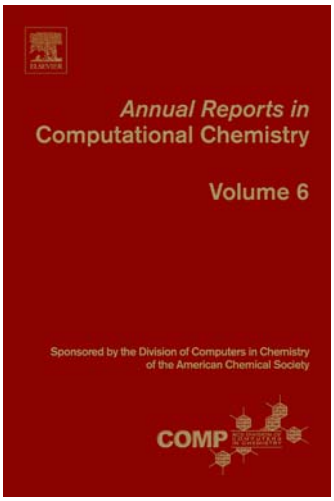

From Yaakov Levy, Folding of Conjugated Proteins.
In: Ralph A. Wheeler, editor,
Annual Reports in Computational Chemistry (Volume 6).
Elsevier, 2010, p. 263.
ISBN: 978-0-44-453552-8

CHAPTER 13

Folding of Conjugated Proteins

Dalit Shental-Bechor, Oshrit Arviv, Tzachi Hagai, and **Yaakov Levy**

Contents

Abstract

This review aims at discussing the molecular details of the folding mechanisms of conjugated proteins using computational tools. Almost all studies of protein folding focus on individual proteins and do not consider how interactions with posttranslational modifications and between domains might affect folding. However, different chemical conjugations may introduce a variety of effects on the protein biophysics. These effects depend both on the chemical characteristics of the protein substrate as well as on the chemical and physical properties of the attachment. We review the folding of various types of conjugated proteins, glycoproteins, proteins with tails, ubiquitinated proteins, and multidomain proteins, to explore the underlying biophysical principles of these complex folding processes and in particular to quantify the cross-talk between the protein and its conjugated polymer.

Keywords: protein folding; multidomain proteins; glycosylation; ubiquitination; coarse-grained models

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

Annual Reports in Computational Chemistry, Volume 6
ISSN: 1574-1400, DOI 10.1016/S1574-1400(10)06013-5

1. INTRODUCTION

The field of protein folding has traditionally focused on the folding of individual proteins in isolation following the paradigm that sequence determines structure and structure determines function. The funnel theory of protein folding hypothesizes that the folding process of a protein is governed by its native structure as was determined by the sequence, since nonnative interactions that may compete with the native interactions and introduce frustration and thus accumulation of traps are minimized [1,2]. However, protein folding in vivo is much more complicated because in the cell there are several factors that may affect the folding. For example, chaperons participate in folding and may change the folding pathway and thermodynamics, as well as the inherently crowded environment of the cell. Another way to affect the folding of a protein is by conjugated moieties as posttranslational modifications (PTMs) (see Figure 1). There is a large variety of PTMs in the cell that serve diverse biological functions. Phosphorylation, for example, is widely used in signal transduction [3]. Ubiquitination, the covalent attachment of the protein ubiquitin, controls the cellular fate of many eukaryotic proteins [4,5]. Sugar trees attached during glycosylation serve as recognition factors to receptors and in protein–protein interactions [6]; the ability of sugar-binding protein receptors (primarily lectins) to recognize carbohydrate conjugates lies at the heart of many central biological processes [7]. Viruses recognize sugars and use them as targets for cell penetration and infection [8]. Myristoylation and palmytoylation, the covalent attachment of a fatty acid, help protein trafficking and membrane association [9].

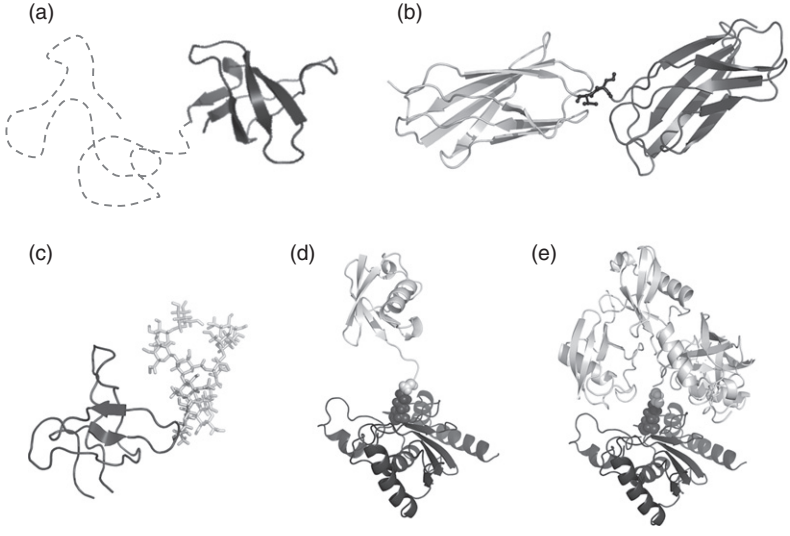

Figure 1 A gallery of conjugated proteins. (a) A tailed SH3 protein. (b) A multidomain protein (FNfn9–FNfn10). (c) A glycosylated SH3 protein. (d) A ubiquitinated Ubc7 protein (monomeric ubiquitin). (e) A ubiquitinated Ubc7 (Lys48-linked tetrameric ubiquitin).

But do the PTMs have a biophysical effect on the protein that may be related to the biological function of the modified protein? In glycoproteins the glycan is added to the unfolded protein while it is in the translocon complex [10], indicating that it may assist in obtaining the correct fold following the recruitment of lectins, calnexin and calreticulin [11–13]. There is evidence for enhanced thermostability imposed by glycosylation [14]; for example, the human immune cell receptor cluster of differentiation CD2 can fold correctly only after glycosylation [15,16]. In other cases, however, elimination of all or some glycans has no effect on folding or protein function, implying that some glycosylation sites are more crucial to folding or function than others and that the effect of glycans on folding is likely to be local.

