

Glycosylation of Threonine of the Repeating Unit of RNA Polymerase II with β -Linked *N*-Acetylglucosamine Leads to a Turnlike Structure

Eric E. Simanek,[†] Dee-Hua Huang,[†] Laura Pasternack,[†] Timothy D. Machajewski,[†] Oliver Seitz,[†] David S. Millar,[‡] H. Jane Dyson,[‡] and Chi-Huey Wong^{*,†}

Contribution from the Departments of Chemistry and Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received July 2, 1998

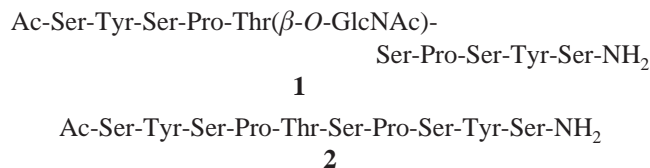
Abstract: Two models of the repeating C-terminal domain of RNA polymerase II (Ac-SYSPTSPSYS-NH₂; Ac-SYSPT(β -O-GlcNAc)SPSYS-NH₂) were prepared and their conformations in water studied using 1-D and 2-D ¹H NMR spectroscopies, CD spectrophotometry, fluorescence anisotropy, and molecular mechanics and dynamics calculations. The data suggest that glycosylation of the native, randomly coiled peptide with a single, biologically relevant sugar leads to the formation of a turn. This report represents the first structural study of a new class of glycoproteins monoglycosylated with *N*-acetylglucosamine on threonine.

Introduction

The polymeric C-terminal domain (CTD) of the largest subunit of the holoenzyme responsible for the synthesis of mRNA, RNA polymerase II, consists of a highly conserved repeating sequence, (YSPTSPS)_{*n*}.^{1,2} The length of the sequence varies with the evolutionary complexity of the organism; mammals have 52 repeats,^{3,4} *Drosophila* have 45 repeats,^{4,5} yeasts have 27 repeats,^{6,7} and plasmodia have 18 repeats.⁸ While it is generally agreed that hyperphosphorylation of the CTD converts the polymerase from an inactive, promoter-bound state to an actively transcribing assembly,⁹ the pattern of phosphorylation remains uncertain.^{10,11} Recently, Hart and colleagues recognized that the CTD is also glycosylated with a single, β -linked *N*-acetylglucosamine (GlcNAc) on threonine.¹² While the role that glycosylation plays is unclear, the modification of the CTD through glycosylation or phosphorylation is mutually exclusive.

The search for general principles that describe the role of glycosylation in peptides^{13–15} and proteins¹⁶ is being pursued

by many groups: examples have been reported where glycosylation leads to differences in structure between native and glycosylated sequences.^{13–15} Herein we report that glycosylation with a single β -linked GlcNAc on the threonine residue of a model of the polymeric domain of RNA polymerase II, **1**,



leads to increased structure over the corresponding sugar-free peptide. We find evidence for a nonrandom, turnlike structure about the site of glycosylation for **1** and random coil structure for **2**. The evidence for the turnlike structure comes primarily from ¹H NMR: nuclear Overhauser effect (NOE) connectivities, variable temperature coefficients of the amide protons, chemical shift differences of amide and H _{α} protons of **1** and **2**, and the rates of ¹H/²H exchange are all consistent with a turn. CD spectrophotometry and fluorescence anisotropy measurements corroborate this conclusion. Computation using NMR-derived constraints produces a model with turnlike structure about the site of glycosylation.

Model System. Molecules **1** and **2** contain the heptad repeat unit and additional amino acids appended to the N- and C-termini. Our choice to include these residues is based on previous studies of models of the native peptide (similar to **2**) by Suzuki,¹⁷ Harding,¹⁸ and Corden.¹⁹

On the basis of data from sedimentation and fluorescence quenching studies, Suzuki concluded that the uncapped sequence, YSPTSPSY, bound to DNA. Noting the similarity of the sequence to the known bisintercalator, Triostin A, Suzuki

[†] Department of Chemistry.

[‡] Department of Molecular Biology.

(1) Corden, J. L. *TIBS* **1990**, *15*, 383–387.
 (2) Woychik, N. A.; Young, R. A. *TIBS* **1990**, 347–351.
 (3) Corden, J. L.; Cadena, D. L.; Ahearn, J. M., Jr.; Dahmus, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 7934–7938.
 (4) Allison, L. A.; Wong, J. K.-C.; Fitzpatrick, V. D.; Moyle, M.; Ingles, C. J. *Mol. Cell. Biol.* **1988**, *8*, 321–329.
 (5) Zehring, W. A.; Lee, J. M.; Weeks, J. R.; Jokerst, R. S.; Greenleaf, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3698–3702.
 (6) Allison, L. A.; Moyle, M.; Shales, M.; Ingles, C. J. *Cell* **1985**, *42*, 599–610.
 (7) Nonet, M.; Sweetser, D.; Young, R. A. *Cell* **1987**, *50*, 909–915.
 (8) Li, W. B.; Bzik, D. J.; Gu, H.; Tanaka, M.; Fox, B. A.; Inselburg, J. *Nucleic Acids Res.* **1989**, *17*, 9621–9636.
 (9) Dahmus, M. E. *Biochim. Biophys. Acta* **1995**, *1261*, 171–182.
 (10) Baskaran, R.; Dahmus, M. E.; Wang, J. Y. *J. Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11167–11171.
 (11) Zhang, J.; Corden, J. L. *J. Biol. Chem.* **1991**, *266*, 2290–2296.
 (12) Kelly, W. G.; Dahmus, M. E.; Hart, G. W. *J. Biol. Chem.* **1993**, *268*, 10416–10424.
 (13) Blanco, F. J.; Jimenez, M. A.; Herranz, J.; Rico, M.; Santoro, J.; Nieto, J. L. *J. Am. Chem. Soc.* **1993**, *115*, 5887.
 (14) O'Connor, S. E.; Imperiali, B. *J. Am. Chem. Soc.* **1997**, *119*, 2295.

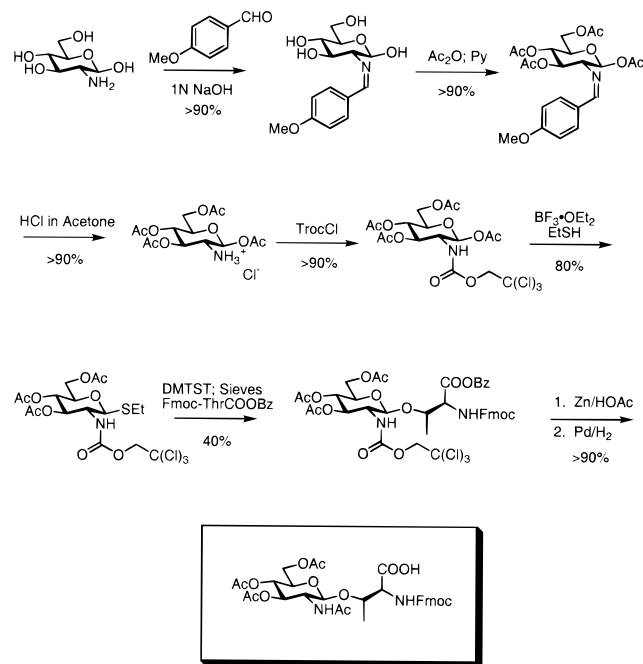
(15) Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 3352.

