Glycoconjugate Journal 17, 835–848, 2000

Glycosylations versus conformational preferences of cancer associated mucin core

Jason Schuman[1], Dongxu Qiu[2], R. Rao Koganty[2], B. Michael Longenecker[2], A. Patricia Campbell[1]*

[1]*Department of Medicinal Chemistry, School of Pharmacy, University of Washington, Seattle, WA 98195*, [2]*Biomira Inc., 2011-94 Street, Edmonton, Alberta, Canada, T6N 1H1*

Synthetic oligosaccharide vaccines based on core STn (sialyl α2-6 GalNAc) carbohydrate epitopes are being evaluated by a number of biopharmaceutical firms as potential immunotherapeutics in the treatment of mucin-expressing adenocarcinomas. The STn carbohydrate epitopes exist as discontinuous clusters, O-linked to proximal serine and threonine residues within the mucin sequence. In an effort to probe the structure and dynamics of STn carbohydrate clusters as they may exist on the cancer-associated mucin, we have used NMR spectroscopy and MD simulations to study the effect of O-glycosylation of adjacent serine residues in a repeating (Ser)n sequence. Three model peptides/glycopeptides were studied: a serine trimer containing no carbohydrate groups ($(Ser)_3$ trimer); a serine trimer containing three Tn (GalNAc) carbohydrates α-linked to the hydroxyls of adjacent serine sidechains ($(Ser.Tn)_3$ trimer); and a serine trimer containing three STn carbohydrates α-linked to the hydroxyls of adjacent serine sidechains ($(Ser.STn)_3$ trimer). Our results demonstrate that clustering of carbohydrates shifts the conformational equilibrium of the underlying peptide backbone into a more extended and rigid state, an arrangement that could function to optimally present the clustered carbohydrate antigen to the immune system. Steric effects appear to drive these changes since an increase in the size of the attached carbohydrate (STn versus Tn) is accompanied by a stronger shift in the equilibrium toward the extended state. In addition, NMR evidence points to the formation of hydrogen bonds between the peptide backbone NH protons and the proximal GalNAc groups in the $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers. The putative peptide-sugar hydrogen bonds may also play a role in influencing the conformation of the underlying peptide backbone, as well as the orientation of the O-linked carbohydrate. The significance of these results will be discussed within the framework of developing clustered STn-based vaccines, capable of targeting the clustered STn epitopes on the cancer-associated mucin.

Keywords: **cancer vaccine, glycopeptide, MUC-1, immunotherapy, NMR**

Abbreviations: **CVFF, consistent valence forcefield; DQFCOSY, double quantum filtered coherence spectroscopy; DSS, 2,2-dimethyl-2-sila-5-pentanesulfonate; KLH, keyhole limpet hemocyanin; MD, molecular dynamics; NMR, nuclear magnetic resonance spectroscopy; ROESY, rotating frame nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy**

Introduction

Mucins are attracting real interest as potential targets for immunotherapy in the development of vaccines for adenocarcinomas [1–4]. Mucin molecules are large transmembrane glycoproteins, in which the extracellular domain is densely covered with highly branched complex carbohydrate structures attached to proximal serine and threonine residues within the peptide sequence [5–8]. In the cancer-associated state, mucins become autoantigens as a result of incomplete glycosylation and sparse distribution of these carbohydrate structures [9]. This abnormal glycosylation results both in the exposure of the core peptide [10] as well as in the exposure of the normally cryptic core Tn, sialyl-Tn (STn) and Thomsen-Friedenriech (TF) carbohydrates [11], which are distributed discontinuously along the peptide backbone [12]. Tn is comprised of the first core sugar GalNAc-α-O-linked to serine or threonine residues in the core peptide of the mucin molecule (Tn, GalNAc-α-O-Ser/Thr). Tn is the precursor to TF, which is generated by the addition of β galactose by galactosyltransferase [13] (TF, Gal β1-3 GalNAc-α-O-Ser/Thr). As human epithelial carcinoma cells may be deficient in a key galactosyltransferase [13,14], Tn is exposed, and premature sialylation of Tn can occur leading to the formation

*To whom correspondence should be addressed: A. Patricia Campbell, Tel.: (206) 685-2468; Fax: (206) 685-3252;
E-mail: apc@u.washington.edu

of STn [11,14] (STn, sialyl α2-6 GalNAc-α-O-Ser/Thr). All three epitopes, Tn, TF, and STn, are strongly expressed on human carcinoma cells [11,15–19], and may be associated with cancer progression and metastasis of human carcinoma [20–22]. Expression of STn, in particular, has been linked to poor prognosis in cancer patients [23–29], suggesting that the STn epitope may have a functional significance in the metastatic cascade [30–32].

Synthetic oligosaccharide vaccines based on the exposed STn core carbohydrate epitopes are being evaluated as potential immunotherapeutics in the treatment of mucin-expressing adenocarcinomas. The choice of STn as a target for immunotherapy is based on its involvement in the aggressive spread of STn positive cancers and its possible role in metastasis [23–32]. Two approaches have been taken in the design of STn carbohydrate vaccines: a monomer approach using isolated STn, and a cluster approach in which dimeric or trimeric clusters of STn are attached to a serine peptide scaffold [33]. The monomer approach has already produced the Theratope® cancer vaccine [34,35], a single STn disaccharide epitope conjugated to KLH through a crotyl linker arm. This vaccine is currently in Phase III clinical trials. The cluster approach is based on the assumption that STn carbohydrates exist also as clusters [33]. Indeed, there is evidence suggesting that a transformation to malignancy is accompanied by the appearance of STn epitopes, both as isolated monomers and as clustered dimers and trimers, on the cancer-associated mucin [36].

From the perspective of immunotherapy, the major difference between a monomer versus cluster approach for STn-based vaccine design is in the specificity of the immune response elicited. For example, Mab B195.3R [37,38], a monoclonal antibody raised against synthetic STn-KLH, and Mab B239 [38], a synthetic trimeric STn cluster specific antibody, both showed reactivity towards human colon carcinoma line LS cells that express STn [36]. These two antibodies display little structural cross reactivity with each other [38]. Their reactivity towards tumor cells, however, indicates the presence of clustered and non-clustered STn disaccharides on the cancer-associated mucin [36–38]. The cluster specific antibody appears to interact with two or more STn epitopes without directly contacting the core protein [38]. This antibody also recognizes STn chemically linked in high densities to protein carriers [38]. Thus, the therapeutic utility of the STn antigen may lie in generating both monomeric and cluster-specific antibodies using their respective synthetic antigens.

A knowledge of the conformation of the tumor cell surface STn is integral to an understanding of the basis for cancer cell specificity of STn-reactive Mabs. Defining the conformation of tumor cell surface STn is also important for the development of Mabs, and ultimately for designing potent immunospecific cancer vaccines against STn-positive cancers. In this paper, we report the results of NMR studies and MD simulations designed to probe the conformation of the tumor cell surface STn as it may exist in a cluster formation. Three model peptides/glycopeptides were studied: (1) a serine trimer containing no carbohydrate groups ($(Ser)_3$ trimer), (2) a serine trimer containing three Tn carbohydrates α-linked to the hydroxyls of adjacent serine sidechains ($(Ser.Tn)_3$ trimer), and (3) a serine trimer containing three STn carbohydrates α-linked to the hydroxyls of adjacent serine sidechains ($(Ser.STn)_3$ trimer). A comparison of the NMR and MD results obtained for these three peptides allowed the effect of carbohydrate clustering and carbohydrate size on peptide backbone conformation and dynamics to be studied, and the effect of carbohydrate clustering on the presentation of the carbohydrate antigen itself to be inferred. The results of these studies are discussed within the framework of developing clustered STn-based vaccines, capable of targeting the clustered STn epitopes on the cancer-associated mucin.

Materials and methods

Sample preparation for NMR

The synthetic glycopeptides, $(Ser)_3$ trimer, $(Ser.Tn)_3$ trimer and $(Ser.STn)_3$ trimer were provided by Biomira Inc., Edmonton, Alberta, Canada. NMR samples were prepared by dissolving the trimer peptides in 500 µl of 90% H_2O/10% D_2O PBS buffer to a concentration of 2 mM. DSS was added as an internal chemical shift reference and the final pH was adjusted to 5.2.

NMR spectroscopy

^1H NMR spectra for the three trimer peptides ($(Ser)_3$ trimer, $(Ser.Tn)_3$ trimer and $(Ser.STn)_3$ trimer) were acquired at 300 MHz using a Varian Unity 300 spectrometer. The hypercomplex method was used for acquisition of two dimensional experiments [39] which included DQFCOSY [40,41], TOCSY [42], and ROESY [43] experiments. All two dimensional experiments were acquired at both 5 and 30°C in order to temperature shift the water peak and allow observation of carbohydrate and peptide resonances between 4.7 and 5.0 ppm. Suppression of the water peak was attained using presaturation for 2.5 sec at a power level of 26 dB. All two dimensional experiments typically incorporated 32 transients for each of 300 increments with 2048 data points along t_2, and a spectral width of 4000 Hz along both dimensions. The TOCSY employed a spin lock field of 7.12 kHz, whereas the ROESY employed a spin lock field of 4.51 kHz and was performed at three mixing times: 150, 300 and 500 msec. The Fourier transformation of the spectra utilized sine bell processing with a phase shift of 90°, and zero filling to 4K × 4K. Chemical shifts were measured from one dimensional experiments, which were acquired with 16,384 data points, and using a spectral width of 4000 Hz.

