Bovine Heart Galectin-1 Selects a Unique (Syn) Conformation of C-Lactose, a Flexible Lactose Analogue

J. L. Asensio,[†] J. F. Espinosa,[†] H. Dietrich,[†] F. J. Cañada,[†] R. R. Schmidt,[‡] M. Martín-Lomas,[§] S. André,[||] H.-J. Gabius,*[,||] and J. Jiménez-Barbero*[,†]

Contribution from the Department Química Orgánica Biológica, Instituto Química Orgánica, CSIC, Juan de la Cierva 3, 28006 Madrid, Spain, Institut für Physiologische Chemie, Tierärztliche Fakultät, Ludwig-Maximilians Universität, Veterinaerstrasse 13, D-80539 München, Germany, Institut Organische Chemie, University Konstanz, Universitaetstrasse 10, D-78464 Konstanz, Germany, and Instituto de Investigaciones Químicas, CSIC, Isla de la Cartuja, Sevilla, Spain[†]

Received February 25, 1999. Revised Manuscript Received July 29, 1999

Abstract: The C-glycoside analogue of lactose harbors a pronounced flexibility in water with three conformers in equilibrium. The bound conformation of C-lactose to bovine heart galectin-1 in solution has been determined by NMR spectroscopy. It is demonstrated that the lectin selects the syn conformation of the structural analogue of natural lactose and not the global minimum, anti conformation. The bound conformer resembles those found in the crystal structures of complexes of galectin-1/*N*-acetyllactosamine-containing oligosaccharides and in solution for an avian galectin. Docking of the analogue within the galectin's binding site furnishes explanations, in structural terms, for the exclusive recognition of the syn conformer.

Introduction

Carbohydrate−protein interactions are assumed to be involved in a wide range of biological activities starting from fertilization and extending to pathological processes such as tumor spread.[1] Among the receptors, galectins, a family of galactoside-binding proteins with conserved binding-site topology embedded in a jelly roll motif, take part in diverse cellular activities including cell recognition, growth control, and apoptosis.[2] The potential for mediation of cellular contacts in the metastatic cascade can render galectins attractive targets for drug design, prompting, for example, the chemical mapping of binding sites in galectins with engineered (deoxy, fluoro) ligands following the methodology described by Lemieux.[3] These studies have provided evidence for interactions that extend to the penultimate sugar unit in line with recent crystallographic studies.[4]

On the way to deduce the ligands properties for a rational drug design,[3a,5] it is pertinent to determine the topological ligand features when they are free in solution and especially in the complex with the galectin. The molecular weight of galectins at 14−16 kDa per carbohydrate recognition domain-bearing monomer places them at the limit of direct ^{1}H NMR observations using current strategies.[6] However, information about the conformation of complexed ligands can be derived from transferred NOE studies (TRNOE), provided that the exchange between the complexed and uncomplexed states is sufficiently fast,[7] as pioneered by Prestegard for studying carbohydrate−protein interactions. Following this methodology, several cases have been described,[7] including the conformational analysis of lactose derivatives bound to an avian galectin.[8] Notably, the conditions required to monitor TR−NOEs appear to be satisfied frequently by sugar receptors.[9] Since many ligand analogues are still objects of hydrolytic attacks, C-glycosides[10] afford the possibility for an improved chemical and biochemical stability evocative of developments of modified ribonucleic acids such as thioates for therapeutic purposes in the antisense technology. However, the methylene-bridged analogues do not simply behave as noncleavable glycosides, and some differences

* Correspondence to Dr. J. Jiménez-Barbero, Instituto Química Orgánica, CSIC, Juan de la Cierva 3, 28006 Madrid, Spain. E-mail iqojj01@pinar.csic.es. Fax: +34-91-5644853. Phone: +34-91-5618806.

[†] Instituto Química Orgánica.

[‡] University Konstanz.

[§] Instituto de Investigaciones Químicas.

[||] Ludwig-Maximilians Universität.

(1) Lee, Y. C.; Lee, R. T. *J. Biomol. Sci.* **1997**, *3*, 221−237. (b) *Glycosciences: Status and Perspectives*; Gabius, H.-J., Gabius, S., Eds.; Chapman & Hall: London, U.K., 1997. (c) Gabius, H.-J., Sinowatz, F., Eds; Special volume on glycosciences. *Acta Anat.* **1998**, *161*, 1−276.

(2) Rini, J. M. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *24*, 551−577. (b) Gabius, H.-J. *Eur. J. Biochem.* **1997**, *243*, 543−576. (c) Hirabayashi, J. *Trends Glycosci. Glycotechnol.* **1997**, *9*, 1−190. (d) Kaltner, H.; Stiersdorfer, B. *Acta Anat.* **1998**, *161*, 162−179. (e) Kopitz, J.; von Reitzenstein, C.; Burchert M.; Cantz M.; Gabius, H. J. *J. Biol. Chem.* **1998**, *273*, 11205−11211. (f) Perillo, N. L.; Marcus, M. E.; Baum, L. G. *J. Mol. Med.* **1998**, *76*, 402−412.

(3) Lemieux, R. U. *Chem. Soc. Rev.* **1989**, *18*, 347−374. (b) Solis, D.; Jimenez-Barbero, J.; Martin-Lomas, M.; Diaz-Mauriño, T. *Eur. J. Biochem.* **1994**, *223*, 107−114. (c) Solis, D.; Romero, A.; Kaltner, H.; Gabius, H.-J.; Diaz-Mauriño, T. *J. Biol. Chem.* **1996**, *271*, 12744−12748.

(4) Bourne, Y.; Bolgiano, B; Liao, D. I.; Strecker, G.; Cantau, P.; Herzberg, O.; Feizi, T.; Cambillau, C. *Nat. Struct. Biol.* **1994**, *1*, 863−870. (b) Liao, D. I.; Kapadia, G.; Ahmed, H.; Vasta, G. R.; Herzberg, O. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1428−1432.

