Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Structural Biology

Folding of glycoproteins: toward understanding the biophysics of the glycosylation code

Dalit Shental-Bechor and Yaakov Levy

Glycosylation is among the most common post-translational modifications that proteins undergo that may affect many of their activities. It may also modify the underlying energy landscape of glycoproteins in a way that their altered biophysical characteristics are linked to their bioactivity. Yet, the capability of glycosylation to modify thermodynamic and kinetic properties varies greatly between glycoproteins. Deciphering the 'glycosylation code' that dictates the interplay between the nature of the carbohydrates or the proteins and the biophysical properties of the glycosylated proteins is essential. In this article, we discuss how the size, number, and structure of glycans, as well as the attachment sites, may modulate the folding of glycoproteins. Understanding the cross-talks between the protein and the attached glycans at the molecular level may assist in tailoring the biophysical properties of proteins in general.

Address
Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel

Corresponding author: Levy, Yaakov (koby.levy@weizmann.ac.il)

Current Opinion in Structural Biology 2009, **19**:524–533

This review comes from a themed issue on
Carbohydrates and glycoconjugates
Edited by Nathan Sharon, John Gallagher

Available online 3rd August 2009

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2009.07.002

Introduction

Post-translational modifications (PTMs) are ubiquitous in the cell. They regulate the function of proteins and provide a mechanism for increasing the diversity of protein structures. In addition, such modifications may result in the localization of proteins to specific cellular organelles, with incorrect targeting being associated with a number of diseases. It is often believed that PTMs may regulate protein activity by changing the protein surface and by introducing new functional groups that act as a signaling tag for binding to other molecules. These modifications may also modulate the thermodynamic and kinetic features of proteins (e.g. stability, structural flexibility, and folding rate). PTMs, therefore, can enrich the repertoire of protein characteristics beyond that dictated by the 20 amino acids. Thus, they efficiently and economically increase the complexity and dimensions of the primary gene products. One may consider the function of a PTM to be related to the modified characteristics it induces in the conjugated protein.

Various PTMs have been shown to affect protein biophysics. For example, phosphorylation contributes negative charges to the protein surface and can induce conformational changes [1], while methylation and acetylation mask charges and play an important role in protein–DNA interactions. Glycosylation has been shown to enhance protein thermal stability and to facilitate folding kinetics [2•,3••,4]. Yet, quantification and generalization of these effects and the molecular components that dictate the capability of the attachments to modulate protein characteristics are currently lacking.

In this article, we focus on the effect of glycosylation on the biophysical properties of proteins, which are a consequence of the features of the underlying energy landscapes that govern their folding. Joint experimental and theoretical efforts [5,6] have led to the conclusion that a protein folds by navigating through an energy landscape that is globally funneled toward a structurally defined native state [7]. The funneled nature of energy landscapes is a product of the evolutionary selection of sequences that are minimally frustrated, in that they have fewer interactions that are in conflict with the native state and thus exhibit relatively high conformational preferences. Schematically, glycosylation may modulate a protein's characteristics by affecting differently the energetics (enthalpy or entropy) of the unfolded and/or folded state.

Glycosylation is a covalent attachment of one or more carbohydrates (also termed 'glycans') to a protein. It is a complex process that commonly involves several of nine different kinds of monosaccharides attached to any of eight types of amino acid residues and is assisted by many enzymes [8,9]. In most cases, however, the attachment is to an Asn residue (i.e. part of the sequence Asn-Xaa-Ser/Thr) in a process termed *N*-glycosylation. Among their repertoire of functional roles, glycans serve as recognition markers [10], mediate interactions with pathogens, modulate immune response, and regulate protein turnover [11]. The ability of sugar binding protein receptors (primarily lectins [12]) to recognize carbohydrate conjugates lies at the heart of many central biological processes. This recognition is remarkable given the flexibility of the

carbohydrates and their high structural variability. *N*-Glycosylation starts during protein synthesis and translocation into the endoplasmic reticulum (ER). The glycan is added to the still unfolded protein while it is in the translocon complex [11], indicating that it may assist in obtaining the correct fold following the recruitment of the lectins, calnexin and calreticulin [4,13,14]. The role of glycans in attaining the correct fold is, however, ambiguous. There is evidence for protein misfolding and aggregation in the absence of glycans (e.g. the folding of the adhesion domain of the human immune cell receptor cluster of differentiation (CD)) [15,16••]. In other cases, however, elimination of some or all glycans has no effect on folding or protein function, implying that some glycosylation sites are more crucial to protein folding or function than others and that the effect of glycans on folding is likely to be local.

Several factors may govern the effect of glycosylation on the protein energy landscape and, consequently, on the biophysical characteristics of the protein. From the oligosaccharide perspective, the size of the glycan, its flexibility, and structure (e.g. the number of branches) may affect the properties of the protein. From the protein perspective, the location and chemical surroundings of the glycosylation sites, as well as the number of occupied glycosylation sites, may influence the effect of glycosylation on the protein. Quantifying how glycosylation can affect protein biophysics by modulating the protein's energy landscape may formulate a 'glycosylation code'. Deciphering such a molecular code, however, is a difficult task for two main reasons. First, both proteins and oligosaccharides are complex and heterogeneous systems, with proteins composed of 20 types of amino acids and glycans assembled from 9 different monosaccharides. Oligosaccharides and polysaccharides are structurally more complex than proteins (although often they lack folding capability) since they can form various topological trees via various linkage positions while the proteins are linear polymers. The glycan chains found on a protein may not only differ between different organisms, but various different glycans can also be present on one type of protein in one single cell or tissue [17]. Second, the number of glycoproteins having a resolved 3D structure is very limited (while ~50% of all proteins are glycosylated only ~3.5% of the proteins in the Protein Database (PDB) carry covalently bound glycan chains and even fewer have full structure of the glycans) [18,19•]. There are several reasons for the limited structural information on glycoproteins, including the fact that the conformational heterogeneity of the oligosaccharides often inhibits their crystallization or makes them barely visible in X-ray crystallography. One may also note that the carbohydrate moieties of PDB entries are subject to a large number of errors (mainly because of the complexity of the oligosaccharides and poor assignment), which hinders efforts to perform a structural survey of glycoproteins. Solving structural aspects of glycoproteins [20–22] (which can be termed the 'glycosylation problem' [20]) may permit a more comprehensive investigation of the 'glycosylation code'.

