BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Molecular Recognition of the Thomsen-Friedenreich Antigen−Threonine Conjugate by Adhesion/Growth Regulatory Galectin-3: Nuclear Magnetic Resonance Studies and Molecular Dynamics Simulations

Austin B. Yongye,[†] Luis Calle,[‡] Ana Ardá,[‡] Jesús Jiménez-Barbero,[‡] Sabine André,[§] Hans-Joachim Gabius,[§] Karina Martínez-Mayorga,*[,†] and Mare Cudic*[,†]

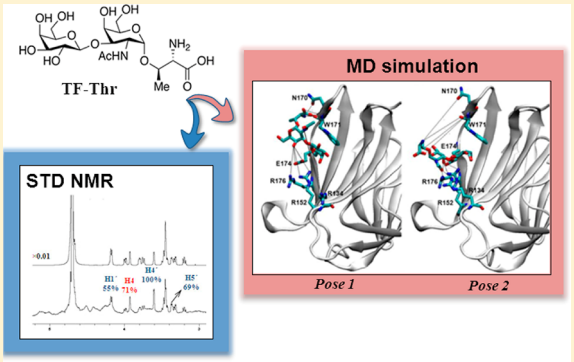
[†]Torrey Pines Institute for Molecular Studies, Port Saint Lucie, Florida 34987-2352, United States

[‡]Chemical and Physical Biology, CIB-CSIC, Madrid, Spain

[§]Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University, Munich, Germany

Supporting Information

ABSTRACT: Nuclear magnetic resonance (NMR) spectroscopy and molecular modeling methods have been strategically combined to elucidate the molecular recognition features of the binding of threonine O-linked Thomsen-Friedenreich (TF) antigen to chimera-type avian galectin-3 (CG-3). Saturation transfer difference (STD) NMR experiments revealed the highest intensities for the H4 protons of both the β-D-Gal*p* and α-D-Gal*p*NAc moieties, with 100 and 71% of relative STD, respectively. The methyl protons of the threonine residue exhibited a small STD effect, <15%, indicating that the interaction of the amino acid with the protein is rather transient. Two-dimensional transferred nuclear Overhauser effect spectroscopy NMR experiments and molecular modeling suggested some differences in conformer populations between the free and bound states. A dynamic binding mode for the TF antigen−CG-3 complex consisting of two poses has been deduced. In one pose, intermolecular interactions were formed between the terminal threonine residue and the receptor. In the second pose, intermolecular interactions involved the internal Gal*p*NAc. The difference in the trend of some shifts in the heteronuclear single-quantum coherence titration spectra indicates some disparities in the binding interactions of CG-3 with lactose and TF antigen. The results obtained from this model of the avian orthologue of human galectin-3 will allow detailed interspecies comparison to give sequence deviations in phylogeny a structural and functional meaning. Moreover, the results indicate that the peptide scaffold presenting TF antigen could be relevant for binding and thus provides a possible route for the design of galectin-3 inhibitors with improved affinity and selectivity.



The growing awareness that the glycan component of cellular glycoconjugates stores biological information, embodied by the term "sugar code",[1] has attracted an increasing amount of attention to the analysis of glycan−lectin interactions. In fact, these sugar receptors provide a route to translate the sugar-encoded information into cellular responses.[2−6] Homing in on branch-end epitopes, the family of galactoside-binding lectins (galectins) elicits a wide range of biological functions involved in immune and inflammatory responses and tumor development, including cell adhesion, migration, and survival.[7−12] This broad activity profile explains the emerging interest in galectins as targets for drug discovery. In principle, galectin-dependent responses are triggered on the basis of the selection of the glycan counter-receptor and the ensuing cross-linking of the selected glycan moiety. While homodimeric and tandem repeat-type family members are natural cross-linking agents, galectin-3 is a unique type. It is monomeric and known to harbor the carbohydrate recognition domain (CRD) in the C-terminal region, a collagenase-sensitive section, and the N-terminal region containing sites for serine phosphorylation.[2,10,13,14] In the presence of suitable multivalent ligands, galectin-3 is known to form pentamers and is thus capable of cross-linking.[15] This tendency to oligomerize has been attributed to the N-terminal non-lectin domain of galectin-3. Of note is the fact that the functional antagonism of galectin-3 to pro-apoptotic galectin-1, blocking negative growth regulation in neuroblastoma and pancreatic carcinoma cells,[16,17] makes structural studies of this protein clinically relevant. Equally importantly, phylogenetic analysis has revealed that despite diversity in the other two classes,

Received: June 7, 2012
Revised: August 22, 2012
Published: August 23, 2012

 © 2012 American Chemical Society dx.doi.org/10.1021/bi300761s | *Biochemistry* 2012, 51, 7278−7289

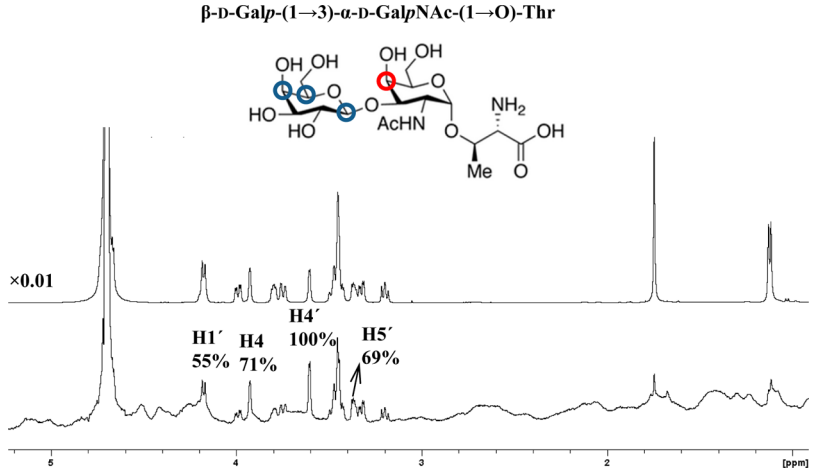


Figure 1. ^{1}H NMR (500 MHz) STD spectrum (saturation time of 2 s with on-resonance irradiation at the aromatic region) of a 1:100 mixture of CG-3 and the TF antigen at 278 K. The bottom trace shows the STD spectrum, while the top trace shows the off-resonance spectrum. The protons showing the strongest STD effect are annotated.

galectin-3 is the only member of the galectin group with a trimodular design.[12,18] Thus, galectin-3 presents itself as a unique model for comparative structure−activity studies between different organisms, with the aim of giving sequence deviations in phylogenesis a functional meaning. With these two perspectives in mind, the therapeutic and the phylogenetic, we here initiate a strategically combined analysis of glycan binding, using experimental and computational approaches. On the glycan side, the focus is on the Thomson-Friedenreich (TF) antigen, the core 1 *O*-glycan disaccharide, Galβ1→3GalNAcα, also known as T-antigen. It has selectivity for galectin-3 binding compared to galectin-1 and is known to account for binding of galectin-3 to mucins, and also the pentasaccharide of ganglioside GM1.[19−25] Structural analysis of the complex by crystallography and flexible docking has been reported for human galectin-3.[24,25] Because its avian orthologue, i.e., chicken galectin-3 (CG-3), has recently been identified,[26] a comparative analysis is possible. Moreover, the TF binding mode can be compared to the complex of this glycan with a fungal galectin from the edible mushroom *Agrocybe aegerita*.[27] It is also intriguing to compare if and how the dynamics and flexibility of the full-length galectin-3 is affected by its interaction with the TF antigen in solution. While X-ray crystal structures provide static representations of molecules arranged in a periodic crystal lattice, solution NMR data provide average information over randomly oriented molecules in solution. The NMR spectroscopy and molecular modeling methods have been an integral component of conformational analysis of carbohydrates, both free and bound to lectin.[28] To assess the potential of the peptide portion to be located in the direct vicinity of the core 1 disaccharide in complex formation with CG-3, a threonine residue attached to the glycan moiety was used in our study. Likewise, previous monitoring of glycopeptides and glycan derivative libraries has indicated a possibility for peptide−aglycone interactions.[29−31]

MATERIALS AND METHODS

Galectin Purification. The chimera-type avian galectin-3 was obtained by recombinant production and isolated by affinity chromatography on lactosylated Sepharose 4B as a crucial step, routinely checked for purity by one- and two-dimensional gel electrophoresis and gel filtration.[26,32] Activity was ascertained by heme agglutination and solid-phase binding assays.[26,32,33] ^{15}N-labeled CG-3 was obtained using $^{15}NH_4Cl$ as the medium additive as described previously.[34]

NMR Experiments. *O*-[2-(Acetamido)-2-deoxy-3-*O*-β-D-galactopyranosyl-α-D-galactopyranosyl]-L-threonine (TF antigen−Thr conjugate) was purchased from Toronto Research Chemicals. The NMR samples were prepared in deuterated phosphate buffer [10 mM phosphate, 137 mM NaCl, and 2.7 mM KCl (pH 7.3)]. STD experiments[35] were performed at a lectin concentration of 50−60 μM on a Bruker Avance 500 MHz NMR spectrometer equipped with a 5 mm TXI inverse probe head at 278 K. Protein saturation (on the resonance frequency in the aromatic region) was achieved by using a series of 50 ms Gaussian pulses with a total saturation time of the protein of 2 s and a maximal field strength of 50 Hz. The ligand:lectin ratio was set between 20:1 and 100:1. TR-NOESY experiments were conducted on a Bruker Avance 600 MHz spectrometer equipped with a triple-channel cryoprobe at 278 K. Different mixing times were employed: 100, 150, and 200 ms. These experiments were conducted for a ligand:protein molar ratio of 10:1 using a 60 μM preparation of lectin as the final concentration in the NMR tube. No purging spin-lock period to remove the NMR signals of the macromolecule background was employed. The nOes were translated into distances using the isolated spin pair approximation and the intraresidual H1′−H5′ (Gal*p*) and H1−H2 (Gal*p*NAc) distances as a reference. Uncertainties in the distances between different mixing times were less than 6%.

