JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Subscriber access provided by ISTANBUL TEKNIK UNIV

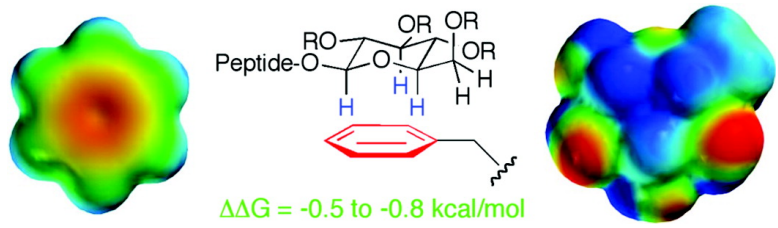
Article

Carbohydrate## Interactions: What Are They Worth?

Zachary R. Laughrey, Sarah E. Kiehna, Alex J. Riemen, and Marcey L. Waters

J. Am. Chem. Soc., **2008**, 130 (44), 14625-14633 • DOI: 10.1021/ja803960x • Publication Date (Web): 10 October 2008

Downloaded from http://pubs.acs.org on February 8, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

ACS Publications
High quality. High impact.

Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

J|A|C|S
ARTICLES

Published on Web 10/10/2008

Carbohydrate−π Interactions: What Are They Worth?

Zachary R. Laughrey, Sarah E. Kiehna,‡ Alex J. Riemen, and Marcey L. Waters*

Department of Chemistry, CB 3290, University of North Carolina, Chapel Hill, North Carolina 27599

Received June 3, 2008; E-mail: mlwaters@email.unc.edu

Abstract: Protein−carbohydrate interactions play an important role in many biologically important processes. The recognition is mediated by a number of noncovalent interactions, including an interaction between the α-face of the carbohydrate and the aromatic side chain of the protein. To elucidate this interaction, it has been studied in the context of a β-hairpin in aqueous solution, in which the interaction can be investigated in the absence of other cooperative noncovalent interactions. In this β-hairpin system, both the aromatic side chain and the carbohydrate were varied in an effort to gain greater insight into the driving force and magnitude of the carbohydrate−π interaction. The magnitude of the interaction was found to vary from −0.5 to −0.8 kcal/mol, depending on the nature of the aromatic ring and the carbohydrate. Replacement of the aromatic ring with an aliphatic group resulted in a decrease in interaction energy to −0.1 kcal/mol, providing evidence for the contribution of CH−π interactions to the driving force. These findings demonstrate the significance of carbohydrate−π interactions within biological systems and also their utility as a molecular recognition element in designed systems.

Introduction

Many biological processes, including bacterial cell wall recognition, viral and bacterial infections, and fertilization, rely on carbohydrate−protein interactions.[1,2] Additionally, glycosylation as a post-translational modification affects the hydration and conformation of a protein.[3,4] Because of its significance in biology, understanding the driving force for binding of carbohydrates in water is an active area of research.[5−7] In addition to hydrogen bonding, a common feature of carbohydrate-binding proteins is the interaction of the α-face of the carbohydrate with the face of an aromatic side chain (Figure 1).[8,9] Carbohydrate−π interactions have been investigated through a variety of analytical techniques, including NMR, IR, molecular modeling, and X-ray analysis.[10−16] The Simons group utilized IR and molecular modeling to examine the interaction.[17] Jimenez-Barbero et al.[18−20] have used NMR and molecular modeling to examine the driving force for binding of oligosaccharides to the hevein domain and variations thereof. These studies indicate that

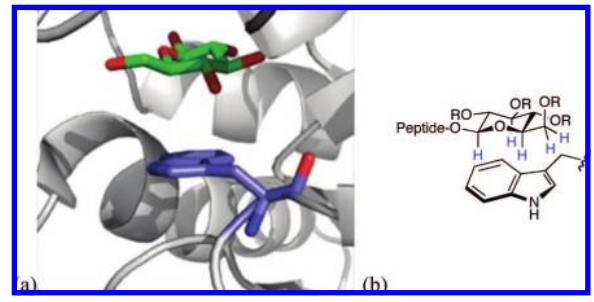


Figure 1. (a) Interaction between glucose and Trp 183 in the *Escherichia coli* chemoreceptor protein (PDB entry 2GBP).[9] (b) Interaction geometry for Trp and Ac4Glc in the context of a β-hairpin peptide.[21]

‡ Current address: Nanotope, Inc., 8025 Lamon Ave., Suite 450, Skokie, IL 60077.

(1) Gabius, H.-J.; Siebert, H.-C.; André, S.; Jiménez-Barbero, J.; Rüdiger, H. *ChemBioChem* **2004**, *5*, 740–764.
(2) Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357–2364.
(3) Liang, F.-U.; Chen, R. P.-Y.; Lin, C.-C.; Huang, K.-T.; Chan, S. I. *Biochem. Biophys. Res. Commun.* **2006**, *342*, 482–488.
(4) Bosques, C. J.; Tschampel, S. M.; Woods, R. J.; Imperiali, B. *J. Am. Chem. Soc.* **2004**, *126*, 8421–8425.
(5) Davis, A. P.; Wareham, R. S. *Angew. Chem., Int. Ed.* **1999**, *38*, 2978–2996.
(6) Klein, E.; Ferrand, Y.; Barwell, N. P.; Davis, A. P. *Angew. Chem., Int. Ed.* **2008**, *47*, 2693–2696.
(7) Ferrand, Y.; Crump, M. P.; Davis, A. P. *Science* **2007**, *318*, 619–622.
(8) Elgavish, S.; Shaanan, B. *Trends Biochem. Sci.* **1997**, *22*, 462–467.
(9) Vyas, N. K.; Vyas, M. N.; Quiocho, F. A. *Science* **1988**, *242*, 1290–1295.
(10) Jiménez-Barbero, J.; Asensio, J. L.; Cañada, F. J.; Poveda, A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 549–555.
(11) Bernardi, A.; Arosio, D.; Potenza, D.; Sánchez-Medina, I.; Mari, S.; Cañada, F. J.; Jiménez-Barbero, J. *Chem.— Eur. J.* **2004**, *10*, 4395–4406.
(12) Terraneo, G.; Potenza, D.; Canales, A.; Jiménez-Barbero, J.; Baldridge, K. K.; Bernardi, A. *J. Am. Chem. Soc.* **2007**, *129*, 2890–2900.
(13) Sujatha, M. S.; Sasidhar, Y. U.; Balaji, P. V. *Protein Sci.* **2004**, *13*, 2502–2514.
(14) Sujatha, M. S.; Sasidhar, Y. U.; Balaji, P. V. *Biochemistry* **2005**, *44*, 8554–8562.
(15) Vandenbussche, S.; Díaz, D.; Fernández-Alonzo, M. C.; Pan, W.; Vincent, S. P.; Cuevas, G.; Cañada, F. J.; Jiménez-Barbero, J.; Bartik, K. *Chem.— Eur. J.* **2008**, *14*, 7570–7578.
(16) Fernández-Alonso, M. C.; Cañada, F. J.; Jiménez-Barbero, J.; Cuevas, G. *J. Am. Chem. Soc.* **2005**, *127*, 7379–7386.
(17) Screen, J.; Stanca-Kaposta, E. C.; Gamblin, D. P.; Liu, B.; Macleod, N. A.; Snoek, L. C.; Davis, B. G.; Simons, J. P. *Angew. Chem., Int. Ed.* **2007**, *46*, 3644–3648.
(18) Asensio, J. L.; Siebert, H.-C.; von der Lieth, C.-W.; Laynez, J.; Bruix, M.; Soedjanaamadja, U. M.; Beintema, J. J.; Cañada, F. J.; Gabius, H.-J.; Jiménez-Barbero, J. *Proteins* **2000**, *40*, 218–236.
(19) Aboitiz, N.; Vila-Perelló, M.; Groves, P.; Asensio, J. L.; Andreu, D.; Cañada, F. J.; Jiménez-Barbero, J. *ChemBioChem* **2004**, *5*, 1245–1255.
(20) Chávez, M. I.; Andreu, C.; Vidal, P.; Aboitiz, N.; Freire, F.; Groves, P.; Asensio, J. L.; Asensio, G.; Muraki, M.; Cañada, F. J.; Jiménez-Barbero, J. *Chem.— Eur. J.* **2005**, *11*, 7060–7074.

