ELSEVIER

Biophysical Chemistry 134 (2008) 157–167

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Modulation of protein stability by *O*-glycosylation in a designed Gc-MAF analog

Justin Spiriti [1], Federica Bogani [1], Arjan van der Vaart *, Giovanna Ghirlanda *

Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85287, United States

Received 14 December 2007; received in revised form 28 January 2008; accepted 8 February 2008
Available online 21 February 2008

Abstract

The post-translational modification of proteins by the covalent attachment of carbohydrates to specific side chains, or glycosylation, is emerging as a crucial process in modulating the function of proteins. In particular, the dynamic processing of the oligosaccharide can correlate with a change in function. For example, a potent macrophage-activating factor, Gc-MAF, is obtained from serum vitamin D binding protein (VDBP) by stepwise processing of the oligosaccharide attached to Thr 420 to the core α-GalNAc moiety. In previous work we designed a miniprotein analog of Gc-MAF, MM1, by grafting the glycosylated loop of Gc-MAF on a stable scaffold. GalNAc-MM1 showed native-like activity on macrophages (Bogani 2006, J. Am. Chem. Soc. 128 7142–43). Here, we present data on the thermodynamic stability and conformational dynamics of the mono- and diglycosylated forms. We observed an unusual trend: each glycosylation event destabilized the protein by about 1 kcal/mol. This effect is matched by an increase in the mobility of the glycosylated forms, as evaluated by molecular dynamics simulations. An analysis of the solvent-accessible surface area shows that glycosylation causes the three-helix bundle to adopt conformations in which the hydrophobic residues are more solvent exposed. The number of hydrophobic contacts is also affected. These two factors, which are ultimately explained with a change in occupancy for conformers of specific side chains, may contribute to the observed destabilization.

Keywords: De novo design; GalNAc; Thermodynamic stability; *O*-glycosylation; Molecular dynamics

1. Introduction

Glycoproteins are a class of naturally occurring bioconjugates in which carbohydrates of varied complexity are post-translationally attached to individual amino acids, either asparagine in N-linked glycosylation or serine/threonine in O-linked glycosylation. The carbohydrate component can affect many disparate functions spanning from cell–cell communication events, to modulation of protein-protein interactions, to the emergence of antigenicity in autoimmune and alloimmumune reactions [1,2].

The most prevalent type of O-linked glycosylation is mucin-type glycosylation, which is initiated by the enzymatic attachment of α-*N*-acetylgalactosamine (GalNAc) from UDP-GalNAc to the β-hydroxyl group of either a serine or a threonine in the cytosol. The nascent oligosaccharide is further decorated with the step-wise addition of monomers in the Golgi [1,2]. Although no universal sequence motif for *O*-glycosylation has been identified, the site of attachment is often preceded at the −1 position and/or followed at the +3 position by a proline, increasing the probability of finding the site in a turn position. Glycosylation can affect biophysical properties such as solubility, thermal stability, aggregation, folding, and structural dynamics. For example, natural glycoproteins generally exhibit slightly higher denaturation temperatures and structure dynamics than their deglycosylated counterparts; these effects correlate with the number of glycans bound to the protein surface rather than with their size [3]. Beyond these observations, it can be difficult to dissect the contributions of each glycosylation event in the context of specific structural motifs using natural

* Corresponding authors. G. Ghirlanda is to be contacted at tel.: +1 480 965 6645; fax: +1 480 965 2747. A. van der Vaart, tel.: +1 480 727 9373; fax: +1 480 965 4669.

E-mail addresses: vandervaart@asu.edu (A. van der Vaart), gghirlanda@asu.edu (G. Ghirlanda).

[1] These authors contributed equally to the work.

0301-4622/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2008.02.005

glycoproteins, which are often characterized by differential utilization of multiple glycosylation sites and/or by heterogeneicity of the complex oligosaccharides. Synthetic model proteins can overcome these obstacles, and have been used by several groups to investigate the effect of simple carbohydrate conjugates, which can be either chemically incorporated during solid-phase synthesis [3–5], or biosynthetically using nonsense codon suppression [6].

Dynamic changes in the composition of the oligosaccharide moieties can trigger the loss or acquisition of function in nuclear and cytosolic proteins, in concert with phosphorylation [7,8], and in secreted proteins [9]. Vitamin D binding protein (VDBP, also known as group complement, Gc) is an abundant serum glycoprotein with functions ranging from the transport of vitamin D, to the scavenging of acting monomers from the bloodstream, to the immunomodulation of macrophage activity [10]. The latter function is mediated by the sequential enzymatic cleavage of its trisaccharide moiety to the core α-GalNAc residue linked to Thr 420 [9]. This processing, which involves surface enzymes on B and T lymphocytes, transforms VDBP into a serum factor (Gc-MAF/VDBP-MAF) that stimulates the Fc mediated phagocytic activity of macrophages [11]. Further processing to cleave the GalNAc moiety results in complete loss of activity on macrophages (REF). Gc-MAF has shown promise for the development of therapeutic agents based on its macrophage-activating properties as adjuvant in cancer therapy [10–13].

In previous work, we reported that a miniaturized model of Gc-MAF, called MM1, exhibits native-like activity on macrophages [14]. MM1 was designed by transferring the glycosylated loop of Gc-MAF, believed to be responsible for its biological activity, onto α3W, a well characterized 3-helix bundle used as a scaffold [15]. The process was repeated for both loops connecting helix 1 to helix 2 and helix 2 to helix 3; initially, only loop 1 was glycosylated to GalNAc-MM1 on Thr 27, which corresponds to Thr 420 in Gc-MAF. The peptide and its monoglycosylated form showed a high content of α-helical secondary structure and thermodynamic stability consistent with proteins of that size, suggesting that the insertion of the hexogeneous loop from Gc-MAF was well tolerated. However, we noticed that GalNAc-MM1 was destabilized by about 1 kcal/mol compared to MM1 in chemical denaturation experiments. It is possible of that the *O*-glycosylation of threonine could play a role in modulating the local secondary structure of MM1, and by analogy of Gc-MAF.

Given the therapeutic promise of MM1 [14], we decided to further investigate the energetic contribution of *O*-glycosylation. Hereto we prepared the diglycosylated form, $GalNAc_2$-MM1, in which in addition to Thr 27 in the first loop the second loop is glycosylated at Thr 52, also corresponding to Thr 420 in Gc-MAF. In this paper, we compare the thermodynamic properties of the three peptides, showing an unusual destabilizing effect correlated with the number of glycans. We used molecular dynamics simulations to explore the conformational dynamics of the miniaturized protein, to investigate the possible role of N-capping, and to relate the differences in motion to the differences in stability of the glycosylated and unglycosylated constructs in a qualitative manner. Our simulations indicate an increased conformational drift throughout the sequence, and a small increase in the exposed hydrophobic surface for the glycosylated form. These results suggest that the destabilization of the protein upon glycosylation may be caused by a disruption in the packing of the protein and an increase of the solvent exposed hydrophobic surface.

