Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 339 (2004) 907–928
Carbohydrate RESEARCH

Review

NMR spectroscopic and molecular modeling studies of protein–carbohydrate and protein–peptide interactions

Margaret A. Johnson[†] and B. Mario Pinto[*]

Departments of Chemistry and of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

Received 26 September 2003; accepted 17 December 2003

This paper is dedicated, with respect, to the memory of Christian Pedersen

Abstract—Investigations of the conformations of carbohydrates, their analogues and their molecular mimics are described, with emphasis on structural and functional information that can be gained by NMR spectroscopic techniques in combination with molecular modeling. The transferred nuclear Overhauser effect (trNOE) has been employed to determine the bound conformations of carbohydrates and other bioactive molecules in complex with protein receptors. The corresponding experiments in the rotating frame (trROE) and selective editing experiments (e.g., QUIET-NOESY) are used to eliminate indirect cross-relaxation pathways (spin diffusion), thereby minimizing errors in the data used for calculation of conformations. Saturation transfer difference NMR experiments reveal detailed information about intermolecular contacts between ligand and protein. Computational techniques are integrated with NMR-derived information to construct structural models of these bioactive molecules and of their complexes with proteins. Recent investigations into the nature of molecular mimicry with regard to protein–ligand interactions are described, along with applications in determining the mode of action of enzyme inhibitors. The results are relevant for the design of the next generation of drug and vaccine candidates.

Keywords: Protein–carbohydrate interactions; Protein–peptide interactions; Bioactive conformations; NMR; Transferred NOE; Transferred ROE; Saturation transfer difference; Epitope mapping; Molecular modeling

Contents

* Corresponding author. Tel.: +1-604-291-4327; fax: +1-604-291-5424; e-mail: bpinto@sfu.ca
† Present address: The Scripps Research Institute, Department of Molecular Biology, 10550 North Torrey Pines Rd., MB-44, La Jolla, CA 92037, USA.

0008-6215/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2003.12.028

1. Introduction

Molecular recognition of small molecules (ligands) by proteins is important in many biological processes. For example, recognition of cell-surface molecules of foreign organisms by antibodies is an important part of immune defense. Protein–carbohydrate recognition occurs in diverse processes such as blood clotting, leukocyte rolling, cell–cell and cell–matrix interactions, and intracellular protein transport.[1–5] Cell-surface carbohydrates of mammalian cells also serve as opportunistic targets for the attachment of viruses, bacteria, and bacterial toxins, mediated by specific binding of proteins.[1–5] Therefore, it is of interest to determine how this molecular recognition occurs, in order to understand how these processes work and to intervene in these processes for therapeutic purposes. Various NMR spectroscopic and molecular modeling techniques are useful in such investigations. This review will describe such studies from our own laboratory aimed at the investigation of protein/carbohydrate and protein/carbohydrate-mimetic interactions. Topics are introduced at a fundamental level suitable for students prior to detailed discussions of case studies. For a more general coverage of the field, including the use of NMR screening in drug discovery, the reader is referred to several excellent accounts.[6–14]

2. Exchanging protein–ligand systems

The binding of a ligand to a protein is depicted below. The ligand, L, a small molecule (such as a peptide or oligosaccharide) binds reversibly to the protein, P (such as an enzyme or antibody).

$$\mathrm{P} + \mathrm{L} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \mathrm{P}\cdot\mathrm{L} \tag{1}$$

The binding is characterized by an association constant, K_{A}:

$$K_{\mathrm{A}} = \frac{k_{\mathrm{on}}}{k_{\mathrm{off}}} = \frac{[\mathrm{P}\cdot\mathrm{L}]}{[\mathrm{P}][\mathrm{L}]}, \tag{2}$$

The association constants of typical binding reactions of biomolecules range from 10^2 to $10^{12}\,\mathrm{M}^{-1}$. However, K_{A} values for protein–carbohydrate interactions are typically weaker, in the 10^2–$10^6\,\mathrm{M}^{-1}$ range.[15] Since the association rate is usually rapid, limited only by molecular diffusion (k_{on} typically $10^8\,\mathrm{M}^{-1}\,\mathrm{s}^{-1}$), dissociation rates are also rapid ($k_{\mathrm{off}} > 100\,\mathrm{s}^{-1}$), leading to fast exchange of the ligand between free and bound states.

3. The nuclear Overhauser effect

The nuclear Overhauser effect (NOE) is defined as the change in intensity of a certain resonance after perturbation of the population of a spin in close spatial proximity (by saturation or inversion). The interaction through space of the dipoles of neighboring nuclei, or cross-relaxation, gives rise to the NOE. The cross-relaxation rate, σ, governs the magnitude and sign of the NOE. For an isolated two-spin system, the change in intensity of spin I after perturbation of spin S is described by the Solomon equation[16]

$$\frac{\mathrm{d}I_Z}{\mathrm{d}t} = -(I_Z - I_Z^0)(W_{0IS} + 2W_{1I} + W_{2IS}) - (S_Z - S_Z^0) \times (W_{2IS} - W_{0IS}), \tag{3}$$

where the W factors are transition probabilities (rate constants) for zero-, single-, and double-quantum transitions between energy levels of the two-spin system, as shown in Figure 1, and I_Z and S_Z are the longitudinal components of the magnetization of spins I and S.

Transitions between the energy levels are stimulated by transient local magnetic fields induced by the dipoles of the spins tumbling past each other. The W transition probabilities (rate constants) therefore depend on the rate of molecular tumbling. The term W_{2IS} is the rate constant for double-quantum transitions and is stimulated by local magnetic fields fluctuating at frequencies of $2\nu^0$ (the frequency difference between the $\alpha\alpha$ and $\beta\beta$ energy levels); therefore, it is greatest if the rate of molecular tumbling is rapid. This term gives rise to positive NOE enhancements. The term W_{0IS} describes zero-quantum transitions, and is stimulated by fields fluctuating at very low frequencies ($\nu_I - \nu_S \sim 0$); this term gives rise to negative NOE enhancements. The

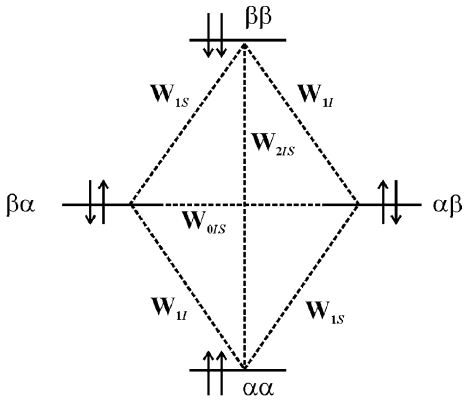


Figure 1. Energy level diagram for a two-spin system, showing transition probabilities (W) and the spin states α and β. The I spin state is written first and the S spin state second. After Ref. 17.

term $(W_{2IS} - W_{0IS})$ in the Solomon equation is the overall rate constant for dipole–dipole transitions giving rise to the NOE, or cross-relaxation; this term is known as the cross-relaxation rate constant, σ_{IS}. The term $(W_{0IS} + 2W_{1I} + W_{2IS})$ is the dipolar longitudinal relaxation rate constant of spin I, or ρ_{IS}.[17] An expression for the steady-state nuclear Overhauser effect (NOE), the change in I spin intensity on saturation of spin S (abbreviated $f_I\{S\}$), may be derived by setting $dI_Z/dt = 0$ and $S_Z = 0$:[17]

$$0 = -(I_Z - I_Z^0)(W_{0IS} + 2W_{1I} + W_{2IS}) + S_Z^0(W_{2IS} - W_{0IS}), \quad (4)$$

$$\frac{I_Z - I_Z^0}{S_Z^0} = \frac{W_{2IS} - W_{0IS}}{W_{0IS} + 2W_{1I} + W_{2IS}}. \quad (5)$$

Since $S_Z^0 = (\gamma_S/\gamma_I)I_Z^0$,

$$f_I\{S\} = \frac{I_Z - I_Z^0}{I_Z^0} = \frac{\gamma_S}{\gamma_I}\frac{W_{2IS} - W_{0IS}}{W_{0IS} + 2W_{1I} + W_{2IS}} = \frac{\gamma_S}{\gamma_I}\frac{\sigma_{IS}}{\rho_{IS}}. \quad (6)$$

When the dependence of the W rate constants on the precession frequencies of the spins (ω) and on the molecular tumbling rate (defined by the rotational correlation time, τ_c), is incorporated into the above expression, the result is,

$$f_I\{S\} = \frac{\gamma_S}{\gamma_I}\frac{5 + \omega^2\tau_c^2 - 4\omega^4\tau_c^4}{10 + 23\omega^2\tau_c^2 + 4\omega^4\tau_c^4}, \quad (7)$$

which reduces to $0.5(\gamma_S/\gamma_I)$ for small molecules tumbling rapidly ($\omega\tau_c \ll 1$); and to $-1(\gamma_S/\gamma_I)$ for large molecules tumbling slowly ($\omega\tau_c \gg 1$). NOE enhancements are expected to be positive for small molecules, and negative for large molecules.

In practice, the most useful information is obtained from transient NOE experiments, rather than from steady-state NOE experiments. Transient experiments include NOESY, or the 1D selective equivalent (illustrated in Fig. 2). From these experiments, NOE buildup

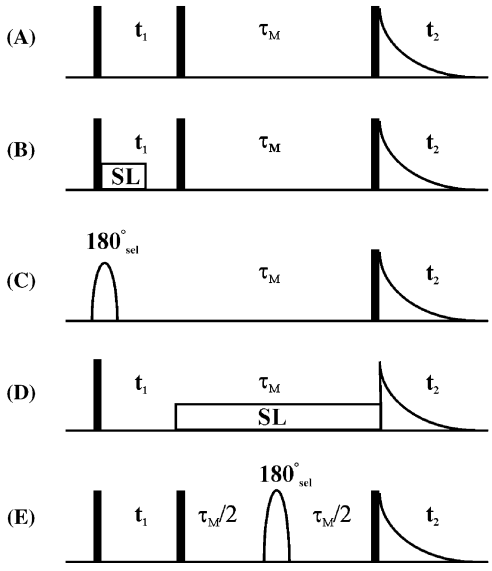


Figure 2. NMR experiments employed for the measurement of NOE effects (some of these experiments are described later in the text). t_1 indicates the incremented delay used to create the first (indirect) frequency dimension; t_2 indicates the time for acquisition of the NMR signal (direct frequency dimension). τ_M indicates the mixing time, during which the NOE effects develop. The dark rectangles represent 90° pulses. (A) The NOESY experiment. (B) The NOESY experiment with a $T_{1\rho}$ filter added after the first 90° pulse, to suppress protein resonances. (C) The 1D selective version of the NOESY experiment (1D transient NOE); $180°_{sel}$ indicates a low-power shaped pulse used to provide selective inversion of one resonance. (D) The ROESY experiment; here, the mixing time is provided by a spin-lock pulse (SL). (E) The QUIET-NOESY experiment; here, $180°_{sel}$ indicates a doubly band-selective 180° pulse, applied halfway through the mixing time, to invert resonances of interest.

rates, governed by σ_{IS}, are measured. The initial rate of NOE buildup is directly proportional to σ_{IS}:

$$\frac{dI_Z}{dt} = -\sigma_{IS}S_Z^0. \quad (8)$$

The relative values of σ_{IS} for different proton pairs within a molecule provide a measure of internuclear distances, because the dipole–dipole interaction, which gives rise to cross-relaxation, has a strong (r^{-3}) dependence on distance.

The dipole–dipole coupling constant is

$$D_{IS} = \frac{-\mu_0}{4\pi}\hbar\gamma_I\gamma_S\frac{1}{2}\left\langle\frac{3\cos^2\theta_{IS} - 1}{r_{IS}^3}\right\rangle, \quad (9)$$

D_{IS}, in combination with the spectral density at certain frequencies, defines the W transition rate constants, as follows:

$$W_{2IS} = \tfrac{3}{5}D_{IS}^2J(2\omega^0) \quad W_{0IS} = \tfrac{1}{10}D_{IS}^2J(\omega_I - \omega_S). \quad (10)$$

The cross-relaxation rate σ_{IS}, defined earlier as $(W_{2IS} - W_{0IS})$, is therefore proportional to D_{IS}^2 and r^{-6}:

$$\sigma_{IS} = W_{2IS} - W_{0IS} = \tfrac{3}{5}D_{IS}^2J(2\omega^0) - \tfrac{1}{10}D_{IS}^2J(\omega_I - \omega_S), \quad (11)$$

$$\sigma_{IS} = \frac{D_{IS}^2}{10}(6J(2\omega^0) - J(0)) = \frac{K}{r^6}(6J(2\omega^0) - J(0)) = \frac{K}{r^6}\left(\frac{6\tau_c}{1+4\omega^{02}\tau_c^2} - \tau_c\right) \quad (12)$$

Eq. 12 describes the homonuclear case, $J(\omega_I - \omega_S) \cong J(0)$. The factor K is defined as

$$K = \left(\frac{\mu^0}{4\pi}\right)^2 \left(\frac{\hbar^2\gamma_I^2\gamma_S^2}{10}\right). \quad (13)$$

It is usually assumed that the molecule is tumbling isotropically and that τ_c is the same for all pairs of spins in the molecule. One may then calculate the distances between pairs of protons by using the NOE buildup rate for a pair of protons of known separation as a calibration factor:

$$\frac{\sigma_{ij}}{\sigma_{\text{known}}} = \frac{r_{ij}^{-6}}{r_{\text{known}}^{-6}}. \quad (14)$$

A set of these internuclear distances allows molecular conformation to be determined.

