Sensitivity of Glycopeptide Conformation to Carbohydrate Chain Length

Rui Liang, Amy Hamilton Andreotti, and Daniel Kahne*

Department of Chemistry, Princeton University
Princeton, New Jersey 08544

Received February 7, 1995

Compelling evidence that glycosylation influences protein structure has been obtained from studies comparing the global conformations of fully glycosylated proteins with their deglycosylated counterparts.[1] A desire to understand in more detail how sugars exert their effects has prompted studies on glycopeptide model systems.[2,3] One conclusion that has been drawn from investigations of O-glycosylated peptides is that the effects of glycosylation on peptide conformation are due to the sugar closest to the peptide backbone.[1a,e,3a,4] We felt that the previous studies were equivocal because the methods used for characterizing the structures did not provide information about the local peptide backbone conformation.[5] We have therefore compared the backbone conformations of two different peptides containing either a mono- or a disaccharide at an internal threonine. Our studies show that in both peptides the backbone conformation is radically different depending on whether a mono- or a disaccharide is attached. This finding could have implications for how glycosylation mediates biological activity in glycoproteins. It certainly has consequences for the design and study of model glycopeptides.

Small peptides are flexible molecules that can adopt a large number of conformations. The flexibility makes structural studies difficult. Nevertheless, we have shown that it is possible to use NMR to evaluate *differences* in backbone conformation between peptides in different glycosylation states.[6] NOEs between sequential amide protons provide a particularly useful indicator of the ensemble average backbone conformation.[7] To evaluate the effects of mono- and disaccharides on peptide backbone conformation, we synthesized glycohexapeptides **1** and **2** (Figure 1) on Rapp TentaGel resin[8] and purified them by RP-HPLC.[9,10] Proton resonances in DMSO were assigned using DQF-COSY and ROESY experiments.[11] Sequential dNN NOE connectivities were then analyzed to characterize the structures.

As shown in Figure 1, the sequential amide−amide ROESY cross peak intensities are very different for glycopeptides **1** and **2**. In peptide **1**, the strongest amide−amide ROESY cross peak is between the amide resonances of Lys and Thr. A weaker NOE cross peak is observed between Phe2 and D-Trp, and there is no NOE cross peak between D-Trp and Lys. In contrast, the NOE cross peak between D-Trp and Lys in peptide **2** is very strong. The Lys−Thr cross peak, which is the strongest one in peptide **1**, is absent in peptide **2**. Thus, the hexapeptide with the disaccharide attached has a profoundly different average backbone conformation from the hexapeptide with the monosaccharide attached.

Our finding that the second sugar has a significant effect on the backbone conformation is not limited to the specific peptide sequence Ac-Phe-Phe-D-Trp-Lys-Thr-Phe-NH_2 or to the DMSO solvent. We also synthesized the Ac-Val-Thr-His-Pro-Gly-Tyr-NH_2 sequence derived from oncofetal fibronectin with both GalNAcα (**3**) and a Galβ(1−3)GalNAcα (**4**) attached to the threonine.[12] Contour plots of the amide region of the ROESY spectra for these peptides in H_2O are shown in Figure 2. Since the proline does not have an amide proton,[13] the possible sequential amide−amide cross peaks are Val-Thr, Thr-His, and Gly-Tyr.[14] In peptide **3**, the Gly-Tyr amide−amide NOE is very strong while the Val-Thr NOE cross peak is absent. In peptide **4** the Val-Thr NOE cross peak is very strong and the Gly-Tyr NOE cross peak is absent. It is also worth noting that the amide−amide NOE between threonine and the sugar, which is strong in peptide **3**, is absent in peptide **4**.

