ELSEVIER

Applications of nuclear magnetic resonance spectroscopy and molecular modeling to the study of protein-carbohydrate interactions

Ana Poveda,* J.L. Asensio,† J.F. Espinosa,† M. Martin-Pastor,† J. Cañada† and Jesús Jiménez-Barbero†

**Servicio Interdepartamental de Investigación, Facultad de Ciencias, Universidad Autónoma de Madrid, Madrid, Spain*

†Grupo de Carbohidratos, Instituto de Química Orgánica, CSIC, Madrid Spain

This work provides an overview of the applications of NMR to the study of protein-carbohydrate interactions. The use of TR-NOE experiments in this context is given. In particular, the study of Ricin/lactose and Hevein/chitobiose complexes is detailed. © *1997 by Elsevier Science Inc.*

Keywords: NMR, molecular modeling, protein-carbohydrate interactions, Ricin, Hevein

INTRODUCTION

Studies have demonstrated that oligosaccharides are involved in a number of recognition events such as cell adhesion, metastasis, fertilization, and embryonic development, among others.[1-3] Therefore, how these molecules are recognized by the binding sites of lectins, antibodies, and enzymes is currently a topic of major interest.[4,5]

From this point of view, it is evident that knowledge of the three-dimensional structure of lectins, antibodies, and enzymes has the potential to assist in the design of new carbohydrate-based therapeutic agents. There have been several encouraging examples in which exploitation of biological structures has led to better inhibitors of enzymes and to a deeper understanding of the molecular basis for drug-target interactions.

Color Plates for this article are on page 53.

Address reprint requests to: Jesús Jiménez-Barbero, Grupo de Carbohidratos, Instituto de Química Orgánica, CSIC, Juan de la Cierva 3, 28006 Madrid, Spain.

Paper submitted to Electronic Conference of the Molecular Graphics and Modelling Society, October 1996

A variety of biophysical,[6-9] biochemical,[10,11] chemical,[12] and spectroscopic[13] techniques have been used to probe the combining site of different lectin-oligosaccharide complexes. Nevertheless, detailed information on the three-dimensional structure of protein-carbohydrate complexes has been obtained from X-ray crystallography data,[14-16] since the usually high molecular weight of lectins has prevented their direct studies by means of nuclear magnetic resonance (NMR) spectroscopy.

However, transferred nuclear Overhauser enhancement (NOE) experiments have proven to be a powerful tool in the investigation of lectin- and antibody-bound conformations of oligosaccharides. For ligands that are not bound tightly and exchange with the free form at reasonably fast rates, the transferred nuclear Overhauser enhancement (TR-NOE) experiment, first proposed by Bothner-By and Gassend,[17] and later developed by Albrand and co-workers[18] and Clore and Gronenborn,[19,20] provides an adequate means to determine the conformation of the bound ligand. On this basis, TR-NOE experiments have become a popular tool in the conformational analysis of bound carbohydrates and analogs to lectins and antibodies.

Nevertheless, it seems necessary to stress here that the TR-NOE technique permits only the derivation of the bound conformation of the oligosaccharide moiety and, in principle, no information on the receptor is available. Therefore, the three-dimensional structure of the protein still needs to be elucidated in an independent or parallel fashion in order to build the three-dimensional structure of the molecular complex. This elucidation may come from X-ray,[21,22] NMR (if the protein is small enough), or modeling analysis. In a second step, once the three-dimensional structures of both entities are known, an efficient modeling procedure is again necessary to dock the oligosaccharide within the binding site of the protein.

Journal of Molecular Graphics and Modelling 15, 9-17, 1997

655 Avenue of the Americas, New York, NY 10010

1093-3263/97/$17.00
PII S1093-3263(97)00012-X

We now present different examples of the determination of the three-dimensional structure of different protein-carbohydrate complexes in solution by means of NOE spectroscopy. In a few examples, direct NOEs between the oligosaccharide and the protein have been observed that permit the unequivocal location of the binding site. However, in most of the cases, the locations of the binding sites have come from X-ray and/or molecular modeling.[23]

Examples of lectin-, antibody-, and enzyme-type receptors are described in the following sections.

BACKGROUND, MOLECULAR MODELING METHODS, AND TRANSFERRED NUCLEAR OVERHAUSER ENHANCEMENT THEORY

As mentioned in the preceding section, X-ray crystallography has allowed workers to deduce which are the major forces that take part in protein-carbohydrate interactions. These are usually hydrogen bonds and van der Waals forces.[15] The relative importance of both types of forces depends on the protein in particular, and this issue remains a topic of discussion. In principle, from the study of the crystal structure of a variety of complexes, the amino acids most commonly involved in hydrogen bonds are Asp, Asn > Glu > Arg, His, Trp, Lys > Tyr, Gln > Ser, Thr. On the other hand, those most usually observed in van der Waals interactions with carbohydrates are Trp, Phe, Tyr, Leu, Val, and Ala; that is, aromatic or aliphatic moieties.[24]

Apart of 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[25,26] or at the surface of the protein,[16] have been shown to provide additional interactions that help to stabilize the complex. In addition, different ions have also been shown to be of crucial importance for the establishment of a protein-carbohydrate molecular complex. Calcium and manganese are encountered in the legume lectins while the animal C-lectins are also calcium dependent.[27–29]

Molecular Modeling

Many questions concerning protein-carbohydrate interactions are associated with conformational behavior. Since, in many cases, this cannot be solved by direct experimental studies, theoretical molecular modeling is required to supplement this lack of experimental data, both in solution and in the solid state. Different modeling protocols have been proposed to locate protein-binding sites. The GRID program[30] permits the calculation of interaction energies between the ligand and the receptor. Other methods are designed to explore in a systematic way all the positions and orientations that the sugar may adopt within the binding site.[31]