(5) Lemieux, R. U. *Acc. Chem. Rev.* **1996**, *29*, 373−380. (b) Gabius, H. J. *Pharm. Res.* **1998**, *15*, 23−30.

(6) a) James, T. L.; Oppenheimer, N. J. Nuclear magnetic resonance. Part C. *Methods Enzymol.* **1994**, *239*, 1−813. (a) For new trends in NMR, see: Prestegard, J. H. *Nat. Struct. Biol.* **1998**, *6*, 517−523.

(7) Ni, F. *Prog. NMR Spectrosc.* **1994**, *26*, 517−606. (b) Bevilacqua, V. L.; Thomson, D. S.; Prestegard, J. H. *Biochemistry* **1990**, *29*, 5529. (c) Bevilacqua, V. L.; Kim, Y.; Prestegard, J. H. *Biochemistry* **1992**, *31*, 9339.

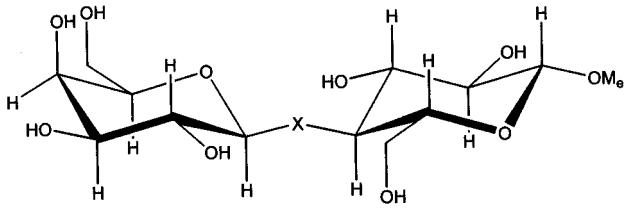
(8) Siebert, H.-C.; Gilleron, M.; Kaltner, H.; von der Lieth, C.-W.; Kozar, T.; Bovin, N. V.; Korchagina, E. Y.; Vliegenthart, J. F. G.; Gabius, H.-J. *Biochem. Biophys. Res. Commun.* **1996**, *219*, 205−212. (b) von der Lieth C. W.; Siebert, H.; Kozar, T.; Burchert, M.; Frank, M.; Gilleron, M.; Kaltner, H.; Kayser, G.; Tajkhorshid, E.; Bovin, N. V.; Vliegenthart, J. F. G.; Gabius, H.-J. *Acta Anat.* **1998**, *161*, 91−109.

(9) Peters, T.; Pinto, B. M. *Curr. Opin. Struct. Biol.* **1996**, *6*, 710−720. (b) Poveda, A.; Jimenez-Barbero, J. *Chem. Soc. Rev.* **1998**, *27*, 133−143.

(10) Levy, D. E.; Tang, C. *The Chemistry of C-glycosides*; Pergamon: New York, 1995.

10.1021/ja990601u CCC: $18.00
Published on Web 09/18/1999

Chart 1. Schematic View of C-lactose **1** and O-lactose **2**.



1) X=CH_2 2) X=O

between the behavior of C- and O-glycosides have been reported.[11,12] Their conformational properties have been debated in a way that can appear as controversial when only being considered superficially.[12,13] This controversy has especially focused on the O-lactose/C-lactose pair (Chart 1). C-lactose and O-lactose display several unambiguous geometric and structural differences (according to MM3* calculations, C1′-Cα, 1.538 versus C′1-O1′, 1.426 Å and O5′-C1′-Cα 107.4° versus O5′-C1′-O1′, 108.3°). In addition, the *exo*-anomeric effect due to the presence of the electroacceptor interglycosidic oxygen atom disappears in the C-glycoside, along with a consequent variation of the steric interactions between both residues.[14]

Basically, we[12,14] and others[13,15] have concluded that both compounds adopt primarily the exo-anomeric conformation around Φ angle (Φ H1′−C1′−x−C4, ∼60°, syn conformation), although we have also postulated that, in addition to this form, a local minimum around this angle does exist only for C-lactose (Φ H1′-C1′-x−C4, ∼180°, anti-Φ conformation, also dubbed gauche−gauche[16]), with a minor population around 5%. Regarding Ψ angle (Ψ, H4−C14−x−C1′), we established[17] for the first time that natural lactose is not monoconformational: most of the population appears to be located in the syn region (Ψ ∼0, >90%), while a 10% of the population is located in the anti-Ψ minimum (Ψ ∼0, <10%). According to our data and our analysis, this situation is altered for C-lactose as a consequence of the chemical change. In this case, apart of the distinct presence of the gauche−gauche conformer described above (Φ/Ψ, 180°:0°, with 5% population), the global minimum is shifted to the anti-Ψ region (Φ/Ψ, 60°:180°, ∼55% population) and the additional 40% of population is located in the syn region, with Φ/Ψ, 60°:0°. On the other hand, and on the basis of almost identical experimental NOE and J data, but using a qualitative analysis of these NMR parameters, it has been recently proposed that C- and O-lactose share the same conformational characteristics in the free state.[13] Finally, other authors[18] have analyzed their NMR off-resonance ROESY data in terms of a fixed Φ, two-state model for Ψ, then proposing a 60:40 syn/anti conformational equilibrium. However, neither a range of variation for Φ/Ψ angles within the syn and anti families nor the presence of the anti-Φ conformation was included in that analysis. Our analysis was performed on the basis of the exclusive[19] inter-residue NOEs that unequivocally characterize the syn, anti, and gauche−gauche regions of the conformational map (see below). The relationship between NOEs and proton−proton distances is well established[20] and can be worked out at least semiquantitatively, when a full matrix relaxation analysis is considered. In this case, it is obvious that the corresponding NOE intensities are sensitive to the respective conformer populations and that, therefore, an indication of the population distribution in free solution and on the galectin-bound conformation can be obtained by focusing on these key NOEs.