4. The transferred nuclear Overhauser effect

For a small-molecule ligand binding reversibly to a protein, fast exchange on the 'relaxation time scale' implies that k_{off} is greater than σ, in both the free and bound states of the ligand. This condition leads to averaging of the cross-relaxation rate, and therefore of the observed NOE, between the ligand free and bound states:

$$\sigma_{\text{OBS}} = p_B\sigma_B + p_F\sigma_F, \quad (15)$$

where p_B and p_F are the fractions of bound and free ligand. This leads to the phenomenon of the transferred nuclear Overhauser effect, which allows the determination of bound-ligand conformations.[18–23]

The transferred nuclear Overhauser effect (trNOE) refers to the observation of negative NOE effects, representative of the bound conformation, for a small-molecule ligand in solution with a protein, in contrast to the positive NOE effects normally expected for a small molecule in the absence of protein. Since σ_B is much greater than σ_F, σ_{OBS} is dominated by σ_B even with an excess of free ligand:

$$p_B\sigma_B \gg p_F\sigma_F, \quad (16)$$

$$\sigma_{\text{OBS}} \cong p_B\sigma_B. \quad (17)$$

The large, negative NOE effects which develop when the ligand is bound to the protein are transferred from the bound to the free state of the ligand by chemical exchange. This allows transferred NOE effects to be measured on samples containing a mixture of the protein and an excess of free ligand. These transferred NOE effects may be used to calculate distances between protons in the ligand, using a known distance as a reference, as described above. A set of several of these internuclear distances allows the bound conformation of the ligand to be calculated.

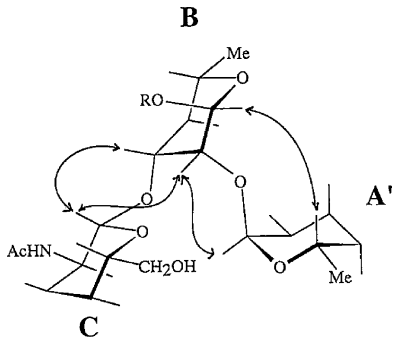


Figure 3. Bound conformation of the branched trisaccharide Rha-α-(1→2)-(Glc*N*Ac-β-(1→3)-)Rha-α-OPr (A′-(C)-B). Some of the major observed trNOE contacts are indicated by arrows, namely H-1A′–H-2B, H-5A′–H-1B, H-1C–H-3B, and H-1C–H-2B. Reproduced with permission from Ref. 24. Copyright © 1995 American Chemical Society.

In an ideal case, a sample containing only free ligand would show positive or zero NOE effects, allowing the negative transferred NOEs to be easily distinguished. A control experiment using another protein should also be run to ensure that the observed changes in NOE sign and intensity do not result simply from the higher viscosity of a solution containing protein.

TrNOE experiments have been used to demonstrate the selection of certain conformers by a protein, from an ensemble of conformers present in solution. For example, the recognition of a branched trisaccharide corresponding to the repeating unit of the cell-wall polysaccharide of Group A *Streptococcus* (GAS), by a monoclonal antibody directed against GAS, was investigated by trNOE experiments.[24] It was demonstrated that the antibody recognizes a local minimum conformation of the trisaccharide, Rha-α-(1→2)-(Glc*N*Ac-β-(1→3)-)Rha-α-OPr (A′-(C)-B). This conformer is slightly higher in energy (estimated at 2 kcal mol^{-1}) than the global minimum conformation, and differs by a rotation around the Rha-α-(1→2)-Rha (A′-B) linkage. The bound conformation is shown in Figure 3; examples of the trNOE buildup curves are shown in Figure 4. In this case, the free trisaccharide also showed negative NOE effects; however, the trNOE effects could be distinguished from these by their more rapid buildup rates (Fig. 4).

5. Experimental complications

The most important source of error in transferred NOE experiments is considered to be spin diffusion, or rapid NOE buildup not only between neighboring spins, but through many spins on the protein and ligand, resulting

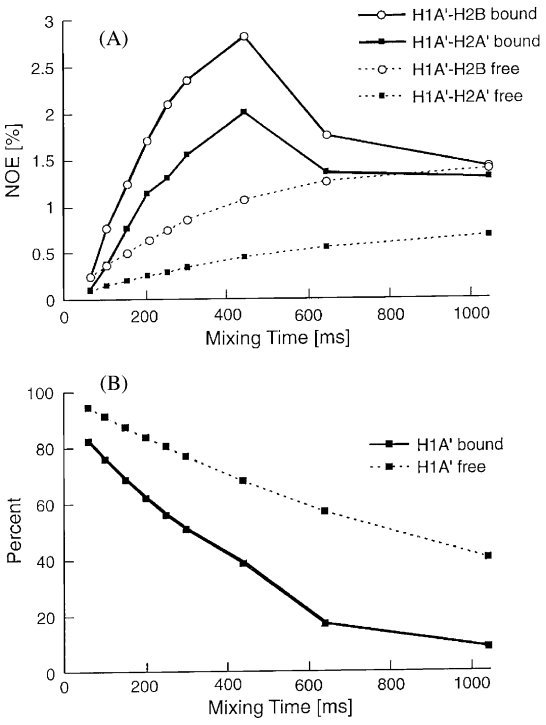


Figure 4. (A) Time dependence of the NOE (free trisaccharide, dotted lines) and trNOE (bound trisaccharide, solid lines) effects for H-1A′–H-2A′ and H-1A′–H-2B. (B) Time dependence of the decay of the selectively inverted H-1A′ resonance; more rapid T_1 relaxation is observed for the bound trisaccharide. Reproduced with permission from Ref. 24. Copyright © 1995 American Chemical Society.

from the rapid cross-relaxation rate, σ_B. Spin diffusion along a network of many spins may lead to the appearance of NOE effects between two spins, which are not actually close in space. This may lead to erroneous internuclear distances and errors in the calculated conformation. Therefore, several methods have been developed to remove spin-diffusion effects from NOE spectra.

6. The transferred ROESY experiment

The ROESY experiment[25,26] is an alternative method of observing distances through space between nuclei in a molecule. The experiment is performed by replacing the mixing time of a NOESY sequence with a mixing time in which a spin-lock pulse is applied (Fig. 2). This is an additional magnetic field pulse along the same axis where the spins are already aligned. This has the effect of abolishing precession in the *xy*-plane and 'locking' the spins together along one axis of the rotating frame. Essentially, the constant B_0 field is replaced by the much smaller B_1 magnetic field.

The relaxation properties of nuclei precessing around the B_1 field are analogous to those of nuclei in the B_0 field, with the exception that the size of the field is much (roughly 10,000 times) smaller. Therefore, essentially all molecules are in the fast-tumbling limit, and the NOEs observed (called ROEs, or rotating frame NOE effects) are positive for all molecules. Therefore, each NOE (ROE) transfer results in a change in sign, and direct effects will have positive signs, while three-spin effects will have negative signs. In a trROESY spectrum, these signs allow one to distinguish between true close distances and distances, which are mediated by spin diffusion. In practice, one often observes cancellation, or reduction in intensity, of cross-peaks which have

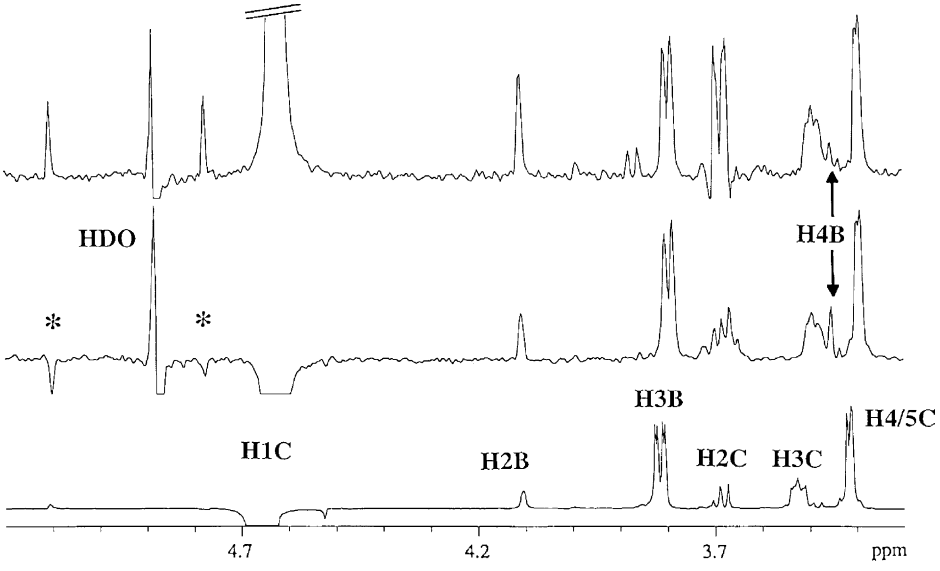


Figure 5. A comparison of the 1D transient NOE spectrum (lower row, mixing time 840 ms) of the free trisaccharide with traces at the frequency of H-1C in F_1 (4.65 ppm) from the trNOESY (upper row, mixing time 200 ms) and trROESY (middle row, mixing time 200 ms) spectra. Peaks denoted with asterisks are artifacts resulting from t_1 noise. Reproduced with permission from Ref. 24. Copyright © 1995 American Chemical Society.

contributions from both direct effects and spin-diffusion effects.

TrROESY also has some disadvantages; for large molecules, it is less sensitive than trNOESY because the rates of buildup and decay are more similar, and resonance offset effects and Hartmann–Hahn effects complicate the calculation of internuclear distances.[17] Therefore, it is usually used as a check: the NOE data are used as the primary data for distance determination and structure calculation, while ROE spectra are used qualitatively, to verify that the NOE cross-peaks represent true close distances; in some cases, the ROE data are also used to furnish distance information.

An example of the use of trROESY experiments is shown in Figure 5 for the branched trisaccharide/anti-GAS antibody system described above. The figure compares NOE effects of the free trisaccharide (lower row) to trNOE effects observed for the bound trisaccharide (upper row) and to trROE effects observed for the bound trisaccharide (middle row). The figure shows traces at the frequency of H-1C; therefore, the peaks observed all represent NOE contacts of H-1C to other protons of the trisaccharide. A relatively strong trNOE effect is observed from H-1C to H-2B; however, the intensity of this peak is considerably reduced in the trROESY spectrum, indicating that spin diffusion contributes to this cross-peak. The distance between these protons was therefore extracted from the trROESY spectra rather than the trNOESY spectra.

Another example of the use of trROESY to identify spin-diffusion effects was provided by Glaudemans and co-workers,[27,28] who initially reported a conformational change in a flexible, $(1 \rightarrow 6)$-linked disaccharide upon binding to an antibody, based on trNOE data.[27] However, the cross-peak defining this conformation disappeared in the trROESY spectra, indicating that spin diffusion had contributed to its intensity, and therefore, the revised data did not support this conformational change.[28]

7. Relaxation network editing

An alternative strategy to reduce or eliminate spin-diffusion effects is to selectively remove certain spins from the relaxation network by applying magnetic field pulses. This is known as 'relaxation network editing': the network is 'edited' by selectively eliminating indirect cross-relaxation pathways, allowing the observation of direct pathways only.[29] Several methods exist for this purpose. In the MINSY experiment (Mixing Irradiation during NOESY),[30] one spin of the ligand is selectively saturated

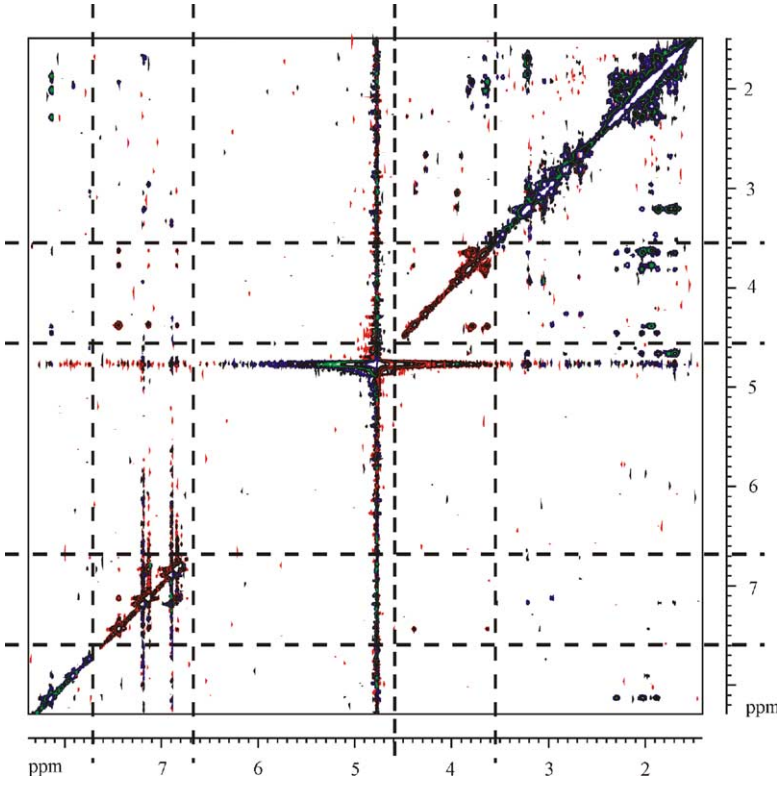


Figure 6. An example of a 2D QUIET-NOESY spectrum. The positions of the inverted bands (3.5–4.5 and 6.7–7.7 ppm) are indicated by dashed lines. The quiet windows occur where the bands intersect. Cross-peaks in these windows are free of spin diffusion from spins outside the bands; their signs are inverted relative to the rest of the spectrum (inverted regions, red; noninverted regions, blue).

using a low-power pulse, during the NOESY mixing time, effectively removing it from the network and neutralizing any NOE transfers mediated by this spin. Alternatively, in the QUIET-NOESY experiment (Quenching of Undesirable External Trouble in NOESY), a doubly band-selective inversion pulse is applied halfway through the NOESY mixing time.[31,32] The positions of the inverted bands are chosen to contain resonances contributing to a cross-peak of interest. After the inversion, NOE buildup occurs as normal between these resonances, since they have both been inverted. However, NOE buildup between resonances within the bands and resonances outside the bands changes sign, and by the end of the mixing time, is cancelled. Therefore, spin diffusion originating from resonances outside the bands is eliminated, and cross-peaks remaining within the quiet window (intersection of the bands) are due to true close distances.[31,32] A sample QUIET-NOESY spectrum, showing the positions of the inverted bands and quiet windows, is shown in Figure 6.

A schematic diagram showing how spin diffusion may arise within a ligand bound to a protein is shown in Figure 7. In addition, schematic traces from 2D spectra, illustrating the expected appearance of trROESY and QUIET-trNOESY spectra, that would allow the identification of these indirect effects, are shown.