(16) Varki, A. *Glycobiology* **1993**, *3*, 97.

(17) Suzuki, M. *Nature* **1990**, *344*, 562–565.

(18) Harding, M. M. *J. Med. Chem.* **1992**, *35*, 4658–4664.

(19) Cagas, P. M.; Corden, J. L. *Proteins* **1995**, 149.

Scheme 1. Synthesis of the Glycosylated Amino Acid

proposed that the tyrosine residues intercalated into the helix. Harding's examination of the conformation of the uncapped sequence YSPTSPSY in water by ^1H NMR found that a small population exists in the conformation proposed by Suzuki. Two β -turns were identified as important for placing the aromatic rings of the tyrosine residues in the correct geometry. Contrary to Harding's results, Corden found that $(\text{YSPTSPS})_8$ exists as a random-coil structure in water but adopted a β -turn structure in trifluoroethanol.

To reconcile these results, we chose to add serine residues to both the N- and C-termini of the Suzuki/Harding model to increase the steric constraints. We also capped the N- and C-termini with acetyl and carboxamide groups, respectively, to reduce ion-pairing interactions.

Results

Synthesis of 1 and 2. Peptides **1** and **2** were synthesized using Fmoc chemistry²⁰ in a peptide synthesis vessel using Rink Amide AM resin with a loading density of 0.65 mmol/g as the solid support. The amino acids were purchased (Fmoc-Ser(O-*t*-Bu)COOH, Fmoc-Tyr(O-*t*-Bu), Fmoc-Pro; Novabiochem) or prepared (Fmoc-Thr(O- β -GlcNAc(OAc)₃)COOH) by a known protocol (Scheme 1).²¹ Details of the iterative synthesis (Scheme 2) are provided in the Experimental Section.

Peptide Purification. Purification of **1** and **2** was effected in three phases. The product oil obtained from the cleavage cocktail was purified by size exclusion on an LH-20 column. Fractions containing peptide were combined, evaporated, and subjected to silica gel chromatography. The fractions containing purified peptide were combined and submitted to HPLC purification using preparative reverse-phase chromatography. These materials gave single ions ($M + \text{Na}^+$) by MALDI-MS. For **1**, Zemplen deprotection afforded a single product by HPLC which gave satisfactory MALDI-MS ($M + \text{Na}^+$) and ^1H NMR. Full details are available in the Experimental Section.

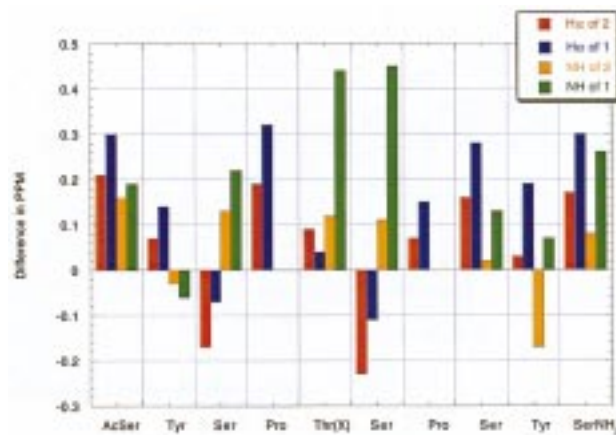


Figure 1. Chemical shift differences from values expected for random coils²³ for the NH and H α of **1** and **2**.

NMR Spectroscopy. One-dimensional ^1H NMR spectroscopy afforded data on coupling constants, chemical shift, VT coefficients, and amide NH exchange rates. Two-dimensional ROESY and NOESY experiments provided a consistent set of NOEs. These data are discussed individually in the following paragraphs.

(a) Coupling Constants. Table 1 summarizes relevant coupling constants and chemical shift information collected for **1** and **2**. For **1**, we observed a $^3J_{\text{HN,H}\alpha}(\text{Thr}^5) = 8.6$ Hz, implying a transoid relationship between these protons: we adopted a dihedral constraint of $-160^\circ < \phi < -80^\circ$ in our computational models.²² The $^3J_{\text{HN,H}\alpha}$ of the remaining amides of **1**, and for those of **2** all fall within a range (6.0–8.0 Hz) where structural predictions are difficult. The very small coupling constant observed between Thr⁵H α and Thr⁵H β ($J = 1.8$ Hz) suggests that there is little free rotation about the C α –C β bond. We adopted dihedral constraints for (H α –C α –C β –H β) in which $55^\circ < \phi < 60^\circ$ and $110^\circ < \phi < 120^\circ$.

(b) Chemical Shift Differences. Local structure can be inferred by comparing the chemical shifts of **1** and **2** to values tabulated for random coil sequences (Figure 1).²³ While there are not substantial differences in the shifts of H α for **1** or **2**, the amide protons of Thr⁵ and Ser⁶ of **1** are markedly different than those of **2** and the values recorded for peptides with random structure.

(c) NOES. Two-dimensional NOESY and ROESY spectroscopy (Figures 2–4) yielded a series of cross-peaks that are represented with double headed arrows in Scheme 3. The existence of proline residues led us to probe for β -turn structures.²⁴ The NOEs that are diagnostic for turns (where $i + 1 = \text{Pro}$) include $d_{\text{NN}}(i+2, i+3)$, $d_{\alpha\text{N}}(i, i+2)$, and either $d_{\alpha\text{N}}(i+1, i+2)$ for type II or $d_{\alpha\text{N}}(i+1, i+2)$ for type I turns. While we see cross-peaks at $d_{\text{NN}}(i+2, i+3)$ and $d_{\alpha\text{N}}(i+1, i+2)$ for both potential turn regions ($i = \text{Ser}^3$ and Ser⁶) in **1**, we see *no* cross-turn NOEs at $d_{\alpha\text{N}}(i, i+2)$ for either region (shown as empty boxes in Figure 5). The absence of these critical NOEs in water prevents us from assigning type II β -turn structure to these regions. We propose, however, that a turnlike structure exists (*vide infra*). Tables 2 and 3 give NOE and dihedral constraints utilized for computational modeling.