2. Materials and methods

2.1. Protein design, synthesis and purification

The model peptides were designed using Insight software (Biosym Technologies, Inc.) as described earlier [14]. Briefly, the glycosylated loop of VDBP-MAF [16,17] was spliced onto the structure of the scaffold 3-helix bundle α3W [15,18]. The PDB accession code is 1LQ7, and the full sequence is: GSRVKALEEKVKALEEKVKALGGGGRIEELKKK-WEELKKKIEELGGGGEVKKVEEEVKKLEEEIKKL.The process was repeated for the second loop and the model was optimized by energy minimization routines using the Discover module (cvff force field). The final sequence of MM1 is: GSRVKALEEKVKALEEKVKALGNATPTELAKLKKK-WEELKKKIEELGNATPTEVKKVEEEVKKLEEEIKKL. In this work, three alternative versions of the peptide were investigated: the aglycosylated control peptide MM1, the monoglycosylated (GalNAc-MM1), and diglycosylated ($GalNAc_2$-MM1) counterparts, and where either the N-terminal loop or both loops were *O*-glycosylated. All peptides were synthesized by standard solid-phase methods using α-GalNAc-threonine as a building block at the desired site. The peptides were manually cleaved from the resin, precipitated with ethyl ether, and purified by reverse HPLC. The protective acetyl groups in GalNAc-MM1 and $GalNAc_2$-MM1 were removed by treatment with sodium methoxide in methanol. After solvent evaporation the deprotected peptide was desalted and further purified by size exclusion chromatography. The molecular mass of pure peptides was confirmed via matrix-assisted laser desorption mass spectrometry (MALDI-TOF).

2.2. Chemical denaturation studies

Guanidine and urea denaturation profiles were obtained by monitoring the change of ellipticity at 222 nm (θ_{222}) as a function of denaturant concentration on JASCO 710 spectropolarimeter equipped with PTC 424S Peltier temperature controller and ATS-429S Automatic titration system as described previously [14]. A 2 μM solution of peptide in 50 mM potassium phosphate, pH = 7.0 was titrated with a second solution identical to the first but containing either 7.5 M GdnHCl or 9 M urea. Observed ellipticity values (θ_{obs}) were fit using the model previously described [19].

2.3. Parameterization of N-Acetylgalactosamine

Since parameters for *N*-acetylgalactosamine (Fig. 1) were not available in either the CHARMM 22 force field for proteins [20] or the CSFF force field for carbohydrates [21] it was necessary to develop parameters for this sugar. Many of the

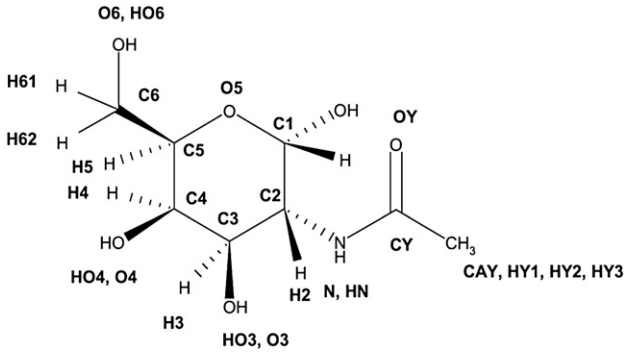

Fig. 1. The chemical structure of GalNAc with the CHARMM atomic names in bold.

parameters were copied from chemically similar entries in the CHARMM or CSFF force fields; others were developed using the method described by MacKerell [22]. In assigning the charges the groups of atoms containing the β carbon of Ser and Thr and their associated hydrogen atoms were kept neutral and the charge of hydrogen atoms bonded to aliphatic carbons was set to 0.09 atomic units. Parameters for the C3-C2-N-CY and CA-CB-O1-C1 dihedral angles were also obtained as described [22], using *N*-acetylgalactosamine and its ethyl acetal as model compounds. For the dihedral angles, a relaxed torsional scan was carried out at the HF/6-31g(d) level, followed by a single point energy calculation at the MP2/6-31g(d) level. All quantum mechanical calculations were carried out with Gaussian 03 [23]. The optimized parameters that are not found in either the CSFF set or the protein set are shown in Tables 1–4.

Table 1
Atom types and partial charges

Group	Atom	Type	Charge[a]	Origin of value
1	GalNAc amide N	NH1	−0.47	CHARMM peptide bond
1	GalNAc amide H	H	0.31	CHARMM peptide bond
1	GalNAc C_2	CTS	0.07	CHARMM peptide bond
1	GalNAc H_2	HAS	0.09	CHARMM peptide bond
2	GalNAc carbonyl C	C	0.51	CHARMM ACE patch
2	GalNAc carbonyl O	O	−0.51	CHARMM ACE patch
3	GalNAc acetyl C	CT3	−0.27	CHARMM ACE patch
3	GalNAc acetyl H	HA	0.09	CHARMM ACE patch
4	Acetal O to Ser or Thr residue	OES	−0.40	CSFF GL14 patch
4	GalNAc C_1	CTS	0.30	CSFF GL14 patch
4	GalNAc H_1	HAS	0.10	CSFF GL14 patch
4	GalNAc C_5	CTS	0.10	CSFF GL14 patch
4	GalNAc H_5	HAS	0.10	CSFF GL14 patch
4	GalNAc O_5 (sugar ring oxygen)	OES	−0.40	CSFF GL14 patch
4	Thr C^{β}	CT1	0.11	Charge neutrality
4	Thr H^{β}	HA	0.09	CHARMM THR residue
4	Ser C^{β}	CT2	0.02	Charge neutrality
4	Ser H^{β}	HA	0.09	CHARMM SER residue

[a]Atomic units.

Table 2
Bonding parameters

Atom types	Force constant (kcal/(mol A^2))	Length (Å)	Origin of values
NH1-CTS	320.0	1.43	CHARMM NH1-CT1
OES-CT1 OES-CT2 OES-CT3	428.0	1.42	CHARMM OH1-CT1

Test simulations of an α-GalNAc-Ser peptide with the optimized parameters in explicit water showed general agreement with the available structural experimental data on GalNAc [24,25]. The intermolecular distances from the simulations were consistent with the experimental NOE data [24]. In addition, the normal mode vibrational frequencies of GalNAc were calculated and showed general agreement with the experimental IR spectrum [25]; however, the degree of agreement was limited by the lack of an assignment for the experimental IR spectrum. There were additional vibrational modes that did not correspond to peaks in the experimental spectrum. These were similar to the additional modes for glucose calculated from the CSFF force field by Brady et al. [21].

