NMR studies of carbohydrate–protein interactions in solution

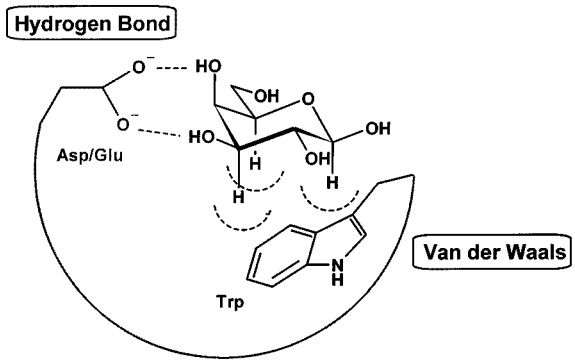
A. Poveda[a] and J. Jiménez-Barbero[b]

[a] *RMN-SIdI, Universidad Autónoma de Madrid, Cantoblanco-28047, Madrid, Spain*
[b] *Instituto Química Orgánica, CSIC, Juan de la Cierva 3, 28006, Madrid, Spain*

This review provides an overview of the current methods of NMR spectroscopy that may be used to obtain information about the conformational aspects, forces, and structural motifs that govern the interactions between proteins and carbohydrates in solution.

1 Introduction

Carbohydrates differ from the other classes of biomolecules in that their constituting moieties (monosaccharides) may be connected to one another by a great variety of linkage types. In addition, they can be highly branched, thus allowing oligosaccharides to provide an almost infinite array of structural variations. The decoding process of the existing information in oligosaccharide structures involves their recognition by other biomolecules. Thus, they are most often specifically recognized by proteins (so called lectins) and these interactions may mediate a particular biological response, such as host–parasite interactions, fertilization, autoimmune disorders and cellular differentiation.[1] Therefore, the study of how oligosaccharides are recognised by the binding sites of lectins, enzymes and antibodies is a topic of major interest. It is evident that knowledge of the three dimensional structure of these biomolecules (proteins and carbohydrates) could assist in the design of new carbohydrate-based therapeutic agents. Current technical facilities and biophysical techniques, mainly X-ray crystallography, have allowed access to detailed information on the three dimensional structure of protein–carbohydrate complexes.[2] These data, complemented mainly by those obtained through titration microcalorimetry, have allowed postulates on the major factors involved in these interactions to be made.[3] Hydrogen bonds and van der Waals forces, often including packing of a hydrophobic sugar face against aromatic amino acid side chains are the usual factors invoked (Scheme 1). The relative importance of each type of force depends on the particular protein and this issue remains a topic of discussion.



Scheme 1 Schematic view of van der Waals and hydrogen bond interactions between amino acid residues of a protein and a monosaccharide

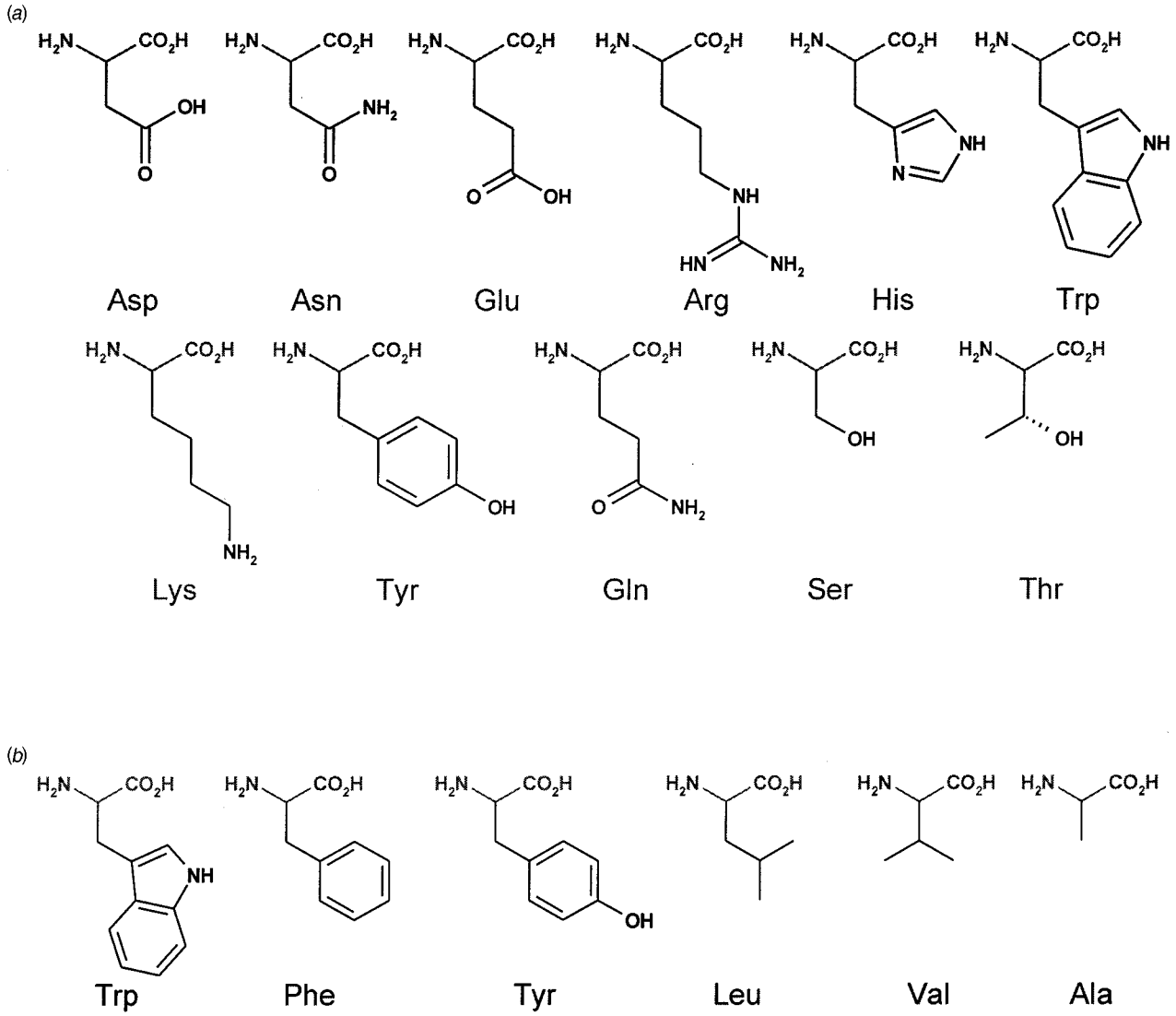
On the basis of crystal structures of a variety of complexes, the amino acids most commonly involved in hydrogen bonds with carbohydrates are known to be: Asp, Asn > Glu > Arg, His, Trp, Lys > Tyr, Gln > Ser, Thr. On the other hand, those most usually observed in van der Waals interactions are Trp, Phe, Tyr, Leu, Val and Ala; that is, those with aromatic or aliphatic side chains (Scheme 2). The ability to bind any one type of sugar has evolved independently in diverse lectin frameworks. In turn, families of lectins that share common structural features often contain members that recognise different groups of sugars. Although lectins bind monosaccharides rather weakly, these proteins employ common strategies for enhancing both the specificity and the affinity of their interactions for more complex oligosaccharide ligands. Thus, by using different lectin subsites and/or subunits it is possible to uncover these enhancements (Scheme 3), as several contacts between a given lectin and several carbohydrates or *vice versa* may take place. In addition, a dramatically increased affinity for oligosaccharides may result from clustering of simple binding

Jesús Jiménez-Barbero (Madrid, 1960) studied Chemistry at University Autonoma, Madrid. He received his PhD with M. Bernabé (1987). He has performed postdoctoral studies at the CERMAV-CNRS, Grenoble; University of Zuerich; National Institute for Medical Research, London; and Carnegie Mellon University, Pittsburgh. He is Senior Research Scientist of the National Research Council of Spain (CSIC). He has published about 100 papers on conformational studies of free and protein-bound carbohydrates by using NMR and molecular mechanics. His research interest are structural studies of biomolecules and molecular recognition processes by using NMR and molecular modeling protocols.

Jesús Jiménez-Barbero

Ana Poveda (Madrid, 1965) studied Chemistry at University Autonoma, Madrid. She received her PhD with J. Jiménez-Barbero on NMR studies of oligosaccharide dynamics. She is in charge of the NMR facility of the University. Her research interests are in applied and methodological NMR.

Ana Poveda



Scheme 2 Lateral chains of the different amino acid residues more frequently encountered in the binding sites of protein–carbohydrate complexes. (*a*) Amino acid residues involved in hydrogen bond interactions. (*b*) Amino acid residues involved in van der Waals interactions.