Whenever the experimental three-dimensional (3D) structure of a protein-carbohydrate complex is known, this information may be employed to derive the structures of other complexes that present structural homology to that known. This approach, the so-called knowledge-based model-building method, has been used with success for several legume lectins.[32] From the point of view of the force fields used in conformational analysis of carbohydrates and in protein-carbohydrate interactions, at present there is not a general force field, although a variety of them have been shown to provide a satisfactory agreement between experimental and modeled data. Force fields, specifically developed for sugars are HSEA[33] and PFOS.[34] General molecular mechanics programs adapted for carbohydrates are MM3,[35,36] CHARMM,[37] AMBER,[38–41] GROMOS,[42] and TRIPOS.[31,43] A general force field that has been shown to be useful for derived three-dimensional structures of carbohydrates is CVFF.[44–46]

Nuclear magnetic resonance and transferred nuclear Overhauser enhancement experiments

Apart from X-ray crystallography, other methods have been applied to study the detailed structure of protein-carbohydrate complexes. The measurement of thermodynamic parameters is essential to understand protein-sugar interactions. Thus, microcalorimetry[9] has become an important tool for the derivation of interaction energies, of the enthalpic and entropic contributions, and of the role of the solvent during the recognition process. One of the most common methods of deriving the three-dimensional structure of biomolecules in solution is NMR spectroscopy, usually assisted by molecular modeling methods. NMR spectroscopy has been used to derive the solution conformation of free oligosaccharides[47] and to study the three-dimensional structure of glycan chains covalently bound to glycoproteins.[48–53] In addition, titration NMR experiments have also been shown to provide an adequate means to observe sugar-induced perturbations of different proteins.[54–56] Moreover, kinetic and thermodynamic parameters may be inferred from these types of experiments.[57,58]

From the three-dimensional structure point of view, for oligosaccharide ligands, which usually are not bound tightly to lectins and exchange with the free form at reasonably fast rates, the transferred nuclear Overhauser enhancement (TR-NOE) experiment provides an adequate means to determine the conformation of the bound ligand.

TR-NOE measurements have been used to detect the protein-bound conformations of different oligosaccharides epitopes.[59–63] The sugar and the protein are usually in fast exchange on the chemical shift scale, so that only a single set of averaged ligand resonances is observed, whose positions are approximately the same as those of the free sugar under the experimental conditions used. In the presence of excess ligand, NOEs between bound oligosaccharide protons appear as negative TR-NOEs on the averaged ligand resonances and give rise to cross-peaks with the same sign as the diagonal in pure phase absorption NOE spectroscopy (NOESY) spectra.

Nevertheless, it must be mentioned that spin diffusion[64] in the bound state may be effective since lectins, antibodies, and enzymes are fairly large molecules. Theoretical calculations[65–67] have shown that spin diffusion becomes important for high ligand/protein ratios. Thus, the use of fairly short experimental mixing times or low ligand concentrations has been suggested to circumvent spin diffusion. The theoretical treatment of spin diffusion on the basis of a complete relaxation matrix approach is not always possible, since the three-dimensional structure of the protein needs to be known and the off rate of the molecular complex is also necessary for a quantitative analysis. Unfortunately, the experimental access to both entities is usually rather difficult.

However, it is well known that spin diffusion is almost abolished or strongly disfavored when acquiring NOEs in the rotating frame, under spin-locking conditions.[68,69] In this case,

three-spin effects show signs opposite to those of effects arising from direct cross-relaxation. Nevertheless, since NOEs in the rotating frame have the same sign with independence of the correlation time of the molecule, it is not evident at first sight whether we are observing transfer NOE peaks for an exchanging protein:ligand system. Fortunately, the cross-relaxation rates for small and for large molecules are different enough to differentiate their behavior at moderately short mixing times.[69,70]

All of these facts indicate that transferred NOEs in the rotating frame are indeed an important complement to NOEs measured in the longitudinal frame, allowing one to differentiate spin-diffusion pathways and to obtain distances in the bound state that are less contaminated by these indirect effects. Disadvantages of TR-ROE are its somewhat lower sensitivity and the fact that cross-relaxation in the free ligand cannot be neglected,[71,72] and must be measured in a separate experiment, before data can be addressed in a quantitative manner. This aspect has led different authors to perform both TR-NOE and rotating frame TR-NOE (TR-ROE) experiments in order to facilitate the derivation of conclusions regarding the three-dimensional structural of protein-bound oligosaccharides. In fact, the application of TR-ROE experiments to an antibody-carbohydrate complex[71] led to a revision of the conclusions reached by the same authors[59] regarding the bound conformation of a flexible (1 → 6) linked disaccharide, since the TR-NOE effect in which the determination of the structure had been based came from an antibody-mediated spin-diffusion effect.

PROTEIN-CARBOHYDRATE INTERACTIONS INVESTIGATED BY NUCLEAR MAGNETIC RESONANCE

Nuclear magnetic resonance is the most prevalent method of studying the solution conformation of free oligosaccharides. As an extension of the application of NMR to the study of the three-dimensional structure of carbohydrates, several studies have addressed the critical issue of determining the protein-bound conformation of oligosaccharides in solution.

Several TR-NOE studies on protein-carbohydrate interactions have studied the complexes between Ricin-B chain and different galactose-containing oligosaccharides. Ricin, a dimeric (A and B chain) galactose-binding lectin isolated from *Ricinus communis* seeds,[73] has been shown to be 10 to 100-fold more toxic to some transformed cell lines than to normal cells and has therefore been considered as a potential antitumoral agent. The B chain is the lectin part, which interacts with galactoses on the cell surface, triggering endocytic uptake of the protein into the cytoplasm.

