Chemical Physics 514 (2018) 95–105

Contents lists available at ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

Entropy-enthalpy compensation in conjugated proteins

Lavi S. Bigman, Yaakov Levy *

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

ARTICLE INFO

Article history:
Available online 12 April 2018

Keywords:
Protein folding
Protein stability
Unfolded state
Multi-domain proteins
Entropy-enthalpy compensation
Coarse-grained simulations

ABSTRACT

Entropy–enthalpy compensation is observed in many reactions, particularly for polymeric biomolecules that often involve large changes in entropy and enthalpy. The imperfect cancelation of entropy and enthalpy dictates many biophysical characteristics, such as protein thermodynamic stability and the free energy barrier for protein folding. In this study, we examine how tethering a conjugate to a protein may affect the thermodynamic stability of the protein. We found that a conjugate mostly affects the unfolded state by eliminating formation of some residual interactions. Consequently, both the enthalpy and the entropy of the unfolded state are affected. We suggest that, because this effect is not localized, the gain in conformational flexibility (*i.e.*, increased entropy) is larger than the loss of some residual interaction (*i.e.*, increased enthalpy). Therefore, the unfolded state of the conjugated protein has a lower free energy than that of the free protein, resulting in thermodynamic destabilization.

1. Introduction

Many proteins, both in prokaryotes and eukaryotes, are composed of several structural modules [1]. In addition to modularity that is encoded in the sequences of proteins, modularity can be introduced by post-translational modifications (PTMs), such as glycosylation [2] and ubiquitination [3]. Various data support the view that, because of their modularity, proteins can evolve using simple domain rearrangement. This way, diverse functional roles are available from a limited number of structural domains [4].

The modular nature of proteins is a fundamental feature used in various laboratory techniques, such as protein design protocols where protein fragments or structural motifs are glued together to construct a larger protein assembly. Moreover, it is common to construct protein chimera, for example, to label proteins with a fluorescent protein such as green fluorescent protein (GFP) to enable visualization of a protein in live cells. However, protein conjugation can also affect the thermodynamic stability of proteins. Several studies suggest that the presence of multiple domains or a conjugate in a protein can lead to stabilization or destabilization of the protein [5]. Accordingly, while some multi-domain proteins can be viewed as more than the sum of their parts, others should be viewed as less than the sum of their parts [5,6]. For example, thermodynamic studies showed that the Engrailed homeodomain undergoes significant destabilization upon its fusion to GFP [7]. Similarly, covalent attachment of a ubiquitin protein to a protein substrate was shown both computationally [8] and experimentally [9,10] to result in a thermodynamic destabilization. In addition, some studies show that multi-domain proteins are more aggregation-prone because of the high effective protein concentration near each domain [11].

The possible effects of a conjugate on the stability of a studied protein are illustrated in the free energy landscape for folding (Fig. 1). In the absence of a conjugate, the thermodynamic stability of a typical protein is measured by the difference between the free energies of the folded and unfolded state ensembles, which, for a thermodynamically stable protein, results in $\Delta G^{wt} = G_F - G_U < 0$ (Fig. 1A and B). Tethering of a conjugate can affect the free energy of either the folded state (Fig. 1A) or the unfolded state ensembles of the protein (Fig. 1B). Protein stabilization upon conjugation (*i.e.*, $\Delta G^{conj} < \Delta G^{wt}$) can be caused by either stabilization of the folded state (Fig. 1B, dashed green line, $\Delta G_F = G_F^{conj} - G_F^{wt} < 0$) or destabilization of the unfolded state (Fig. 1A, dashed red line, $\Delta G_U = G_U^{conj} - G_U^{wt} > 0$). Similarly, protein destabilization (*i.e.*, $\Delta G_{conj} > \Delta G_{wt}$) can be caused by either destabilization of the folded state (Fig. 1B, dashed red line, $\Delta G_F > 0$), or stabilization of the unfolded state (Fig. 1A, dashed green line, $\Delta G_U < 0$).