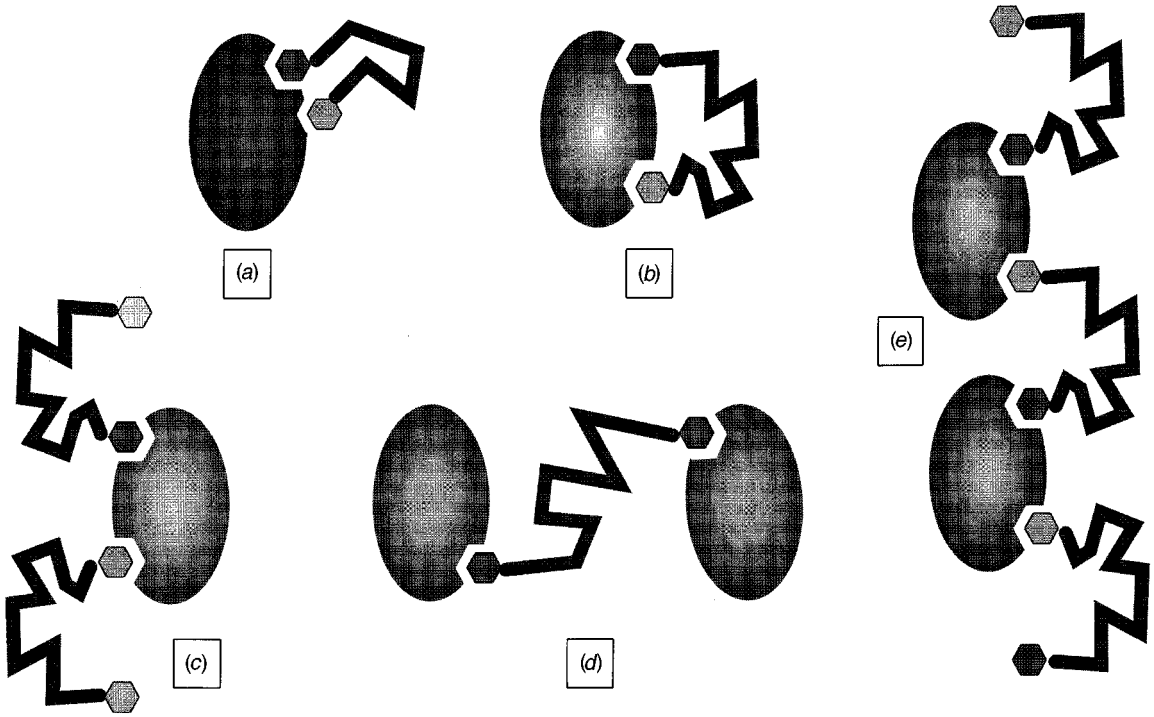
sites in oligomers of lectins. Apart from the direct interactions between atoms of both the protein and carbohydrate, other molecules are often involved in the formation of the complex. Water molecules, either located in the binding site or at the surface of the protein have also been shown to provide additional interactions which help to stabilize the complex and to achieve higher selectivity.[2] In addition, different ions have also been shown to be of crucial importance for the establishment of protein–carbohydrate molecular complexes. Calcium and manganese are encountered in the legume lectins, while the animal C-lectins are also calcium-dependent.[2,4]

Until very recently, the large size of carbohydrate-binding proteins (lectins, antibodies and enzymes) has precluded direct study using NMR. However, in the last few years, researchers have begun to apply NMR to the study of the molecular recognition processes involved at different levels of complexity. Although it has already been mentioned that X-ray crystallography has achieved many results in this field, oligosaccharides often present on glycoproteins or on protein–carbohydrate complexes are less amenable to crystallography due to their intrinsic flexibility. Moreover, carbohydrates also exhibit greater dynamic fluctuations than proteins and, therefore, NMR measurements may well offer new insights into the conformation of bound oligosaccharides and/or the corresponding dynamic timescales.

Thus, using NMR (Table 1), it is possible to deduce the specificity and affinity of binding, association constants and equilibrium thermodynamic parameters (titration experiments). It may also be possible to deduce which amino acid residues are involved in binding (titration, NOESY, CIDNP experiments), even the three dimensional structure of the bound carbohydrate and/or the protein (NOESY experiments). From the dynamic point of view, it could also be possible to say something about the ligand exchange timescale and the size of the complex (relaxation measurements).

Table 1 NMR protocols to study protein–carbohydrate complexes. The information provided from these studies is also shown

	P⊗	P C	P C
Titration	★	⋆	
CIDNP	★		
TR-NOE		★	
Relaxation	⋆	★	⋆
3D Structure	★	★	★



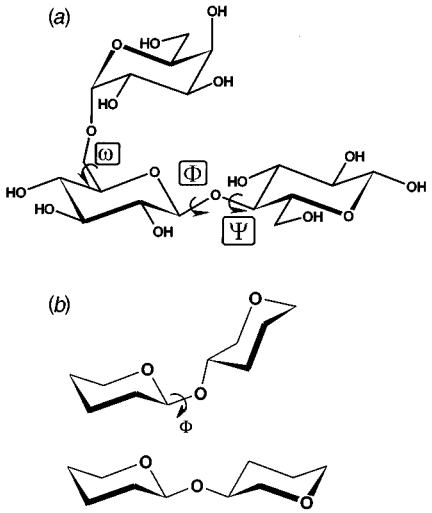
Scheme 3 Schematic view of the different ways in which carbohydrate binding proteins enhance the affinity towards their ligands. (*a*) Existence of different subsites. (*b*) Different domains of a given lectin may be involved in the binding of the same sugar. (*c*) Different domains of a given lectin may be involved in the binding of different sugars. (*d*) Different lectin molecules may bind different regions of a multivalent oligosaccharide. (*e*) Clustering of lectins and oligosaccharides.

Many questions concerning protein–carbohydrate interactions are therefore associated with conformational behavior. In many cases, not all of the required information can be obtained directly by experimental studies, so theoretical molecular modeling is usually required to supplement experimental data, both in solution and in the solid state. Different modeling protocols have recently been proposed to locate protein binding sites. Some of them can be used to calculate the interaction energies between ligand and receptor. Other methods are designed to explore systematically all the positions and orientations that the sugar may adopt within the binding site.[5]

Whenever the experimental 3D structure of a protein–carbohydrate complex is known, this information may be employed to derive the structures of other complexes which present structural homology. This so-called knowledge-based model building approach, has been used with success for several legume lectins.[6] From the point of view of the force fields used in conformational analysis of carbohydrates and in protein–carbohydrate interactions, there is no general force field at present although a variety of them have been shown to provide satisfactory agreement between experimental and modeled data.[7] Force fields specifically developed for sugars are HSEA and PFOS. General molecular mechanics programs adapted for carbohydrates are MM3, CHARMM, AMBER, GROMOS, and TRIPOS.[5–7] A general force field which has also been shown to be useful for deriving three dimensional structures of free carbohydrates is CVFF. In several cases, different modifications of a given force field are present in the literature.[7]

2 NMR investigations of protein–carbohydrate complexes

Different examples of the study of the structural events that mediate the molecular recognition processes between proteins and carbohydrates have recently been presented, using NMR spectroscopy, to examine interactions involving lectin-, antibody-, and enzyme-type receptors. Specific comparisons of the differential binding of natural and modified analogues have also been reported. In some cases, protein-induced conformational changes in the oligosaccharide ligand have been observed (Scheme 4). Nevertheless, in other cases the lectin selects one of the conformers present in solution or a structure close to the major one existing in solution. Finally, there are also cases in which the ligand retains at least part of the flexibility it has in the isolated state.



Scheme 4 (*a*) Schematic view of the glycosidic torsion angles which define the three dimensional shape of an oligosaccharide. (*b*) Schematic view of putative conformational changes around the φ glycosidic torsion angle of a disaccharide.

In principle, different types of information may be deduced for protein–carbohydrate complexes in solution by NMR. Although different types of clasification could be envisaged, we have decided to distinguish between those methods which permit us to deduce (*a*) structural information on the protein residues involved in the interaction; (*b*) information on the bound carbohydrate; and (*c*) information on the entire protein–carbohydrate complex.

Table 2 Summary of the studies described in the text, showing the method used by their authors

System	Method	Ref.
WGA/Sialic acid	Titration	8
Hevein/GlcNAc-containing oligosaccharides	Titration	9
UDA/GlcNAc-containing oligosaccharides	Titration	10
Ac-AMP2/GlcNAc-containing oligosaccharides	Titration	11
CBD *C. fimi* Cen C/cello oligosaccharides	Titration	12, 13
CBD *T. Reesei*/cello oligosaccharides	Titration	14
Tripeptides/heparin disaccharide	Titration	16
Macrophage mannose receptor/mannose	Titration	15
Hevein, pseudohevein, UDA, WGA/chitooligosaccharides	CIDNP	17
Ricin-B/galactose derivatives	TR-NOESY	20
IgA X24 antibody/fluorinated disaccharide	TR-NOESY	19
Galectin/galactose disaccharide	TR-NOESY	21
AAA/fucose disaccharide	TR-NOESY	22
Antibody/*Salmonella* trisaccharide	TR-NOESY	25
Ricin/C-lactose	TR-NOESY	27
Strep-9 antibody/GlcNAc-containing trisaccharide	TR-NOESY	18
E-selectin/SLeX	TR-NOESY	18
P-,L-selectin/SLeX	TR-NOESY	23
Dolichos biflorus lectin/Blood group A trisaccharide	TR-NOESY	6
Dolichos biflorus lectin/Forsmann pentasaccharide	TR-NOESY	24
Lentil lectin/sucrose	TR-NOESY	18
Acidic fibroblast growth factor/sucrose octasulfate	TR-NOESY	26
A. niger glucoamylase/N, S-heteroanalogues of maltose	TR-NOESY	18
Polyclonal IgG/GD1a ganglioside	Relaxation	33
CBD *T. Reesei*/cello oligosaccharides	Relaxation	14
Immunoglobulin G glycoforms (glycoprotein)	Relaxation	1
HCGα glycans (glycoprotein)	Relaxation	29, 30
Hevein/GlcNAc-containing oligosaccharides	complete 3D	9
CBD *C. fimi*/cello oligosaccharides	complete 3D	12, 13
Verotoxin VT-1/globotriaosyl ceramide	complete 3D	35
RNAse B glycoforms (glycoprotein)	complete 3D	1
Adhesion domain of human CD2 (glycoprotein)	complete 3D	18
Human granulocyte colony stimulating factor (glycoprotein)	complete 3D	31
Fucosylated peptide (glycoprotein)	complete 3D	32
HCG α subunit (glycoprotein)	complete 3D	29, 30