This review discusses the role that some molecular features of *N*-glycosylation may play in the energy landscape of glycoproteins.

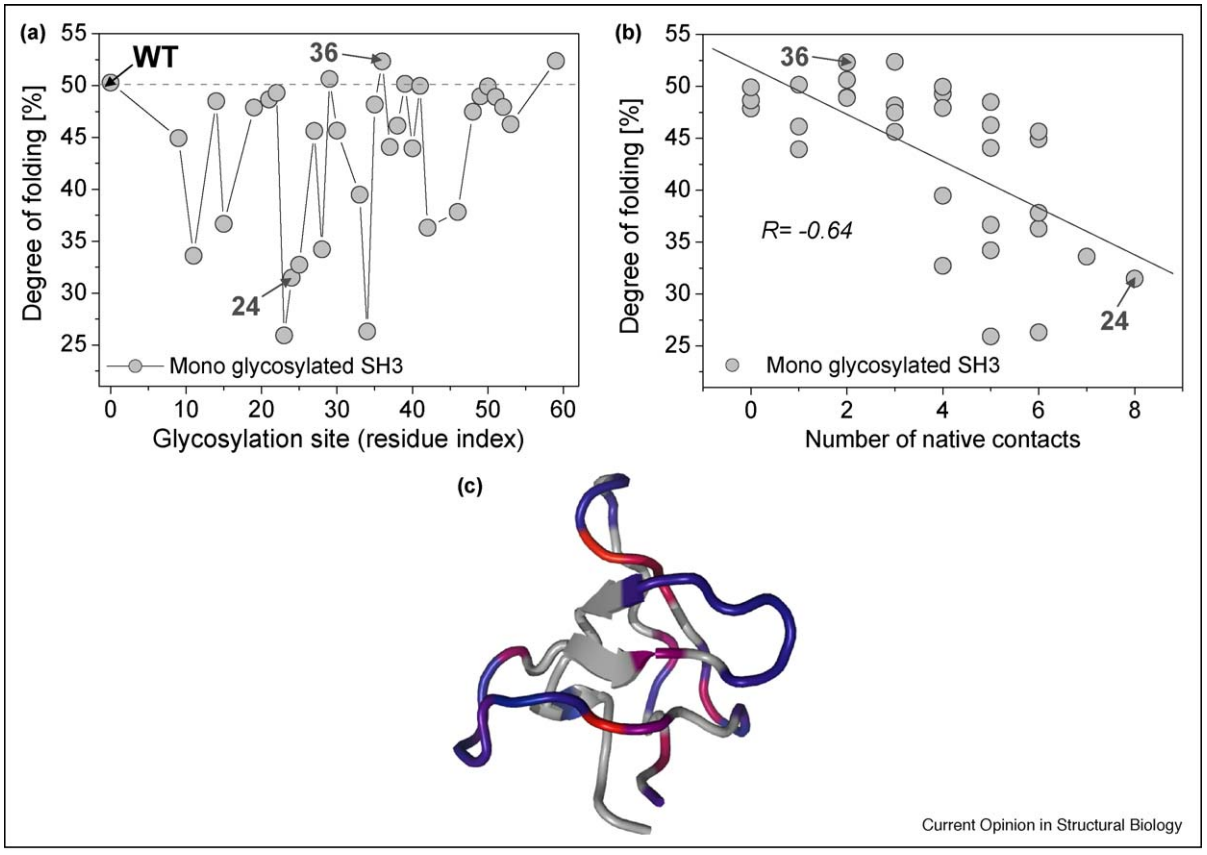
Glycosylation biophysics: increased stability and accelerated kinetics

Glycans, which are bulky hydrophilic polymers, often increase solubility of the protein to which they are attached as well as its resistance to proteolysis. Moreover, the covalent binding of glycans to protein surfaces may enhance their thermodynamic and kinetic stability. Greater thermodynamic stability is expressed by an increase in the melting temperature of the glycosylated protein and/or in the free energy difference between the unfolded and folded state ensembles being greater for the glycosylated protein than for its nonglycosylated wild-type counterpart. Although numerous studies report enhanced thermodynamic stability for *N*-linked proteins [15,23–25], contradictory results were reported regarding the effect of *O*-glycosylation on protein stability [26]. The limited data available make it difficult to generalize about the relation between the type of protein glycosylation and the thermodynamics of glycoproteins. Several studies of the folding and stability of glycoproteins in their glycosylated and nonglycosylated forms show that the sugars have diverse effects on both folding kinetics and pathway [27•].

Effect of the position of the glycosylation site

Thermodynamic investigations of glycosylated proteins suggest that, while enhanced stability is achieved by the carbohydrates when they are conjugated at certain sites of the protein, attachment to other sites improves stability to lesser extent, if at all. We designed *in silico* artificial glycosylated variants of the Src Homology 3 domain (SH3, a small domain of 64 amino acid residues [3••]). SH3 was used in our case study, although it is not glycosylated, since its folding was extensively studied using both experimental and computational tools. To obtain a comprehensive understanding of the effect of the glycosylation site on the biophysical effect of glycosylation, a glycan (either $Man_3GlcNAc_2$ or $Man_9GlcNAc_2$) was individually attached to all residues of SH3 that are exposed to the solvent (35 out of the 64 residues of SH3), resulting in 35 variants of monoglycosylated SH3. A full description of the molecular dynamics simulations and thermodynamic stability calculations we performed appears in [3••]. A large difference in the effect of a single glycan on protein stability as a function of glycosylation site was observed (Figure 1A). Namely, attachment of a sugar can either increase or decrease protein stability (reflected by the relative population of the folded and unfolded states) depending on the position