The chemical properties of both the oligosaccharide and the protein may govern the effect of glycosylation on the protein energy landscape and hence the biophysical properties of the conjugated protein. On the one hand, the size of the sugar tree, its chemical composition, and its structure (e.g., the number of branches) have an effect on the overall change in the nature of the conjugated protein. On the other hand, the specific glycosylation site and its chemical environment as well as the number of conjugated glycans may modulate the effect of glycosylation on the protein. A quantification of this relationship between the properties of both the sugar and the protein, and the overall influence on the glycoprotein's energy landscape may formulate a "glycosylation code". Deciphering such a molecular code, however, is a difficult task for two main reasons. First, there is large variety in the composition and structure of oligosaccharides. Second, the structural information about glycans in the context of the folded protein is very limited—while about 50% of all proteins are glycosylated only 3.5% of the proteins in the Protein Data Bank (PDB) contain the glycan chains and even fewer entries include the full structure of the glycans [17,18]. By studying the folding of glycosylated proteins in silico we can try to formulate the main characteristics of the interplay between a protein and conjugated sugars.

A similar kind of conjugation is the existence of an unstructured tail at the termini of proteins (see Figure 1). In nature, there are intrinsically disordered proteins that remain very flexible until they interact with a companion protein that induces structure [19,20]. Disordered tails, in particular, have a role in interacting with other biomolecules such as DNA [21,22] and can accelerate binding kinetics via the fly-casting mechanism [23–25]. Using computational methods we can ask if the attachment of a flexible polymer can modulate the biophysical properties of the protein. As with glycoproteins, we can also ask how the characteristics of the tail (length and flexibility) affect protein characteristics. Furthermore, tails can be conjugated to the proteins not only at the termini but as a branch via the side chain. In this case, one can ask how the number of tails and their conjugation sites modulate the protein's biophysical properties.

The modification of proteins by attaching chains of ubiquitin (known as ubiquitination) can serve as another example to study interface effects between domains. Ubiquitination is a unique PTM in that the conjugated modification is a protein or a polymer of proteins, which can be viewed as a special case of

multidomain protein. The covalent attachment of ubiquitin molecules to the substrate protein is done using an isopeptide bond between the C-terminal of ubiquitin and a Lys of the substrate. The conjugated ubiquitin itself can be further connected, by one of its seven lysine residues, to other ubiquitin molecules. The points of attachment in the polymer determine the shape and topology of the modification, as well as the fate of the modified protein [26,27]. The most well-characterized chain topology is that in which the isopeptide link is formed using lysine on position 48 in each of the attached ubiquitin molecules. This type of attachment creates a densely packed tetramer having a large interface with the ubiquitinated substrate (see Figure 1). These Lys48-linked chains are commonly related to protein degradation. Another, very different, elongated topology is obtained by the attachment of subsequent ubiquitin units using the lysine in position 63. The biological function of this ubiquitin tree is nondegenerative, but related to other processes, such as DNA repair. A fundamental question is the effect of ubiquitination on the protein's thermodynamics and kinetics. The ubiquitin attachment may significantly affect the substrate in various ways that may support the function introduced by the conjugation [28].

Another kind of conjugated entity may be a protein domain in the context of multidomain proteins (see Figure 1) [29,30]. Multidomain proteins are very common in genomes and the folding of the tethered domain might be different to that of the isolated domains. A domain is defined as a structural, functional, and evolutionary component of proteins that can often be expressed as a single unit [31]. In fact, implementation of sequence analyses had shown that most eukaryotic and a cardinal part of prokaryotic proteins are composed of more than one domain [32] and that proteins have evolved through vast duplication and shuffling of domains. However, only a small fraction of possible domain combinations can be found in wild-type multidomain proteins. This modular character of a limited set of domain families supported the emergence of complex protein functions. Yet, the existing domain combinations must have also met constraints of folding in the native operative context, in which the domains fold in the presence of their tethered neighboring domains. Folding in a multidomain architecture suggests a conservation of energetically favorable folding pathways also in the perspective of these conjugated constructs.

Multidomain proteins may be viewed as conjugated proteins in which each domain may affect the folding dynamics and thermodynamic properties of its counterpart domain. Experimentally, the thermodynamics and kinetics of both isolated domains and conjugated constructs from several multidomain proteins were studied (a very detailed and fairly current report can be found in Reference [29]). A computational characterization of the mechanistic principles of the folding of multidomain proteins [33], utilizing native structure-based models, provides a reduced microscopic description of their folding, which in turn may enable the formulation of the forces involved in the interplay between neighboring domains.

In this paper, we discuss the biophysical effects that are imposed on a protein by conjugation. Specifically, we ask how the nature of the conjugation (its size, shape, flexibility, and conjugation site) affects protein folding.

2. METHODS

The various conjugated protein systems were studied using coarse-grained models. The protein moiety was studied using a native topology-based model and the conjugate moieties were modeled in various ways to capture their polymeric nature. Both the tail and glycan [34,35] were modeled as flexible polymers. The glycan was represented as a tree of beads where each bead represents a single sugar ring. The rigidity of the glycan was introduced by including an angle potential term between the sugar beads and by the excluded volume effect. The tail was modeled as an entropic chain of beads connected with bonds. The flexibility of the chain was represented solely by the excluded volume. The conjugation of a protein, as in multidomain proteins or ubiquitinated proteins [28], was modeled by the native topology-based model as well, yet one can control the relative stability of the conjugates by constraining the protein dynamics. The details of the models can be found in previous publications [36,37].