2.4. Molecular dynamics simulation

Three molecular dynamics simulations of the Gc-MAF analog were undertaken, two with and one without attached sugars, but otherwise identical. Each simulation used the CHARMM 22 force field [20] with CMAP enhancements [26], the Brady force field [21] and the additional optimized parameters as described above. SHAKE constraints [27] were used on all bonds containing hydrogen atoms, and a 2 fs time

Table 3
Angle parameters

Atom types	Force constant (kcal/(mol rad^2))	Equilibrium value (degrees)	Origin of values
H-NH1-CTS	35.0	117.0	CHARMM H-NH1-CT1
CTS-NH1-C	50.0	120.0	CHARMM CT1-NH1-C
HAS-CTS-NH1	51.5	109.5	CHARMM HA-CT1-NH1
CTS-CTS-NH1	70.0	113.5	CHARMM
CBS-CTS-NH1			CT1-CT1-NH1
CT1-OES-CTS CT1-OES-CBS CT2-OES-CTS CT2-OES-CBS CT3-OES-CTS CT3-OES-CBS	92.6	111.5	CSFF CTS-OES-CTS
OES-CT1-CT1 OES-CT1-CT3 OES-CT2-CT1 OES-CT2-CT2 OES-CT2-CT3	75.7	110.1	CHARMM OH1-CT1-CT1
OES-CT1-HA OES-CT2-HA OES-CT3-HA	45.9	108.9	CHARMM OH1-CT1-HA

Table 4
Dihedral parameters

Atom types	Mult. 1 amplitude (kcal/mol)	Mult. 2 amplitude (kcal/mol)	Mult. 3 amplitude (kcal/mol)	Origin of values
CTS-CTS-NH1-C CBS-CTS-NH1-C	2.2	−1.0	0	Fitted to quantum calculations for GalNAc as described
H-NH1-CTS-CTS H-NH1-CTS-CBS	0.0	0.0	0.0	Cleared to avoid double-counting rotation around N–C2 bond
NH1-CTS-CTS-HAS NH1-CTS-CBS-HAS	0.0	0.0	0.0	Cleared because only small 3rd order dihedral in CSFF
NH1-CTS-CTS-OHS NH1-CTS-CBS-OHS	−4.9	0.3	0.5	CSFF OHS-CTS-CTS-OHS
OES-CTS-CTS-NH1 OES-CBS-CTS-NH1	−3.8	0.6	0.4	CSFF OES-CTS-CTS-OHS
CTS-CTS-CTS-NH1	−1.9	0.3	0.0	CSFF CTS-CTS-CTS-OHS
HAS-CTS-NH1-C	0	0	0	CHARMM HB-CT1-NH1-C
HAS-CTS-NH1-H	0	0	0	CHARMM HB-CT1-NH1-H
CT3-C-NH1-CTS	1.6	2.5	0	CHARMM CT3-C-NH1-CT1
O-C-NH1-CTS	0	2.5	0	CHARMM O-C-NH1-CT1
CT3-CT2-OES-CTS CT3-CT2-OES-CBS CT3-CT1-OES-CTS CT3-CT1-OES-CBS CT1-CT2-OES-CTS CT1-CT2-OES-CBS CT1-CT1-OES-CTS CT1-CT1-OES-CBS	−0.3	−0.3	0.4	Fitted to quantum calculations for the ethyl acetal of GalNAc as described
CT1-OES-CTS-CTS CT1-OES-CBS-CTS CT2-OES-CTS-CTS CT2-OES-CBS-CTS CT3-OES-CTS-CTS CT3-OES-CBS-CTS	−0.8	−0.3	0.4	CSFF CTS-OES-CTS-CTS
OES-CTS-OES-CT1 OES-CBS-OES-CT1 OES-CTS-OES-CT2 OES-CBS-OES-CT2 OES-CTS-OES-CT3 OES-CBS-OES-CT3	0.2	1.0	0.9	CSFF OES-CTS-OES-CTS
CT1-OES-CTS-HAS CT1-OES-CBS-HAS CT2-OES-CTS-HAS CT2-OES-CBS-HAS CT3-OES-CTS-HAS CT3-OES-CBS-HAS	0	0	0	Cleared because only small 3rd order dihedral in CSFF
HA-CT1-OES-CTS HA-CT1-OES-CBS HA-CT2-OES-CTS HA-CT2-OES-CBS HA-CT3-OES-CTS HA-CT3-OES-CBS	0	0	0	Cleared because only small 3rd order dihedral in CHARMM force field

step was employed. The simulations were done in explicit solvent [28] with periodic boundary conditions, and the particle mesh Ewald method [29] for the long range electrostatics; a switching function between 8 Å and 12 Å was used for the Van Der Waals interactions. All simulations were done with the CHARMM program [30]. The conformation of the simulated protein was recorded each 0.5 ps. Proteins were first minimized and solvated with explicit water in a 49×49×65 Å box, with two chloride ions to achieve charge neutrality. The system was minimized and the water was equilibrated by heating to 600 K and cooling to 300 K while holding the protein fixed. The system was then minimized again and reheated to 300 K with harmonic constraints on the protein backbone with a force constant of 1.0 kcal/(mol Å). Once the system reached 300 K, these harmonic constraints were gradually removed in 5 steps over 50 ps. Each simulation was continued at constant pressure and temperature using the Nosé–Hoover algorithm [31,32]. The last 20 ns was used for analysis. One of the simulations was used to examine the role of the Thr 50 side chain in the anti conformation, which was not sampled in the other simulation

α_3W (37-52) LKKKIEEL GGGG EVKK

VDBP(loop) AKL PDAT PTELAK

MM1 (39-56) LKKKIEEL GNAT PTEVKK

Fig. 2. Model of GalNac$_2$-MM1, built by grafting the glycosylated loop of Gc-MAF (orange) onto the structure of α3W (grey); the GalNAc moieties attached to Thr 27 and Thr 52 are represented in sticks. Figure generated with Pymol [45]. The lower panel shows the sequence alignment of the glycosylated loop of Gc-MAF with the second loop of α3W and of MM1, with the residues grafted from Gc-MAF in orange. The site of glycosylation, a threonine, is underlined. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of the glycosylated protein. In this simulation the χ angle of Thr 50 was held near −175° by a harmonic constraint on the dihedral angle. After 10 ps of equilibration, the restraint was removed and the unrestrained system was simulated for 20 ns during which the Thr 50 side chain remained in the anti conformation.

The convergence of the simulations was verified by the calculation of the RMSD with the initial state, and the radius of gyration along the trajectories. In addition, quasi-harmonic analyses at various time intervals, followed by projections of the trajectories onto several of the lowest modes were performed as a check of the sampling. The secondary structure was calculated using STRIDE [33] and the Thr 50 side chain conformations were classified according to their χ angle as either "anti" (for χ angles greater than 120° or less than −120°), "gauche+" (for χ angles between 0 and 120°) or "gauche−" (for χ angles between −120 and 0°). For the hydrogen bonding analysis, a distance cutoff of 2.4 Å and an angle cutoff of 150° was used.