2.1 Methods which allow us to obtain information on the protein residues involved in the interaction

2.1.1 Titration experiments

NMR is the prevalent method used to study molecular conformation and dynamics in solution, in contrast to X-ray crystallography methods. NMR spectroscopy provides detailed information on the chemical surroundings of a given nucleus through chemical shifts and therefore, titration NMR experiments may provide an adequate means to analyse sugar-induced perturbations of proteins and *vice versa*. Usually, the specific binding of carbohydrates to a lectin is monitored by recording the ^{1}H NMR spectra of a series of samples with variable sugar concentration (six to eleven different concentrations). The concentration of the protein during the experiments is held constant. A first sample is used to obtain the ^{1}H NMR chemical shifts of the free-sugar lectin sample (δ_{free}). A second sample is prepared by dissolving a large amount of the corresponding sugar in a similar protein solution. The titration curve is built by adding small aliquots of the sugar-concentrated protein solution over the free-sugar protein sample. This protocol is a way to verify the existence of complexes between a given lectin and the corresponding carbohydrates, and, in addition, the alterations in the chemical shifts of the amino acid proton resonances may be used to determine the equilibrium association constant, K_a, considering the equilibrium eqn. (1)–(3) where δ_{bound} corresponds to

$$\text{Protein} + \text{Sugar} \longleftrightarrow \text{Protein sugar} \quad (1)$$

$$K_a = \frac{[\text{Protein} \cdot \text{Sugar}]}{[\text{Protein}] \times [\text{Sugar}]} \quad (2)$$

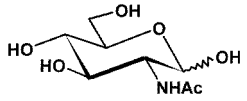
$$\delta_{obs} = \delta_{free} + (\delta_{bound} - \delta_{free}) \cdot \frac{[\text{Protein} \cdot \text{Sugar}]}{[\text{Protein}] + [\text{Protein} \cdot \text{Sugar}]} \quad (3)$$

the NMR chemical shifts of the sugar-bound form of the lectin. The values of K_a and δ_{bound} may be obtained by non-linear least-squares fitting of the observed NMR chemical shifts δ_{obs} of different selected protons of the lectin as a function of the total sugar concentration. δ_{free} may be introduced as an adjustable parameter to control the goodness of fit by comparing its value with that obtained experimentally.

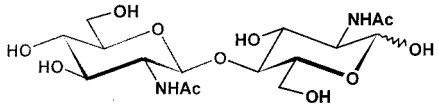
Moreover, provided that the protein resonances have been assigned, this methodology may be used to locate, at least qualitatively, the sugar binding site around certain protein residues. Thus, either qualitative or quantitative estimations of binding affinity and specificity can be derived by using 1D or 2D spectroscopy. Kinetic and thermodynamic parameters may also be inferred, at least qualitatively, by running these titration experiments at different temperatures and following a van't Hoff type of analysis, using a representation of $R \log K_a$ *vs.* 1/T.

2.1.1.1 GlcNAc-binding proteins The pioneering work in this field was performed by Kronis and Carver,[8] who analysed the interaction and thermodynamics of the binding of sialyl oligosaccharides by the lectin wheat germ agglutinin (WGA).

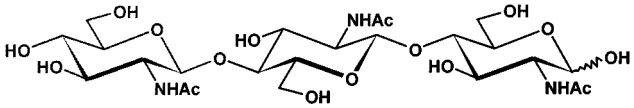
As a recent example of the application of this protocol, Asensio *et al.*[9] have recently reported the determination of the binding site of hevein by using NMR spectroscopy and different *N*-acetyl glucosamine-derived (GlcNAc) ligands. The GlcNAc, chitobiose, and chitotriose (Scheme 5) specific binding constants were also determined by 1D NMR spectroscopy. These constants increase by one order of magnitude when passing from the mono- to the di- and to the tri-saccharide. In addition, the thermodynamic parameters for chitotriose–hevein and chitobiose–hevein interactions were obtained from a van't Hoff analysis, indicating that the association process is enthalpy driven, while entropy opposes binding. This behaviour is usually observed for protein–carbohydrate interactions.[1–3] The deduced negative signs indicated that hydrogen bonding and/or van der Waals forces are the major interactions stabilizing the complex. The differences in binding constants were explained in terms of the three-dimensional structure of the complexes,



N-Acetylglucosamine



Chitobiose



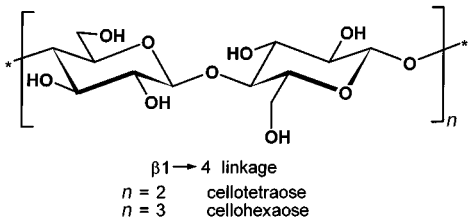
Chitotriose

Scheme 5 Structure of *N*-acetyl glucosamine (GlcNAc) containing sugars. From top to bottom: GlcNAc, chitobiose, chitotriose.

also obtained from NOESY NMR spectroscopy (see below in Section 2.3).

A similar study has also been performed to deduce the chitotriose-induced perturbations in *Urtica dioica* agglutinin (UDA), which contains two homologous hevein domains. The data confirmed the presence of two binding sites of non-identical affinities, since sugar induced perturbations occur in one domain of the lectin at sugar concentrations below equimolar. Residues in the second domain are shifted at higher trisaccharide concentrations.[10] The interaction between chitotriose and a related antifungal and antimicrobial peptide, Ac-AMP-2, has also been studied by ^{1}H NMR, showing that, as observed for hevein and UDA, three aromatic residues are involved in binding.[11]

2.1.1.2 Cellulose-binding proteins The binding specificity of the interaction of different glucans with the 152 amino acid *N*-terminal cellulose binding domain (CBD) of *Cellulomonas fimi* CenC has also been studied by ^{1}H NMR, showing that at least four β-(1 → 4) linked glucopyranosides (Scheme 6) are



Scheme 6 Structure of β-(1 → 4)-linked glucose oligomers, n = 2, cellotetraose, n = 3, cellohexaose

required to detect binding.[12] The cellulose binding domain spans five glucosyl units. Using an NMR model of the protein (see below in Section 2.3), it was deduced that the interaction takes place primarily through hydrogen bonding and van der Waals stacking.

Titration experiments of a related binding domain with cellohexaose allowed it to be established that Trp54 and Trp72 participate in cellulose binding. Using an NMR derived structure of this polypeptide it was deduced that both residues are adjacent in space and exposed to solvent, forming a ligand binding cleft, which is a feature common to the cellulose binding domains of the same family.[13]

The identification of the functionally relevant amino acids of the cellulose binding domain from *Trichoderma reesei* cellobiohydrolase I has been performed by recording the NMR spectra of synthetically modified peptides. Although in general, the structural effects of substitutions were minor, in some cases decreased binding could clearly be ascribed to conformational perturbations. At least one glutamine and two tyrosine residues were found to be essential for tight binding (see below in Section 2.3).[14]

2.1.1.3 Other examples One recent example of the application of titration NMR experiments has focused on the mechanism of calcium and sugar binding to a C-type carbohydrate recognition domain of the macrophage mannose receptor. The authors varied the nature of several key amino acid residues of the protein using site directed mutagenesis. Titration NMR experiments were performed for every mutant and from the corresponding magnitudes of the affinity constants, it was possible to deduce that a stacking interaction between mannose and Tyr729 contributes about 25% of the total free energy of binding.[15]

^{13}C NMR has also been used for titration experiments. For instance, the interaction of the heparin disaccharide with tripeptides has been studied by titration and ^{13}C NMR relaxation measurements. Relaxation rates (see below under 2.2.2) for the disaccharide are significantly higher in the presence of the peptide. The analysis of the data in terms of molecular diffusion constants indicated that the peptide is oriented proximal to the uronic acid ring.[16]

2.1.2 Chemically induced dynamic nuclear polarization (CIDNP)

In many cases sugar recognition by proteins involves the side chains of tryptophan, tyrosine and histidine moieties (Scheme 2), so the photo CIDNP method may profitably be applied to monitoring the effect of ligand binding on the receptor. These aromatic moieties are able to produce CIDNP signals after laser irradiation in the presence of a suitable radical pair-generating dye. Elicitation of such a response in lectins implies accessibility of the respective aromatic groups to the light-absorbing dye. Therefore, this protocol may be suitable for monitoring surface properties of a protein receptor and the effect of sugar binding provided that CIDNP-active amino acid residues are involved in the recognition site. Experimentally, the intensity and the shape of the CIDNP signals are therefore determined in the absence and in the presence of different carbohydrate ligands.

