Supporting Information for

Evaluation of a Carbohydrate-π Interaction in a Peptide Model System

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

CONTENTS

Synthetic Procedures.

Fmoc-Ser(Ac_4Glucose)-OH. Synthesis of the glycosylated amino acid was achieved *via* a modified procedure by Elofsson, et al.[i] To a dry round bottom flask purged with N_2 was added 2.0 g (5.1 mmol) peracylated glucose, 1.85 g (5.61 mmol) Fmoc-Ser-OH, 0.8 mL (6.63 mmol) $SnCl_4$, and 15 mL dry acetonitrile. The mixture was allowed to stir at room temperature and the reaction progress was monitored by TLC (visualized by UV and *p*-anisaldehyde stain). When the reaction was complete, the solution was diluted with CH_2Cl_2, washed with 10% HCl (aq) and dried over magnesium sulfate. The solvent was removed by reduced pressure leaving a white solid. The pure product was obtained in a 64% yield (2.15 g) after column chromatography utilizing 90% CH_2Cl_2/9% MeOH/1% NH_4OH as the mobile phase. The NMR spectrum of the product was identical to that previous reported.[1]

General Procedure for Peptide Synthesis and Purification. All 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 0.45M HBTU/HOBt in DMF. The following protecting groups were used: Arg(Pbf), Asn(Trt), Cys(Trt), Gln(Trt), Lys(Boc), Ser(tBu), Thr(tBu), Trp(Boc).

Deprotection of the Fmoc groups was achieved with 2% piperidine, 2% 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in DMF. All 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 was performed simultaneously by treatment with 95% trifluoroacetic acid (TFA)/2.5% triisopropylsilane (TIPS)/2.5% H_2O, for 2-3 hours under nitrogen. The TFA was removed by distillation under vacuum. The crude peptides were precipitated with cold ether, extracted into water and lyophilized.

Crude peptides were then dissolved and purified by reverse phase HPLC using a Vydac C_{18} semipreparative column. Peptides were eluted with a linear gradient of 95% H_2O /5% acetonitrile with 0.1% TFA (Solvent A) and 95% acetonitrile/5% water with 0.1% TFA (Solvent B) from 0-30% B. Peptides were detected by monitoring at 220 and 280 nm. Molecular weights were determined using MALDI mass spectrometry. Disulfide bonds were formed by air oxidation of purified peptides in methanol. Peptides were then filtered through Celite and repurified.

*Synthesis of Peptides Containing Glc orAc_4Glc (**1, 2, 4, 5, 8, and 9**).* Peptides were made via solid phase peptide synthesis on the automated synthesizer up to the addition of Fmoc-Ser(Ac_4Glc)-OH, which was added manually in order to ensure complete coupling via the Kaiser test. Deprotection of the Fmoc and addition of the following amino acid was also completed by hand, at which point the peptide was placed back on the synthesizer and the synthesis was completed as described above. Deacetylation of peptides **2**, **5**, and **9** was achieved by treatment of the lyophilized pure peptide with 6mM NaOMe for 30 minutes.[2] The desired peptide was repurified by HPLC and was obtained in greater than 90% yield.

NMR Data Acquisition.

NMR samples were made to concentrations of 3-8 mM and 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 between 8-64 scans using a 1.5 s presaturation. All 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 1st dimension and 256 in the 2nd dimension. ROESY experiments were performed with 32 scans in the 1st dimension and 256-512 in the 2nd dimension. All spectra were analyzed using standard window functions (sinebell and gaussian with shifting). Assignments were made using standard methods.[3] Thermal denaturations were performed in duplicate in 5-10 degree increments. The temperature was calibrated with methanol and ethylene glycol standards using Varian macros.

Determination of fraction folded.