Molecular Modeling. The threonine conjugate of the TF antigen and lactose were generated using the GLYCAM web carbohydrate builder[36] and the tleap module distributed with ANTECHAMBER version 1.27.[37] The GLYCAM06 (Glycam_06g.dat downloaded from the Woods group Web site)[38] and PARM99SB[39] parameter sets were employed to model the ligand and receptor, respectively. The threonine residue of the conjugate was in the zwitterionic state. Automated docking of the TF antigen to CG-3 was performed using the GlideXP version 5.7 scoring function of Schrödinger, Inc.[40] The three-dimensional model for chicken galectin-3 was generated through the SWISS-MODEL repository database[41,42] from the sequence A4GTP0 (UniProt database) and by using Protein Data Bank (PDB) entry 1A3K (carbohydrate

dx.doi.org/10.1021/bi300761s | *Biochemistry* 2012, 51, 7278−7289

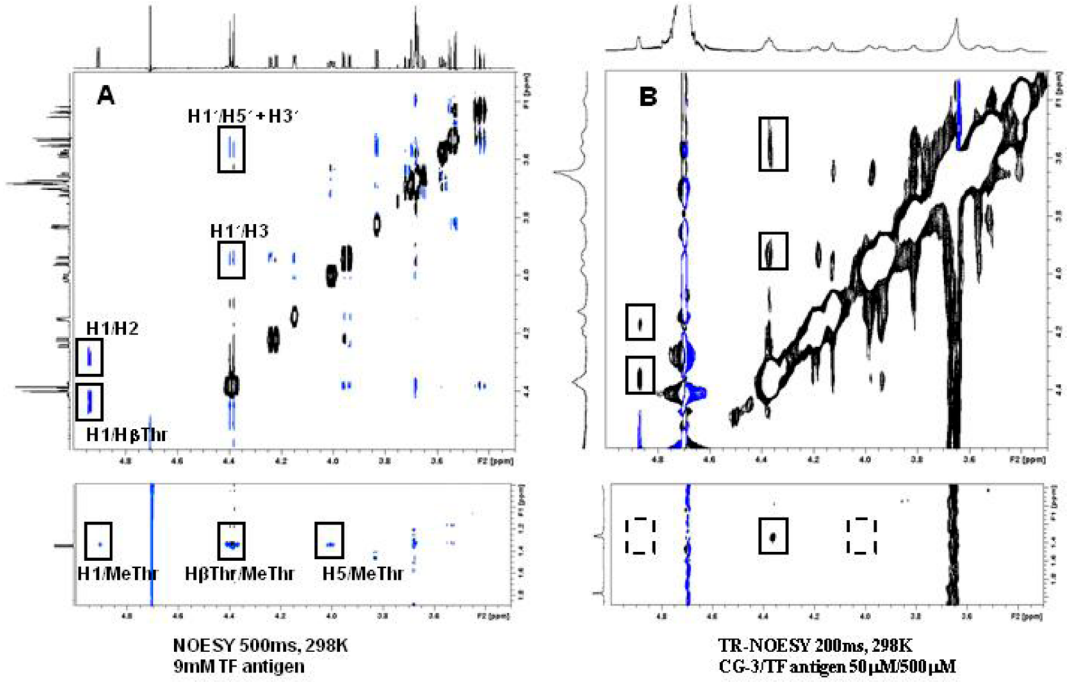


Figure 2. (A) NOESY spectrum (500 ms) for the TF antigen. (B) TR-NOESY spectrum (200 ms mixing time) for the TF antigen (50 μM) bound to CG-3 (500 μM). Insets are shown at the bottom.

recognition domain of human galectin-3) as a template. The NMR structure of CG-3 is still to be completed; however, all the available data are in agreement with the employed model. A survey of 38 galectin−carbohydrate complexes (Table S1 of the Supporting Information) was performed to assess the most frequently occurring position of the terminal galactosyl residue during binding. On the basis of this analysis, manual docking was conducted by superimposing the β-D-Gal*p* residue of the TF antigen with that of *N*-acetyllactosamine in the crystal structure of human galectin-3 in complex with this ligand (PDB entry 1A3K), which was subsequently energy minimized. Details of the docking, molecular dynamics simulations, and postprocessing of trajectories are provided as Supporting Information.

RESULTS AND DISCUSSION

The interaction of CG-3 (tested as a CRD mimicking the natural form after proteolytic truncation) with the Thr−TF antigen conjugate was studied by NMR spectroscopy and molecular modeling. Combining these two approaches afforded the structural and conformational characterization of the TF antigen in the free and in lectin-bound states. In addition, the interacting epitope of the TF antigen was identified, and the mobility of CG-3 when accommodating the TF antigen and lactose was determined and compared. This comparison allowed us to propose the mode of recognition of the TF antigen by the galectin.

Binding of the TF Antigen by CG-3. STD experiments provide key information about the regions of the ligand that interacts with the protein. The STD spectrum of the TF antigen bound to CG-3 is shown in Figure 1. The ^{1}H and ^{13}C NMR assignments of the ligand resonances are summarized in Figure S1 of the Supporting Information. The protons that present the highest STD intensities are both H4 of β-D-Gal*p* and α-D-Gal*p*NAc, with 100 and 71% of the relative STD, respectively. The next most intense STD effects correspond to H5′ and H1′ of β-D-Gal*p*. The rest of the protons present less than 50% of the relative STD. This pattern is in contrast with the STD pattern usually observed for the recognition of lactose by galectins.[43] In this well-studied case, H4, H5, and H6 of the terminal galactose residue receive most of the saturation from the galectin, while the rest of protons are poorly saturated. The protons of the Thr residue are overlapping with the protons of the sugar moiety, but the methyl group of the Thr shows a small STD effect, <15%, indicating that the interaction of the amino acid with the protein is rather transient.

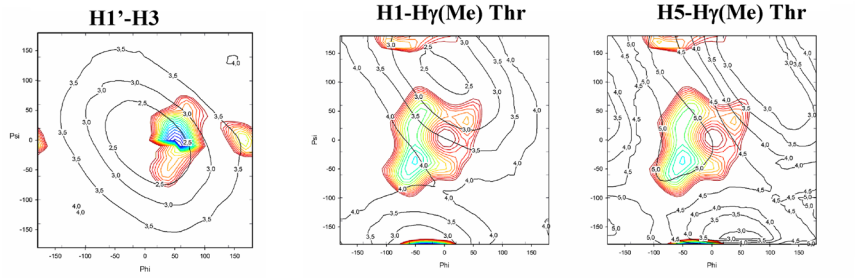
Conformation of the TF Antigen in the Free and Lectin-Bound States. To determine the conformational preference of the TF antigen in the bound and free states, we present an evaluation of proton nuclear Overhauser enhancements (nOes). Typically, nOes are observed for proton pairs within a maximal distance of 4.0 Å. Therefore, changes in the presence or intensity of nOes are evidence of variations in the relative spatial orientation of specific proton pairs, which can be related to specific dihedral angles. The comparison of the nOe signals from NOESY (free) and TR-NOESY (bound) was used to identify changes in the conformational preference of the TF antigen from the free to the bound state. The left-hand side of Figure 2 shows the NOESY spectrum (free state), whereas the right-hand side shows the TR-NOESY spectra (bound state). Insets included in the spectra specify the key nOes. The observed intraresidue nOe between H1′ and H5′ of the β-D-Gal*p* unit as well as the absence of an H1−H5 nOe or a strong H1−H2 nOe of the α-D-Gal*p*NAc moiety permitted assessment of the typical low-energy chair conformation. These nOes were maintained when the ligand was bound to CG-3. In addition, the inter-residue H1′−H3 nOe strongly suggested the presence of the *exo* anomeric orientation around the $\beta(1\rightarrow3)$ glycosidic bond, with a ϕ of 60°, in both the free and bound states. With regard to the α-D-Gal*p*NAc−Thr linkage, in the free state the H1−Hβ Thr nOe was stronger than the intraresidue H1−H2 nOe, indicating a very close contact

dx.doi.org/10.1021/bi300761s | *Biochemistry* 2012, 51, 7278−7289

Table 1. Relevant nOes (percent) Derived from the Experimental Data and Corresponding Distances[a]