This method has been elegantly applied recently by Siebert *et al.* to study the complexation of GlcNAc-containing oligosaccharides to a series of plant lectins of increasing structural complexity.[17] In particular, the binding of chitosugars (Scheme 5) to hevein, pseudohevein, *Urtica dioica* agglutinin (UDA), wheat germ aglutinin and its B domain were investigated. When the sugar is bound, CIDNP signals of the aromatic moieties of Tyr, Trp, or His are altered with respect to those of the free protein: they may be broadened, appear with reduced intensity, or even disappear completely. Thus, their involvement in sugar binding may be deduced. The results obtained were in agreement with those previously reported.[9,10] In addition, a conformational change of an indole ring of a Trp residue was also detected for UDA.

2.2 Methods which allow us to obtain structural information on the bound carbohydrate

2.2.1 Transferred nuclear Overhauser enhancement (TR-NOE) studies

It is obvious that knowledge of the recognised conformation of a biologically active carbohydrate presents considerable implications for rational drug design. From the three dimensional point of view, TR-NOE may allow the assessment of the conformation of protein-bound oligosaccharides.[18] NOESY experiments provide information about which protons are close in space and, therefore, they may be used to deduce conformational information. The TR-NOESY is a regular NOESY experiment, but it is applied to a protein–ligand system in dynamic exchange in which the ligand is present in excess. For ligands which are not bound tightly and exchange with the free form at a reasonably fast rate, as usually observed for carbohydrates, TR-NOE provides an adequate means to determine their bound conformation (Scheme 7). In complexes involving large molecules, cross relaxation rates of the bound compound (σ^B) are opposite in sign to those of the free one (σ^F) and produce negative NOEs. Therefore, the existence of binding may be easily deduced by visual inspection, since NOEs for small molecules are positive (Scheme 8). The conditions for the applicability of this approach are well established, considering the well known equilibrium and the molar fractions of free and protein bound sugar eqn. (4)–(6)

$$\text{Protein} + \text{Sugar}_{(\text{excess})} \longleftrightarrow \text{Protein sugar} \quad (4)$$

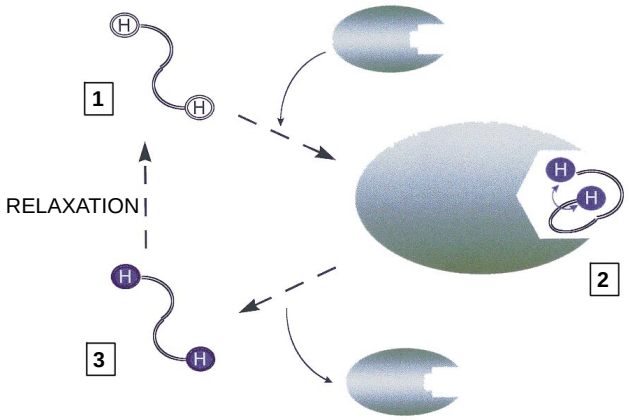
$$K_a = \frac{[\text{Protein} \cdot \text{Sugar}]}{[\text{Protein}] \times [\text{Sugar}]} \quad (2)$$

$$p_b\sigma^B > p_f\sigma^F \quad (5)$$

$$K_1 \gg \sigma^B \quad (6)$$

where p_b and p_f are the fractions of bound and free sugar ligand and σ^B and σ^F the cross relaxation rates for the bound and the free ligand, respectively. K_{-1} is the off-rate constant.

Under these conditions, it can be considered that eqn. (7) holds.

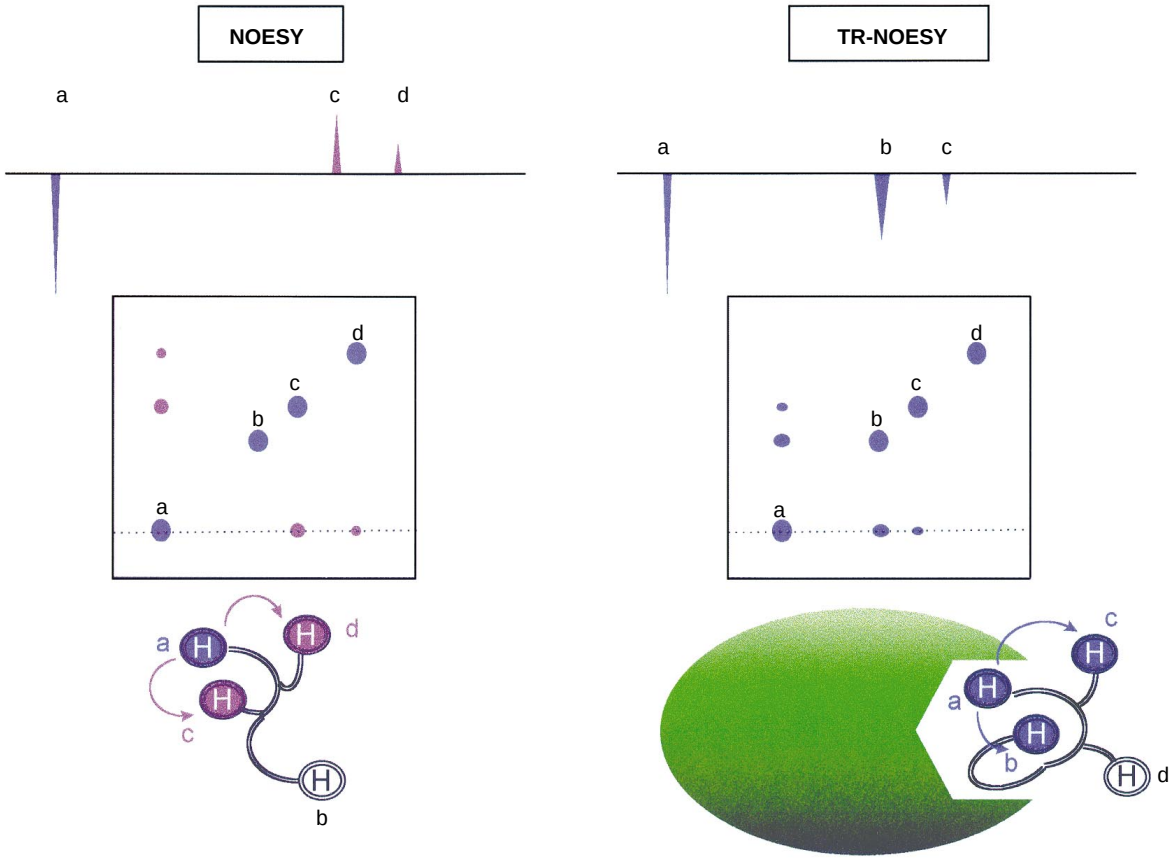


Scheme 7 Schematic representation of the facts which take place during TR-NOESY experiments. (1) Initial state: free sugar. (2) Formation of the complex: bound sugar. NOEs between protons close in space are developed. (3) Dissociation of the complex. The free sugar maintains the information acquired in the bound state for a given period of time, which depends on its relaxation times. After this time, the system reverts to the initial state.