(d) Variable Temperature Coefficients. The extent to which an amide proton is exposed to solvent is reflected in the dependence of its chemical shift on temperature. Exposed

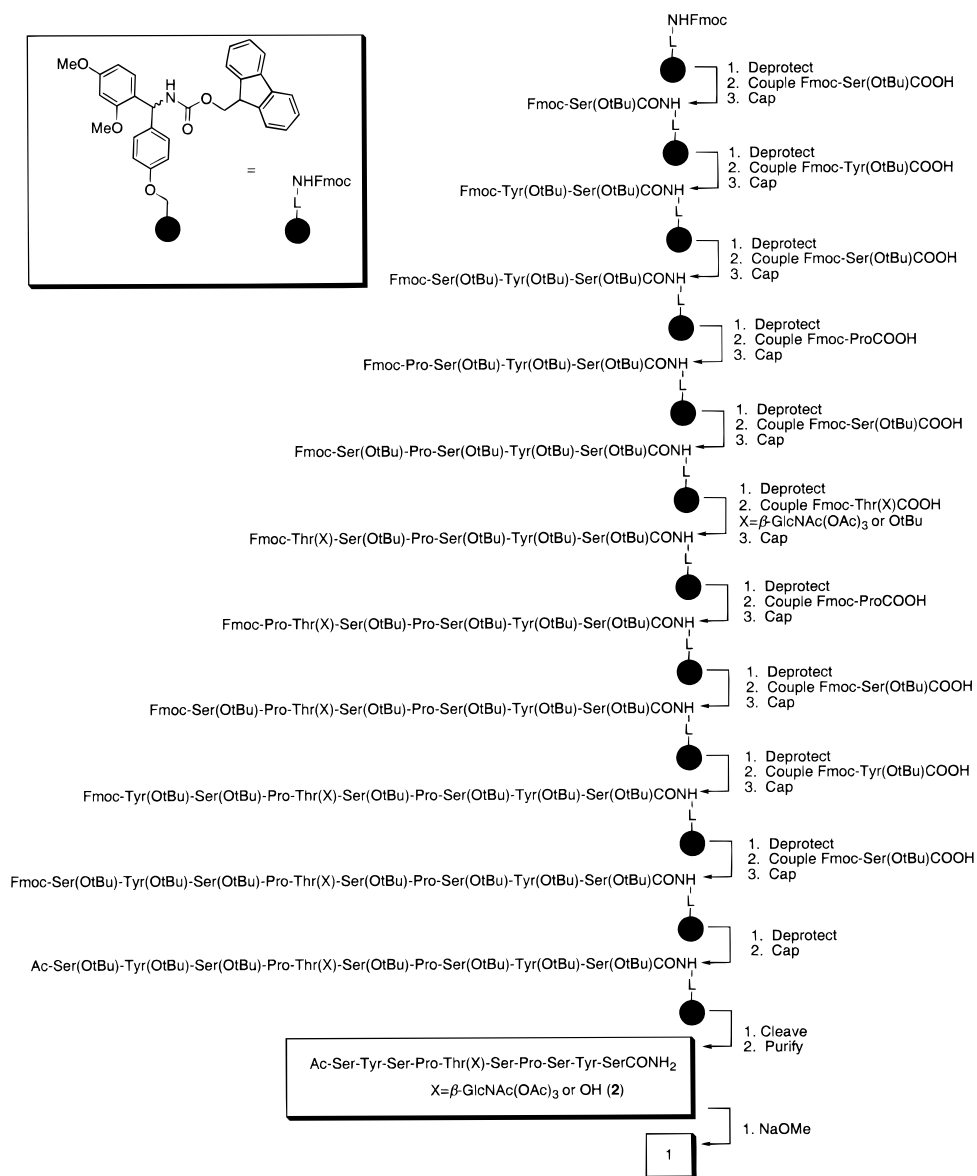
(20) Sheppard, R. C.; Williams, B. J. *J. Chem. Soc., Chem. Commun.* **1982**, 587.

(21) Meinjohanns, E.; Meldal, M.; Bock, K. *Tetrahedron Lett.* **1995**, 36, 9205.

(22) Botuyan, M. V.; Toy-Palmer, A.; Chung, J.; Blake, R. C.; Beroza, P.; Case, D. A.; Dyson, H. J. *J. Mol. Biol.* **1996**, 263, 752.

(23) Wishart, D. S.; Sykes, B. D. *Methods Enzymol.* **1994**, 239, 363.

(24) Nelsoney, C. L.; Kelly, J. W. *Bioorg. Med. Chem.* **1996**, 4, 739.

Scheme 2. Iterative Synthesis of **1** and **2**

amides such as those in a randomly coiled peptide shift more than those in an ordered structure.²⁵ The variable temperature spectra for **1** (Figure 6) reveal that the VT coefficients of Thr⁵ and Ser⁶ of **1** are significantly lower than those recorded for the other amide protons (Table 1), suggesting that these protons are involved in structure either directly through hydrogen bonding or indirectly by being buried within the protein. The values for the remaining amide protons of **1**, as well as those recorded for **2**, reveal that these protons are not effectively protected from solvent.

(e) Rates of ¹H/²H Exchange. Amide NH involved in structure or engaged in hydrogen bonding show different rates of exchange than those amides of a randomly coiled peptide.²⁵ Adding 25% (v/v) D₂O to a solution of **1** in H₂O at 278 K reveals that the all amide protons except those of Thr⁵ and Ser⁶ exchange rapidly. Exchange for these two residues is >20× slower at 278 K, thus offering further evidence for their involvement in a structured domain resulting from glycosylation. All amide protons of **2** exchange at similar rates.

CD and Fluorescence Measurements. CD spectrophotometry (Figure 7) and fluorescence anisotropy measurements corroborate a structural change upon glycosylation. A minimum around 198 nm in the CD spectra is consistent with a peptide having a random-coil structure. While this feature is the only feature in the spectrum of **2**, the traces of **1** show local maxima at 190 nm—a feature consistent with a turnlike structure.²⁶ Fluorescence anisotropy measurements for **1** and **2** suggest that tyrosines of **1** are held more rigidly in place than those of **2** (0.0065 vs 0.0045). While these values are significant, their magnitude suggests that the ends of both peptides remain highly flexible.

Computation. Independent computational analyses using Monte Carlo conformation searches and simulated annealing calculations reveal a common low-energy structure (Figure 8).

(a) Monte Carlo Calculations using Macromodel. Monte Carlo conformational searches were performed on an SGI Octane workstation using Macromodel v5.5. Our initial modeling was carried out on the complete model of **1**. The NOE distance and dihedral constraints (Tables 2 and 3) were applied

(25) Wutrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.

(26) Woody, R. W. *The Peptides*; Academic Press: New York, 1985; Vol. 7.

