NOESY experiments were performed with mixing times of 50, 125, and 200 ms.[27–29]

3.3. Laser photo CIDNP

The laser light for laser photo CIDNP experiments was generated by a continuous-wave argon ion laser that operates in the multi-line mode with principal emission wavelengths of 488.0 and 514.5 nm, close to the edge of the 450 nm absorption band of the used dye. The laser light was directed to the tube by an optical fiber. Mechanical chopping controlled by the spectrometer prevented a heating of the lectin-containing solution, precluding a denaturation of the protein. The CIDNP characteristic radical reaction was initiated by the flavin mononucleotide as laser-responsive dye. The irradiation leads to the generation of protein dye radical pairs by the surface-exposed Tyr-, Trp-, and His-ring systems. Nuclear spin-polarization is achieved from back-reactions of the radical pairs. Chemical shifts can be assessed relative to acetone (2.225 ppm) and HDO (4.76 ppm at a defined temperature and pH/pD-value). Two spectra were recorded in each laser photo CIDNP experiment. The resulting light spectrum was subtracted from the dark spectrum, thereby establishing the laser photo CIDNP difference spectrum only showing signals of polarized residues.

3.4. Molecular modeling

The lectin–ligand interaction profiles were studied by molecular docking using three different docking protocols, AUTODOCK,[30] FLEXX[31] and GLIDE.[32]

The atomic coordinates of the *S. huwena* spider lectin-I[1] (SHL-I) were obtained from the protein data bank[20] at www.pdb.org. The 20 NMR geometries in the 1QK7.pdb file were split into individual files for each conformation using the SWISS-PDB-VIEWER program.[33] In order to consider protein flexibility, all 20 SHL-I geometries were considered for docking calculations.

The SITEMAP program[18] was used in order to localize the possible binding sites of SHL-I.