Temperature coefficients and coupling constants

Temperature coefficients ($-\Delta\delta/\Delta T$ ppb) for all backbone amide (NH) protons in the three trimer peptides were calculated from linear plots of NH chemical shift versus temperature. The NH chemical shifts were measured from DQFCOSY spectra acquired at 5 and 30°C and from one dimensional spectra acquired at 5, 10, 15, 20, 25 and 30°C. $^3J_{N\alpha}$ coupling constants were obtained from the backbone NH proton region of the one dimensional spectra acquired at 5°C, where the individual resonances were curve-fitted using a program written by R. Boyko (University of Alberta; xcrvfit program available at www.pence.ualberta.ca.ftp) which utilizes an iterative fitting procedure. Where overlap of the NH resonances in the one dimensional spectrum precluded analysis, $^3J_{N\alpha}$ coupling constants were obtained from the αCH-NH fingerprint region of the $2K \times 2K$ ($F2 \times F1$) DQFCOSY spectrum acquired at 5°C. The DQFCOSY spectrum was zero filled to 16K in the F2 dimension and processed using 90° shifted sinebell weighting in the F1 dimension and no weighting in the F2 dimension. Traces were taken in $\omega 2$ and then curve-fitted as described above.

Molecular dynamics simulations

MD simulations were performed both *in vacuo* and with explicit water using the Insight II/Discover software package (Molecular Simulations Inc., San Diego, CA) on an SGI Octane. All simulations were performed for unglycosylated and glycosylated serine hexamers ($(Ser)_6$, $(Ser.Tn)_6$ and $(Ser.STn)_6$), as opposed to trimers, since a conformational preference for extended versus helical secondary structure is more easily detected for sequences greater than five residues in length. The CVFF forcefield was used in each simulation, as it contains the appropriate parameters and charges for the two glycosides, Tn and STn [44–46].

MD simulations *in vacuo* used the following protocol. The Biopolymer module of Insight II was first used to generate three different sets of starting coordinates for each hexamer; one set of coordinates corresponding to α-helix (ϕ, $\psi = -65°$, $-40°$), one set corresponding to β-sheet (ϕ, $\psi = -120°$, $+120°$), and one set corresponding to extended strand (ϕ, $\psi = 180°$, $-180°$). Each of the nine starting structures were then energy minimized using conjugate gradients until the maximum derivative was < 0.05 Kcal-mol^{-1}. This was followed by MD for 1 ns at constant volume and temperature (278°K) with a 1 fs time step. A distance dependent dielectric was used to simulate solvent effects, and a 12 Å cutoff was used for non-bonded interactions. Structures were saved every 10 ps.

When the trajectories were compared for the *in vacuo* simulations, it became evident that the lowest energies were consistently observed in the trajectories starting from the β-sheet coordinates (the 'β-sheet trajectories'). This was true for each hexamer examined ($(Ser)_6$, $(Ser.Tn)_6$ or $(Ser.STn)_6$). Thus, only one water simulation was performed for each hexamer, using the lowest energy structure observed in the *in vacuo* β-sheet trajectory as the starting structure. These three starting structures were first placed in a 10 Å water box, invoked by the SOAK option in InsightII. Each component of the system (both water and hexamer) was then energy minimized using 100 steps of steepest descents, followed by 500 interations of conjugate gradient minimization. Finally, MD was performed for 1 ns at constant volume and temperature (278°K), with a 1.5 fs time step. Periodic boundary conditions were used with a 10 Å non-bonded cutoff and a dielectric of 1. Structures were saved every 10 ps.

Results

Resonance assignments

Table 1 displays the ^{1}H NMR resonance assignments for the $(Ser)_3$ trimer, $(Ser.Tn)_3$ trimer and $(Ser.STn)_3$ trimer peptides, including assignments for the Tn and STn carbohydrate resonances as well as for the alkyl ether group N-linked to the Ser3 C-terminal carbonyl of each peptide. Figure 1 shows the structure of the $(Ser)_3$ trimer (1A), the structure of the Tn (GalNAc-α-O-) antigen in the $(Ser.Tn)_3$ trimer (1B), and the structure of the STn (sialyl α2-6 GalNAc-α-O-) antigen in the $(Ser.STn)_3$ trimer (1C). The GalNAc proximal sugar of both the Tn and STn moieties is attached to the serine sidechain hydroxyls via α-linkages. The Tn and STn resonances were assigned on the basis of previously published ^{1}H NMR assignments of mucous glycoproteins [47] and other MUC-1 derived peptides [48].

Protection from solvent exchange

Figure 2 shows the backbone amide region of the one dimensional spectra of the $(Ser)_3$ trimer (2A), $(Ser.Tn)_3$ trimer (2B), and $(Ser.STn)_3$ trimer (2C) at 5°C and 30°C. The NH proton resonances of all three peptides are observed to titrate upfield with increasing temperature (a feature of all exchangeable protons). However, the resonances at the higher temperature vary in intensity for each peptide. For the unglycosylated $(Ser)_3$ trimer (2A), the NH proton resonances decrease in intensity with temperature, indicating that they are relatively unprotected from solvent exchange at the higher temperature. For the glycosylated $(Ser.Tn)_3$ (2B) and $(Ser.STn)_3$ (2C) trimers, the NH proton resonances retain their intensity, and even increase in intensity and resolution in the case of the $(Ser.STn)_3$ trimer. This retention of NH intensity at higher temperatures indicates that the presence of a bulky carbohydrate on the serine side chains protects the backbone NH protons from solvent exchange at higher temperatures, which is an indication of decreased backbone exposure. This phenomenon must involve a steric effect, since the larger STn (a disaccharide, versus Tn, a monosaccharide) provides the greatest protection from solvent exchange. However, other effects may also contribute to the increased lifetimes of these exchangeable backbone

Table 1. ^{1}NMR resonance assignments of trimer peptides (A) $(ser)_3$ trimer, (B) $(ser.Tn)_3$ trimer and (C) $(ser.STn)_3$ trimer

Residue	*NH*	*α-H*	*β*-H	*Others*
A. $(ser)_3$ trimer, pH 5.2, 5°C				
Ser1	8.50, 8.31(30°C)	4.48	3.89	NAcCH3 2.07
Ser2	8.65, 8.45(30°C)	4.53	3.92	
Ser3	8.41, 8.26(30°C)	4.44	3.87	
alkyl ether	8.08, 7.93(30°C)			Ha 3.40(2); Hb 3.57(2); Hc 3.96(2); Hd 5.58(1); He 5.82(1); Hf 1.69(3)
B. $(ser.Tn)_3$ trimer, pH 5.2, 5°C				
Ser1	8.57, 8.39(30°C)	4.66	3.78, 3.96	NAcCH3 2.06
Ser2	8.86, 8.63(30°C)	4.74	3.80, 4.00	
Ser3	8.86, 8.62(30°C)	4.61	3.74, 3.93	
alkyl ether	8.46, 8.26(30°C)			Ha,a′ 3.32, 3.48(2); Hb 3.56(2); Hc 3.95(2); Hd 5.58(1); He 5.82(1); Hf 1.69(3)

Tn chemical shifts

GalNAc	*H1*	*H2*	*H3*	*H4*	*H5*	*H6,H6′*	*Ac*	*AcNH*
(Ser 1)	4.85	4.16	3.90	3.97	3.86	3.75	2.05	8.06, 7.87(30°C)
(Ser 2)	4.85	4.14	3.90	3.97	3.86	3.75	2.05	8.10, 7.91(30°C)
(Ser 3)	4.89	4.17	3.91	3.99	3.89	3.75	2.05	8.17, 7.95(30°C)

Residue	*NH*	*α-H*	*β-H*	*Others*
C. $(ser.STn)_3$ trimer, pH 5.2, 5°C				
Ser1	8.56, 8.40(30°C)	4.73	3.83, 4.00	NAcCH3 2.09
Ser2	8.74, 8.51(30°C)	4.75	3.81, 4.06	
Ser3	8.83, 8.63(30°C)	4.61	3.70, 3.99	
alkyl ether	8.44, 8.24(30°C)			Ha,a′ 3.29, 3.45(2); Hb 3.52(2); Hc 3.95(2); Hd 5.58(1); He 5.82(1); Hf 1.69(3)

STn chemical shifts

GalNAc	*H1*	*H2*	*H3*	*H4*	*H5*	*H6,H6′*	*Ac*	*AcNH*
(Ser 1)	4.89	4.15	3.88	3.99			2.03	8.03, 7.85(30°C)
(Ser 2)	4.89	4.14	3.91	3.99			2.04	8.03, 7.86(30°C)
(Ser 3)	4.84	4.16	3.92	3.99			2.06	8.18, 7.94(30°C)

Sialate	*H3,H3′*	*H4*	*H5*	*H6*	*H7*	*H8*	*H9,H9′*	*Ac*	*AcNH*
(Ser1)	2.73, 1.69	3.68	3.81	3.69	3.57	3.88	3.87, 3.63	2.03	8.18, 8.01(30°C)
(Ser2)	2.73, 1.69	3.68	3.81	3.69	3.57	3.88	3.87, 3.63	2.03	8.18, 8.01(30°C)
(Ser3)	2.73, 1.69	3.68	3.81	3.69	3.57	3.88	3.87, 3.63	2.03	8.18, 8.01(30°C)