Figure 1



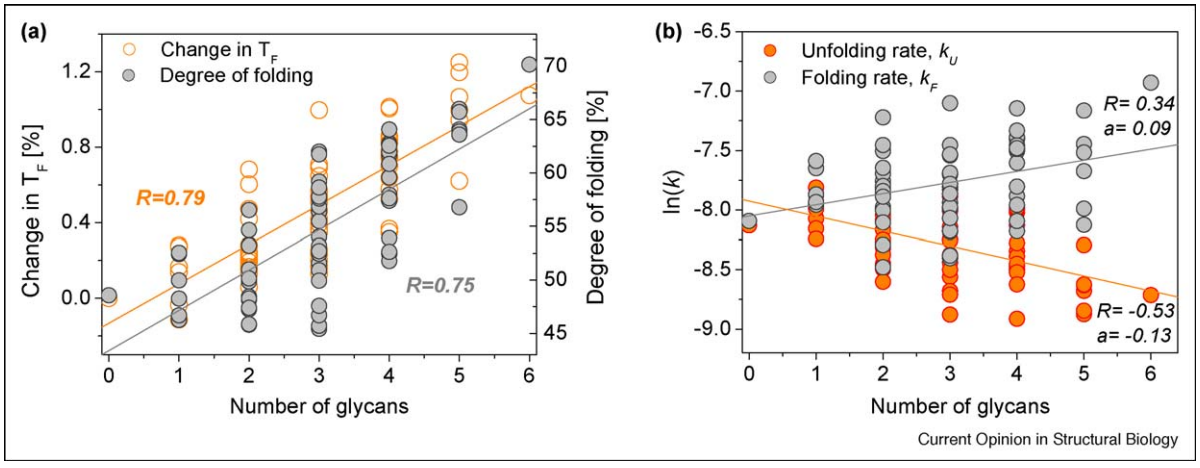
Effect of position of glycosylation site on the thermodynamic stability of proteins. SH3 was taken as a study case and 35 out of its 64 residues were monoglycosylated proteins with $Man_9GlcNAc_2$. Glycosylation at position *i* was achieved after mutating the residues at position *i* and *i* + 2 to Asn and Ser, respectively. Glycosylation was not introduced in the other 29 sites either because these residues were invisible in the NMR or because they were buried and therefore could not host a glycan. Stabilization is reflected by a degree of folding larger than 50% (the calculations were performed at the temperature at which wild-type SH3, which is nonglycosylated, has a degree of folding of 50%; namely, its folded and unfolded state are equally populated). The thermodynamic stability calculations are based on molecular dynamics simulations using native topology-based models [3••]. The degree of folding is shown versus **(A)** the residue index of the glycosylation sites and **(B)** the number of native contacts each glycosylated residue participates in (the glycans act as a bulky polymer that does not form any attractive interaction with the protein). For illustration, the variants with attachment at position 24 and 36 are highlighted. Glycosylation at position 24 results in destabilization as reflected by a shift in the population of the folded state from 50% to about 30%. Glycosylation at position 36, results in stabilization that is shown by the mild increase in the population of the folded state from 50% to about 52%. Panel B suggests that the effect on stability by attachment at different positions is correlated with the number on intramolecular native contacts each position participates in. Residues 24 and 36 participate in 8 and 2 native contacts, respectively. **(C)** Location of the 35 designed glycosylation sites of SH3, color-coded according to their effect on folding temperature, illustrates that they are located in loops (blue; high degree of folding) are more effective in enhancing protein stability than other sites (red; low degree of folding).

of the glycosylation site. It was found that stabilization or destabilization by monoglycosylation results in a shift in the degree of folding to 52% or 25%, respectively (at the temperature at which the nonglycosylated protein is 50% folded).

We then analyzed the structural characteristics of the sites with the aim of correlating the properties of the glycosylation sites with the effect of glycosylation on the stability of SH3. Figure 1B shows the degree of folding of the 35 monoglycosylated variants of SH3 as a function of the number of native intramolecular contacts made by the residue that constitutes the glycosylation site. A negative correlation was found between the degree of folding and the number of native contacts, indicating that glycosylation at a more-structured region results in destabilization, while glycosylation at a more-disordered region increases protein stability. The location of the 35 selected glycosylation sites of SH3, color-coded according to their effect on folding temperature (T_F), is depicted in Figure 1C, and illustrates that glycosylation sites located on loops (blue) are more effective in enhancing protein stability in contrast to other sites which are more structured.

Owing to the limited number of glycoproteins whose structure was fully resolved by X-ray crystallography or

Figure 2



Effect of degree of glycosylation on folding thermodynamics and kinetics. **(A)** Effect of the degree of glycosylation (number of glycans, ranging from 0 to 6) of the SH3 domain on the thermodynamic stability of the protein, measured by the change in folding temperature relative to the nonglycosylated protein (T_F; orange; left *y*-axis) or the degree of folding, which indicates the population of the folded state (gray, right *y*-axis). The degree of folding of the glycosylated variants of SH3 was calculated at a temperature where the folded state of the nonglycosylated SH3 is 50% populated. Note that a 1.2% change in T_F corresponds to a folded state population 20% larger than that of the nonglycosylated protein. **(B)** Effect of the number of glycans on the folding and unfolding rates of the glycosylated SH3 domain. All the data correspond to the simulations presented in Ref. [3••].

NMR [18,19•,28], statistical analysis of the structural features of glycosylation sites must be treated with caution. Yet, several structural analyses of glycoproteins provide some interesting observations on the nature of glycosylation sites. Petrescu *et al.*, find that occupied *N*-glycosylation sites can occur on all forms of secondary structure, but turns and bends are favored [18]. Combining the latter observation with the finding of higher stabilization by glycans attached at less structured regions (i.e. residues that are involved in fewer native contacts) may suggest that natural glycosylations are involved in protein stabilization.

The observation that the effect of glycosylation on protein stability is highly sensitive to the location of the glycosylation site in the 3D structure of proteins is supported by previously reported protein conjugations. Tethering a protein to an immobile surface can result in either stabilization or destabilization, depending on the location of the residue used as the tether point [29]. Crosslinking of protein dimers via various covalent bridges introduces a molecular crowding effect that is highly dependent on the position of the crosslinking sites [30].