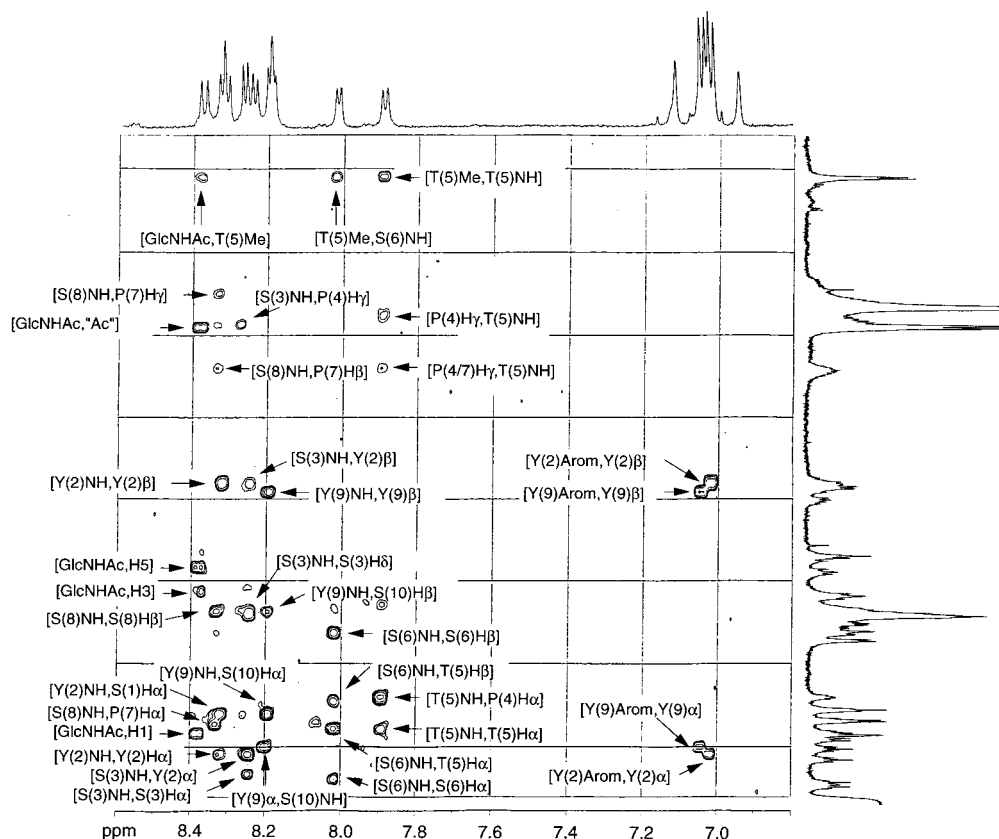
Table 1. Assignment of Glycopeptide, VT Coefficients^a

residue of 1	NH	³ J _{αN}	VT ^a	VT ^b	H _α	H _{β1,2}	Hγ _{1,2}	other
AcSer ¹	8.19	6.9	3.9	4.7	4.20	3.59		Me, 1.74
Tyr ²	8.24	7.5	4.8	5.5	4.46	3.60, 2.81		<i>o</i> , 6.92; <i>m</i> , 6.62
Ser ³	8.16	7.4	3.5	4.4	4.57	3.60		
Pro ⁴					4.12	2.10	1.80, 1.77	δ _{1,2} , 3.55, 3.45
Thr ⁵	7.80	8.5	1.0	1.9	4.31	4.15		Me, 0.91
Ser ⁶	7.93	6.4	1.6	2.6	4.61	3.73		
Pro ⁷					4.29	2.10	1.84, 1.65	δ _{1,2} , 3.68, 3.58
Ser ⁸	8.25	7.3	3.2	3.3	4.22	3.60		
Tyr ⁹	8.11	6.5	3.3	4.4	4.41	3.63, 2.87		<i>o</i> , 6.95; <i>m</i> , 6.64
SerNH ₂ 10	8.12	7.0	3.2	4.0	4.20	3.62		NH _{a,b} , 7.10, 6.85

GlcNAc: NH, 8.29; H1, 4.41; H2, 3.56; H3, 3.40; H4, 3.31; H5, 3.31; H6a, 3.65; H6b, 3.78

residue of 2	NH	³ J _{αN}	VT ^b	H _α	H _{β1,2}	Hγ _{1,2}	other
AcSer ¹	8.22	8.1	3.5	4.29	3.76, 3.73		Me, 1.84
Tyr ²	8.21	7.8	4.2	4.53	2.99		<i>o</i> , 6.90; <i>m</i> , 6.60
Ser ³	8.25	7.7	3.9	4.67	3.71		
Pro ⁴				4.25	2.23	1.90–1.93	δ _{1,2} , 3.55, 3.64_
Thr ⁵	8.12	7.0	3.0	4.26	4.18		Me, 1.12
Ser ⁶	8.27	7.0	3.9	4.73	3.83		
Pro ⁷				4.39	2.21	1.96, 1.77	δ _{1,2} , 3.79, 3.68
Ser ⁸	8.36	7.0	4.4	4.34	3.71		
Tyr ⁹	8.35	6.7	5.0	4.57	2.93		<i>o</i> , 7.02; <i>m</i> , 6.70
SerNH ₂ 10	8.30	7.1	4.4	4.33	3.71		NH _{a,b}

^a The assignment of the ¹H NMR spectra of **1** and **2** acquired in 90%:10% H₂O:D₂O at pH 3.25. Chemical shifts are reported in parts per million. ³J_{αN} are reported in hertz. VT coefficients measured (a) between 278 and 298 K and (b) between 278 and 308 K are included (−Δppb/ΔK).

**Figure 2.** NH–H_α region of the 2-D ¹H NMR spectrum of **1** recorded in 90%:10% H₂O:D₂O at pH 3.5.

with the default force constant of 100 kJ/(mol·Å²). In those cases where the distance constraint could be applied to more than one hydrogen (methylene or methyl hydrogens), the constraint was set to the corresponding carbon with an appropriate increase in the acceptable deviation from the mean distance. The Monte Carlo simulation was set to vary all dihedral angles except the amide bonds, which were constrained in the trans orientation. A total of 6200 Monte Carlo steps using

AMBER^{*27}/PRCG minimization with the GB/SA solvent (H₂O) treatment²⁸ to a gradient convergence (<0.05 kJ/mol) were conducted. A low-energy structure was found which contained a turn at the site of glycosylation.

(27) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078.

(28) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.