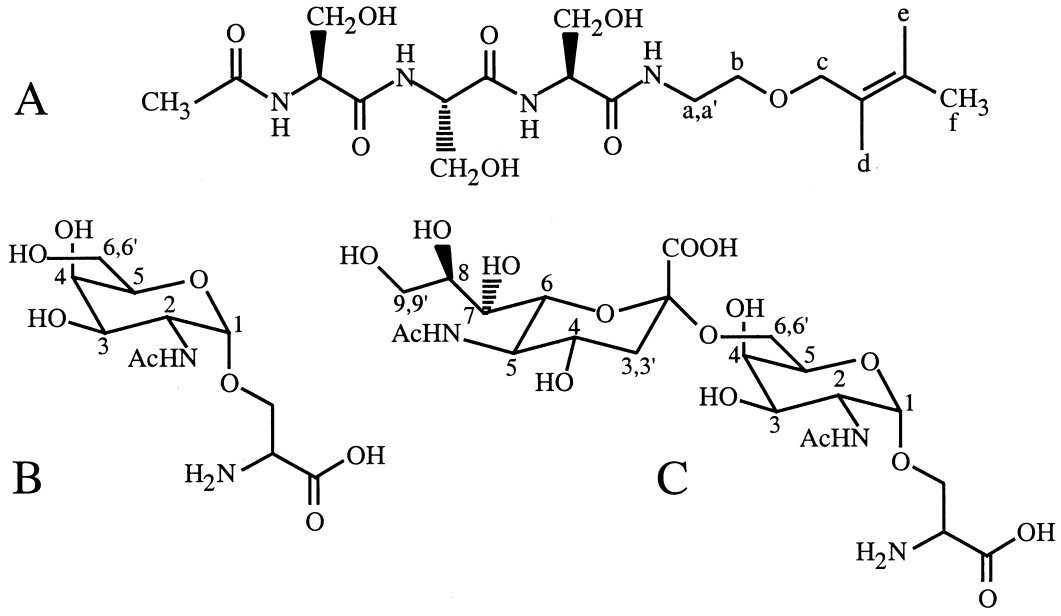

Figure 1. Structures of the (A) N-acetylated $(Ser)_3$ trimer (Ser-Ser-Ser) showing the alkyl ether group N-linked to the Ser3 C-terminal carbonyl (B) Tn (GalNAc-α-O-Ser) antigen attached to the serine β-hydroxyls in the $(Ser.Tn)_3$ trimer, (C) STn (sialyl α2-6 GalNAc-α-O-Ser) antigen attached to the serine β-hydroxyls in the $(Ser.STn)_3$ trimer. The protons of each sugar group are numbered conventionally, and the protons of the alkyl ether group are alphabetically lettered a–f.

NH protons – for example, the existence of hydrogen bonds between the backbone NH protons and the proximal GalNAc groups in the $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers. There is evidence from other groups to support this hypothesis. In their NMR studies of the glycosylated trimer Ac-Thr(GalNAc-α-O)-Ala-Ala-OMe, Mimura and co-workers [49] found coupling constant and temperature coefficient evidence for intramolecular hydrogen bonding between the backbone NH protons of the peptide and the N-acetyl carbonyl of the carbohydrate moiety. In addition, Butenhof and Gerken [50] found evidence of similar peptide-carbohydrate hydrogen

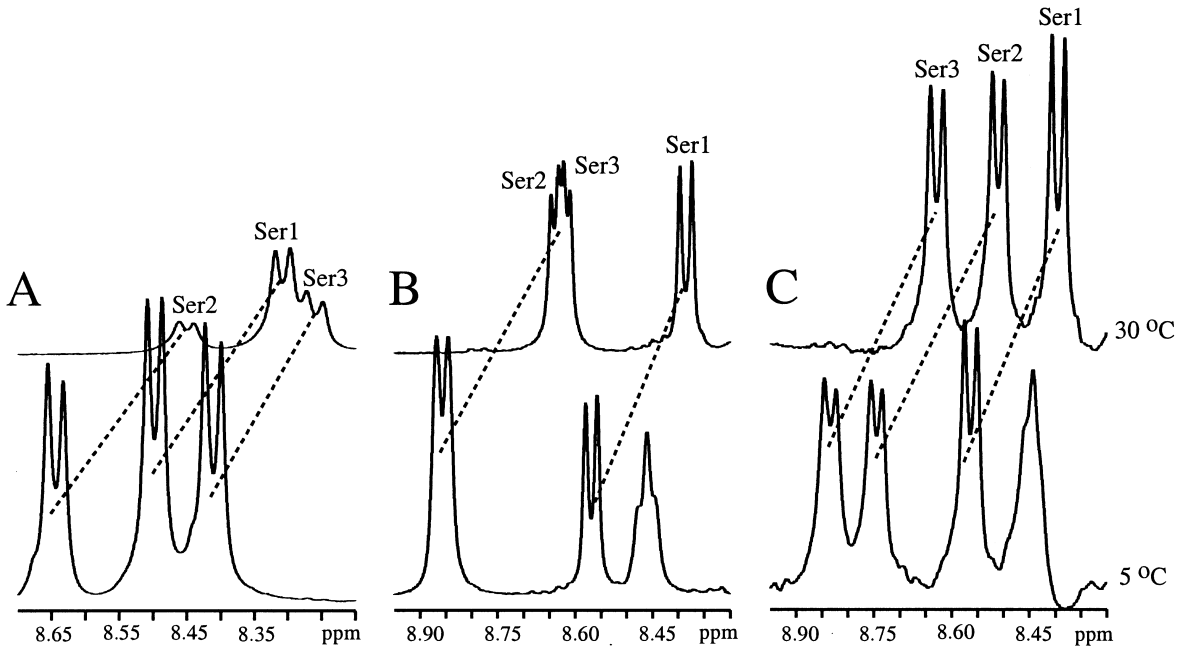

Figure 2. Backbone amide region of the one dimensional spectra of trimer peptides (A) $(Ser)_3$ trimer, (B) $(Ser.Tn)_3$ trimer, and (C) $(Ser.STn)_3$ trimer at 5°C (bottom spectra) and 30°C (top spectra). All three peptides are in 90% H_2O/10% D_2O, PBS buffer, pH 5.2.

Table 2. Comparison of temperature coefficients ($-\Delta\ \delta/\Delta$T)[1], coupling constants ($^3J_{N\alpha}$)[2] and alpha proton chemical shift index ($\Delta\delta\alpha$)[3] measured for the $(ser)_3$ trimer, $(ser.Tn)_3$ trimer and $(ser.STn)_3$ trimer peptides

	$(ser)_3$ trimer			*$(ser.Tn)_3$ trimer*			*$(ser.STn)_3$ trimer*		
Residue	*$\Delta\delta\alpha$ (ppm)*	*$^3J_{N\alpha}$ (Hz)*	*$-\Delta\delta/\Delta T$ (ppb/K)*	$\Delta\delta\alpha$	$^3J_{N\alpha}$	$-\Delta\delta/\Delta T$	$\Delta\delta\alpha$	$^3J_{N\alpha}$	$-\Delta\delta/\Delta T$
Ser1	0.01	6.6	8.3	0.19	6.8	7.7	0.26	7.3	7.1
Ser2	0.05	6.8	8.7	0.27	6.9	10.0	0.28	7.0	10.0
Ser3	−0.03	7.1	6.5	0.14	7.2	10.5	0.13	7.3	8.9
alkyl ether	–	–	6.5	–	–	8.7	–	–	8.7

[1]Uncertainty in the measured temperature coefficient values is ±0.2 ppb/K.
[2]Uncertainty in the measured coupling constants is ±0.3 Hz.
[3]The chemical shift index for the ser α proton is calculated as the difference between the observed chemical shift and the random coil chemical shift ($\Delta\delta\alpha = \delta\alpha$ obs $-\delta\alpha$ coil) where the random coil shift is taken as 4.7 ppm for a serine residue in a random coil environment at 25°C [53].

bonds in their stochastic dynamics simulations of model mucin glycopeptides.

Chemical shifts and secondary structure

Backbone NH and α proton chemical shifts are sensitive to chemical environment and local secondary structure. Generally, downfield shifts from random coil (more positive values) are observed for the backbone amide and α protons in extended configurations such as β-sheet, and upfield shifts (more negative values) are observed for helical configurations [51]. A preference for the glycosylated $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers to adopt an extended conformation is suggested by the amide proton chemical shifts measured for the three trimers (see Table 1). For example, the amide proton shift of Ser3 increases from 8.41 ppm for the $(Ser)_3$ trimer to 8.86 ppm for the $(Ser.Tn)_3$ trimer and 8.83 ppm for the $(Ser.STn)_3$ trimer, and similar trends are observed for the amide protons of the other two serines in the sequence. However, care must be taken not to overinterpret NH chemical shifts on the basis of secondary structure alone, as both solvent effects and hydrogen bonding effects can contribute to the observed chemical shift [52] (see the 'Temperature coefficient' subsection, under Results).

Unlike the backbone NH proton, the backbone α proton appears to be uniquely sensitive to local secondary structure [53]. Thus, the chemical shift of the backbone α proton is often used to assess the 'conformational preference' for a defined secondary structure. The conformational preference for extended strand versus helix has been roughly quantitated by calculating the difference between the observed and random coil serine α proton shift ($\Delta\delta\alpha$) for each residue in the three peptides (see Table 2). Wishart *et al.* [51] showed that the α proton of all 20 naturally occurring amino acids experiences a mean shift of $\Delta\delta\alpha = -0.39$ ppm when placed in a helical configuration and a mean shift of $\Delta\delta\alpha = +0.37$ ppm when placed in an extended configuration. From Table 2 it can be seen that all $\Delta\delta\alpha$ values[1] for the $(Ser)_3$ trimer are close to zero, suggesting no 'conformational preference' in solution for extended strand versus helix within a dynamic conformational equilibrium[2]. However the large positive $\Delta\delta\alpha$ values seen for the $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers suggest that the addition of Tn and STn carbohydrate groups to the serine side chains shifts the conformational equilibrium away from helix and toward extended strand, presumably as a result of the steric limitations the bulky carbohydrates impose on helical conformational space. Furthermore, a comparison of the $\Delta\delta\alpha$ values seen for the first serine in the $(Ser.STn)_3$ trimer (0.26 ppm) versus the $(Ser.Tn)_3$ trimer (0.19 ppm) suggests that the bulkier STn group produces a more pronounced shift in equilibrium toward extended conformation.