The extent of folding of these β-hairpins was determined by examining the perturbation of the alpha hydrogen chemical shifts,[4,5,6] which have been reported to shift with increased folding of the peptide.[7] The alpha hydrogens in the strand typically shift downfield from random coil due to increased hydrogen bonding, while the turn residues typically shift upfield, with a shift of > 0.1 ppm considered significant. To determine the random coil chemical shifts, two 7-residue peptides were synthesized for each mutant, one containing the seven N-terminal residues and the other containing the seven C-terminal residues (Figure 1, peptides **7-10**). Shifting of alpha hydrogens was consistent with beta-hairpin structure, with the exception of Leu, which was shifted upfield due to proximity to the aromatic ring of Trp (Figure 2). Additionally, the β-hairpin structure of peptides **1** and **2** was confirmed by the observation of long-range cross strand

NOEs (Figure 3). Some cross strand NOEs were not observed due to spectral overlap (Val_{α}-Lys_{α} in **1,** Arg_{α}-Gln_{α} in **1**, Arg_{α}-Gln_{α} in **2**).

4: CRWVTVNGKS(Ac_4Glc)ILQC

5: CRWVTVNGKS(Glc)ILQC

6: CRWVTVNGKSILQC

7: RWVTVNG

8: NGKS(Ac_4Glc)ILQ

9: NGKS(Glc)ILQ

10: NGKSILQ

Figure 1. Sequences of fully folded and random coil peptides.

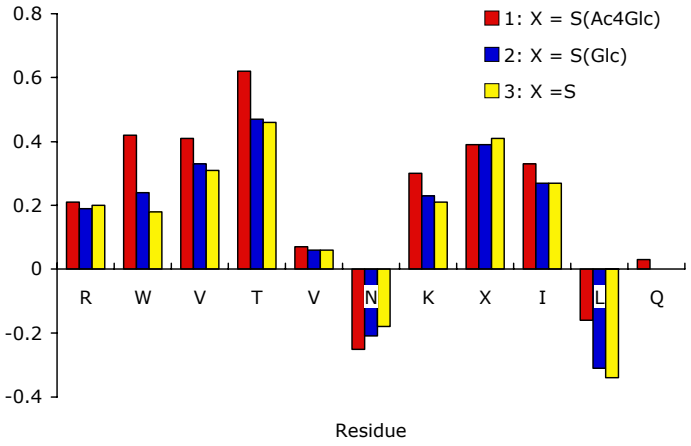

Figure 2. Downfield shifting of alpha hydrogens in peptides **1**, **2** and **3.**

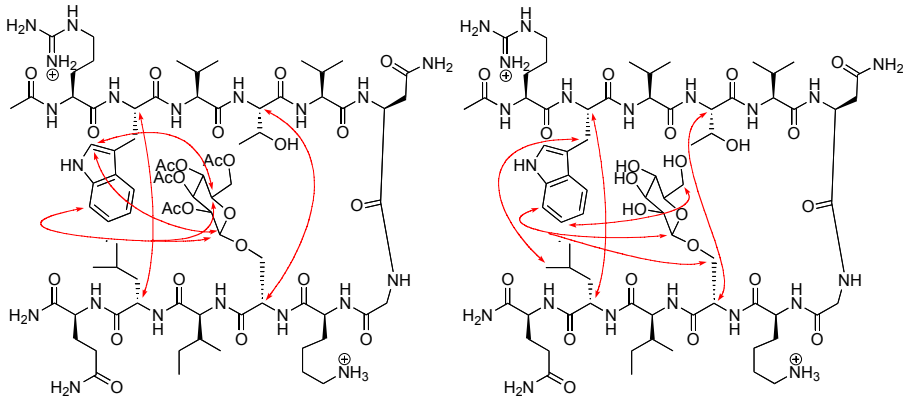

Figure 3. Cross-strand NOEs for peptides **1** and **2**.