Effect of the degree of glycosylation

Many glycoproteins contain several glycosylation sites, which may be occupied alternatively or concomitantly. It is therefore of great interest to understand how the biophysical characteristics of glycoproteins are affected by the number of oligosaccharides covalently attached to the proteins and whether there is a cooperative effect between the various glycans. To address this question, we selected six positions on SH3 that stabilized the protein when they were monoglycosylated (i.e. six positions were selected from the 35 sites that were studied) (Figure 1A). For these six selected glycosylation sites, we designed all possible glycosylated variants using the undecasaccharide $Man_9GlcNAc_2$ (structure M9 in [4]) [3••]. This design resulted in 63 variants: 6 variants with a single glycan (one at each of the glycosylation sites), 15 with two glycans, 20 with three glycans, 15 with four glycans, 6 with five glycans, and a single fully glycosylated variant in which all six positions were glycosylated. An increase in the transition temperature [defined in the simulations as the T_F at which the protein has a stability of 0 (i.e. $\Delta G^{Glyco} = \Delta G^{WT} \sim 0$)] is observed as the degree of glycosylation increases (Figure 2A). On average, each glycan increases the transition temperature by about 0.6–0.9°C. The transition temperature of the SH3 domain with six glycans is, accordingly, higher than that of wild-type SH3 by about 3–4°C. A similar increase in thermal stabilization per additional glycan was demonstrated experimentally by the chemical glycosylation of α-chymotrypsin [2•,31] and subtilisin Carlsberg [32] using either the disaccharide lactose or dextran ($Glc(\alpha1–6)_{\sim60}$, a 10 kDa polysaccharide), with the increase in their melting temperature depending on the number of glycans attached [2•,31]. The melting temperatures of α-chymotrypsin and of subtilisin Carlsberg increased by 1°C and 2°C per added glycan, respectively. Accordingly, the experiments and simulations not only share a common stabilization trend as a function of the degree of glycosylation, but also quantitatively predict similar magnitudes of stabilization.

To obtain a more quantitative understanding of the effect of glycosylation on protein stability, we calculated the degree of folding of the 63 variants of the glycosylated SH3 domain. These calculations were performed at the temperature at which wild-type SH3 is 50% folded. Figure 2A (right-hand *y*-axis) illustrates that the degree of folding increases significantly with the degree of glycosylation by about 4% per added glycan. Glycosylation thus profoundly shifts the population of the states that comprise the energy landscape of proteins, with the folded state being more populated than the unfolded state.

The degree of glycosylation also affects protein folding kinetics (Figure 2B). The unfolding rate decreases as the degree of glycosylation increases, indicating kinetic stabilization of the protein. The folding rate, by contrast, accelerates with increasing degree of glycosylation, but still the unfolding rate is affected more substantially than the folding rate. The effect of glycosylation on folding and unfolding kinetics is supported by experimental results, which show a milder effect (if any) on the folding rate than on the unfolding rate [33,34].

Effect of the oligosaccharide size

While the number of glycans conjugated to the proteins at the glycosylation sites significantly affects the magnitude of protein stabilization, the size of the oligosaccharides has a minor effect. Conjugation of α-chymotrypsin with dextran, produces a similar thermal stabilization to that obtained from the attachment of lactose [2•,31]. The minor effect of glycan size was also observed in a computational study of SH3 when the pentasaccharide $Man_3GlcNAc_2$ and undecasaccharide $Man_9GlcNAc_2$ were shown to produce similar effects on protein stability [3••]. One may note that attachment, of polyethylene glycol to proteins similarly to glycosylation, also increases protein thermostability independently of the size of the polyethylene glycol [35•].

Recently, a detailed thermodynamic and kinetic study of folding was performed on a mono-*N*-glycosylated variant of the adhesion domain of the human immune cell receptor CD2 (an Ig domain) [16••]. Of the 14 monosaccharides of $Glc_3Man_9GlcNAc_2$, the first unit was found to account for two-thirds of the protein stabilization effect relative to the nonglycosylated form. The remaining third of the stabilization is contributed mostly by the next two sugar units. The core trisaccharides $ManGlcNAc_2$ of the above glycan was responsible for an overall enhancement of 3.1 kcal/mol in the stability of the glycosylated protein and for a fourfold increase in folding kinetics. The glycosylation of this protein, which does not fold in its deglycosylated form (ΔG = 0.4 kcal/mol), stabilizes the protein more efficiently than do mutations based on homologs of this protein with higher stability.

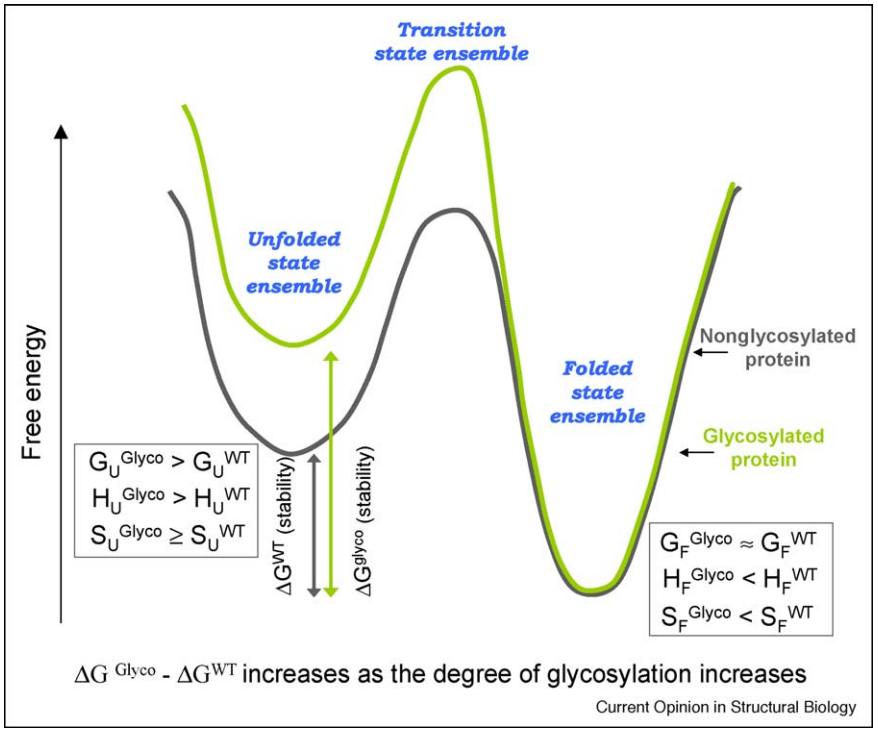
Stabilization by glycosylation: enthalpic or entropic effect?