	β-D-Galp	β-D-Galp-(1→3)-α-D-GalpNAc	α-D-GalpNAc-(1→O)-Thr		
	H1′−H5′	H1′−H3	H1−Hβ Thr	H5−Hγ(Me) Thr	H1−Hγ(Me) Thr
NOESY[b]	1% (2.5 Å)	0.6% (2.7 Å)	1.1% (2.4 Å)	0.5% (2.7 Å)	0.1% (3.5 Å)
TR-NOESY[b]	1% (2.5 Å)	1.4% (2.4 Å)	1.2% (2.3 Å)	no nOe	no nOe
unbound model 1, α-D-GalpNAc-OThr ($\phi_\alpha = -43°$; $\psi_\alpha = -150°$)	2.5 Å	2.6 Å	3.6 Å	5.3 Å	2.6 Å
unbound model 2, α-D-GalpNAc-OThr ($\phi_\alpha = -43°$; $\psi_\alpha = -20°$)	2.5 Å	2.6 Å	2.5 Å	2.9 Å	4.4 Å
manualDock					
Figure 4A	2.5 Å	2.5 Å	2.6 Å	3.3 Å	4.6 Å
best GlideXP pose in Figure 4B	2.4 Å	2.4 Å	3.1 Å	3.1 Å	4.3 Å

[a]Distances from molecular models at the defined ϕ and ψ angles (unbound models 1 and 2) are also included. Distances derived from automated docking structures are listed at the bottom (bound models 1 and 2). [b]nOe enhancements are shown, along with the derived distances in parentheses. nOes were translated into distances using the isolated spin pair approximation and the intraresidual H1−H5 distance for the Gal residue and H1−H2 distance for GalNAc as references.



	ϕ_β(°) (H1′-C1′-O3-C3)	ψ_β(°) (C1′-O3-C3-H3)	ϕ_α(°) (H1-C1-Oγ1-Cβ)	ψ_α(°) (C1-Oγ1-Cβ-Hβ)
Adiabatic map	180, 60 *anti, +gauche*	-60, 0, 60 *-gauche, syn, +gauche*	-60, 60 *-gauche, +gauche*	-60, 0, 60, 180 *-gauche, syn, +gauche, anti*
unbound model 1	63	7	-43	-150
unbound model 2	63	7	-43	-20
bound model 1	51	14	-38	-39
bound model 2	-17	-43	-70	-58

Figure 3. Adiabatic maps obtained from the rotation of ϕ and ψ angles at each glycosidic bond. Contour levels are drawn with high- to low-energy regions colored gradually from red to blue. Superimposed on every map are plotted the interproton distances: H1′−H3 for the β-D-GalpNAc-(1→3)-α-D-Galp linkage and H1−Hγ(Me) Thr and H5−Hγ(Me) Thr for the α-D-Galp-Thr linkage. The table underneath shows ϕ and ψ angles predicted from the adiabatic maps, those derived from nOe information (unbound main, unbound minor, and bound) and those derived from automated docking structures.

between H1 and Hβ of Thr. Moreover, there were nOes between H1 of GalpNAc and Hγ(Me) of Thr (weak) and H5 of GalpNAc and Hγ(Me) (medium). The simultaneous presence of all these nOes suggests the existence of a conformational equilibrium around ψ of the α-D-GalpNAc−Thr linkage in the free state. In the bound state, the relative intensity of the H1−Hβ Thr nOe versus that of the intraresidue H1−H2 nOe was maintained. In this case, the H1−Hγ(Me) Thr and H5−Hγ(Me) nOe cross-peak intensities were below the noise level. This fact suggests the existence of certain conformational selection. Because the intensities of the observed cross-peaks are very similar in both the free and bound states (although of different sign), the lack of the two peaks mentioned above, H1−Hγ(Me) Thr and H5−Hγ(Me), points toward the recognition of the $\phi = -60°$ and $\psi = -60°$ conformer. Table 1 summarizes the pertinent nOe intensities, their corresponding distances, and those obtained from molecular models constructed with ϕ_α values of −60° and −60° and ψ_α values of 180° and −60° for the α-D-GalpNAc-(1→O)−Thr bond. Unbound model 1 agrees with the presence of the H1−Hγ(Me) Thr nOe but does not explain the nOes between the H1 and Hβ of Thr and between H5 and Hγ(Me) of Thr. In contrast, unbound model 2 accounts for the nOes observed between H1 and Hβ of Thr and H5 and Hγ(Me) of Thr, but not the nOe from H1 to Hγ(Me) of Thr. It is worth noting that all these nOes could arise from *−gauche* and *+gauche* values of ϕ_α and ψ_α, respectively, except for a narrow range of ψ_α below 14° for which the distance from H1 to Hγ(Me) of Thr is larger than 4 Å. The fact that in the free state, nOes were observed simultaneously for the H1−Hγ(Me) Thr, H1−Hβ Thr, and H5−Hγ(Me) Thr proton pairs could be reconciled only with the presence of multiple conformations. Therefore, in the free state, three interconverting rotamer states may exist for the ψ_α angle: *−gauche*, *+gauche*, and *anti*. If the ψ_α *anti* state were predominant, the H5−Hγ(Me) Thr nOe would not have been observed and the H1−Hβ proton pair would have been weakened with a distance of ∼3.6 Å. Hence, either or both of the other two plausible rotamers for the ψ_α angle (*−gauche* and *gauche*) should be present. In contrast, the H1−Hγ(Me) Thr nOe and H5−Hγ(Me) Thr nOe were not

dx.doi.org/10.1021/bi300761s | *Biochemistry* 2012, 51, 7278−7289

observed in the bound state. The absence of these nOes strongly suggests that the −*gauche* ψ_α rotamer is more populated in the bound state.

Thus, molecular dynamics simulations (see below) were employed to rationalize these nOe data. Figure 3 shows the adiabatic maps obtained by variations in the ϕ and ψ angles, for each glycosidic bond. These maps revealed the global and local minima, providing information about preferred conformations and flexibility. The corresponding distances are mapped atop each plot. A summary of the allowed dihedral angles, predicted by the adiabatic maps, and the corresponding dihedral angles for TF antigen in the bound and free models, is presented in the table underneath Figure 3. It should be noted that a relatively high degree of conformational freedom is associated with the ψ angles of both glycosidic bonds. Moreover, an even flatter energy well was obtained for the ϕ and ψ angles of the reducing-end glycosidic bond. The maps coupled with the experimental nOes demonstrated the preference of the *exo* anomeric conformation around the $\beta(1\rightarrow3)$ glycosidic bond, in the free and bound states.

To rationalize the TR-NOESY data, molecular docking of TF antigen to the binding site of a homology model of CG-3 was performed. The docking was done in two ways: manual and automated. The manual docking involved the overlay of the β-D-galactosyl residues of the crystal structure of *N*-acetyllactosamine (PDB entry 1A3K) and the constructed bound-state conformation of TF antigen (bound model 1), followed by refinement and GlideXP scoring. The resultant binding pose is shown in Figure 4A, along with its GlideXP score. Panels B and C show the binding poses obtained from automated docking. Panel B (bound model 2) corresponds to the best scoring pose, while panel C shows a binding pose that is consistent with the NMR TR-NOESY data (see Figure 3) for the ϕ and ψ angles of the TF antigen in the binding poses obtained from GlideXP. Indeed, the ϕ and ψ angles of the nonreducing linkage for the conformer in panel C provided H1′−H3 distances very similar to those derived from the nOe data. The dihedral angles of the nonreducing linkage were thus consistent with the experimentally observed conformation. In panels B and C, the terminal nonreducing β-D-Gal*p* unit has been translated one sugar unit to the right with reference to the corresponding residue of *N*-acetyllactosamine in the crystal structure of PDB entry 1A3K. Nonetheless, the position of the internal α-D-Gal*p*NAc residue of the TF antigen was consistent with that of the terminal β-D-Gal*p* residue of *N*-acetyllactosamine.