The involvement of ligand residues beyond the terminal galactose in binding to ricin has been a matter of controversy for some time. Although lactose is bound more effectively to ricin than is galactose, no specific interactions of the glucose residue with the lectin were detected in the X-ray structure of the ricin-lactose complex.[22] Similarly, a crystal structure of a biantennary oligosaccharide with ricin showed only the same galactose interactions with the polypeptide chain,[22] notwithstanding that ricin binds 1000-fold better to terminal galactose-containing oligosaccharides than to the monosaccharide galactose itself.[74]

Previous studies from our group, using monodeoxy, O-methyl, halodeoxy, and other modified lactose derivatives,[75–77] showed that the galactose HO-4, HO-3, and HO-6 were key polar groups in the interaction with ricin B chain (ricin-B), while the glucose moiety, having the $^4C_1$ conformation, was also important for recognition and binding. In particular, a nonpolar interaction involving position 3 was demonstrated to be operative.[77]

The first NMR study of ricin-B-disaccharide complexes used monodimensional transferred NOE (1D-TR-NOE) experiments[60,61] to study the binding of ricin-B by methyl β-lactoside. In this example it was demonstrated, by using a selectively deuterated substrate,[60] that there were minor changes in the conformation of free methyl β-lactoside on binding to ricin-B. In a second example, the same research group deduced the ricin-B-bound conformation of melibiose [Gal-β(1 → 6)-Glc] and compared it to its conformation in free solution[61] (Figure 1).

TR-NOESY experiments and molecular modeling (docking) studies, using the published X-ray structure of ricin,[22] allowed the demonstration that only one of the solution conformations of melibiose is recognized by the lectin. Since this disaccharide exists predominantly as a conformational equilibrium between two major conformers in free solution, the experimental TR-NOESY data supported the conclusion that the protein selects only one set of glycosidic torsion angles, and causes a shift in the solution equilibrium toward the bound conformation during the recognition process. The 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 β-galactosides, and since the orientation of the glucose residue in α-substituted galactosyl glucosides differs from that existing

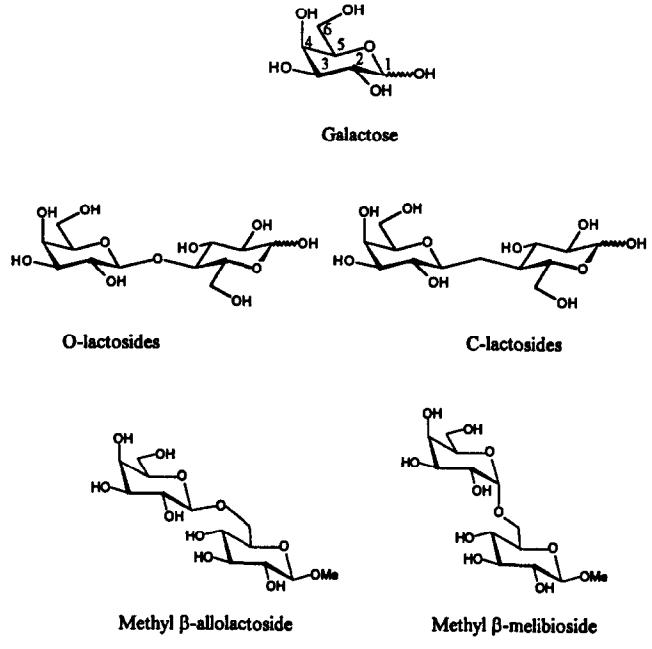

Figure 1. Schematic view of different compounds related to lactose studied by NMR spectroscopy to deduce their protein-bound conformations.

in their β-analogs, the conclusions reached for the melibioside could not be extrapolated in a general way.

Thus, in an attempt to generalize this structural problem, the conformational changes that occur when methyl α-lactoside was bound to ricin-B in aqueous solution were then studied.[78] TR-NOESY and TR-ROESY experiments were performed using several protein/ligand ratios. NOESY spectra acquired with a variety of mixing times were analyzed using a full matrix relaxation approach.[66] The observed data indicated that the protein causes a conformational variation in the glycosidic torsion angles of methyl α-lactoside, changing toward smaller angle values, although the recognized conformer was still within the lowest energy region.[46,72]

In the same article, a second complex was also analyzed, that between methyl β-allolactoside [Galβ(1 → 6)Glcβ-OMe] and ricin-B. In this case, and contrary to the observations for methyl β-lactoside, 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 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.

Molecular modeling using molecular dynamics, minimization, and docking of both disaccharides within the binding sites of ricin-B strongly suggested that, for methyl α-lactoside, 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.[76] Thus, as also deduced from the crystallographyc analysis, both van der Waals contacts and hydrogen bonding contribute to the stability of the complex (Color Plate 1).

This result was in contrast with the X-ray analysis of the ricin-B/lactose and ricin-B/galactose-containing biantennary hexasaccharide complexes.[22] On the other hand, for methyl β-allolactoside, the experimental data and the modeling studies demonstrated that only the galactose moiety was bound by the lectin and, therefore, the conformation around the glycosidic angles could probably be described by a distribution similar to that existing in free solution.

From the NMR experimental point of view, this work[78] also demonstrated that the measurement of transferred NOEs in the rotating frame (TR-ROESY) permits one to distinguish direct enhancements from spin-diffusion effects, and thus complements those measured under regular conditions, providing conformational information that is less contaminated by spin-diffusion effects.[70,71]

Ricin-B has also been used as a model to study the bound conformation of potential glucosidase inhibitors such as C-glycosides.[79,80] These carbon-bridged compounds are thought to affect the activity of glucosidases as competitive inhibitors of the natural compounds. Therefore, it is extremely important that both compounds, the natural substrate and the corresponding C-glycoside, are recognized by the enzyme in a similar way. Although many reports have usually assumed that the conformation of free C-glycosides was the same as that of the corresponding O-analogs,[81] it has been reported that, at least for O- and C-lactoses, this is not the case.[82]

A second report[21] studied whether both types of compounds were recognized by a sugar-binding protein in the same conformation. Apart from the presence of key groups in the non-natural substrate, this would be an essential requirement for an important inhibitory activity to occur. Thus, 2D transferred NOE (TR-NOE) and ROE (TR-ROE) experiments were reported to study the complexation of C-lactose by ricin-B.