ROESY connectivities

Analysis of NOE (or ROE) connectivities can be used to assess regions of defined secondary structure in peptides and proteins. However, even for peptides too short to adopt a defined secondary structure, NMR solution state techniques can still be used to assess 'conformational preferences' for helical dihedral space (roughly $-80° < \phi < -45°$, $-70° < \psi < -45°$) versus extended dihedral space (roughly $-180° < \phi < -120°$, $+120° < \psi < +180°$). This is based on the observation of key sequential dNN and dαN connectivities. The observation of strong dαN(i,i+1) NOEs

[1]The chemical shift from random coil value is calculated for each serine α proton as $\Delta\delta\alpha = \delta\alpha_{obs} - \delta\alpha_{rc}$, where $\delta\alpha_{obs}$ is the observed chemical shift, and $\delta\alpha_{rc}$ is the random coil chemical shift, obtained from a table of random coil values in [53].

[2]The flexible $(Ser)_3$ trimer is undergoing a dynamic equilibrium in solution between at least two conformations which are approximated as either helical or extended. No 'conformational preference' means that the $(Ser)_3$ trimer spends a proportionately equal amount of time between helical and extended conformation. Whereas helical and extended conformation are not the only regions of dihedral space available to a short flexible peptide, they are the regions of dihedral space most clearly discriminated by conventional NMR measurements (chemical shifts, NOEs and coupling constants). Within the context of NMR analysis, extended conformation is most commonly understood to mean β-sheet (ϕ, $\psi = -120°$, $+120°$), but could equally mean extended strand (ϕ, $\psi = -180°$, $+180°$) or any other conformation within the allowable extended dihedral space ($-180° < \phi < -120°$, $+120° < \psi < +180°$).

coupled with weak or absent dNN(i,i+1) NOEs suggests a conformational preference for extended dihedral space [54]. In contrast, the observation of weak-medium dαN(i,i+1) NOEs coupled with strong dNN(i,i+1) NOEs suggests a conformational preference for helical dihedral space [54]. For short flexible peptides, both strong dαN(i,i+1) NOEs and strong dNN(i,i+1) NOEs may be observed together. The observation of both types of NOEs in the same NOESY experiment indicates that the peptide has no conformational preference for either helical or extended dihedral space, and spends a proportionately equal amount of time in each conformation.

Figure 3 shows the dNN region of the ROESY spectra of the $(Ser)_3$ trimer (3A), $(Ser.Tn)_3$ trimer (3B) and $(Ser.STn)_3$ trimer (3C) peptides. The spectrum of the $(Ser)_3$ trimer (3A) displays medium dNN(i,i+1) ROEs between Ser1 and Ser2, and Ser2 and Ser3, respectively. Strong sequential dαN ROEs are also observed in the dαN region of the $(Ser)_3$ trimer spectrum (data not shown). This suggests that the $(Ser)_3$ trimer has no conformational preference for either helical or extended dihedral space, and spends a proportionately equal amount of time in each conformation. By contrast, the spectrum of the $(Ser.Tn)_3$ trimer (3B) displays only a very weak dNN(i,i+1) ROE between Ser1 and Ser2, and no sequential dNN ROEs are visible in the spectrum of the $(Ser.STn)_3$ trimer (3C). The absence of sequential dNN ROEs coupled with the observation of strong sequential dαN ROEs in the dαN regions of both the $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimer spectra (data not shown) indicates that the two glycosylated peptides display a strong preference for extended conformation, a conclusion supported by the α proton chemical shift data discussed earlier. Steric effects appear to drive this conformational preference since the weak dNN(i,i+1) ROE observed between Ser1 and Ser2 in the $(Ser.Tn)_3$ spectrum is completely absent in the $(Ser.STn)_3$ trimer spectrum. Thus, an increase in the size of the attached carbohydrate is accompanied by a stronger shift in the equilibrium of the peptide backbone toward the extended state.

Also observed in the ROESY spectra of the $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers are ROEs between the backbone NH protons of these peptides and the methyl and NH protons of the GalNAc N-acetyl groups on the Tn and STn moieties. Similar peptide-sugar connectivities have been observed in other NMR studies of α-linked glycosylated peptides. Live and co-workers observed strong NOEs between the methyl protons of GalNAc acetyl groups and the peptide backbone NH protons of a series of α-linked glyco-pentapeptides [55]. Liang and co-workers observed strong NOEs between the NH proton of GalNAc acetyl groups and the threonine NH proton of an α-linked monoglycosylated hexapeptide [56]. The presence of these specific peptide-sugar ROEs in the glycosylated serine trimers suggests that the N-acetyl of the proximal GalNAc interacts directly with the peptide backbone, possibly through hydrogen bonding between the NH protons of the peptide backbone and the carbonyl on the N-acetyl group of the GalNAc (as suggested earlier). These hydrogen bonds may play a role in influencing both the orientation of the O-linked carbohydrate as well as the conformation of the underlying peptide backbone. In this regard, Mimura and co-workers' molecular modeling studies of their glycosylated trimer Ac-Thr(GalNAc-α-O)-Ala-Ala-OMe [49] demonstrated that the existence of a hydrogen bond between the peptide NH protons

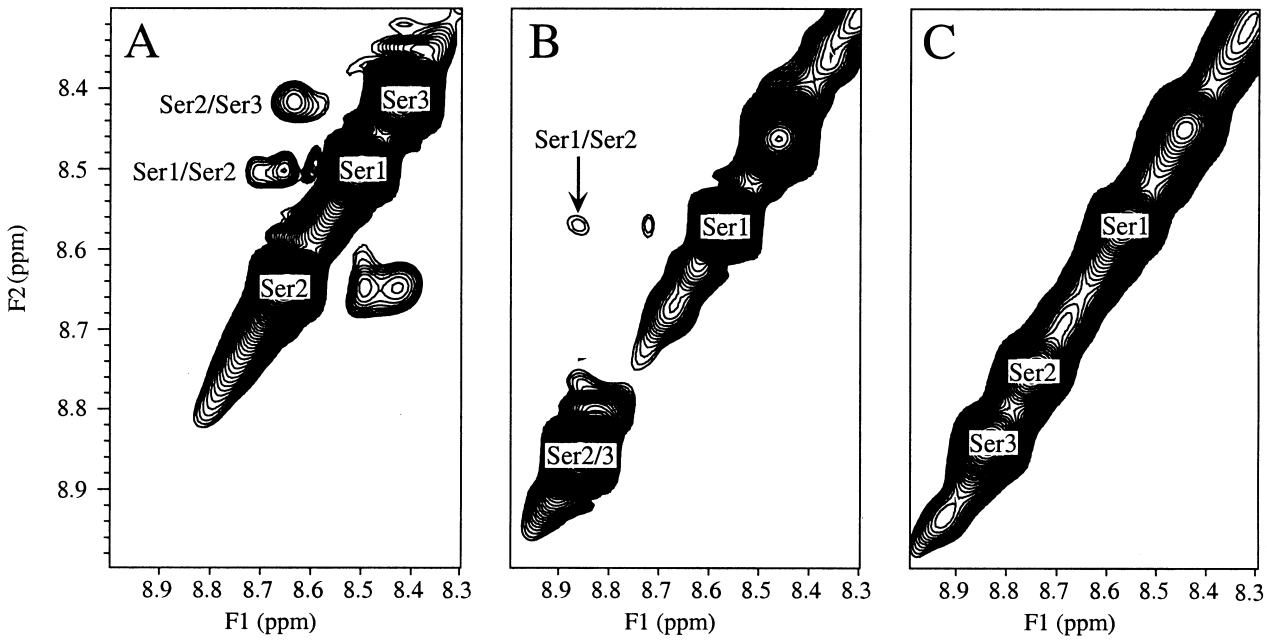

Figure 3. dNN region of the 300 MHz ROESY spectra (150 ms mixing time) of trimer peptides (A) $(Ser)_3$ trimer, (B) $(Ser.Tn)_3$ trimer, and (C) $(Ser.STn)_3$ trimer. All three peptides are in 90% H_2O/10% D_2O, PBS buffer, pH 5.2 at 5°C.

and the GalNAc N-acetyl carbonyl would force the plane of the GalNAc ring and the attached peptide backbone into a roughly perpendicular arrangement. In the case of our glycosylated serine trimers, a similar arrangement would imply an extended peptide backbone conformation, with Tn or STn carbohydrate moieties sticking out at right-angles from the peptide scaffold.

In summary, the ROESY data suggest two mechanisms whereby the attached carbohydrate might influence the conformation of the underlying peptide backbone: (1) a mechanism involving non-specific steric effects which limit the conformational space available to the peptide backbone; and (2) a mechanism involving specific hydrogen bonds that 'lock' the orientation of the proximal GalNAc with respect to the peptide backbone. Either or both of these mechanisms would lead to restricted motions of the underlying peptide backbone, forcing it into an extended conformation so as to minimize steric problems associated with crowding of adjacent O-glycosidic groups.