The NMR data clearly show that the monosaccharyl peptides have very different ensemble average conformations from the disaccharyl peptides. Although none of the peptides has a single discrete structure, an analysis of the NMR data for peptides **1** and **2** suggests an interesting possibility. Peptides **1** and **2** are glycosylated linear analogues of a cyclic peptide designed to contain a type II′ β turn with D-Trp and Lys in the $i + 1$ and $i + 2$ positions, respectively.[6,15] The NMR data for **1**, showing a large NOE between Lys and Thr and no NOE between D-Trp and Lys, is consistent with this kind of turn (Figure 3). Peptide **2**, in contrast, has a large NOE between D-Trp and Lys and no NOE between Lys and Thr. These NOEs are inconsistent with a type II′ turn containing D-Trp and Lys in the $i + 1$ and $i + 2$

(1) (a) Rose, M. C.; Voter, W. A.; Sage, H.; Brown, C. F.; Kaufman, B. *J. Biol. Chem.* **1984**, *259*, 3167−3172. (b) Gerken, T. A.; Dearborn, D. G. *Biochemistry* **1984**, *23*, 1485−1497. (c) Gerken, T. A.; Jentoft, N. *Biochemistry* **1987**, *26*, 4689−4699. (d) Shogren, R. L.; Jentoft, N.; Gerken, T. A.; Jamieson, A. M.; Blackwell, J. *Carbohydr. Res.* **1987**, *160*, 317−327. (e) Shogren, R.; Gerken, T. A.; Jentoft, N. *Biochemistry* **1989**, *28*, 5525−5536. (f) Gerken, T. A.; Butenhof, K. J.; Shogren, R. *Biochemistry* **1989**, *28*, 5536−5543. (g) Butenhof, K. J.; Gerken, T. A. *Biochemistry* **1993**, *32*, 2650−2663.

(2) (a) Dill, K.; Hardy, R. E.; Daman, M. E.; Lacombe, J. M.; Pavia, A. A. *Carbohydr. Res.* **1982**, *108*, 31−40. (b) Ishii, H.; Inoue, Y.; Chujo, R. *Int. J. Pept. Protein Res.* **1984**, *24*, 421−429. (c) Bush, C. A.; Feeney, R. E. *Int. J. Pept. Protein Res.* **1986**, *28*, 386−397. (d) Maeji, N. J.; Inoue, Y.; Chujo, R. *Int. J. Pept. Protein Res.* **1987**, *29*, 699−707. (e) Maeji, N. J.; Inoue, Y.; Chujo, R. *Biopolymers* **1987**, *26*, 1753−1767. (f) Mimura, Y.; Inoue, Y.; Maeji, N. J.; Chujo, R. *Int. J. Pept. Protein Res.* **1989**, *34*, 363−368. (g) Hollosi, M.; Perczel, A.; Fasman, G. D. *Biopolymers* **1990**, *29*, 1549−1564.

(3) (a) Otvos, L., Jr.; Thurin, J.; Kollat, E.; Urge, L.; Mantsch, H. H.; Hollosi, M. *Int. J. Pept. Protein Res.* **1991**, *38*, 476−482. (b) Urge, L.; Gorbics, L.; Otvos, L., Jr. *Biochem. Biophys. Res. Commun.* **1992**, *184*, 1125−1132. (c) Jansson, A. M.; Meldal, M.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1992**, 1699−1707. (d) Kessler, H.; Matter, H.; Gemmecker, G.; Kottenhahn, M.; Bats, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 4805−4818. (e) Davis, J. T.; Hirani, S.; Bartlett, C.; Reid, B. R. *J. Biol. Chem.* **1994**, *269*, 3331−3338. (f) Mouritsen, S.; Meldal, M.; Christiansen-Brams, I.; Elsner, H.; Werdelin, O. *Eur. J. Immunol.* **1994**, *24*, 1066−1072.

(4) Jentoft, N. *Trends Biochem. Sci.* **1990**, *15*, 291−294.

(5) For example, in an elegant series of experiments, Otvos and co-workers showed using CD that glycosylation of a helical peptide with a monosaccharide disrupted the helix.[3a] Glycosylation with a disaccharide disrupted the helix to a similar extent. This study demonstrated the profound effect glycosylation can have on peptide structure, but it did not provide structural information at a high enough resolution to establish whether there were differences in local conformation between the mono- and disaccharyl peptides.

(6) Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 3352−3353.

(7) (a) Wagner, G.; Neuhaus, D.; Worgotter, E.; Vasak, M.; Kagi, J. H. R.; Wuethrich, K. *J. Mol. Biol.* **1986**, *187*, 131−135. (b) Wuethrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons, Inc.: New York, 1986; Chapter 9. (c) Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1988**, *201*, 161−200.