The conformational study of C-lactose in the free state was performed using NMR and MM3 calculations[82] and showed that the exoanomeric conformation around the C-glycosidic bond was adopted. However, the conformation around the aglyconic bond differed from that of the natural compound. For O-lactose, ca. 90% of the population was located in the central low-energy region (minimum syn, ϕ/ψ: 54/18) and ca. 10% of population with similar values but with values around −180° (minimum B, ϕ/ψ: 36/180). These two different dispositions around the aglyconic torsion were named syn and anti conformations, respectively.[83] However, C-lactose was shown to exhibit much higher flexibility than its O-analog and three conformational regions were significantly populated in solution. The anti conformer as well as the syn and gauche-gauche orientations were detected.

The observed NOEs for C-lactose in the presence of ricin-B were negative, thus indicating binding. The comparison between the NOESY spectra of C-lactose, recorded in the absence and in the presence of the lectin, indicated that minima syn and gauche-gauche were not bound.

To confirm the existence of important differences between the free and bound forms, and to differentiate spin-diffusion effects from direct NOE enhancements, TR-ROEs were also recorded. A quantitative analysis of the TR-NOE data obtained for C-lactose/ricin-B was attempted using simulation of the NOE buildup curves according to the full relaxation matrix method.[66] The obtained results evidenced the presence of mobility of C-lactose even in the binding site, so that several conformers within the valley defined by minimum B, anti, were still possible.

More important, the experimental results indicated, unequivocally and unexpectedly, the ricin-B selects different conformers of C-lactose (syn), and its O-analog (anti). 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 observed that the increase in the amount of C-lactose added to the NMR tube induced a decrease in the negative NOE cross-peak intensities of O-lactose, and when a 1:1 ratio was reached, the NOEs of C-lactose became positive, whereas strong negative NOEs were clearly observed for O-lactose. Two different conclusions arose from these results: (1) Both ligands compete for the same binding sites on the lectin and (2) the affinity constant of C-lactose is smaller than that of its O-analog.

From the study of the crystal structure of the complex of ricin-B with lactose, no unambiguous evidence could be taken as responsible for the change in conformation of C-lactose with respect to O-lactose. 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, and, for both compounds, the bound conformation is the most populated in water for the corresponding free molecule.

As stated previously, protein-mediated, indirect TR-NOE effects may lead to interpretation errors in the analysis of the ligand-bound conformation. As a leading example, one of the

first reported uses of TR-NOE experiments to determine the derivation of oligosaccharide-bound conformations concluded[59] that a fluorinated Gal-β-(1 → 6)-Gal-β-OMe disaccharide underwent major conformational changes around the glycosidic linkages when binding to its specific antibody. The conclusion was based on the detection of an NOE cross-peak between two protons located on the two different pyranoid moieties. However, the reevaluation of the problem by the same authors, now by use of TR-ROE experiments[71] for a sample containing a 40:1 ratio of the same disaccharide and the Fab′ derived from the monoclonal antibody IgA X24, demonstrated that this cross-peak is dominated by an indirect effect, mediated by a protein proton.

Other cases of protein-induced conformational changes have also been reported. Bundle and co-workers[62] have reported TR-NOE evidence that shows 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 demonstrated that only this trisaccharide portion of the complete polysaccharide was bound by the antibody, with one of the monosaccharide rings deeply buried within the binding site. Although the TR-NOE 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 free solution conformation of the oligosaccharide.

The monoclonal antibody Strep 9-bound conformation of a branched but flexible trisaccharide, namely GlcNac-β-(1 → 3)-α-Rha-(1 → 2)-α-Rha-OMe, has been[84] investigated to TR-NOE and TR-ROE 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 on the orientation of the angle of the 1 → 2 linkage when compared to the global minimum conformation. From the experimental point of view, it was reported that TR-ROE experiments were essential to solve the bound conformation of the sugar. Finally, it was also necessary to include repulsive constraints, derived from the absence of TR-ROEs, to deduce the three-dimensional structure of the trisaccharide in the binding site of the antibody.

Nevertheless, there are examples in which the protein does not select a single conformer. The *Aurelia aurantia* agglutinin (AAA) lectin recognizes different conformations[63] of Fuc-α(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.

The conformation of the Lewis blood group antigens and their derivatives has been a topic of interest, and the quest for the active conformation of the Lewis X oligosaccharide has stimulated different research groups. The sialyl Lewis X tetrasaccharide exists in solution as an equilibrium of several conformations, which are mainly characterized by the orientations of the *N*-acetylneuraminic acid residue. One report on the investigation of the bound conformation of sialyl Lewis X to E-selectin[85] showed a conformational change around the sialyl-galactose torsions between the free and bound oligosaccharide on binding. Docking studies of the bound conformation of the tetrasaccharide onto a model of the protein suggested that the observed variation around the sialic-galactose glycosidic linkage resulted primarily from steric interactions. A different set of observed NOEs was reported by Hensley et al.,[86] which claimed that the conformation in the bound state is identical to that in free solution. Finally, and perhaps the preeminent study on this topic, has been reported by Scheffler and co-workers[87] using spin-locked filtered NOESY and Metropolis Monte Carlo calculations.

The most relevant conclusion is that E-selectin complexes exclusively with one conformation of sialyl Lewis X among those constituting the conformational equilibrium in aqueous solution, that one being the sialic acid showing an orientation defined by (ϕ/ψ: −76/6), already reported to be present in free solution.[88,89] On the other hand, the orientation of the fucose residue (ϕ,ψ: 38/26) differs from that preferred in aqueous solution.