Temperature coefficients

The temperature dependence of the backbone NH proton chemical shift, or temperature coefficient ($-\Delta\delta/\Delta T$), has been interpreted as a measure of solvent shielding in proteins and folded peptides [57,58]. However, this 'solvent shielding' interpretation of $-\Delta\delta/\Delta T$ assumes that the conformation of the protein or peptide is unaffected by increases in temperature (i.e. does not 'unfold' at higher temperatures). This is not *a priori* true for partially folded peptides that display conformational averaging, as the equilibrium between conformational states changes with increasing temperature as the peptide goes from a folded to a random coil state. In these partially folded systems, no direct correlation between solvent-shielding and $-\Delta\delta/\Delta T$ is observed [59][3]. Instead, a significant correlation is found to exist between $-\Delta\delta/\Delta T$ and the deviation of the NH chemical shift from the random coil reference shift [59]. Upon warming, downfield NH resonances (large deviations) tend to shift upfield more rapidly (large $-\Delta\delta/\Delta T$ values) than NHs with an upfield shift (small deviations; small $-\Delta\delta/\Delta T$ values). Here, the upfield shift of the NH proton reflects temperature-induced shifts in the equilibrium from folded state to random coil state, and $-\Delta\delta/\Delta T$ values become larger as NHs are shifted downfield relative to their random coil shift values due to 'structuring' [59]. Thus, the temperature dependence of the NH proton chemical shift in partially folded peptides is a function of at least two variables: the temperature-dependent equilibrium between the folded and random coil states, and the degree of structuring of the folded state at the lower temperature.

[3]This is true for the serine trimer system studied here, as the solvent shielded backbone NH protons of the glycosylated $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers display larger $-\Delta\delta/\Delta T$ values than the solvent exposed backbone NH protons of the unglycosylated $(Ser)_3$ trimer.

Temperature coefficients were measured for the backbone NH resonances of the $(Ser)_3$, $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimer peptides and these values are displayed in Table 2. As expected for these very short peptides, a good correlation exists between $-\Delta\delta/\Delta T$ and the downfield shift of the NH chemical shift relative to some upfield random coil reference shift (see Table 1). Specifically, the largest $-\Delta\delta/\Delta T$ values correlate with the most downfield NH resonances. As the most downfield NH resonances correspond to serine residues in the $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers, this implies a higher degree of structuring in the glycosylated serine trimers versus the unglycosylated serine trimer (at least at lower temperatures). However, it is difficult to reconcile the presence of downfield shifted NH resonances with structuring in this case, as the concept of a dominant folded state conformer at low temperatures seems inappropriate for peptides too short to adopt a defined secondary structure.

In order to address this apparent incongruity, it is necessary to step back and examine what elements of structuring actually lead to deviations in the observed NH chemical shift from random coil values. The main determinants of NH proton chemical shift in folded peptides and proteins are ring current effects, magnetic anisotropy of the peptide group, hydrogen bonding (electrostatic effects) and solvent effects [52]. Of these, the relative strength of the hydrogen bond seems to be the single most important factor in determining the NH proton chemical shift, with the presence of a strong hydrogen bond correlating to a marked downfield shift (NHs with strong bonding to a backbone carbonyl are far downfield both in β sheets and in the central portions of curved helices [59]). Thus, one way to interpret structuring in the serine trimers is to relate it to the presence or absence of a hydrogen bond – specifically, the hydrogen bond proposed to exist between the NH protons of the peptide backbone and the carbonyl on the N-acetyl group of the GalNAc. In agreement with this interpretation of structuring, the unglycosylated $(Ser)_3$ trimer, which lacks the intramolecular hydrogen bond, displays more upfield shifted NH resonances and therefore smaller $-\Delta\delta/\Delta T$ values. The more 'structured' $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers, with their putative peptide-sugar hydrogen bonds, display more downfield shifted NH resonances and larger $-\Delta\delta/\Delta T$ values. It should be noted that although this interpretation is consistent with the 'structuring' interpretation of $-\Delta\delta/\Delta T$ presented here and elsewhere [59], it does contradict the conventional 'solvent shielding' interpretation of $-\Delta\delta/\Delta T$ [57], which assumes a correlation between small $-\Delta\delta/\Delta T$ and the presence of a hydrogen bond.

Finally, it is worth noting that a comparison of the Ser3 $-\Delta\delta/\Delta T$ for the $(Ser.Tn)_3$ versus the $(Ser.STn)_3$ trimer shows a decrease from 10.5 to 8.9 ppb, respectively (see Table 2). This cannot be interpreted solely on the basis of the structuring arguments presented above, since the Ser3 NH is presumably hydrogen bonded to the N-acetyl carbonyl of the proximal GalNAc sugar in both the $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers (an assumption supported by the similar chemical shifts of Ser3

NH in the $(Ser.Tn)_3$ versus the $(Ser.STn)_3$ trimer; 8.86 ppm versus 8.83 ppm, respectively, at 5°C). This may therefore be a case where solvent shielding effects also contribute to $-\Delta\delta/\Delta T$, so that the significantly smaller $-\Delta\delta/\Delta T$ measured for Ser3 NH in the STn versus Tn glycosylated trimer results from the increased coverage afforded the backbone by the larger STn. Increased solvent shielding by STn would lead to a smaller decrease in Ser3 NH intermolecular versus intramolecular hydrogen bonding with temperature, a smaller shift upfield upon warming, and a smaller $-\Delta\delta/\Delta T$. This interpretation is consistent with the increased intensity of the Ser3 NH proton resonance in the $(Ser.STn)_3$ trimer spectrum at 30 versus 5°C (see Figure 2).

Coupling constants

Table 2 shows the measured $^3J_{N\alpha}$ coupling constants for the serine residues in the various trimers. All $^3J_{N\alpha}$ coupling constant values are greater than 6 Hz and less than 8 Hz, where the measured coupling constants reflect a populated weighted average of the $^3J_{N\alpha}$ values individually associated with the extended (> 9 Hz) and helical (< 5 Hz) conformations [60] of these structurally dynamic peptides in solution. Although the uncertainty in the measured coupling constant is ± 0.3 Hz, a significant increase from 6.6 to 7.3 Hz can be seen for the coupling constant of Ser1 in the $(Ser)_3$ trimer versus $(Ser.STn)_3$ trimer, indicating a shift in the equilibrium population of peptide toward extended dihedral space in the glycosylated peptide.

Using a method developed by Yao *et al*. [61] it is possible to make a rough calculation to quantitate helical versus extended populations of the trimers from the $^3J_{N\alpha}$ coupling constants measured for Ser1. We use a simplistic two-site model in which Ser1 can find itself in either a helical conformation or in an extended conformation. Assuming 100% helical conformation, the coupling constants should be $^3J_{N\alpha} = 4$ Hz corresponding to a $\phi 2 = -60$ [60]. Assuming 100% extended conformation, coupling constants of $^3J_{N\alpha} = 9$ Hz are expected corresponding to a $\phi < -100$ [60]. From these two extreme values of $^3J_{N\alpha}$, the measured $^3J_{N\alpha}$ values for Ser1 in all three peptides is used to estimate the percentage of extended conformation around this residue. In such a way, values are obtained of 50% extended conformation[4] for the $(Ser)_3$ trimer, 60% extended conformation for the $(Ser.Tn)_3$ trimer and 70% extended conformation for the $(Ser.STn)_3$ trimer. This rough estimate suggests that the introduction of bulky carbohydrate groups on the side chain of the serine residues pushes the conformational equilibrium toward extended conformational space (at least around Ser1), as supported by the NH and α proton chemical shift analyses and ROE connectivities.

[4]If the peptide spends equal amounts of time in helical versus extended dihedral space, 50% of the total population of the peptide will be helical (at any instantaneous moment of time), whereas 50% of the population will be extended. Thus, a peptide that displays no conformational preference is, by definition, 50% helix and 50% extended. A peptide that displays a preference for extended conformation is > 50% extended.

Molecular dynamics simulations

In order to further investigate the influence of Tn and STn carbohydrate clustering on the conformational preference of the underlying serine backbone, MD simulations were performed on unglycosylated and glycosylated serine hexamers ($(Ser)_6$, $(Ser.Tn)_6$ and $(Ser.STn)_6$). Hexamers were studied as opposed to trimers, since a conformational preference for extended versus helical secondary structure is more easily detected for sequences greater than five residues in length. Preliminary MD simulations were performed *in vacuo* using three different sets of starting coordinates for each hexamer; one corresponding to α-helix ($\phi, \psi = -65°, -40°$), one to β-sheet ($\phi, \psi = -120°, +120°$), and one to extended strand ($\phi, \psi = -180°, +180°$). MD simulations were then performed in explicit water. These water simulations used the lowest energy structure observed in the *in vacuo* 'β-sheet trajectory' as the starting structure. The MD simulations in water were performed for 1 ns, with structures saved every 10 ps throughout the trajectory. This resulted in 100 structures or 'time points'. A flattening of the trajectory (plotted as backbone Cα RMSD to the starting structure; data not shown) at approximately 200 ps into each simulation suggested that each system had come to equilibration. An analysis of the ensemble of structures sampled after this initial equilibration time then allowed a sampling of the conformational space explored by each hexamer.

Figure 4 shows representative structures taken every 250 ps (250, 500, 750 and 1000 ps) in the 1 ns MD water simulations of the $(Ser)_6$ hexamer (4A), $(Ser.Tn)_6$ hexamer (4B) and $(Ser.STn)_6$ hexamer (4C). Each structure is superposed to the lowest energy structure of the four. From the ensemble of $(Ser)_6$ hexamer structures (4A), it can be seen that the serine backbone is not conformationally restrained and displays no preference in solution for extended structure (includes β-sheet and extended strand) versus helix. In contrast, the ensemble of $(Ser.Tn)_6$ hexamer (4B) and $(Ser.STn)_6$ hexamer structures (4C) demonstrate that the presence of the Tn and STn sugars attached to the serine backbone shifts the conformational equilibrium toward extended structure.