Protein stabilization by glycosylation, indicated either by a higher melting temperature or by a larger free energy difference between the folded and unfolded states for the glycosylated protein than for its nonglycosylated counterpart, has been reported for several proteins [2•], such as avidin [23], ribonuclease [36], dihydrofolate reductase [37], and the CD2 protein [16••]. Yet, currently the mechanism of this enhanced stabilization is unclear. Several different glycosylation-induced changes in the energy landscape of a protein can explain the increased stability, which may originate from effects on the folded or unfolded states. More specifically, stabilization of the folded state or destabilization of the unfolded state will both result in enhanced stability for the folded protein. Furthermore, one has to ask whether the enhanced stability (i.e. the lower or higher free energy of the folded or unfolded states, respectively) is of enthalpic or entropic origin.

In some structural studies of glycoproteins, specific interactions between the sugar and amino acids of the folded protein were observed [15,38], namely, the stabilization effect in these cases is enthalpic in origin. Earlier studies suggested that the origin of glycoprotein's thermal stabilization is largely entropic, rather than enthalpic [23,39–41] proposing a 'chaperone-like' activity of glycans [11]. Accordingly, there is a controversy as to whether glycans affect the entropy or enthalpy and whether these effects are mainly on the folded or unfolded state (or a combination of both). Experimental investigations often do not provide direct insight into the origin of this stabilization and theoretical studies therefore serve a complementary role [42,43].

Decomposing the folding free energy into the free energies of the folded and unfolded states of the 63 variants of SH3 with different degrees of glycosylation reveals interesting insights into the origin of stabilization by glycosylation [3••]. The decrease in the free energy for folding as a function of the degree of glycosylation is associated with a clear increase in the free energy of the ensemble of the unfolded state (i.e. destabilization of the unfolded state) [3••]. The free energy of the folded state ensemble remains unchanged, independently of the degree of glycosylation. Further decomposition of the free energy of the folded state into its enthalpy and entropy components reveals that it is actually not altered by the glycosylation. Both the enthalpy and the entropy of the folded state decrease as the degree of glycosylation increases. The reduced entropy induced by glycosylation is caused by the suppression of protein fluctuations in the folded state while its conformation remains unchanged [15,40,44,45••]. As was suggested by several experimental studies [2•], the glycans can therefore be viewed as imposing pressure on the protein that reduces the

Figure 3

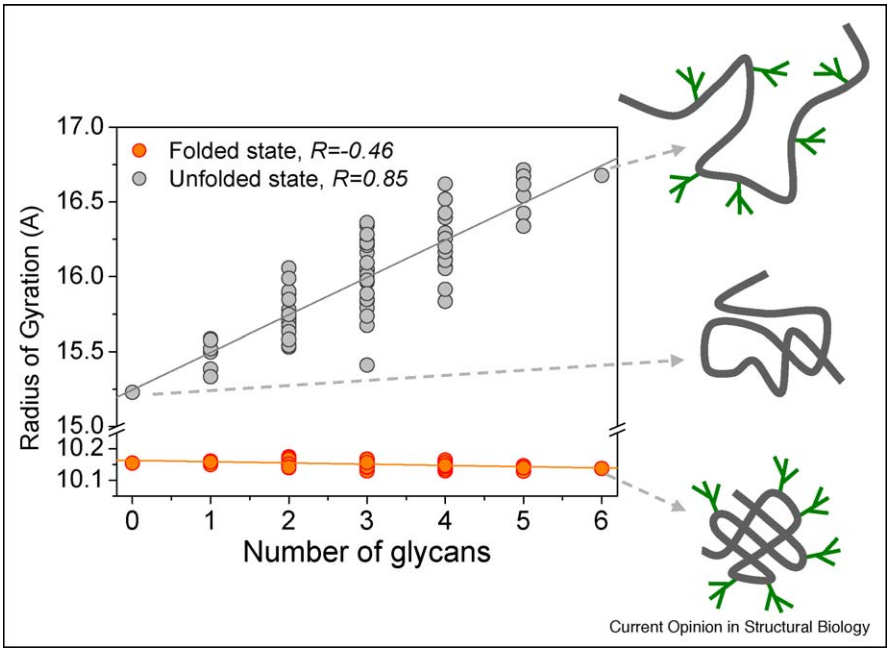


Schematic free energy profile of a glycoprotein. Carbohydrates often stabilize proteins by destabilizing the unfolded state. The destabilization correlates with higher enthalpy because of the inhibition of residual structure formation in the unfolded state. The folded state of the glycoprotein is characterized by lower entropy and enthalpy because of the suppression of its native flexibility compared to the nonglycosylated variant, but overall the free energy of the folded state is mildly affected by the glycosylation. Destabilization of the unfolded state increases together with the degree of glycosylation and, to a small extent, with the size of the glycans.

flexibility of the folded state. For example, it was found that glycosylation affects the structural dynamics of α-chymotrypsin and subtilisin Carlsberg [32]. We emphasize, however, that a change in dynamics (i.e. entropy) can be coupled with a change in energy (i.e. enthalpy). In our theoretical study, we found that the rigidified structure induced by glycosylation is characterized by tertiary interactions that are more localized in their optimal distances; thus, this rigidified structure and thus has a lower enthalpy [3••]. The magnitude of the decrease in the enthalpy and entropy of the folded state is very similar, and since these two effects operate in opposing directions, they compensate for each other and, therefore, the net change in the free energy of the folded state is negligible.

Decomposing the free energy of the unfolded state into its enthalpic and entropic components also reveals that the major contribution to the increase in free energy when more glycans are added is enthalpic in origin. The glycans affect the dynamics of the unfolded proteins. Owing to their bulkiness, the glycans restrict the conformational space of the unfolded glycosylated polypeptide chain in comparison to that of the unfolded nonglycosylated protein. Numerous experimental and theoretical studies support the understanding that, although proteins in the unfolded state can be described by polymer physics as having the statistics of a random coil [46,47], they are often not without structure. Residual structure supported by native contacts or by the transient formation of non-native contacts [48,49] may result in an unfolded state with nonzero enthalpy (and its entropy is expected to be lower than that of a random coil polymer of the same size). Attachment of oligosaccharides can obviously affect the enthalpy and entropy of the unfolded protein (Figure 3). Prohibiting formation of transient interactions in the unfolded state results in an unfolded state with a higher enthalpy (i.e. in destabilization of the unfolded state) that increases with the degree of glycosylation. The effect of the oligosaccharides on the entropy of the unfolded state is more complicated. The glycans may confine the conformational space of the unfolded state, which will shrink its entropy, or the restriction on the formation of residual structures may contribute to an increase in the entropy of the folded state. In our study, we found a mild increase in the entropy of the unfolded state when more glycans are added, namely the stability of the unfolded state is governed by the change in enthalpy.