 © 2008 American Chemical Society

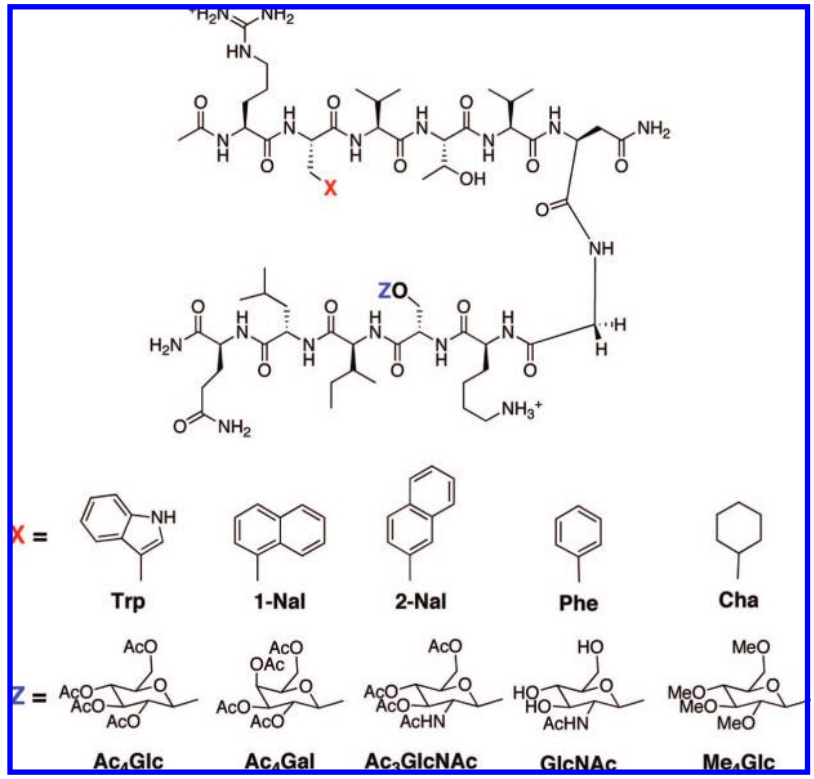


Figure 2. Structures of the β-hairpin and the **X** and **Z** side chains.

carbohydrate$-\pi$ interactions are important for the recognition of carbohydrates and that these interactions are dependent on the electronic nature of the aromatic group. However, there is a limited amount of experimental data investigating the favorable contribution of carbohydrate$-\pi$ interactions in isolation.[15,16] Given the importance of carbohydrate recognition in biology, a better understanding of the role of carbohydrate$-\pi$ interactions is warranted.

We previously reported an attractive interaction between a tryptophan and a diagonally cross-strand tetraacetylglucoserine [Ser(Ac_4Glc)] that stabilized the folding of a β-hairpin.[21] An examination of the proton NMR shifts of the carbohydrate protons demonstrated that the interaction was primarily through the α-face of the carbohydrate and the face of the Trp side chain, suggesting a carbohydrate$-\pi$ interaction. Upfield shifting and NOE data were consistent with the geometry shown in Figure 1b. The magnitude of the interaction was found to be -0.8 kcal/mol, which is greater than the magnitude of a Lys$-$Trp cation$-\pi$ interaction measured in the same model system.[22] However, when the acetyl groups on the glucose were removed, the interaction between the carbohydrate and tryptophan was lost. The reduction was attributed to the increased desolvation cost of the unprotected glucose.

To further investigate the efficacy of an isolated carbohydrate$-\pi$ interaction in aqueous solution, two series of peptides were synthesized and studied. In the first series, aromatic or aliphatic residues (**X**) were incorporated in close proximity to a Ser(Ac_4Glc) on the face of a β-hairpin (Figure 2). In the second series, Trp was used as the aromatic residue at position **X** and the nature of the carbohydrate (**Z**) in Ser(**Z**) was varied.

Results and Discussion

Design. The 12-residue sequence used in this study was based on previously described peptides in which a stabilizing carbohydrate$-\pi$ interaction between a Trp at position 2 and tetraacylated glucose at position 9 was explored.[21] Several key features were maintained, including a +3 charge to provide solubility and discourage aggregation, an Asn$-$Gly turn nucleator sequence, and hydrophobic clusters on both the HB face (Val 3, Val 5, and Ile 10) and the NHB face (**X** 2 and Leu 11) of the hairpin. Aromatic amino acids and carbohydrates were placed in positions 2 (**X**) and 9 (**Z**), respectively. These positions have been shown to allow diagonal cross-strand interactions and provide ample room for the bulky side chains.[22−25] The glycosylated series were synthesized via literature methods and introduced into the peptide chain as Fmoc-protected amino acids (see the Supporting Information).[21,26−31]

Characterization of Structure. β-Hairpin structure characterization was accomplished by a number of NMR measurements, including carbohydrate chemical shifts, α-hydrogen (Hα) chemical shifts, glycine splitting, and cross-strand nuclear

(21) Kiehna, S. E.; Laughrey, Z. R.; Waters, M. L. *Chem. Commun.* **2007**, 4026–4028.
(22) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2004**, *126*, 2028–2034.
(23) Hughes, R. M.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 6518–6519.
(24) Hughes, R. M.; Waters, M. L. *J. Am. Chem. Soc.* **2006**, *128*, 13586–13591.
(25) Tatko, C. D.; Waters, M. L. *Protein Sci.* **2003**, *12*, 2443–2452.
(26) Salvador, L. A.; Elofsson, M.; Kihlberg, J. *Tetrahedron* **1995**, *51*, 5643–5656.
(27) Macmillan, D.; Daines, A. M.; Bayrhuber, M.; Slitsch, S. L. *Org. Lett.* **2002**, *4*, 1467–1470.
(28) Takasu, A.; Houjyou, T.; Inai, Y.; Hirabayashi, T. *Biomacromolecules* **2002**, *3*, 775–782.
(29) Wang, H.; Sun, L.; Glazeback, S.; Zhao, K. *Tetrahedron Lett.* **1995**, *36*, 2953–2956.
(30) Mitchell, S. A.; Pratt, M. R.; Hruby, V. J.; Polt, R. *J. Org. Chem.* **2001**, *66*, 2327–2342.
(31) Sjolin, P.; Kihlberg, J. *J. Org. Chem.* **2001**, *66*, 2957–2965.

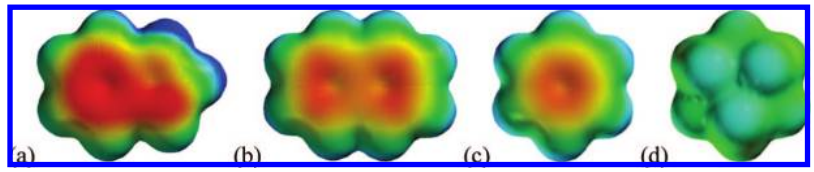


Figure 3. Electrostatic potential maps of the side chains at position **X** in the β-hairpin: (a) indole; (b) naphthalene; (c) benzene; (d) cyclohexane. The maps were generated using MacSpartan at the HF/6-31G* level, with an isodensity value of 0.02 and a range of −25 kcal/mol (red, electron-rich) to 25 kcal/mol (blue, electron-poor).