Figure 5 shows 'temporal' Ramachandran (ϕ, ψ) plots of structures taken every 250 ps in the 1 ns MD water simulations for the (A) $(Ser)_6$ hexamer, (B) $(Ser.Tn)_6$ hexamer, and (C) $(Ser.STn)_6$ hexamer. The four time points correspond to the four superimposed structures in Figure 4, where only the internal serines in each peptide (Ser2, Ser3, Ser4 and Ser5) are represented. The Ramachandran plot of the $(Ser)_6$ hexamer (5A) shows that the unglycosylated peptide explores all allowable dihedral space, as evidenced by the large temporal deviations in both dihedral angles. In contrast, a preference for extended structure in the $(Ser.Tn)_6$ and $(Ser.STn)_6$ hexamers (5B and 5C) is observed as a clustering of ϕ, ψ angles within

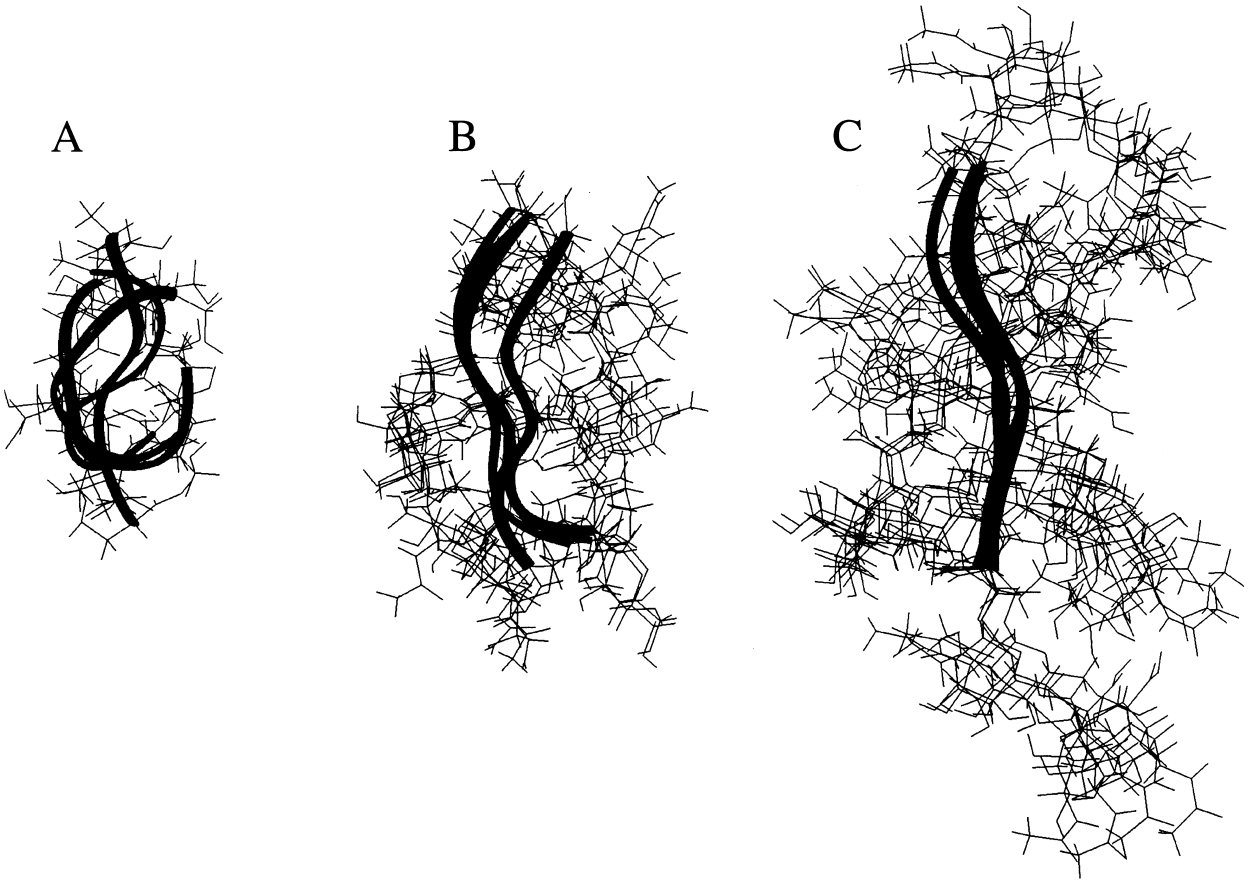

Figure 4. Superposition of structures taken every 250 ps in the 1 ns MD water simulations for the (A) $(Ser)_6$ hexamer, (B) $(Ser.Tn)_6$ hexamer, and (C) $(Ser.STn)_6$ hexamer. Each structure is superposed to the lowest energy structure of the four. Serine backbone is represented in ribbon model, whereas serine sidechain as well as Tn and STn carbohydrate groups are represented in stick model.

the allowable extended dihedral space ($-180^\circ < \phi < -120^\circ$, $+120^\circ < \psi < +180^\circ$). This clustering demonstrates that the addition of Tn and STn carbohydrate groups to the serine side chains restricts the serine backbone to exploring only extended dihedral space. In addition, as the effect of increasing the size of the attached sugar (STn versus Tn; 5C versus 5B) is to further shrink the limits of the dihedral space explored by the serine backbone in the MD simulations, this suggests that steric effects are the major determinants of conformational preference. It should be noted here that although only four time points are represented in Figure 5, these four points accurately reflect the larger ensemble of time points sampled throughout each trajectory. For example, the ϕ, ψ values measured for Ser4 across the latter 80 time points of the $(Ser)_6$ trajectory reveal a roughly equal split between helical and extended conformational space. In contrast, the average ϕ, ψ values measured for Ser4 across the latter 80 time points of the $(Ser.Tn)_6$ and $(Ser.STn)_6$ trajectories are $\phi = -110^\circ \pm 20^\circ$, $\psi = +110^\circ \pm 40^\circ$ and $\phi = -130^\circ \pm 30^\circ$, $\psi = +160^\circ \pm 20^\circ$, respectively.

Finally, the relative exposure of the serine backbone to solvent was assessed from MD structures of the unglycosylated versus glycosylated serine hexamers. Figure 6 shows CPK space-filled representations (panels A–C) and solvent accessible surfaces (panels D–F) of the lowest energy structure in the ensemble for the $(Ser)_6$ hexamer (6A,D), $(Ser.Tn)_6$ hexamer (6B,E), and $(Ser.STn)_6$ hexamer (6C,F). These representations demonstrate the decrease in surface area exposure of the serine backbone as a function of the increasing size of the sugar moiety. The decrease in backbone solvent exposure upon glycosylation is consistent with the significantly increased protection of the backbone NH protons from solvent exchange observed in the NMR experiments of the glycosylated serine trimers. However, it should be noted that the MD simulations detected no evidence of the putative hydrogen bonds between the peptide NH protons and the GalNAc N-acetyl carbonyl earlier proposed to have contributed to this increased solvent protection.

Discussion

Discontinuous carbohydrate clusters, found on the mucins expressed by cancer cells, create a class of distinct spatial epitopes that are not usually found on normal cells [36–38]. These glycosylation sites often exist in succession of two or more [38]. In an effort to probe the structure and dynamics of

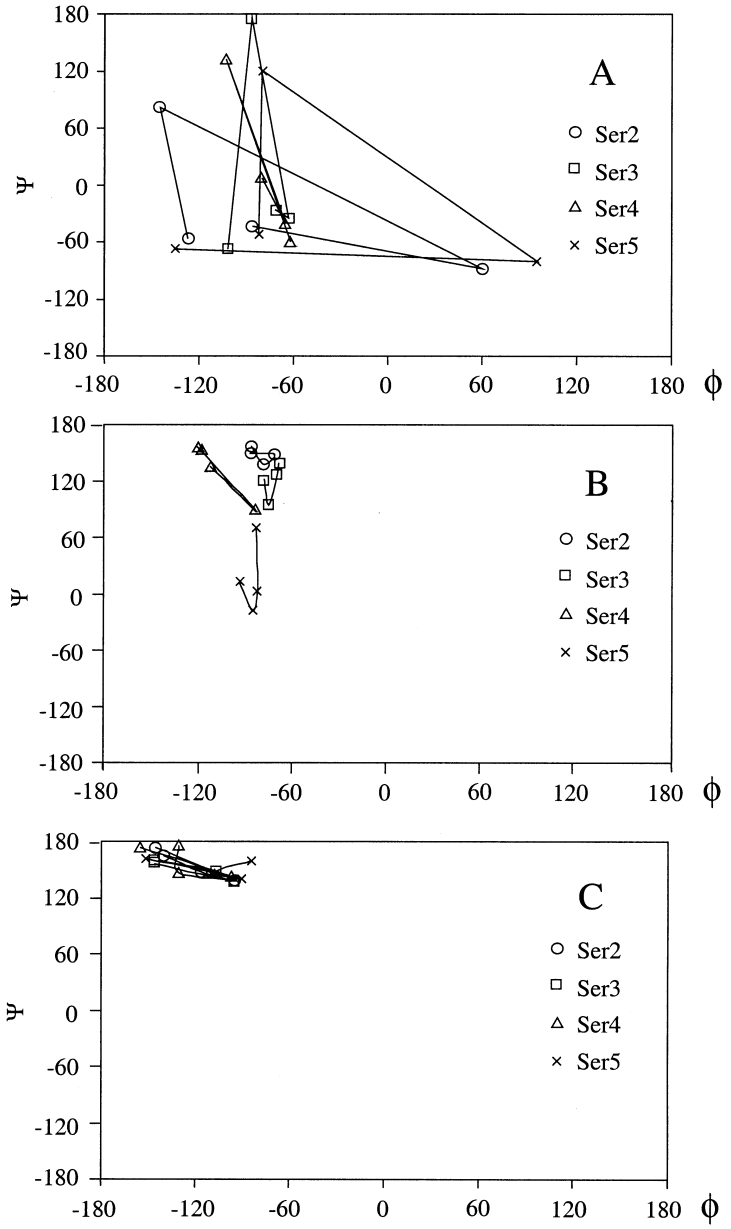

Figure 5. Temporal Ramachandram plots of structures taken every 250 ps in the 1 ns MD water simulations for the (A) $(Ser)_6$ hexamer, (B) $(Ser.Tn)_6$ hexamer, and (C) $(Ser.STn)_6$ hexamer. The four time points correspond to the four superimposed structures in Figure 4. Only the central four serines in each peptide (Ser2, Ser3, Ser4 and Ser5) are represented.