We would like to point out that, because conjugated proteins may have inhomogeneous degrees of freedom (i.e., part of the system is significantly more flexible than the rest), special care is required in choosing the thermostat for the molecular dynamics simulations. We have recently reported that the Berendsen and the Langevin thermostats show different abilities to regulate the temperature of systems that include flexible and more rigid regions [38]. In simulations performed using the Berendsen thermostat, the flexible tail is significantly hotter than the protein, both in its folded and unfolded states. Upon weakening the strength of the Berendsen thermostat, the temperature gradient between the fast and the slow degrees of freedom is significantly decreased, yet linkage between the temperatures of the flexible tail and the protein remains. The Langevin thermostat is proven to regulate the temperature of these inhomogeneous systems reliably, without discriminating between the slow and fast degrees of freedom (Figure 2).

3. RESULTS AND DISCUSSION

3.1 Folding of glycoproteins

A thorough investigation of the effect of glycosylation on the stability of the conjugated protein demonstrated that, while in some cases the oligosaccharide increased its thermodynamic stability, in other cases protein stability was not affected or was even reduced. This was observed from folding–unfolding simulations of 35 glycoconjugated variants of the Src Homology domain 3 (SH3). In these simulations the glycan was attached to 35 different solvent-exposed positions on the protein's surface to obtain 35 variants of glycoconjugated proteins with a single oligosaccharide attached. A detailed description of the simulations can be found in References [34,35]. In general, it was observed that the change in protein stability is tightly related to the location of the attached glycan. The influence of the glycan varies between stabilization to significant destabilization, which is reflected by the relative population of the folded and unfolded states.

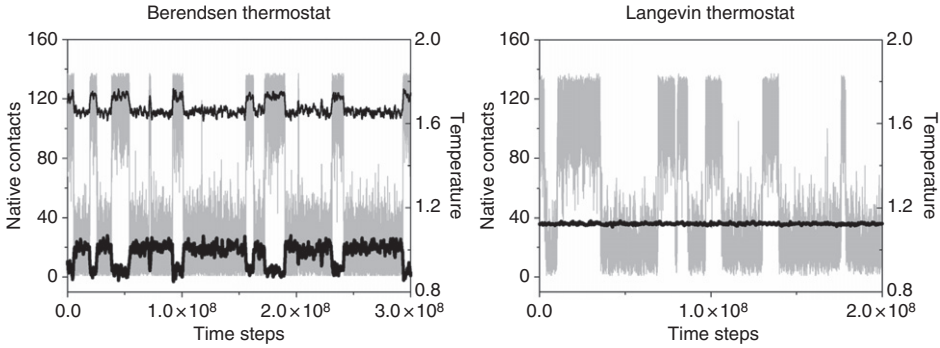

Figure 2 The effect of the thermostat on the temperature of the conjugate. Time evolution of the temperatures of the flexible polymeric tail (composed of 80 residues) and the SH3 domain simulated with the Berendsen thermostat (right panel) and the Langevin thermostat (left panel). The gray lines correspond to the time evolution of the number of native contacts and show several folding/unfolding events. The temperatures of the tail (thin black line) and the SH3 domain (thick black line) illustrate that the Langevin thermostat reliably regulates the temperature of the inhomogeneous system without discriminating between the slow and fast degrees of freedom while the Berendsen thermostat yields temperature gradients between the fast and slow degrees of freedom.

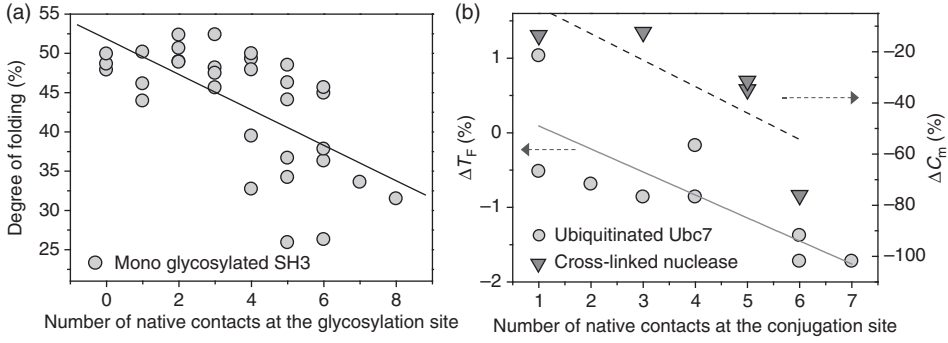

Figure 3 The linkage between the position of the conjugation sites and the effect introduced by the conjugation. The glycosylation (a) and ubiquitination (b) sites are characterized by the number of native contacts the modification site is involved in. Both glycosylation and ubiquitination will show destabilization if the modification is made at a more structured position. Experimentally, it was shown that cross-linked dimers will be destabilized compared to the isolated monomers if the cross-linking is made through a structured residue (b, triangles).

The thermostability effect at each of the 35 selected glycosylation sites of SH3 is depicted in Figure 3a, and illustrates that glycosylation sites located on loops (less structured positions) are more effective in enhancing protein stability than other sites that are more structured.