Table 1. Fraction Folded and $\Delta G°$(folding) at 298 K for β-Hairpins[a]

X	Z	$\Delta\delta_{Gly}$ (ppm)[b]	fraction folded (Gly splitting)[c]	fraction folded (Hα)[d]	$\Delta G°$(folding) (kcal mol^{-1})[e]
Trp[f]	Ac_4Glc	0.484	0.85	0.83 (0.02)	−1.03
1-Nal	Ac_4Glc	0.484	0.86	0.83 (0.03)	−1.08
2-Nal	Ac_4Glc	0.471	0.83	0.81 (0.04)	−0.94
Phe	Ac_4Glc	0.325	0.57	0.51 (0.13)	−0.17
Cha	Ac_4Glc	0.242	0.45	0.42 (0.12)	0.12

[a] Conditions: 50 mM sodium acetate-d_4, pH 4.0 (uncorrected) at 298 K, referenced to DSS. [b] Error is ±0.005 ppm. [c] Error is ±0.01. [d] The Hα fraction folded was determined from the average of the values for Val 3, Val 5, Lys 8, and Ile 10. Standard deviations are shown in parentheses. [e] $\Delta G°$ was determined using glycine-splitting values; the error is ±0.05 kcal/mol. [f] Previously reported in ref 21.

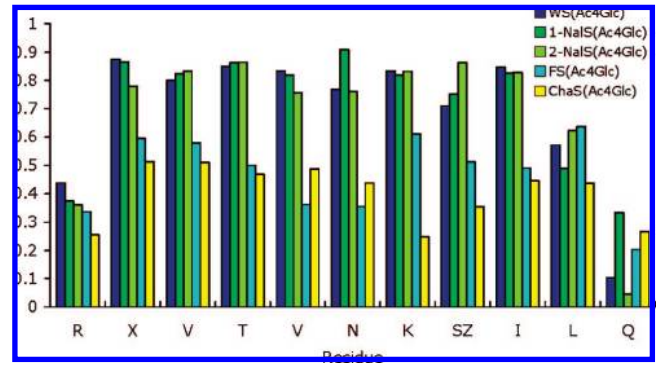


Figure 4. Fraction folded as determined from Hα chemical shifts. Values for WS(Ac_4Glc) were originally reported in ref 21.

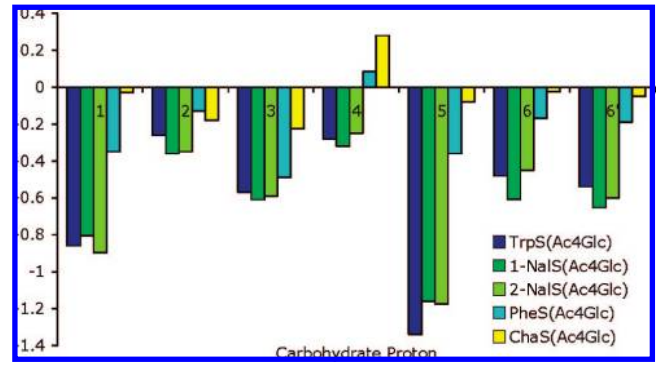


Figure 5. (a) Upfield shifting of carbohydrate protons in peptides WS(Ac_4-Glc), 1-NalS(Ac_4Glc), 2-NalS(Ac_4Glc), PheS(Ac_4Glc), and ChaS(Ac_4Glc). Conditions: 50 mM sodium acetate-d_4, pH 4.0 (uncorrected) at 298 K, referenced to DSS. Values for WS(Ac_4Glc) were originally reported in ref 21.

Overhauser effects (NOEs), as described below. NMR spectroscopy provides insight into the geometry of the interaction, as the carbohydrate protons are shifted upfield when in close proximity to the face of the aromatic side chains because of ring-current effects.[32] The extent of downfield shifting of Hα relative to the random coil is an indicator of the extent of β-sheet conformation at each position along the strand. Downfield shifting of Hα by >0.1 ppm is indicative of β-hairpin formation.[33] The fraction folded at each residue can be determined by comparing the observed Hα chemical shifts to those in the unfolded and fully folded states (as obtained from an unfolded control peptide and a cyclic control peptide, respectively; see the Experimental Section).[56] Alternatively, glycine splitting, when compared to a cyclic control, acts as a global indicator of β-hairpin conformation.[34] Fraction-folded values determined from Hα shifting and Gly splitting were generally in good agreement. Finally, long-distance cross-strand NOEs between side chains are consistent with β-hairpin structure and were observed for all peptides.

Variation of the Aromatic Side Chain (X). To examine the role of the aromatic side chain, **X**, a series of peptides were synthesized in which Trp was replaced with other aromatic or hydrophobic side chains, including 1-Nal, 2-Nal, Phe, and Cha (Figure 2), while the carbohydrate **Z** was maintained as Ac_4Glc.

1-Nal was substituted for Trp to investigate the significance of the NH group in Trp. In addition, 1-Nal has a greater surface area than Trp (161 Å^2 for 1-Nal compared with 147 Å^2 for Trp), and the electron density on the face of the ring is not as great (Figure 3). This substitution results in a well-folded peptide that is folded to the same extent as WS(Ac_4Glc) within experimental error (Table 1 and Figure 4). The protons on the α-face (H1, H3, and H5) are all shifted upfield, indicating that the α-face packs against the aromatic face of 1-Nal (Figure 5), similar to the shifting observed for WS(Ac_4Glc). Protons H6 and H6′ were also found to be shifted to a lesser extent, indicating that the exocyclic CH_2 interacts with the aromatic face. This interaction and geometry has precedent in galactose-binding lectins. While the acetyl groups could not be assigned, the maximum upfield shift was ≤0.07 ppm (assuming that the methyl group peak that was farthest downfield in the control peptide was the farthest upfield in the β-hairpin). This is significantly less than the shifts of the protons of the α-face of the carbohydrate, which ranged from −0.6 to −1.2 ppm, and indicates that the acetyl groups play little or no direct role in the stabilizing interaction. NOESY NMR displayed long-distance cross-strand interactions in the peptide, indicating that the peptide is properly folded in a β-hairpin structure. Strong NOEs were also observed between the carbohydrate and 1-Nal (Figure 6), although not as extensively as seen with Trp.

Because the β-sheet propensity of each amino acid influences the stability of the folded state, one cannot directly compare the extents of folding of two peptides in which **X** has been varied and attribute differences exclusively to side chain−side chain interactions. To determine the energetic contribution of the

(32) Pople, J. A. *J. Chem. Phys.* **1956**, *24*, 1111.
(33) Maynard, A. J.; Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 1996–2007.
(34) Searle, M. S.; Griffiths-Jones, S. R.; Skinner-Smith, H. *J. Am. Chem. Soc.* **1999**, *121*, 11615–11620.