Although, intuitively, protein tethering is expected to result in thermodynamic stabilization or no major thermodynamic effect, several studies have shown recently that tethering may lead to thermodynamic destabilization [7,12]. Moreover, it has been suggested that this thermodynamic destabilization is related to the stabilization of the unfolded state ensemble of the protein [7]. However, the molecular origin of the stabilization of the unfolded state is still not entirely clear. The question of whether the

* Corresponding author.
E-mail address: Koby.Levy@weizmann.ac.il (Y. Levy).

https://doi.org/10.1016/j.chemphys.2018.04.007
0301-0104/

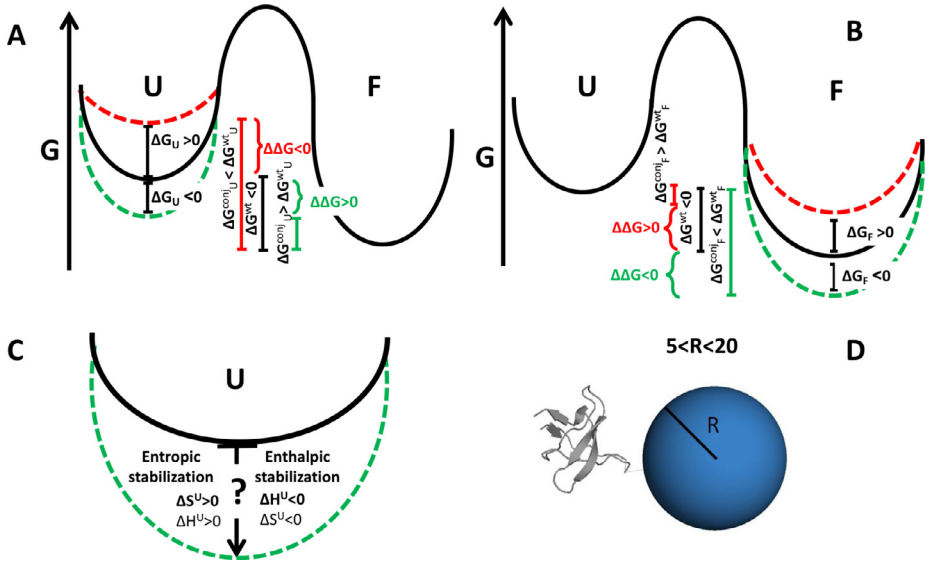

Fig. 1. Illustration of the possible effects of tethering on the free energy landscape of protein folding. (A) Tethering can lead to destabilization of the unfolded state ($\Delta G_U = G_U^{conj} - G_U^{wt} > 0$, dashed red line) and thus stabilization of the protein ($\Delta\Delta G = \Delta G^{conj} - \Delta G^{wt} < 0$, red line), or to stabilization of the unfolded state ($\Delta G_U < 0$, dashed green line) and thus destabilization of the protein ($\Delta\Delta G > 0$, green line). (B) Tethering can also lead to destabilization of the folded state ($\Delta G_F > 0$, dashed red line) and thus destabilization of the protein ($\Delta\Delta G > 0$, red line), or to stabilization of the folded state ($\Delta G_F < 0$, dashed green line) and thus stabilization of the protein ($\Delta\Delta G < 0$, green line). (C) Stabilization of the unfolded state can result either from an increase in entropy ($\Delta S_U = S_U^{conj} - S_U^{wt} > 0$, left), which is larger than ΔH_U. Alternatively, the stabilization can originate from a decrease in enthalpy ($\Delta H_U = H_U^{conj} - H_U^{wt} < 0$, right), which is larger than the decrease in entropy. (D) In this study, we explored the effect of tethering by studying the src homology domain (SH3, PDB ID 1SRL) as a model protein. The effect of tethering was studied by attaching a sphere that represents a conjugated moiety with a varied radius, R, in the range of 5 < R < 20 Å. The sphere was tethered to the protein at the C terminus, at residue 14 (loop 1), or at residue 44 (loop 2).

destabilization has an enthalpic (Fig. 1C, right) or entropic (Fig. 1C, left) origin remains open.

Changes in the free energy landscape upon tethering can be attributed to changes in entropy or enthalpy. Very often the changes in entropy and enthalpy are coupled. In many cases, a perturbation that leads to a change in the enthalpy of protein folding is correlated with a similar change in entropy in what is commonly referred to as "entropy–enthalpy compensation" [13,14], so that the overall net effect on the free energy, $\Delta\Delta G^{WT} = \Delta G^{conj} - \Delta G^{WT} \sim 0$. Entropy–enthalpy compensation (EEC) is reported for many chemical reactions, and is often accounted for as a general thermodynamic principle. In the context of protein folding, it was suggested that the physical origin of EEC is the rearrangement of water molecules [13,14]. That is, a perturbation that leads to the formation of more water-mediated hydrogen bonds in the protein will lead to a decrease in enthalpy, but at the same time, the new constraints on the water molecules will lead to a decrease in entropy. Therefore, the overall change in the free energy in such cases will be minute [14,15].