2.5. Umbrella sampling

The relative free energies of the Thr 50 rotamers and the free energy barriers for interconversion were obtained from an umbrella sampling simulation [34] of the Thr 50 χ angle. These simulations were performed for the glycosylated and unglycosylated protein under conditions identical to those described above. The umbrella biasing potential was expanded in a 12-order trigonometric polynomial, and the bin size of the histograms was 1°. The umbrella potential was readjusted every 200 ps. A total of 60 readjustments of the biasing potential were made (12 ns total per run). The weighted histogram analysis method (WHAM) [35] was used to combine the data from all umbrella simulations for the determination of the potential of mean force (PMF) as a function of Thr 50 χ angle. The relative free energies of the minima corresponding to anti, gauche+ or gauche− conformations were used to calculate the relative populations of the conformers.

3. Results

3.1. Protein design

MM1is a Gc-MAF analog glycopeptide derived from α3W, a de novo designed three-helix bundle, by swapping the original glycine-rich loops with the functional loop of the native protein, which is glycosylated at Thr 420 with a GalNAc moiety (Fig. 2). Initially, we synthesized and characterized only the monoglycosylated adduct, GalNAc-MM1. GalNAc-MM1 showed an activity comparable with that of Gc-MAF in enhancing Fc-receptor mediated phagocytosis in macrophages in vitro [14]. The highly α-helical secondary structure of α3W was preserved in MM1 and GalNac-MM1, although the exchange of about 20% of the sequence resulted in a small destabilization to chemical denaturation of MM1 relative to α3W. Here, we explore the effect of glycosylation on the thermodynamic stability of the MM1 scaffold by examining both the mono- and the diglycosylated forms.

3.2. Loss of thermodynamic stability in mono- and diglycosylated MM1

The circular dichroism spectra of MM1, GalNAc-MM1 [14], and GalNAc$_2$-MM1 are typical of α-helical proteins, showing characteristic minima at 208 and 222 nm, and are indistinguishable from each other (data not shown), suggesting that the secondary

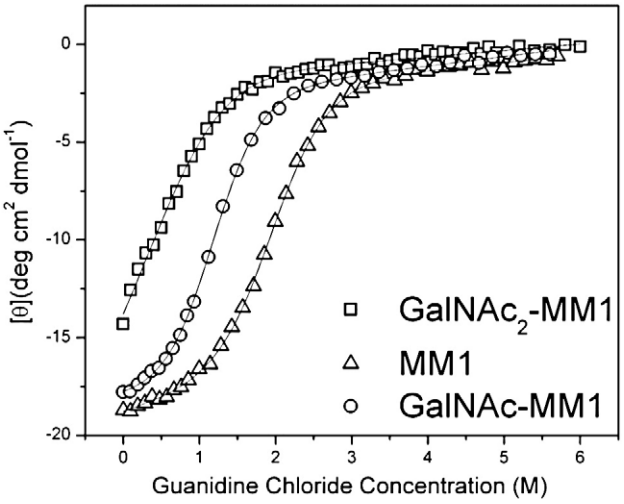

Fig. 3. Guanidinium hydrochloride denaturation curves of MM1 (triangles), GalNAc-MM1 (circles) and GalNAc$_2$-MM1 (squares) monitored by molar ellipticity at 222 nm as a function of GdnCl concentration. The solid curves represent the best fit to the data.

structure is globally well preserved after insertion of the GalNAc residues. The peptides are resistant to thermal denaturation in the range between 20 °C and 96 °C, showing a slope in the temperature dependency of the folded baseline but not a clear unfolding transition. The thermodynamic stability of the series was investigated by chemical denaturation. Guanidinium denaturation profiles, obtained by monitoring the change of ellipticity at 222 nm (θ_{222}) as a function of denaturant concentration (Fig. 3), show that all the model proteins undergo reversible two-state denaturation transitions with unique midpoints across the series. The thermodynamic parameters extracted from the curves reveal a clear trend in free energy of folding, ΔG_{H2O}, as a function of glycosylation. The most stable protein of the series is MM1, which has a ΔG_{H2O} of -3.23 ± 0.1 kcal mol^{-1} in good agreement with the -4.6 to -6.2 kcal mol^{-1} range of ΔG_{H2O} values for de novo designed three-helix bundles in the α3 family [15,18,36] which includes the original scaffold protein. The two glycosylated forms are less stable, with each glycosylation site destabilizing MM1 by about one kcal mol^{-1}: the ΔG_{H2O} of GalNAc-MM1 is -2.16 ± 0.12 kcal mol^{-1}, whereas the ΔG_{H2O} of GalNAc$_2$-MM1 is -0.84 ± 0.19 kcal mol^{-1}.

These results were confirmed by urea denaturation profiles (Fig. 4), in which the milder denaturant allows for a more accurate determination of the folded baseline. To investigate whether the hydrophilic character of the hydroxyl moieties contributed to the destabilizing effect of GalNAc substitution [5], we assessed the free energy of folding of Ac$_3$GalNAc-MM1, the protected version of GalNAc-MM1. The urea denaturation profiles of GalNAc-MM1 and Ac$_3$GalNAc-MM1 overlap, indicating that the effect observed in the GalNAc adducts is not due to a specific role of the hydroxy moiety. The results from the chemical denaturation studies are summarized in Table 5.

3.3. Molecular dynamics studies

We performed molecular dynamics simulations of GalNAc$_2$-MM1 and of the non-glycosylated form, MM1, in order to compare their conformational dynamics and, in particular, to explore the possible effect of glycosylation on N-capping of the helix and on the burial of the hydrophobic core. An analysis of C^{α} atoms RMSD variations, reported in Fig. 5, reveals a more pronounced conformational drift in the glycosylated protein, especially in the loops near the glycosylation sites. To provide a more detailed insight into the flexibility of the proteins, we examined several factors individually.

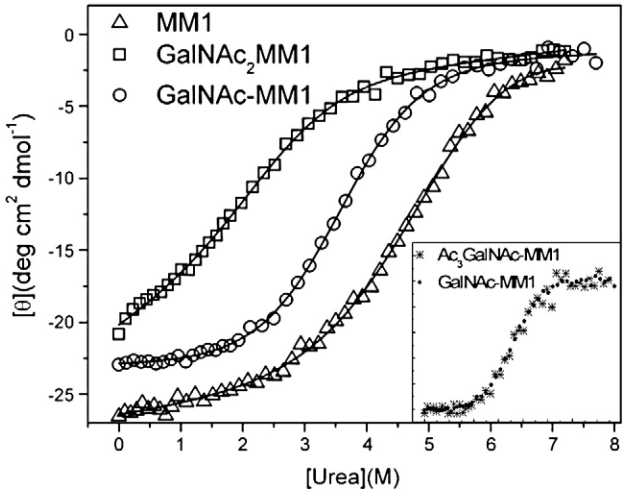

Fig. 4. Urea denaturation curves of MM1 (triangles), GalNAc-MM1 (circles) and GalNAc$_2$-MM1 (squares) monitored by molar ellipticity at 222 nm as a function of urea concentration. The solid curves represent the best fit to the data.