Since the structures of only a small fraction of natural glycoproteins have been fully resolved by either X-ray crystallography or NMR, statistical analysis of the structural features of favored glycosylation sites is limited. Yet, several structural

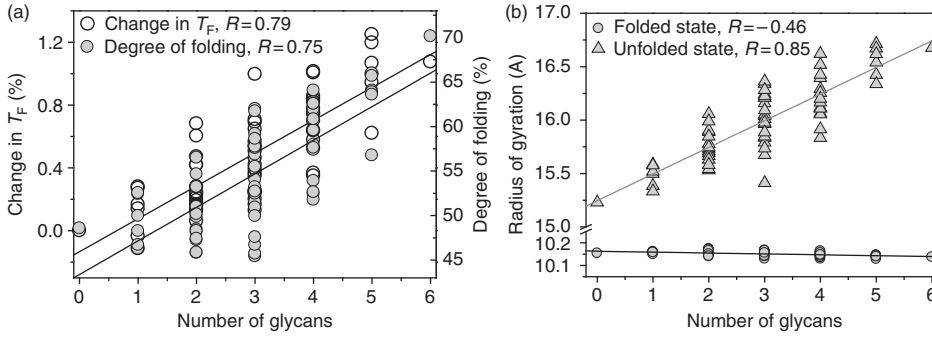

Figure 4 Effect of degree of glycosylation on protein biophysics. (a) The effect of degree of glycosylation on thermal stability as measured either by the change in T_F relative to the unmodified protein ($\Delta T_F = (T_F^{\text{glycosylated}} - T_F^{\text{unglycosylated}})/T_F^{\text{unglycosylated}}$) or by the degree of folding (calculated at the temperature at which the folded state of the unmodified protein is 50% populated). (b) The effect of degree of glycosylation on the size of folded and unfolded conformations.

analyses of glycoproteins that provide some insight regarding the tendency of the potential site containing the consensus sequon of N-glycosylation to accept glycans [18] found that while occupied N-glycosylation sites can occur on all forms of secondary structure, turns and bends are favored. 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.

Since many glycoproteins contain several glycans, it is of great interest to understand how the biophysical characteristics of glycoproteins are affected as a function of the number of the covalently attached oligosaccharides. Specifically, we asked whether there is a cooperative effect between the various attached glycans. To address this question, six positions on SH3, which stabilized the protein, were selected from the 35 sites that were studied. For these six selected glycosylation sites, we designed all possible glycosylated variants using the dodecasaccharide $Man_9GlcNAc_2$. 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 folding temperature (T_F) at which the protein has a stability of zero (i.e., $\Delta G^{\text{Glyco}} = \Delta G^{\text{WT}} \sim 0$)) is observed as the degree of glycosylation increases (Figure 4). 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 [39,40] and subtilisin Carlsberg [34,35] using either the disaccharide lactose or dextran (Glc(α1–6)n, 10 kDa oligosaccharide), with the increase in their melting

temperature depending on the number of glycans attached. The melting temperatures of α-chymotrypsin and of subtilisin Carlsberg increased by 1 and 2°C per added glycan, respectively. Accordingly, not only do the experiments and simulations share a common stabilization trend as a function of the degree of glycosylation, but they also quantitatively predict similar magnitudes of stabilization.

Comprehensive thermodynamic analyses of the simulations demonstrate that changes in the unfolded state cause the thermal stabilization [34]. This observation is in accord with the idea that the unfolded state is not just a random coil but, rather, retains some residual structures, and it was observed that the conjugated glycans interfered with the formation of these structures in the unfolded state. This interference destabilized the unfolded state, shifted the thermodynamic equilibrium toward the folded state, and resulted in an overall thermodynamic stabilization.

3.2 Folding of proteins with flexible tails

To examine the effect of flexible tails on the stability of the protein we attached tails of various lengths to various positions of the SH3 domain, and simulated the folding/unfolding of each variant to decipher the effect of the tail's length on the stability and kinetics of the protein. We found that a short tail of few beads stabilized the protein and the stabilization was increased with the length of the tail (Figure 5). However, longer tails destabilize the protein and reduce the T_F even below the T_F of the unmodified proteins. It seems that the first few beads of the tail are responsible for the stabilization, and a question arises, why do longer tails destabilize the protein? Figure 4b presents the change in the R_g of the protein

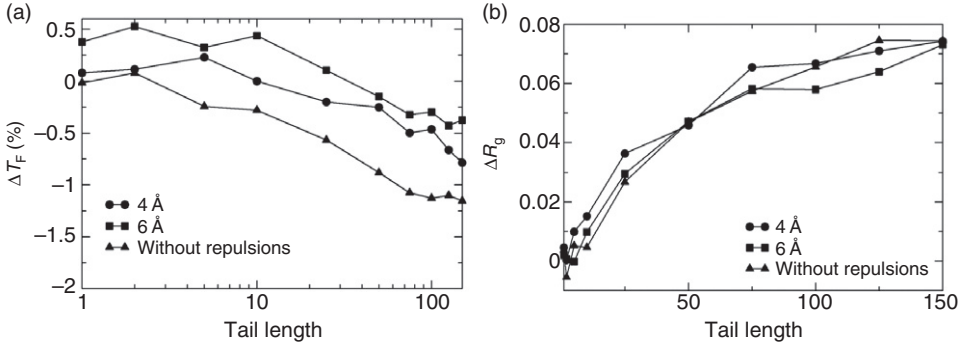

Figure 5 Folding characteristics of tailed proteins. The effects of the length of the attached flexible tail on the protein's thermostability (a) and the protein's radius of gyration in the unfolded state (the tails were attached to an SH3 domain at residue 36). The stability and radius of gyration changes are indicated by $\Delta T_F = (T_F^{\text{with tail}} - T_F^{\text{unmodified}})/T_F^{\text{unmodified}}$ and $\Delta R_g = (R_g^{\text{with tail}} - R_g^{\text{unmodified}})/R_g^{\text{unmodified}}$, respectively. Tails with three different kinds of repulsive interactions with the protein were studied: repulsion distance of 4 or 6 Å between the tail and the protein, as well as cases in which the tail had no repulsive interactions with the SH3 protein.

with respect to that of the wild-type SH3 domain. The R_g of the protein was not altered by the tail in the folded state (not shown); it remained constant and similar to that of the wild-type SH3 domain. In the unfolded state, however, the R_g of the protein was increased with the length of the tail. The inert tail, which was not in a specific interaction with the protein, interfered with the structure of the unfolded state. This observation is in accord with the increased enthalpy and entropy of the conjugated variants in the unfolded state.