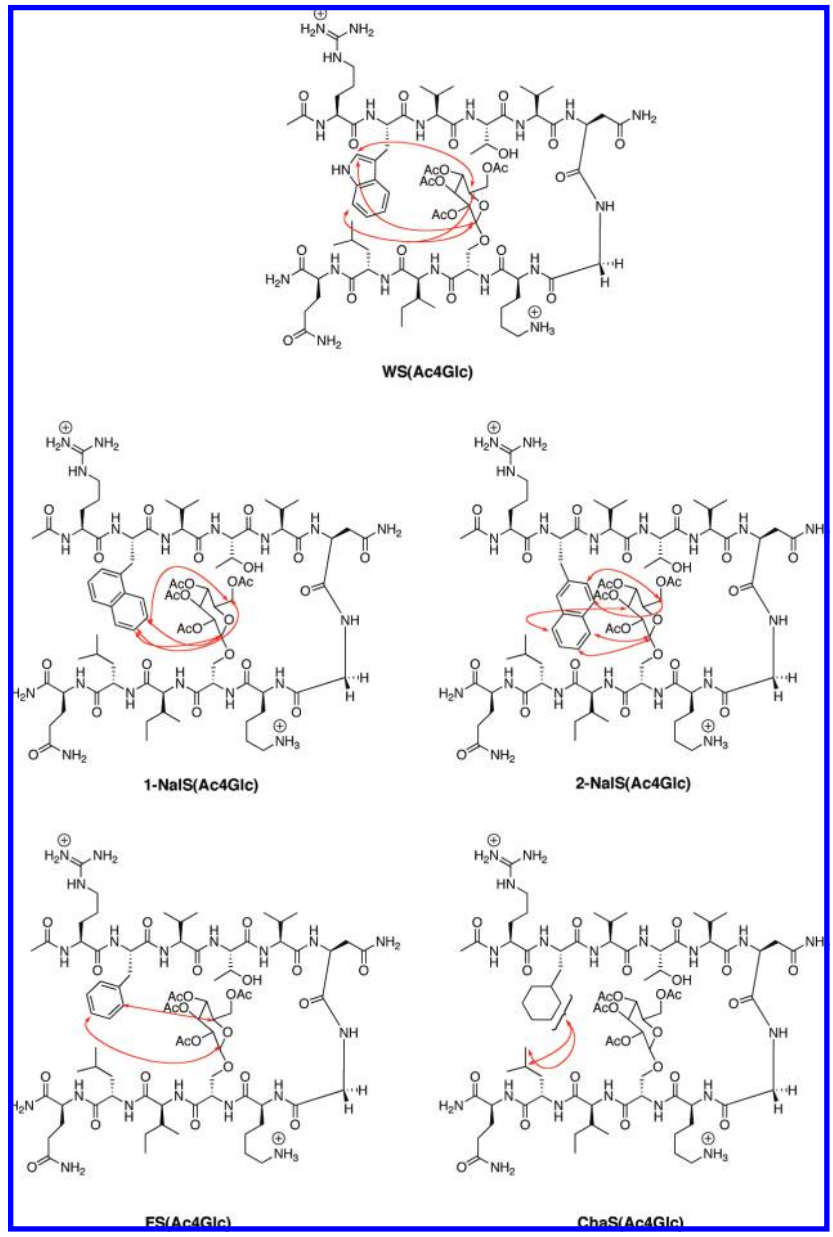


Figure 6. Unambiguous NOEs (red arrows) observed between **X** (Trp, 1-Nal, 2-Nal, Phe, or Cha) and the carbohydrate side chain. Unambiguous NOEs are defined as ones that can be definitively assigned to a particular set of protons.

diagonal side chain−side chain interaction alone, a double mutant cycle was completed.[33−39] A double mutant cycle replaces two interacting side chains with two noninteracting ones. A single mutation disrupts the interaction of interest but could additionally cause other changes to stability (i.e., β-sheet propensity, hydrogen bonding, etc.). The double mutant corrects for these unintentional changes, leaving only the noncovalent interaction of interest. In this study, **X** and Ser(**Z**) were exchanged for Leu and Ser, respectively. Leu was chosen because it has a high β-sheet propensity that minimizes the net loss of β-hairpin stability. Ser was chosen because it has a small polar side chain that does not interact diagonally. The double-mutant cycle revealed that the interaction of 1-Nal with Ser(Ac_4Glc) has a magnitude similar to that of the Trp−Ser(Ac_4Glc) interaction (Table 2).

Table 2. Diagonal Interaction Energies between Residues 2 and 9 As Determined by Double Mutant Cycles

X	Z	$\Delta\Delta G$ (kcal mol^{-1})[a]
Trp[b]	Ac_4Glc	−0.8
1-Nal	Ac_4Glc	−0.7
2-Nal	Ac_4Glc	−0.7
Phe	Ac_4Glc	−0.5
Cha	Ac_4Glc	−0.1

[a] Error is ±0.1 kcal mol^{-1}. [b] Previously reported in ref 21.

2-Nal was substituted at the **X** position to determine the influence of the orientation difference between 1-Nal and 2-Nal on the carbohydrate−π interaction. The results were similar to

(35) Serrano, L.; Bycroft, M.; Fersht, A. R. *J. Mol. Biol.* **1991**, *218*, 465–475.
(36) Schreiber, G.; Fersht, A. R. *J. Mol. Biol.* **1995**, *248*, 478–486.
(37) Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 5291–5300.
(38) Blanco, F. J.; Serrano, L. *Eur. J. Biochem.* **1995**, *230*, 634–649.
(39) Carver, F. J.; Hunter, C. A.; Jones, P. S.; Livingstone, D. J.; McCabe, J. F.; Seward, E. M.; Tiger, P.; Spey, S. E. *Chem.— Eur. J.* **2001**, *7*, 4854–4862.

Table 3. Thermodynamic Parameters for Folding at 298 K Obtained from Thermal Denaturation of the Peptides

peptide	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} K^{-1})	ΔC_p (cal mol^{-1} K^{-1})
WS(Ac_4Glc)[a]	−5.9	−16.4	−112
FS(Ac_4Glc)	−4.23	−13.77	−77
ChaS(Ac_4Glc)	−2.96	−10.32	−88

[a] Previously reported in ref 21.

those for both 1-Nal and Trp, indicating that the carbohydrate interacts in a favorable manner with 2-Nal via stacking with the aromatic side chain despite the differences in orientation of the two side chains (Table 1 and Figure 4). The upfield shifting of the peptide Hα protons demonstrates that the β-hairpin is well-formed throughout the peptide with a stability similar to those of WS(Ac_4Glc) and 1-NalS(Ac_4Glc). Additionally, the carbohydrate shifts are similar in magnitude to those for the other two peptides with large aromatic groups (Figure 5): the protons of the carbohydrate's α-face are shifted significantly upfield, indicating an interaction with the aromatic face. Long-distance NOEs between the carbohydrate's α-face and 2-Nal also support a stacking geometry (Figure 6). The magnitude of the interaction of 2-Nal with Ac_4Glc also was found to be similar to those involving Trp and 1-Nal, as determined from the double mutant cycle (Table 2).

The situation changed when the smaller Phe group was placed in the sequence at position 2. The fraction folded decreases from 0.85 to 0.57 (Figure 4 and Table 1). Phe is known to have a lower β-sheet propensity than Trp,[40,57] but the double mutant cycle indicates that the loss in β-hairpin stability is due in part to a weakening of the carbohydrate−π interaction (Table 2). Additionally, the carbohydrate's α-face protons are not shifted upfield nearly as much as with the larger aromatic **X** groups. There are significantly fewer unambiguous NOEs between the Phe side chain and the face of the carbohydrate than between the larger aromatic side chains and the carbohydrate. The smaller changes in the chemical shifts of both the Hα and the carbohydrate protons indicate a less folded hairpin and a less favorable carbohydrate−π interaction.

When the aromatic nature of the **X** side chain was removed by replacing Phe with Cha, the stability of the hairpin was further reduced (Table 1), despite the fact that Cha has been shown to have a higher β-sheet propensity than Phe.[41] A double mutant cycle indicates that the interaction of Cha with Ac_4Glc is weaker than that of Phe with Ac_4Glc (Table 2), despite the similar facial surface area. In contrast to the results for the aromatic peptides in this series, there are no unambiguous NOEs between the cyclohexane side chain and the carbohydrate (Figure 6). Since Cha is not aromatic, no shifting of the carbohydrate protons is observed.

To provide additional insight into the effect of the **X** group on the interaction with Ac_4Glc, we performed thermal denaturations on WS(Ac_4Glc), FS(Ac_4Glc), and ChaS(Ac_4Glc) using variable-temperature NMR.[33] Fitting of the data provided $\Delta H°$, $\Delta S°$, and ΔC_p values for folding (Table 3). Since the only change in the peptide sequence is the **X** group at position 2, changes in the driving force for folding can be attributed to the role of that residue in stabilizing the folded state. One can see that for Trp and Phe, folding is enthalpically more favorable than for Cha, which does not significantly interact with Ac_4Glc.