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.[90] Comparison between experimental and simulated TR-NOE volumes, using a full matrix relaxation approach, led to the conclusion that only one conformation of the trisaccharide was bound[91] by the GalNAc-specific lectin isolated from *Dolichos biflorus*. Spin-diffusion NOEs were detected by means of TR-ROE experiments. The proposed bound conformation was in agreement with one of the two deduced from previous modeling studies.[92] 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.[93] TR-NOE and TR-ROE 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 nonreducing 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 that reflected the size of the carbohydrate epitope and thus also of the binding pocket of the lectin. To detect contacts between the protein and the carbohydrate chain, T_2-filtered NOESY spectra were performed, which allowed the detection of NOEs between the terminal disaccharide fragment and protein protons, most likely belonging to leucine residues. These observations were in agreement with the previously reported molecular modeling study of the complex.[92]

The conformational features of one of the most relevant of food industry sugars, sucrose, in the combining site of lentil lectin in solution have been characterized through TR-NOE experiments and molecular modeling.[94] 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 using X-ray crystallography. It is important to mention that major differences in the hydrogen-bonding network of sucrose were found, since neither of the two interresidue hydrogen bonds detected in crystalline sucrose[95,96] were conserved in the complex with the lectin. Stacking interactions between a phenylalanine residue and the hydrophobic face of the glucose residue, as well as between the same phenylalanine 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.

Passing to the recognition of modified carbohydrates by enzymes, Andrews and co-workers[97] have investigated *Aspergillus niger* glucoamylase. This enzyme catalyzes the hydrolysis of maltose-type molecules with inversion of configuration. In an elegant manner, and using novel hetero-analogs of maltose containing sulfur in the nonreducing ring and nitrogen in the interglycosidic linkage, they have demonstrated that methyl 5′-thio-4-*N*-α-maltoside is bound by the enzyme in a conformation close to its global minimum. The characteristic NOEs observed for a second conformer that 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 dihydroglucocarbose[98] indicated that the bound conformation of this ligand resembles the global minimum and that the local minimum conformer cannot be readily accommodated by the enzyme because of adverse van der Waals interactions. In addition, they also showed that this compound is a potent enzyme inhibitor.

It is not frequent that carbohydrate-binding proteins are directly amenable to NMR analysis. However, hevein is a cysteine-rich single-chain protein of 43 amino acids,[99] whose structure has independently been solved by X-ray at 0.28-nm resolution,[100] and by NMR methods.[101] Interestingly, although the structure of hevein in water-dioxane solution[101] differs significantly from that observed in the crystal,[100] it closely resembles the solid-state structures of the domains of wheat germ agglutinin (WGA).[102,103]

Hevein and WGA are two chitin-binding proteins that have a high degree of homology.[104] In fact, according to high-resolution X-ray analysis,[102,103] the four domains of WGA have similar three-dimensional structures and, in terms of primary structure, hevein shows a 56% sequence identity to the C domain of WGA. This so-called hevein domain (43 amino acids) also has partial identity to several chitin-binding domains in other lectins[56,105,106] and enzymes.[107] On the other hand, it has been proposed that hevein is involved in the coagulation of latex[108] by interacting with a 22-kDa glycoprotein through binding to an *N*-acetylglucosamine (GlcNAc) residue, and that this interaction is mediated by Ca^{2+}.

Asensio et al.[109] have reported on the determination of the ligand-binding site of hevein as well as the structure of its complex with chitobiose, by using NMR spectroscopy. Using NOESY and restrained molecular dynamics, they also presented a refined NMR structure (0.092-nm backbone rmsd over residues 3–41 for the 20 best structures) of free hevein in water (Color Plate 2). The structure of the complex of hevein with methyl β-chitobiose has also been derived.[110] In this case, the backbone rmsd is 0.055 nm. The GlcNAc-, chitobiose-, and chitotriose-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 disaccharide, and from the dito the trisaccharide (Figure 2). The binding constant for the disaccharide is ca. 1000.

In addition, the thermodynamic parameters for chitotriose-hevein and chitobiose-hevein interactions have been obtained from a van't Hoff analysis, indicating that the association process is enthalpy driven, while entropy opposes binding. This

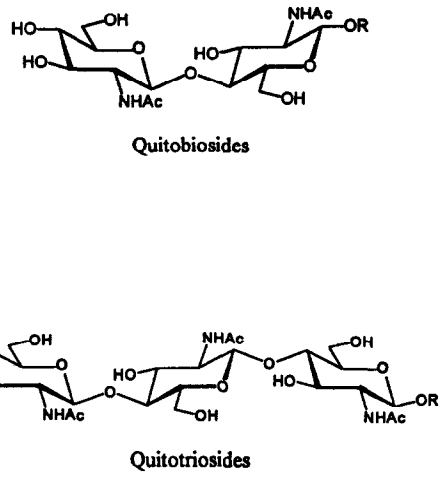

Figure 2. Schematic view of chitobiose derivatives.

negative sign indicated that hydrogen bonding and 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, also obtained from NOESY and NMR spectroscopy.

Protein-carbohydrate nuclear Overhauser enhancements measured for the hevein-chitobiose and hevein-methyl β-chitobiose[110] complexes allowed the building of three-dimensional models of these complexes. 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 nonreducing GlcNAc moiety of β-chitobiose displayed NOE contacts with Tyr-30 and Trp-21 residues and appeared strongly shielded. From an inspection of the model, a hydrogen bond between Ser-19 and the nonreducing *N*-acetyl carbonyl group was suggested as well as between Tyr-30 and HO-3 of the same sugar residue. The previously mentioned *N*-acetyl methyl group of the nonreducing GlcNAc displayed nonpolar contacts with the aromatic Tyr-30 and Trp-21 residues. Moreover, the higher affinities deduced for the β-linked oligosaccharides with respect to GlcNAc and GlcNAcα-(1 $\rightarrow$ 6)-Man could be explained by favorable stacking of the second β-linked GlcNAc moiety and Trp-21.

Finally, since the measured K_a values for chitobiose binding are almost identical with and without calcium ions, it is shown that these cations are not required for sugar binding.

The final 3D structures derived by NMR were compared to those of WGA, Ac-AMP II (which is also a chitin-binding protein, solved by NMR; Martins et al.),[56] and to the crystal structure of hevein. The corresponding average rmsd values 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).