Indeed, and on this basis, we have previously shown that the conformations selected by ricin-B,[12] a toxic galactose-binding lectin, differ between O-lactose (syn) and C-lactose (anti-Ψ). Moreover, we have also demonstrated experimentally that the high-energy gauche−gauche or anti-Φ conformer of C-lactose (nondetectable for natural O-lactose and 5% populated for C-lactose) is selected by a hydrolytic enzyme, namely, *E. coli* β-galactosidase.[16] If the binding site architectures of proteins indeed drive the recognition of oligosaccharides, then flexible analogues such as C-lactose should also be expected to be bound by a galectin (i.e., galectin-1) in the syn conformation, despite its mobility in solution, on the basis of the contacts between the glucose unit and the binding site which have been predicted by X-ray crystallography.[4] Therefore, we here present TRNOE-basede NMR studies assisted by molecular modeling to address this issue, which has relevance for drug design. We unequivocally demonstrate that the syn conformation of C-lactose is the only one recognized by this lectin. In our opinion, this fact shows that the galectin-1 binding site is designed to select this conformation of lactose and lactose analogues through the establishment of key syn exclusive hydrogen bonds between the glucose residue and several amino acids. From a general point of view of C-glycoside flexibility, it is shown that the three major minima of C-lactose can be bound by different proteins through the establishment of key van der Waals and/or hydrogen bond interactions. Therefore, similar or even identical saccharides can thus be bound in different conformations depending on the protein binding sites architectures.

Results and Discussion

NMR Studies. The addition of galectin-1 to a D_2O solution of **1** induced broadening in the resonance signals of the ligand, especially on those corresponding to the galactose moiety (Figure 1). This fact is a clear indication of binding. Therefore, TRNOESY and TRROESY experiments (Figures 2 and 3d) were performed to deduce the bound conformation of **1**. For bound ligands that exchange with the free state at a fast rate, this experiment provides an adequate means to determine the conformation of the bound ligand.[7,9] In complexes of large molecules, cross relaxation rates of the bound compound are opposite in sign to those of the free ligand and produce negative NOEs.

TR−NOESY experiments (Figure 3d) produced strong and negative NOEs, as expected for ligand binding. These signals are the basis for attributing the properties of the bound state to any of the three different conformational families which coexist in water[12] (see also above). A single Φ/Ψ value is given for each family (Figure 4), although obviously rapid fluctuations

(11) Weatherman, R. V.; Kiessling, L. L. *J. Org. Chem.* **1996**, *61*, 534. (b) Weatherman, R. V.; Mortell, K. H.; Chervenak, M.; Kiessling, L. L. *Biochemistry* **1996**, *35*, 3619.

(12) Espinosa, J.-F.; Asensio, J. L.; Cañada, F. J.; Martin-Pastor, M.; Dietrich, H.; Martin-Lomas, M.; Schmidt, R. R.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **1996**, *118*, 10862−10871.

(13) Ravishankar, R.; Surolia, A.; Vijayan, M.; Lim, S.; Kishi, Y. *J. Am. Chem. Soc.* **1998**, *120*, 11297−11303.

(14) Martín-Pastor, M.; Espinosa, J. F.; Asensio, J. L.; Jimenez-Barbero, J. *Carbohydr. Res.* **1997**, *298*, 15−47.

(15) Wei, A.; Boy, K. M.; Kishi, Y. *J. Am. Chem. Soc.* **1995**, *117*, 9432, and references therein.

(16) Espinosa, J. F.; Montero, E.; Vian, A.; Garcia, J. L.; Dietrich, H.; Martin-Lomas, M.; Schmidt, R. R.; Cañada, F. J.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **1998**, *120*, 1309.

(17) Asensio, J. L.; Jiménez-Barbero, J. *Biopolymers* **1995**, *35*, 55.

(18) Rubinstenn, G.; Sinay, P.; Berthault, P. *J. Phys. Chem. A* **1997**, *101*, 2536.

(19) Dabrowski, J.; Kozar, T.; Grosskurth, H.; Nifant'ev, N. E. *J. Am. Chem. Soc.* **1995**, *117*, 5534.

(20) Neuhaus, D.; Williamson, M. P. *The nuclear Overhauser effect in structural and conformational analysis*; VCH: New York, 1988.

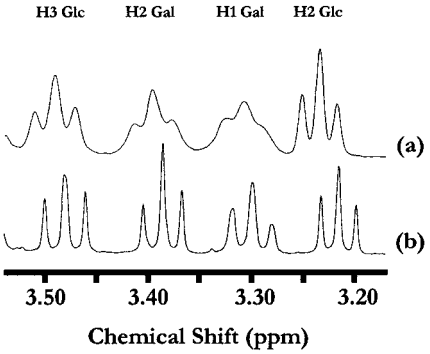


Figure 1. Partial section of the 500 MHz ^{1}H NMR spectrum recorded for free C-lactose at 303 K in D_2O in the presence (A) and absence (B) of bovine heart galectin-1 in 12:1 molar ratio. The broadening of all signals (especially those pertaining to the galactose moiety) may be observed.