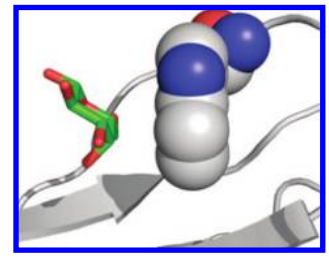


Figure 7. Trp−Gal interaction in the binding pocket of galectin.[43]

Table 4. Fraction Folded and $\Delta G°$(folding) for β-Hairpins at 298 K[a]

X	Z	$\Delta\delta_{Gly}$ (ppm)[b]	fraction folded (Gly splitting)[c]	fraction folded (Hα)[d]	$\Delta G°$(folding) (kcal/mol)[e]
Trp[f]	Ac_4Glc	0.484	0.85	0.83 (0.02)	−1.03
Trp	Ac_4Gal	0.466	0.82	0.77 (0.04)	−0.90
Trp	$Ac_3GlcNAc$	0.421	0.73	0.66 (0.09)	−0.59
Trp	GlcNAc	0.392	0.70	0.64 (0.02)	−0.50
Trp	Me_4Glc	0.50	0.88	0.81 (0.01)	−1.18
Trp[f]	Glc	0.383	0.65	0.63 (0.02)	−0.37
Trp	OH	0.375	0.64	0.60 (0.01)	−0.34

[a] Conditions: 50 mM sodium acetate-d_4, pH 4.0 (uncorrected) at 298 K, referenced to DSS. [b] Error is ±0.005 ppm. [c] Error is ±0.01. [d] The Hα fraction folded was determined from the average of the values for Val 3, Val 5, Lys 8, and Ile 10. Standard deviations are shown in parentheses. [e] $\Delta G°$ was determined using glycine-splitting values; the error is ±0.05 kcal/mol. [f] Previously reported in ref 21.

This is consistent with an enthalpic driving force for the interaction of the carbohydrate with the aromatic ring, as has been observed in other systems,[5,42] and is suggestive of CH(δ+)−π and dispersion forces as major contributors to the interaction.

Variation of the Carbohydrate. The carbohydrate was also varied while Trp was used as **X**, and the resulting impact on the side chain−side chain interaction was explored. Previous studies comparing Ac_4Glc and Glc suggested that Ac_4Glc formed a favorable carbohydrate−π interaction, but the desolvation cost appeared to be too high for Glc to interact favorably with Trp.[21] In this work, two other acetylated carbohydrates, Ac_4Gal and $Ac_3GlcNAc$ (Figure 2) were substituted for Ac_4Glc. Ac_4Gal was used to examine the effect of the stereochemistry at C4 on the carbohydrate−π interaction, and $Ac_3GlcNAc$ was chosen to investigate the replacement of oxygen with nitrogen at C2. We also studied the deprotected counterpart, GlcNAc, in which only the nitrogen at C2 is acetylated. Lastly, we investigated Me_4Glc to further explore the role of desolvation and determine the role (if any) of the acetyl groups.

The only difference between Ac_4Glc and Ac_4Gal is the orientation of the alcohol at C4 (equatorial vs axial, repectively) (Figure 2). The binding sites of many galactose-binding proteins (galectins) contain an aromatic residue that interacts with the "hydrophobic cluster" made up of C4, C5, and C6 (Figure 7).[43] The upfield shift of the 6/6′ protons of Ac_4Glc in WS(Ac_4Glc) suggests that such an interaction at C4, C5, and C6 may be feasible in the β-hairpin. Thus, we replaced Ac_4Glc with Ac_4Gal and investigated its interaction with Trp. There is only a small change in the fraction folded for WS(Ac_4Gal) relative to WS(Ac_4Glc), as measured by glycine splitting and Hα shifts (Table 4 and Figure 8). NOEs between the sugar and Trp indicate that the interaction occurs on the α-face of the sugar,

(40) Minor, D. L. J.; Kim, P. S. *Nature* **1994**, *371*, 264–267.
(41) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 9372–9373.
(42) Lemieux, R. U. *Acc. Chem. Res.* **1996**, *29*, 373–380.
(43) Leonidas, D. D.; Vatzaki, E. H.; Vorum, H.; Celis, J. E.; Madsen, P.; Acharya, K. R. *Biochemistry* **1998**, *37*, 13930–13940.

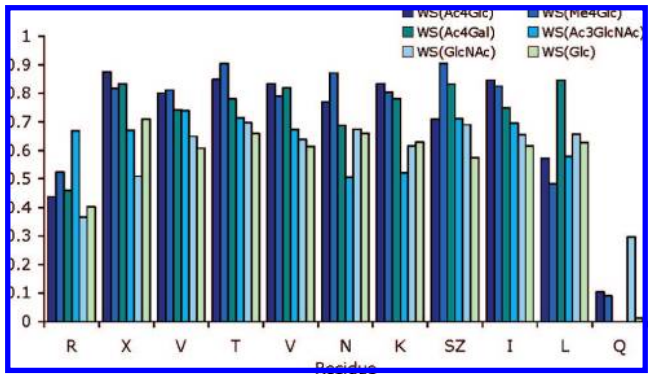


Figure 8. Fraction folded as determined from Hα chemical shifts. Values for WS(Ac_4Glc) were originally reported in ref 21.

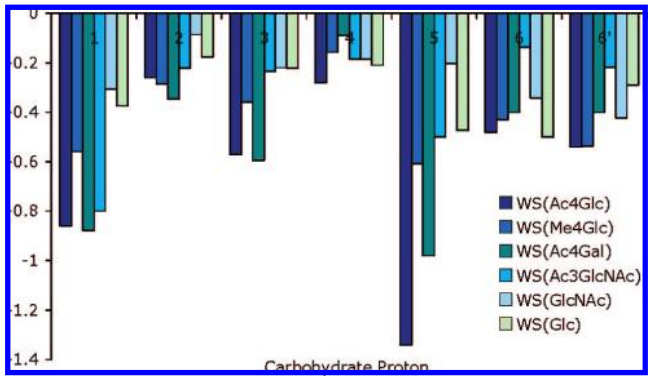


Figure 9. Upfield shifting of carbohydrate protons in peptides. Values for WS(Ac_4Glc) were originally reported in ref 21.

Table 5. Diagonal Interaction Energies As Determined by Double Mutant Cycles

X	Y	$\Delta\Delta G$ (kcal mol^{-1})[a]
Trp[b]	Ac_4Glc	−0.8
Trp	Ac_4Gal	−0.7
Trp	Ac_3GlcNAc	−0.6[c]
Trp	GlcNAc	−0.5[c]
Trp	Me_4Glc	−0.8[c]

[a] Error is ±0.1 kcal mol^{-1}. [b] Previously reported in ref 21. [c] Although the interaction energy for this mutant is similar to that of the Trp−Ac_4Glc interaction, the NMR data suggest that an interaction other than the carbohydrate−π interaction is contributing. See text for details.

as was seen for Ac_4Glc. Inspection of the carbohydrate chemical shifts reveals that the protons of the α-face (C1, C3, and C5) are shifted by the greatest amount relative to a random coil and that C4 is not shifted significantly (Figure 9), indicating the same geometry as for WS(Ac_4Glc) rather than that seen in galectins. This may be due to conformational restrictions of the system rather than a specific preference for one geometry over the other. The extent of shifting at positions 1, 3, and 5 of Ac_4Gal is similar to that observed in Ac_4Glc, suggesting that the interaction with the α-face is equally as favorable. This is consistent with the interaction energy determined from double mutant cycles (Table 5), which is within the experimental error of that measured for Ac_4Glc.