CONCLUDING REMARKS

Carbohydrate-protein interactions are involved in a variety of recognition phenomena related to the function of living organisms. The three-dimensional structure of both entities are of prime importance in the interaction process. Knowledge of the

bound conformation of the oligosaccharide ligand may nowadays be deduced from TR-NOE experiments, assisted by molecular modeling. Although new examples of the application of homonuclear NOE spectroscopy to the study of protein-carbohydrate complexes are still necessary, continuous effort in molecular biology methods and access to labeled protein[48] and carbohydrates will dramatically modify in the near future the way in which these molecular complexes are studied. In addition, new experimental NMR methods that permit the detection of NOEs to water molecules[111] will also probably allow the detection of bound water molecules in the binding sites of lectins in order to help clarify their specific roles in different complexes.

ACKNOWLEDGMENTS

We thank DGICYT for financial support (PB 93 1027).

We thank M. Martin-Lomas, A. Imberty, F. Casset, S. Perez, and T. Peters for fruitful discussions. We thank Dr. Martins for the coordinates of Ac-AMP2. We also thank SIdI-UAM for the facilities provided.

REFERENCES

1 Rademacher, T.W., Parekh, R.B., and Dwek, R.A. *Annu. Rev. Biochem.* 1988, **57,** 785–823
2 Phillips, M.L., Nudelman, E., Gaeta, F.C.A., Perez, M., Shingal, A.K., Hakomori, S., and Paulson, J.C. *Science* 1990, **250,** 1130–1133
3 Harlan, J.M. and Liu, D.Y. *Adhesion: Its Role in Inflammatory Disease.* W.H. Freeman, New York, 1992
4 Sharon, N. and Lis, H. *Science* 1989, **246,** 227–234
5 Shaanan, B., Lis, H., and Sharon, N. *Science* 1991, **254,** 862–866
6 Williams, B.A., Chervenak, M.C., and Toone, E.J. *J. Biol. Chem.* 1992, **207,** 22907–22911
7 Mandal, D.K., Kishore, N., and Brewer, C.F. *Biochemistry* 1994, **33,** 1149–1156
8 Mandal, D.K., Bhattacharyya, L., Koenig, S.H., Brown, R.D., Oscarson, S., and Brewer, C.F. *Biochemistry* 1994, **33,** 1157–1162
9 Toone, E. *Curr. Opin. Struct. Biol.* 1994, **4,** 719–728
10 Drickamer, K. *Nature (London)* 1992, **360,** 183–186
11 Bundle, D.R. and Young, N.M. *Curr. Opin. Struct. Biol.* 1992, **2,** 666–673
12 Lemieux, R.U. *Chem. Soc. Rev.* 1989, **18,** 347–374
13 Troganis, A. and Stassinopoulou, C.I. *Biochim. Biophys. Acta* 1994, **1206,** 215–224
14 Quiocho, F.A. *Annu. Rev. Biochem.* 1986, **55,** 287–315
15 Vyas, N.K. *Curr. Opin. Struct. Biol.* 1991, **1,** 732–740
16 Bourne, Y., van Tielbeurgh, H., and Cambillau, C. *Curr. Opin. Struct. Biol.* 1993, **3,** 681–686
17 Bothner-By, A.A. and Gassend, R. *Ann. N.Y. Acad. Sci.* 1972, **222,** 668–676
18 Albrand, J.P., Birdsall, B., Feeney, J., Roberts, G.C.K., and Burgen, A.S.V. *Int. J. Biol. Macromol.* 1979, **1,** 37–41
19 Clore, G.M. and Gronenborn, A.M. *J. Magn. Reson.* 1982, **48,** 402–417
20 Clore, G.M. and Gronenborn, A.M. *J. Magn. Reson.* 1983, **53,** 423–442
21 Espinosa, J.-F., Cañada, J., Asensio, J.L., Dietrich, H., Martin-Lomas, M., Schmidt, R.R., and Jiménez-Barbero, J. *Angew. Chem. Int. Ed. Engl.* 1996, **35,** 303–306
22 Rutenber, E. and Robertus, J.D. *Proteins* 1991, **10,** 260–269
23 Hardman, K.D., Agarwal, R.C., and Freiser, M.J. *J. Biol. Chem.* 1982, **157,** 69–86
24 Qian, M., Haser, R., Buisson, G., Due, E., and Payan, F. *Biochemistry* 1994, **33,** 6284–6289
25 Quiocho, F.A., Wilson, D.K., and Vyas, N.K. *Nature (London)* 1989, **340,** 404–407
26 Cygler, M., Rose, D.R., and Bundle, D.R. *Science* 1991, **253,** 442–445
27 Loris, R., Casset, F., Bouckaert, J., Pletinckx, J., Dao-Thi, M., Poortmans, F., Imberty, A., Perez, S., and Wyns, L. *Glycoconj. J.* 1994, **11,** 507–517
28 Weis, W.I., Drickamer, K., and Hendrickson, W.A. *Nature (London)* 1992, **360,** 127–134
29 Bourne, Y., Rouge, P., and Cambilleau, C. *J. Biol. Chem.* 1992, **267,** 197–203
30 Goodford, P.J. *J. Med. Chem.* 1985, **28,** 849–857
31 Imberty, A., Hardman, K.D., Carver, J.P., and Perez S. *Glycobiology* 1991, **1,** 456–484
32 Imberty, A. and Perez, S. *Glycobiology* 1994, **4,** 351–366
33 Thogersen, H., Lemieux, R.U., Bock, K., and Meyer, B. *Can. J. Chem.* 1982, **60,** 44–57
34 Perez, S. Ph.D. thesis. Grenoble University, Grenoble, France, 1978
35 Allinger, NL, Rahman, M, and Lii, J-H. *J. Am. Chem. Soc.* 1990, **112,** 8293–8307
36 French, A.D. and Dowd, M.K. *J. Mol. Struct. Theochem.* 1990, **286,** 183–201
37 Ha, S.N., Giammona, A., Field, M., and Brady, J.W. *Carbohydr. Res.* 1988, **180,** 207–221
38 Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P.J. *J. Am. Chem. Soc.* 1984, **106,** 765–776
39 Homans, S.W. *Biochemistry* 1990, **29,** 9110–9118.
40 Woods, R.J., Dwek, R.A., Edge, C.J., and Fraser-Reid, B. *J. Phys. Chem.* 1995, **99,** 3832–3846
41 Senderowitz, H., Parish, C., and Still, W.C. *J. Am. Chem. Soc.* 1996, **118,** 2078–2086
42 Koehler, J.E.H., Saenger, W., and van Gunsteren, W.F. *Biophys. J.* 1987, **15,** 211–214
43 Perez, S. *Sucrose: Properties and Application.* (M. Mathlouthi and P. Reiser, eds.). Blackie, Academic and Professional, London, 1995, pp 11–32
44 Hagler, A.T., Lifson, P., and Dauber, P. *J. Am. Chem. Soc.* 1979, **101,** 5122–5130
45 Seibert, H.C., Reuter, G., Schauer, R., von der Lieth, C.W., and Dabrowski, J. *Biochemistry* 1992, **31,** 6962–6971
46 Asensio, J.L., Martin-Pastor, M., and Jiménez-Barbero, J. *Int. J. Biol. Macromol.* 1995, **17,** 137–148
47 Woods, R.J. *Curr. Opin. Struct. Biol.* 1995, **5,** 591–598
48 Weller, C.T., Lustbader, J., Seshadri, K., Brown, J.M., Chadwick, C.A., Kolthoff, C.E., Ramnarain, S., Pollak, S., Canfield, R., and Homans, S.W. *Biochemistry* 1996, **35,** 8815–8823
49 Fletcher, C.M., Harrison, R.A., Lachmann, P.J., and Neuhaus, D. *Structure* 1994, **2,** 185–199
50 Withka, J.M., Wyss, D.F., Wagner, G., Arulanandam, A.R.N., Reinherz, E.L., and Recny, M.A. *Structure* 1993, **1,** 69–81