Table 5
Summary of chemical denaturation results

Protein	Guanidine chloride		Urea	
	m	Δ*G* (kcal/mol)	*m*	Δ*G* (kcal/mol)
MM1	−1.64±0.05	−3.23±0.11	−0.68±0.05	−3.47±0.17
GalNAc-MM1	−1.85±0.06	−2.16±0.12	0.90±0.04	−3.23±0.16
Ac$_3$GalNac-MM1	NA	NA	−0.91±0.12	−3.15±0.45
GalNAc$_2$-MM1	−1.60±0.13	−0.70±0.49	−0.79±0.08	−1.92±0.35

A possible explanation for the loss of stability is that the glycosylation of Thr 52 at the start of the third alpha helix might interfere with N-capping of the third helix by Thr 50, which is the last residue in the loop with non-helical dihedral angles [37]. The size of the $\Delta\Delta G$ corresponding to each glycosylation event, about 1 kcal/mol, is consistent with the effect expected for each N-cap. Molecular modeling showed that Thr 50, which precedes Pro 51 in the TPT sequence, could act as N-capping of the third helix through hydrogen bonding to the backbone Glu 53. This model raised the possibility that the glycosylation of the second threonine, Thr 52, could interfere with the N-capping process.

The N-capping hypothesis was ruled out by our simulations, however. We analyzed our trajectories for the effects of glycosylation on the conformation of the Thr 50 side chain, and on its hydrogen bonding patterns. Table 6 shows that N-capping only takes place when Thr 50 is in the gauche+ conformation, and that it occurs more frequently when the system is glycosylated than when it is not glycosylated. In the glycosylated gauche+ state, N-capping occurs 14.5% of the time, while in the unglycosylated gauche+ state N-capping occurs 5.8% of the time. In both cases most Thr 50 hydrogen bonds are formed with the carboxylate group of Glu 53 instead. Moreover, our potential of mean force calculations showed that the gauche+ conformation is strongly favored in the glycosylated state (Table 6 and Fig. 6). The equilibrium fraction of the gauche+ state is 56.3% for the glycosylated protein and 4.9% for the unglycosylated protein (in the unglycosylated state the gauche− conformation is favored). This means that contrary to the hypothesis, N-capping is more prevalent for the glycosylated state. Despite these differences in N-capping upon glycosylation, N-capping happens at most 15% of the time. Consequently, N-capping of helix 3 is not deemed to be important for the stabilization of the protein. A possible direct involvement of Thr 52 appears to be unlikely: Thr 52 rarely forms hydrogen bonds in the unglycosylated system, suggesting that the direct blockage of hydrogen bonds by attachment of the sugar is of no consequence. These findings are in agreement

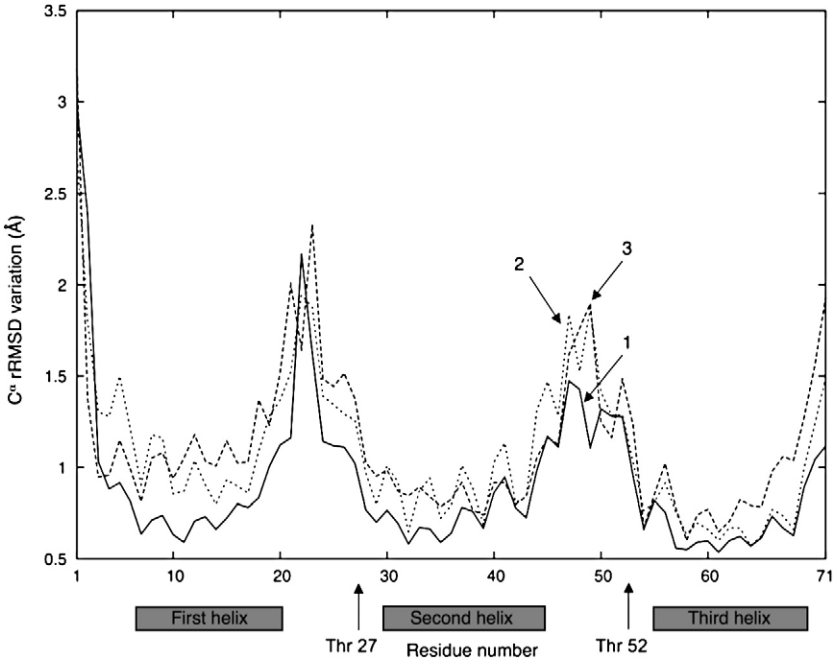

Fig. 5. Alpha carbon RMSD fluctuation as a function of residue number. Arrows indicate glycosylation sites. (1) MM1 (black), (2) $GalNAc_2$-MM1 with Thr 50 side chain in gauche+ or gauche− conformation (dashed), (3) $GalNAc_2$-MM1 with Thr 50 side chain in anti conformation (dotted).

with the experimental results on the destabilization of the $Ac_3GalNAc$-MM1 construct.

We also correlated the secondary structure of the protein with the Thr 50 side chain conformation in the glycosylated and unglycosylated states. These percentages are shown in Table 7. Compared to MM1, $GalNAc_2$-MM1 exhibits a slightly larger content of turn secondary structure and a slightly smaller amount of 3_{10} helical secondary structure. In addition, depending on the Thr 50 conformation, there is a change in the types of turns that are favored, from type I turns to type I′, II, and IV turns. These changes are consistent with other results reported in the literature. For example, it was previously observed in an NMR study that glycosylation of a 9-amino acid analog of the epitope of MUC1 shifts its equilibrium away from type I turns and toward type II turns [38]. There is very little change in the amount of α helical structure present. In contrast, in calcitonin glycosylation near helix ends (CT6 or CT21) was observed to cause a loss of helical structure [39]. Clearly, the effect of