Another common carbohydrate found in nature is GlcNAc. This carbohydrate has two distinctive features: nitrogen replaces oxygen at C2, and this nitrogen is acetylated. Both of these differences affect how the carbohydrate interacts with the face of Trp. The tetraacetylated sugar Ac_3GlcNAc was used for direct comparison with Ac_4Glc. The presence of the amide reduces the interaction relative to WS(Ac_4Glc): the fraction folded decreases from 0.85 to 0.73 (Figure 8 and Table 4), and the double mutant cycle demonstrates that the interaction energy is reduced by 0.2 kcal mol^{-1} (Table 5). However, the protons of the Ac_3GlcNAc α-face are not nearly as upfield-shifted as those of WS(Ac_4Glc) (Figure 9). The proton at C1 is the only one that was shifted significantly, suggesting a change in geometry due to the amide at C2. This in turn suggests that the energetic term from the double mutant cycle arises from a favorable interaction other than the carbohydrate−π interaction. Indeed, the downfield shifting of the NH group of Ac3GlcNAc (9.34 ppm) suggests that it may participate in a hydrogen bond. The Trp NH is not significantly shifted, however [10.16 ppm in WS(Ac_3GlcNAc) vs 10.11 ppm in an unfolded control peptide].

When the acetyl groups were removed from the Ac_3GlcNAc, the fraction folded decreased only slightly (0.73 for Ac_3GlcNAc vs 0.70 for GlcNAc). However, the carbohydrate protons are significantly less shifted than those of the other carbohydrates, with the greatest shift occurring at C6 (Figure 9). In fact, the chemical shifts of GlcNAc are similar to those of Glc, which do not display any significant interaction with Trp.[21] There are several weak unambiguous NOEs between Trp and the α-face of the carbohydrate (Figure 10). The double mutant cycle indicates that the interaction energy for GlcNAc is comparable to that for Ac_3GlcNAc. Thus, it appears that some sort of favorable interaction is present, but it occurs via a different geometry than seen with other carbohydrates. However, NMR data provides no evidence of hydrogen bonding involving either the GlcNAc NH (7.82 ppm in this peptide vs 7.88 ppm in the unfolded control peptide) or the Trp NH (10.17 ppm in this peptide vs 10.11 ppm in the unfolded control peptide).

We also investigated the peptide in which the acetyl protecting groups of Ac_4Glc were replaced with methyl groups to probe the role of desolvation and determine if there is a specific influence of the acetyl groups. The peptide WS(Me_4Glc) is equally well folded as WS(Ac_4Glc) (Table 4, Figure 8) and exhibits numerous NOEs between the Trp residue and the α-face of the sugar (Figure 10), indicating that Me_4Glc also forms a favorable interaction with Trp. Double mutant cycles indicate that the magnitudes of the interactions for Me_4Glc and Ac_4Glc are within the experimental error of each other. This appears to suggest that protection of the hydroxyl groups, and hence reduction of the desolvation cost, is indeed the primary reason for the difference between the carbohydrate−π interactions for peptides WS(Ac_4Glc) and WS(Me_4Glc) and that for WS(Glc) containing the unprotected glucose. However, Ac_4Glc and Me_4Glc do not behave identically: peptide WS(Me_4Glc) does not demonstrate the same extent of upfield shifting of the carbohydrate protons at positions C1, C3, and C5 as does peptide WS(Ac_4Glc) [0.4−0.6 ppm for WS(Me_4Glc) versus 0.6−1.35 ppm for WS(Ac_4Glc)], despite the similar stability of the β-hairpins. This may be due to competition of the methyl groups for interaction with the Trp, as the methyl groups were also shifted upfield by as much as 0.25 ppm. Indeed, the magnitude of the upfield shift of the methyl groups was very similar to that observed by Cuevas and co-workers[44] in their study of carbohydrate−π interactions with Me_5Glc. Thus, it appears that while Me_4Glc forms a favorable interaction with Trp, the interaction of the α-face of the sugar is not the only contributor to the interaction.

(44) Bautista-Ibáñez, L.; Ramírez-Gualito, K.; Quiroz-García, B.; Rojas-Aguilar, A.; Cuevas, G. *J. Org. Chem.* **2008**, *73*, 849–857.

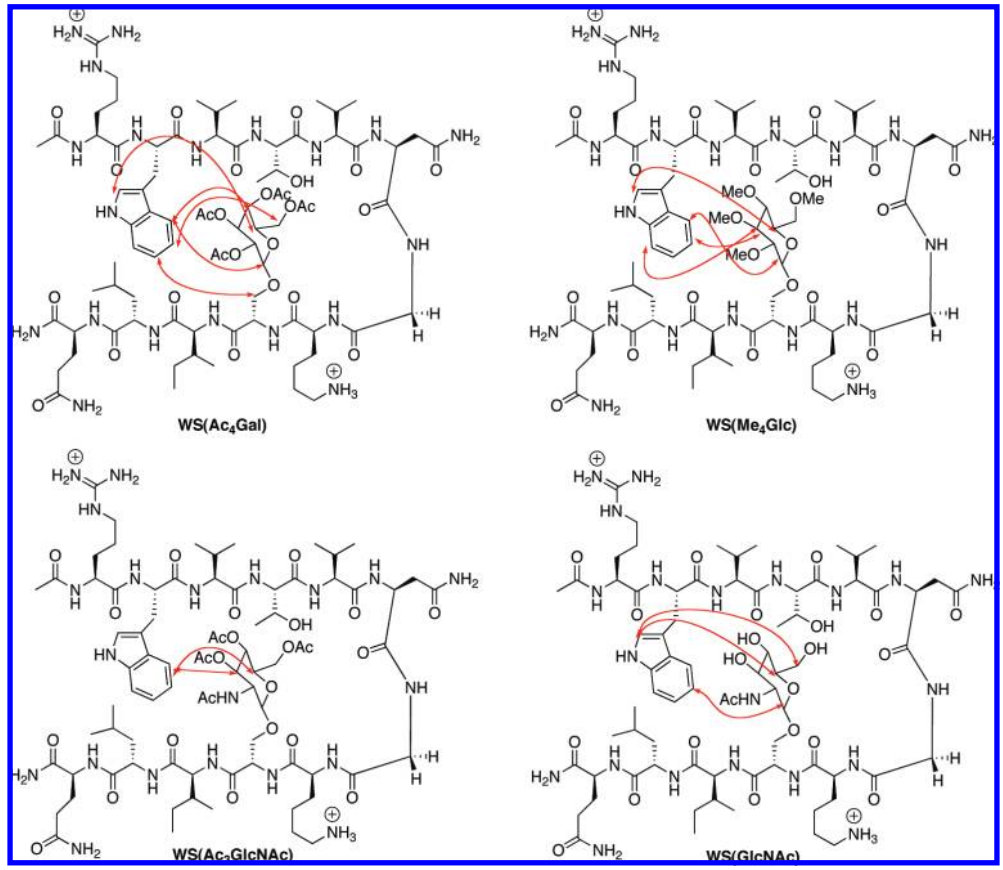


Figure 10. Unambiguous NOEs (red arrows) observed between the carbohydrate side chain (Ac_4Gal, $Ac_3GlcNAc$, GlcNAc, or Me_4Glc) and Trp. Unambiguous NOEs are defined as ones that can be definitively assigned to a particular set of protons.

Discussion

The system described here has made possible the systematic investigation of carbohydrate−π interactions in the absence of other cooperative noncovalent interactions, such as hydrogen bonds, thus allowing for the quantification of the binding energy and an exploration of the factors that contribute to these interactions. Variation of the **X** side chain provides significant insight into the nature and driving force of the carbohydrate−π interaction. The similar interactions of Ac_4Glc with Trp, 1-Nal, and 2-Nal confirm that Ac_4Glc interacts primarily with the face of the aromatic ring and that any hydrogen bonding to the NH of Trp is at best a minor contributor to the interaction (≤0.1 kcal/mol). The fact that 1-Nal and 2-Nal interact similarly indicates that this model system has enough flexibility to optimize the interaction when the orientation of the aromatic ring is varied. Comparison of Trp to Phe indicates that the surface area of the aromatic ring impacts the magnitude of the interaction (−0.8 kcal/mol for Trp vs −0.5 kcal/mol for Phe).