Alternatively, EEC may originate from properties of the protein itself [15]. For example, heating a protein may lead to an increase in enthalpy due to a breakage of contacts, while simultaneously leading to an increase in the configurational entropy of the protein. It is noteworthy that, although EEC was reported for many cases, it is possible that in some of them it is purely an artifact of either experimental errors or data analysis [16]. Despite criticism of the validity of all reported cases of EEC [15,16], it is clear that, in many cases, entropy and enthalpy can have opposing effects on protein stability. The magnitude of this effect may not always lead to complete compensation, but competition between the two should be taken into account when considering a change in protein stability. In the case of protein folding reactions, the major contribution to the entropy and enthalpy is from the unfolded and folded state, respectively. However, because of the saddle difference between entropy and enthalpy in protein folding, the thermodynamic stability can be modulated not only by following intuitive approaches in which the enthalpy of the folded state (e.g., by mutation) or the entropy of the unfolded state (e.g., by cyclization or disulfide bonds) are changed. Indeed, the EEC in protein folding was shown to be affected by a less common approach in which the entropy of the folded state can be modulated [17,18].

The EEC in conjugated proteins will be identical to that of the unmodified protein if there are no cross-talks between the protein and the conjugate. A thermodynamic destabilization due to tethering [5] can be interpreted in terms of a change in the EEC. In this context, if tethering leads to destabilization of a protein via stabilization of the unfolded state, this effect can be dominated by either an increase in configurational entropy or a decrease in effective enthalpy for the unfolded state (Fig. 1C). However, an increase in the configurational entropy of the unfolded state is likely to be coupled with an increase in its effective enthalpy. The argument is true for a decreased effective enthalpy that most likely is coupled to a decreased configurational entropy.

In this study, we used coarse-grained molecular dynamics (MD) simulations to directly characterize the effect of a conjugate on the thermodynamic stability of the unfolded ensemble of a protein. Specifically, we examined whether changes in the stability of the unfolded state of the protein are enthalpic or entropic in origin. Although one may expect that tethering of a conjugate will act as an intrinsic crowder on the protein leading to a decrease in its unfolded state configurational entropy, our results surprisingly suggest that tethering of a conjugate leads to an increase in the configurational entropy of the unfolded state of the protein. The increase in entropy is not perfectly compensated for by an increase in enthalpy, and therefore it leads to overall destabilization of the protein. Our results also suggest that this effect depends on the size and position of the tethered conjugate.

2. Methods

2.1. Coarse-grained model of folding of tethered proteins

We studied the effect of protein conjugation by using the crystal structure of SH3 as a study case (Protein Data Bank (PDB) ID: 1SRL)

to investigate its folding thermodynamics when it is alone or tethered to a conjugate. The protein is modeled using a coarse-grained model in which the backbone atoms of each residue are represented by a single bead at the position of the Cα atom. The force-field applied in our simulations uses a native-topology-based potential [19–22]. The potential in this model rewards conformations that resemble the native fold and ensures a funnel-like energy landscape [23–29] by excluding nonnative interactions. The potential of a particular conformation $V(\Gamma, \Gamma_0)$, where Γ denotes a particular conformation and Γ_0 denotes the native conformation along the coarse-grained simulation trajectory, consists of the following terms:

$$V(\Gamma,\Gamma_0) = \sum_{bonds} K_{bonds}(b_{ij} - b_{ij}^0)^2 + \sum_{angles} K_{angles}(\theta_{ijk} - \theta_{ijk}^0)^2 + \sum_{dihedrals} K_{dihedrals}[1 - \cos(\phi_{ijkl} - \phi_{ijkl}^0) - \cos 3(\phi_{ijkl} - \phi_{ijkl}^0)] + \sum_{i\neq j} K_{contacts}[5(\frac{A_{ij}}{r_{ij}})^{12} - 6(\frac{A_{ij}}{r_{ij}})^{10}] + \sum_{i\neq j} K_{repulsion}(\frac{C_{ij}}{r_{ij}})^{12} \quad (1)$$