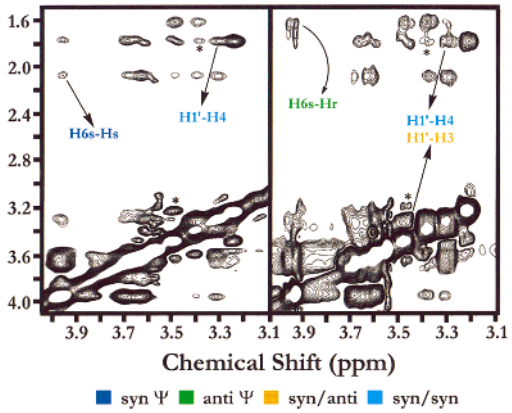


Figure 2. (Right) Expansion of the key regions of the 500 MHz ^{1}H NMR ROESY spectrum recorded for free C-lactose at 303 K (mixing time, 200 ms). Relevant cross-peaks are indicated. A strong H1′-H3 NOE indicates the major presence of the anti-Ψ conformer. The presence of the syn-Φ (H1′-H4) and anti-Φ (H2′-H4) conformers is also detected. (Left) Expansion of the key regions of the 500 MHz ^{1}H NMR ROESY spectrum recorded for C-lactose bound to galectin-1 molar ratio (12:1) under the same experimental conditions. Relevant cross-peaks are indicated. The anti-Ψ (H1′-H3) and anti-Φ (H2′-H4) cross-peaks have basically disappeared. The strong H1′-H4 NOE indicates the major presence of the syn-Φ,Ψ conformer. Spureous Hartmann−Hahn effects are noted with an asterisk.

are allowed around the different minima (about ±20°). The global minimum belongs to the so-called anti conformation (ϕ/ψ 36°:180°), where the glycosidic torsion angles ϕ and ψ are defined as H1′−C1′−Cα−C4 and C1′−Cα−C4−H4, respectively. Two more local minima are also detectable. The first one (destabilized in 3 kJ/mol) belongs to the syn conformation (ϕ/ψ 54°:18°), which is the major conformation reported by NMR for free[17] natural lactose and by NMR and X-ray for lectin-bound lactose and *N*-acetyllactosamine-containing oligosaccharides.[4,21] A high-energy local minimum (gauche−gauche or anti-ϕ, ϕ/ψ, 180°:0°), which is slightly populated (approximately 5%, destabilized in 9 kJ/mol), completes the conformational space. It represents a conformational family only observed in C-glycosides,[12,22,23] so far never detected for free $\beta(1 \rightarrow 4)$ natural disaccharides.

(21) Asensio, J. L.; Cañada, F. J.; Jimenez-Barbero, J. *Eur. J. Biochem.* **1995**, *233*, 618−630
(22) Espinosa, J. F.; Dietrich, H.; Martin-Lomas, M.; Schmidt, R. R.; Jiménez-Barbero, J. *Tetrahedron Lett.* **1996**, *37*, 1467.
(23) Espinosa, J. F.; Bruix, M.; Jarreton, O.; Skrystrup, T.; Beau, J. M.; Jimenez-Barbero, J. *Chem. Eur. J.* **1999**, *5*, 442.

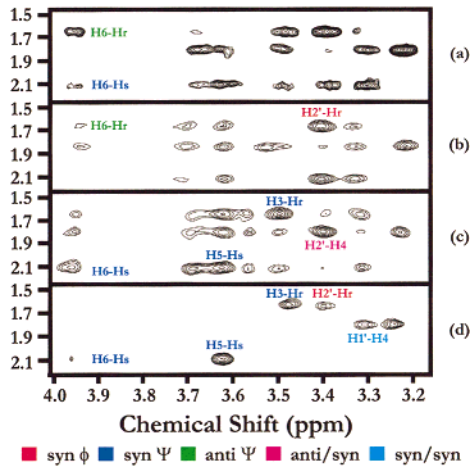


Figure 3. Expansion of the key region (with Glc H-4 and Hs,r protons at the methylene pseudoglycosidic linkage) of the 500 MHz ^{1}H NMR NOESY spectra recorded for C-lactose at 303 K under different conditions. (A) Free C-lactose[6] (mixing time, 700 ms). All of the cross-peaks for the three conformers are observed, including the H-1′/H-3 cross-peak in the other part of the spectrum (see Figure 2). (B) C-lactose/Ricin-B chain,[6] molar ratio 24:1 (mixing time 200 ms). The strongest peak in the spectrum (H-1′/H-3, anti-Ψ conformer) does not appear in this region.[6] (C) C-lactose/*E. coli* β-galactosidase,[16] molar ratio 46:1 (mixing time 150 ms). The strongest peak in the spectrum (H-2′/H-4, anti-Φ conformer) is evident.[16] (D) C-lactose/galectin-1, molar ratio 12:1 (mixing time 200 ms). The strongest peak in the spectrum (H-1′/H-4, syn-Φ,Ψ conformer) is evident.[16] Other weak interactions are observed due to spin diffusion effects.

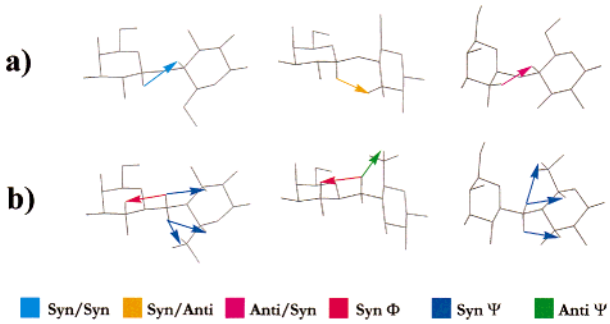


Figure 4. Stereoviews of the three conformers of C-lactose that are present in equilibrium in D_2O solution. (A) Exclusive interresidue NOEs and (B) contacts between the pyranose rings and methylene Hs,r protons that may be correlated with a major conformation at either linkage are indicated. Only the relevant protons are shown. From left to right, syn-Φ/syn-Ψ (galectin-1 case), syn-Φ/anti-Ψ (ricin case[6]), anti-Φ/syn-Ψ (β-galactosidase case[16]). All of the marked eight NOEs are observed for free C-lactose.