The interaction between carbohydrates and aromatic groups has been variously described in terms of the hydrophobic effect, dispersion forces, and CH−π interactions.[20,45,46] Comparison of Phe versus Cha at position **X** indicates that the carbohydrate−π interaction is more favorable than an equivalent hydrophobic interaction between Ac_4Glc and an aliphatic side chain (−0.5 kcal/mol for Phe vs −0.1 kcal/mol for Cha). This is similar to what has been seen in protein mutation studies, in which mutation of an aromatic residue in the binding pocket abolishes binding of the carbohydrate.[20,47] Because naturally occurring aliphatic residues have different sizes and shapes than aromatic residues, the results from protein mutation studies have been difficult to attribute solely to the loss of aromaticity. Since Phe and Cha have the same facial surface area, the comparison within the β-hairpin model system is more direct, and it clearly indicates that aromaticity influences the interaction energy. The preference for interaction of Ac_4Glc with Phe over Cha and the greater enthalpic driving force for folding of FS(Ac_4Glc) relative to ChaS(Ac_4Glc) point to CH(δ+)−π interactions as a significant contributor to the driving force of the interaction: since cyclohexane is more polarizable than benzene, dispersion forces would be expected to be stronger for Cha than for Phe.[48] Moreover, Cha is also more hydrophobic than Phe, arguing against the hydrophobic effect as the primary driving force for the interaction. This is consistent with the finding of Jiménez-Barbero et al.[20] that variation of the electronic structure of the aromatic ring influences carbohydrate binding in the hevein domain.

Variation of the carbohydrate provides insight into the balance of features that influence this interaction. Within the β-hairpin model system, it appears that interaction on the α-face of the carbohydrate is most favorable, even when another "hydrophobic" surface is present, as in Ac_4Gal. For $Ac_3GlcNAc$ and GlcNAc, the interaction energy decreased and the geometry of

(45) Muraki, M.; Morii, H.; Harata, K. *Protein Pept. Lett.* **1998**, *5*, 193–198.
(46) Muraki, M.; Morii, H.; Harata, K. *Protein Eng.* **2000**, *13*, 385–389.
(47) Flint, J.; Bolam, D. N.; Nurizzo, D.; Taylor, E. J.; Williamson, M. P.; Walters, C.; Davies, G. J.; Gilbert, H. J. *J. Biol. Chem.* **2005**, *280*, 23718–23726.
(48) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303–1324.

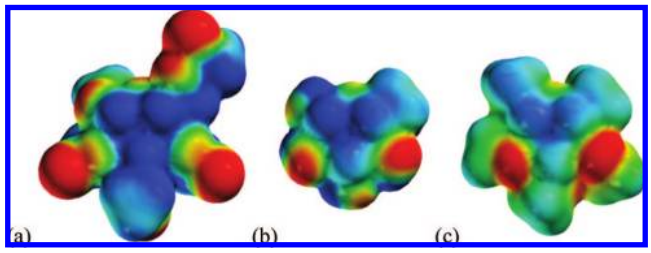


Figure 11. Electrostatic potential maps of the side chains at position **Z** in the β-hairpin: (a) Ac_4Glc; (b) Glc; (c) Me_4Glc. The maps were generated using MacSpartan at the HF/6-31G* level, with an isodensity value of 0.02 and a range of −25 kcal/mol (red, electron-rich) to 25 kcal/mol (blue, electron-poor).

the interaction changed, likely because of the presence of the amide at C2, which is expected to have stronger interactions with the solvent than the corresponding ester does.

A comparison of Ac_4Glc to Me_4Glc was made to address the roles of desolvation and electrostatics in the interaction. We have previously shown that Ac_4Glc interacts favorably with Trp, with an interaction energy of approximately −0.8 kcal/mol, and that removal of the acetyl groups results in the loss of the favorable interaction. We attributed this to differences in desolvation cost, although the electron-withdrawing nature of the acetyl groups also results in differences in the partial charges on the α-faces of Ac_4Glc and Glc, as indicated by electrostatic potential maps (Figure 11). Thus, we investigated the interaction of Me_4Glc with Trp because it has an electrostatic potential map similar to that of Glc but its desolvation cost is significantly reduced. The observed stabilizing interaction ($\Delta\Delta G = -0.8$ kcal/mol) indicates that paying the desolvation cost is indeed enough to allow for a favorable carbohydrate−π interaction and that there is nothing unique about the acetylated glucose. However, the NMR shifts of Me_4Glc indicate that interaction of Trp with the α-face is reduced and that direct interaction with the polarized methyl groups also occurs. Thus, a direct comparison of the role of electrostatics in the interactions of Trp with Ac_4Glc and Me_4Glc is not possible, as there are other contributors to the interaction energy. Nonetheless, the NMR data suggest that the weaker polarization of the α-face of the sugar may reduce the carbohydrate−π interaction and that the interaction between the methyl groups and Trp provides a compensating interaction.

Since the carbohydrate−π interaction is only observed in this system when the hydroxyl groups are protected, the question arises as to whether this interaction is significant in carbohydrate-binding proteins, where the carbohydrate is unprotected. We have shown that the role of the protecting groups is to desolvate the sugar to allow for interaction with the aromatic ring. Within a carbohydrate-binding protein, hydrogen-bonding groups are preorganized for the same task. Thus, it appears that nature uses cooperative interactions between the aromatic ring and hydrogen-bonding groups to desolvate and bind the carbohydrate.[42] Indeed, obtaining this sort of cooperative binding may be the primary challenge in designing synthetic receptors for carbohydrates in water.[7]

We have measured a wide range of noncovalent interactions within the same peptide model system, so a direct comparison can be made between them. Surprisingly, the energy of the interaction between Ac_4Glc and Trp is larger than that of the cation−π interaction between Lys and Trp (−0.4 kcal/mol)[22,25] but similar in magnitude to that of the interaction between KMe3 and Trp (−1.0 kcal/mol).[23] Thus, the carbohydrate−π interaction is considerable. In contrast, the interaction of Cha with Ac_4Glc has the same magnitude as its interaction with Lys (−0.1 kcal/mol).[22]

Lastly, these studies also provide evidence for a novel method of influencing protein structure. In structural studies of glycosylated proteins and peptides, it has generally been found that stabilization of the folded state occurs because glycosylation rigidifies the peptide backbone, thereby destabilizing the unfolded state.[49−55] In contrast, in the system reported here, incorporation of a carbohydrate−π interaction results in enthalpic stabilization of the folded structure through a specific interaction.

Conclusion

This study provides insight into the role of carbohydrate−π interactions in carbohydrate recognition by proteins. The energetic contribution of the carbohydrate−π interaction between the α-face of the pyranose ring and the face of an aromatic ring was found to range from −0.5 to −0.8 kcal mol^{-1} and is dependent on the nature of both the aromatic ring and the carbohydrate. Of significance is the fact that a favorable interaction was only observed when the hydroxyl groups of the carbohydrate were protected using either acetyl groups or methyl groups. This implies a significant cost for desolvation of the sugar. However, the NMR data suggests that the interaction with Ac_4Glc is more favorable than that with Me_4Glc, which may imply an electronic tuning of the interaction. Moreover, the preferential interaction of the pyranose ring with the face of a phenyl group as opposed to a cyclohexyl ring suggests that CH(δ+)−π interactions play a measurable role in the interaction. These studies provide a better physical understanding of the driving force behind carbohydrate−π interactions as well as insight into their magnitudes and significance relative to other noncovalent interactions that have been measured for the same model system. In addition to providing insight into the recognition of carbohydrates by proteins, we expect that the findings of this study will be useful in the development of new and improved receptors for carbohydrate recognition.