where K_{bonds} = 100 kcal mol^{-1} Å^{-2}, K_{angles} = 20 kcal mol^{-1}, and $K_{dihedrals}$, $K_{contacts}$, $K_{repulsion}$ are each valued at 1 kcal mol^{-1}. The term b_{ij} is the distance (in Å) between bonded beads *i-j*, and b_{ij}^0 is the optimal distance (in Å) between bonded beads *i-j*. The term θ_{ijk} is the angle (in radians) between sequentially bonded beads *i-j-k* and θ_{ijk}^0 is the optimal angle between subsequently bonded beads *i-j-k*. The term φ_{ijkl} is the dihedral angle (in radians) between subsequently bonded backbone beads *i-j-k-l* and φ_{ijkl}^0 is the optimal dihedral angle between subsequently bonded backbone beads *i-j-k-l*. The native contact interactions are modeled using the Lennard-Jones potential. A_{ij} is the optimal distance (in Å) between beads *i-j* that are in contact with each other and r_{ij} is the distance (in Å) between beads *i-j* in a given conformation along the trajectory. Optimal values were calculated from the atomic coordinates of the X-ray structure. C_{ij} is the sum of radii for any two beads not forming a native contact; the repulsion radius of the backbone bead was 2.0 Å. Electrostatic interactions [30] between charged residues of the proteins are not included in this study.

The conjugate was represented by a spherical bead whose radius was in the range 5 < *R* < 20 Å and positioned either at the C terminus or in one of the two loops of the SH3 protein (*i.e.*, at position number 14 or 44). The sphere represents a typical conjugate tethered to the protein and can mimic a small moiety such as glycan or a protein domain. The conjugated sphere interacts with the protein via excluded volume interactions only (i.e., the repulsion term). The size of the sphere is introduced via the repulsive forces it has with any protein bead. The bond length between the sphere and the bead it is conjugated to was taken to be *R* + *r*, where R is the radius of the sphere and *r* is the radius of the protein bead (typically ~2 Å).The native-state topology of SH3 was assumed to be the same when tethered to a sphere or alone. Similar structure-based models have been used previously to successfully capture the essential details of folding of multidomain or conjugated proteins [7,8,31–35]. The dynamics of the isolated and tethered systems is simulated with the Langevin equation [19,30,36]: $m_i\dot{V}_i = F_i - \gamma m_i V_i + R_i(t)$, where m_i, V_i and $\dot{V}_i$ are the mass, velocity, and acceleration of the *i*th bead, respectively. $\mathbf{F}_i$ is the force applied on the *i*th bead, and $\mathbf{R}_i$ is a stochastic variable drawn from a Gaussian distribution with zero mean and variance:

$\langle R_i(t)R_j(t+\tau)\rangle = 2m_i\gamma k_B T\delta(\tau)$, where γ was set to 0.01.

Using this setup, the temperature was varied from k_BT = 1.06 to k_BT = 1.3 with increments of 0.01 k_BT. At each temperature, three independent constant temperature simulations were performed for each system studied. These simulations were used to estimate statistical errors. Each simulation included at least 10^7 elementary MD steps and included many folding/unfolding events of the SH3 domain (at least 10 transition in each trajectory). The simulated trajectories yielded distributions of the number of native contacts, *Q*, in SH3 for each temperature, and these distributions were then analyzed using the weighted histogram analysis method (WHAM) [37]. The WHAM analyses provided a complete thermodynamic description of each system that was used to plot the specific heat capacity, C_V, as a function of temperature and the free energy (PMF) as a function of *Q* [38]. A conformation was assigned as folded if its number of native contacts was larger than the number of contacts at the highest point of the transition barrier, Q_{TS}, and as unfolded if otherwise. The folding temperature, T_F, which is defined as the temperature at the peak of the specific heat capacity curve, provides a measure of the relative stability of SH3 in the presence and absence of a tethered conjugate.

2.2. Coarse-grained model of the unfolded state of conjugated proteins

To obtain a detailed thermodynamic characterization of the effect of tethering on the unfolded state and particularly, the balance between entropy and enthalpy of the unfolded state, we simulated the protein at high temperatures (k_BT = 1.25), while ensuring that the average number of contacts was in the range of the unfolded state ensemble of the SH3 domain. At this selected temperature, about 10% of the native protein structure is formed. The configurational entropy of the unfolded state was estimated based on the entropy of dihedral angles. First, we calculated the entropy of each dihedral angle *j* by $S_j = -k_B T\sum_i \log(p_i)$ where p_i is the probability of the dihedral angle to sample a conformation *i*. Then, the entropy of the protein was determined by a summation over the entropy of all dihedral angels *j*: $S_{dihedral} = \sum_{j=1}^{N-3} S_j$.