To understand the effect of the repulsive interaction between the tail and the protein on thermodynamic stability, we repeated the simulations while canceling these repulsive interactions. As a result, we obtained two chains, a protein and an entropic chain, that could penetrate each other's spaces. Shutting down the repulsion between the tail and the protein reduced the stability of the protein (Figure 5a). Increasing the repulsion distance to 6Å resulted in enhanced thermal stability. These results imply that the protein–tail repulsive interactions are responsible for the alteration in thermodynamic stability. Interestingly, the R_g of the protein during the unfolded state increased with the length of the tail even when the repulsions where shut down (Figure 5b).

The entropy of the protein is affected by two opposing factors. First, the tail is very flexible and can disrupt the structure of the unfolded state. This is because the tail increases the R_g of the protein and so increases its enthalpy. This necessarily increases the entropy of the protein, because when the residual structure is reduced, more conformations are available and the entropy increases. On the other hand, one may assume that the tail confines the available space of the protein because it reduces the dynamics of the unfolded chain. The repulsive interactions between the protein and the tail restrict the expansion of the protein and hence reduce its entropy. It is evident that the number of repulsive interactions between the tail and the protein levels off when the tail contains 25 beads and longer tails do not contribute to additional repulsive interactions. As a result of these two opposing factors, the entropy increases as the tail gets longer, but when the confining effect of the protein reaches its saturation level, the remaining effect becomes dominant and the entropy increases more rapidly than the enthalpy. Then, the free energy of the unfolded state becomes lower and, as a result, the protein is destabilized.

3.3 Folding of ubiquitinated proteins

Recently, we studied the thermodynamic effects of attaching an ubiquitin moiety to a protein, and suggested that these effects may facilitate the cellular process that this specific signal controls [28]. One of the processes ubiquitination mediates is protein degradation: a highly regulated process in which proteins are first recognized by specific cellular machinery, ubiquitinated by a specific ubiquitin polymer (Figure 1e), and then delivered to the proteasome, where they first undergo unfolding and later are degraded into small fragments. We have speculated that, in addition to its recognition role, the ubiquitin attachment may enhance the degradation process by thermally destabilizing the protein. To address this question, we selected the enzyme Ubc7 (Figure 6a), which is ubiquitinated for degradation by a specific ubiquitin polymer (Lys48-linked

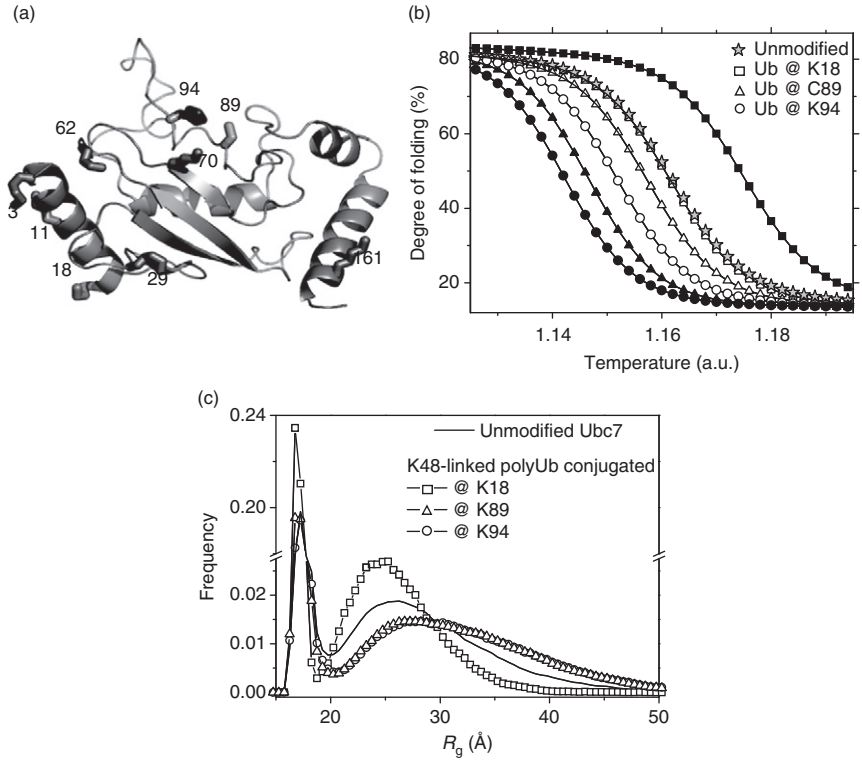

Figure 6 Thermostability of ubiquitinated proteins. The effect of conjugating various ubiquitin polymers (e.g., a monomeric ubiquitin or Lys48-linked polyubiquitin) at the conserved in vivo sites for degradation of Ubc7 protein (residues 89 and 94) as well as at other lysine residues was investigated (a). The computational melting curve of ubiquitinated Ubc7 at positions 89 and 94 by Lys48-linked polyubiquitin indicates strong destabilization compared to the unmodified Ubc7 while its ubiquitination at Lys18 results in stabilization. To ease the comparison, the melting curve of monoubiquitinated Ubc7 at these positions is shown as well (ubiquitination by monomeric ubiquitin or by Lys-48 linked tetra-ubiquitin are shown by empty and filled symbols, respectively) (b). Distributions of the radii of gyration of the ubiquitinated src-SH3 domain systems in the folded and unfolded state illustrate that the change in thermal stability is correlated with changes in the structure of the unfolded state (c).