51 Wyss, D.F., Choi, J.S., and Wagner, G. *Biochemistry* 1995, **34,** 1622–1634
52 Lommersee, J.P.M., Kroon-Batenburg, L.M.J., Kamerling, J.P., and Vliegenthart, J.F.G. *Biochemistry* 1995, **34,** 8196–8206
53 Ruud, P.M., Woods, R.J., Wormald, M.R., Opdenakker, G., Downing, A.K., Campbell, I.A., and Dwek, R.A. *Biochim. Biophys. Acta* 1995, **1284,** 1–10
54 Verheyden, P., Pletinckx, J., Maes, D., Pepermans, H.A.M., Wyns, L., Willem, R., and Martins, J. *FEBS Lett.* 1995, **370,** 245–249
55 Hom, K., Gochin, M., Peumans, W.J., and Shine, N. *FEBS Lett.* 1995, **361,** 157–161
56 Martins, J., Maes, D., Loris, R., Pepermans, H.A.M., Wyns, L., Willem, R., and Verheyden, P. *J. Mol. Biol.* 1996, **258,** 322–333
57 Kronis, K.A. and Carver, J.P. *Biochemistry* 1985, **24,** 834–840
58 Kronis, K.A. and Carver, J.P. *Biochemistry* 1985, **24,** 826–833
59 Glaudemans, C.P.J., Lerner, L., Daves, G.D., Kovac, P., Venable, R., and Bax, A. *Biochemistry* 1990, **29,** 10906–10911
60 Bevilacqua, V.L., Thomson, D.S., and Prestegard, J.H. *Biochemistry* 1990, **29,** 5529–5537
61 Bevilacqua, V.L., Kim, Y., and Prestegard, J.H. *Biochemistry* 1992, **31,** 9339–9349
62 Bundle, D.R., Baumann, H., Brisson, J.-R., Gagne, S.M., Zdanov, A., and Cygler, M. *Biochemistry* 1994, **33,** 5183–5192
63 Weimar, T. and Peters, T. *Angew. Chem. Int. Ed. Engl.* 1994, **33,** 88–91
64 Neuhaus, D. and Williamson M.P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis.* VCH, New York, 1989
65 Campbell, A.P. and Sykes, B.D. *Annu. Rev. Biophys. Biomol. Struct.* 1993, **22,** 99–122
66 London, R.E., Perlman, M.E., and Davis, D.G. *J. Magn. Reson.* 1992, **97,** 79–98
67 Ni, F. and Scheraga, H.A. *Acc. Chem. Res.* 1994, **9,** 257–264
68 Bothner-By, A.A., Stephens, R.L., Lee, J.-M., Warren, C.D., and Jeanloz, R.W. *J. Am. Chem. Soc.* 1984, **106,** 811–813
69 Brown, L.R. and Farmer, B.T., II. *Methods Enzymol.* 1989, **176,** 199–216
70 Perlman, M.E., Davis, D.G., Koszalka, G.W., Tuttle, J.V., and London, R.E. *Biochemistry* 1994, **33,** 7547–7559
71 Arepalli, S.R., Glaudemans, C.P.J., Daves, G.D., Kovac, P., and Bax, A. *J. Magn. Reson. B* 1995, **106,** 195–198
72 Asensio, J.L. and Jiménez-Barbero, J. *Biopolymers* 1995, **35,** 55–73
73 Olsnes, S. and Pihl, A. *The Molecular Actions of Toxins and Viruses* (P. Cohen, and S. Van Heyningen, eds.). Elsevier Biomedical Press, New York, 1982, pp. 52–105
74 Baenzinger, J.U. and Fiete, D. *J. Biol. Chem.* 1979, **254,** 9795–9799
75 Rivera-Sagredo, A., Solis, D., Diaz-Mauriño, T., Jiménez-Barbero, J., and Martín-Lomas, M. *Eur. J. Biochem.* 1991, **197,** 217–228
76 Solis, D., Fernandez, P., Diaz-Mauriño, T., Jiménez-Barbero, J., and Martin-Lomas, M. *Eur. J. Biochem.* 1993, **214,** 677–683
77 Fernandez, P., Jiménez-Barbero, J., Martin-Lomas, M., Solis, D., and Diaz-Mauriño, T., *Carbohydr. Res.* 1994, **256,** 223–244
78 Asensio, J.L., Cañada, F.J., and Jiménez-Barbero, J. *Eur. J. Biochem.* 1995, **233,** 618–630
79 Dietrich, H., and Schmidt, R.R. *Liebigs Ann. Chem.* 1994, 975–981
80 Wei, A., Haudrechy, A., Audin, C., Hyuk-Sang, J., Haudrechy-Bretel, N., and Kishi, Y. *J. Org. Chem.* 1995, **60,** 2160–2170
81 Wang, Y., Goekjian, P.G., Ryckman, D.V., Miller, W.H., Babirad, S.A., and Kishi, Y. *J. Org. Chem.* 1992, **57,** 482–489
82 Espinosa, J.-F., Martin-Pastor, M., Asensio, J.L., Dietrich, H., Martin-Lomas, M., Schmidt, R.R., and Jiménez-Barbero, J. *Tetrahedron Lett.* 1995, **36,** 6329–6332
83 Dabrowski, J., Kozár, T., Grosskurth, H., and Nifant'ev, N.E. *J. Am. Chem. Soc.* 1995, **117,** 5534–5539
84 Weimar, T., Harris, S.L., Pitner, J.B., Bock, K., and Mario Pinto, B. *Biochemistry* 1995, **34,** 13672–13681
85 Cooke, R.M., Hale, R.S., Lister, S.G., Shah, G., and Malcolm, P.W. *Biochemistry* 1994, **33,** 10591–10596
86 Hensley, P., McDevitt, P.J., Brooks, I., Trill, J.J., Feild, J.A., McNuilty, D.E., Connor, J.R., Griswold, D.E., Kumar, N.V., Kopple, K.D., Carr, S.A., Dalton, B.J., and Johanson, K. *J. Biol. Chem.* 1994, **269,** 23949–23958
87 Scheffler, K., Ernst, B., Katopodis, A., Magnani J.L., Wang, W.T., Weisemann, R., and Peters, T. *Angew. Chem. Int. Ed. Engl.* 1995, **34,** 1841–1844
88 Ishikawa, Y., Lin, Y.-C., Dumas, D.P., Shen, G.-J., Garcia-Junceda, E., Williams, M.A., Bayer, R., Ketcham, C., Walker, L.E., Paulson, J.C., and Wong, C.-H. *J. Am. Chem. Soc.* 1992, **114,** 9283–9298
89 Rutherford, T.J. and Homans, S.W. *Biochemistry* 1994, **33,** 9606–9614
90 Imberty, A., Mikros, E., Koca, J., Mollicone, R., Oriol, R., and Perez, S. *Glycoconj. J.* 1995, **12,** 331–349
91 Casset, F., Peters, T., Etzler, M., Korchagina, E., Nifant'ev, N., Perez, S., and Imberty, A. *Eur. J. Biochem.* 1996, **239,** 710–719
92 Imberty, A., Casset, F., Gegg, C.V., Etzler, M.E., and Perez, S. *Glycoconj. J.* 1994, **11,** 400–413
93 Casset, F., Imberty, A., Perez, S., Etzler, M., Paulsen, H., and Peters, T. 1997 (in press)
94 Casset, F., Hamelryck, T., Loris, R., Brisson, J.-R., Tellier, C., Dao-Thi, M., Wyns, L., Poortmans, F., Perez, S., and Imberty, A. *J. Biol. Chem.* 1995, **270,** 5619–25628
95 Brown, G.M. and Levy, H.A. *Acta Crystallogr. Sect. B.* 1973, **29,** 790–797
96 Hanson, J.C., Sieker, L.C., and Jensen, L.H. *Acta Crystallogr. Sect. B.* 1973, **29,** 797–808
97 Andrews, J.S., Weimar, T., Frandsen, T.P., Svensson, B., and Mario Pinto, B. *J. Am Chem. Soc.* 1995, **117,** 10799–10804
98 Stoffer, B., Aleshin, A.E., Firsov, L.M., Svensson, B., and Honzatko, R.B. *FEBS. Lett.* 1995, **358,** 57–61
99 Rodriguez, A., Tablero, M., Barragan, B., Lara, P., Rangel, M., Arreguin, B., Possani, L., and Soriano-Garcia, M. *J. Crystal Growth* 1986, **76,** 710–714