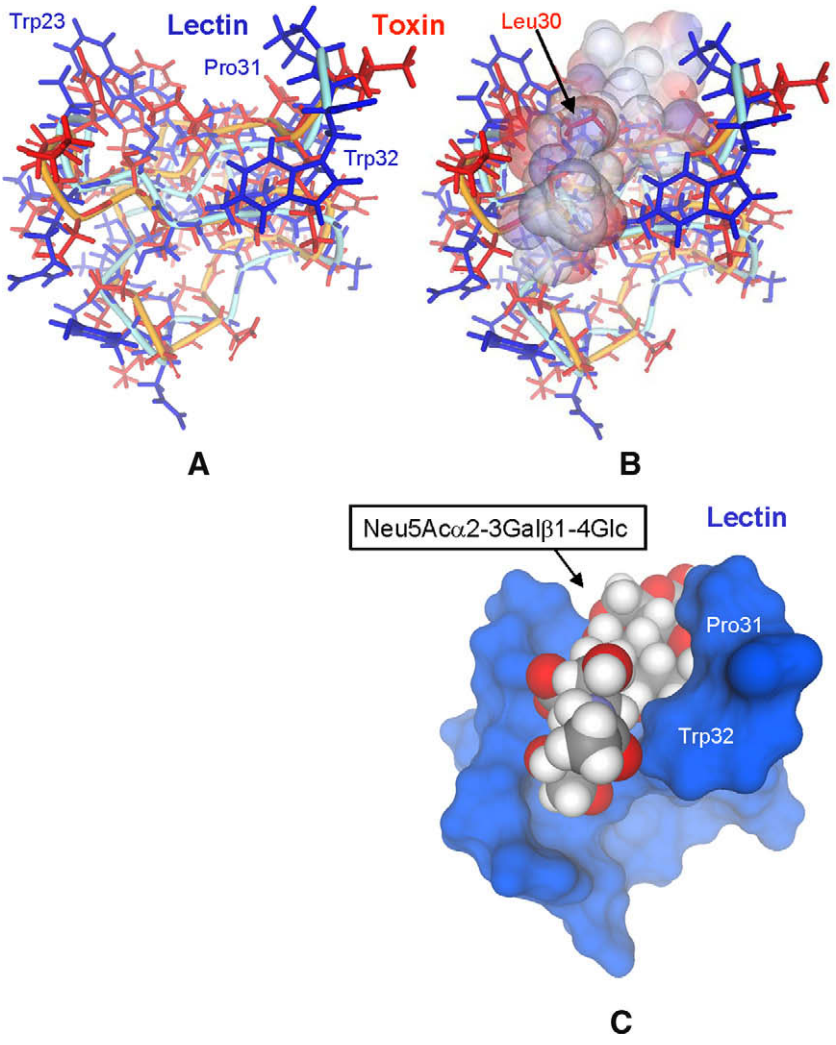

Figure 12. (A) and (B): Overlay of SHL-I (dark blue bonds; light blue backbone) and HWTX-I (red bonds, orange backbone). The semitransparent vdW surface area of the Neu5Acα2→3Galβ1→4Glc ligand is also shown in (B). The presence of highlighted Leu30 penetrating into overlaid ligand surface area disallows thus the preferred lectin-like binding mode of Neu5Acα2→3Galβ1→4Glc shown in (C) in the case of the lectin. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Semi-empirical minimizations on selected SHL-I conformations were performed with the HYPERCHEM 8.0[19] software package using AM1. Protein data bank searches[21] were carried out in order to identify amino acid residues that are preferably involved in interactions with sialic acids.

Conformational searches were performed on the trisaccharide ligands Neu5Acα2→3Galβ1→4Glc and Neu5Acα2→6Galβ1→4Glc using the MACROMODEL[34] program. The geometries of these two and the other monomeric ligands (Neu5Ac and Kdn) were further refined in accordance with our former ab initio study[35] by DFT B3LYP 6-31** geometry optimization using the JAGUAR[36] program.

The AUTODOCK[30] program version 4.0 was selected to evaluate the lectin–ligand binding energies on our first track of binding studies. The SHL-I geometries were converted into format suitable for AUTODOCK molecular docking. The affinity grids were calculated with 0.375 Å spacing around the binding site of the lectins using the AUTOGRID 4 program. The parameters of Lamarckian genetic algorithm with local searches were set as follows: ga_pop_size 150; ga_num_evals 1,500,000; ga_run 100 and ga_num_generations 27,000 in the AUTODOCK 4 program. The AUTODOCKTOOLS program was used for setting up the docking runs, that is, to assign the conformational variables for Neu5Acα2→3Galβ1→4Glc, Neu5Acα2→6Galβ1→4Glc (the names further abbreviated to Neu5Acα2→3Lac and Neu5Acα2→6Lac), Neu5Ac and Kdn and for assigning the atomic charges to the SHL-I lectin and for the ligand structures.

The docked geometries were sorted according to their binding energies and clustered with RMS tolerance of 2 Å. All geometries falling into the lowest energy clusters were on-screen visualized and analyzed with the LIGPLOT program.[37]

The FLEXX[31] program[38] was used on the second track of SHL-I–ligand interaction studies. The ab initio minimized ligand structures were considered to be suitable molecular structures for these runs. The default value of 6.5 Å to define the binding site was enlarged to 10 Å in order to better accommodate the trisaccharide structures.

The use of the GLIDE[32] docking program[39] was our last option to get deeper insight into how SHL-I binds carbohydrate ligands. Since GLIDE is one of the newest programs for docking studies and has not yet been widely tested for carbohydrate binding, we performed several runs:

(a) ligand inputs based on ab initio DFT B3LYP 6-31** optimized structures,
(b) GLIDE redocking of AUTODOCK preferred ligand conformations (docking altogether 80 input geometries for each SHL-I conformation),

(c) QPLD (Quantum Polarized Ligand Docking)[40] refinement of docking runs resulting from step (a). QPLD means standard docking in the first step, then ab initio partial atomic charge calculation on the ligand atoms and redocking with the refined charge set reflecting the polarization resulting from the protein–ligand interaction.