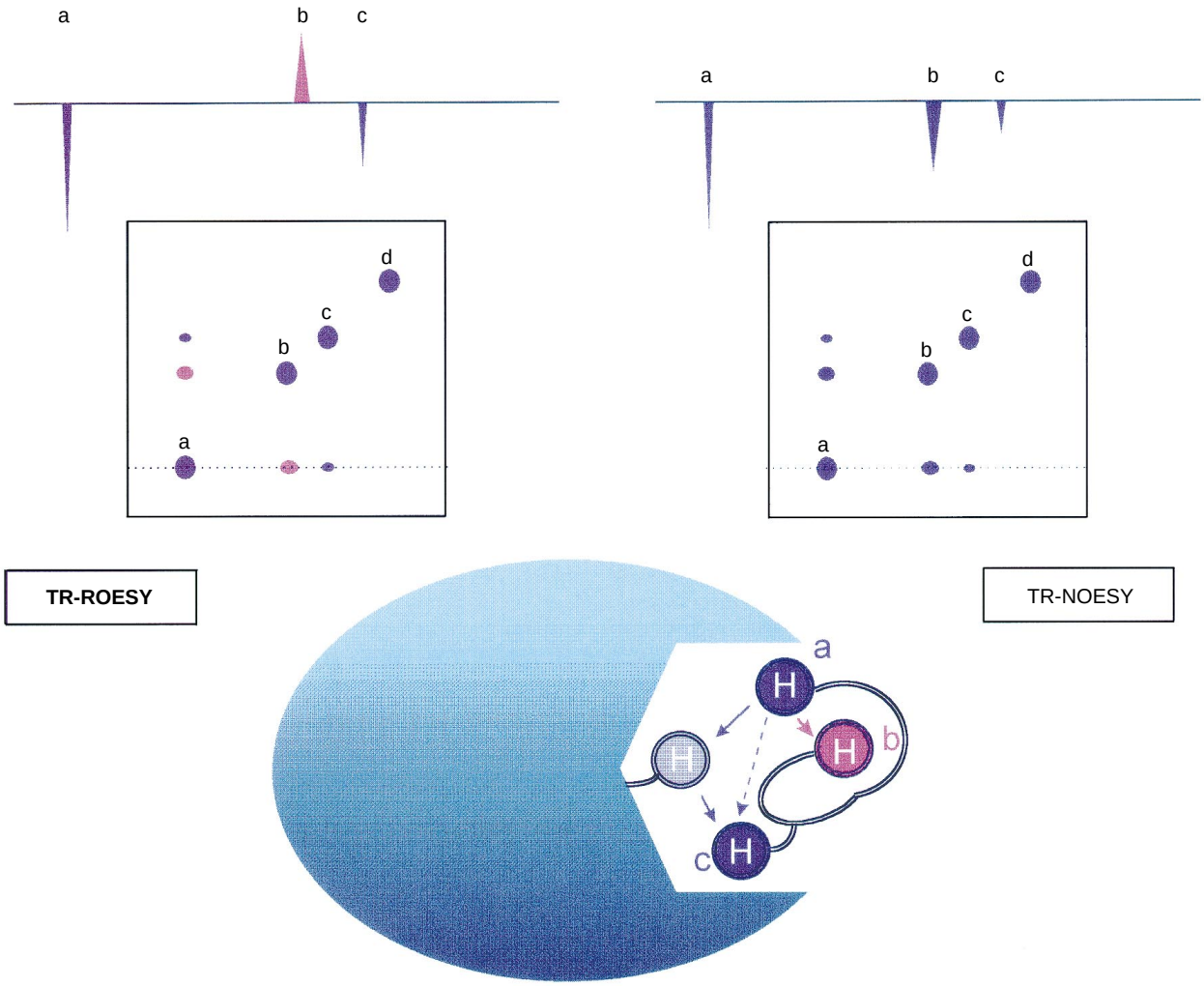
$$\sigma^{obs} = p_b\sigma^B + p_f\sigma^F \quad (7)$$

TR-NOESY experiments are usually performed at different mixing times and ligand–protein ratios and produce strong, negative NOEs on ligand binding. However, one of the major drawbacks of this experiment is the possible existence of spin diffusion effects, which are typical for large molecules. In this case, apart from direct enhancements between protons close in space, other spins may mediate the exchange of magnetization, thus producing negative cross peaks between protons far apart in the macromolecule. Thus, protein-mediated, indirect TR-NOE effects may lead to interpretation errors in the analysis of the ligand bound conformation. As a prime example, one of the first reported applications of TR-NOE experiments was to the derivation of oligosaccharide bound conformations, which concluded that a fluorinated Gal-β-(1 → 6)-Gal-β-OMe disaccharide underwent major conformational changes around the glycosidic linkages when bound to a specific antibody. The conclusion was based on the detection of an NOE cross peak between two protons located on two different pyranoid moieties. However, the reevaluation of the problem by the same authors, using TR-NOEs in the rotating frame (TR-ROESY) experiments[19] demonstrated that this cross peak was dominated by an indirect effect, mediated by a protein proton. In TR-ROESY, spin-diffusion (three spin) effects appear as positive cross peaks and therefore, the application of this experiment permits one to distinguish direct from indirect enhancements, and thus complements those measured under regular conditions, providing conformational information which is less contaminated by artifacts (Scheme 9). Different examples which have focused on the study of the structural events that mediate the molecular recognition processes between proteins and carbohydrates have recently been presented, through examples of lectin-, antibody- and enzyme-type receptors. Specific comparisons of the differential binding of natural and modified analogues have also been reported.

2.2.1.1 Recognition of the global minimum conformation Although there is not any general rule, in many cases, protein binding sites are well preorganized to recognise a conformation of the oligosaccharide which is located close to the its global minimum energy region. Several TR-NOE studies on protein–carbohydrate interactions have studied the complexes between Ricin-B chain and different oligosaccharides.[20] Ricin is a



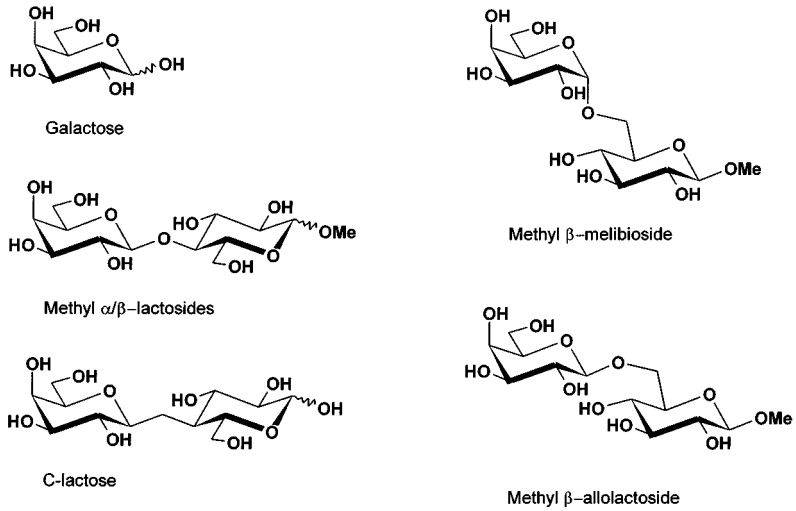
Scheme 8 Left. Schematic representation of a NOESY spectrum for a free sugar. Cross peaks and diagonal peaks have different signs. Right. Schematic representation of a TR-NOESY spectrum recorded for an exchanging sugar–protein system. Cross peaks and diagonal peaks have the same signs, as expected for a large molecule, thus indicating binding to the protein. The relative sizes of the peaks and the appearance of new ones may be used to detect conformational variations.



Scheme 9 Schematic representation of TR-ROESY (left) and TR-NOESY (right) spectra for an exchanging sugar–protein system. In TR-ROESY spin diffusion (three spin) cross peaks (*i.e.* a/c) and diagonal peaks have the same signs. On the other hand, direct cross peaks (a/b) show different sign to diagonal peaks (a). In TR-NOESY, all direct and spin diffusion-mediated cross peaks have the same sign as diagonal peaks.

dimeric (A and B chains) galactose-binding lectin seeds which has been shown to be 10–100 fold more toxic to some transformed cell lines than to normal cells and has therefore been considered as a potential antitumor agent.

The first NMR study of Ricin-B/disaccharide complexes used monodimensional (1D)-TR-NOE experiments to study the binding of Ricin-B by methyl β-lactoside. In this example, it was demonstrated, by using a selectively deuterated substrate that there were minor changes in the conformation of free methyl β-lactoside (Scheme 10) upon binding to Ricin-B. In a second example, the Ricin-B-bound conformation of melibiose [Gal-β(1 → 6)-Glc] was deduced and compared to its conformation in free solution demonstrating that only one of the two solution conformations of melibiose was recognized by the



Scheme 10 Structure of different galactose-containing disaccharides and analogues

lectin. Therefore, the protein causes a shift in the solution equilibrium towards the bound conformation during the recognition process. Docking studies indicated that the protein chain excluded binding of certain ligand conformations on the basis of unfavorable interactions between the protein surface and remote parts of the carbohydrate. However, since Ricin-B preferentially binds β-galactosides rather than α-galactosides and, since the orientation of the glucose residue in α-substituted galactosyl glucosides is very different from that existing in their β-analogues, the conclusions reached for the melibioside could not be extrapolated in a general way. Thus, in an attempt to generalise this structural problem, the conformational changes that occur when methyl α-lactoside was bound to the Ricin-B chain in aqueous solution were then studied.[20] The observed data indicated that the protein causes a slight conformational variation in the glycosidic torsion angles of methyl α-lactoside, although the recognized conformer was still within the lowest energy region. Molecular modeling using molecular dynamics, minimization, and docking of the disaccharide within the binding site of Ricin B strongly suggested that, apart from the expected contacts between the galactose moiety and different amino acid residues, there were also van der Waals contacts between the protein and the remote glucose moiety, as previously deduced from binding studies using modified lactoside derivatives. Thus, both van der Waals contacts and hydrogen bonding contribute to the stability of the complex.[20]

As with Ricin–lactose, there are other cases in which there are no major variations in the conformational behavior of the oligosaccharide upon protein binding. For instance, the TR-NOESY study of the binding of Galβ-(1 → 2)Galβ-(1 → R) to the galectin of chicken liver showed that the conformation of the disaccharide in the bound state is very close to its global energy minimum state in solution.[21]

2.2.1.2 Simultaneous recognition of different conformations There are examples in which the protein does not select a single conformer. The *Aurelia aurantia* agglutinin (AAA) lectin recognises, simultaneously, different conformations[22] of Fucα(1 → 6)GlcNAcβ-(1 → OMe). This disaccharide, which is fairly flexible when free in solution, appears to remain, to a certain extent, flexible around the glycosidic linkage within the lectin binding site. An analogous case has been reported for the complex between methyl β-allolactoside [Galβ(1 → 6)Glcβ-OMe] and Ricin B.[20] In this case, and contrary to the observations for lactose, different conformations around the φ, ψ, and ω glycosidic bonds of methyl β-allolactoside were recognized by the lectin. In fact, for this complex, only the TR-NOESY cross-peaks corresponding to the protons of the galactose residue were negative, as expected for a molecule in the slow motion regime. In contrast, the corresponding cross peaks for the glucose residue were *ca.* zero, as expected for a molecule whose motion is practically independent of the protein.