As mentioned above, there are exclusive NOEs that unequivocally characterize the syn, anti, and gauche−gauche regions of the conformational map (Figure 4). For C-lactose, these are H1′−H4, H1′−H3, and H4−H2′, respectively. In addition, the presence of two methylene protons for C-lactose, in contrast to the natural glycosides, may provide additional NOEs (Table 1, Figure 4) that can be correlated with a major orientation around either linkage (Φ or Ψ). In total, eight NOEs with conformational information are observed in water solution: these three exclusive NOEs and five more involving the interglycosidic methylene protons. Thus, since the corresponding NOE intensities will be sensitive to their respective populations, at least qualitatively, a first indication of the bound conformation can be obtained by focusing on these key NOEs. Different mixing times and protein/ligand molar ratios were systematically used.

Table 1. Interresidue Exclusive NOEs and Contacts between the Pyranose Rings and the Methylene Hs,r Protons That Define the Major Conformation at Either Linkage (A) Syn-Φ/Syn-Ψ, (B) Syn-Φ/Anti-Ψ, or (C) Anti-Φ/Syn-Ψ

Exclusive Interresidue NOEs: Conformer		
syn-Φ/syn-Ψ	syn-Φ/anti-Ψ	anti-Φ/syn-Ψ
Gal H-1-Glc H-4[a]	Gal H-1-Glc H-3[b]	Gal H-2-Glc H-4[c]
Key NOEs Involving the Methylene Hs,r Protons: Conformer		
syn-Φ	syn-Ψ	anti-Ψ
Gal H-2-H−r[a,b]	Glc H-3-H−r[a,c] Glc H-5-H−s[a,c] Glc H-6-H−s[a,c]	Glc H-6-H−r[b]

[a] Observed for Galectin-1/C-lactose complex. [b] Observed for Ricin/C-lactose complex.[6] [c] Observed for *E. coli* β-galactosidase/C-lactose complex.[16] All of these cross-peaks are observed for free C-lactose.

The comparison between the NOESY/ROESY spectra of C-lactose (Figures 2 and 3d) recorded in the absence and in the presence of galectin-1 shows important and clear differences. Some of the cross-peaks in the NOESY spectrum of the free ligand are no longer displayed in the TR−NOESY spectrum of the complex. It is important to stress the disappearance of both H1′−H3 and H2′−H4 NOEs (intensity smaller than 0.5% of their diagonal peaks). This result provides evidence that neither the anti conformation nor the gauche−gauche conformation is recognized by the lectin. In contrast, the H4−H1′ NOE that displayed a medium intensity for the free ligand corresponds now to the strongest interresidual contact in the spectrum (with intensity of 8.1% of the diagonal peak). These findings indicate that the major bound conformation belongs to the syn family. Four additional cross-peaks involving the methylene protons also point to the recognition of the syn conformation (Figures 2, 3d, and 4). In particular, the presence of a major conformer with a syn-Φ angle is characterized by a H-2′/Hr (4.4%) contact, while the presence of a syn-Ψ torsion is evidenced by H-3/Hr (8.8%), H-5/Hs (9.6%), and H-6/Hs (2.1%) NOEs. TR−ROESY experiments[17] also permitted the exclusion of spin diffusion effects. Indeed, the above-mentioned cross-peaks showed a different sign relative to the diagonal peaks, thus excluding the possibility of protein-relayed or spin-diffusion-mediated correlations. A drawback of the TRNOE method is that it is not directly able to separate NOEs when more than one conformation is present. However, the use of a full matrix relaxation approach,[20] including exchange between the free and bound forms, may be used to estimate the expected NOEs for the binding of the other possible anti-Ψ and anti-Φ families. Thus, TRNOE calculations for different ensemble average conformations were performed and compared to the experimental data. Fortunately, the very small H1′−H3 (2.0 Å) and H2′−H4 (2.3 Å) distances for the corresponding anti-Ψ and anti-Φ would give rise to detectable H1′−H3 and H2′−H4 TRNOE cross-peaks (above 1% intensity of the diagonal peak), if populated above 5−10%. Thus, the full matrix relaxation calculations indicate that the presence of the anti-Ψ and anti-Φ families above a 5−10% of population can be safely excluded. Otherwise, the corresponding cross-peaks would be effectively detected. Of course, molecular motion within the valley corresponding to this local syn-Φ minimum could still be maintained.[12] The existence of specific binding was deduced from competitive TRNOE experiments[12] in which the corresponding O-glycoside, methyl-β-lactoside, was added to the NMR tube containing the C-lactose/galectin-1 solution. It was observed that, at equimolar ratio between the C/O-glycoside ligands, the cross-peaks corresponding to C-lactose changed their sign to positive, while those pertaining to the O-lactoside appear as negative, indicating that both ligands compete for the same binding site and that the affinity for lactose is higher than for the C-analogue.

The given data unquestionably lead to the conclusion that C-lactose is bound by bovine galectin-1 in a major conformation that is different from the global minimum of this ligand analogue. Inspection of the crystallographic data for this galectin in complex with *N*-acetyllactosamine or lactose[4] and NMR data with an avian galectin[8] unveils the resemblance of the galectin-1-bound conformation of the C-glycoside to the low-energy syn conformations of O-lactose and related O-glycosides (ϕ from 33° to 63°/ψ from 7° to 18°, according to the X-ray data).

In this case, the balance between entropic and enthalpic factors has guided the ligand into a preferred syn-Φ conformation to yield the optimal $\Delta G°$ value. The situation is different for C-lactose binding to ricin-B[12,24] and β-galactosidase.[16] A comparison among the key regions of the corresponding NOESY spectra recorded for C-lactose in different conditions (free, with ricin-B, with *E. coli* β-galactosidase, and with galectin-1) is given in Figure 3. Obviously, since the different proteins have different sizes and therefore very distinct correlation times, different concentrations, protein/sugar ratios, and mixing times were used to get the best TRNOE results for every case. Nevertheless, the difference in relative intensities of the exclusive NOEs expected for the different bound conformers may be appreciated (see also Table 1). For the galectin-1 case, taking into account the energetic differences in solution between the syn and the global minimum anti conformer, which amounts to about 3 kJ/mol, it can be assumed that the distortion will be easily compensated at least by an enthalpic gain, since further consideration of entropic factors involving solvent molecules is beyond the present reach.