The DISCOVERY STUDIO VISUALIZER[41] was used for superposition and consequent visualization of the lowest energy protein–ligand complexes in addition to the MAESTRO[42] visualization tools.

Acknowledgments

The authors acknowledge Professor A. J. Olson for providing the AUTODOCK program. T. K. acknowledges the Slovak VEGA granting agency for providing support for related Project 2/7053/27.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.carres.2009.06.002.

References

1. Lu, S.; Liang, S.; Gu, X. *J. Protein Chem.* **1999**, *18*, 609–617.
2. Qu, Y.; Liang, S.; Ding, J.; Liu, X.; Zhang, R.; Gu, X. *J. Protein Chem.* **1997**, *16*, 565–574.
3. Liang, S. P. *Toxicon* **2004**, *43*, 575–585.
4. Shu, Q.; Liang, S. P. *J. Pept. Res.* **1999**, *53*, 486–491.
5. Peng, K.; Shu, Q.; Liu, Z. H.; Liang, S. P. *J. Biol. Chem.* **2002**, *277*, 47564–47571.
6. Shu, Q.; Lu, S. Y.; Gu, X. C.; Liang, S. P. *Protein Sci.* **2002**, *11*, 245–252.
7. Shu, Q.; Huang, R. H.; Liang, S. P. *Eur. J. Biochem.* **2001**, *268*, 2301–2307.
8. Siebert, H. C.; von der Lieth, C. W.; Kaptein, R.; Beintema, J. J.; Dijkstra, K.; vanNuland, N.; Soedjanaatmadja, U. M. S.; Rice, A.; Vliegenthart, J. F. G.; Wright, C. S.; Gabius, H. J. *Proteins: Struct. Funct. Genet.* **1997**, *28*, 268–284.
9. Asensio, J. L.; Canada, F. J.; Siebert, H. C.; Laynez, J.; Poveda, A.; Nieto, P. M.; Soedjanaamadja, U. M.; Gabius, H. J.; Jiménez-Barbero, J. *Chem. Biol.* **2000**, *7*, 529–543.
10. Espinosa, J. F.; Asensio, J. L.; Garcia, J. L.; Laynez, J.; Bruix, M.; Wright, C.; Siebert, H. C.; Gabius, H. J.; Canada, F. J.; Jiménez-Barbero, J. *Eur. J. Biochem.* **2000**, *267*, 3965–3978.
11. Asensio, J. L.; Siebert, H. C.; von Der Lieth, C. W.; Laynez, J.; Bruix, M.; Soedjanaamadja, U. M.; Beintema, J. J.; Canada, F. J.; Gabius, H. J.; Jiménez-Barbero, J. *Proteins* **2000**, *40*, 218–236.
12. Rudiger, H.; Siebert, H. C.; Solis, D.; Jiménez-Barbero, J.; Romero, A.; von der Lieth, C. W.; Diaz-Marino, T.; Gabius, H. J. *Curr. Med. Chem.* **2000**, *7*, 389–416.
13. Jiménez-Barbero, J.; Canada, F. J.; Asensio, J. L.; Aboitiz, N.; Vidal, P.; Canales, A.; Groves, P.; Gabius, H. J.; Siebert, H. C. *Adv. Carbohydr. Chem. Biochem.* **2006**, *60*, 303–354.
14. Siebert, H. C.; Tajkhorshid, E.; Vliegenthart, J. F. G.; von der Lieth, C. W.; Andre, S.; Gabius, H. J. Laser Photo CIDNP Technique as a Versatile Tool for Structural Analysis of Inter- and Intramolecular Protein–Carbohydrate Interactions. In *NMR Spectroscopy and Computer Modeling of Carbohydrates Recent Advances*; Vliegehhart, J. F. G., Woods, R., Eds.; American Chemical Society Symposium Series: Washington, DC, 2006; Vol. 930, pp 81–113.
15. Siebert, H. C.; Kaptein, R.; Beintema, J. J.; Soedjanaatmadja, U. M.; Wright, C. S.; Rice, A.; Kleineidam, R. G.; Kruse, S.; Schauer, R.; Pouwels, P. J.; Kamerling, J. P.; Gabius, H. J.; Vliegenthart, J. F. *Glycoconjugate J.* **1997**, *14*, 531–534.