2.2.1.3 Protein-induced conformational selections Lectins and antibodies may select just one of the conformers present in the conformational equilibrium for the free state. The quest for the active conformation of the Lewis X oligosaccharide has stimulated different research groups. The sialyl Lewis X (SLeX, {αNeuNAc-(1 → 3)β-Gal(1 → 4)[αFuc-(1 → 3)]Glc} tetrasaccharide exists in solution as an equilibrium of several conformations, which are mainly characterized by the orientations of the *N*-acetylneuraminic acid residue. Perhaps the pre-eminent study on this topic has been reported by Peters and coworkers[18] using spin-locked filtered NOESY and Metropolis Monte Carlo calculations. The most relevant conclusion is that E-selectin complexes exclusively to a conformation of sialyl Lewis from the conformational equilibrium in aqueous solution in which the sialic acid shows an orientation, defined by (φ/ψ: 76/6), already reported to be present in free solution. On the other hand, the orientation of the fucose residue (φ,ψ: 38/26) differs from that preferred in aqueous solution. This work clarified previous discussions on the conformational changes of sLeX upon binding to E-selectin. Recently, the bound conformation of sLeX bound to E-selectin was also compared to those recognised by P- and L-selectin.[23] In all cases, it was demonstrated that the conformation of the branched trisaccharide remained close to the conformation of the free ligand. However, E- and P-selectins recognised a different conformation around the sialic acid glycosidic linkage than L-selectin.

The blood group A trisaccharide {αGalNAc-(1 → 3)[αFuc-(1 → 2)]β-Gal} exists in solution as an equilibrium between two families of low energy conformers. Comparison between experimental and simulated TR-NOESY volumes, lead to the conclusion that only one conformation of the trisaccharide was bound[6] by the GalNAc-specific lectin isolated from *Dolichos biflorus*. Spin diffusion NOEs were detected by means of TR-ROESY experiments. The proposed bound conformation was in agreement with one of the two deduced from previous modeling studies. As with other lectins, complementary forces emanate from hydrogen bonding and van der Waals forces including hydrophobic interactions.

A second report on the application of TR-NOE experiments to the molecular recognition of oligosaccharides by the seed lectin of *Dolichos biflorus* has been completed.[24] TR-NOESY and TR-ROESY experiments collected for a mixture of this lectin and the Forssman pentasaccharide GalNAcα-(1 → 3)GalNAcβ-(1 → 3)Galα-(1 → 4)Galβ-(1 → 4)Glc revealed close contacts between the non-reducing disaccharide moiety of the carbohydrate and the lectin binding site. In addition, and using an elegant protocol of recording experiments at different lectin:sugar ratios, the authors deduced two distinct classes of NOE cross peaks which reflected the size of the carbohydrate epitope and thus also of the binding pocket of the lectin. In order to detect contacts between the protein and the carbohydrate chain, T_2-filtered TR-NOESY spectra were performed which permitted the detection of NOEs between the terminal disaccharide fragment and protein protons, most likely belonging to Leu residues, in agreement with the previously reported molecular modeling study of the complex.

2.2.1.4 Protein-induced conformational variations Several cases of protein-induced major conformational changes have also been reported. Bundle and coworkers[25] have presented TR-NOESY evidences which show that a branched trisaccharide {α-Galp(1 → 2)[α-Abep(1 → 3)]-Manp-1 → OMe}, related to the antigenic determinant of a *Salmonella* polysaccharide, undergoes an antibody-induced conformational shift about one glycosidic linkage (Gal-Man) when bound in solution. Previous data have demonstrated that only this trisaccharide portion of the complete polysaccharide was bound by the antibody. Although the TR-NOESY distance constraints were compatible with two different bound conformations, one of them was shown to be consistent with the X-ray structure of the same molecular complex, but none with the free solution conformation of the oligosaccharide.

The Strep 9 antibody-bound conformation of a branched trisaccharide, namely, GlcNAc-β-(1 → 3)-α-Rha-(1 → 2)-α-Rha-OMe has been[18] investigated by TR-NOESY and TR-ROESY experiments and Metropolis Monte Carlo calculations. It was deduced that the monoclonal antibody Strep 9 selects only one defined conformation of the carbohydrate hapten. This bound conformation, which is a local energy minimum on the potential energy maps of the free ligand, undergoes a change in the orientation of one glycosidic linkage when compared to the global minimum conformation in the free solution state. It was also necessary to include repulsive constraints, derived from the absence of NOEs, to deduce the three dimensional structure of the trisaccharide in the binding site of the antibody.

The conformational features of one of the most important of the food industry-relevant sugars, sucrose, in the combining site of lentil lectin in solution have been recently characterized through TR-NOESY experiments and molecular modeling.[18] The experimental NMR data, which indicated that the bound

sucrose has a unique conformation for the glycosidic linkage, were in agreement with the results obtained for the complex using X-ray crystallography. It is important to mention that major differences with respect to the hydrogen bonding network of free sucrose were found, since none of the two inter-residue hydrogen bonds detected in crystalline sucrose were conserved in the complex with the lectin. Stacking interactions between a Phe residue and the hydrophobic face of the glucose residue as well as between the same Phe and H-4 and H-6 of the fructose moiety were deduced both experimentally (X-ray) and by modeling. A variety of protein–sugar hydrogen bonds were also detected. In addition, the NMR study provided insight into the residual conformational flexibility of sucrose in the lectin binding site. On the other hand, it has also been shown that free sucrose octasulfate appears to assume a conformation significantly different from any of the X-ray conformations determined for sucrose when bound to the acidic fibroblast growth factor. In this case, strong electrostatic interactions between guest and host may be the dominant factor in the deformation of sucrose octasulfate.[26]

2.2.1.5 Use of structurally modified carbohydrate analogues Obviously, not only natural ligands, but also structurally-modified oligosaccharides may be used as lectin ligands or as inhibitors of carbohydrate-processing enzymes, and thus, these analogues may be employed to deduce enzymatic mechanisms. Moreover, conformational differences between free and protein-bound natural carbohydrates and synthetic analogues may also be assessed by TR-NOE experiments.

Ricin B has also been used as a model to study the bound conformation of potential glycosidase inhibitors such as C-glycosides.[27] Although many reports have usually assumed that the conformation of free C-glycosides was the same as that of the corresponding *O*-analogues, it has recently been reported that, at least for *O*- and *C*-lactoses, this is not the case.[27]

Thus, 2D TR NOESY experiments were recorded to study the complexation of C-lactose by Ricin B. The conformational study of C-lactose in the free state showed that the *exo*-anomeric conformation around the C-glycosidic bond was adopted. However, the conformation around the aglyconic bond was rather different to that of the natural compound. For *O*-lactose, *ca.* 90% of the population was located around the so called minimum *syn*, ϕ/ψ: 54/18 and *ca.* 10% of population around minimum *anti*, ϕ/ψ: 36/180.[20] However, C-lactose was shown to exhibit much higher flexibility than its *O*-analogue and three conformational regions (*syn*, *anti* and *gauche–gauche*) were significantly populated in solution (Scheme 4). The comparison between the NOESY and ROESY spectra of C-lactose, recorded in the absence and in the presence of the lectin, indicated that conformers *syn* and *gauche–gauche* were not bound. Therefore, the experimental results indicated that Ricin B selects different conformers of C-lactose, (*anti*), and its *O*-analogue, (*syn*).[20] In order to estimate the relative binding affinities of the C- and *O*-glycosides, competitive TR-NOEs, with different *O*-lactose/C-lactose ratios, were also performed. It was demonstrated that both ligands compete for the same binding sites of the lectin and that the affinity constant of C-lactose is smaller than that of its *O*-analogue. Although merely speculative, and since the flexibility of C-lactose in the free state is much higher than that of *O*-lactose, the cause of the recognition of different conformations could be of entropic origin.

Other modified carbohydrates have been employed to study the structure of enzyme–inhibitor complexes. Mario Pinto and coworkers[18] have investigated *A. Niger* glucoamylase. This enzyme catalyzes the hydrolysis of maltose-type molecules with inversion of configuration. In an elegant manner, and using novel heteroanalogues of maltose containing sulfur in the non-reducing ring and nitrogen in the interglycosidic linkage, they have recently demonstrated that methyl 5′-thio-4-*N*-α-maltoside is a potent enzyme inhibitor and that it is bound by the enzyme in a conformation close to its global minimum. The characteristic NOEs observed for a second conformer which is also present in free solution, as a minor form, were not detected in the presence of glucoamylase G1. It is noteworthy that the crystal structure of a complex of a closely related glucoamylase with dihydroglucoacarbose indicated that the bound conformation of this ligand resembles the global minimum and that an existing local minimum conformer cannot be readily accommodated by the enzyme because of adverse van der Waals interactions.