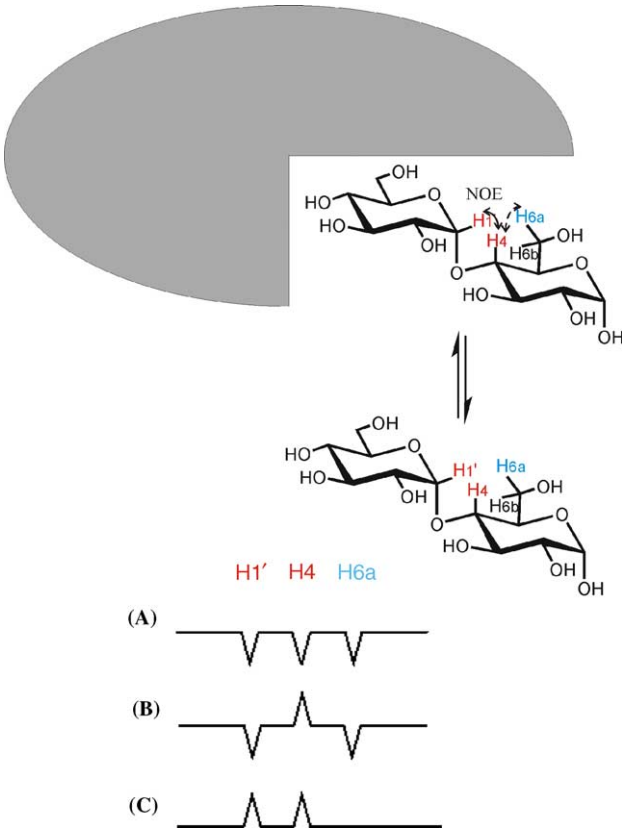


Figure 7. A hypothetical representation of spin-diffusion effects in maltose in complex with a protein. The solid arrow represents a direct NOE between the H-1′ and H-4 protons (red), while the dashed arrow represents spin diffusion from H-4 to H-6a; this would result in the observation of a cross-peak between H-1′ and H-6a in the trNOESY spectrum. This cross-peak would not represent a true close distance between H-1′ and H-6a, but rather would arise from spin diffusion, as described. The lower panel shows the expected appearance of traces from the NOE spectra taken at the frequency of H-1′. The traces are shown with the sign of H-1′ negative. (A) The trNOESY spectrum would show negative trNOE cross-peaks for both H-4 and H-6a. (B) The trROESY spectrum would show a true (positive) trROE cross-peak to H-4, but a negative trROE cross-peak to H-6a, indicating a three-spin NOE transfer (spin-diffusion pathway). (C) The QUIET-NOESY spectrum would show a cross-peak to H-4, but no cross-peak to H-6a; in this case, the spin-diffusion pathway would have been eliminated. (In this case, the signs of H-1′ and H-4 are positive, as the observation of direct cross-peaks between them requires each to be inverted halfway through the mixing time.)

8. Epitope mapping

8.1. Line-broadening effects

Selective line broadening of certain resonances of the ligand that are in close contact with the protein may occur due to enhanced T_2 relaxation. The large size of the ligand–protein complex, and its slow tumbling in

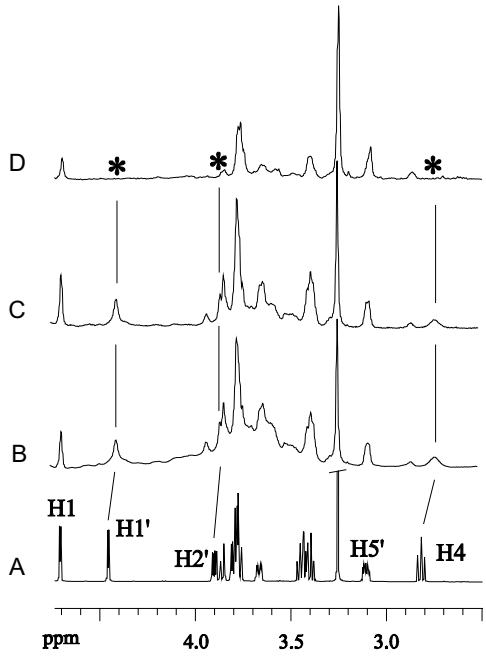
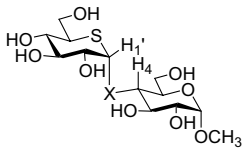


Figure 8. 1D NMR spectra of **2**, in the absence and presence of glucoamylase G1. (A) 1D NMR spectrum of **2** in the absence of the enzyme. (B) 1D spectrum of **2** in the presence of the enzyme. (C, D) 1D spectra with $T_{1\rho}$ filters of 16 and 200 ms, respectively. The signals H-1′, H-2′, and H-4 (marked with asterisks) show faster relaxation than the rest of the signals; these signals are reduced to near-zero intensity in the spectrum with the 200 ms filter (D). Reproduced with permission from Ref. 34.

solution, will increase the rate of T_2 relaxation, giving rise to greater line widths for the ligand resonances. Enhanced $T_{1\rho}$ relaxation (the analogue of T_1 relaxation in the rotating frame) may also occur, and may be observed in ROESY spectra as a decrease in intensity, or even complete disappearance, of certain resonances. This is also the basis of the use of a $T_{1\rho}$ filter (a short spin-lock pulse, applied after the first 90° pulse of the NOESY sequence) to relax protein resonances, for clearer observation of ligand resonances in trNOESY spectra (Fig. 2). For example, Scherf and Anglister[33] used the $T_{1\rho}$ filter in NOESY to remove intramolecular NOEs of the protein from the trNOESY spectrum of a peptide, in a rapidly exchanging peptide–antibody system. In this case, a 10-fold difference in $T_{1\rho}$ between the antibody protons (10 ms) and the peptide protons (100–500 ms) allowed the application of a 20 ms spin-lock pulse to rapidly relax the protein protons, while minimizing the effect on the peptide resonances.

Enhanced $T_{1\rho}$ relaxation in both $T_{1\rho}$-filtered 1D spectra and trROESY spectra was observed in complexes of glucoamylase with the inhibitors **1** and **2**.[34]



1 X = NH
2 X = S

Line broadening in excess of that observed for the other resonances, and faster $T_{1\rho}$ relaxation, were observed for the H-4 and H-1′ resonances of **1** and **2** (Fig. 8). In addition, the diagonal- and cross-peaks of these resonances disappeared in the trROESY spectra, due to rapid $T_{1\rho}$ relaxation (Fig. 9). This enhanced relaxation originated from especially close distances of these protons to protons of aromatic residues of the enzyme, as revealed by molecular modeling of the inhibitors in the enzyme active site (Fig. 10).[34]

Similar effects were observed in the complex of D-*gluco*-dihydroacarbose (**3**), a potent glycosidase inhibitor, with glucoamylase G2.[34] Despite a high affinity

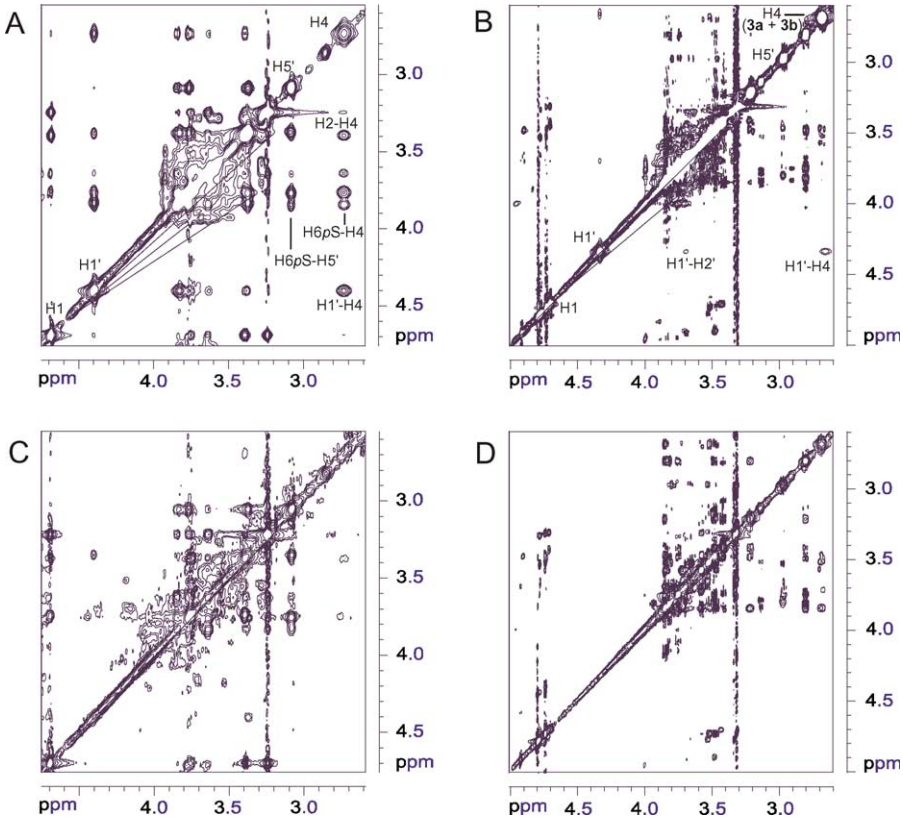


Figure 9. 2D NMR spectra of **1** and **2**, in the presence of glucoamylase G1. (A) TrNOESY spectrum of **2** in the presence of the enzyme, with a mixing time of 200 ms. (B) TrNOESY spectrum of **1** in the presence of the enzyme, with a mixing time of 250 ms. Compound **1** exists as an anomeric mixture, and weak trNOE effects are visible for the binding α-anomer; the β-anomer does not bind. (C) TrROESY spectrum of **2** in the presence of the enzyme, with a spin-lock pulse duration of 200 ms; the diagonal- and cross-peaks of H-1′ and H-4 are missing, due to rapid $T_{1\rho}$ relaxation. (D) TrROESY spectrum of a mixture of **1** and **2** in the presence of the enzyme, with a spin-lock pulse duration of 200 ms; in this case, the diagonal- and cross-peaks of the H-1′ and H-4 signals of both **1**α and **2** are missing due to rapid $T_{1\rho}$ relaxation. The H-4 signal of the nonbinding compound, **1**β, is still visible. Reproduced with permission from Ref. 34.

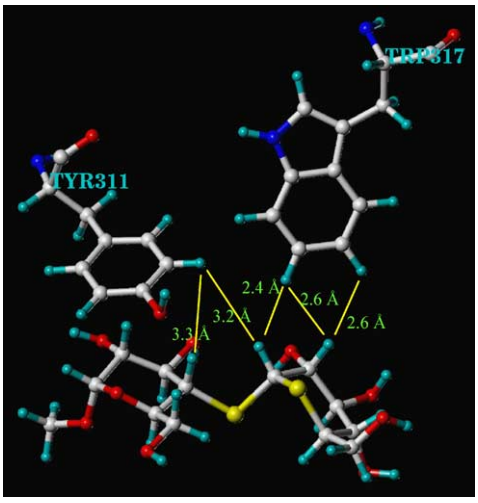


Figure 10. Close contacts between H-1′, H-2′, and H-4 and aromatic amino acids Tyr-311 and Trp-317 of the protein, in the complex of **2** with the catalytic domain of glucoamylase. Reproduced with permission from Ref. 34. Copyright © 2000 American Chemical Society.

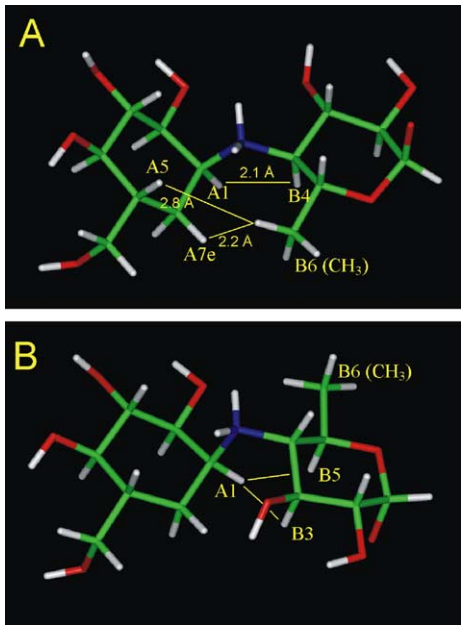
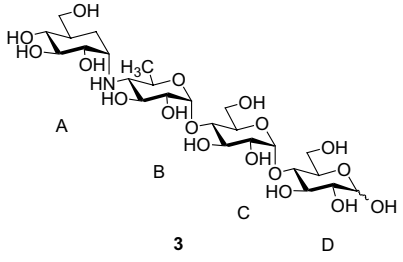


Figure 11. Conformations of D-*gluco*-dihydroacarbose (**3**) at the N-glycosidic linkage between rings A and B. Panel (A) shows the model of **3** observed in an X-ray crystal structure of the complex with glucoamylase (conformation A, preferred at basic pH).[36] Panel (B) shows the model of the 'inverted' conformation preferred at acidic pH (conformation B), which differs from A by a rotation around the ψ_H angle of the N-glycosidic linkage. Distances less than 3 Å are labeled in panel (A); these correspond to trNOEs observed for the glucoamylase/**3** complex. Short distances are also indicated in panel (B); the corresponding trNOEs were not observed, indicating that conformer B is not bound by the enzyme. Reproduced with permission from Ref. 35. Copyright © 2000 Elsevier Science.

constant ($K_A = 3 \times 10^7\ M^{-1}$), the off-rate for the complex was still high enough to allow the observation of transferred NOEs. Two conformational families about the N-glycosidic linkage of **3** are populated in solution. Thus, conformation A (Fig. 11A), defined as (ϕ_H, $\psi_H = -44°, -27°$) is preferred in basic solution, while the 'inverted' conformation B, which differs from A by a rotation around the ψ_H angle (ϕ_H, $\psi_H = -22°, -179°$) is preferred in acidic solution (Fig. 11B).[36] It was determined that only one of these families is selected and bound by the enzyme (Fig. 11). Thus, trNOEs representative of conformation A (Fig. 11A), for example, A H-1–B H-4 and A H-7e–B H-6, were observed, but no trNOEs representative of conformation B (Fig. 11B), for example, A H-1–B H-3 or A H-1–B H-5, were observed. Interestingly, even at acidic pH (pH 4.5), the enzyme selects conformation A, which is not the predominant conformer for the free ligand **3** at acidic pH.



3

The trNOE NMR study[34] was in agreement with the results of crystallographic analysis, which had showed conformer A bound in the enzyme active site.[37] The comparison of molecular recognition in the solution phase with that observed in solid-phase crystal structures is of general significance.