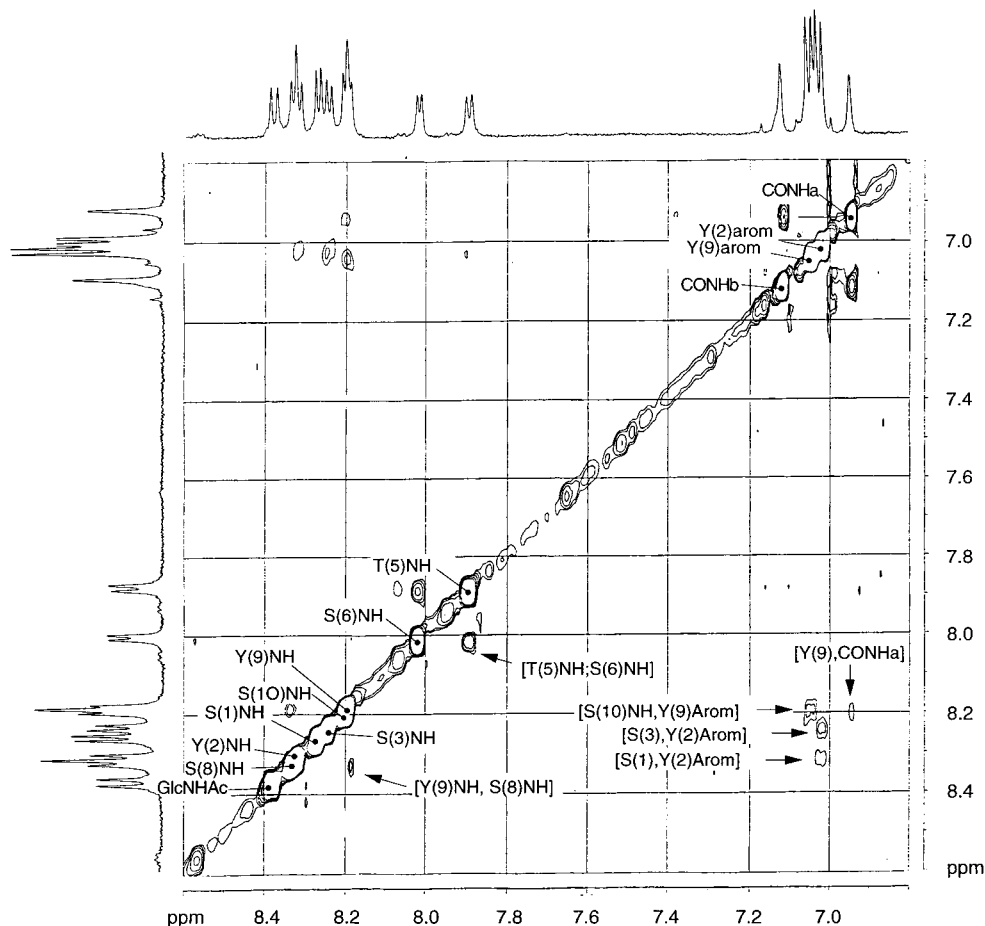
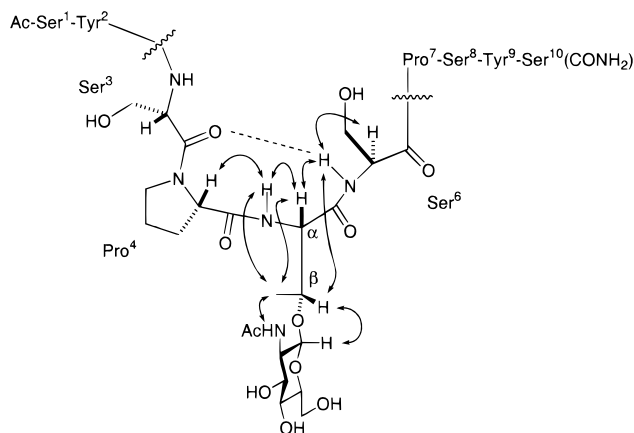


Figure 3. NH–NH region of the 2-D ^1H NMR spectrum of **1** recorded in 90%:10% $\text{H}_2\text{O}:\text{D}_2\text{O}$ at pH 3.5.

Scheme 3. Structural Model for **1**^a



^a NOEs are shown as double headed arrows. The dashed line indicates a *potential* hydrogen-bonding interaction. The cross-turn NOE at $d_{\alpha\text{N}}(\text{Ser}^3, \text{Thr}^5)$ is not observed. The actual position of the carbohydrate is not shown: it lies above the plane of the turn. See Figure 8 for more details.

To reduce the computational time, we examined the conformational search of a truncated model of **1**, Ac-Ser-Pro-Thr(β -O-GlcNAc)-Ser-Pro-NH₂. The nOe distance and dihedral constraints were applied as before. A total of 15 000 Monte Carlo steps using AMBER*/PRCG minimization *without* solvent treatment were conducted. Of the 645 unique structures obtained, the lowest energy conformer was found 18 times: an additional 42 related conformers were found to be within 2.0 kcal/mol. All of these structures exhibited a turn at the site of glycosylation. These conformations superimpose with an aver-

Table 2. NOEs Recorded for **1** and the Assigned Distances

atom 1	atom 2	mean distance (Å)	deviation (\pm Å)
GlcNAc-H1	Thr5- β Methyl	2.0	1.1
GlcNAc-H1	Thr5-H β	2.0	1.0
GlcNAc-H1	GlcNAc-H3	2.0	1.0
GlcNAc-NH	GlcNAc-H5	3.0	1.0
GlcNAc-NH	Thr5- β Methyl	4.0	1.1
GlcNAc-NH	GlcNAc-H1	3.0	1.0
GlcNAc-NH	GlcNAc-NAc	2.0	1.1
GlcNAc-NH	GlcNAc-H3	3.0	1.0
Ser3-H α	Ser3-C β	3.0	1.5
Ser3-H α	Pro4-C δ	3.0	1.5
Thr5-NH	Pro4-C β	4.0	1.5
Thr5-NH	Pro4-H α	2.0	1.0
Thr5-NH	Thr5-H α	3.0	1.0
Thr5-NH	Thr5- β Methyl	4.0	1.1
Thr5-NH	Thr5-H β	2.0	1.0
Thr5-H β	Thr5- β Methyl	3.0	1.1
Ser6-NH	Thr5-H α	3.0	1.0
Ser6-NH	Thr5- β Methyl	4.0	1.1
Ser6-NH	Thr5-NH	2.0	1.0
Ser6-NH	Thr5-H β	3.0	1.0
Ser6-NH	Ser6-H α	3.0	1.0
Ser6-H α	Ser6-C β	2.0	1.5
Ser6-NH	Ser6-C β	2.0	1.5
Ser6-H α	Pro7-C δ	2.0	1.5
Ser3-NH	Ser3-C β	3.0	1.5
Ser3-NH	Ser3-H α	3.0	1.0

age backbone root mean square deviation of 0.560 Å from the lowest energy conformer. The lowest 13 structures are shown in Figure 9.

As a test of the validity of the computation on the truncated model, the lowest energy conformations obtained from both