polyubiquitin) on two residues only, although many other residues are theoretically available for use. Using native-state simulation models, we studied the thermodynamics of this protein with and without a ubiquitin attached at these two residues as well as on other residues that are not used by the cellular machinery. We used a variety of ubiquitin polymers in our study (for example, the tetrameric ubiquitin polymer that is used to tag proteins for degradation and a monomeric ubiquitin that is used by the cell to mediate other nondegradative processes). We observed a range of ubiquitination effects that varied according to the location of the ubiquitin attachment and the type of the ubiquitin polymer we used [28]. These results varied from overstabilizing the protein to different degrees of

destabilization (Figure 6b). Interestingly, we observed a significant destabilization when attaching the ubiquitin polymer that is responsible for signaling degradation to the specific two residues that are used by the cell for Ubc7 degradation. This suggests that ubiquitin may directly modulate the attached protein's properties in a manner that aids the regulated cellular process.

Why do we observe such diverse thermodynamics when attaching a protein to another protein moiety, such as seen in the ubiquitin case? Clearly, various factors affect the thermal stability of the attached protein. One such factor is the well-studied confinement effect, which reduces the entropy of the unfolded state and thereby stabilizes the overall folding reaction (Figure 6c) [41,42]. In the case of ubiquitination, we have observed overstabilization of the protein due to a confinement effect; however, this effect is relatively minor and rare. Another effect that is largely responsible for the varying degree of destabilization observed in our study arises when two protein moieties move in different directions in the solvent, thereby pulling each other. This pulling results in a distortion of the folded state and in the destruction of residual structures in the unfolded state. The pulling effect leads to overall destabilization, mostly because of the increase in the entropy of the unfolded state, due to the residual structures that are unwound near the ubiquitination attachment. We have demonstrated that the degree of destabilization becomes greater when the ubiquitin moiety is attached to a more structured region (Figure 3b). Regions that are structured cannot easily accommodate the attachment of the ubiquitin moiety and its independent movements, and therefore are prone to disruption of the folded state, and to a decrease in the residual structure of the unfolded state near the ubiquitination site. This is evident in a strong correlation observed between the degree of structure in the region and the thermodynamic outcome of attaching an ubiquitin to this region. Our observations of these correlations are augmented by experimental studies in which a nuclease protein was cross-linked in vitro, thus forming a dimer from two monomers using introduced cysteine residues [43,44]. In these studies, different dimers were formed by using different linkage locations, and similarly to our observations, a degree of destabilization was observed. This degree of destabilization correlates well with the density of the structure near the modification site (Figure 3b). Therefore, from these two different systems—ubiquitination of Ubc7 and nuclease cross-linking—we can conclude that the covalent attachment of a protein to another protein may lead to a significant change in the thermal stability of the conjugated protein, and that the thermodynamic outcome is largely dependent on the properties of the modification site. These effects can be used by the cell to facilitate important processes, such as mediating degradation, as in the case of ubiquitination, and can be exploited by experimentalists to alter the properties of the studied system.

3.4 Folding of multidomain proteins

Multidomain proteins are widespread in genomes. The tethering of domains may play a biophysical role in addition to enriching functional diversity. To explore the underlying biophysical principles of the complex folding processes of

multidomain proteins and, in particular, to quantify the cross-talk between the domains, a reduced coarse-grained model based on the native topology was used. The method applied involved a comparison between a two-domain conjugated construct and its isolated domain components. We will concentrate on the FNfn9 domain and its natural conjugated neighbor FNfn10 (the ninth and tenth fnIII domains of Fibronectin, PDB code 1fnf). Experiments have shown that FNfn9, which appeared to be unstable on its own, was significantly stabilized by its conjugated neighbor FNfn10 [45]. However, when FNfn9 was lengthened by two residues, its stability was found to be independent of the presence of FNfn10 [46]. Therefore, it was concluded that the two residues at the C-terminus of FNfn9 and the N-terminus of FNfn10 belong to both domains. Following this domain boundaries definition, the isolated domains and the two-domain conjugated construct were studied and their thermodynamic properties were calculated using the weighted histogram analysis method (WHAM). Figure 7 shows plots of specific heat capacity (C_v) vs. temperature. The peak of these curves corresponds to the transition folding temperature (T_F) at which the protein has zero stability (i.e., $\Delta G=0$). A significant destabilization is demonstrated by the tethering of FNfn9 to FNfn10 i.e., the T_F of the FNfn9 tethered variant is smaller than that of isolated FNfn9. Moreover, if one does not include the interfacial contacts between the two adjacent domains, the decrease in stability is significantly larger. It seems that, in the framework of our model, the tethering by itself causes considerable thermal destabilization. Additional simulations point to the involvement of the structure and flexibility of the linker region (marked as balls and sticks in Figure 1b). The contacts in the interface between domains may compensate for this destabilization; however, in the case of FNfn9-FNfn10 construct this was not sufficient. Next to be considered is the effect of the relative stabilities of the domains. In order to account for the immense difference in thermal stability