Contributions from spin diffusion were evaluated by transferred ROESY experiments, as described above. For example, the contact between the methyl group of ring B (B H-6) and one of the methylene protons of ring A (A H-7a), which spans the N-glycosidic linkage, showed a change of sign in the trROESY spectrum (Fig. 12). This observation demonstrates that this contact results from spin diffusion, probably mediated by the short distance between B H-6 and A H-7e (Fig. 11A). The B H-6–A H-1 contact did not change sign, but was reduced in intensity in the trROESY spectrum, also indicating a significant contribution from spin diffusion. These results were consistent with the rest of the NMR data defining the bound conformer.

Line-broadening effects were also observed. Thus, the H-4 resonance of ring B had a much greater line width than the other ligand resonances (Fig. 12). Inspection of the crystal structure of the glucoamylase/**3** complex revealed a close distance between this proton and Hε1 of Tyr-311. Such close contacts to protons of the protein receptor will increase the rate of T_2 and $T_{1\rho}$ relaxation of certain ligand protons in the complex, and will lead to line broadening of the ligand resonances involved.[34,35]

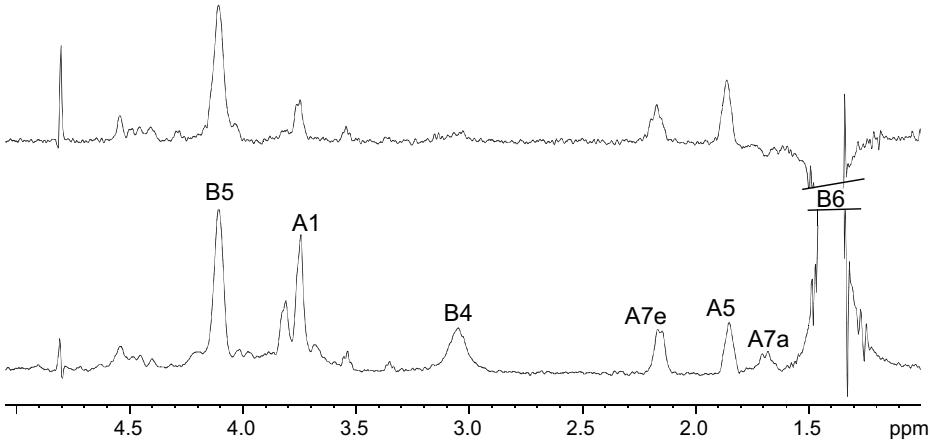


Figure 12. Cross-sections of trNOESY (lower trace, mixing time 200 ms) and trROESY (upper trace, spin-lock time 200 ms) spectra at the position of the B6 methyl resonance of **3**, in the presence of glucoamylase G2. The proton signals are labeled with a number, and a letter to indicate the ring (i.e., B6 instead of B H-6). The reduction of the peak intensity for the A1 resonance, and the change in sign of the A7a resonance, in the trROESY trace indicate spin-diffusion effects. In addition, pronounced line broadening of the B4 resonance in the trNOESY trace is observed; in the trROESY trace, this peak has completely relaxed during the spin-lock time in the ROESY experiment. The A1 resonance also shows line broadening, although it is not as pronounced as that of the B4 resonance. Reproduced with permission from Ref. 35. Copyright © 2000 Elsevier Science.

The studies described above illustrate the use of enhanced relaxation to identify ligand protons that are in close contact with protons of the protein receptor; this information contributes to knowledge of the bound topography or epitope (epitope mapping).

8.2. Saturation transfer difference NMR (STD-NMR)

A related concept to that described above is that of saturation transfer difference NMR (STD-NMR).[6,38–41] In this experiment, resonances of the protein are selectively saturated, and intramolecular NOE effects build up rapidly and are transferred throughout the protein by spin diffusion. Intermolecular NOE transfer to a binding ligand also occurs, resulting in a decrease of the ligand resonance intensities due to the buildup of negative NOE (Fig. 13). This decrease may be observed as enhancements in a difference spectrum, that is, the result of subtracting a 1D ^{1}H NMR spectrum of the ligand from a 1D spectrum recorded without saturation of the protein.[38–41] This allows determination of the areas of the ligand actually contacting the binding site (epitope), or in other words, epitope mapping.

This experiment was first described by Akasaka, who observed a decrease in the resonance intensities of adenosine 5′-diphosphate (ADP) in the presence of myosin, on saturation of the protein.[42] Recently, STD-NMR has been applied to determine the epitopes of oligosaccharides binding to plant lectins,[41,43] of bacterial carbohydrate antigens to antibodies,[44,45] of glycopeptide tumor-associated antigens to an anti-tumor antibody,[46] of a carbohydrate tumor-associated antigen to a lectin,[47] and of peptide ligands to human cell-surface glycoproteins embedded in liposome membranes.[48] Vogtherr and

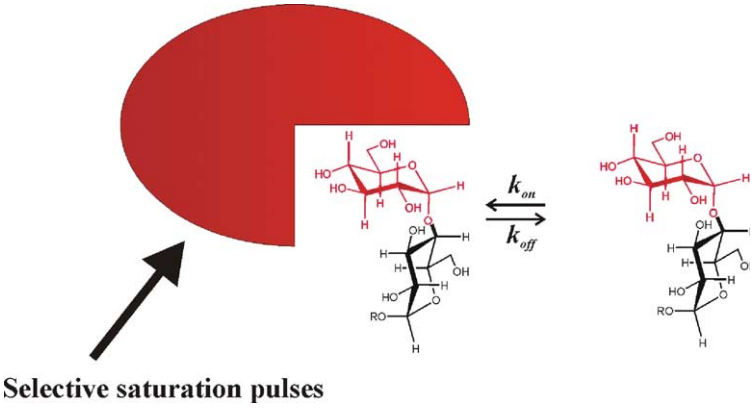


Figure 13. The STD-NMR experiment. Resonances of the protein are selectively saturated, and this saturation is quickly transferred by spin diffusion (buildup of negative NOE effects; represented by the red color) throughout the protein. Intermolecular NOE transfer also results in a decrease in the intensity of ligand resonances in close contact with the protein (red). These changes may be observed in a difference spectrum, allowing the identification of parts of a ligand contacting the protein (epitope).

Peters[49] have described the use of STD-NMR, in combination with chemical modification, for the determination of the binding specificity of a plant lectin. This provides information not only on which ligands are active, but also identifies hydroxyl groups of the saccharide residues that make important intermolecular contacts. This work also combined the saturation transfer difference technique with 2D NMR methods, for example, STD-TOCSY and STD-HMQC.

Another important application of STD-NMR is the identification of compounds with binding affinity to a receptor directly from complex mixtures of many compounds.[6,38,39]

9. Recent examples

9.1. Epitope mapping of a carbohydrate-mimetic peptide of the *Shigella flexneri* Y O-polysaccharide

The comparison of epitope mapping data obtained by STD-NMR and those obtained by X-ray crystallography is especially important both in validating the STD technique and also in identifying differences that may exist between complexes in the solution phase and in the crystal. STD-NMR epitope mapping of the carbohydrate-mimetic peptide MDWNMHAA (**4**) in complex with an anti-carbohydrate antibody directed against the O-polysaccharide of *S. flexneri* Y revealed certain features consistent with the high-resolution crystal structure, and certain features that were inconsistent.[50] For example, STD enhancements were observed for the side chains of residue Trp-3 (Figs. 14 and 15), consistent with its position buried within a hydrophobic pocket in the antibody combining site. Many other enhancements consistent with the crystal structure were also observed (Fig. 16). In contrast, certain other enhancements are difficult to explain: for example, the Met-1 γ-methylene proton at 2.3 ppm is not within 5 Å of any antibody proton in the crystal structure. Interestingly, no enhancements of the amide protons of the peptide were observed, although His-6 and Met-5 HN are both close to several protons of the antibody in the crystal structure. This observation represents a point of difference between the crystal structure and NMR data. A complicating factor is the presence in the crystal structure of ordered water molecules, several of which form hydrogen-bonded bridges involving amide protons of the

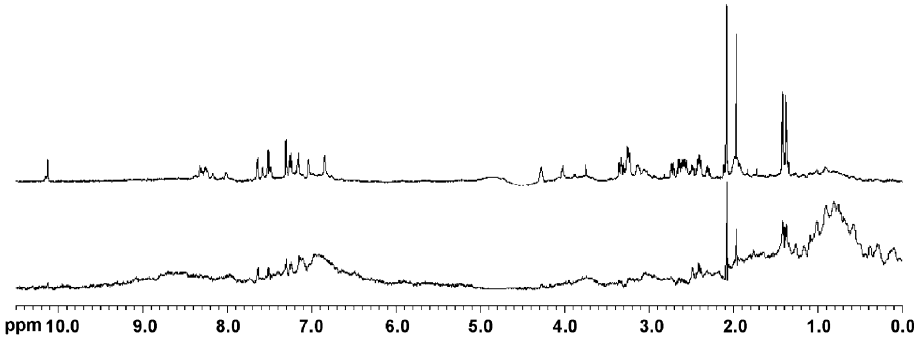


Figure 14. 1D ^{1}H NMR spectrum (top) and 1D STD-NMR spectrum (bottom) of the octapeptide (**4**) in the presence of the antibody. Reproduced with permission from Ref. 50. Copyright © 2004 Elsevier Science.

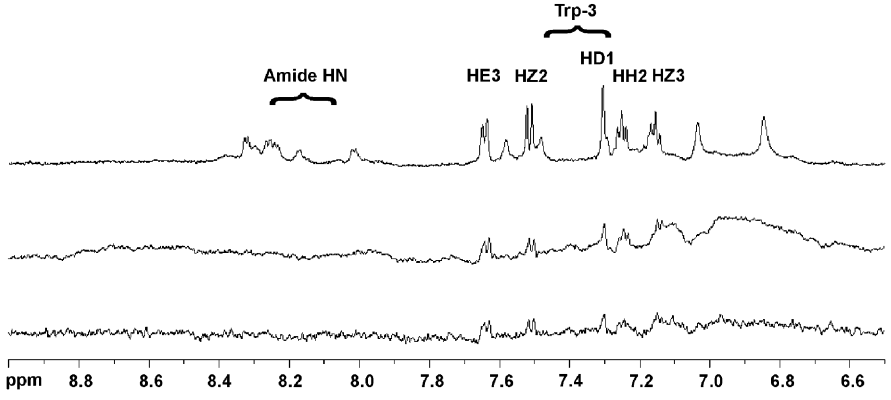


Figure 15. Expansions of the amide HN and aromatic region of the 1D ^{1}H NMR spectrum (top), 1D STD-NMR spectrum (middle), and 1D STD-NMR spectrum with spin-lock pulse (bottom) of the octapeptide (**4**) in the presence of the antibody. Reproduced with permission from Ref. 50. Copyright © 2004 Elsevier Science.

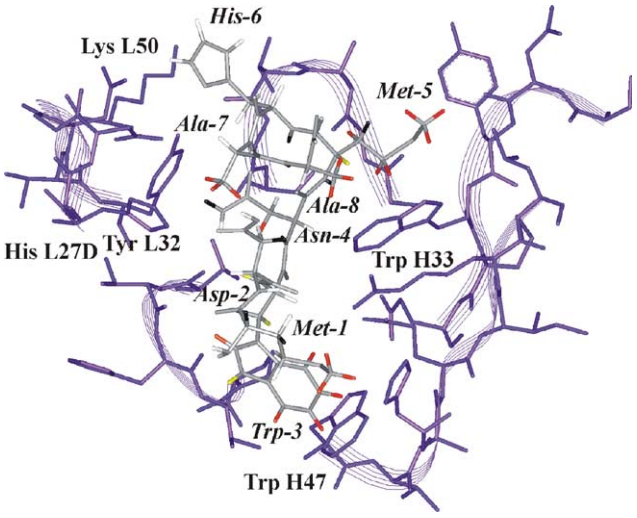


Figure 16. The crystal structure of the (**4**)/antibody complex, with STD-NMR intensities mapped onto the bound peptide. Residues of the antibody combining site are shown in purple, with selected residues labeled, and the direction of the backbone indicated in ribbon representation. Residues of the peptide are labeled in italics. Heavy atoms of the peptide are shown in gray, while the default color for hydrogen atoms is white. Observed STD-NMR intensities are mapped onto hydrogen atoms of the peptide by color, with red indicating 50–100% enhancement, orange 30–50% enhancement, and yellow <30% enhancement. Protons that are definitely *not* enhanced are shown in black; those for which no enhancement could be determined (due to interference by other resonances, or not observable in the 1D spectrum) remain white. Reproduced with permission from Ref. 50. Copyright © 2004 Elsevier Science.

peptide ligand. Thus, exchange effects mediated by water may be responsible for the lack of STD effects; alternatively, the peptide conformation in solution may differ somewhat from that observed in the crystal structure.

9.2. Epitope mapping of the Group A *Streptococcus* (GAS) cell-wall oligosaccharides

The epitope mapping of biologically active ligands by STD-NMR is extremely useful in the design of next-generation vaccines consisting of chemically well-defined structures presented as protein conjugates. An experiment combining selective 1D-TOCSY with the saturation transfer difference technique has been developed by our group and has been applied to the epitope mapping of oligosaccharides representing portions of the cell wall of Group A *Streptococcus* (GAS).[51] A branched trisaccharide and hexasaccharide (Fig. 17) were found to share the same pattern of STD enhancements (Figs. 18 and 19) with strongest enhancements for the branching Glc*N*Ac residue and weaker enhancements for the backbone rhamnose residues. Selective STD-1D-TOCSY experiments were employed to investigate enhancements observed in the 3–4 ppm region where many resonances overlap. All enhancements investigated in this manner revealed resonances of several spin systems, indicating that the rhamnose and Glc*N*Ac residues are all involved in contact with the antibody.