Then, to visualize how the C-glycoside can be placed into galectin-1's binding site, we performed docking studies on the template of X-ray data sets for this galectin.[4]

Docking Studies. This galectin is a homodimer, and each of the two monomers possess almost identical binding features. Without the existing X-ray information for other analogous complexes, it is clear that the docking study would be a major project in itself. Nevertheless, since the TRNOE method has allowed the demonstration that C- and O-lactose compete for the same binding site, the X-ray template of the published structure was considered. For the first subunit, the binding site is located between amino acids His-44, Asn-46, and Arg-48, which provide the specificity for galactose residues, and Trp-68, which provides additional stacking with H-1, H-3, H-4, and H-5 of the galactose residue. The second subunit displays the same type of interactions. Therefore, only one subunit was further considered for energy minimizations. Since the final structure is largely defined by the starting conformation, different conformations of the pseudodisaccharide were used as input. In all starting structures, the galactose moiety of **1** was manually docked within the receptor's binding site to mimic its orientation in the X-ray structure.[4] Then, different conformations of C-lactose were generated and the resulting protein−sugar complex was surrounded by a sphere containing 863 water molecules (25 Å, see experimental). The calculations of the complexes yielded different orientations. The obtained results for the anti-Ψ and syn-Φ/Ψ starting conformations are shown in Figure 5. Independently of the starting structure, minor movements of the galactose ring were observed during the

(24) Espinosa, J. F.; Asensio, J. L.; Cañada, F. J.; Dietrich, H.; Martin-Lomas, M.; Schmidt, R. R.; Jimenez-Barbero, J. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 303−306.

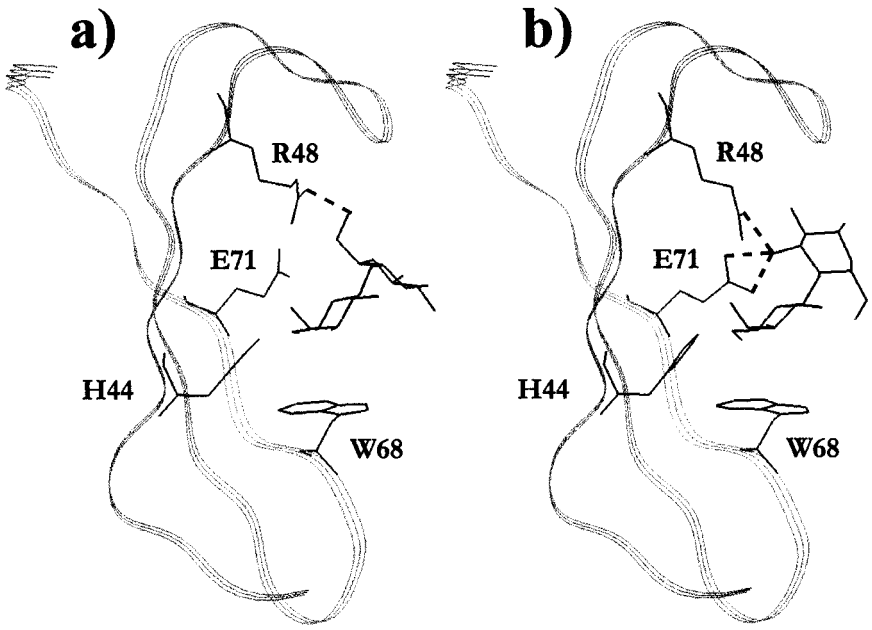


Figure 5. Stereoview of the calculated three-dimensional structures of galectin-1 complexed with the two more stable conformers of C-lactose after docking studies (A) with the anti-Ψ conformer and (B) with the syn-Ψ conformer. The relevant intermolecular hydrogen bonds and stacking interactions are observable.

energy mimimization and the polypeptide chain remained fairly close to the crystallographic 3D structure. In all cases, no additional forces were included to keep the ligand within the binding site to try to minimize bias. Several intermolecular hydrogen bonds are established between the saccharide and the protein, which have been shown to be classical in protein−carbohydrate interactions.[25] Nevertheless, important differences between the two more stable complexes are evident from the docking studies: In the syn conformation, three hydrogen bonds are formed between O-3 of the Glc residue: two with Glu-71 (2.5 Å) and one with Arg-48 (2.6 Å). In contrast, for the anti-Ψ conformation (Figure 5a), only one hydrogen bond between Arg-48 and Glc O-6 is possible (2.7 Å). The stabilizing interactions with Glc O-3 are not formed for the anti-Ψ conformation, since the involved atoms point now into different spatial orientations. Although only qualitative, the enthalpic difference between both complexes, according to the AMBER force field, amounts to more than 100 kJ/mol, favoring the complex with the syn conformer. Since this value includes the different solvation of both complexes, additional minimizations were carried out without solvent using a distance-dependent dielectric constant ($4r$). In this case, the syn complex was also favored more than 2.5 kcal/mol. In addition, the freezing of the Glc hydroxymethyl group in a particular rotamer would also engender an entropic penalty. For both cases, the hydrogen bond and van der Waals interactions mentioned above, which provide the galactose specificity, are kept. Notably, interactions of Trp rings with apolar carbohydrate faces frequently occur in protein/carbohydrate complexes, both in the solid state[25] and in solution.[9,26] By using the laser photo CIDNP approach, we have convincingly documented the importance of the surface-exposed and binding-site located Trp-residue for galectins.[27] Indeed, in both models shown here, as in the X-ray structures for the natural compounds,[4] the galactose methine hydrogen atoms H1, H3, and H5 create a flat hydrophobic surface, which leads to excellent stacking interactions with Trp-68.