We then sought to determine the possibility of observing the poses in panels B and C of Figure 4 in reported crystal structures. Thirty-eight carbohydrate−galectin complexes were analyzed (see Table S1 of the Supporting Information), from monosaccharides to hexasaccharides. Thirty-six ligands contained at least one galactosyl residue, either as a terminal or as an internal residue. Fittingly, the position of one of these galactosyl residues always coincided with that of β-D-Gal*p* observed in PDB entry 1A3K, as shown in Figure 4A. The preference of a galactosyl residue for this position was also shown in cases confined to monosaccharides (PDB entries 2GAL and 2ZGN), a thioglycoside analogue (PDB entry 1A78), or the α-anomer of lactose (PDB entry 2D6M). Interestingly, the ligand from PDB entries 3AFK and 3AYA was the TF antigen, and the position of the β-D-Gal*p* residue also matched that in PDB entry 1A3K. However, the binding modes in 3AFK and 3AYA were different relative to that determined from our TR-NOESY data: in the crystal structures, the methyl

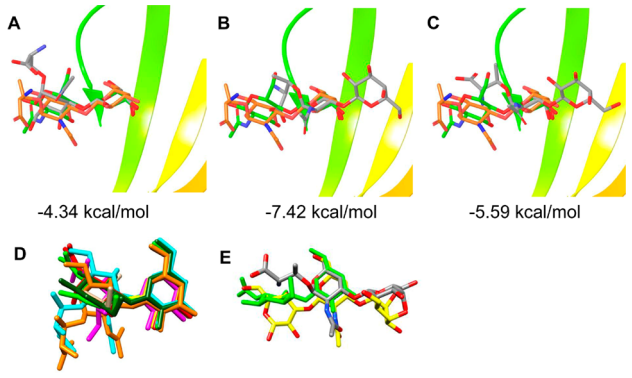


Figure 4. Docking poses of the TF antigen, colored by element type, with the CG-3 homology model: (A) manual docking, (B) best scoring pose, and (C) high-scoring pose in agreement with experimental TR-NOESY data. The GlideXP scores are shown below each panel. The carbon atoms of the X-ray structure of *N*-acetyllactosamine and the TF antigen extracted from PDB entries 1A3K and 3AYA, respectively, are colored lime green and orange, respectively. (D) Overlays of mono- and disaccharides from eight galectin−saccharide complexes with PDB entries 1A3K (lime green), 3OY8 (magenta), 2GAL (yellow), 3NV2 (red), 3NAJ (tan), 3AYE (dark green), 3AYA (orange), and 3AFK (cyan). (E) Overlay of trisaccharide α-D-Gal*p*NAc-(1→3)-β-D-Gal*p*NAc-(1→3)-β-D-Gal*p* [PDB entry 2EAL (yellow)] with *N*-acetyllactosamine [PDB entry 1A3K (lime green)] and the TF antigen from docking to the homology model (element type).

group of the threonine residue is oriented toward the solvent, while the TR-NOESY data also indicate a transitory orientation of this group toward the surface of CG-3. The overlay of eight mono- and disaccharides is depicted in Figure 4D. Even though there are positional variations in the reducing-end residues, the overlap of the nonreducing β-D-Gal*p* residues is evident. This overlap suggests that this pose is predominant in galectins in complexes with saccharides bearing a nonreducing-end galactosyl residue. We also found one case of a trisaccharide fragment, α-D-GalNAc-(1→3)-β-D-Gal*p*NAc-(1→3)-β-D-Gal (PDB entry 2EAL), from Forssman pentasaccharide, where the nonreducing terminal residue was shifted such that the internal β-D-Gal*p*NAc residue now overlaid with the terminal β-D-Gal*p* residue of *N*-acetyllactosamine in PDB entry 1A3K. The poses obtained from our automated dockings closely mimic this crystal structure, as outlined in Figure 4B. Analysis of the stabilizing interactions between the ligands and galectins from the crystal structure of PDB entry 2EAL and from the docking (data not shown) revealed that the majority of the contacts involve the internal β-D-Gal*p*NAc residue (of 2EAL), with only one hydrogen bond interaction between the nonreducing-end O6 atom and the NE atom of Trp82. Hence, this docking pose was also included in our MD simulations. The pose was stabilized by a salt bridge between the NH_3^+ group of threonine and the side chain of E174, which would be absent if the nonreducing terminal β-D-Gal*p* residue were fully equivalently positioned with that in 1A3K. It is noteworthy that this was the best scoring binding pose from GlideXP docking studies. Also consistent with the NMR data, the methyl group of the threonine side chain, although transient, is oriented toward the protein. Thus, the alternative binding pose might also be plausible.

Distinctive Molecular Dynamics of CG-3 upon Binding of the TF Antigen. Initial analysis of the protein backbone flexibility of human galectin-3 has produced evidence of

dx.doi.org/10.1021/bi300761s | *Biochemistry* 2012, 51, 7278−7289

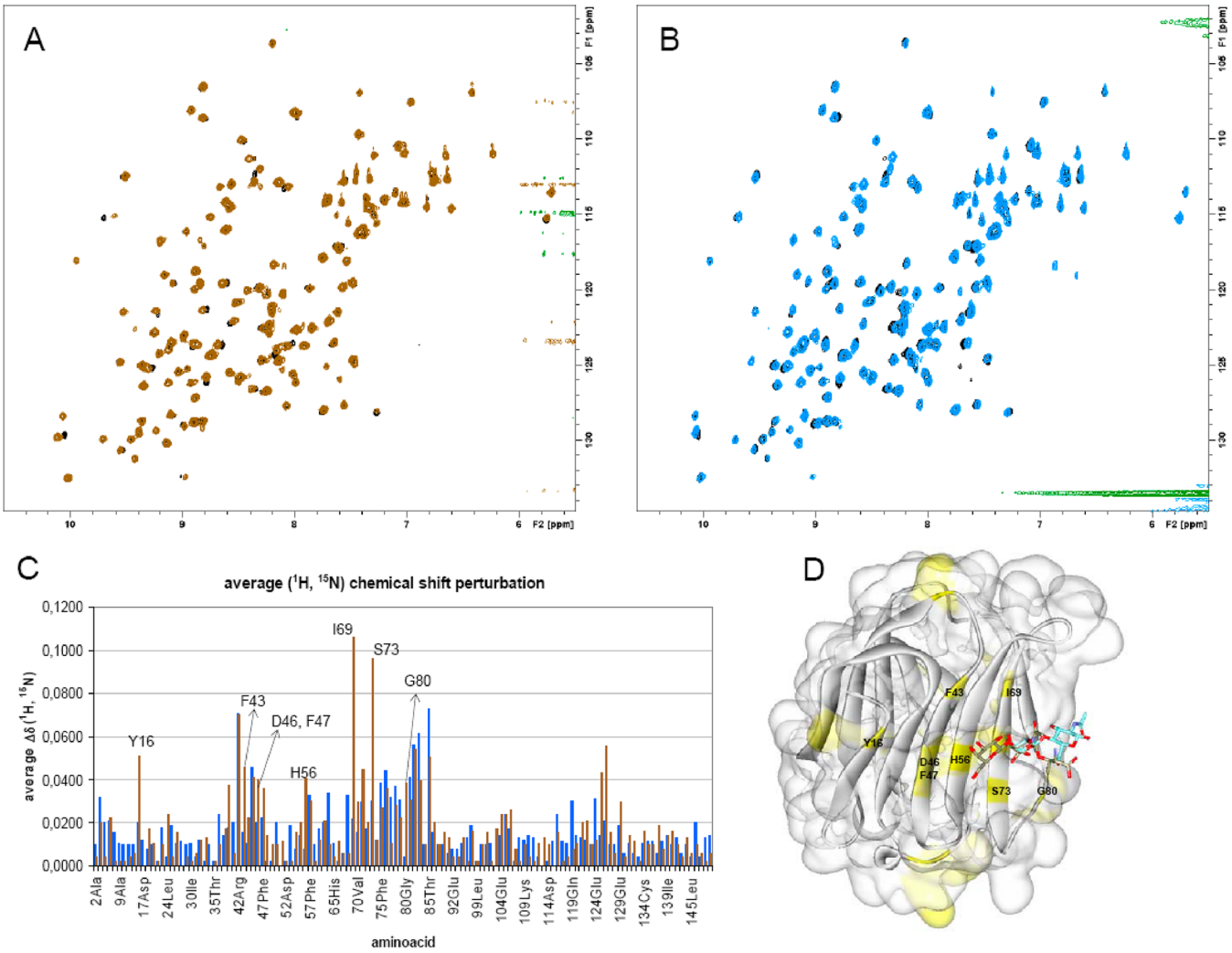


Figure 5. (A) Superposition of the HSQC spectra: CG-3 (black) and 1:20 CG-3:TF antigen molar ratio (brown). (B) Superimposition of HSQC spectra: CG-3 (black) and 1:20 CG-3:lactose molar ratio (blue). (C) Average (^{1}H and ^{15}N) chemical shift perturbation of CG-3 upon addition of TF antigen (brown bars) and lactose (blue bars). The residues that are affected by the addition of the TF antigen and not by lactose are annotated. (D) Model of CG-3 with the TF antigen docked in pose 2 (brown) and lactose docked (blue). CG-3 residues affected only by the addition of the TF antigen are highlighted in yellow. The proteins are numbered as follows: Y16 (Y108), F43 (F135), D46 (D138), F47 (F139), H56 (H148), I69 (I161), S73 (S165), and G80 (G172).

increased conformational entropy upon ligand binding.[44,45] To evaluate the effect of the TF antigen−Thr conjugate on the structure of CG-3, we performed ^{1}H−^{15}N heteronuclear single-quantum coherence (HSQC) titration experiments and compared them to the effect of lactose. In addition, we conducted molecular dynamics simulations in an attempt to distinguish between the different binding poses obtained from docking and to obtain insights into the observed differences in residue fluctuations in TF antigen and lactose complexes. Changes in the HSQC spectra (in the absence and presence of a given ligand) facilitate the identification of the protein residues that are mostly affected as a result of binding to the added ligand. Moreover, comparisons of the changes derived by different ligands permit identification of residues that are selectively affected. The HSQC spectrum of apo-CG-3 is colored black in panels A and B of Figure 5. The HSQC spectra after the addition of the TF antigen−Thr conjugate (1:20 CG-3:TF antigen−Thr conjugate molar ratio) are superimposed in brown in Figure 5A. The spectrum recorded after addition of lactose (1:20 CG-3:lactose molar ratio) is colored blue in Figure 5B. Figure 5C shows the average chemical shift perturbation derived from these HSQC experiments, due to the presence of each ligand, with brown bars for the addition of TF antigen and blue bars for the presence of lactose. It can be deduced that those residues affected by lactose binding were also affected by the addition of the TF antigen. However, some additional variations were observed when the TF antigen was added. These extra residues are mapped in yellow in the protein model, as depicted in Figure 5D. The most evident changes are shown in the brown bars of the figure, especially for Y108, F135, H148, I161, and S165 and to a lesser extent for residues D138, F139, G172, K217, and K218. All these residues that are affected by only the TF antigen are located along the β-sheet surface and at the loops connecting them. Because these residues are relatively far from the typical lactose binding site, these observations may suggest the presence of an alternative binding mode, for example, that observed in PDB entry 2EAL.