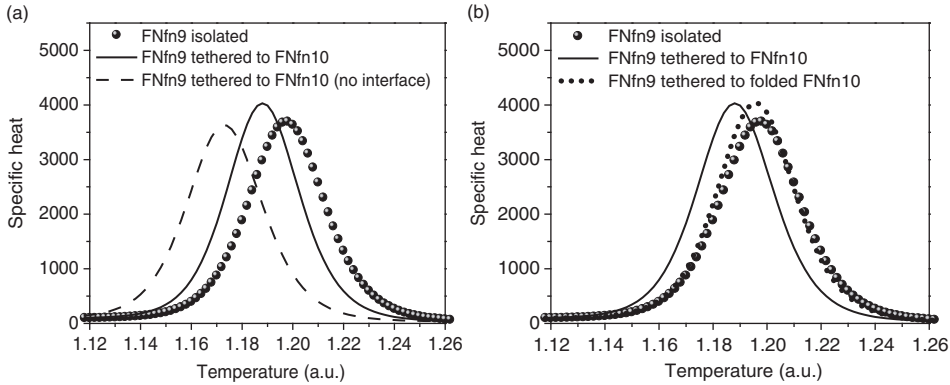

Figure 7 The thermal stability of the multidomain FNfn9–FNfn10 protein. (a) The specific heat curve of the FNfn9 domain in isolation, when it is tethered to FNfn10, and when it is tethered to FNfn10 but no interfacial interactions between the two domains are allowed. (b) The specific heat of the isolated FNfn9 domain is compared to that of an FNfn9 tethered to an infinitely (i.e., permanently) stable FNfn10.

between FNfn9 and FNfn10 [46,47], we designed FNfn10 to be petrified in the folded state. This means that, during folding, FNfn9 "meets" its tethered neighbor when the latter is always folded. This situation, which better distributes the thermal stabilities of the components of this two-domain construct, seems to compensate for the original decrease in stability. Now, the construct in which FNfn10 is folded (Figure 7) shows very similar stability to that of isolated FNfn9, as was also found experimentally.

4. CONCLUSIONS

Many proteins are composed of several domains. These domains may be in direct contact with each other or linked via a flexible linker. One may ask whether the biophysical characteristics of the domains are modified because of the tethering. We show that the properties of the tethered domains can be significantly affected by conjugation to another domain and these effects depend on the properties of the two domains: their flexibility, relative stability, size, and shape. Accordingly, a multidomain protein should not be viewed as a protein that can be described as "sum of its parts". While the tethering in natural multidomain proteins always takes place via the termini and the protein remains a linear polymer, PTMs often result in branched proteins in which a conjugate is attached to the protein through the side chains of various amino acids. The conjugate can have a polymeric nature. For example, in glycosylation and ubiquitination, polysaccharides or ubiquitin proteins are attached to the protein substrate, respectively. In this article, we showed that glycosylated and ubiquitinated proteins can be either stabilized or destabilized by the conjugation depending on the degree of conjugation, its position on the protein, and the molecular details of the conjugate. We conclude that conjugation can enrich the properties of proteins that are encoded in the genome and that nature may take advantage of this venue to modulate protein biophysics. We show that ubiquitination can induce destabilization and unfolding and thus assists degradation by the proteasome.

REFERENCES

1. Oliveberg, M., Wolynes, P.G. The experimental survey of protein-folding energy landscapes. Q. Rev. Biophys. 2005, 38, 245–88.
2. Onuchic, J.N., Wolynes, P.G. Theory of protein folding. Curr. Opin. Struct. Biol. 2004, 14, 70–5.
3. Narayanan, A., Jacobson, M.P. Computational studies of protein regulation by post-translational phosphorylation. Curr. Opin. Struct. Biol. 2009, 19, 156–63.
4. Hershko, A., Ciechanover, A. The ubiquitin system. Annu. Rev. Biochem. 1998, 67, 425–79.
5. Varshavsky, A. Discovery of cellular regulation by protein degradation. J. Biol. Chem. 2008, 283, 34469–89.
6. Sharon, N., Lis, H. Carbohydrates in cell recognition. Sci. Am. 1993, 268, 82–9.
7. Lis, H., Sharon, N. Lectins: Carbohydrate-specific proteins that mediate cellular recognition. Chem. Rev. 1998, 98, 637–74.
8. Harrison, S. Viral membrane fusion. Nat. Struct. Mol. Biol. 2008, 15, 690–8.
9. Resh, M. Trafficking and signaling by fatty-acylated and prenylated proteins. Nat. Chem. Biol. 2006, 2, 584–90.