(8) Manuscript in preparation.

(9) All new compounds were characterized by NMR and FAB MS. The conditions for synthesis, RP-HPLC purification, and spectral characterization for compounds **1**−**4** are provided in supporting information.

(10) The monoglycosylated peptide has limited solubility in water.

(11) (a) Billeter, M.; Braun, W.; Wuthrich, K. *J. Mol. Biol.* **1982**, *155*, 321−346. (b) Rance, M.; Sorensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479−485. (c) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−813.

(12) (a) Matsuura, H.; Hakomori, S. I. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 6517−6521. (b) Matsuura, H.; Takio, K.; Titani, K.; Greene, T.; Levery, S. B.; Salyan, M. E. K.; Hakomori, S. *J. Biol. Chem.* **1988**, *263*, 3314−3322. (c) Matsuura, H.; Greece, T.; Hakomori, S. *J. Biol. Chem.* **1989**, *264*, 10472−10476. (d) Lüning, B.; Norberg, T.; Tejbrant, J. *J. Chem. Soc., Chem. Commun.* **1989**, 1267. (e) Gobbo, M.; Biondi, L.; Filira, F.; Rocchi, R. *Int. J. Pept. Protein Res.* **1991**, *38*, 417.

(13) Both glycopeptides **3** and **4** have the trans configuration around the His-Pro amide bond. See supporting information for details.

(14) Unfortunately, the Thr-His NOEs cannot be assessed due to overlap in both peptides **3** and **4**.

(15) Veber, D. F.; Holly, F. W.; Paleveda, W. J.; Nutt, R. F.; Bergstrand, S. J.; Torchiana, M.; Glitzer, M. S.; Saperstein, R.; Hirschmann, R. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2636.

0002-7863/95/1517-10395$09.00/0

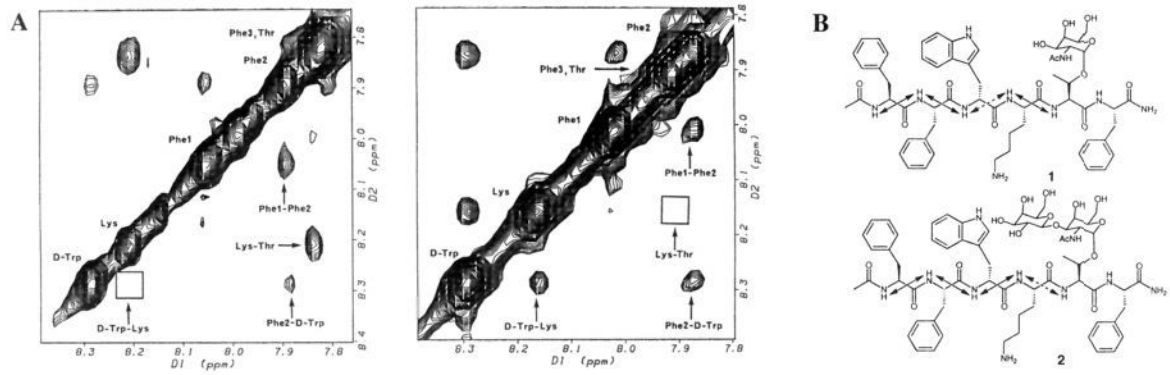

Figure 1. (A) Amide regions of the 500 MHz ROESY spectra of **1** (left) and **2** (right) in DMSO-d_6. Conditions: 30 °C, 3.7 mM, 175 ms mixing time. Empty boxes indicate absent amide−amide ROESY cross peaks. (B) Sequential amide−amide contacts in **1** and **2** are shown by the solid arrows. The broken arrows correspond to the empty boxes in the spectra.