Table 6
Hydrogen bonding of the Thr 50 side chain

			Glycosylated				Unglycosylated			
			Anti [a] (%)	Gauche+ [b] (%)	Gauche− [b] (%)	Average (%)	Anti [c] (%)	Gauche+ [c] (%)	Gauche− [c] (%)	Average (%)
HN	GLY	47	0.0	0.1	0.0	0.0				
O	GLY	47	0.0	0.0	0.0	0.0				
O	ASN	48	0.0	0.0	0.1	0.0				
HN	ASN	48	0.0	0.0	0.0	0.0				
HN	THR	52	5.4	0.1	0.0	2.3	0.1	0.1	0.0	0.1
HN	GLU	53	0.9	14.5	0.0	8.5	0.0	5.8	0.0	0.3
OG1	THR	52	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HG1	THR	52	n/a [d]	n/a [d]	n/a [d]	n/a [d]	0.0	5.8	0.0	0.3
OE	GLU	53	0.0	84.6	10.2	47.8	0.0	93.9	30.7	22.1
HO6	NGA	72	0.0	0.1	0.0	0.1				
O6	NGA	72	0.4	0.9	0.0	0.6				
O5	NGA	72	0.0	0.0	0.0	0.0				
OY	NGA	72	0.0	0.0	0.0	0.0				
HN	NGA	72	2.3	0.0	0.0	1.0				
Equilibrium fraction			42.1	56.3	1.6		38.0	4.9	57.1	

[a] From the glycosylated state simulation that started in the Thr 50 anti conformation. No gauche+ or gauche− conformations were observed in this simulation.
[b] From the glycosylated state simulation that started in the Thr 50 gauche+ conformation. No anti conformations were observed in this simulation.
[c] From the unglycosylated state simulation that started in the Thr 50 gauche+ conformation.
[d] In the glycosylated protein, Thr 52 H^{γ} is removed and replaced with the sugar.

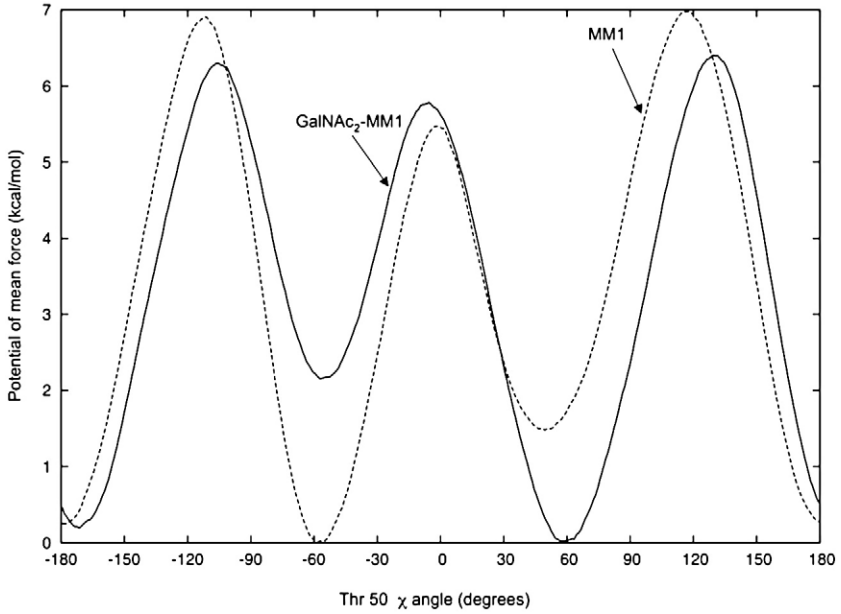

Fig. 6. Potential of mean force as a function of the Thr 50 χ angle for MM1 and GalNAc$_2$-MM1.

glycosylation on the stability of helical proteins is very much dependent on the specific position of the glycosylated residue.

Histograms of the solvent-accessible surface area are shown in Fig. 7. There are long tails on the histograms of hydrophobic surface area of GalNAc$_2$-MM1, especially for the most populated gauche+ and anti conformations. This shows that glycosylation causes the three-helix bundle to adopt conformations in which the hydrophobic residues are more solvent exposed. Analysis showed that the increased solvent exposure affected almost all hydrophobic residues: those near the loops connecting the helices as well as those in the core of the protein. Moreover, the increased exposure correlated with a small increase in the radius of gyration of the protein. In addition, there are between 1 and 8 more hydrophobic contacts among atoms in hydrophobic residues in unglycosylated MM1 compared to GalNAc$_2$-MM1 (data not shown).

The findings indicate two possible reasons for the destabilization of the protein upon glycosylation. The first is a solvation penalty due to the increased exposure of hydrophobic residues [40] upon glycosylation. We expect this solvation penalty to be small, since the increases in solvent exposed areas are small and only concern a fraction of the configurations (Fig. 7). Second, the small increase in the radius of gyration and the decrease in contacts between hydrophobic residues in the glycosylated state indicate a

Table 7
Secondary structure analysis

	Glycosylated				Unglycosylated			
	Anti [a]	Gauche+ [b]	Gauche− [b]	Average	Anti [c]	Gauche+ [c]	Gauche− [c]	Average
Thr 50 conformation								
Total time observed (ns)	20.00	18.97	1.03		14.54	2.33	3.12	
Sec. structure type								
Total turns	9.9%	7.0%	8.1%	8.2%	5.0%	9.3%	7.2%	6.5%
α Helix	78.6%	79.1%	79.2%	78.9%	80.2%	78.5%	78.5%	79.2%
3_{10} Helix	0.1%	0.2%	0.0%	0.2%	2.0%	0.3%	1.4%	1.6%
Random coil	11.2%	13.6%	12.4%	12.6%	12.4%	11.0%	12.2%	12.2%
Others	0.2%	0.1%	0.3%	0.1%	0.4%	0.9%	0.7%	0.6%
Type I turns	0.4%	0.5%	0.1%	0.5%	3.1%	5.4%	4.9%	4.2%
Type I′ turns	4.4%	0.0%	0.0%	1.9%	0.0%	0.0%	0.0%	0.0%
Type II turns	1.4%	1.4%	0.2%	1.4%	0.0%	0.0%	0.0%	0.0%
Type IV turns	3.7%	5.1%	7.8%	4.5%	1.8%	3.8%	2.3%	2.2%
Type VIII turns	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%
Minimum free energy (kcal/mol) [d]	0.2	0.0	2.2		0.2	1.5	0.0	
Equilibrium fraction	42.1%	56.3%	1.6%		38.0%	4.9%	57.1%	

[a] From the glycosylated state simulation that started in the Thr 50 anti conformation. No gauche+ or gauche− conformations were observed in this simulation.
[b] From the glycosylated state simulation that started in the Thr 50 gauche+ conformation. No anti conformations were observed in this simulation.
[c] From the unglycosylated state simulation that started in the Thr 50 gauche+ conformation.
[d] From the umbrella sampling simulations (Fig. 6).