Figure 4



Effect of glycosylation on the dimensions of the folded and unfolded states. The radii of gyration of the folded and unfolded states of the glycosylated variants of SH3 (see Figure 1) are drawn as a function of the degree of glycosylation. While the size of the folded conformations is hardly affected by glycosylation, the unfolded conformation expands and becomes less compact because of the unwinding of the residual structure by the bulky sugar groups.

Glycosylation affects the unfolded state ensemble

As explained in the foregoing, glycosylation increases the enthalpy of the unfolded state, which results in a higher free energy. The destabilization of the unfolded state ensemble because of reduced residual structure is reflected in a significant increase in the radius of gyration of the unfolded protein (Figure 4). The conformations in the unfolded state ensemble of the glycosylated proteins are more extended in comparison to the nonglycosylated form and their dimensions correlate with the degree of glycosylation. SH3 to which to six glycans are attached is ~10% larger in its unfolded state than is the nonglycosylated SH3. The conformations of the folded state ensemble, by contrast, remain very compact regardless of the degree of glycosylation (the folded state actually becomes slightly more compact, in agreement with the observed reduction in the flexibility induced by the glycans). Recent small angle X-ray scattering measurements of *Pheniophora lycii* phytase, a glycoprotein, show that the dimensions of the unfolded and folded conformations are respectively 32% and 11% larger than those of the deglycosylated protein, respectively (note that the calculated radii of gyration in these measurements include contributions from both the protein and the glycans [45••]). The expansion of the unfolded conformations supports the notion that the glycans affect the dynamics of the unfolded state.

The expansion of the unfolded state for different numbers of glycans occurs in an almost additive manner. Each glycan restricts the formation of interactions in the surroundings of its attachment sites and therefore disturbs the local structure in its vicinity. The local effect of each glycan depends on the topological characteristics of its glycosylation site (e.g. on the number of native contacts the residue at the site is involved in, as discussed above) and can be treated as the 'fingerprint' of the glycosylation site [3••]. The local perturbation of the unfolded state caused by the glycans is in accord with the observation that the first sugars are more effective in affecting protein stability (i.e. the length of the glycans has a minor effect). Furthermore, the finding that the effect of the glycans on protein stability is additive supports the notion that each glycan locally affects the dynamics of the unfolded state and that there is no significant long-range communication between the glycans.

Conclusions

The roles of glycosylation as well as other PTMs in controlling protein function are widely acknowledged, yet their effects on the physico-chemical nature of proteins are poorly understood and have not received sufficient attention. In this article, we discuss the thermodynamic and kinetic effects of glycosylation on protein folding, motivated by a desire to formulate the molecular

components that dictate the cross-talk between the carbohydrates and the protein. It is possible that cross-talk between different PTMs (e.g. between glycosylation and ubiquitination [50]) is mediated via changes in the biophysical characteristics of the protein. There is evidence that mutating the glycosylation sites of a protein *in vivo* reduces the efficiency of folding and increases degradation [51]. It is possible that the folding is less efficient because the interaction with the ER lectins chaperons is interrupted. According to this scenario the chances of correct folding are lower because the direct lectin–glycan interaction, which has a chaperoning effect, is abolished and therefore results in misfolded polypeptides that are easily degraded [51]. Alternatively, the inherent instability that occurs as a result of elimination of the glycans, evidence for which was presented in this article, may promote degradation rather than correct folding of the naked protein and its secretion.

The magnitude of thermodynamic protein stabilization by glycosylation depends on the properties of both the carbohydrate and the protein moieties. From the carbohydrate perspective, enhanced stability is correlated with degree of glycosylation and, to a much smaller extent, with the size of the attached glycans. Actually, the first two to three saccharides of large glycans often contribute almost all the stabilization effect. A similar phenomenon was observed for PEGylated proteins where increasing the length of the polyethylene glycol very mildly affected protein stability [52]. From the protein perspective, the degree of glycosylation strongly depends on the position of the attached glycans. Greater stabilization effect is achieved when the attached glycan is located at a more flexible region (e.g. a loop) of the protein. Accordingly, the same degree of glycosylation may result in a different thermal effect depending on the location of the sugars. When the glycosylation site is located at a highly structured region, a significant destabilization occurs. The finding that occupied glycosylation sites tend to be located in flexible regions of the protein suggests that these sites are selected to yield increase its stability.

Stabilization by glycosylation often correlates with reduced flexibility of the folded protein, yet this reduced entropy is compensated for by reduced enthalpy. The free energy of the folded state, therefore, may be only mildly affected by the conjugation of the carbohydrates. The free energy of the unfolded state increases upon glycosylation and this may result from an increase in enthalpy and/or decrease in entropy. Our computational study indicates a significant change in the enthalpy of the unfolded state because of the inhibition of the formation of residual structures, which are normally formed in the absence of the conjugates [3••]. This provides further evidence for the importance of the unfolded state in modulating folding biophysics and that its polymer physics is richer than that of a random coil polymer.

The stabilization effect of glycosylation is reminiscent of the protein stabilization introduced by molecular crowding or by confinement. Although glycosylation changes the folding transition temperature by just ~1–2% (for comparison, molecular crowding and confinement may yield a change of ~7% [53] and ~25% [54–57], respectively), it is associated with a significant population shift from the unfolded to the folded state. It is likely that glycoproteins in a crowded or confined environment will exhibit an enhanced stabilization effect, yet the synergism between the carbohydrates and the crowded or confined environments is unclear.

The effect of glycosylation on the biophysical characteristics of proteins may suggest that perturbation of the protein structure by conjugation may be a general mechanism by which proteins may be modulated and enriched. Furthermore, the modulation of protein characteristics may support the biological activity of the PTM.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Johnson LN, Lewis RJ: **Structural basis for control by phosphorylation**. *Chem Rev* 2001, **101**:2209-2242.

2. • Sola RJ, Rodriguez-Martinez JA, Griebenow K: **Modulation of protein biophysical properties by chemical glycosylation: biochemical insights and biomedical implications**. *Cell Mol Life Sci* 2007, **64**:2133-2152.