2.2.1.6 Other developments TR-NOESY experiments have also been recently applied to the identification and characterization of biologically active molecules from a mixture.[28] As already stated above, the sign of transferred NOEs is opposite to that of NOEs of small molecules that do not bind to a protein and, thus, an unequivocal and fast identification of molecules with binding properties is possible.

2.2.2 NMR relaxation measurements

NMR relaxation properties depend on the spectral density functions [$J(\omega)$], which in turn are sensitive to molecular motion. Spectrometer frequency, molecular size, and internuclear distances are also important parameters. In fact, spectral densities are related to the motional correlation times (τ_c), which measure the rate of molecular tumbling. The functional form of the basic longitudinal (T_1) and transversal (T_2) relaxation times are given below in eqn. (8)–(11).

$$T_1^{-1} = (\Omega/20)\,[J(\omega_H - \omega_c) + 3J(\omega_c) + 6J(\omega_H + \omega_c)] \quad (8)$$

$$T_2^{-1} = (\Omega/20)\,[4J(0) + J(\omega_H - \omega_c) + 3J(\omega_c) + 6J(\omega_H) + 6J(\omega_H + \omega_c)] \quad (9)$$

$$J(n\omega) = \frac{2\tau_c}{1 + (n\omega\tau_c)^2} \quad (10)$$

$$\Omega = N\left(\frac{\gamma_C \gamma_H h}{2\pi r_{CH}^3}\right) \quad (11)$$

Exchange and dynamic processes affect heavily these relaxation parameters. Since T_2 values are directly related to linewidths, ($T_2^* = 1/\pi\upsilon_{1/2}$) the simple measurement or estimation of linewidths ($\upsilon_{1/2}$) may serve as a basis to deduce the occurrence of a dynamic process (such as binding or recognition) in the vicinity of a given nucleus (Fig. 1). Therefore, for instance, provided that the NMR signals of the protein have been assigned, ^{15}N and/or ^{13}C NMR relaxation measurements at one or several magnetic fields may be used to probe the change of mobility of specific amino acid residues of carbohydrate-recognizing proteins upon sugar binding. However, to the best of our knowledge, this aproach has not yet been used. Nevertheless, it is obvious that the relaxation properties of the oligosaccharide will also be affected upon protein binding, due to their dependence on molecular motion. Therefore, provided that the ^1H or ^{13}C NMR signals of the sugar have been assigned, NMR relaxation measurements may be used to probe the change of mobility of specific carbohydrate moieties upon binding.

As an example of this application, the interaction between cellohexaose and cellulose binding domains from *Trichoderma reesei* cellulases has been studied by T_2 relaxation analysis of the oligosaccharide resonances. In addition, and using an NMR derived structure of the polypeptide, a model for the molecular complex has been proposed in which three aligned aromatic residues (tyrosines) with a precise spatial arrangement stack onto every other glucose ring of the cellulose polymer.[14] The glycoproteins approach has recently been used by Dwek and coworkers[1] to probe the relative mobility of different glycoforms of immunoglobulin G, observing that mobility is dependent on the primary sequence of the glycan. Homans and coworkers have analysed line widths and ^1H and ^{13}C chemical shift changes of the glycan moieties on isotopically ^{13}C, ^{15}N

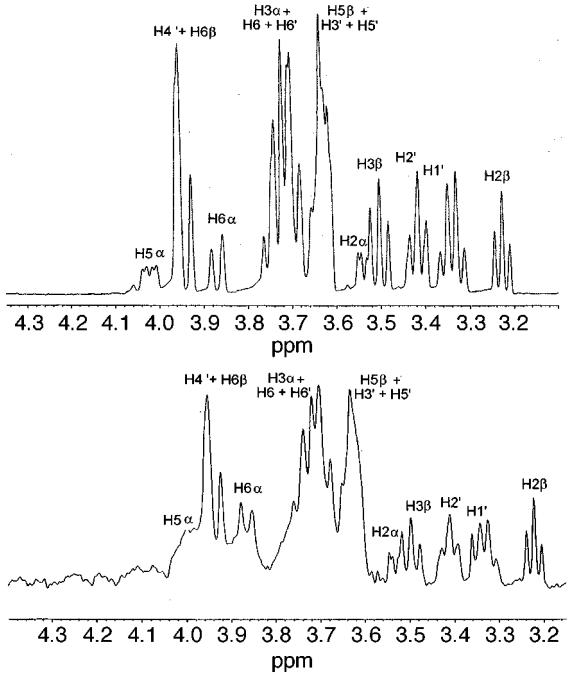


Fig. 1 Comparison between the ^{1}H NMR spectrum recorded for a free disaccharide (top) and that recorded for the same sugar in the presence of a specific lectin (bottom). In the case of the spectrum of the bottom part (sugar/protein, 20/1 ratio), the change in linewidths indicates that an interaction is taking place, since the relaxation properties of the disaccharide are heavily affected in the bound state.

enriched recombinant human chorionic gonadotropin α-subunit.[29] They have shown that the biologically relevant glycan at Asn 52 appears to extend into solution both in the isolated α subunit and in the complex with the β subunit. Similar conclusions have been deduced by the group at Utrecht, in this case, using a natural abundance sample.[30] Other studies of the motion of the glycan moieties in different glycoproteins have been performed, using either relaxation measurements or line width analysis. These include the adhesion domain of human CD2,[18] the human granulocyte-colony-stimulating factor,[31] a fucosylated peptide,[32] several glycoforms of RNAseB.[1] In all cases, the overall mobility of the glycan chain is reduced as compared to that of the corresponding small glycopeptides or oligosaccharides. The results obtained also suggest that the carbohydrate moiety reduces the local mobility around the glycosylation site. In addition, the carbohydrate provides more kinetic and thermodynamic thermal stability to the protein.

The influence of 9-*O*-acetylation of GD1a ganglioside on the recognition by a natural human antibody has been analysed by molecular dynamics simulations and NMR.[33] Although acetylation did not influence the overall conformation of the ganglioside, the NMR spectrum of the acetylated GD1a in the presence of the polyclonal IgG showed the disappearance of the 9-*O*-acetyl signal, indicating a variation in the value of T_2, thus its involvement in an exchange process, and consequently that the interaction with the human protein takes place on this site.

2.3 Complete determination of the 3D structure of protein–carbohydrate complexes

In a few favorable cases, dealing with protein receptors small enough to be amenable to direct analysis to NMR methods, ^{1}H NMR techniques have been used to deduce the three dimensional structure of protein–carbohydrate complexes. Very recently, and following the impressive development in NMR methodology and in molecular biology methods (allowing the obtention of ^{15}N- and ^{13}C-labelled molecules), detailed structural information on the 3D structure of carbohydrate–protein complexes and glycoproteins in solution has become available by using modern NMR pulse sequences especially designed for NMR active heteroatoms.[29]

Hevein[9] is a protein of 43 amino acids, whose structure has independently been solved by X-ray at 0.28 nm resolution, and by NMR methods. Interestingly, although the structure of hevein in water–dioxane and water solutions differs significantly from that observed in the crystal, it closely resembles the solid state structures of the domains of wheat germ agglutinin (WGA). Asensio *et al.*[9] have recently reported on the determination of the structure of the complex of hevein with chitobiose (Scheme 5, see above in section 2.1.1.1), by using NMR spectroscopy. Using NOESY spectroscopy and restrained molecular dynamics, they also presented a refined NMR structure of free hevein in water. The structure of the complex of hevein with methyl β-chitobiose has also been derived recently (Asensio *et al.*, in press).

Protein–carbohydrate NOEs measured for the hevein–chitobiose and hevein–methyl β-chitobiose complexes (Asensio *et al.*, submitted) allowed the deduction of the conformation of these complexes. Obviously, the presence of NOESY cross peaks between certain protons of the sugar and the protein permit one to infer that these atoms are close in space and therefore, to derive the three dimensional structure of the complex. No important changes in the protein NOEs were observed, indicating that carbohydrate-induced conformational changes in the protein are small. The *N*-acetyl methyl signal of the non-reducing GlcNAc moiety of β-chitobiose displayed NOE contacts with Tyr30 and Trp21 residues and appeared strongly shielded. From the inspection of the model, a hydrogen bond between Ser19 and the non-reducing *N*-acetyl carbonyl group was suggested as well as one between Tyr30 and HO-3 of the same sugar residue. The previously mentioned *N*-acetyl methyl group of the non-reducing GlcNAc displayed non polar contacts to the aromatic Tyr30 and Trp21 residues. Moreover, the higher affinities deduced for the β-linked oligosaccharides with respect to GlcNAc and GlcNAcα-(1 → 6)-Man could be explained by favorable stacking of the second β-linked GlcNAc moiety and Trp21. The final 3D structures derived by NMR were compared to those of WGA, Ac-AMP II (which is also a GlcNAc-binding protein) recently solved by NMR, Martins *et al.*[34] and to the crystal structure of hevein. The corresponding average rmsd are 0.060 nm (B domain of WGA, residues 16–32), 0.100 nm, (Ac-AMP2, residues 12–32), and 0.269 nm (crystal of hevein, residues 16–41). The structure of the 152 amino acid *N*-terminal cellulose binding domain of *Cellulomonas fimi* CenC has also been derived by multidimensional heteronuclear NMR, in the presence of saturating concentrations of cellotetraose (Scheme 6).[12] The polypeptide is composed of ten beta strands, folded into two antiparallel beta sheets with the topology of a jellyroll beta sandwich. These strands form the face of the protein previously determined by titration experiments to be responsible for cellulose binding (see above in Section 2.1). There is a binding cleft, which contains a central strip of hydrophobic residues that is flanked on both sides by polar hydrogen bonding groups. The existence of this cleft provides a structural explanation for the selectivity of this binding domain.