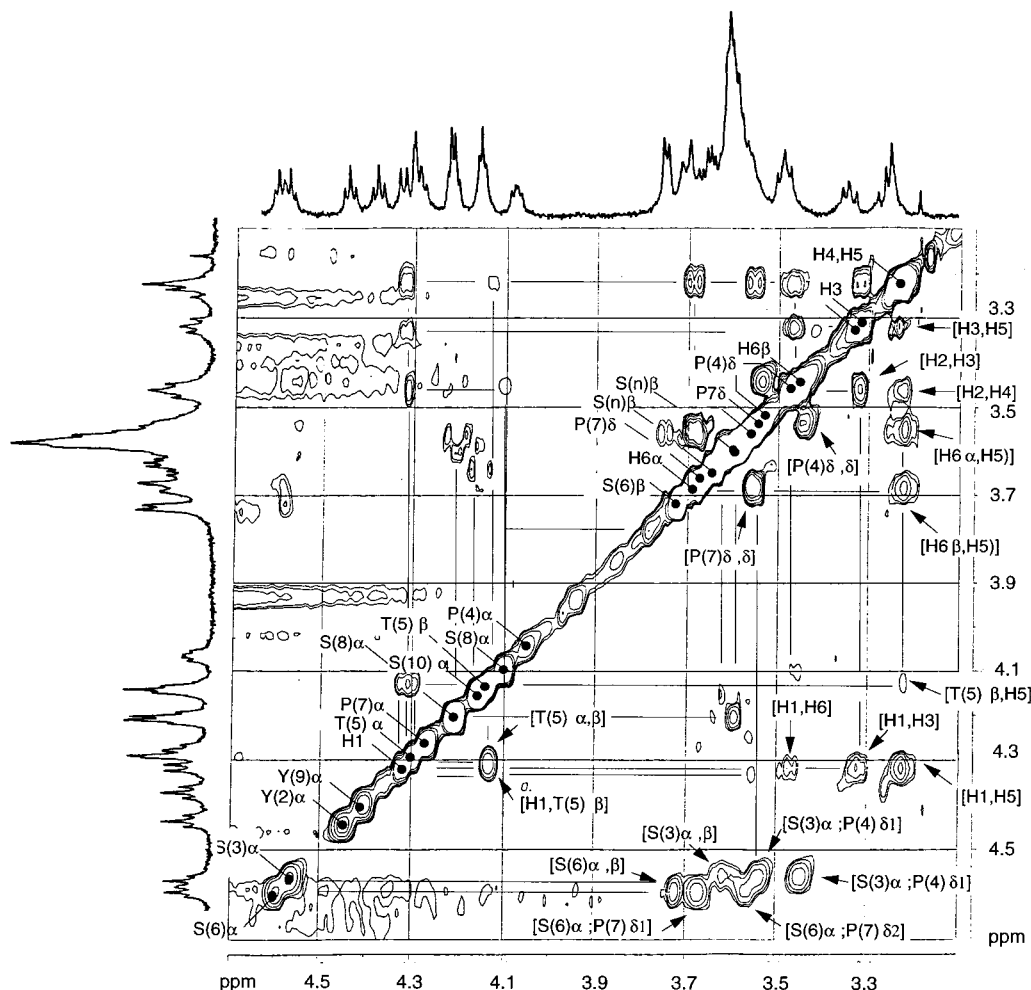


Figure 4. Detail of the aliphatic region of the 2-D ^1H NMR spectrum of **1** recorded in 90%:10% $\text{H}_2\text{O}:\text{D}_2\text{O}$ at pH 3.5.

Table 3. Coupling Constant Dihedral Angle Constraints for **1**

atom 1	atom 2	atom 3	atom 4	mean angle (deg)	deviation (\pm deg)
Thr5-NH	Thr5-N	Thr5-C α	Thr5-H α	-120.0	40
Ser6-NH	Ser6-N	Ser6-C α	Ser6-H α	-120.0	40

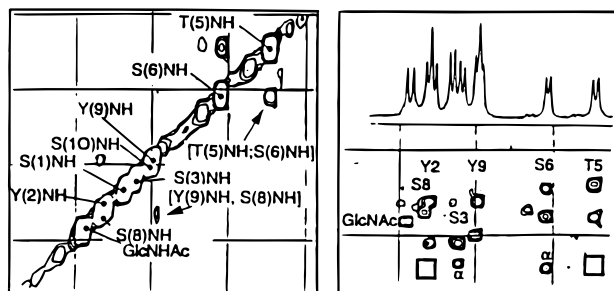


Figure 5. Portions of the 2-D NOESY spectra of **1** recorded at 5 $^\circ\text{C}$ in 90%:10% $\text{H}_2\text{O}:\text{D}_2\text{O}$ at pH 3.5: (a) NH-NH region; (b) NH-H α region. The boxes indicate the positions where $d_{\text{aN}}(i,i+2)$ cross-peaks might have confirmed the existence of type II β -turns.

searches were superimposed (Figure 10). The turn is clearly defined in both structures with only minor deviations in the peripheral residues.

(b) Molecular Dynamics. Molecular dynamics calculations utilizing the Amber force field with Homans' carbohydrate potentials²⁹ were performed on an SGI Indigo 2 workstation

(29) Homans, S. W. *Biochemistry* **1990**, *29*, 9110.

using the Discover module of InsightII. Models for **1** and **2** were constructed using the biopolymer builder module. Initial molecular dynamics (MD) calculations of the entire glycopeptide indicated high variability of the end residues. Thus subsequent calculations used the Ac-Ser-Pro-Thr-Ser-Pro-NH₂ section and carbohydrate.

Dynamics runs included minimization, randomization of coordinates, equilibration at 1000 K, ramping of force constants at 1000 K, cooling, and finally minimization using conjugate gradients (to a maximum derivative of less than 0.01 kcal/Å). A total of 26 experimentally determined constraints were applied during the MD runs. Calculations were carried out on structures in vacuo without the explicit inclusion of water molecules. A distance-dependent dielectric constant of $4r$ was used. A total of 60 final structures were obtained and categorized by visual inspection (Figure 11). Of the 60 models 25 fell into a single group whose relative energy was 2.6 kcal lower than any of the other groups. This group was chosen for further analysis. Superimposition of all 25 models gave an all atom root mean square deviation of 1.18 Å with a maximum violation of NMR constraints of 0.510 Å. A second structure of higher energy was also found in which the carbohydrate was placed in the plane of the turn (identified as the "open" structure of Figure 11).

Discussion

What emerges from these studies is a model of the glycosylated CTD of RNA polymerase II. Using computation guided by ^1H NMR-derived distance constraints, we find that glyco-

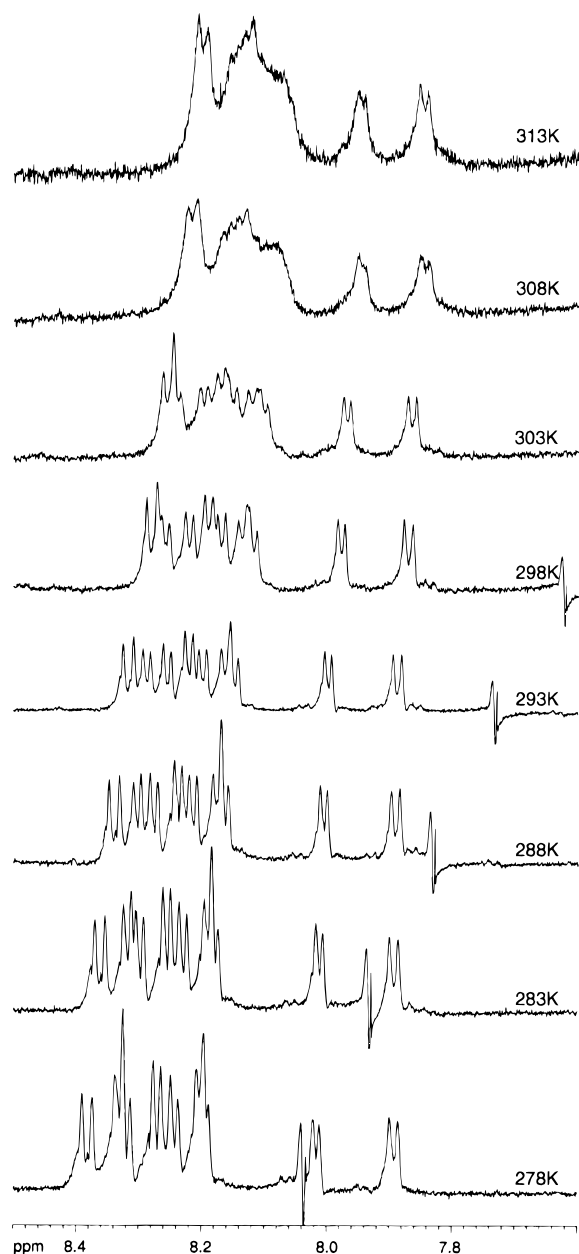


Figure 6. Variable temperature spectra for **1**.