Therefore, we can infer that the binding of the syn conformer with the structure shown in Figure 5b represents the most likely binding mode of C-lactose by galectin-1. It is not only in agreement with the presented NMR data. It also is in accordance with the galectin-1/O- lactosides X-ray data, with ϕ angles from 33° to 63° and ψ angles from 7° to 18°.

Conclusions

The same conformation (syn) of O- and C-lactose is recognized by bovine heart galectin-1. Since both compounds present different conformational population distributions in water solution, this observation represents a case of conformer selection by a lectin's binding site. Evidently, the topological features of the protein binding sites may restrict ligand mobility, then shifting the conformational equilibrium of the flexible C-glycoside. In principle, if a protein establishes interactions to several sugar units within a given oligosaccharide, then only certain favored conformers will snugly fit into the binding site. Alternatively, the possibilities still exist that the conformational equilibrium could be maintained, if the entropic penalty to freeze it exceeds the enthalpic gain when weak sugar−protein interactions are established. Also the global minimum conformation could reach an optimal $\Delta G°$ value, as seen in the case of ricin.[12] Along this reasoning, the recently published 2.7 Å resolution data on a C-lactoside/peanut agglutinin crystal[13] could readily be reconciled with the current status of interpretation. In this particular case, two hydrogen bonds exist between glucose O-3 and Ser 211 and glucose O-3 and Gly 213, which are only possible if the C-lactose moiety adopts the syn conformation.[13] In fact, these interactions between Glc O-3 of the syn conformer of **1** and this agglutinin[13] are identical to those described above for galectin-1. Therefore, for both galectin-1 and peanut agglutinin,[13] despite topological differences and no homology,

(25) Rini, J. M. *Curr. Opin. Struct. Biol*. **1995**, *5*, 617−21. (b) Weis, W. I.; Drickamer, K. *Annu. Rev. Biochem*. **1996**, *65*, 441−473. (c) Lis, H.; Sharon, N. *Chem. Rev*. **1998**, *98*, 637−674. (d) Loris, R.; Hamelryck, T.; Bouckaert, J.; Wyns, L. *Biochim. Biophys. Acta* **1998**, *1383*, 9−36.

(26) Asensio, J. L.; Cañada, F. J.; Bruix, M.; Khiar, N.; Gonzalez, C.; Rodríguez-Romero, A.; Jiménez-Barbero, J. *Glycobiology* **1998**, *8*, 621−633.

(27) Siebert, H.-C.; von der Lieth, C.-W.; Kaptein, R.; Beintema, J. J.; Dijkstra, K.; van Nuland N.; Soedjanaatmadja, U. M.; Rice, A.; Vliegenthart, J. F. G.; Wright, C. S.; Gabius H.-J. *Proteins* **1997**, *28*, 268.

the binding site topology does not allow the global energy minimum of the C-glycoside to be tightly accommodated, thus changing the torsion angles their low-energy position to a local minimum conformer. Nevertheless, the distortion is here limited to accommodating one conformation into the binding site, which is also populated in solution. No drastic rotation of a glycosidic angle is required, as seen for pentasaccharide binding to ConA.[28]

From a general point of view, and in relation with other results for different protein-bound saccharides, there are cases in which proteins bind oligosaccharides near their global minimum conformation, although there are also examples of either major or local conformational variations upon binding.[9,29] In fact, the anti-Ψ and anti-Φ (gauche−gauche) conformations of C-lactose are selected by ricin-B (a galactose-binding lectin) and *E. coli* β-galactosidase, respectively (Figures 2−4). Ricin-B differs substantially from galectin-1 in both the folding topology and the shape of the binding site. When probed at physiological pH, only a weak hydrogen bond of the C-6−OH of Glc is possible for ricin.[30] Overall, the energetically preferred anti conformer of C-lactose is well tolerated for binding.[12] *E coli* β-galactosidase can even bind the high-energy conformer (destabilized in 9 kJ/mol).[16] We have speculated that the recognition of this high-energy conformation of C-lactose (anti-Φ or gauche−gauche) could have implications for the catalytic mechanism, lowering the energy barrier necessary for hydrolysis. Obviously, no comparison with the natural compound (O-lactose, **2**) is possible, since this is readily hydrolyzed.

As summary, the description of the cases described herein with those reported for Ricin-B and *E coli* β-galactosidase indicates that the three conformational families of C-lactose can be selected by different sugar receptors and that the formation of distinct van der Waals and/or hydrogen bond interactions between the binding sites and the pseudodisaccharide atoms can drive the selection of a particular conformer. The same oligosaccharide can thus be bound in different conformations, depending on the protein binding site.[8,9,27,31] A similar situation is encountered even in related lectins, as documented for the selectin subfamily of C-type lectins and the sLex tetrasaccharide.[32] From the NMR point of view, the study of carbohydrate−lectin interactions is an area where the TR−NOE methodology, despite the above-mentioned drawbacks, and as pioneered by Prestegard, has a remarkable applicability. The reason for this favorable situation probably rests in different facts: these interactions are not extremely strong, there is fast exchange between the free and the bound states of the ligand, and the perturbations of the conformational equilibrium of a given oligosaccharide upon binding to a protein are accessible to observation by TRNOE.