Recently, the solution structure of the carbohydrate-binding B-subunit homopentamer of verotoxin VT-1 from *E. coli* complexed to globotriaosylceramide [αGal-(1 → 3)βGal-(1 → 4)βGlc → R] has been deduced by Homans and coworkers,[35] using a ^{13}C/^{15}N doubly labelled protein sample. Unlike the crystal structure, in solution, there is no evidence of anomalous association between two of the subunits, which may be an artifact of crystallization. TR-NOEs obtained for the complex compare satisfactorily with those predicted from a previous molecular modeling study of the complex.

Although they are not protein–carbohydrate complexes, it has to be noted that recently, several glycoprotein structures in solution along with their structure-related properties have also been derived by NMR. These have been mentioned in Section

2.2.2. and are the adhesion domain of human CD2,[18] the human granulocyte-colony-stimulating factor,[31] a fucosylated peptide,[32] several glycoforms of RNAseB,[1] and human chorionic gonadotropin.[29,30]

3 Perspectives

With no doubt, present and future developments, including expression systems for glycoproteins,[29] will allow us to produce these biomolecules in the required amounts for detailed analysis of NMR data. Current NMR methodologies,[29,35] which permit one to deduce dynamic parameters through relaxation measurements will also find their application in the derivation of differential flexibility of the protein binding site before and after complex formation. A similar methodology may be applied to the ligand molecule, provided it is also ^{13}C-labelled. Moreover, the use of new methods to detect long lived protein-bound water molecules by NMR, in combination with other biophysical techniques, will surely allow us to dissect the relative contribution of van der Waals, hydrogen bond, water-mediated and entropy contributions to the stabilization of the carbohydrate–protein supramolecule.

4 Acknowledgments

We thank Drs J. L. Asensio, J. F. Espinosa, M. Martín-Pastor, J. Cañada, M. Bernabé and M. Martín-Lomas for helpful discussions and their contributions. We also thank SIdI-UAM for the facilities provided and DGICYT (PB96-0833) and the Mizutani Glycoscience Foundation for financial support.

5 References

1 R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683.
2 W. I. Weis and K. Drickamer, *Ann. Rev. Biochem.*, 1996, **65**, 441.
3 E. Toone, *Curr. Opin. Struct. Biol.*, 1994, **4**, 719, and references therein.
4 J. M. Rini, *Curr. Opin. Struct. Biol.*, 1995, **5**, 617, and references therein.
5 A. Imberty, K. D. Hardman, J. P. Carver and S. Perez, *Glycobiology*, 1991, **1**, 456.
6 F. Casset, T. Peters, M. Etzler, E. Korchagina, N. Nifant'ev, S. Perez and A. Imberty, *Eur. J. Biochem.*, 1996, **239**, 710.
7 R. J. Woods, *Curr. Opin. Struct. Biol.*, 1995, **5**, 591, and references therein.
8 K. A. Kronis and J. P. Carver, *Biochemistry*, 1985, **24**, 826.
9 J. L. Asensio, F. J. Cañada, M. Bruix, A. Rodriguez-Romero and J. Jimenez-Barbero, *Eur. J. Biochem.*, 1995, **230**, 621, and references therein.
10 K. Hom, M. Gochin, W. J. Peumans and N. Shine, *FEBS Lett.*, 1995, **361**, 157.
11 P. Verheyden, J. Pletinckx, D. Maes, H. A. M. Pepermans, L. Wyns, R. Willem and J. Martins, *FEBS Lett.*, 1995, **370**, 245.
12 P. E. Johnson, M. D. Joshi, P. Tomme, D. G. Kilburn and L. P. McIntosh, *Biochemistry*, 1996 **35**, 14381, and references therein.
13 G. Y. Xu, E. Ong, N. R. Gilkes, D. G. Kilburn, D. R. Muhandiram, M. Harris-Brandts, J. P. Carver, L. E. Kay and T. S. Harvey, *Biochemistry*, 1995, **34**, 6993.
14 M. L. Mattinen, M. Kontelli, J. Kerovuo, M. Linder, A. Annila, G. Lindeberg, T. Reinikainen and T. Drakenberg, *Protein Sci.*, 1997, **6**, 294, and references therein.
15 N. P. Mullin, P. G. Hitchen and M. E. Taylor, *J. Biol. Chem.*, 1997, **272**, 5668.
16 D. Mikhailov, K. H. Mayo, A. Pervin and R. J. Lindhardt, *Biochem J*, 1996, **315**, 447.
17 H. C. Siebert, C. W. Von der Lieth, R. Kaptein, J. J. Beintema, K. Dijkstra, N. van Nuland, U. M. Soedjanaamadja, A. Rice, J. F. G. Vliegenthart, C. S. Wright and H. J. Gabius, *Proteins*, 1997 **28**, 268.
18 T. Peters and B. M. Pinto, *Curr. Opin. Struct. Biol.*, 1996, **6**, 710, and references therein.
19 S. R. Arepalli, C. P. J. Glaudemans, D. G. Davis, P. Kovac and A. Bax, *J. Magn. Reson. B*, 1995, **106**, 195, and references therein.
20 J. L. Asensio, F. J. Cañada and J. Jimenez-Barbero, *Eur. J. Biochem.*, 1995, **233**, 618, and references therein.
21 H. C. Siebert, M. Guilleron, H. Kaltner, C. W. Von der Lieth, T. Kozar, N. Bovin, E. Y. Korchagina, J. F. G. Vliegenthart and H. J. Gabius, *Biochem. Biophys. Res. Comm.*, 1996, **219**, 205.
22 T. Weimar and T. Peters, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 88.
23 L. Poppe, G. S. Brown, J. S. Philo, P. V. Nikrad and B. H. Shah, *J. Am. Chem. Soc.*, 1997, **119**, 1727, and references therein.
24 F. Casset, A. Imberty, S. Perez, M. Etzler, H. Paulsen and T. Peters, *Eur. J. Biochem.*, 1997, **244**, 242.
25 D. R. Bundle, H. Baumann, H. Brisson, S. Gagne, A. Zdanov and M. Cygler, *Biochemistry*, 1994, **33**, 5183.
26 J. Shen and L. Lerner, *Carbohydr Res.*, 1995, **273**, 115.
27 J. F. Espinosa, J. Cañada, J. L. Asensio, M. Martín-Pastor, H. J. Dietrich, R. R. Schmidt, M. Martín-Lomas and J. Jiménez-Barbero, *J. Am. Chem. Soc.*, 1996, **118**, 10682, and references therein.
28 B. Meyer, T. Weimar and T. Peters, *Eur. J. Biochem.*, 1997, **246**, 705.
29 C. T. Weller, J. Lustbader, K. Seshadri, J. M. Brown, C. A. Chadwick, C. E. Kolthoff, S. Ramnarain, S. Pollak, R. Canfield and S. W. Homans, *Biochemistry*, 1996, 35, 8815.
30 T. de Beer, C. W. E. M. van Zuylen, B. R. Leeflang, K. Hard, R. Boelens, R. Kaptein, J. P. Kamerling and J. F. G. Vliegenthart, *Eur. J. Biochem.*, 1996, **241**, 229.
31 V. Gervais, A. Zerial and H. Oschkinat, *Eur. J. Biochem.*, 1997, **247**, 286.
32 G. Mer, H. Hietter and J. F. Lefevre, *Nature Struct. Biol.*, 1996, **3**, 45.
33 H. C. Siebert, C. W. Von der Lieth, X. Dong, G. Reuter, R. Schauer, H. J. Gabius and J. F. G. Vliegenthart, *Glycobiology*, 1996, **6**, 561.
34 J. Martins, D. Maes, R. Loris, H. A. M. Pepermans, L. Wyns, R. Willem and P. Verheyden, *J. Mol. Biol.*, 1996, **258**, 322.
35 J. M. Richardson, P. D. Evans, S. W. Homans and A. Donohue-Rolfe, *Nature Struct. Biol.*, 1997, **4**, 190.

Received, 14th August 1997
Accepted, 7th November 1997