Molecular Dynamics Simulations of the Free TF Antigen and Complexes of CG-3 with the TF Antigen and Lactose. The nOe- and TR-NOESY-derived interproton

dx.doi.org/10.1021/bi300761s | *Biochemistry* 2012, 51, 7278−7289

distances for the β-D-Gal*p*-(1→3)-α-D-Gal*p*NAc linkage suggested the existence of one major conformer at this site in both the free and bound states. However, a single conformer could not account for all the experimental data of the α-D-Gal*p*NAc-(1→*O*)-Thr linkage: two or more ψ angles could be populated. Thus, MD simulations were performed to understand the dynamics of the TF antigen in the free state and also to provide atomic-level insights into the modes of binding of the TF antigen to CG-3. However, in the course of generation of the TF antigen−CG-3 complex, two modes were spawned: *manualDock* and *glideXPDock* (see Materials and Methods). It was therefore necessary to determine which of these poses was more consistent with the available experimental NMR HSQC, TR-NOESY, and STD data. To this end, the following analyses were performed on the entire trajectories of the complexes: nOe-derived interproton distances, per residue root-mean-square fluctuations (rmsfs), per residue/atom solvent accessible surface areas (SASA), intra- and intermolecular hydrogen bond interactions, and molecular mechanics generalized Born surface area calculations (MM-GBSA). Furthermore, the entire trajectories were separated into NMR-consistent and NMR-inconsistent populations and reanalyzed. From these analyses, the *manualDock* complex was selected as the more consistent, and on this basis, a dynamic binding mode for the TF antigen−CG-3 complex comprising two poses was proposed.

The root-mean-square deviations showing equilibration of the free and bound ligand trajectories are shown in Figure S2 of the Supporting Information. The ring atoms and backbone atoms of the threonine residue were employed for the free ligand, while the backbone atoms of the receptor were used for the complexes.

nOe-Derived Interproton Distances. A comparison of the experimental and MD-derived interproton distances for both the free and bound TF antigen is presented in Table 2.

Table 2. Experimental and Computational[a] nOe-Derived Distances (angstroms)

	NOESY[b]	TR-NOESY[c]
	experimental/ MD	experimental/ *glideXPDock*/*manualDock*
H1′−H5′	2.5	2.5
	2.57 ± 0.20	2.50 ± 0.18/2.46 ± 0.19
H1′−H3	2.72	2.36
	2.34 ± 0.22	2.27 ± 0.46/2.23 ± 0.17
H1−Hβ	2.36	2.33
	2.26 ± 0.19	2.30 ± 0.20/2.19 ± 0.19
H5−Hγ(Me)	2.69	no[d]
	3.28 ± 0.38	3.23 ± 0.27/3.42 ± 0.51
H1−Hγ(Me)	3.52	no[d]
	3.84 ± 0.35	4.20 ± 0.20/3.51 ± 0.46

[a]The value of $\langle 1/r^6\rangle^{-1/6}$ averaged 80 and 20 ns of simulation time for the ligand and each complex, respectively, for comparison with the NOESY and TR-NOESY data, respectively. [b]Unbound ligand. [c]Bound ligand. [d]Not observed nOe.

For the free TF antigen, the nOe distances showed good overall agreement with the experimental data, including interglycosidic nOe distances that are important for determining the conformational properties of carbohydrates. For example, the H1−Hγ(Me) Thr distances of 3.52 and 3.84 ± 0.35 Å for the experimental and MD data, respectively, are in good agreement considering the α-D-Gal*p*NAc-(1→O)-Thr linkage. The H1′−H3 distance is slightly underestimated by the simulation (2.61 Å vs 2.34 ± 0.22 Å), while the H5−Hγ(Me) Thr distance is overestimated (2.69 Å vs 3.28 ± 0.38 Å). When CG-3 binds, the majority of the interglycosidic nOes observed in the free state of the TF antigen (H1′−H3 and H1−Hβ Thr) are preserved, suggesting similar conformational properties for those dihedral angles that influence these interproton distances (ϕ_β, ψ_β, and ϕ_α). The MD interproton distances for both the *glideXPDock* and *manualDock* complexes agreed favorably with the respective experimental distances. While the H1−Hγ(Me) Thr and H5−Hγ(Me) Thr nOes were observed in the free state, they were not detected in the complex, which is evident from the TR-NOESY data. These interproton distances are largely dependent on the values adopted by ψ_α and pointed to differences in the conformational preferences of this dihedral angle for the complexed and free TF antigen. The H1−Hγ(Me) Thr distances computed from the *glideXPDock* and *manualDock* trajectories were 4.20 ± 0.20 and 3.51 ± 0.46 Å, respectively. Given the 4 Å limit of nOe detection,[46] the *glideXPDock* trajectory seemed to be more consistent in this case, but as will be shown later, other considerations ruled out the *glideXPDock* complex. However, neither *glideXPDock* (3.23 ± 0.27 Å) nor *manualDock* (3.42 ± 0.51 Å) trajectories could explain the absence of the H5−Hγ(Me) Thr nOe.

The average dihedral angles and populations sampled in the course of the MD simulations are listed in Table S2 of the Supporting Information. The rotamer populations of the ϕ_β, ψ_β, and ϕ_α dihedral angles in the *glideXPDock* and *manualDock* complexes echoed those of the free ligand, though the −*gauche* ψ_β population was remarkably larger (90%) in the *manualDock* complex. For the ψ_α torsion angle, the partition was based on values that could result in a distance of greater than or less than 4.0 Å between the α-D-Gal*p*NAc H1 and Thr Hγ protons. This cutoff was determined to be 14° by performing rotations about the ψ_α angle while monitoring the H1−Hγ(Me) distances. Thus, in the free state, two ψ_α rotamers, 32 ± 11° and −4 ± 13°, were populated with populations of 58 and 42%, respectively. The disappearance of the H1−Hγ(Me) nOe in the complex indicated a preference for the latter conformer. This conformer was predominant (88%) in the *glideXPDock* complex and existed for only 30% of the time in the *manualDock* complex, which explained the H1−Hγ(Me) distances of 4.20 ± 0.20 and 3.51 ± 0.46 Å computed from the *glideXPDock* and *manualDock* simulations, respectively.

Root-Mean-Square Fluctuations (rmsfs). As shown earlier (Figure 5), chemical shift differences in the backbone NH moiety were observed in the spectra of free CG-3, as well as CG-3 loaded with the TF antigen or lactose. Remarkably, changes were also observed for residues situated away from the binding site (Figure 5C and Table S3 of the Supporting Information). For the purpose of determining the extent to which the *glideXPDock* and *manualDock* complexes were consistent with the observed changes, per residue rmsfs were calculated from both trajectories and compared with those computed from a trajectory of apo-CG-3 (Figure S3 of the Supporting Information). Differences in fluctuations were observed for the specified residues between apo- and holo-CG-3, especially for the *glideXPDock* complex. In particular, for residues 161, 165, 217, and 218, located on a loop very distal from the ligand binding cavity, differences in their fluctuations were more obvious in the *glideXPDock* system than in the *manualDock* complex. Noteworthy, these MD fluctuations echoed the trends observed in the experimental data: the

Table 3. Binding Free Energy Changes (kilocalories per mole) Computed from the Trajectories of the GlideXP and Manually Docked TF Antigen to CG-3

	GlideXP docking			manual docking		
energy	igb = 1	igb = 2	igb = 5	igb = 1	igb = 2	igb = 5
$\langle \Delta E_{vdw} \rangle^a$	−24.70 ± 4.57	−24.70 ± 2.98	−24.70 ± 4.42	−19.35 ± 4.27	−19.35 ± 4.27	−19.35 ± 4.27
$\langle \Delta E_{coul} \rangle^b$	−53.51 ± 13.99	−53.51 ± 13.99	−53.51 ± 13.99	−60.25 ± 23.51	−60.25 ± 23.51	−60.25 ± 23.51
$\langle \Delta G_{polar,GB} \rangle^c$	55.01 ± 11.18	62.26 ± 11.89	60.80 ± 11.63	53.14 ± 18.03	60.28 ± 18.93	58.21 ± 18.57
$\langle \Delta G_{nonpolar} \rangle^d$	−3.94 ± 0.32	−3.94 ± 0.32	−3.94 ± 0.32	−3.43 ± 0.38	−3.43 ± 0.38	−3.43 ± 0.38
$\langle \Delta E_{bat} \rangle^e$	0.00	0.0	0.0	0.0	0.0	0.0
$\langle \Delta G_{TOT} \rangle^f$	−27.14 ± 4.57	−19.89 ± 4.01	−21.35 ± 4.42	−29.89 ± 6.69	−22.75 ± 6.05	−24.82 ± 6.69
$-T\Delta S^g$	0.80	0.80	0.80	0.54	0.54	0.54

[a]Molecular mechanics van der Waals energy. [b]Molecular mechanics coulombic energy. [c]Polar contribution to the solvation energy. [d]Nonpolar contribution to the solvation energy. [e]Contribution from the bond, angle, and dihedral angle strain energies. [f]Total energy excluding contributions from entropy. [g]Temperature × conformational entropy.

backbone NH group of S165 showed larger differences in the complex with the TF antigen than in that with lactose.