The fraction folded was quantified using equation 1, in which the chemical shifts were compared to those of a fully folded and random coil peptide. The fully folded chemical shifts were obtained from the cyclic peptides in which Cys residues were incorporated at both the N- and C- terminal positions (Figure 1, peptides **4-6**). Cyclization was performed via disulfide formation. The β-hairpin structure of the cyclic peptides was confirmed by the presence of cross-strand NOEs down the length of the peptide (Figure 4).

$$\text{Fraction Folded} = f = [\delta_{obs} - \delta_0]/[\delta_{100} - \delta_0] \qquad \text{(eqn 1)}$$

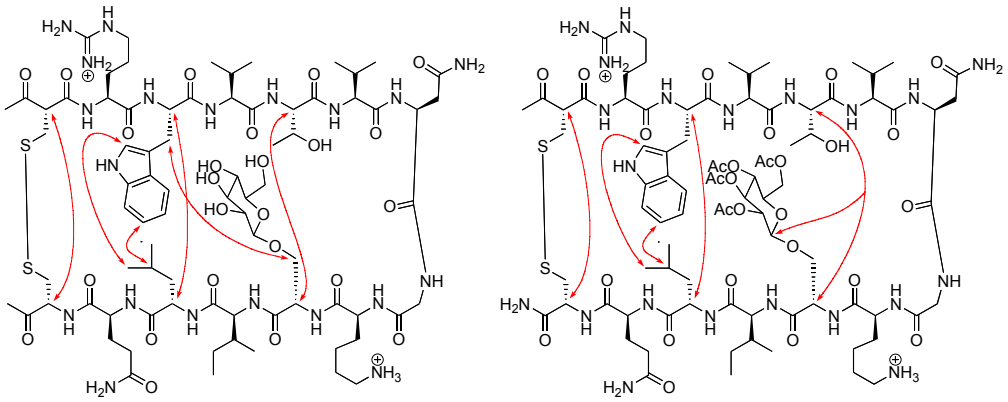

Figure 4. NOEs of fully folded cyclic peptides **4** and **5**.

The difference in chemical shift of the alpha hydrogens of Gly7 (located in the turn) is also indicative of folding.[6,7] The difference in chemical shift of the glycine hydrogens relative to the fully folded state was related to fraction folded as in equation 2.[6] The fraction folded was then related to an equilibrium constant (equation 3), which was then used to determine the change in free energy (equation 4).

$$\text{Fraction Folded} = f = [\Delta\delta_{obs}]/[\ \Delta\delta_{100}] \qquad \text{(eqn 2)}$$

$$K = f/(1-f) \qquad \text{(eqn 3)}$$

$$\Delta G = -\ RT \ln K \qquad \text{(eqn 4)}$$

Thermal Denaturation.

Variable temperature NMR was used to perform the thermal denaturation experiments. A temperature range of 275 to 330K was explored in five-degree increments. The change in glycine chemical shift difference was used to determine the fraction folded at each temperature. The fraction folded of the peptide was plotted against temperature, and the curve was fitted with equation 5. [4]

$$\text{Fraction folded} = (\exp[x / RT]) / (1 + \exp[x / RT]) \qquad \text{(eqn 5)}$$

$$x = (T[\Delta S_{298}^{o} + \Delta C_p^{o} \ln \{T/298\}] - [\Delta H_{298}^{o} + \Delta C_p^{o}\{T-298\}])$$

Double Mutant Cycle.

A double mutant cycle was examined in order to quantify the interaction between the peracetylated glucose and Trp. Single mutant peptides were synthesized in which Ser(Ac_4Glc) and Trp were replaced by Ser and Leu, respectively (Figure 5, peptides **3** and **11**). The double mutant peptide contained both substitutions (peptide **12**). The energy of folding for each peptide was determined from the difference in chemical shift of the glycine hydrogens. The side chain interaction energy was then determined to be –0.8 kcal/mol using equation 6 (Table 1).

$$\Delta G_{Ac4Glc\text{-}Trp} = \Delta G_{\mathbf{1}} - \Delta G_{\mathbf{3}} - \Delta G_{\mathbf{11}} + \Delta G_{\mathbf{12}} \qquad \text{(eqn 6)}$$

1: R**W**VTVNGK **S(Ac$_4$Glc)**ILQ
3: R**W**VTVNGK**S**ILQ
11: R**L**VTVNGK **S(Ac$_4$Glc)**ILQ
12: R**L**VTVNGK**S**ILQ

Figure 5. Sequences of peptides for the double mutant cycle.