sylation leads to the formation of a turn. Both the VT coefficients for the amides of Thr⁵ and Ser⁶ and their lower rates of ¹H/²H exchange in comparison to the other amides are consistent with the structural model.²⁰

While the biological relevance is unclear, we draw attention to three aspects of the model that result from the close proximity of the side chains of Ser³ and Ser⁶. In the lowest energy structure they lie on the same face of the turn beneath the carbohydrate group. First, in the lowest energy conformer these two residues appear to make hydrogen bonding contacts with hydroxyl groups and the glycosidic oxygen. The downfield shift of the β -protons of Ser⁶ is consistent with such interactions. Second, these two serine residues have been identified as sites for phosphorylation.¹¹ The existence of negative charges on these two groups would presumably lead to severe ion-ion repulsion, resulting in a dramatic conformational change. We are currently exploring this hypothesis. Third, the model also explains one aspect of the exclusivity of glycosylation and phosphorylation: the carbohydrate shields the hydroxyl groups of Ser³ and Ser⁶ from solvent and soluble factors (i.e. enzymes).

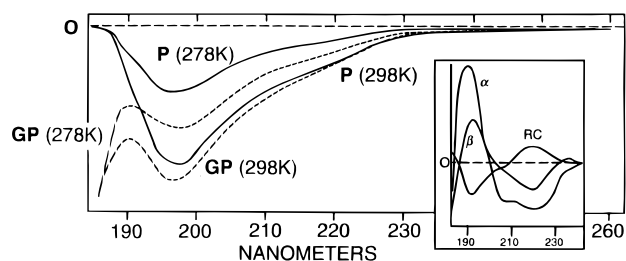


Figure 7. CD spectrogram of **1** (bold lines) and **2** (dashed lines) at 278 and 298 K. The inset shows the traces expected for α -helix (α), β -sheets (β), and random coils (RC).

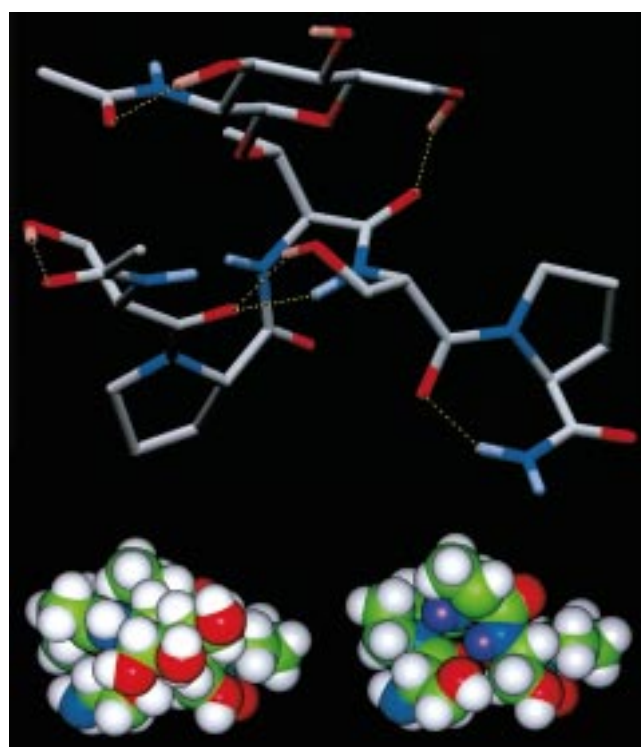


Figure 8. (Top) Computational model of **1**. Analysis of the computation-derived structures from Ac-Ser-Pro-Thr(GlcNAc)-Ser-Pro-NH₂ revealed a stable conformation for the carbohydrate residue. The model obtained from computation is consistent with experimental data. Interestingly, the sites identified for phosphorylation, Ser³ and Ser⁶, are close enough that a large conformational change might be expected upon phosphorylation.¹¹ (Bottom) CPK model of **1**. The structure on the left shows the carbohydrate over the analyzed domain, Ser-Pro-Thr(GlcNAc)-Ser. The carbohydrate has been removed in the structure on the right in order to show the buried amides of Thr⁵ and Ser⁶ (purple).

Of greater interest perhaps is the comparison of two independent calculations. Both Monte Carlo and molecular dynamics calculations provide the same low-energy structure in which the carbohydrate lies over the plane of the turn.

In summary, we have investigated, for the first time, the effect of glycosylation on the local structure of an important class of glycoproteins that are glycosylated with a single GlcNAc residue. We have chosen water as a solvent, as it is likely to best mimic the environment which is biologically significant. Capping both N- and C-termini prevents ion-pairing interactions. While a turn structure is revealed, the spectrum remains too complex to interpret additional structural details: the region around δ 3.6 ppm which contains signals for the β -protons of five serine residues, some carbohydrate protons, and the δ -protons of the prolines is too crowded for detailed analysis. Additional model compounds are warranted. Work is in progress to investigate further details of this monoglycosylation-

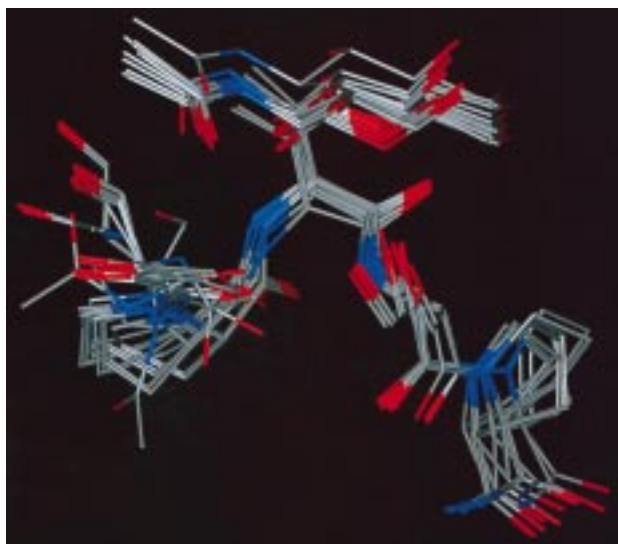


Figure 9. Superimposition of the 13 lowest energy conformations of the truncated model of **1** as determined from Monte Carlo simulations.

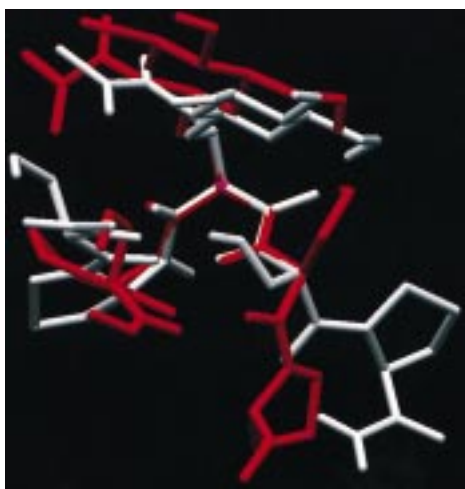


Figure 10. Superimposition of the lowest energy conformations from computation using the entire model of **1** (red; for clarity only Ser-Pro-Thr(GlcNAc)-Ser-Pro is shown) with the truncated model of **1** (white).