16. Gabius, H. J.; Siebert, H. C.; Andre, S.; Jiménez-Barbero, J.; Rudiger, H. *Chembiochem* **2004**, *5*, 741–764.
17. Siebert, H. C.; Tajkhorshid, E.; Dabrowski, J. *J. Phys. Chem. A* **2001**, *105*, 8488–8494.
18. SiteMap, version 2.2. New York, NY, USA, Schrödinger, LLC; 2008.
19. HYPERCHEM, version 8.0. Gainesville, FL, USA, Hypercube, 2007.
20. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
21. Lutteke, T.; Frank, M.; von der Lieth, C. W. *Nucleic Acids Res.* **2005**, *33*, D242–D246.
22. Inoue, S.; Kitajima, K. *Glycoconjugate J.* **2006**, *23*, 277–290.
23. Siebert, H. C.; Andre, S.; Asensio, J. L.; Canada, F. J.; Dong, X.; Espinosa, J. F.; Frank, M.; Gilleron, M.; Kaltner, H.; Kozar, T.; Bovin, N. V.; von Der Lieth, C. W.; Vliegenthart, J. F.; Jiménez-Barbero, J.; Gabius, H. J. *Chembiochem* **2000**, *1*, 181–195.
24. Siebert, H. C.; Lu, S. Y.; Frank, M.; Kramer, J.; Wechselberger, R.; Joosten, J.; Andre, S.; Rittenhouse-Olson, K.; Roy, R.; von der Lieth, C. W.; Kaptein, R.; Vliegenthart, J. F.; Heck, A. J.; Gabius, H. J. *Biochemistry* **2002**, *41*, 9707–9717.
25. Siebert, H. C.; Andre, S.; Lu, S. Y.; Frank, M.; Kaltner, H.; van Kuik, J. A.; Korchagina, E. Y.; Bovin, N.; Tajkhorshid, E.; Kaptein, R.; Vliegenthart, J. F.; von der Lieth, C. W.; Jiménez-Barbero, J.; Kopitz, J.; Gabius, H. J. *Biochemistry* **2003**, *42*, 14762–14773.
26. Siebert, H. C.; Born, K.; Andre, S.; Frank, M.; Kaltner, H.; von der Lieth, C. W.; Heck, A. J. R.; Jiménez-Barbero, J.; Kopitz, J.; Gabius, H. J. *Chem. Eur. J.* **2005**, *12*, 388–402.
27. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355–360.
28. Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.
29. Marion, D.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967–974.
30. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
31. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470–489.
32. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. *J. Med. Chem.* **2006**, *49*, 6177–6196.
33. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. *Nucleic Acids Res.* **2003**, *31*, 3381–3385.
34. MACROMODEL, version 9.6. Schrödinger, LLC, New York, 2008.
35. van Lenthe, J. H.; den Boer, D. H. W.; Havenith, R. W. A.; Schauer, R.; Siebert, H. C. *J. Mol. Struct.-Theochem.* **2004**, *677*, 29–37.
36. JAGUAR, version 7.0. Schrödinger, LLC, New York, 2008.
37. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Eng.* **1995**, *8*, 127–134.
38. FLEXX, version 3.1. BioSolveIT, GmbH: St. Augustin, 2008.
39. GLIDE, version 5.0. Schrödinger, LLC, New York, 2008.
40. Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. *J. Comput. Chem.* **2005**, *26*, 915–931.
41. Discovery Studio Visualizer, version 1.6. ACCELRYS Software, San Diego, ©2002–2006.
42. MAESTRO, version 8.5. Schrödinger, LLC, New York, 2008.