Table 1. Glycine chemical shift data and ΔG° at 298K for double mutant cycle peptides.[a]

Peptide	$\Delta\delta^{Gly}$ (ppm)	Fraction folded	ΔG° (kcal/mol)
1	0.48	0.85	-1.03
3	0.38	0.64	-0.34
11	0.15	0.28	0.56
12	0.18	0.31	0.48

(a) Error is ± 0.05 kcal mol^{-1}

References

i) Salvador, L. A., Elofsson, M., Kihlberg, J.; *Tetrahedron* **1995**, *51*, 5643-5656.
2) Sjolin, P.; Elofsson, M.; Kihlberg, J. *J. Org. Chem.* **1996**, *61*, 560-565.
3) Wuthrich, K. NMR of Proteins and Nucleic Acids, Wiley Interscience, New York, 1986.
4) Maynard, A.J.; Sharman, G.J.; Searle, M.S. *J. Am. Chem. Soc.* **1998,** *120,* 1996-2007.
5) Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S. *J. Mol. Biol.* **1999,** *292,* 1051-1069.
6) Searle, M. S.; Griffiths-Jones, S. R.; Skinner-Smith, H. *J. Am. Chem. Soc.* **1999,** *121,* 11615-11620.
7) Wishart, D. S.; Sykes, B.D.; Richards, F.M. *Biochemistry* **1992,** *31,* 1647-1651.

Peptide 1: Ac-RWVTVNGKS(Ac_4Glc)ILQ-NH_2

Residue	α	β	γ	δ	ε		
R	4.34	1.63	1.49	3.12			
W	5.18	3.15	7.25 (2)	7.49(4)	7.03(5)	7.23(6)	7.42(7)
V	4.53	2.12	0.88				
T	4.96	3.90	1.05				
V	4.20	1.92	0.90				
N	4.44	3.06, 2.77					
G	4.08, 3.60						
K	4.63	1.82	1.43	1.72	3.02		
S(AcGlc)*	5.00	3.89	3.74				
I	4.52	1.87	1.41, 1.21, 0.88	0.88			
L	4.21	1.44	1.18	0.53,0.37			
Q	4.34	2.02, 1.89	2.26				
***AcGlc**	4.05 (1)	4.68 (2)	4.78 (3)	4.82(4)	2.74(5)	3.95(6)	3.69(6')

Peptide 2: Ac-RWVTVNGKS(Glc)ILQ-NH_2

Residue	α	β	γ	δ	ε		
R	4.32	1.67, 1.64	1.46	3.10			
W	5.00	3.09, 3.14	7.21 (2)	7.47(4)	7.02(5)	7.21(6)	7.41(7)
V	4.45	2.07	0.87				
T	4.81	3.99	1.08				
V	4.19	1.96	0.92				
N	4.48	3.01, 2.77					
G	4.06,3.67						
K	4.59	1.88, 1.79	1.43	1.70	3.01		
S(Glc)*	5.02	4.02, 3.85					
I	4.44	1.85	1.43,1.18,0.87	0.87			
L	4.06	1.35	0.80	0.55,0.35			
Q	4.30	2.05, 1.91	2.28				
***Glc**	4.11 (1)	3.10 (2)	3.28 (3)	3.15 (4)	2.98(5)	3.43(6)	3.42(6')

Peptide 3: Ac-RWVTVNGKSILQ-NH_2

Residue	α	β	γ	δ	ε		
R	4.33	1.62	1.48	3.10			
W	4.94	3.20, 3.09	7.21 (2)	7.46(4)	7.05(5)	7.21(6)	7.45(7)
V	4.43	2.04	0.88				
T	4.80	4.02	1.10				
V	4.19	1.99	0.91				
N	4.51	3.01, 2.78					
G	4.06,3.68						
K	4.57	1.84	1.44	1.73	3.00		
S	4.88	3.78					
I	4.46	1.81	1.42,1.17,0.88	0.88			
L	4.01	1.29, 1.10	0.77	0.51, 0.35			
Q	4.29	2.05, 1.93	2.28				