Experimental Section

Peptide Synthesis and Purification. All of the peptides were synthesized on Fmoc−PAL−PEG−PS amide resin using standard solid-phase protocols on a continuous flow Pioneer Peptide Synthesizer (Applied Biosystems). Fmoc-amino acids (4−6 equiv) were activated and coupled with 0.45 M HBTU/HOBt in dimethylformamide (DMF). The following protecting groups were used: Arg(Pbf), Asn(Trt), Cys(Trt), Gln(Trt), Lys(Boc), Ser(tBu), Thr(tBu), and Trp(Boc). Deprotection of the Fmoc groups was achieved with 2% piperidine/2% 1,8-diazabicyclo[5.4.0]undec-7-ene in DMF.

(49) Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 3352–3353.
(50) Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R. N.; Smolyar, A.; Reinherz, E. L.; Wagner, G. *Science* **1995**, *269*, 1273–1278.
(51) Live, D. H.; Kumar, R. A.; Beebe, X.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12759–12761.
(52) O'Connor, S. E.; Imperiali, B. *J. Am. Chem. Soc.* **1997**, *119*, 2295–2296.
(53) Imperiali, B.; O'Connor, S. E. *Chem. Biol.* **1998**, *5*, 427–437.
(54) Imperiali, B.; O'Connor, S. E. *Curr. Opin. Chem. Biol.* **1999**, *3*, 643–649.
(55) Bann, J. G.; Peyton, D. H.; Bächinger, H. P. *FEBS Lett.* **2000**, *473*, 237–240.
(56) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 11577–11578.
(57) Smith, C. K.; Regan, L. *Acc. Chem. Res.* **1997**, *30*, 153–161.

All of the peptides were acylated at the N-terminus using a 5% acetic anhydride/6% lutidine/DMF solution and amidated at the C-terminus. Peptide resin cleavage and deprotection were performed simultaneously by treatment with 95% trifluoroacetic acid (TFA)/2.5% triisopropylsilane/2.5% H_2O for 2−3 h under nitrogen. The TFA was removed by distillation under vacuum. The crude peptides were precipitated with cold ether, extracted into water, lyophilized, and then dissolved and purified by reverse-phase (RP) HPLC using a Vydac C_{18} semipreparative column. The peptides were eluted with a linear gradient of solvent A (95% H_2O/5% acetonitrile with 0.1% TFA) and solvent B (95% acetonitrile/5% water with 0.1% TFA) from 0 to 30% B and detected by monitoring at 220 and 280 nm. Molecular weights were determined using electrospray-ionization mass spectrometry. Disulfide bonds were formed by dimethyl sulfoxide oxidation of purified peptides in phosphate-buffered saline (pH 7.4). The peptides were then repurified by RP HPLC.

NMR Spectroscopy. Solutions for NMR spectroscopy with concentrations of 1−3 mM were analyzed on a Varian Inova 600 MHz instrument. Samples were dissolved in D_2O buffered with 10 mM acetate-d_3 (pH 4.2) and referenced to DSS. NMR spectra were collected with 8−64 scans using a 1.5 s presaturation. All of the 2D NMR experiments used pulse sequences from the Chempack software, including TOCSY, gCOSY, and ROESY. TOCSY and gCOSY experiments were performed with 4−8 scans in the first dimension and 256 scans in the second dimension. ROESY experiments were performed with 32 scans in the first dimension and 256−512 scans in the second dimension. All of the spectra were analyzed using standard window functions (sinebell and Gaussian with shifting). Assignments were made using standard methods. Thermal denaturations were performed in duplicate in 5−10 °C increments. The temperature was calibrated with methanol and ethylene glycol standards using Varian macros.

Determination of Fraction Folded. To determine the chemical shifts of the fully folded state, 14-residue disulfide-linked analogues of the peptides were synthesized with the sequence Ac−CR**X**VTVNGKS(**Z**)ILQC−NH_2, where X = Trp, 1-Nal, 2-Nal, Cha, or Phe and S(**Z**) = Ser(Ac_4Glc), Ser(Ac_4Gal), Ser(Ac_3GlcNAc), Ser(GlcNAc), or Ser(Me_4Glc), and characterized by NMR. To determine the chemical shifts of the unfolded state, a series of seven-residue peptides were synthesized and characterized. The sequences of these peptides were Ac−RXVTVNG−NH_2 (X = 1-Nal, 2-Nal, Cha, or Phe) and Ac−NGKZILQ−NH_2 [Z = Ser(Ac_4Gal), Ser(Ac_3GlcNAc), Ser(GlcNAc), or Ser(Me_4Glc)]. The 7-mers with either Trp or Ser(Ac_4Glc) had previously been described.[21] The fraction folded based on Hα chemical shifts was determined using eq 1:

$$\text{fraction folded} = \frac{\delta_{obs} - \delta_0}{\delta_{100} - \delta_0} \tag{1}$$

where δ_{obs} is the observed chemical shift, δ_{100} is the chemical shift for the cyclic peptide, and δ_0 is the chemical shift for the unfolded control peptide. The fraction folded based on glycine splitting was determined with the eq 2:

$$\text{fraction folded} = \frac{(\Delta\delta_{Gly})_{hairpin}}{(\Delta\delta_{Gly})_{cycle}} \tag{2}$$

Double Mutant Cycles. Double-mutant cycles were performed in order to quantify the interaction between the series of carbohydrates and the side chain **X**. Single mutant peptides in which Ser(Ac_4Glc), Ser(Ac_4Gal), Ser(Ac_3GlcNAc), Ser(GlcNAc), or Ser(Me_4Glc) was replaced by Ser or 1-Nal, 2-Nal, Phe, or Cha was replaced by Leu were prepared. The double mutant contained both substitutions. The singly-mutated peptides RWVTVNGKSILQ and RLVTVNGKS(Ac_4Glc)ILQ as well as the double mutant RLVTVNGKSILQ were previously reported.[21] Difficulties arose in the synthesis of the cyclic RLVTVNGKS(Me_4Glc)ILQ control, so the glycine-splitting value for the cyclic RLVTVNGKS(Ac_4Glc)ILQ control was used instead. The energy of folding for each peptide was determined from the difference in the chemical shifts of the glycine hydrogens. The side chain interaction energy was then determined using eq 3:

$$\Delta\Delta G(\text{X}-\text{Z}) = \Delta G_{XZ} - \Delta G_{XS} - \Delta G_{LZ} + \Delta G_{LS} \tag{3}$$

Thermal Denaturation. Variable-temperature NMR was used to perform the thermal denaturation experiments. A temperature range of 275−330 K was explored in 5 K increments. The temperature was calibrated using methanol and ethylene glycol standards. The change in glycine chemical-shift difference was used to determine the fraction folded at each temperature. The fraction folded was then plotted against temperature and the data fitted with eq 4:

$$\text{fraction folded} = \frac{\exp\left(\frac{x}{RT}\right)}{1 + \exp\left(\frac{x}{RT}\right)} \tag{4}$$

where

$$x = T\left[\Delta S_{298}^\circ + \Delta C_p^\circ \ln\left(\frac{T}{298\text{ K}}\right)\right] - \Delta H_{298}^\circ + \Delta C_p^\circ (T - 298\text{ K}) \tag{5}$$

Acknowledgment. This work was supported by funding from the National Institutes of Health, Institute of General Medical Sciences (GM072691 and GM071589).

Supporting Information Available: Synthetic procedures for glycosylated amino acids, NMR assignments, data from double mutant cycles, and thermodynamic data. This material is available free of charge via the Internet at http://pubs.acs.org.

JA803960X