A survey on experimental studies of the effect of natural and chemical glycosylation on biophysical characteristics of proteins.

3. •• Shental-Bechor D, Levy Y: **Effect of glycosylation on protein folding: a close look at thermodynamic stabilization**. *Proc Natl Acad Sci U S A* 2008, **105**:8256-8261.

An extensive computational study of the effect of glycosylation on protein folding. Using coarse-grained models, a quantitative analysis of the effect of the degree of glycosylation on folding thermodynamics and kinetics is performed. This study indicates that the origin of stabilization is the destabilization of the unfolded state by eliminating the formation of residual structures.

4. Lederkremer GZ: **Glycoprotein folding, quality control and ER-associated degradation**. *Curr Opin Struct Biol* 2009, **19**:515-523.

5. Daggett V, Fersht AR: **Protein folding and unfolding at atomic resolution**. *Cell* 2002, **108**:573-582.

6. Oliveberg M, Wolynes P: **The experimental survey of protein-folding energy landscapes**. *Q Rev Biophys* 2005, **38**:245-288.

7. Onuchic JN, Wolynes PG: **Theory of protein folding**. *Curr Opin Struct Biol* 2004, **14**:70-75.

8. Varki A: **Biological roles of oligosaccharides: all of the theories are correct**. *Glycobiology* 1993, **3**:97-130.

9. Spiro RG: **Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds**. *Glycobiology* 2002, **12**:43R-56R.

10. Sharon N, Lis H: **Carbohydrates in cell recognition**. *Sci Am* 1993, **268**:82-89.

11. Helenius A, Aebi M: **Roles of *N*-linked glycans in the endoplasmic reticulum**. *Annu Rev Biochem* 2004, **73**:1019-1049.

12. Lis H, Sharon N: **Lectins: carbohydrate-specific proteins that mediate cellular recognition**. *Chem Rev* 1998, **98**:637-674.

13. Trombetta E, Parodi A: **Quality control and protein folding in the secretory pathway**. *Annu Rev Cell Dev Biol* 2003, **19**:649-676.

14. Molinari M: ***N*-Glycan structure dictates extension of protein folding or onset of disposal**. *Nat Chem Biol* 2007, **3**:313-320.

15. Wyss D, Choi J, Knoppers LIJ, Willis M, Arulanandam K, Smolyar A, Reinherz A, Wagner EG: **Conformation and function of the *N*-linked glycan in the adhesion domain of human CD2**. *Science* 1995, **269**:1273-1278.

16. •• Hanson SR, Culyba EK, Hsu TL, Wong CH, Kelly JW, Powers ET: **The core trisaccharide of an *N*-linked glycoprotein intrinsically accelerates folding and enhances stability**. *Proc Natl Acad Sci U S A* 2009, **106**:3131-3136.
A detailed thermodynamic and kinetic investigation of the effect of a single glycan on the folding characteristics of CD2 protein studied at denaturation conditions. By a gradual trimming of the oligosaccharide tree the authors elegantly show that most of the stabilization effect is contributed from the first three sugars. This result supports previous studies reporting that the length of the oligosaccharide has only a minor effect of stability and that the effect of glycan is local.

17. Rudd PM, Dwek RA: **Glycosylation: heterogeneity and the 3D structure of proteins**. *Crit Rev Biochem Mol Biol* 1997, **32**:1-100.

18. Petrescu AJ, Milac AL, Petrescu SM, Dwek RA, Wormald MR: **Statistical analysis of the protein environment of *N*-glycosylation sites: implications for occupancy, structure, and folding**. *Glycobiology* 2004, **14**:103-114.

19. • Lutteke T: **Analysis and validation of carbohydrate three-dimensional structures**. *Acta Crystallogr D Biol Crystallogr* 2009, **65**:156-168.
A comprehensive review on common structural features of glycoproteins. The author discusses difficulties and challenges in structural determination of glycosylated proteins.

20. Chang VT, Crispin M, Aricescu AR, Harvey DJ, Nettleship JE, Fennelly JA, Yu C, Boles KS, Evans EJ, Stuart DI *et al.*: **Glycoprotein structural genomics: solving the glycosylation problem**. *Structure* 2007, **15**:267-273.

21. Petrescu AJ, Wormald MR, Dwek RA: **Structural aspects of glycomes with a focus on *N*-glycosylation and glycoprotein folding**. *Curr Opin Struct Biol* 2006, **16**:600-607.

22. DeMarco ML, Woods RJ: **Structural glycobiology: a game of snakes and ladders**. *Glycobiology* 2008, **18**:426-440.

23. Wang C, Eufemi M, Turano C, Giartosio A: **Influence of the carbohydrate moiety on the stability of glycoproteins**. *Biochemistry* 1996, **35**:7299-7307.

24. DeKoster G, Robertson A: **Thermodynamics of unfolding for kazal-type serine protease inhibitors: entropic stabilization of ovomucoid first domain by glycosylation**. *Biochemistry* 1997, **36**:2323-2331.

25. Pham VT, Ewing E, Kaplan H, Choma C, Hefford MA: **Glycation improves the thermostability of trypsin and chymotrypsin**. *Biotechnol Bioeng* 2008, **101**:452-459.

26. Spiriti J, Bogani F, van der Vaart A, Ghirlanda G: **Modulation of protein stability by *O*-glycosylation in a designed Gc-MAF analog**. *Biophys Chem* 2008, **134**:157-167.

27. • Mitra N, Sinha S, Ramya TN, Surolia A: ***N*-Linked oligosaccharides as outfitters for glycoprotein folding, form and function**. *Trends Biochem Sci* 2006, **31**:156-163.
A comprehensive review that describes the capability of glycans to affect protein folding, oligomerization, and stability. A number of examples are discussed to illuminate the various mechanisms used to modulate the properties of the protein.

28. Ben-Dor S, Esterman N, Rubin E, Sharon N: **Biases and complex patterns in the residues flanking protein *N*-glycosylation sites**. *Glycobiology* 2004, **14**:95-101.

29. Friedel M, Baumketner A, Shea JE: **Stability of a protein tethered to a surface**. *J Chem Phys* 2007, **126**:095101.

30. Kim YH, Stites WE: **Effects of excluded volume upon protein stability in covalently cross-linked proteins with variable linker lengths**. *Biochemistry* 2008, **47**:8804-8814.