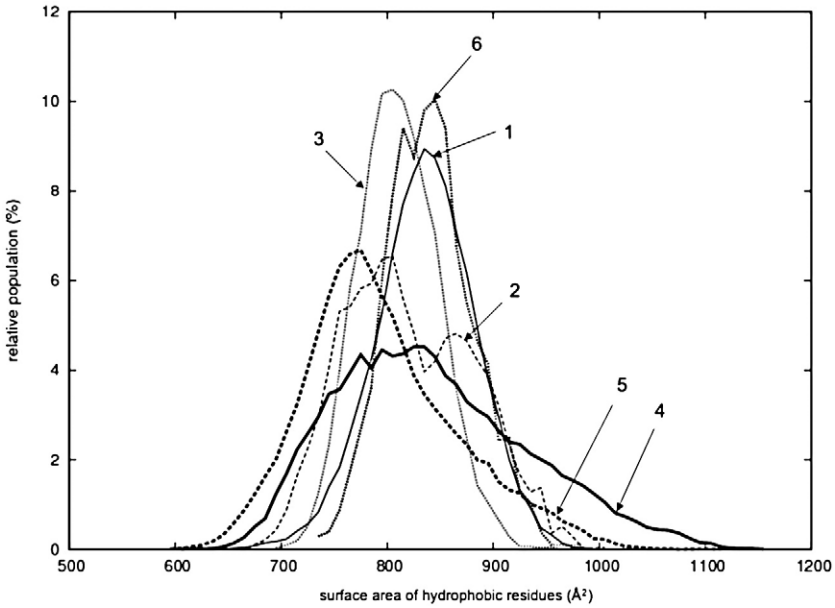

Fig. 7. Histogram of solvent-accessible surface area of MM1, by level of glycosylation and Thr 50 side chain conformation. (1) MM1 with Thr 50 anti (black, thin, 14.54 ns total sampling time); (2) MM1 with Thr 50 gauche+ (dashed, thin, 2.33 ns); (3) MM1 with Thr 50 gauche− (dotted, thin, 3.12 ns); (4) $GalNAc_2$-MM1 with Thr 50 anti (black, thick, 20 ns); (5) $GalNAc_2$-MM1 with Thr 50 gauche+ (dashed, thick, 18.97 ns); (6) $GalNAc_2$-MM1 with Thr 50 gauche− (dotted, thick, 1.03 ns).

disruption in the packing of the hydrophobic core of the glycosylated protein. Consistent with the experimental findings, the simulations indicate that the combined energetic effects are small. The small magnitude of the destabilization is also consistent with the relatively small increase in protein fluctuations observed for the glycosylated state (Fig. 5).

Combining the data on the exposed hydrophobic surface (Fig. 7) with the data on the N-capping and the data from the umbrella sampling simulations (Table 6 and Fig. 6), the following model for the destabilization of the glycosylated state emerges. Glycosylation of Thr 52 favors a different rotamer population of Thr 50. This shift triggers a change in the stability of the system: in the new equilibrium, hydrophobic residues become more solvent exposed, and the packing of the hydrophobic core is slightly disrupted. Both of these effects may destabilize the glycosylated state.

4. Discussion

Model peptide-based systems have been used extensively to evaluate the effect of glycosylation on the biophysical properties of proteins, as their use allows for precise control of the site and type of glycosylation [4,5,38,41]. The results of these studies have been mixed: glycosylation can have opposite effects, for example stabilize [4] or destabilize [5] a simple helix-turn-helix motif that self assemble into four helix bundle, depending on the type of attachment to the host amino acid. Extensive studies on calcitonin, which is not glycosylated in nature, show that threonine glycosylation perturbs the local structure, resulting in a destabilizing effect observed at the helical termini, and almost no effect at mid-helix positions [39,42]. The glycosylation of residues at or near turn structure can affect the conformational equilibrium, favoring an extended conformation or a type II β-turn. An effect on local structure could be important in modulating the interaction with receptors for MUC1 [38,43,44].

In Gc-MAF, changes in glycosylation status are correlated with the gain of macrophage-activating activity when the core GalNAc is exposed, and with its loss as GalNAc is cleaved by GalNAc-ases. We demonstrated that function is conserved in a miniaturized protein displaying solely the glycosylated loop of Gc-MAF, and to a much lesser extent is preserved in an unstructured peptide derived from the loop [14]. We now show a correlation between local structure and glycosylation, hinting at a possible biological role for structure modulation. In MM1, glycosylation favors a different rotamer population of Thr50. This local structure change then induces a slight increase in solvent exposure of hydrophobic residues. The increase in exposed hydrophobic surface is consistent with our experimental data on the loss of about 1 kcal/mol per glycosylation event in the folded–unfolded transition. Moreover, the lack of a cooperative temperature dependent melting transition is consistent with a molten globule character of the glycosylated protein model, in which high secondary structure content is accompanied by high conformational dynamics. Given the sizable thermodynamic effect of loop glycosylation, future model protein should utilize well folded scaffolds that would be tolerant to destabilization.

Acknowledgements

We thank Alex MacKerell for guidance with the parameterization process, and Hiqmet Kamberaj for technical assistance. Computer time was provided by the Fulton High Performance Computing Initiative at Arizona State University.

References

[1] M.R. Pratt, C.R. Bertozzi, Synthetic glycopeptides and glycoproteins as tools for biology, Chem. Soc. Rev. 34 (2005) 58–68.

[2] A. Varki, R. Cummings, J. JEsko, H. Freeze, G. Hart, J. Marth, Essentials of Glycobiology, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999.

[3] R.J. Sola, J.A. Rodriguez-Martinez, K. Griebenow, Modulation of protein biophysical properties by chemical glycosylation: biochemical insights and biomedical implications, Cell Mol. Life Sci. (2007).

[4] L.K. Andersson, G.T. Dolphin, J. Kihlberg, L. Baltzer, The effect of glycosylation on the structure of designed four-helix bundle motifs, J. Chem. Soc., Perkin Trans. 2 (2000) 459–464.

[5] S. Vijayalekshmi, S.K. George, L.K. Andersson, J. Kihlberg, L. Baltzer, A surface exposed O-linked galactose residue destabilises the structure of a folded helix-loop-helix dimer, Org. Biomol. Chem. 1 (2003) 2455–2460.

[6] N.E. Fahmi, L. Dedkova, B. Wang, S. Golovine, S.M. Hecht, Site-specific incorporation of glycosylated serine and tyrosine derivatives into proteins, J. Am. Chem. Soc. 129 (2007) 3586–3597.

[7] G.W. Hart, K.D. Greis, L.Y. Dong, M.A. Blomberg, T.Y. Chou, M.S. Jiang, E.P. Roquemore, D.M. Snow, L.K. Kreppel, R.N. Cole, et al., O-linked N-acetylglucosamine: the "yin-yang" of Ser/Thr phosphorylation? Nuclear and cytoplasmic glycosylation, Adv. Exp. Med. Biol. 376 (1995) 115–123.

[8] F.I. Comer, G.W. Hart, O-Glycosylation of nuclear and cytosolic proteins. Dynamic interplay between O-GlcNAc and O-phosphate, J. Biol. Chem. 275 (2000) 29179–29182.