Peptide 4: Disulfide Ac-CRWVTVNGKS(Ac$_4$Glc)ILQC-NH$_2$

Residue	α	β	γ	δ	ε		
C	5.25	3.11					
R	4.61	1.81	1.51	1.73	3.17		
W	5.24	3.12, 3.00	7.31 (2)	7.23 (4)	6.94(5)	7.22(6)	7.32 (7)
V	4.61	2.18	0.88				
T	5.07	3.86	1.03				
V	4.22	2.02	0.88				
N	4.37	3.11, 2.77					
G	4.10, 3.54						
K	4.69	1.86	1.45	1.89	3.03		
S(AcGlc)	5.16	3.93, 3.78					
I	4.58	1.88	1.70,1.54,0.88	0.88			
L	4.09	1.40	0.45	0.41, -0.28			
Q	4.60	2.15, 1.85	2.28				
C	5.06	2.96					

Peptide 5: Disulfide Ac-CRWVTVNGKS(Glc)ILQC-NH$_2$

Residue	α	β	γ	δ	ε		
C	5.24	3.01					
R	4.61	1.84	1.49	1.71	3.18		
W	5.10	3.06	7.25 (2)	7.47 (4)	6.89(5)	7.20 (6)	7.25 (7)
V	4.66	2.14	0.91				
T	5.05	3.91	1.06				
V	4.24	1.93	0.95				
N	4.38	3.12, 2.79					
G	4.12, 3.54						
K	4.73	1.90	1.47	1.79	3.06		
S(Glc)	5.31	4.03, 3.89					
I	4.61	1.83	1.50,1.21,0.88	0.88			
L	3.87	1.32	0.05	0.39, -0.32			
Q	4.56	2.11, 1.85	2.23				
C	5.04	2.94					

Peptide 6: Disulfide Ac-CRWVTNGKSILQC-NH$_2$

Residue	α	β	γ	δ	ε		
C	5.22	3.01, 2.40					
R	4.62	1.80, 1.69	1.50	3.19			
W	5.04	3.08, 3.00	7.21 (2)	7.45(4)	6.95(5)	7.19(6)	7.31(7)
V	4.62	2.07	0.89				
T	5.09	3.95	1.04				
V	4.23	1.94	0.90				
N	4.38	3.10, 2.78					
G	4.13, 3.54						
K	4.69	1.85	1.46	1.73	3.06		
S	5.17	3.84, 3.72					
I	4.67	1.85	1.46,1.18,0.90	0.90			
L	3.81	1.25, 0.73	0.38	-0.09, -0.30			
Q	4.55	2.11,1.83	2.25				
C	5.04	3.08, 3.00					

Peptide 7: Ac-RWVTVNG-NH_2

Residue	α	β	γ	δ	ε		
R	4.13	1.55	1.37	3.05			
W	4.76	3.34, 3.22	7.21 (2)	7.62 (4)	7.08 (5)	7.22 (5)	7.44 (7)
V	4.12	2.00	0.87				
T	4.34	4.16	1.19				
V	4.13	2.10	0.93				
N	4.69	2.86, 2.75					
G	3.88						

Peptide 8: Ac-NGKS(AcGlc)ILQ-NH_2

Residue	α	β	γ	δ	ε		
N	4.69	2.83, 2.77					
G	3.95						
K	4.33	1.78	1.39	1.67	2.99		
S(AcGlc)*	4.61	4.02					
I	4.19	1.87	1.46,1.17,0.90	0.90			
L	4.37	1.65	1.65	0.90			
Q	4.31	2.15, 2.00	2.39				
***AcGlc**	4.91 (1)	4.94 (2)	5.35 (3)	5.10 (4)	4.08 (5)	4.43 (6)	4.23 (6')

Peptide 9: Ac-NGKS(Glc)ILQ-NH_2

Residue	α	β	γ	δ	ε		
N	4.68	4.87, 2.74					
G	3.95						
K	4.36	1.83	1.48	1.66	3.00		
S(Glc)*	4.63	4.14, 3.93					
I	4.17	1.87	1.48,1.20,0.91	0.91			
L	4.37	1.66	1.66	0.90			
Q	4.30	2.10, 1.99	2.37				
***Glc**	4.48 (1)	3.28 (2)	3.50 (3)	3.36 (4)	3.46(5)	3.93(6)	3.17(6')