Solvent Accessible Surface Area (SASA). The STD data pointed to a binding pose in which the H4 protons of β-D-Gal*p* and α-D-Gal*p*NAc are the closest residues from the surface of the protein. The per residue and per atom SASA values are presented in Figure S4 and Table S4 of the Supporting Information.

Hydrogen Bonding Interactions. Hydrogen bonding interactions were also determined from the trajectories of the complexes, using donor−acceptor distance and donor−H−acceptor angle cutoffs of 3.5 Å and 100°, respectively. A hydrogen bond was deemed important if it persisted for at least 5% of the simulation time. The objective was to determine essential features of the different modes of binding of the TF antigen and to try to understand the basis of the rotamer population changes of the ψ_α angle observed in the TR-NOESY data. The hydrogen bond distances and percent occupancies from the *glideXPDock* and *manualDock* complexes are provided as Tables S5 and S6, respectively, of the Supporting Information. Interestingly, the salt bridge that was formed between the NH_3^+ group of threonine and the carboxyl side chain of E174 in the GlideXP docked complex was absent during the explicit-solvent simulation under the constraints employed here to define hydrogen bonds. A qualitative description of the intermolecular hydrogen bonds for the threonine, α-D-Gal*p*NAc, and β-D-Gal*p* residues was performed using the following notation: none, high, and intermediate, respectively, for the *glideXPDock* trajectory and low, intermediate, and high, respectively, for the *manualDock* trajectory. A larger conformer selectivity of the NMR-consistent ψ_α dihedral angle was observed for the *glideXPDock* pose (see Table S2 of the Supporting Information). However, the absence of any intermolecular hydrogen bonds between the threonine residue and the receptor in the *glideXPDock* structure, in conjunction with the mutual presence of all intramolecular bonds involving the threonine residue in both trajectories, made it difficult to determine the origin of the conformational preference. With regard to the different binding poses, the hydrogen bonding data coupled with the per residue SASA data suggested that the shape of the TF antigen was "concave" and "convex" for the *manualDock* and *glideXPDock* complexes, respectively, from the perspective of the receptor. The β-D-Gal*p* and threonine residues occupied the tips, while the α-D-Gal*p*NAc residue was at the center of the curvature.

Molecular Mechanics Generalized Born Surface Area (MM-GBSA) Calculations of Interaction Energies. As a further step, the *glideXPDock* and *manualDock* trajectories were subjected to MM-GBSA analyses to quantify the interaction energies between the TF antigen and CG-3 in both complexes. In this study, only the trajectories of the complexes were employed, as reported previously for a polysaccharide−antibody complex.[47] The advantages of utilizing MM-GBSA analyses were (1) the ability to decompose the interaction energies into contributions from molecular mechanics energies, polar and nonpolar contributions to solvation and desolvation, van der Waals, and electrostatics and (2) the fact that the use of several structures derived from MD simulations provided an average of ensembles as opposed to a single docked structure. The computed interaction energies and decomposition into different energy components are listed in Table 3. To determine the extent to which the results of the computed free energy changes could be influenced by the choice of solvation model, igb values of 1, 2, and 5 were employed. Overall, the ΔG_{TOT} indicated that the *manualDock* complex was preferred in solution. The conformational entropic penalties were <1 kcal/mol, although they were slightly higher for the *glideXPDock* complex. The presence of several hydrogen bonds between the TF antigen and CG-3 in both complexes strongly suggests the predominance of polar interactions for determining the overall binding free energies. The Coulomb energy contributions also lent credence to this statement. However, as expected for carbohydrates, the polar solvation term was very unfavorable in both complexes and almost neutralized the Coulomb energy term, thus resulting in the major net favorable contributions arising from the van der Waals components. A similar observation has been noted in other studies involving oligosaccharide−protein complexes.[47−49] Given the absence of unanimity in the outcomes of the analyses in determining which complex was more consistent with the available NMR data, the *manualDock* and *glideXPDock* trajectories were partitioned into subtrajectories whose ψ_α angles would result in the H1−Hγ(Me) Thr distance being greater than or less than 4.0 Å (nOe not observed in the NMR analysis). The trajectories for which the H1−Hγ(Me) Thr distance was greater than 4.0 Å were denoted NMR consistent, while those for which the distance was less than 4.0 Å were called NMR inconsistent. The previous analyses conducted for the entire trajectories were also performed on these subtrajectories. In the interest of brevity, the majority of these data are provided as Supporting Information, and we have highlighted the results from the NMR consistent trajectory of the *manualDock* complex.

dx.doi.org/10.1021/bi300761s | *Biochemistry* 2012, 51, 7278−7289

Table 4. Binding Free Energy Changes Computed (igb = 2) from Subsets of the Trajectories of the GlideXP and Manually Docked TF Antigen to CG-3

	GlideXP docking		manual docking	
energy (kcal/mol)	NMR consistent	NMR inconsistent	NMR consistent	NMR inconsistent
$\langle \Delta E_{vdw} \rangle^a$	−24.77 ± 2.96	−24.76 ± 3.07	−17.76 ± 3.84	−20.15 ± 4.22
$\langle \Delta E_{coul} \rangle^b$	−54.20 ± 13.81	−49.08 ± 14.41	−65.63 ± 26.01	−57.71 ± 21.71
$\langle \Delta G_{polar,GB} \rangle^c$	62.91 ± 11.72	58.15 ± 12.40	62.45 ± 22.06	59.21 ± 17.14
$\langle \Delta G_{nonpolar} \rangle^d$	−3.94 ± 0.32	−3.92 ± 0.35	−3.24 ± 0.32	−3.51 ± 0.37
$\langle \Delta E_{bat} \rangle^e$	0.00	0.00	0.00	0.00
$\langle \Delta G_{TOT} \rangle^f$	−20.00 ± 3.94	−19.61 ± 4.12	−24.18 ± 8.65	−22.16 ± 5.81
$-T\Delta S^g$	1.24	0.79	1.28	0.80

[a]Molecular mechanics van der Waals energy. [b]Molecular mechanics coulombic energy. [c]Polar contribution to the solvation energy. [d]Nonpolar contribution to the solvation energy. [e]Contribution from the bond, angle, and dihedral angle strain energies. [f]Total energy excluding contributions from entropy. [g]Temperature × conformational entropy.

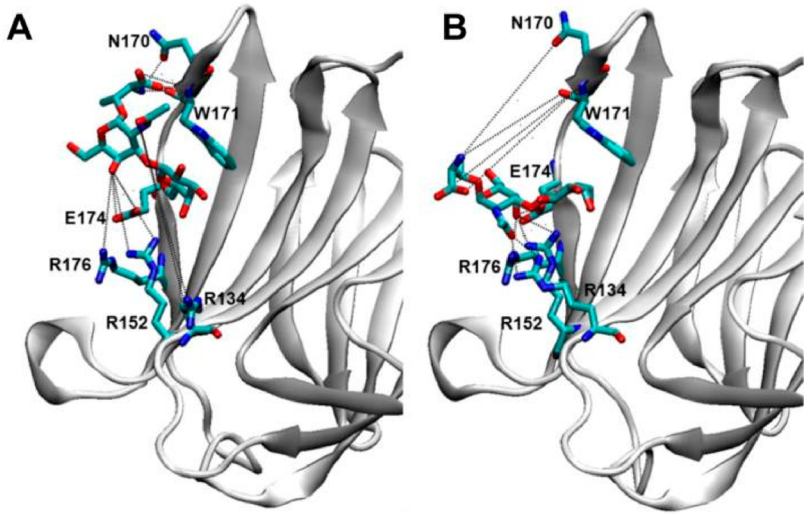


Figure 6. Two poses for the TF antigen bound to CG-3 identified in the NMR consistent trajectory of the manually docked ligand: (A) pose 1 and (B) pose 2.