31. Sola RJ, Al-Azzam W, Griebenow K: **Engineering of protein thermodynamic, kinetic, and colloidal stability: chemical glycosylation with monofunctionally activated glycans**. *Biotechnol Bioeng* 2006, **94**:1072-1079.

32. Pagan M, Sola RJ, Griebenow K: **On the role of protein structural dynamics in the catalytic activity and thermostability of serine protease subtilisin Carlsberg**. *Biotechnol Bioeng* 2009, **103**:77-84.

33. Kwon K, Yu M: **Effect of glycosylation on the stability of alpha1-antitrypsin toward urea denaturation and thermal deactivation**. *Biochim Biophys Acta* 1997, **1335**:265-272.

34. Yamaguchi H: **Chaperone-like functions of *N*-glycans in the formation and stabilization of protein conformation**. *Trends Glycosci Glycotechnol* 2002, **14**:139-151.

35. • Rodriguez-Martinez JA, Sola RJ, Castillo B, Cintron-Colon HR, Rivera-Rivera I, Barletta G, Griebenow K: **Stabilization of alpha-chymotrypsin upon PEGylation correlates with reduced structural dynamics**. *Biotechnol Bioeng* 2008, **101**:1142-1149.
An investigation of the effect of PEGylation of protein folding showing that PEGylation and glycosylation share many features in common.

36. Xu G, Narayan M, Scheraga HA: **The oxidative folding rate of bovine pancreatic ribonuclease is enhanced by a covalently attached oligosaccharide**. *Biochemistry* 2005, **44**:9817-9823.

37. Swanwick RS, Daines AM, Tey LH, Flitsch SL, Allemann RK: **Increased thermal stability of site-selectively glycosylated dihydrofolate reductase**. *Chembiochem* 2005, **6**:1338-1340.

38. Fusetti F, Schröter K, Steiner R, van Noort P, Pijning T, Rozeboom H, Kalk K, Egmond M, Dijkstra B: **Crystal structure of the copper-containing quercetin 2,3-dioxygenase from *Aspergillus japonicus***. *Structure* 2002, **10**:259-268.

39. Riederer MA, Hinnen A: **Removal of *N*-glycosylation sites of the yeast acid phosphatase severely affect protein folding**. *J Bacteriol* 1991, **173**:3539-3546.

40. Rudd P, Joao H, Coghill E, Fiten P, Saunders M, Opdenakker G, Dwek R: **Glycoforms modify the dynamic stability and functional activity of an enzyme**. *Biochemistry* 1994, **33**:17-22.

41. Sinha S, Surolia A: **Attributes of glycosylation in the establishment of the unfolding pathway of soybean agglutinin**. *Biophys J* 2007, **92**:208-216.

42. Bosques C, Tschampel S, Woods R, Imperiali B: **Effects of glycosylation on peptide conformation: a synergistic experimental and computational study**. *J Am Chem Soc* 2004, **126**:8421-8425.

43. Hansia P, Dev S, Surolia A, Vishveshwara S: **Insight into the early stages of thermal unfolding of peanut agglutinin by molecular dynamics simulations**. *Proteins* 2007, **69**:32-42.

44. Wormald MR, Petrescu AJ, Pao YL, Glithero A, Elliott T, Dwek RA: **Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling**. *Chem Rev* 2002, **102**:371-386.

45. •• Hoiberg-Nielsen R, Westh P, Arleth L: **The effect of glycosylation on interparticle interactions and dimensions of native and denatured Phytase**. *Biophys J* 2009, **96**:153-161.
An investigation of the effect of glycosylation on the dimensions of the folded and unfolded states. Using SAXS measurements the authors show that glycosylation of *Pheniophora lycii* phytase results in a much more expanded unfolded state while the dimension of folded state is less affected by the conjugation. This result supports the restriction of formation of residual structure in the unfolded state by glycosylation as reported in Ref. [3••].

46. Fitzkee NC, Rose GD: **Reassessing random-coil statistics in unfolded proteins**. *Proc Natl Acad Sci U S A* 2004, **101**:12497-12502.

47. Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR *et al.*: **Random-coil behavior and the dimensions of chemically**

unfolded proteins. *Proc Natl Acad Sci U S A* 2004, **101**:12491-12496.

48. Cho JH, Sato S, Horng JC, Anil B, Raleigh DP: **Electrostatic interactions in the denatured state ensemble: their effect upon protein folding and protein stability**. *Arch Biochem Biophys* 2008, **469**:20-28.

49. Mor A, Haran G, Levy Y: **Characterization of the unfolded state of repeat proteins**. *HFSP J* 2008, **2**:405-415.

50. Lederkremer GZ, Glickman MH: **A window of opportunity: timing protein degradation by trimming of sugars and ubiquitins**. *Trends Biochem Sci* 2005, **30**:297-303.

51. de Virgilio M, Kitzmüller C, Schwaiger E, Klein M, Kreibich DG, Ivessa NE: **Degradation of a short-lived glycoprotein from the lumen of the endoplasmic reticulum: the role of *N*-linked glycans and the unfolded protein response**. *Mol Biol Cell* 1999, **10**:4059-4073.

52. Castillo B, Sola RJ, Ferrer A, Barletta G, Griebenow K: **Effect of PEG modification on subtilisin Carlsberg activity, enantioselectivity, and structural dynamics in 1,4-dioxane**. *Biotechnol Bioeng* 2008, **99**:9-17.

53. Cheung MS, Klimov D, Thirumalai D: **Molecular crowding enhances native state stability and refolding rates of globular proteins**. *Proc Natl Acad Sci U S A* 2005, **102**:4753-4758.

54. Thirumalai D, Lorimer GH: **Chaperonin-mediated protein folding**. *Annu Rev Biophys Biomol Struct* 2001, **30**:245-269.

55. Takagi F, Koga N, Takada S: **How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: molecular simulations**. *Proc Natl Acad Sci U S A* 2003, **100**:11367-11372.

56. Xu WX, Wang J, Wang W: **Folding behavior of chaperonin-mediated substrate protein**. *Proteins* 2005, **61**:777-794.

57. Friedel M, Sheeler DJ, Shea JE: **Effects of confinement and crowding on the thermodynamics and kinetics of folding of a minimalist beta-barrel protein**. *J Chem Phys* 2003, **118**:8106-8113.