Peptide 10: Ac-NGKSILQ-NH_2

Residue	α	β	γ	δ	ε
N	4.68	2.81			
G	3.95				
K	4.36	1.85	1.43	1.66	3.00
S	4.47	3.87			
I	4.19	1.90	1.47, 1.21, 0.92	0.92	
L	4.35	1.62	1.62	0.91	
Q	4.29	2.12, 1.98	2.37		

Peptide 11: Ac-RLVTVNGKS(Ac$_4$Glc)ILQ-NH$_2$

Residue	α	β	γ	δ	ε
R	4.29	1.77	1.64	3.20	
L	4.53	1.60	1.60	0.89	
V	4.31	2.04	0.93		
T	4.59	4.12	1.15		
V	4.17	2.04	0.94		
N	4.65	2.78			
G	4.00, 3.85				
K	4.45	1.82	1.42	1.64	3.00
S(Ac4Glc)	4.76	4.04, 4.00			
I	4.26	1.84	1.43,1.16	0.89	
L	4.43	1.60	0.87		
Q	4.33	2.10, 2.00	2.37		

Peptide 12: Ac-RLVTVNGKSILQ-NH$_2$

Residue	α	β	γ	δ	ε
R	4.29	1.77, 1.70	1.61	3.20	
L	4.52	1.59	1.59	0.88	
V	4.28	2.04	0.94		
T	4.58	4.09	1.13		
V	4.16	2.02	0.94		
N	4.62	2.79			
G	4.00, 3.82				
K	4.45	1.80	1.43	1.68	2.99
S	4.62	3.81			
I	4.26	1.88	1.43,1.18,0.89	0.89	
L	4.42	1.61	0.90		
Q	4.32	2.11, 1.99	2.36		

Peptide 13: Ac-CRLVTVNGKSILQC-NH$_2$

Residue	α	β	γ	δ	ε
C	5.26	3.16			
R	4.62	1.82	1.53	3.19	
L	4.89	1.49	1.49	0.77	
V	4.48	2.02	0.86		
T	5.03	3.91	0.98		
V	4.21	1.92	0.92		
N	4.38	2.75			
G	4.12, 3.54				
K	4.67	1.87	1.43	1.69	3.01
S	5.03	3.70			
I	4.46	1.82	1.41,1.17,0.87	0.87	
L	4.66	1.43	1.43	0.79	
Q	4.66	2.15, 1.90	2.28		
C	5.11	3.02			

Peptide 14: Ac-CRLVTVNGKS(Ac$_4$Glc)ILQC-NH$_2$

Residue	α	β	γ	δ	ε	others
C	5.12	3.11/2.60				
R	4.59	1.79	1.59	3.17		
L	4.92	1.59	0.90			
V	4.20	2.08	0.99			
T	5.04	3.89	1.02			
V	4.08	1.98	0.87			
N	4.38	3.15/2.74				
G	4.09, 3.56					
K	4.64	2.24	1.42	1.82	3.02	
S(AcGlc)	5.22	4.01/3.815				
I	4.44	1.78	1.445	0.82		
L	4.66	1.55	0.83			
Q	4.58	2.12	1.89/1.55			
C	5.27	3.08/3.00				
AcGlc	4.42 (H1)	4.85 (H2)	5.38 (H3)	5.12 (H4)	3.15 (5)	4.01/4.02 (H6, H6')

Peptide 15: Ac-RLVTVNG-NH$_2$

Residue	α	β	γ	δ	ε
R	4.26	1.80, 1.74	1.62	3.17	
L	4.42	1.61	1.61	0.92, 0.87	
V	4.14	2.08	0.93		
T	4.39	4.16	1.17		
V	4.17	2.05	0.92		
N	4.69	2.86, 2.77			
G	3.90				