	"closed"		"open"		
					
number	25	9	4	15	7
average relative energy	3.11	5.60	6.67	7.21	11.24

Figure 11. Visual inspection of the 60 lowest energy structures of **1** by molecular dynamics reveals two classes of structures: the low-energy "closed" structure and a higher energy "open" structure. The low-energy structure obtained is very similar to that obtained from Monte Carlo simulations.

induced conformational change, and how glycosylation affects protein function.

Experimental Section

Fmoc-protected amino acids, hydroxybenztriazole (HOBT), *o*-benzotriazolyltetramethyluronium hexafluorophosphate (HBTU), and

Rink Amide Resin were purchased from Novabiochem. Dimethylformamide (DMF) was obtained from Fluka. Other reagents for synthesis were purchased from Aldrich Chemical Co. The peptide synthesis vessel was obtained from Chem Glass.

Peptide Synthesis. (a) **Deprotection.** To 250 mg of resin, 10 mL of a solution of 1:1 DMF:morpholine was added. The reaction vessel was placed on an Innova 2000 platform shaker (160 rpm) for 50 min. After removal of solvent by vacuum filtration, and repetitive rinsing with DMF, the coupling sequence was followed.

(b) **Coupling.** First, protected amino acid (4 equiv, based on resin loading), HOBT (6 equiv), and *N*-methylmorpholine (4 equiv) were added to 10 mL of DMF. Upon dissolution, this solution was transferred to a vial containing HBTU (4 equiv). Upon dissolution, this solution was transferred to the reaction vessel. The vessel was placed on the platform shaker for 4 h (160 rpm). After removal of solvent by vacuum filtration, and repetitive rinsing with DMF, the capping sequence was followed.

(c) **Capping.** Unreacted amine groups were capped by adding a 1:1 mixture of Ac₂O and pyridine to the reaction vessel. The vessel was placed on the platform shaker for 10 min. After removal of solvent by vacuum filtration, and repetitive rinsing with DMF, the deprotection sequence was followed.

(e) **Cleavage.** Prior to cleavage, the N-terminus was capped following the protocols established for deprotection and capping. To cleave the peptide from the solid support, the resin was first washed with copious amounts of CH₂Cl₂ and dried thoroughly under vacuum. Next, a cleavage cocktail (10 mL) was prepared containing 90% (v:v) trifluoroacetic acid (TFA), 5% H₂O, 2.5% ethanedithiol, and 2.5% anisole. The solution was added to the resin and the vessel placed on the platform shaker for 3 h. The resulting solution was collected using vacuum filtration, and the resin was washed with TFA.

Peptide Purification. A product oil was obtained upon removal of cleavage cocktail with a rotary evaporator. This oil was loaded onto a LH-20 column 50% methanol in chloroform. Fractions were screened by thin layer chromatography (TLC, using 10% MeOH in CHCl₃ as an eluent) using ceric ammonium molybdate as a stain. Those fractions containing peptide were combined, and the solvent was removed. Silica gel chromatography using the same conditions provide a white solid upon evaporation of the eluent. The purity of this material was assessed using analytical HPLC.

Using a Hitachi HPLC system (L-6200A Intelligent Pump; L-4000 UV Detector; D-2500 Chromato-Integrator), the peptides were separated using an acidic gradient of MeCN (98% MeCN; 2% H₂O; 0.1% TFA) and water (98% H₂O; 2% MeCN; 0.1% TFA) on a protein and peptide C-18 column from VYDAC. For analytical separation the gradient started with 0% MeCN and grew to 10% MeCN over 30 min under a constant 1 mL/min flow rate. Both **1** and **2** were judged by integration to be >90% pure. Preparative scale HPLC was carried out using the same acidic eluents at flow rates of 3 mL/min and a gradient of 0–20% MeCN over 60 min using a VYDAC C-18 prep column. By collecting only the "top" of the peaks, material that appeared >95% pure was obtained. No visible impurities were observed in the analytical trace or in the ¹H NMR spectrum. These materials gave single ions (M + Na⁺) by MALDI-MS.

The peracetylated glycopeptide was subjected to Zemplen deprotection. To a solution of dry methanol, Na⁺ was added until the pH was 8.5. This solution was then added to the peracetylated glycopeptide, and the reaction was monitored by analytical HPLC. In 3 h, a distinct, single line was observed in the HPLC trace. Acidification of the reaction mixture with AcOH and removal of solvent provided **1**, which gave satisfactory MALDI-MS (M + Na⁺) and ¹H NMR.

¹H NMR Spectroscopy. NMR experiments were carried out on 3–5 mM solutions of **1**, using a Bruker AMX600. Data were recorded in 100% D₂O and 90%:10% H₂O:D₂O at pH = 3.5 and pH = 7.4. All spectra were recorded at 5 °C. Adjusting the pH did not affect C–H lines or NOE connectivities. No differences were observed when the 2-D spectra were recorded in the presence of 20 mM phosphate. Spectra were recorded over spectral widths of 7200 Hz with quadrature detection employed throughout. Two-dimensional spectra were acquired in the phase-sensitive mode using time-proportional phase

incrementation. Two-dimensional data sets typically containing 512 increments of t_1 were acquired and zero filled to 1024 points of a 2048 point experiment. The number of scans varied from 32 to 64 for 2-D experiments with a recycle delay of N ms. TOCSY spectra of **1** and **2** in 90%:10% H₂O:D₂O were acquired using the WATERGATE pulse sequence for water suppression and a 65 ms mixing time. NOESY and ROESY spectra of **1** and **2** in 90%:10% H₂O:D₂O were acquired using the WATERGATE pulse sequence for water suppression and three different mixing times (200, 400, and 600 ms). Data were subjected to sine-bell weighting functions in f_1 and f_2 , and a base-line correction was applied.

Fluorescence Anisotropy. Values for fluorescence anisotropy were collected on a SimAmicon fluorimeter using an excitation wavelength of 275 nm and monitoring at an emission wavelength of 310 nm. Data were collected at pH = 7 in a buffered phosphate solution (100 mM).

CD Spectrophotometry. CD traces were recorded on an Aviv 60I spectrophotometer fitted with a Peltier block. Sample concentration

was 2.5 mM. Traces recorded at pH = 7 and pH = 3.5 were indistinguishable.

Acknowledgment. The authors wish to thank Dr. Linda Tenant who assisted in the collection of CD spectra and Dr. Jeff Kelly for providing a thoughtful reading of the manuscript. This work was supported by NSF Grants CHE-9700391 (C.-H.W.) and CA 27489 (H.J.D.).

Supporting Information Available: Text and figures describing peptide synthesis and purification, computation, Monte Carlo conformational search, and ¹H NMR of **1** (17 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA982312W