The interproton distances for these subtrajectories are listed in Table S7 of the Supporting Information. As expected, the MD-derived H1−Hγ(Me) Thr distance compared favorably and unfavorably with the experimental nOe-derived distances for the NMR consistent and NMR inconsistent trajectories, respectively. Interestingly, no significant changes were observed in the other interproton distances.

In the case of the rmsfs, improvements were observed in the *manualDock* NMR consistent trajectory. In particular, residues K217 and K218 that were very similar in the entire trajectory of the *manualDock* complex now showed distinct separations from values computed from apo-CG-3 (Figure S5 of the Supporting Information). The corresponding rmsfs for the partitioned *glideXPDock* trajectories are given in Figure S6 of the Supporting Information. The SASA values (see Table S8 of the Supporting Information) led to the same conclusions derived from the entire *manualDock* trajectory, though the SASA for the H1′ proton showed a slight 1.42 Å^2 increase. Notably, splitting the *glideXPDock* complex did not alleviate the high solvent accessibility of H4′ in either subtrajectory. The MM-GBSA-derived interaction energies for these subtrajectories are listed in Table 4. As was the case with the entire trajectories, the *manualDock* complexes were predicted to be more stable than the *glideXPDock* complexes, whether NMR consistent or inconsistent. Interestingly, the *manualDock* NMR consistent interaction energies were more favorable than the NMR inconsistent counterparts. Of note is the slight increase in the entropic penalties of these subtrajectories relative to the complete trajectories: 0.54 kcal/mol for the full trajectory compared to 1.28 and 0.80 kcal/mol for the *manualDock* NMR consistent and inconsistent trajectories, respectively. This is to be expected given that the conformational space of the bound TF antigen is further restricted by the partitioning when compared to the free state. On these bases, two binding poses were proposed by analyzing the NMR consistent trajectory of the *manualDock* complex, as depicted in Figure 6. The nonreducing-end β-D-Gal*p* residue served as an anchor in both poses, making strong intermolecular hydrogen bonds with the receptor. The O4 atom of this residue formed the characteristic hydrogen bond with the NE2-HE2 proton of H148 in the binding pocket. In the first pose, pose 1, hydrogen bond interactions were formed primarily between the threonine residue and N170 and W171 of CG-3 (Figure 6). This binding mode can be used for the design of new peptide−carbohydrate ligands aiming to enhance the participation of the peptide in the binding recognition. In the second pose, pose 2, hydrogen bonds were formed predominantly between the internal α-D-Gal*p*NAc and R134, R152, and R176 of CG-3 (Figure 6). Pose 2 is in agreement with the orientation observed in the X-ray crystal structure of the TF antigen in complex with the carbohydrate recognition domain of galectin-3.[24] Common to both poses was a hydrogen bond between the O4 atom of α-D-

dx.doi.org/10.1021/bi300761s | *Biochemistry* 2012, 51, 7278−7289

GalpNAc and E174. E174 may serve as an arm that holds α-D-GalpNAc close to the protein surface during conversions between the poses. Pose 1 may rationalize the weak STD enhancement of the Hγ(Me) Thr protons. A complete listing of the hydrogen bonding data is provided in Tables S9 and S10 of the Supporting Information.

CONCLUSIONS

In this study, we have analyzed the molecular recognition features of the binding of the threonine O-linked TF antigen to chimera-type avian galectin-3. The most intense STD effects correspond to the H4 protons of the β-D-Galp and α-D-GalpNAc moieties, which is in contrast with the usually observed STD pattern for the recognition of lactose by galectins. The weak saturation of the methyl protons of the Thr residue observed by STD NMR suggests the existence of fleeting contacts with CG-3. TR-NOESY NMR experiments and molecular modeling revealed possible differences in conformational populations between free and bound states. Analysis of heteronuclear single-quantum coherence titration (HSQC) binding data indicated that the residues of CG-3 affected upon binding with lactose and the TF epitope were very similar, being located in the same region around the expected lactose-binding site. However, the trend of some shifts in the HSQC spectra was different, indicating disparities in binding interactions. The differences were observed in residues localized far from the lactose-binding site along the β-sheet surface and at the loops connecting them. Molecular dynamics simulations revealed two possible binding poses of which one, pose 2, coincides with that of the X-ray resolved crystal structure of the TF antigen in complex with galectin-3. Interestingly, pose 1 (Figure 6A) revealed unique low-occupancy intermolecular interactions formed primarily between the terminal threonine residue and galectin-3. These findings are an indication of the fine binding specificity of the TF antigen toward galectin-3 and demonstrate the importance of multidisciplinary strategies in revealing the structural and molecular basis of carbohydrate−lectin interactions. Respective structural work on human galectin-3 is underway. We are also undertaking binding studies of human galectin-3 with MUC1 20-mer repeats that have the TF antigen attached at one of the glycosylation sites to delineate any contribution of the protein backbone to binding. This way, interspecies differences in these aspects (interaction with the glycan and peptide parts) are likely to be identified, facilitating the relation of deviations in sequence between the two galectins to changes in binding properties. Moreover, this study may further the development of novel approaches for the structure-based design of galectin-3 inhibitors with improved affinity and selectivity.

ASSOCIATED CONTENT

Supporting Information

Molecular modeling and molecular dynamics protocol, PDB entries of 38 carbohydrate−galectin complexes, tables, and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Torrey Pines Institute for Molecular Studies, 11350 S.W. Village Parkway, Port Saint Lucie, FL 34987. K.M.-M.: phone, (772) 345-4688; e-mail, kmartinez@tpims.org. M.C.: phone, (772) 345-4597; e-mail, mcudic@tpims.org.

Funding

This study was supported by the State of Florida, Executive Officer of the Governor's Office of Tourism, Trade and Economic Development (K.M.-M. and M.C.), and the European Union Seventh Framework Program (FP7/2007-2013) under Grant 260600 ("GlycoHIT"). The group at Madrid also thanks the Ministry of Economy and Competitiveness of Spain for funding (CTQ2009-08536).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Matthew Tessier for discussions about O-linked threonine calculations, Dr. Gregg B. Fields for critical reading of the manuscript, and Karen Gottwald for editing.

ABBREVIATIONS

TF, Thomsen-Friedenreich oncofetal; CG-3, chimera-type avian galectin-3; STD, saturation transfer difference; TR-NOESY, two-dimensional transferred nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single-quantum coherence titration; CRD, carbohydrate recognition domain; MD, molecular dynamics; rmsf, per residue root-mean-square fluctuation; SASA, per residue/atom solvent accessible surface area; MM-GBSA, molecular mechanics generalized Born surface area calculations.

REFERENCES

(1) Gabius, H. J. (2009) *The sugar code: Fundamentals of glycosciences*, Wiley-VCH, Weinheim, Germany.
(2) Gabius, H. J. (1997) Animal lectins. *Eur. J. Biochem. 243*, 543−576.
(3) Lis, H., and Sharon, N. (1998) Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition. *Chem. Rev. 98*, 637−674.
(4) Kilpatrick, D. C. (2000) *Handbook of Animal Lectins. Properties and Biomediacal Applications*, Wiley, Chichester, U.K.
(5) Ambrosi, M., Cameron, N. R., and Davis, B. G. (2005) Lectins: Tools for the molecular understanding of the glycocode. *Org. Biomol. Chem. 3*, 1593−1608.
(6) Gabius, H. J., Andre, S., Jimenez-Barbero, J., Romero, A., and Solis, D. (2011) From lectin structure to functional glycomics: Principles of the sugar code. *Trends Biochem. Sci. 36*, 298−313.
(7) Di Lella, S., Sundblad, V., Cerliani, J. P., Guardia, C. M., Estrin, D. A., Vasta, G. R., and Rabinovich, G. A. (2011) When Galectins Recognize Glycans: From Biochemistry to Physiology and Back Again. *Biochemistry 50*, 7842−7857.
(8) Liu, F.-T., and Rabinovich, G. A. (2005) Galectins as modulators of tumour progression. *Nat. Rev. Cancer 5*, 29−41.
(9) Norling, L. V., Perretti, M., and Cooper, D. (2009) Endogenous galectins and the control of the host inflammatory response. *J. Endocrinol. 201*, 169−184.
(10) Kasai, K., and Hirabayashi, J. (1996) Galectins: A family of animal lectins that decipher glycocodes. *J. Biochem. 119*, 1−8.
(11) Cooper, D. N. (2002) Galectinomics: Finding themes in complexity. *Biochim. Biophys. Acta 1572*, 209−231.
(12) Kaltner, H., and Gabius, H. J. (2012) A toolbox of lectins for translating the sugar code: The galectin network in phylogenesis and tumors. *Histol. Histopathol. 27*, 397−416.
(13) Hughes, R. C. (1994) Mac-2: A versatile galactose-binding protein of mammalian tissues. *Glycobiology 4*, 5−12.
(14) Dumic, J., Dabelic, S., and Flögel, M. (2006) Galectin-3: An open-ended story. *Biochim. Biophys. Acta 1760*, 616−635.
(15) Ahmad, N., Gabius, H. J., Andre, S., Kaltner, H., Sabesan, S., Roy, R., Liu, B., Macaluso, F., and Brewer, C. F. (2004) Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates

and forms heterogeneous cross-linked complexes. *J. Biol. Chem. 279*, 10841−10847.
(16) Kopitz, J., von Reitzenstein, C., Andre, S., Kaltner, H., Uhl, J., Ehemann, V., Cantz, M., and Gabius, H. J. (2001) Negative regulation of neuroblastoma cell growth by carbohydrate-dependent surface binding of galectin-1 and functional divergence from galectin-3. *J. Biol. Chem. 276*, 35917−35923.
(17) Sanchez-Ruderisch, H., Fischer, C., Detjen, K. M., Welzel, M., Wimmel, A., Manning, J. C., Andre, S., and Gabius, H. J. (2010) Tumor suppressor p16 INK4a: Downregulation of galectin-3, an endogenous competitor of the pro-anoikis effector galectin-1, in a pancreatic carcinoma model. *FEBS J. 277*, 3552−3563.
(18) Houzelstein, D., Goncalves, I. R., Fadden, A. J., Sidhu, S. S., Cooper, D. N., Drickamer, K., Leffler, H., and Poirier, F. (2004) Phylogenetic analysis of the vertebrate galectin family. *Mol. Biol. Evol. 21*, 1177−1187.
(19) Yu, L.-G., Andrews, N., Zhao, Q., McKean, D., Williams, J. F., Connor, L. J., Gerasimenko, O. V., Hilkens, J., Hirabayashi, J., Kasai, K., and Rhodes, J. M. (2007) Galectin-3 Interaction with Thomsen-Friedenreich Disaccharide on Cancer-associated MUC1 Causes Increased Cancer Cell Endothelial Adhesion. *J. Biol. Chem. 282*, 773−781.
(20) Senapati, S., Chaturvedi, P., Chaney, W. G., Chakraborty, S., Gnanapragassam, V. S., Sasson, A. R., and Batra, S. K. (2011) Novel Interaction of MUC4 and Galectin: Potential Pathobiological Implications for Metastasis in Lethal Pancreatic Cancer. *Clin. Cancer Res. 17*, 267−274.
(21) Sparrow, C. P., Leffler, H., and Barondes, S. H. (1987) Multiple soluble β-galactoside-binding lectins from human lung. *J. Biol. Chem. 262*, 7383−7390.
(22) Bresalier, R. S., Byrd, J. C., Wang, L., and Raz, A. (1996) Colon cancer mucin: A new ligand for the β-galactoside-binding protein galectin-3. *Cancer Res. 56*, 4354−4357.
(23) Andre, S., Kaltner, H., Lensch, M., Russwurm, R., Siebert, H. C., Fallsehr, C., Tajkhorshid, E., Heck, A. J., von Knebel Doeberitz, M., Gabius, H. J., and Kopitz, J. (2005) Determination of structural and functional overlap/divergence of five proto-type galectins by analysis of the growth-regulatory interaction with ganglioside GM1 in silico and in vitro on human neuroblastoma cells. *Int. J. Cancer 114*, 46−57.
(24) Bian, C.-F., Zhang, Y., Sun, H., Li, D.-F., and Wang, D.-C. (2011) Structural Basis for Distinct Binding Properties of the Human Galectins to Thomsen-Friedenreich Antigen. *PLoS One 6*, e25007.
(25) Krzeminski, M., Singh, T., André, S., Lensch, M., Wu, A. M., Bonvin, A. M. J. J., and Gabius, H.-J. (2011) Human galectin-3 (Mac-2 antigen): Defining molecular switches of affinity to natural glycoproteins, structural and dynamic aspects of glycan binding by flexible ligand docking and putative regulatory sequences in the proximal promoter region. *Biochim. Biophys. Acta 1810*, 150−161.
(26) Kaltner, H., Kübler, D., López-Merino, L., Lohr, M., Manning, J. C., Lensch, M., Seidler, J., Lehmann, W. D., André, S., Solís, D., and Gabius, H.-J. (2011) Toward Comprehensive Analysis of the Galectin Network in Chicken: Unique Diversity of Galectin-3 and Comparison of its Localization Profile in Organs of Adult Animals to the Other Four Members of this Lectin Family. *Anat. Rec., Part A 294*, 427−444.
(27) Feng, L., Sun, H., Zhang, Y., Li, D.-F., and Wang, D.-C. (2010) Structural insights into the recognition mechanism between an antitumor galectin AAL and the Thomsen-Friedenreich antigen. *FASEB J. 24*, 3861−3868.
(28) Martin-Santamaria, S., Gabius, H.-J., and Jimenez-Barbero, J. (2012) Structural studies on the interaction of saccharides and glycomimetics with galectin-1: A 3D perspective using a combined molecular modeling and NMR approach. *Pure Appl. Chem. 84*, 49−64.
(29) André, S., Maljaars, C. E. P., Halkes, K. M., Gabius, H.-J., and Kamerling, J. P. (2007) Discovery of galectin ligands in fully randomized combinatorial one-bead-one-compound (glyco)peptide libraries. *Bioorg. Med. Chem. Lett. 17*, 793−798.
(30) Andre, S., Giguere, D., Dam, T. K., Brewer, F., Gabius, H.-J., and Roy, R. (2010) Synthesis and screening of a small glycomimetic library for inhibitory activity on medically relevant galactoside-specific lectins in assays of increasing biorelevance. *New J. Chem. 34*, 2229−2240.
(31) Giguère, D., André, S., Bonin, M.-A., Bellefleur, M.-A., Provencal, A., Cloutier, P., Pucci, B., Roy, R., and Gabius, H.-J. (2011) Inhibitory potential of chemical substitutions at bioinspired sites of β-D-galactopyranose on neoglycoprotein/cell surface binding of two classes of medically relevant lectins. *Bioorg. Med. Chem. 19*, 3280−3287.
(32) Kaltner, H., Solis, D., Kopitz, J., Lensch, M., Lohr, M., Manning, J. C., Murnseer, M., Schnolzer, M., Andre, S., Saiz, J. L., and Gabius, H. J. (2008) Prototype chicken galectins revisited: Characterization of a third protein with distinctive hydrodynamic behaviour and expression pattern in organs of adult animals. *Biochem. J. 409*, 591−599.
(33) Gabius, H. J., Engelhardt, R., Rehm, S., and Cramer, F. (1984) Biochemical characterization of endogenous carbohydrate-binding proteins from spontaneous murine rhabdomyosarcoma, mammary adenocarcinoma, and ovarian teratoma. *J. Natl. Cancer Inst. 73*, 1349−1357.
(34) Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR 20*, 71−75.
(35) Mayer, M., and Meyer, B. (1999) Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem., Int. Ed. 38*, 1784−1788.
(36) Woods Group 2005−2012 GLYCAM Web Complex Carbohydrate Research Center, University of Georgia, Athens, GA (http://www.glycam.com).
(37) Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell. 25*, 247−260.
(38) Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008) GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem. 29*, 622−655.
(39) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf. 65*, 712−725.
(40) *Glide*, version 5.7 (2011) Schrödinger, LLC, New York.
(41) Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res. 37*, D387−D392.
(42) Kopp, J., and Schwede, T. (2004) The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Res. 32*, D230−D234.
(43) Muñoz, F. J., Santos, J. I., Ardá, A., André, S., Gabius, H. J., Sinisterra, J. V., Jiménez-Barbero, J., and Hernáiz, M. J. (2010) Binding studies of adhesion/growth-regulatory galectins with glycoconjugates monitored by surface plasmon resonance and NMR spectroscopy. *Org. Biomol. Chem. 8*, 2986−2992.
(44) Diehl, C., Genheden, S., Modig, K., Ryde, U., and Akke, M. (2009) Conformational entropy changes upon lactose binding to the carbohydrate recognition domain of galectin-3. *J. Biomol. NMR 45*, 157−169.
(45) Diehl, C., Engström, O., Delaine, T., Håkansson, M., Genheden, S., Modig, K., Leffler, H., Ryde, U., Nilsson, U. J., and Akke, M. (2010) Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3. *J. Am. Chem. Soc. 132*, 14577−14589.
(46) Wormald, M. R., Petrescu, A. J., Pao, Y.-L., Glithero, A., Elliott, T., and Dwek, R. A. (2002) Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography and molecular modelling. *Chem. Rev. 102*, 371−386.
(47) Kadirvelraj, R., Gonzalez-Outeirino, J., Foley, B. L., Beckham, M. L., Jennings, H. J., Foote, S., Ford, M. G., and Woods, R. J. (2006) Understanding the bacterial polysaccharide antigenicity of *Streptococcus agalactiae* versus *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. U.S.A. 103*, 8149−8154.

(48) Bryce, R. A., Hillier, I. H., and Naismith, J. H. (2001) Carbohydrate-protein recognition: Molecular dynamics simulations and free energy analysis of oligosaccharide binding to concanavalin A. *Biophys. J. 81*, 1373−1388.
(49) Ford, M. G., Weimar, T., Kohli, T., and Woods, R. J. (2003) Molecular dynamics simulations of galectin-1-oligosaccharide complexes reveal the molecular basis for ligand diversity. *Proteins: Struct., Funct., Genet. 53*, 229−240.