Tn and STn carbohydrate clusters as they may exist on the cancer-associated mucin, we have used NMR spectroscopy and MD simulations to study the effect of O-glycosylation of adjacent serine residues in a repeating (Ser)n sequence. We have found that glycosylation protects the backbone amides from solvent exchange and shifts the conformational equilibrium of the peptide backbone in solution toward extended conformational space. Steric effects appear to drive these changes since an increase in the size of the carbohydrate (and accompanying loss of degrees of conformational freedom) is accompanied by a stronger shift in the equilibrium toward the extended state. These results suggest that the flexible peptide backbone, in spite of its own normal folding preference, adapts to a more rigid and 'rod-like' glycosylation dependent orientation, in order to accommodate a surface covering of carbohydrate structures that are overwhelmingly large in comparison to the surface area of the peptide itself. NMR evidence also points to the formation of hydrogen bonds between the peptide backbone NH protons and the proximal GalNAc groups in the $(Ser.Tn)_3$ and $(Ser.STn)_3$ trimers. The putative peptide-sugar hydrogen bonds may also play a role in influencing the conformation of the underlying peptide backbone, as well as the orientation of the O-linked carbohydrate.

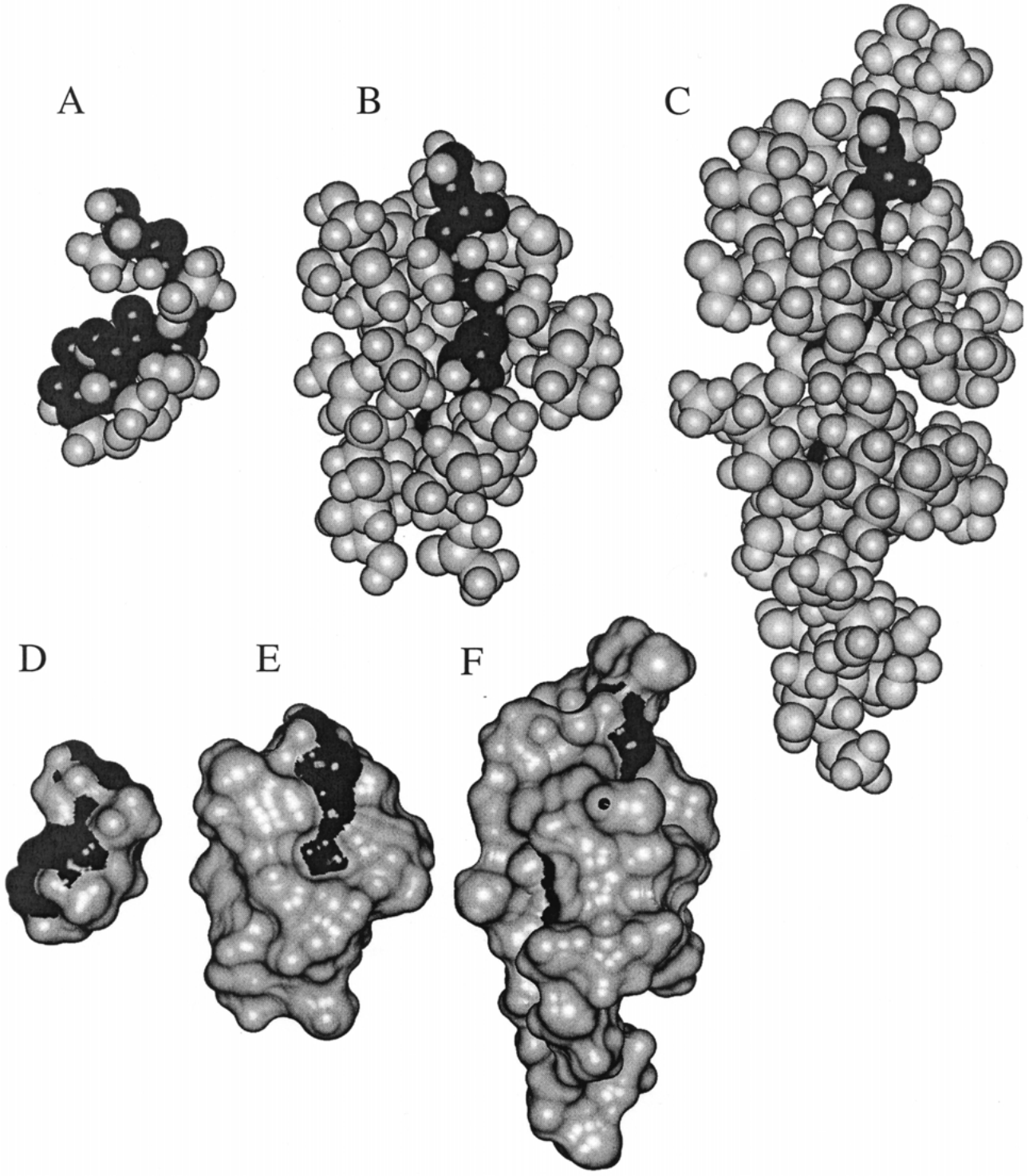

Figure 6. Representations of the lowest energy structure within the entire ensemble of 100 structures for the (A, D) $(Ser)_6$ hexamer, (B, E) $(Ser.Tn)_6$ hexamer, and (C, F) $(Ser.STn)_6$ hexamer from the MD water simulations. Panels A, B and C show CPK-spaced filled representations and panels D, E and F show solvent accessible surfaces of these lowest energy structures. The peptides are shown in side view, with the N-terminus oriented at the top, and the C-terminus at the bottom of each figure. The hexamer backbones are colored black, whereas the sidechains and glycosides are colored light gray. Surfaces in panels D, E and F were created using a test sphere with a 1.4 Å radius.

These findings are in agreement with the results of other NMR and molecular modeling studies of peptides with clustered carbohydrate groups. An NMR study of a synthetic clustered carbohydrate tumor antigen from mucin CD43 showed that glycosylation induced the pentapeptide backbone into a 'rigid scaffold', 'approaching the stability of motifs found in globular proteins' [55]. An NMR/MD study of a decapeptide from human glycophorin A showed that clustering of O-glycosylation sites on adjacent serine and threonine residues induced a rigid 'wave-type' extended structure to the peptide backbone [62]. In general, clustered O-glycosylation of Ser/Thr(/Pro)-rich has been found to impart rigidity to otherwise flexible polypeptide backbones, increasing the stability to local unfolding events in both proteins [63] and peptides [64]. However, even a single site of glycosylation can significantly alter the local conformation and dynamics of the polypeptide backbone. For example, O-glycosylation of a single threonine in a linear hexapeptide was found to 'favor conformations in which the (peptide) backbone bends away from the sugar' [65]. In another example, O-glycosylation of a single threonine in the antimicrobial peptide, drosocin, leads to an increase in the population of extended local structure [66]. The authors of both studies attributed the shift in the conformational equilibrium to a reduction in the number of conformations available to the otherwise flexible peptide [65,66], a conclusion which is consistent with the results presented herein for the mucin system.

The rigid, rod-like' state of the STn-glycosylated serine trimer backbone probably corresponds in secondary conformation to the elongated and densely glycosylated normal-state mucin glycoprotein [67,68]. However, as the distribution of carbohydrate clusters is discontinuous in the cancer-associated state, these regions of elongated and densely glycosylated structure are likely to be well separated from each other by regions of flexible and solvent-exposed peptide backbone. This distribution gives rise to distinct carbohydrate epitopes, spatially separated down the length of the mucin peptide backbone. One can envision the immunological significance of this arrangement in a number of ways. First, an optimal separation of individual carbohydrate clusters might allow immunodetection of multiple carbohydrate epitopes on the same mucin molecule at the same time, which could serve to amplify the immune response. Secondly, the clustering of the carbohydrates into groups of two or more might allow the carbohydrate epitope to be presented as a multivalent ligand to the immune system, which could serve to increase the specificity of the immune response. Finally, increased rigidity of the peptide scaffold supporting the carbohydrate cluster might reduce the entropy loss associated with binding to components of the immune system. The resulting increases in binding affinities could ultimately translate into a stronger immune response targeting the cancer-associated mucin.

In summary, this study has shown that carbohydrate clustering shifts the conformational equilibrium of the underlying peptide backbone into a more extended and rigid state, so as to minimize the steric crowding of the proximal carbohydrates structures. A more rigid and extended peptide platform probably also serves to present the clustered carbohydrate antigen to the immune system. Thus, a design strategy that incorporates conformational constraints into the peptide scaffold so as to present clustered carbohydrate epitope may be key to the design of STn-based vaccines, capable of targeting clustered STn epitopes on the cancer-associated mucin.

Acknowledgements

The authors would like to thank Dr. Brian D. Sykes for the gracious use of his NMR, computer and wet laboratory facilities. The authors would also like to thank Roger Armen for helpful suggestions regarding the MD simulations, and Valerie Daggett for critical reading of the manuscript. This work was supported by a U.S. Army Breast Cancer Grant (DAMD17-99-1-9437) and by the American Association of Colleges of Pharmacy.