[9] N. Yamamoto, R. Kumashiro, Conversion of vitamin D3 binding protein (group-specific component) to a macrophage activating factor by the stepwise action of beta-galactosidase of B cells and sialidase of T cells, J. Immunol. 151 (1993) 2794–2802.

[10] P.T. Gomme, J. Bertolini, Therapeutic potential of vitamin D-binding protein, Trends Biotechnol. 22 (2004) 340–345.

[11] N. Yamamoto, V.R. Naraparaju, S.O. Asbell, Deglycosylation of serum vitamin D3-binding protein leads to immunosuppression in cancer patients, Cancer Res. 56 (1996) 2827–2831.

[12] O. Kisker, S. Onizuka, C.M. Becker, M. Fannon, E. Flynn, R. D'Amato, B. Zetter, J. Folkman, R. Ray, N. Swamy, S. Pirie-Shepherd, Vitamin D binding protein-macrophage activating factor (DBP-maf) inhibits angiogenesis and tumor growth in mice, Neoplasia 5 (2003) 32–40.

[13] N. Yamamoto, V.R. Naraparaju, Immunotherapy of BALB/c mice bearing Ehrlich ascites tumor with vitamin D-binding protein-derived macrophage activating factor, Cancer Res. 57 (1997) 2187–2192.

[14] F. Bogani, E. McConnell, L. Joshi, Y. Chang, G. Ghirlanda, A designed glycoprotein analogue of Gc-MAF exhibits native-like phagocytic activity, J. Am. Chem. Soc. 128 (2006) 7142–7143.

[15] C. Tommos, J.J. Skalicky, D.L. Pilloud, A.J. Wand, P.L. Dutton, De novo proteins as models of radical enzymes, Biochemistry 38 (1999) 9495–9507.

[16] C. Verboven, A. Rabijns, M. De Maeyer, H. Van Baelen, R. Bouillon, C. De Ranter, A structural basis for the unique binding features of the human vitamin D-binding protein, Nat. Struct. Biol. 9 (2002) 131–136.

[17] J.F. Head, N. Swamy, R. Ray, Crystal structure of the complex between actin and human vitamin D-binding protein at 2.5 A resolution, Biochemistry 41 (2002) 9015–9020.

[18] Q.H. Dai, C. Tommos, E.J. Fuentes, M.R. Blomberg, P.L. Dutton, A.J. Wand, Structure of a de novo designed protein model of radical enzymes, J. Am. Chem. Soc. 124 (2002) 10952–10953.

[19] M.M. Santoro, D.W. Bolen, Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants, Biochemistry 27 (1988) 8063–8068.

[20] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (1998) 3586–3616.

[21] M. Kuttel, J.W. Brady, K.J. Naidoo, Carbohydrate solution simulations: producing a force field with experimentally consistent primary alcohol rotational frequencies and populations, J. Comput. Chem. 23 (2002) 1236–1243.

[22] A.D. MacKerell, in: O.M. Becker, A.D. MacKerell, B. Roux, M. Watanabe (Eds.), Computational Biochemistry and Biophysics, Marcel Dekker, New York, 2001, pp. 7–35.

[23] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ezra, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian, Inc, 2003.

[24] F. Corzana, J.H. Busto, G. Jimenez-Oses, J.L. Asensio, J. Jimenez-Barbero, J.M. Peregrina, A. Avenoza, New insights into alpha-GalNAc-Ser motif: influence of hydrogen bonding versus solvent interactions on the preferred conformation, J. Am. Chem. Soc. 128 (2006) 14640–14648.

[25] SDBSWeb. National Institute of Advanced Industrial Science and Technology.

[26] A.D. Mackerell, M. Feig, C.L. Brooks, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, J. Comput. Chem. 25 (2004) 1400–1415.

[27] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical-integration of Cartesian equations of motion of a system with constraints – molecular-dynamics of N-Alkanes, J. Comput. Phys. 23 (1977) 327–341.

[28] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.

[29] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.

[30] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, Charmm – a program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. 4 (1983) 187–217.

[31] W.G. Hoover, Canonical dynamics: equilibrium phase-space distributions, Phys. Rev. A 31 (1985) 1695–1697.

[32] S. Nose, A molecular-dynamics method for simulations in the canonical ensemble, Mol. Phys. 52 (1984) 255–268.

[33] D. Frishman, P. Argos, Knowledge-based protein secondary structure assignment, Proteins Struct. Funct. Genet. 23 (1995) 566–579.

[34] C. Bartels, M. Karplus, Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations, J. Comput. Chem. 18 (1997) 1450–1462.

[35] A.M. Ferrenberg, R.H. Swendsen, Optimized Monte-Carlo data-analysis, Phys. Rev. Lett. 63 (1989) 1195–1198.

[36] S.T. Walsh, H. Cheng, J.W. Bryson, H. Roder, W.F. DeGrado, Solution structure and dynamics of a de novo designed three-helix bundle protein, Proc. Natl. Acad. Sci. U S A 96 (1999) 5486–5491.

[37] R. Aurora, G.D. Rose, Helix capping, Protein Sci. 7 (1998) 21–38.

[38] J. Schuman, A.P. Campbell, R.R. Koganty, B.M. Longenecker, Probing the conformational and dynamical effects of O-glycosylation within the immunodominant region of a MUC1 peptide tumor antigen, J. Pept. Res. 61 (2003) 91–108.

[39] M. Tagashira, H. Iijima, Y. Isogai, M. Hori, S. Takamatsu, Y. Fujibayashi, K. Yoshizawa-Kumagaye, S. Isaka, K. Nakajima, T. Yamamoto, T. Teshima, K. Toma, Site-dependent effect of O-glycosylation on the conformation and biological activity of calcitonin, Biochemistry 40 (2001) 11090–11095.

[40] A. Fersht, Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, W. H. Freeman, 1998.
[41] F.C. Liang, R.P. Chen, C.C. Lin, K.T. Huang, S.I. Chan, Tuning the conformation properties of a peptide by glycosylation and phosphorylation, Biochem. Biophys. Res. Commun. 342 (2006) 482–488.
[42] M. Tagashira, H. Iijima, K. Toma, An NMR study of O-glycosylation induced structural changes in the alpha-helix of calcitonin, Glycoconj. J. 19 (2002) 43–52.
[43] A.H. Andreotti, D. Kahne, Effects of glycosylation on peptide backbone conformation, J. Am. Chem. Soc. 115 (1993) 3352–3353.
[44] A. Kuhn, H. Kunz, Saccharide-induced peptide conformation in glycopeptides of the recognition region of LI-cadherin, Angew. Chem., Int. Ed. Engl. 46 (2007) 454–458.
[45] W.L. Delano, Delano Scientific, LLC, South San Francisco, CA, USA, 2005.