Mapping of the observed enhancements onto the bound conformation of the trisaccharide, previously determined by trNOE studies,[24] allowed the overall pattern of enhancements to be observed (Fig. 20). Importantly, the NMR data complemented immuno-

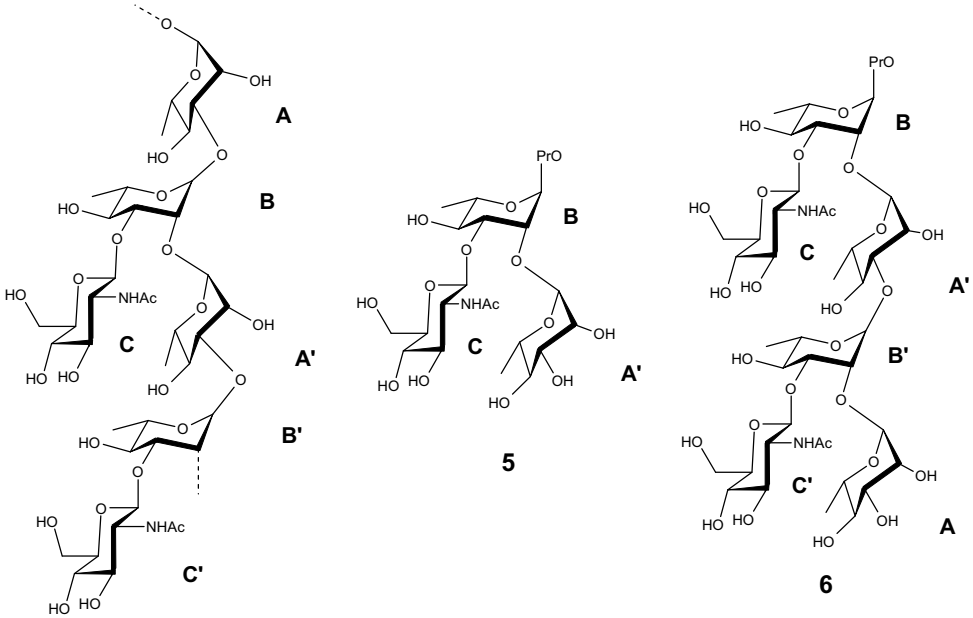


Figure 17. Structures of the Group A *Streptococcus* cell-wall polysaccharide, and of branched trisaccharide (**5**) and hexasaccharide (**6**) fragments recognized by antibodies.

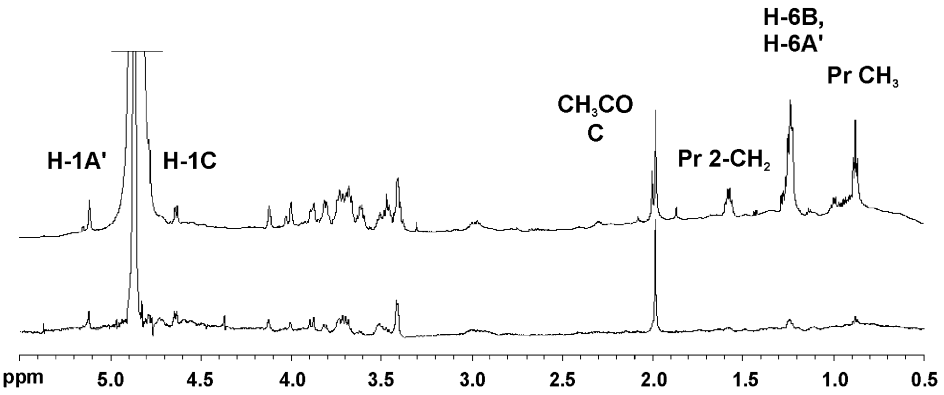


Figure 18. 1D ^{1}H NMR spectrum (top) and 1D STD-NMR spectrum (bottom) of the trisaccharide (**5**) in the presence of the antibody. Reproduced with permission from Ref. 51. Copyright © 2002 American Chemical Society.

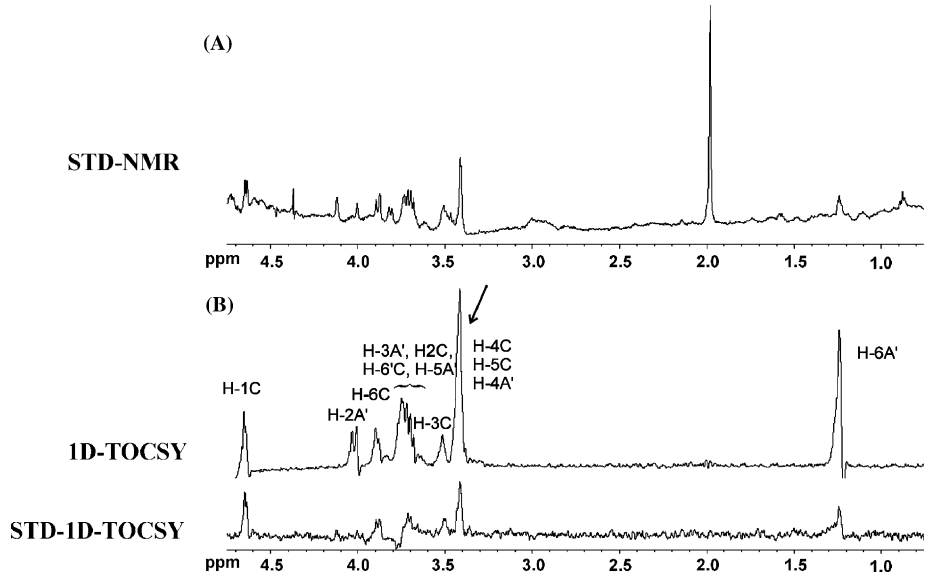


Figure 19. An enhancement at 3.4 ppm (A) in the 1D STD-NMR spectrum of (**5**) may indicate involvement of the H-4/5 protons of the C Glc*N*Ac residue, and/or the H-4 proton of the A′ residue. The STD-1D-TOCSY spectrum (B) reveals contributions from both residues. A 1D-TOCSY spectrum is shown for reference. Reproduced with permission from Ref. 51. Copyright © 2002 American Chemical Society.

logical studies that had shown that, although anti-GAS antibodies may bind weakly to Glc*N*Ac, and Glc*N*Ac is the immunodominant monosaccharide residue, larger oligosaccharides—minimally, a pentasaccharide—were required for effective production of anti-GAS antibodies.[52] Therefore, the topography presented by the cell-wall polysaccharide (CWPS) (Fig. 21)[53] and uniquely recognized by antibodies is an extended surface such as that shown for the hexasaccharide **6** (Fig. 20C) with backbone Rha and branching Glc*N*Ac residues all involved in defining the spatial features characteristic of GAS CWPS. It follows that conjugate vaccines must include fragments of the CWPS large enough to comprise such an extended surface.

Furthermore, the NMR data allowed the refinement of a previously constructed theoretical model of the Ab-oligosaccharide complex in which the C residues were buried most deeply in the site, while the B Rha residues contacted the antibody more peripherally and no contact was made by the A Rha residues.[53] The STD data allowed us to observe that in fact both A and B Rha residues are involved in the epitope.

9.3. Bound conformation and epitope mapping of a peptide mimic of the GAS cell-wall polysaccharide

The combination of several of the above techniques is illustrated by studies of the binding of a peptide molecular mimic of the CWPS to another anti-GAS antibody. Transferred nuclear Overhauser effect (trNOE) experiments were performed at 800 MHz to investigate the bound conformation of the hexapeptide DRPVPY **7**, a functional molecular mimic of the Group A *Streptococcus* cell-wall polysaccharide (Fig. 22).[54] The

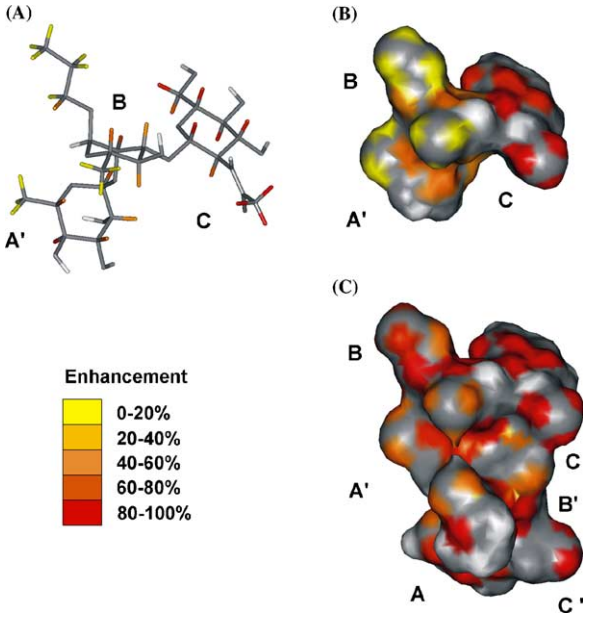


Figure 20. (A) Bound conformation of the trisaccharide (**5**) derived by trNOE, with protons enhanced in the STD-NMR and STD-1D-TOCSY spectra, which form the epitope, colored according to the degree of enhancement in the STD spectra. (B) A molecular surface representation of the epitope. (C) A molecular surface representation of the hexasaccharide (**6**), with the top branched trisaccharide unit in the same orientation as those in parts A and B, colored according to the degree of enhancement in the STD spectra. Reproduced with permission from Ref. 51. Copyright © 2002 American Chemical Society.

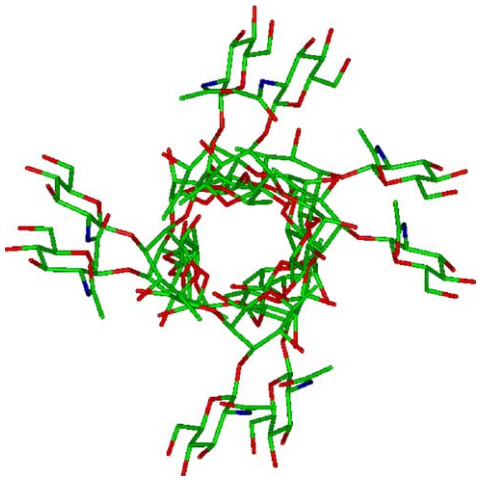


Figure 21. A model of the cell-wall polysaccharide (CWPS) of the Group A *Streptococcus* (GAS) looking down the barrel of the helix. The core of the helix is composed of α-L-rhamnose residues with the β-D-*N*-acetylglucosamine moieties displayed on the periphery.

hexapeptide **7** binds to the monoclonal antibody SA-3, mimicking the branched trisaccharide repeating unit, L-Rha-α-(1 → 2)-(D-Glc*N*Ac-β-(1 → 3))-α-L-Rha. Transferred NOE experiments revealed that the peptide adopts a tight turn conformation in the bound state, with close contacts between the side chains of valine and tyrosine (Fig. 23). QUIET-trNOESY experiments were used to confirm the validity of the observed contacts and

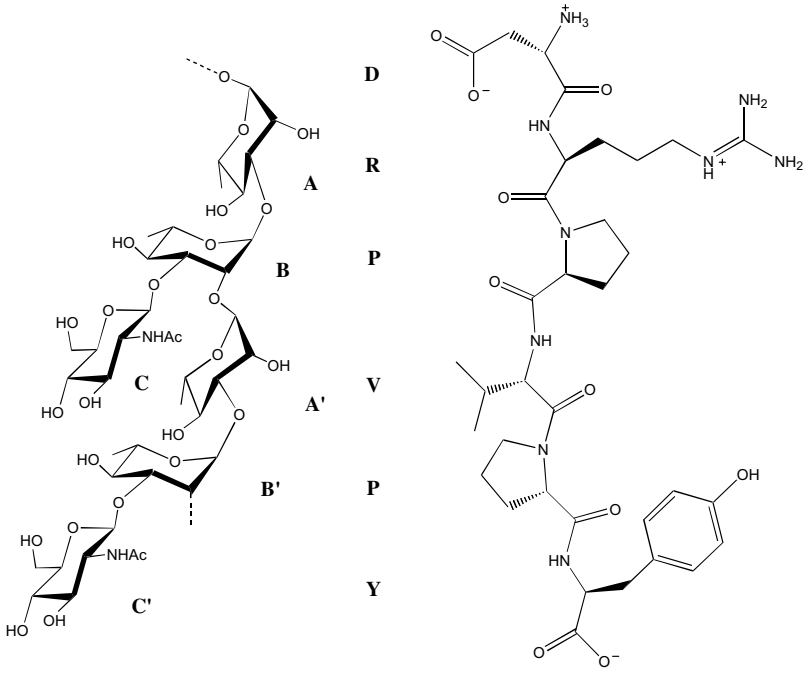


Figure 22. Structures of the cell-wall polysaccharide (CWPS) of Group A *Streptococcus* (left, with monosaccharide residues of the branched trisaccharide repeating unit labeled A, B, and C); and of a peptide molecular mimic of the CWPS, the hexapeptide DRPVPY **7** (single letter amino acid code).

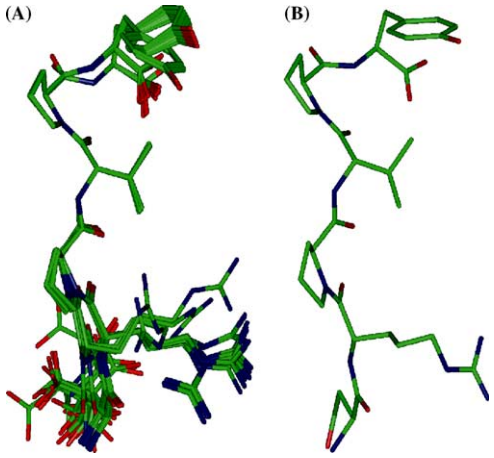


Figure 23. Bound conformation of the hexapeptide DRPVPY **7** in complex with the monoclonal anti-carbohydrate antibody SA-3. (A) A family of 27 calculated structures, with backbone atoms of residues 4–6 superimposed. (B) Average of the 27 structures shown in panel (A). Residues 4–6 adopt a well-defined tight turn. Reproduced with permission from Ref. 54. Copyright © 2002 American Chemical Society.

to evaluate the presence of spin-diffusion pathways. Interestingly, only one weak trNOE contact was found to disappear in the QUIET-trNOESY spectrum, indicating that it had been caused by spin diffusion (Fig. 24); removal of the distance restraint corresponding to this contact did not affect the calculated structure.

Interestingly, the QUIET-trNOESY spectra described above showed residual trNOE peaks within the bands connecting the quiet windows. The absence of such peaks in particular spectra, depending on which bands had been inverted, was proposed as an indicator of spin-diffusion pathways. For example, Figure 25 depicts traces taken at the F_1 frequency of Pro-3 Hα. A residual contact to Val-4 Hγ is present in all of the spectra except that shown in trace A, in which Val HN is also inverted; the disappearance of this effect indicates a possible spin-diffusion pathway from Pro-3 Hα to Val-4 HN to Val-4 Hγ.