2.3. Analytical model for the entropy and enthalpy of unfolded state ensembles

To obtain a quantitative measure of the competition between the changes in entropy and enthalpy of the unfolded state of proteins, we used an analytical model that is based on geometrical considerations of the proteins. This model was applied specifically for simulations of the SH3 domain, so that it can be compared to the results of the coarse-grained modeling [39–41]. The entropy of a specific state of the protein, defined by the fraction of contacts in that state, *Q*, is given by the following expression:

$$S_{tot} = NS_0 + S_{bond} + S_{route} \quad (2)$$

where *N* is the number of residues in the protein (in SH3, *N* = 57), S_0 is the entropy of a residue when no contacts are formed and S_{bond} is the entropic cost due to formation of native contacts along the folding pathway. S_{bond} is given by:

$$S_{bond} = S_{MF} - \frac{3}{2}k_B M\delta Q\delta(\log)l)) \quad (3)$$

where S_{MF} is the entropy calculated using a mean filed approximation and the second term in Eq. (3) reflects decrease in entropy due to loop-loop fluctuations. The exact term for S_{MF} is:

$$S_{MF} = -QNS_0 - \frac{3k_B}{2}MQ\frac{\bar{l}\log\bar{l}}{\bar{l}-1} + \frac{3k_B}{2}M\frac{1}{\bar{l}-1}[1+(\bar{l}-1)Q]\log[1+(\bar{l}-1)Q] \quad (4)$$

Here, *Q* is the fraction of contacts in a specific state (0 < *Q* < 1), k_B is the Boltzmann constant, *M* is the total number of contacts

in the protein (in SH3, $M = 137$) and $\bar{l}$ is the mean loop length. The second term in Eq. (3) is given by:

$$M\delta Q\delta(\log(l)) = \sum_{i=1}^{M}(Q_i - Q)(logl_i - \overline{logl}) \quad (5)$$

where Q_i is the probability of formation of a specific contact, and l_i is the loop length of a specific contact. S_{route} is the entropy gained from all different ways to arrange a specific set of MQ contacts in a specific state, and is given by:

$$S_{route} = k_B\lambda(Q)\sum_{i=1}^{M}[-Q_i logQ_i - (1 - Q_i)\log(1 - Q_i)] \quad (6)$$

The function $\lambda(Q)$ accounts for the decrease in entropy due to the connectivity of the polypeptide chain, and is given in details elsewhere [42–44]. The enthalpy of a specific state is governed by the fraction of contacts in that state, and given by:

$$H(Q) = -\sum_{i=1}^{M_i}\varepsilon_i Q_i \quad (7)$$

The exact values of S_0 and ε_i where tailored by Suzuki et al. [44] specifically for native-topology-based simulations of SH3, and therefore we used the same values: $\varepsilon_i = 1.1$ kcal/mol and $S_0 = 2.49$ kcal/mol.

It is noteworthy that the original theory [44] included a correction for the enthalpic gain due to three-body interactions that add cooperativity to the problem. We did not use that correction in our calculations, since we focused on the unfolded state of the protein. Cooperative effects are important mainly when simulating a full folding pathway, or when focusing on the transition state of the protein.

3. Results and discussion

3.1. Description of the model used in this study

To study directly the thermodynamic stability of a protein with a conjugate, we studied the thermodynamics of protein folding of the src homology domain SH3 as a model protein with a spherical conjugate. We represent the conjugate as a sphere with radius R ($5 < R < 20$ Å), linked covalently to the protein at the C terminus or at one of two different loops (residues number 14 or 44) (See *Methods* for details). The conjugated sphere interacts with the protein via excluded volume interactions only. Accordingly, the sphere interacts with the protein only entropically (i.e., excluded volume interactions) and does not contribute to the effective enthalpy of the protein via direct contact formation. The size of the sphere is introduced via the repulsive forces it has with any protein bead (Fig. 1D).