10. Helenius, A., Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. Annu. Rev. Biochem. 2004, 73, 1019–49.
11. Lederkremer, G.Z. Glycoprotein folding, quality control and ER-associated degradation. Curr. Opin. Struct. Biol. 2009, 19, 515–23.
12. Molinari, M. *N*-glycan structure dictates extension of protein folding or onset of disposal. Nat. Chem. Biol. 2007, 3, 313–20.
13. Trombetta, E., Parodi, A. Quality control and protein folding in the secretory pathway. Annu. Rev. Cell Dev. Biol. 2003, 19, 649–76.
14. Wang, C., Eufemi, M., Turano, C., Giartosio, A. Influence of the carbohydrate moiety on the stability of glycoproteins. Biochemistry 1996, 35, 7299–307.
15. Hanson, S.R., Culyba, E.K., Hsu, T.L., Wong, C.H., Kelly, J.W., Powers, E.T. The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. Proc. Natl. Acad. Sci. U.S.A. 2009, 106, 3131–6.
16. Wyss, D., Choi, J., Li, J., Knoppers, M., Willis, K., Arulanandam, A., Smolyar, A., Reinherz, E., Wagner, G. Conformation and function of the N-linked glycan in the adhesion domain of human CD2. Science 1995, 269, 1273–8.
17. Lutteke, T. Analysis and validation of carbohydrate three-dimensional structures. Acta Crystallogr. D Biol. Crystallogr. 2009, 65, 156–68.
18. Petrescu, A.J., Milac, A.L., Petrescu, S.M., Dwek, R.A., Wormald, M.R. Statistical analysis of the protein environment of N-glycosylation sites: Implications for occupancy, structure, and folding. Glycobiology 2004, 14, 103–14.
19. Wright, P.E., Dyson, H.J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. J. Mol. Biol. 1999, 293, 321–31.
20. Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., et al. Intrinsically disordered protein. J. Mol. Graph. Model. 2001, 19, 26–59.
21. Crane-Robinson, C., Dragan, A.I., Privalov, P.L. The extended arms of DNA-binding domains: A tale of tails. Trends Biochem. Sci. 2006, 31, 547–52.
22. Vuzman, D., Azia, A., Levy, Y. Searching DNA via a "monkey bar" mechanism: The significance of disordered tails. J. Mol. Biol. 2010, 396, 674–84.
23. Shoemaker, B.A., Portman, J.J., Wolynes, P.G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. Proc. Natl. Acad. Sci. U.S.A. 2000, 97, 8868–73.
24. Levy, Y., Onuchic, J.N., Wolynes, P.G. Fly-casting in protein-DNA binding: Frustration between protein folding and electrostatics facilitates target recognition. J. Am. Chem. Soc. 2007, 129, 738–9.
25. Tóth-Petróczy, A., Simon, I., Fuxreiter, M., Levy, Y. The role of disordered tails in specific DNA binding of homeodomains. J. Am. Chem. Soc. 2009, 131, 15084–5.
26. Finley, D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. Annu. Rev. Biochem. 2009, 78, 477–513.
27. Hochstrasser, M., Deng, M., Kusmierczyk, A.R., Li, X., Kreft, S.G., Ravid, T., Funakoshi, M., Kunjappu, M., Xie, Y. Molecular genetics of the ubiquitin-proteasome system: Lessons from yeast. Ernst Schering Found Symp. Proc. 2008, 1, 41–66.
28. Hagai, T., Levy, Y. Ubiquitin not only serves as a tag but also assists degradation by inducing protein unfolding. Proc. Natl. Acad. Sci. U.S.A. 2010, 107, 2001–6.
29. Han, J.H., Batey, S., Nickson, A.A., Teichmann, S.A., Clarke, J. The folding and evolution of multidomain proteins. Nat. Rev. Mol. Cell. Biol. 2007, 8, 319–30.
30. Batey, S., Nickson, A.A., Clarke, J. Studying the folding of multidomain proteins. HFSP J. 2008, 2, 365–77.
31. Murzin, A., Brenner, S., Hubbard, T., Chothia, C. Scop—A structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. 1995, 247, 536–40.
32. Apic, G., Gough, J., Teichmann, S. Domain combinations in archaeal, eubacterial and eukaryotic proteomes. J. Mol. Biol. 2001, 310, 311–25.
33. Itoh, K., Sasai, M. Cooperativity, connectivity, and folding pathways of multidomain proteins. Proc. Natl. Acad. Sci. U.S.A. 2008, 105, 13865–70.
34. 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, 8256–61.

35. Shental-Bechor, D., Levy, Y. Folding of glycoproteins: Toward understanding the biophysics of the glycosylation code. Curr. Opin. Struct. Biol. 2009, 19, 524–33.
36. Clementi, C., Nymeyer, H., Onuchic, J.N. Topological and energetic factors: What determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. J. Mol. Biol. 2000, 298, 937–53.
37. Hills, R.D., Brooks, C.L. Insights from coarse-grained go models for protein folding and dynamics. Int. J. Mol. Sci. 2009, 10, 889–905.
38. Mor, A., Ziv, G., Levy, Y. Simulations of proteins with inhomogeneous degrees of freedom: The effect of thermostats. J. Comput. Chem. 2008, 29, 1992–8.
39. Sola, R.J., Al-Azzam, W., Griebenow, K. Engineering of protein thermodynamic, kinetic, and colloidal stability: Chemical glycosylation with monofunctionally activated glycans. Biotechnol. Bioeng. 2006, 94, 1072–9.
40. Sola, R.J., Rodriguez-Martinez, J.A., Griebenow, K. Modulation of protein biophysical properties by chemical glycosylation: Biochemical insights and biomedical implications. Cell Mol. Life Sci. 2007, 64, 2133–52.
41. Mittal, J., Best, R.B. Thermodynamics and kinetics of protein folding under confinement. Proc. Natl. Acad. Sci. U.S.A. 2008, 105, 20233–8.
42. 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, 11367–72.
43. Kim, Y.H., Stites, W.E. Effects of excluded volume upon protein stability in covalently cross-linked proteins with variable linker lengths. Biochemistry 2008, 47, 8804–14.
44. Byrne, M.P., Stites, W.E. Chemically crosslinked protein dimers: Stability and denaturation effects. Protein Sci. 1995, 4, 2545–58.
45. Spitzfaden, C., Grant, R., Mardon, H., Cambell, I. Module-module interactions in the cell binding region of fibronectin: Stability, flexibility and specificity. J. Mol. Biol. 1997, 265, 565–79.
46. Steward, A., Adhya, S., Clarke, J. Sequence conservation in Ig-like domains: The role of highly conserved proline residues in the fibronectin type III superfamily. J. Mol. Biol. 2002, 318, 935–40.
47. Clarke, J., Cota, E., Fowler, S.B., Hamill, S.J. Folding studies of immunoglobulin-like β-sandwich proteins suggest that they share a common folding pathway. Structure 1999, 7, 1145–53.