The residual contacts observed, and their intensities, varied depending on the positions of the inversion bands. This phenomenon may be due to second-order

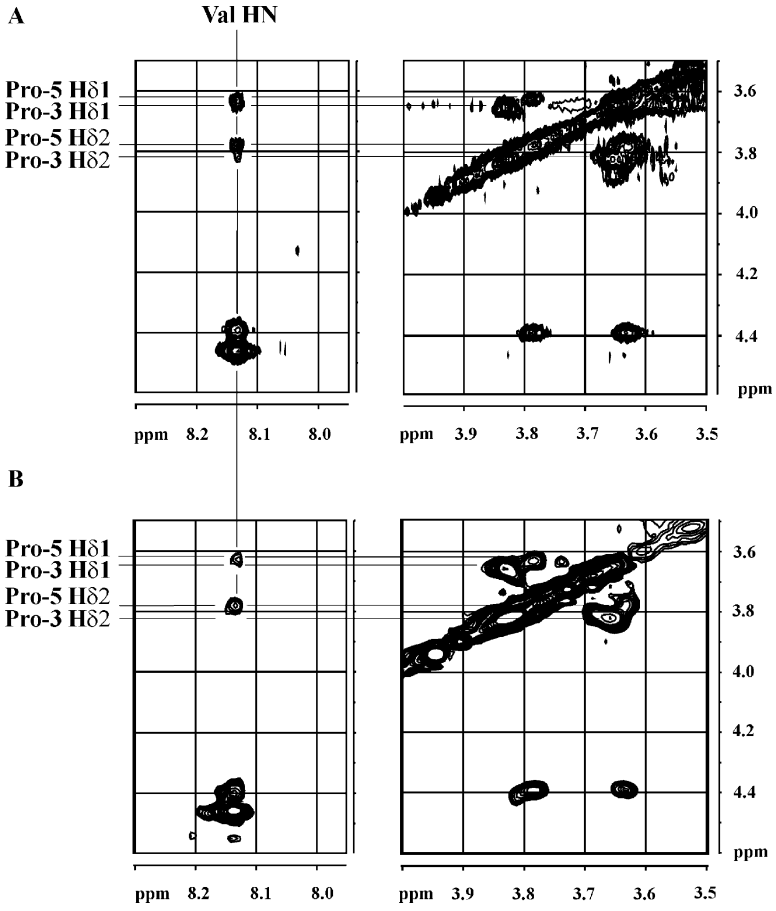


Figure 24. TrNOESY and QUIET-trNOESY spectra of the hexapeptide DRPVPY **7** in the presence of the monoclonal anti-carbohydrate antibody SA-3. (A) Regions of a trNOESY spectrum ($\tau_m = 200$ ms) showing Val-4 HN–Pro-3,-5 Hδ cross-peaks and Pro-3,-5 Hδ–Hδ cross-peaks. (B) Corresponding regions of a QUIET-trNOESY spectrum ($\tau_m = 200$ ms) with inversion of 0.75 ppm wide bands centered at 4.15 and 8.1 ppm; within the intersection of the quiet bands (quiet window), the Val HN–Pro-5 Hδ cross-peaks are still present, while the Val HN–Pro-3 $H\delta_2$ cross-peak is absent, indicating cancellation due to spin diffusion. Reproduced with permission from Ref. 54. Copyright © 2002 American Chemical Society.

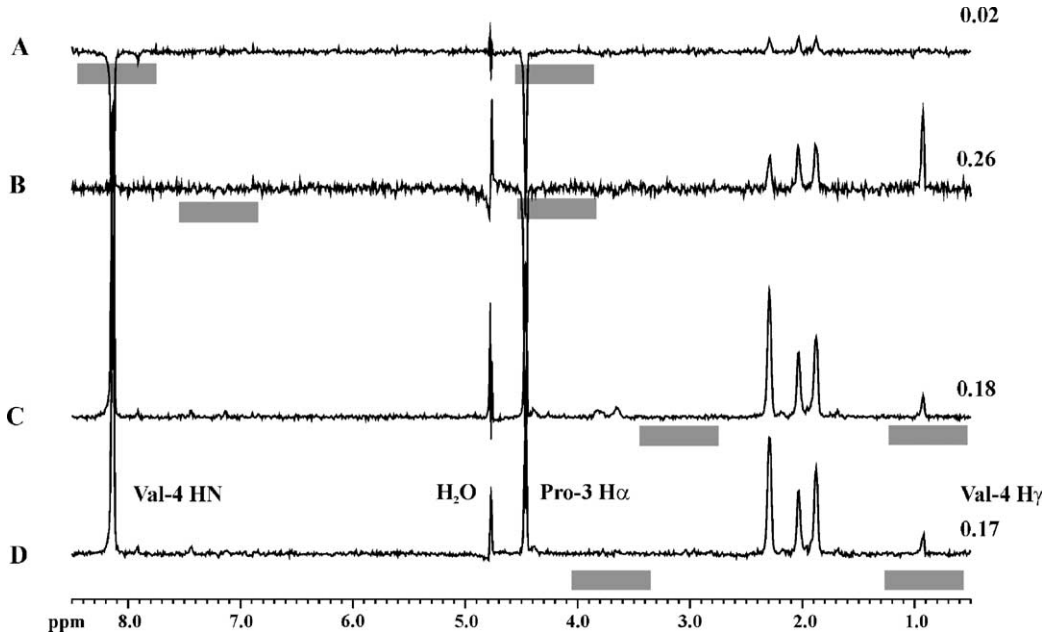


Figure 25. Traces from 2D QUIET-trNOESY spectra (DRPVPY/mAb SA-3) taken at 4.46 ppm in F_1, Pro-3 Hα resonance frequency. Inverted bands are indicated by shaded bars below the traces. In each trace the absolute intensity of the Val-4 Hγ resonance, relative to the Pro-3 Hα diagonal resonance in panel (C), is given. This signal is present in panels (B)–(D) due to incomplete cancellation (second-order effects). However, in trace A, when Val-4 HN is also inverted, the Val-4 Hγ signal becomes completely cancelled, indicating a possible spin-diffusion pathway from Pro-3 Hα to Val-4 HN to Val-4 Hγ. Reproduced with permission from Ref. 54. Copyright © 2002 American Chemical Society.

effects: incomplete cancellation may result because the rate of NOE buildup is not identical in the first and second halves of the mixing time; consequently, for most purposes the shortest mixing time possible should be used. Relaxation effects during the doubly band-selective inversion pulse may also contribute.

More information about the mode of peptide binding was provided by STD-NMR experiments. Enhancements of nearly all residues of the peptide were observed (Fig. 26), with the exception of Asp-1. Since the contribution of Asp-1 to binding is minimal, it appeared to be an excellent position at which to conjugate the peptide to proteins for the preparation of a conjugate vaccine. Synthesis of such peptide conjugates and confirmation of strong binding to an anti-GAS antibody have recently been achieved (unpublished results).

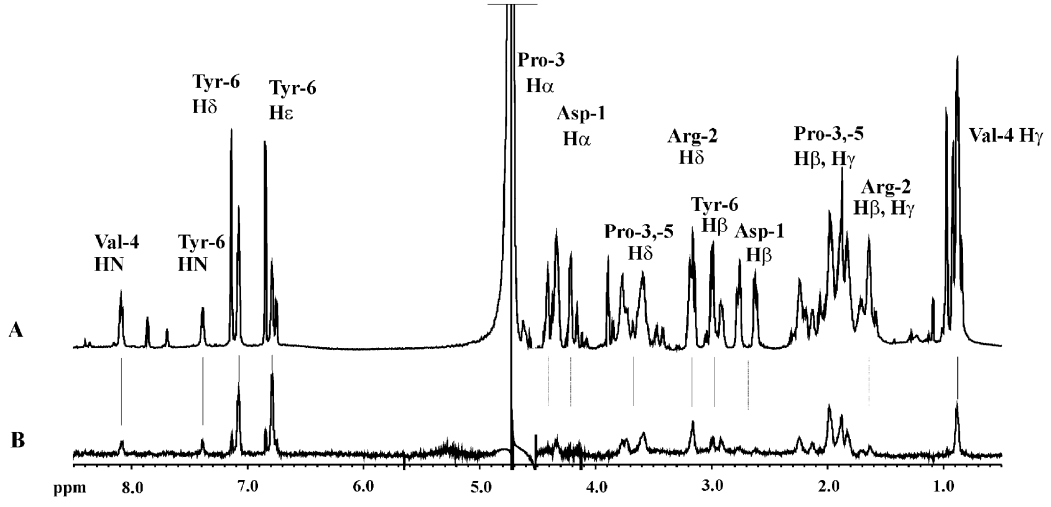


Figure 26. (A) 1D ^{1}H NMR spectrum of the hexapeptide DRPVPY **7** in the presence of the monoclonal anti-carbohydrate antibody SA-3. (B) 1D STD-NMR spectrum of the peptide in the presence of the antibody, showing enhancements of resonances of protons making close contacts with the antibody combining site. Reproduced with permission from Ref. 54. Copyright © 2002 American Chemical Society.

9.4. Bioactive conformation of a glycosidase inhibitor bound to glucoamylase G2

Molecular modeling approaches may be applied to gain further information about these molecular systems. The advantages of combining NMR spectroscopic and molecular modeling approaches are illustrated by a recent study of the binding of a sulfonium-ion glycosidase inhibitor **8** to glucoamylase G2.[55] Transferred NOE experiments were performed and a relaxation matrix approach was employed to calculate a bound conformation. The sulfonium ion **8**, a novel compound designed as an analogue of castanospermine, was found to bind in an unusual, high-energy $^{1,4}B$ conformation (Fig. 27), while the major conformer populated in the free state had a $^{1}C_4$ conformation of the six-membered ring. Thus, neither the bound nor the free states of the compound resemble the ground-state conformation of castanospermine ($^{4}C_1$). The conformation of castanospermine when bound to GA is unknown; however, it adopts a $^{1}S_3$ skew conformation when bound to a fungal glucanase.

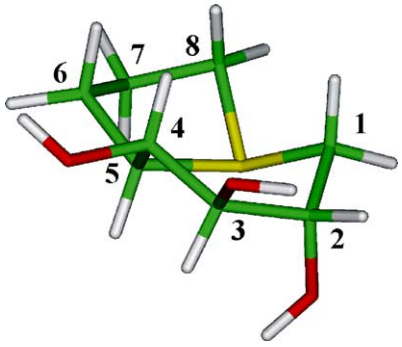


Figure 27. Refined $^{1,4}B$ conformation of the inhibitor **8** bound to glucoamylase G2, determined by trNOE NMR spectroscopy. Reproduced with permission from Ref. 55. Copyright © 2003 American Chemical Society.

The reasons for the selection of this unusual $^{1,4}B$ conformation were explored by computational docking using the program AUTODOCK 3.0, using coordinates from a crystal structure of glucoamylase. A possible binding mode within subsite 1 was proposed, with an electrostatic interaction between the sulfonium-ion center and the catalytic base Glu 400 (Fig. 28). The physical constraints of the shape of this site can only accommodate a relatively flat molecule, such as **8** in the

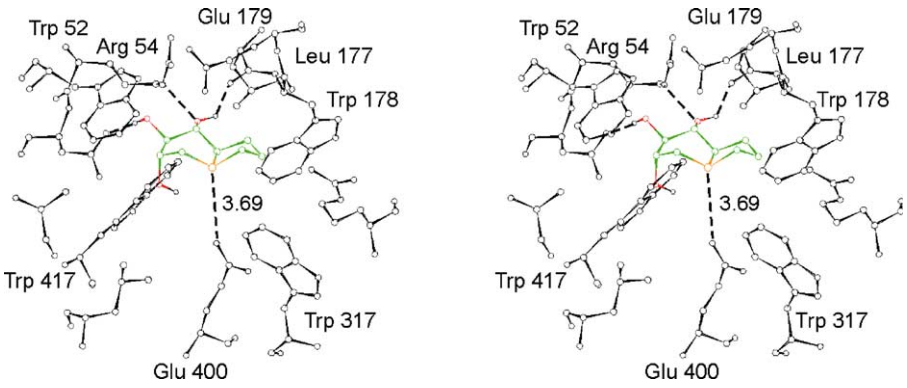


Figure 28. Putative binding modes of the inhibitor **8** to glucoamylase G2. Hydrogen bonds and important electrostatic interactions are shown as dashed black lines. Residues of the binding site are shown in black. (A) Predicted binding mode of **8** in the $^{1,4}B$ conformation determined by trNOE NMR spectroscopy. An electrostatic interaction between S^{+} and the catalytic base Glu 400 is predicted. Reproduced with permission from Ref. 55. Copyright © 2003 American Chemical Society.

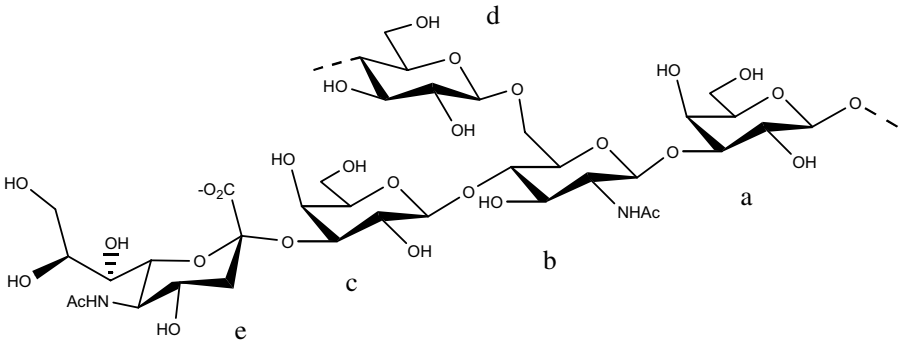


Figure 29. Repeating unit of the Group B streptococcal type III capsular polysaccharide (GBSPIII).