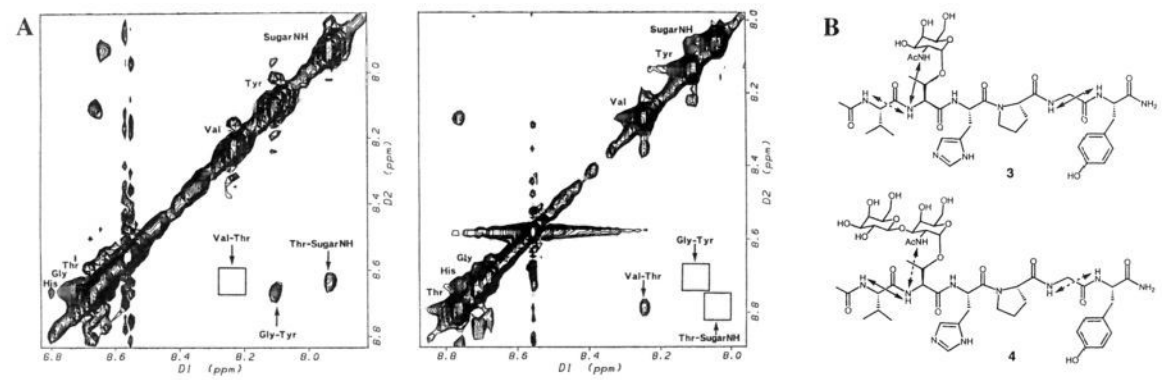

Figure 2. (A) Amide regions of the 500 MHz ROESY spectra of **3** (left) and **4** (right) in 90:10 H_2O:D_2O. Conditions: 15 °C, 6 mM (**3**), 7 mM (**4**), pH 4.1, 175 ms mixing time. Empty boxes indicate absent amide−amide ROESY cross peaks. (B) Sequential amide−amide contacts in **3** and **4** are shown by the solid arrows. The broken arrows correspond to the empty boxes in the spectra.

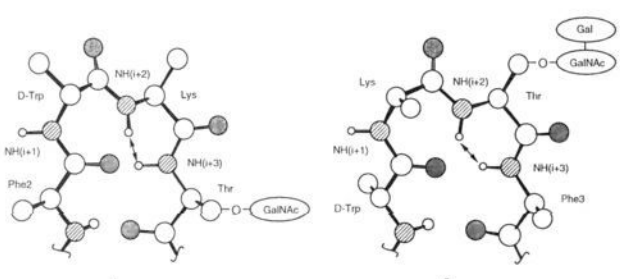

Figure 3. Computer-generated models of the preferred turns for peptides **1** and **2** based on the dNN NOEs. The oxygen atoms are represented by shaded circles, the nitrogen atoms by striped circles, and the carbons and hydrogens by large and small open circles, respectively. In solution the peptides actually adopt an ensemble of conformations of which these turns may comprise only a small fraction.

positions but are compatible with a type II′ turn with Lys in the $i + 1$ and Thr in the $i + 2$ position (Figure 3).[16] It is noteworthy that the second sugar overwhelms the effect of D-Trp in disfavoring a turn with the D-amino acid in the $i + 1$ position.[17]

Recent experimental evidence has suggested that glycosylation causes flexible peptides to bend, *i.e.*, to turn.[6,18] Our results provide the first evidence that the type of bend depends on the structure of the attached sugar. It has been proposed that Nature uses post-translational glycosylation, like phosphorylation, as a general way to modulate protein structure and thus activity.[19] Our findings provide a glimpse into how glycosylation with different sugars could be used as a biological switch since they show that different sugars can interact with the same peptide backbone to stabilize profoundly different conformations.

Acknowledgment. This work was supported by the National *Science Foundation (PYI program).*

Supporting Information Available: Procedures for the synthesis and purification of compounds **1**−**4**, 1-D and 2-D NMR spectra for **1**−**4**, and FAB MS for **1**−**4** (22 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA950403C

(16) The mono- and disaccharyl fibronectin peptides **3** and **4** are also compatible with two different frame-shifted turn conformations. However, the presence of proline reduces the information available from NMR and makes modeling more difficult.

(17) Chandrasekaran, R.; Lakshminarayanan, A. V.; Pandya, U. V.; Ramachandran, G. N. *Biochim. Biophys. Acta* **1973**, *303*, 14−27.

(18) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 97−101.

(19) (a) Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993** *218*, 1−27. (b) Varki, A. *Glycobiology* **1993**, *3*, 97−130. (c) 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.