Methods

Source of the Lectin. The galactose-specific lectin from bovine heart (galectin-1) was purified by affinity chromatography on lactosylated Sepharose 4B, obtained after divinyl sulfone activation as a decisive step and checked for purity by one- and two-dimensional gel electrophoresis and for activity after processing by solid-phase assays with (neo)glycoproteins, as described.[33]

Molecular Modeling. Protein coordinates and glycosidic torsion angles for the syn conformer were taken from the published crystal structures[4] of different galectin-1/oligosaccharide complexes (PDB codes for bovine 1SLA, 1SLB, 1SLC, 1SLT and bufo galectins 1A78 and 1GAN). Glycosidic torsion angles for other complexes were also obtained for galectin-2 (1HLC) and galectin-3 (1A3K). C-lactose was built using the biopolymer module within the INSIGHTII program. Atomic charges were AMBER charges. The starting orientation of the galactose residue was chosen to match that of the crystal structure. A region close to the protein's recognition site was considered that involved all of the amino acid residues from Asn-40 to Phe-79. All energy calculations were done using the AMBER force field.[34] The complexes were inmersed into a sphere of 863 water molecules. The sphere was centered on the Gal moiety to perfectly solvate the interface and the binding site. The presence of water molecules was essential to keep the sugar within the binding site. A template force potential was introduced to avoid major movements of the polypeptide backbone during the calculations. The pseudodisaccharide and the amino acid lateral chains were left free during the minimization processes. No cutoffs for nonbonding interactions were used. The three major conformers, syn-Φ, anti-Ψ, and anti-Φ, were generated with two initial Φ and Ψ values. Energy minimizations were then conducted on the six complexes using 2000 conjugate gradient iterations. The anti-Φ conformer generated important steric conflicts with the polypeptide chain and gave rise to a final syn-Φ conformation.

NMR Experiments. NMR spectra were recorded at 30 °C in D_2O, on a Varian Unity exposed to repeated cycles of freeze-drying with D_2O, and transferred to the NMR tube to give a final concentration of 0.037 mM. TR−NOESY experiments were performed with mixing times of 200 and 300 ms for 12:1 and 24:1 molar ratios of β-methyl-C-lactoside/lectin. In all cases, line broadening of the sugar protons was monitored after the addition of the ligand. The theoretical analysis of the TRNOEs was performed according to the protocole employed by London,[35] using a full relaxation matrix with exchange as described.[12] Different exchange-rate constants, k, defined as $pfk = K_{-1}$ (where pf is the fraction of the free ligand) and leakage relaxation times were employed to get the best match between experimental and theoretical results of the intraresidue H-1′/H-3′, H-1′/H-5′, H-1/H-3, and H-1/H-5 cross-peaks for the given protein/ligand ratio. Normalized intensity values were used since they allow correction for spin relaxation effects. The overall correlation time τ_c for the free state was always set to 0.15 ns,[6] and the τ_c for the bound state was estimated as 30 ns according to the molecular weight of the lectin ($\tau_c = 10^{-12}$ W_M). The association constant was approximated as half of that calculated for lactose.[3b,c] To fit the experimental TRNOE intensities, exchange-rate constants between 100 and 1000 s^{-1} and external relaxation times ρ^* for the bound state of 0.5, 1, and 2 s were tested. The best agreement was achieved when using $k = 200$ s^{-1} and $\rho^* = 1$ s.

TR−ROESY experiments were also carried out to exclude spin diffusion effects. A continuous wave spin lock pulse was used during the 250 ms mixing time. Key NOEs were shown to be direct cross-peaks, since they showed different sign to diagonal peaks.

Acknowledgment. We thank DGICYT (PB96-0837), TMR(FMRX-CT98-0231), and CARENET-2 (EU) for financial support. J.L.A. thanks CAM and MEC for postdoctoral fellowships. Dedicated to Prof. Dr. J. Elguero on occasion of his 65 birthday.

JA990601U

(28) Moothoo, D. N.; Naismith, J. *Glycobiology* **1998**, *8*, 173−181.
(29) Bundle, D. R.; Young, N. M. *Curr. Opin. Struct. Biol.* **1992**, *2*, 666. (b) Imberty, A. *Curr. Opin. Struct. Biol.* **1997**, *7*, 617. (c) Milton, M. J.; Bundle, D. R. *J. Am. Chem. Soc.* **1998**, *120*, 10547.
(30) Solis, D.; Fernandez, P.; Diaz-Mauriño, T.; Jimenez-Barbero, J.; Martin-Lomas, M. *Eur. J. Biochem.* **1993**, *214*, 677−683.
(31) Gilleron, M.; Siebert, H.-C.; Kaltner, H.; von der Lieth, C.-W.; Kozar, T.; Halkes, K. M.; Korchagina, E. Y.; Bovin, N. V.; Gabius, H.-J.; Vliegenthart, J. F. G. *Eur. J. Biochem.* **1998**, *249*, 27−38.
(32) Scheffler, K.; Brisson, J. R.; Weisemann, R.; Magnani, J. L.; Wong, W. T.; Ernst, B.; Peters, T. *J. Biomol. NMR* **1997**, *9*, 423, and references therein. (b) Poppe, L.; Brown, G. S.; Philo, J. S.; Nikrad, P. V.; Shah, B. H. *J. Am. Chem. Soc.* **1997**, *119*, 1727−1736.
(33) Gabius, H.-J.; Engelhardt, R.; Rehm, S.; Cramer, F. *J. Natl. Cancer Inst.* **1984**, *73*, 1349−1357. (b) Kohnke-Godt; Gabius, H.-J. *Biochemistry* **1989**, *28*, 6351−6358. (c) Andre, S.; Unverzagt, C.; Kojima, S.; Dong, X.; Fink, C.; Kayser, K.; Gabius, H.-J. *Bioconjugate Chem.* **1997**, *8*, 845−855.
(34) Homans, S. *Biochemistry* **1990**, *29*, 9110.
(35) London, R. E.; Pearlman, M. E.; Davis, D. G. *J. Magn. Reson.* **1992**, *97*, 79−98.