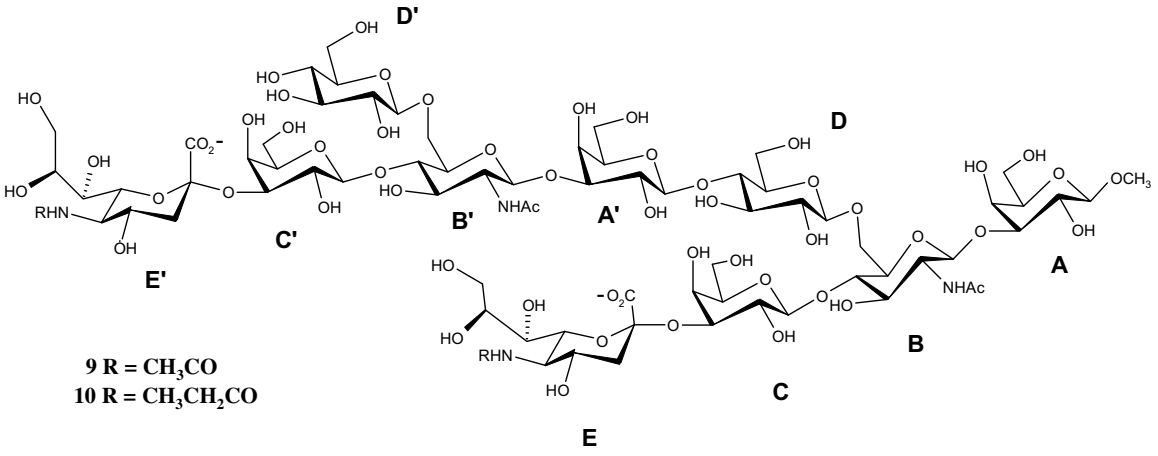


Figure 30. Decasaccharides **9** and **10** representing two repeating units of the Group B *Streptococcus* type III capsular polysaccharide.

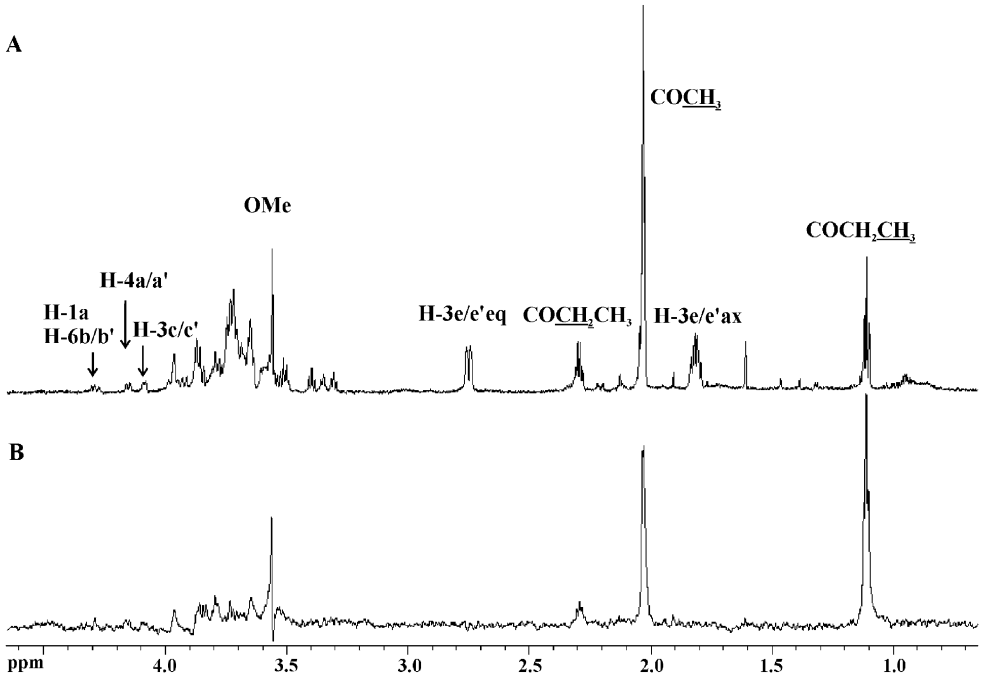


Figure 31. Epitope mapping of decasaccharides **9** and **10** in the presence of mAb S9. (A) 1D NMR spectrum (800 MHz, 298 K) of decasaccharides **9** and **10** in the presence of mAb S9. (B) 1D-STD-NMR spectrum (800 MHz, 298 K) of decasaccharides **9** and **10** in the presence of mAb S9. Selected resonances are labeled in spectrum (A); all labeled resonances are enhanced in spectrum (B), with the important exception of the H-3 proton resonances of sialic acid (e/e′). For the H-3 resonances, eq and ax denote the equatorial and axial H-3 resonances, respectively. Reproduced with permission from Ref. 56. Copyright © 2003 American Society for Biochemistry and Molecular Biology, Inc.

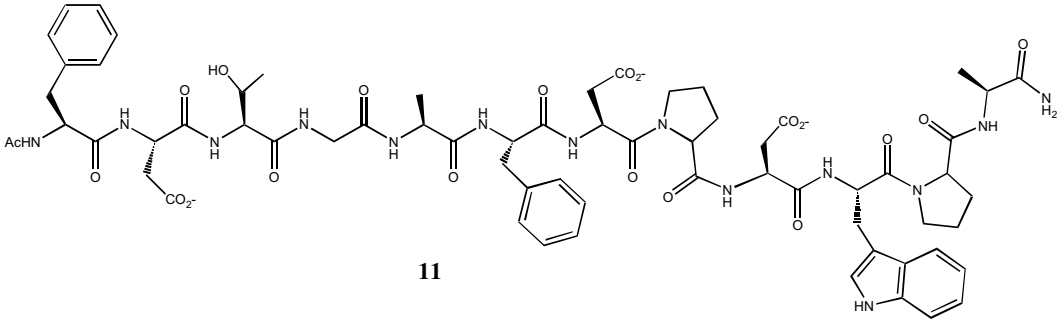


Figure 32. The dodecapeptide FDTGAFDPDWPA (**11**), an immunologically functional molecular mimic of the GBS type III capsular polysaccharide.

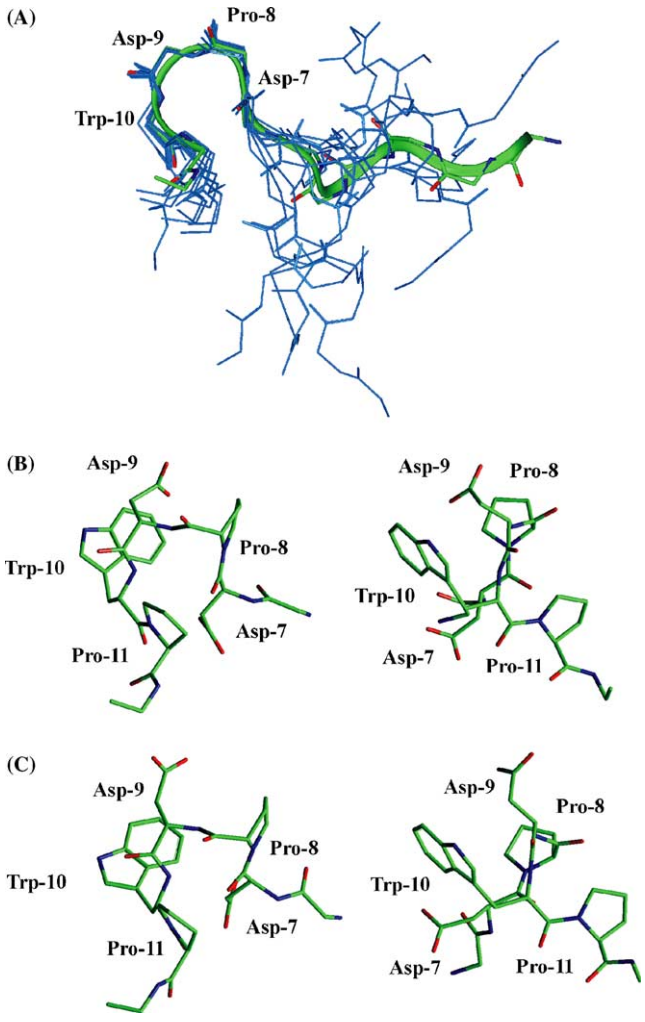


Figure 33. Views of the calculated bound structure of the peptide FDTGAFDPDWPA (**11**). (A) Backbone (N, C_α, C, O) atoms of the 10 structures with lowest NOE restraint energy produced by simulated annealing refinement, superimposed using backbone atoms of residues 7–10. The lowest-energy structure is colored by atom (C, green; N, blue; O, red) and a ribbon representation is shown in green, while the other structures are shown in light blue. Residues 7–10 comprise a type I β-turn and are labeled. (B) Close-up views of residues 6–12 of the lowest-energy calculated structure (colored by atom). Backbone atoms only are shown for residues 6 and 12. In the left panel, the turn is shown in a similar orientation to the structures in (A); in the right panel, it is shown in an orthogonal orientation, so that the exposed position of the Trp-10 side chain, on one face of the turn, is visible. (C) Close-up views of residues 6–12, as in (B), but in a typical type I β-turn conformation, as observed for the free peptide in solution. The conformation is very similar to the calculated bound conformation (B). Reproduced with permission from Ref. 56. Copyright © 2003 American Society for Biochemistry and Molecular Biology, Inc.

$^{1,4}B$ conformation. Other possible conformations such as the global minimum $^{1}C_4$ conformation or the alternative chair $^{4}C_1$ conformation were both too bulky to fit within this subsite. Thus, molecular modeling provided a potential explanation for the mode of action of this inhibitor. The observations, based on NMR data, that the bioactive conformer comprises only a small fraction of the population of the free compound, suggest a possible reason for its weak inhibition, and suggest avenues for the design of new inhibitors. More generally, these observations, taken together with other recent structural studies, suggest that the modes of action of alkaloid and sulfonium glycosidase inhibitors cannot be inferred simply by configurational similarity, either to sugars or to each other.

9.5. Epitope mapping of oligosaccharides corresponding to the capsular polysaccharide of type III Group B *Streptococcus* (GBS)

NMR spectroscopic techniques have also been used to investigate the interactions of ligands with an antibody directed against Group B *Streptococcus*.[56] The capsular polysaccharide of type III GBS has a complex branched structure with terminal sialic acid residues (Fig. 29), and it has been proposed that antibodies directed against it recognize a conformational epitope, comprising secondary structures that, as in the case of the sulfonium ion described above, comprise a very minor portion of the conformations populated in solution. Complementarity to these minor conformations would avoid cross-reaction with sialic-acid-containing structures commonly found in human tissues, and the resulting harmful effects. Interestingly, although sialic acid is not an immunodominant monosaccharide in this case, it is essential for the formation of this conformational epitope. Epitope mapping by STD-NMR of 'natural' decasaccharides and decasaccharides modified by N-propionylation on sialic acid (Fig. 30) revealed an interesting pattern whereby strong enhancements were observed for the *N*-propionyl groups of sialic acid, and of the *N*-acetyl groups of Glc*N*Ac and/or sialic acid, but no enhancements for the H-3 protons of sialic acid, indicating that the sialic acid was required to present the *N*-acyl functionality to the antibody, but its monosaccharide ring was not involved in binding (Fig. 31). This represents some evidence for presentation of a conformational epitope.

Transferred NOE effects were not observed in this system, likely due to unusual binding kinetics.

9.6. Bound conformations and epitope mapping of peptide mimics of the capsular polysaccharide of type III Group B *Streptococcus* (GBS)

TrNOE experiments were used to investigate the binding of a peptide mimic of the GBSIII capsular polysaccharide (Fig. 32).[56] The dodecapeptide FDTGAFDPDWPA (**11**) behaves as an immunologically functional molecular mimic; immunization of mice with this peptide induces an anti-polysaccharide immune response. The peptide was found to adopt a type I β-turn conformation spanning residues Asp-7–Trp-10; in addition,

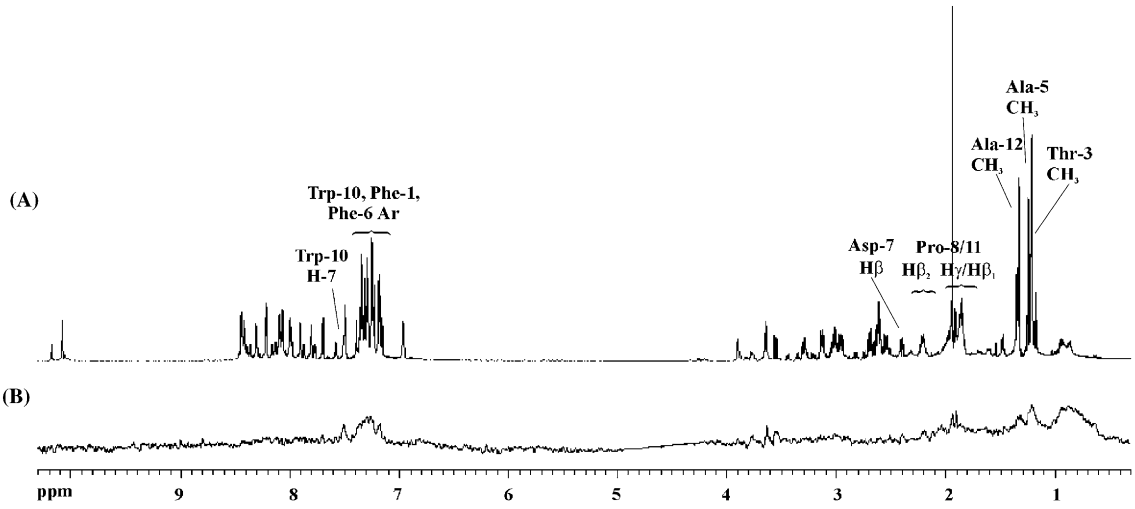


Figure 34. Epitope mapping of FDTGAFDPDWPA (**11**) in the presence of mAb S9. Resonances enhanced in the STD-NMR spectrum are labeled. (A) 1D NMR spectrum (800 MHz, 298 K) of the peptide with mAb S9. (B) 1D-STD-NMR spectrum (600 MHz, 310 K) of the peptide with mAb S9. Reproduced with permission from Ref. 56. Copyright © 2003 American Society for Biochemistry and Molecular Biology, Inc.