References

1 Taylor-Papadimitriou J, Burchell J, Miles DW, Dalziel M, *Biochim Biophys Acta* **1455**, 301–13 (1999).
2 Miles DW, Taylor-Papadimitriou J, *Pharmacol Ther* **82**, 97–106 (1999).
3 Apostolopoulos V, Sandrin MS, McKenzie IF, *J Mol Med* **77**, 427–36 (1999).
4 Apostolopoulos V, Pietersz GA, McKenzie IF, *Curr Opin Mol Ther* **1**, 98–103 (1999).
5 van den Steen P, Rudd PM, Dwek RA, Podenakker G, *Crit Rev Biochem Mol Biol* **33**, 151–208 (1998).
6 Rudd PM, Dwek RA, *Crit Rev Biochem Mol Biol* **32**, 1–100 (1997).
7 Carlstedt I, Davies JR, *Biochem Soc Trans* **25**, 214–9 (1997).
8 Koganty RR, Reddish MA, Longenecker BM. In *Glycopeptides and Related Compounds: Synthesis, Analysis and Application*, edited by Large DG, Warren CD (Dekker, New York, 1997), pp. 707–43.
9 Apostolopoulos V, McKenzie IF, *Crit Rev Immunol* **14**, 293–309 (1994).
10 Girling A, Bartkova J, Burchell J, Gendler S, Gillett C, Taylor-Papadimitriou J, *Int J Cancer* **43**, 1072–6 (1989).
11 Itzkowitz SH, Yuan M, Montgomery CK, Kjeldsen T, Takahashi HK, Bigbee WL, Kim YS, *Cancer Res* **49**, 197–204 (1989).
12 Ogata S, Koganty R, Reddish M, Longenecker BM, Chen A, Perez C, Itzkowitz SH, *Glycoconj J* **15**, 29–35 (1998).
13 Zhuang D, Yousefi S, Dennis JW, *Cancer Biochem Biophys* **12**, 185–98 (1991).
14 Brockhausen I, Yang J, Dickinson N, Ogata S, Itzkowitz SH, *Glycoconj J* **15**, 595–603 (1998).
15 Cao Y, Karsten U, Otto G, Bannasch P, *Virchows Arch* **434**, 503–9 (1999).
16 Cao Y, Schlag PM, Karsten U, *Virchows Arch* **431**, 159–66 (1997).
17 Springer GF, *J Mol Med* **75**, 594–602 (1997).

18 Springer GF, *Crit Rev Oncog* **6**, 57–85 (1995).
19 Springer GF, *Science* **224**, 1198–1206 (1984).
20 Kishikawa T, Ghazizadeh M, Sasaki Y, Springer GF, *Jpn J Cancer Res* **90**, 326–32 (1999).
21 Terasawa K, Furumoto H, Kamada M, Aono T, *Cancer Res* **56**, 2229–32 (1996).
22 David L, Nesland JM, Clausen H, Carneiro F, Sobrinho-Simoes M, *APMIS Suppl* **27**, 162–72 (1992).
23 Takao S, Uchikura K, Yonezawa S, Shinchi H, Aikou T, *Cancer* **86**, 1966–75 (1999).
24 Imada T, Rino Y, Hatori S, Takahashi M, Amano T, Kondo J, Suda T, *Hepatogastroenterology* **46**, 208–14 (1999).
25 Terashima S, Takano Y, Ohori T, Kanno T, Kimura T, Motoki R, Kawaguchi T, *Surg Today* **28**, 682–6 (1998).
26 Werther JL, Tatematsu M, Klein R, Kurihara M, Kumagai K, Llorens P, Guidugli Neto J, Bodian C, Pertsemlidis D, Yamachika T, Kitou T, Itzkowitz S, *Int J Cancer* **69**, 193–9 (1996).
27 Miles DW, Linehan J, Smith P, Filipe I, *Br J Cancer* **71**, 1074–6 (1995).
28 Takahashi I, Maehara Y, Kusumoto T, Kohnoe S, Kakeji Y, Baba H, Sugimachi K, *Br J Cancer* **69**, 163–6 (1994).
29 Kobayashi H, Terao T, Kawashima Y, *J Clin Oncol* **10**, 95–101 (1992).
30 Soares R, Marinho A, Schmitt F, *Pathol Res Pract* **192**, 1181–6 (1996).
31 Terasawa K, Furumoto H, Kamada M, Aono T, *Cancer Res* **56**, 2229–32 (1996).
32 Bresalier RS, Ho SB, Schoeppner HL, Kim YS, Sleisenger MH, Brodt P, Byrd JC, *Gastroenterology* **110**, 1354–67 (1996).
33 Ragupathi G, Howard L, Cappello S, Koganty RR, Qiu D, Longenecker BM, Reddish MA, Lloyd KO, Livingston PO, *Cancer Immunol Immunother* **48**, 1–8 (1999).
34 MacLean GD, Miles DW, Rubens RD, Reddish MA, Longenecker BM, *J Immunother Emphasis Tumor Immunol* **19**, 309–16 (1996).
35 Reddish MA, MacLean GD, Poppema S, Berg A, Longenecker BM, *Cancer Immunol Immunother* **42**, 303–9 (1996).
36 Ogata S, Koganty R, Reddish M, Longenecker BM, Chen A, Perez C, Itzkowitz SH, *Glycoconj J* **15**, 29–35 (1998).
37 Reddish MA, Jackson L, Koganty RR, Qiu D, Hong W, Longenecker BM, *Glycoconj J* **14**, 549–60 (1997).
38 Zhang S, Walberg LA, Ogata S, Itzkowitz SH, Koganty RR, Reddish M, Gandhi SS, Longenecker BM, Lloyd KO, Livingston PO, *Cancer Res*, **55**, 3364–8 (1995).
39 States DJ, Haberkorn RA, Ruben DJ, *J Magn Reson* **48**, 286–92 (1982).
40 Piatini U, Sorenson OW, Ernst RR, *J Am Chem Soc* **104**, 6800–1 (1982).
41 Rance M, Sorenson OW, Bodenhausen G, Wagner G, Ernst RR, Wüthrich K, *Biochem Biophys Res Commun* **117**, 479–85 (1983).
42 Bax A, Davis DG, *J Magn Reson* **65**, 355–60 (1985).
43 Kessler H, Griesinger C, Kerssebaum R, Wagner K, Ernst RR, *J Am Chem Soc* **109**, 607–9 (1987).
44 Dauber-Osguthorpe P, Roberts VA, Osguthorpe DJ, Wolff J, Genest M, Hagler AT, *Proteins* **4**, 31–47 (1988).
45 Balaji PV, Qasba PK, Rao VS, *Glycobiology* **4**, 497–515 (1994).
46 Kozar T, Tvaroska I, Carver JP, *Glyconjugate J* **15**, 187–91 (1998).
47 Gerken TA, *Arch Biochem Biophys* **247**, 239–53 (1986).
48 Liu X, Sejbal J, Kotovych G, Koganty RR, Reddish MA, Jackson L, Gandhi SS, Mendonca AJ, Longenecker BM, *Glyconjugate J* **12**, 607–17 (1995).
49 Mimura Y, Inoue Y, Maeji NJ, Chujo R, *Int J Pept Prot Res* **34**, 363–8 (1989).
50 Butenhof KJ, Gerken TA, *Biochemistry* **32**, 2650–63 (1993).
51 Wishart DS, Sykes BD, Richards FM, *J Mol Biol* **222**, 311–33 (1991).
52 Osapay K, Case DA, *J Biomol NMR* **4**, 215–30 (1994).
53 Wishart DS, Bigam, CG, Holm A, Hodges RS, Sykes BD, *J Biol NMR* **5**, 67–81 (1995).
54 Dyson HJ, Wright PE, *Annu Rev Biophys Chem* **20**, 519–38 (1991).
55 Live DH, Williams LJ, Kuduk SD, Schwarz JB, Glunz PW, Chen XT, Sames D, Kumar RA, Danishefsky SJ, *Proc Natl Acad Sci USA* **96**, 3489–93 (1999).
56 Liang R, Adreotti AH, Kahne D, *J Am Chem Soc* **117**, 10395–6 (1995).
57 Rose GD, Gierasch LM, Smith JA, *Adv Protein Chem* **37**, 1–109 (1985).
58 Watts CR, Tessmer MR, Kallick DA, *Lett Pept Sci* **2**, 59–70 (1995).
59 Andersen NH, Neidigh JW, Harris SM, Lee GM, Liu Z, Tong H, *J Am Chem Soc* **119**, 8547–61 (1997).
60 Pardi A, Billeter M, Wüthrich K, *J Mol Biol* **180**, 741–51 (1984).
61 Yao J, Feher VA, Espejo BF, Reymond MT, Wright PE, Dyson HJ, *J Mol Biol* **243**, 736–53 (1994).
62 Schuster O, Klich G, Sinnwell V, Kranz H, Paulsen H, Mayer B, *J Biomol NMR* **14**, 33–45 (1999).
63 Van den Steen P, Rudd PM, Dwek RA, Opdenakker G, *Crit Rev Biochem Mol Biol* **33**, 151–208 (1998).
64 Bailey D, Renouf DV, Large DG, Warren CD, Hounsell EF, *Carbohyd Res* **324**, 242–54 (2000).
65 Andreotti AH, Kahne D, *J Am Chem Soc* **115**, 3352–3 (1993).
66 McManus AM, Otvos L, Hoffmann R, Craik DJ, *Biochemistry* **38**, 705–14 (1999).
67 Gururaja TL, Ramasubbu N, Venugopalan P, Reddy MS, Ramalingam K, Levine MJ, *Glycoconj J* **15**, 457–67 (1998).
68 Shogren R, Gerken TA, Jentoft N, *Biochemistry* **28**, 5525–36 (1989).

Received 30 June 2000, revised 8 March 2001, accepted 28 March 2001