100 Rodriguez-Romero, A., Ravichandran, K.G., and Soriano-Garcia, M. *FEBS. Lett.* 1991, **291,** 307–309
101 Andersen, N.H., Cao, B., Rodríguez-Romero, A., and Arreguin, B. *Biochemistry* 1993, **32,** 1407–1422
102 Wright, C.S. *J. Mol. Biol.* 1990, **215,** 635–651
103 Wright, C.S. *J. Biol. Chem.* 1992, **267,** 14345–14352
104 Drenth, J., Low, B.W., Richardson, J.S., and Wright, C.S. *J. Biol. Chem.* 1980, **255,** 2652–2655
105 Verheyden, P., Jaspers, H., Martins, J.C., Maes, D., Pletinckx, J., van Binst, G., Wyns, L., and Willem, R. *Arch. Int. Phys. Biochim. Biophys.* 1994, **102,** B43
106 Kieliszewski, M.J., Showalter, A.M., and Leykam, J.F. *Plant J.* 1994, **5,** 849–861
107 Beintema, J.J. *FEBS. Lett.* 1994, **350,** 159–163
108 Gidrol, X., Chrestin, H., Tan, H.-L., and Kush, A. *J. Biol. Chem.* 1994, **209,** 9278–9283
109 Asensio, J.L., Cañada, J., Bruix, M., Rodriguez-Romero, A., and Jiménez-Barbero, J. *Eur. J. Biochem.* 1995, **230,** 621–633
110 Asensio, J.L., et al. 1997 (in preparation)
111 Piotto, M., Saudek, V., and Sklenar, V. *J. Biomol. NMR* 1992, **2,** 661–665