STD-NMR revealed the involvement of Trp-10 and several other residues in this region in contact with the antibody (Figs. 33 and 34). Importantly, this type I β-turn was also significantly populated in the absence of antibody, indicating a strong intrinsic propensity to adopt the bioactive conformation (Fig. 33B). This strong preference likely contributes to the peptide's immunological effectiveness.

9.7. A novel protein modeling protocol

Finally, we would like to describe a novel application of the knowledge of bound-ligand conformations to the problem of protein structure prediction.[57] Comparative modeling (homology modeling) is widely used to generate theoretical models of protein structures, based on known 3D crystal or NMR structures of related proteins. Comparative models may be quite accurate when the entire protein structure is considered, having rmsds of ~1 Å. However, this accuracy decreases with the level of sequence similarity, and importantly, certain regions such as loops, which vary greatly in sequence and structure, may differ widely from the known protein and may therefore be predicted incorrectly. These regions often contribute to binding or active sites, and accurate modeling of these regions is very important if models are to be used to draw conclusions about protein–ligand interactions and to design new ligands. We have intro-

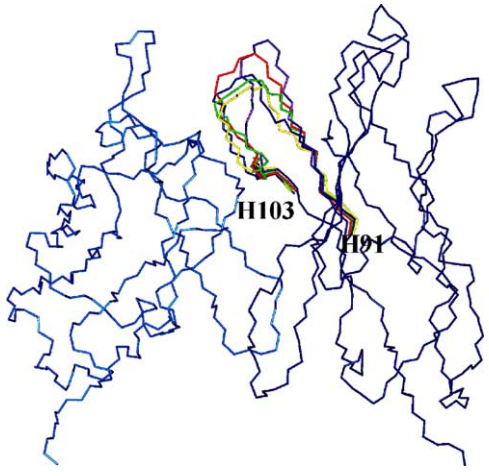


Figure 35. Representative members of a library of protein models that were subjected to screening by computational docking with ligands. The framework held constant for all models is shown in blue, while the loop region which differed between the models is shown in colors (five loops are shown out of a library of 12). Reproduced with permission from Ref. 57. Copyright © 2003 American Chemical Society.

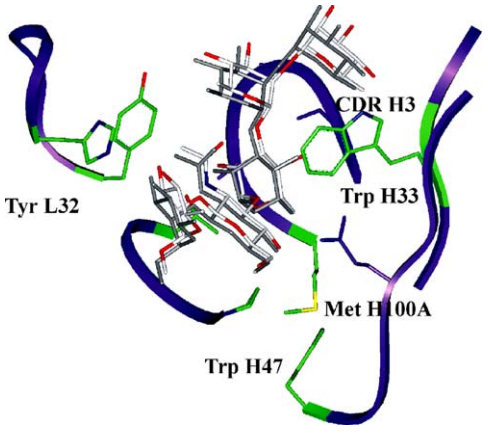


Figure 36. A view of the pentasaccharide ligand docked in the binding site of the most accurate antibody model, with the correct H-3 loop conformation (labeled CDR H3; loop rmsd of Cα atoms = 0.69 Å). The ligand is positioned accurately (the position determined experimentally by crystallography is shown in gray for reference). Reproduced with permission from Ref. 57. Copyright © 2003 American Chemical Society.

duced a method in which the known bound conformation of a ligand, which may be derived by trNOE experiments, is used to screen a library of protein models, each with a different loop conformation, by molecular docking. It was shown that the accuracy of fit (complementarity) between the ligand and the protein model is an effective scoring function to identify the most accurate models out of a series, allowing discrimination between a series of protein models that were all identical if conventional scoring functions such as sequence homology or force field energy were applied. This concept was demonstrated using an anti-carbohydrate antibody (illustrated in Figs. 35–38).

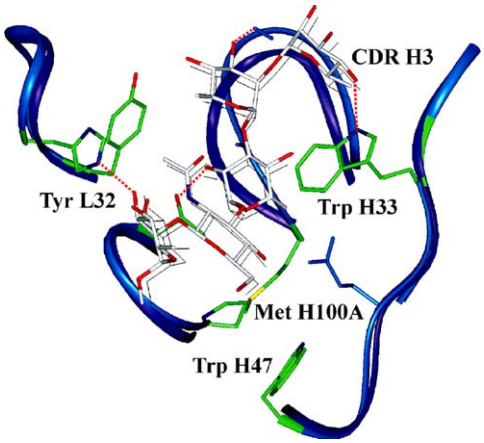


Figure 37. A view of the pentasaccharide ligand docked in the binding site of a slightly less accurate antibody model (loop rmsd of C_α atom positions = 1.72 Å). The backbone of the model's binding site is shown in blue with the crystal structure in purple; hydrogen bonds are indicated by red dashes. The ligand is positioned accurately (the position determined experimentally by crystallography is shown in thin gray lines). Many of the important interactions are predicted correctly, as the change in the conformation of CDR H3 is relatively small. A favorable calculated binding energy results, allowing this antibody model to be identified as being one of the more accurate in the library. Reproduced with permission from Ref. 57. Copyright © 2003 American Chemical Society.

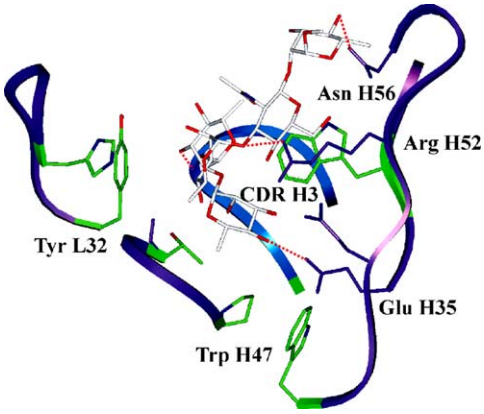


Figure 38. A view of the pentasaccharide ligand docked in the binding site of a much less accurate antibody model (loop rmsd of C_α atom positions = 2.88 Å). The very different conformation of CDR H3, highlighted in blue, results in a much smaller binding site that can accommodate only two residues of the pentasaccharide. None of the correct hydrogen bonds can be formed, and the Rha B and C and Glc*N*Ac D methyl groups are exposed to the solvent. These effects are correctly evaluated by the docking program as resulting in a very poor binding energy, allowing this antibody model to be identified as being inaccurate. Reproduced with permission from Ref. 57. Copyright © 2003 American Chemical Society.

10. Summary

The combination of NMR experiments and molecular modeling techniques described above provides important information regarding protein–ligand interactions. Transferred NOE experiments allow the determination of bound-ligand conformations, and epitope mapping data provide information about which parts of the ligand contact the protein. The latter data are critical for accurate molecular modeling of the ligand in the binding site of the protein. Appropriate control experiments should be performed to ensure the accuracy of the trNOE data. In combination with information about the protein structure, this information allows the mode of interaction within the complex to be determined.

References

1. Varki, A. *Glycobiology* **1993**, *3*, 97–130.
2. Yarema, K. J.; Bertozzi, C. R. *Curr. Opin. Chem. Biol.* **1998**, *2*, 49–61.
3. Drickamer, K.; Taylor, M. E. *Trends Biochem. Sci.* **1998**, *23*, 321–324.
4. Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995**, *28*, 321–327.
5. Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637–674.
6. Meyer, B.; Peters, T. *Angew. Chem., Int. Ed.* **2003**, *42*, 864–890.
7. Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Quart. Rev. Biophys.* **1999**, *3*, 211–240.
8. Kogelberg, H.; Solís, D.; Jiménez-Barbero, J. *Curr. Opin. Struct. Biol.* **2003**, *13*, 646–653.
9. Post, C. B. *Curr. Opin. Struct. Biol.* **2003**, *13*, 581–588.
10. Pellecchia, M.; Sem, D. S.; Wüthrich, K. *Nat. Rev. Drug Discov.* **2002**, *1*, 211–219.
11. Zartler, E. R.; Yan, J.; Mo, H.; Kline, A. D.; Shapiro, M. J. *Curr. Top. Med. Chem.* **2003**, *3*, 25–37.
12. Yan, J.; Kline, A. D.; Mo, H.; Zartler, E. R.; Shapiro, M. J. *J. Am. Chem. Soc.* **2002**, *124*, 9984–9985.
13. Zartler, E. R.; Hanson, J.; Jones, B. E.; Kline, A. D.; Martin, G.; Mo, H.; Shapiro, M. J.; Wang, R.; Wu, H.; Yan, J. *J. Am. Chem. Soc.* **2003**, *125*, 10941–10946.
14. *NMR Spectroscopy of Glycoconjugates*; Jimenez-Barbero, J., Peters, T., Eds.; Wiley-VCH: Weinheim, 2003.
15. Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555–578.
16. Solomon, I. *Phys. Rev.* **1955**, *99*, 559–565.

17. Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, 2nd ed.; Wiley-VCH: Weinheim, 2000.
18. Balaram, P.; Bothner-By, A. A.; Dadok, J. *J. Am. Chem. Soc.* **1972**, *94*, 4015–4016.
19. Balaram, P.; Bothner-By, A. A.; Breslow, E. *J. Am. Chem. Soc.* **1972**, *94*, 4017–4018.
20. Albrand, J. P.; Birdsall, B.; Feeney, J.; Roberts, G. C. K.; Burgen, A. S. V. *Int. J. Biol. Macromol.* **1979**, *1*, 37–41.
21. Cayley, P. J.; Albrand, J. P.; Feeney, J.; Roberts, G. C. K.; Piper, E. A.; Burgen, A. S. V. *Biochemistry* **1979**, 3886–3895.
22. Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1982**, *48*, 402–417.
23. Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1983**, *53*, 423–442.
24. Weimar, T.; Harris, S. L.; Pitner, J. B.; Bock, K.; Pinto, B. M. *Biochemistry* **1995**, *34*, 13672–13681.
25. Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811–813.
26. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *63*, 207–213.
27. Glaudemans, C. P. J.; Lerner, L.; Daves, G. D., Jr.; Kovac, P.; Venable, R.; Bax, A. *Biochemistry* **1990**, *29*, 10906–10911.
28. Arepalli, S. R.; Glaudemans, C. P. J.; Daves, G. D., Jr.; Kovac, P.; Bax, A. *J. Magn. Reson. B* **1995**, *106*, 195–198.
29. Macura, S.; Fejzo, J.; Hoogstraten, C. G.; Westler, W. M.; Markley, J. L. *Isr. J. Chem.* **1992**, *32*, 245–256.
30. Massefski, W.; Redfield, A. G. *J. Magn. Reson.* **1988**, *78*, 150–155.
31. Zwahlen, C.; Vincent, S. J. F.; Di Bari, L.; Levitt, M. H.; Bodenhausen, G. *J. Am. Chem. Soc.* **1994**, *116*, 362–368.
32. Vincent, S. J. F.; Zwahlen, C.; Bodenhausen, G. *J. Biomol. NMR* **1996**, *7*, 169–172.
33. Scherf, T.; Anglister, J. *Biophys. J.* **1993**, *64*, 754–761.
34. Weimar, T.; Stoffer, B.; Svensson, B.; Pinto, B. M. *Biochemistry* **2000**, *39*, 300–306.
35. Weimar, T.; Petersen, B. O.; Svensson, B.; Pinto, B. M. *Carbohydr. Res.* **2000**, *326*, 50–55.
36. Bock, K.; Meldal, M.; Refn, S. *Carbohydr. Res.* **1991**, *221*, 1–16.
37. Aleshin, A. E.; Stoffer, B.; Firsov, L. M.; Svensson, B.; Honzatko, R. B. *Biochemistry* **1996**, *35*, 8319–8328.
38. Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784–1788.
39. Klein, J.; Meinecke, R.; Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **1999**, *121*, 5336–5337.
40. Peters, T.; Meyer, B. Ger. Patent DE19649359; Swiss Patent 690695; U.K. Patent GB23211401; U.S. Patent pending.
41. Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108–6117.
42. Akasaka, K. *J. Magn. Reson.* **1979**, *36*, 135–140.
43. Haselhorst, T.; Weimar, T.; Peters, T. *J. Am. Chem. Soc.* **2001**, *123*, 10705–10714.
44. Maaheimo, H.; Kosma, P.; Brade, L.; Brade, H.; Peters, T. *Biochemistry* **2000**, *39*, 12778–12788.
45. Kooistra, O.; Herfurth, L.; Lüneberg, E.; Frosch, M.; Peters, T.; Zähringer, U. *Eur. J. Biochem.* **2002**, *269*, 573–582.
46. Möller, H.; Serttas, N.; Paulsen, H.; Burchell, J. M.; Taylor-Papadimitriou, J.; Meyer, B. *Eur. J. Biochem.* **2002**, *269*, 1444–1455.
47. Weimar, T.; Bukowski, R.; Young, N. M. *J. Biol. Chem.* **2000**, *275*, 37006–37010.
48. Meinecke, R.; Meyer, B. *J. Med. Chem.* **2001**, *44*, 3059–3065.
49. Vogtherr, M.; Peters, T. *J. Am. Chem. Soc.* **2000**, *122*, 6093–6099.
50. Johnson, M. A.; Pinto, B. M. *Bioorg. Med. Chem.* **2004**, *12*, 295–300.
51. Johnson, M. A.; Pinto, B. M. *J. Am. Chem. Soc.* **2002**, *124*, 15368–15374.
52. Reimer, K. B.; Gidney, M. A. J.; Bundle, D. R.; Pinto, B. M. *Carbohydr. Res.* **1992**, *232*, 131–142.
53. Pitner, J. B.; Beyer, W. F.; Harris, S. L.; Marino-Albernas, J. R.; Auzanneau, F.-I.; Forooghian, F.; Venetta, T.; Nycz, C.; Mitchell, M. J.; Pinto, B. M. *Carbohydr. Res.* **2000**, *324*, 17–29.
54. Johnson, M. A.; Rotondo, A.; Pinto, B. M. *Biochemistry* **2002**, *41*, 2149–2157.
55. Johnson, M. A.; Jensen, M. T.; Svensson, B.; Pinto, B. M. *J. Am. Chem. Soc.* **2003**, *125*, 5663–5670.
56. Johnson, M. A.; Jaseja, M.; Zou, W.; Jennings, H. J.; Copié, V.; Pinto, B. M.; Pincus, S. H. *J. Biol. Chem.* **2003**, *278*, 24740–24752.
57. Johnson, M. A.; Höög, C.; Pinto, B. M. *Biochemistry* **2003**, *42*, 1842–1853.