3.2. Tethering leads to entropic destabilization of the protein and expansion of the unfolded state

Using WHAM analysis of Langevin dynamic simulations (see *Methods* for details) of the free SH3 domain or with a tethered conjugate, we found that tethering of a conjugate of size $R = 5$ Å at various positions, decreases the stability of the protein. This destabilization is represented by a decrease in the folding temperature, T_F (Fig. 2A, see *Methods* for details). In addition, the difference in the free energy between the folded and unfolded state, ΔG, represented by the potential of mean force (PMF), increased upon conjugation (Fig. 2B, see *Methods* for details).

The degree of destabilization varied for conjugates of different sizes (*i.e.*, $R = 5$, 10, 15 or 20 Å, Fig. 3A). Interestingly, the maximal destabilization occurred for relatively small conjugates (i.e., $R = 5$ or 10 Å, depending on the tethering position), indicating that destabilization is size dependent. The observation that the destabilization is size dependent, may hint that the origin of the destabilization is entropic.

In order to directly determine if the destabilization is entropic or enthalpic in origin, we decomposed the PMF profile at T_F to entropic ($-T\Delta S$) and enthalpic (ΔH) components. We found that while $-T\Delta S$ increases significantly due to tethering, the changes in ΔH are minor (Fig. 3B shows representative results for a conjugate at loop 1). In addition, it is noteworthy that the size dependence of the changes in entropy is in concert with the trends we observed for the T_F (i.e., a larger change for conjugates of size $R = 5$ or 10 Å and a smaller effect for larger conjugates).

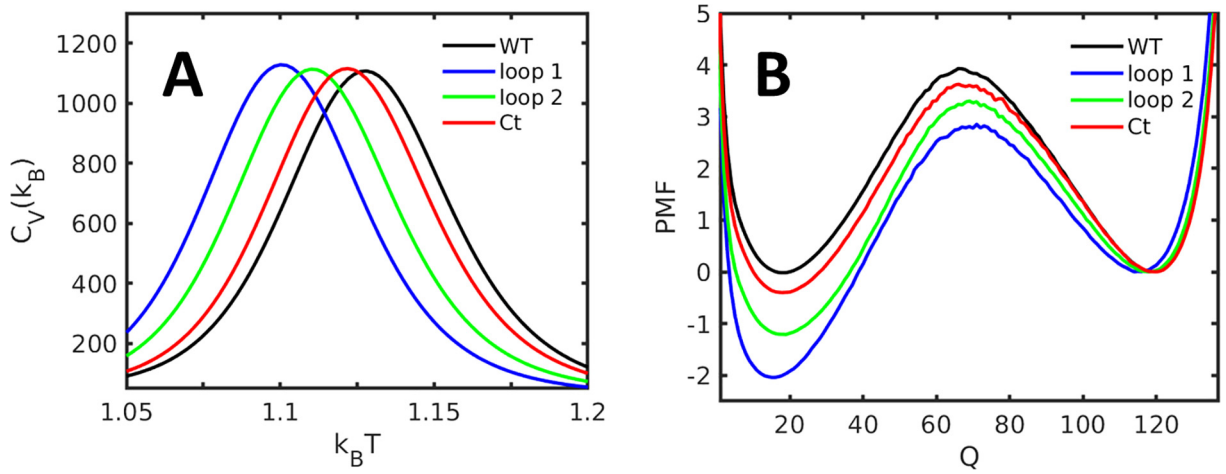

Fig. 2. Conjugation leads to thermodynamic destabilization of the protein. (A) The specific heat capacity versus temperature is shown for the unmodified protein (wild type, WT) and for the protein with a conjugate of size $R = 5$ Å at loop 1, loop 2 and the C-terminus (color code indicated on the figure) studied using native topology-based models. The peaks of the specific heat correspond to the transition folding temperature (T_F) at which the folded and unfolded state populations of the protein are roughly the same (*i.e.*, $\Delta G \sim 0$). The values of the specific heat capacity were calculated using $C_V = (\langle E^2\rangle - \langle E\rangle^2)/k_BT^2$ where $1.06 < k_BT < 1.3$ and the energy, E, is taken to be temperature-independent. Tethering leads to a decrease in the stability of the protein. (B) The free energy profile, represented by the potential of mean force (PMF), of the protein versus the number of contacts (Q) for the protein without and with a conjugate of size R = 5 Å at different positions. The PMF shown are for the T_F of the protein without a conjugate. The minima at low and high Q correspond to the unfolded and folded state ensembles, respectively, and the barrier that separates them is the free energy barrier for folding.

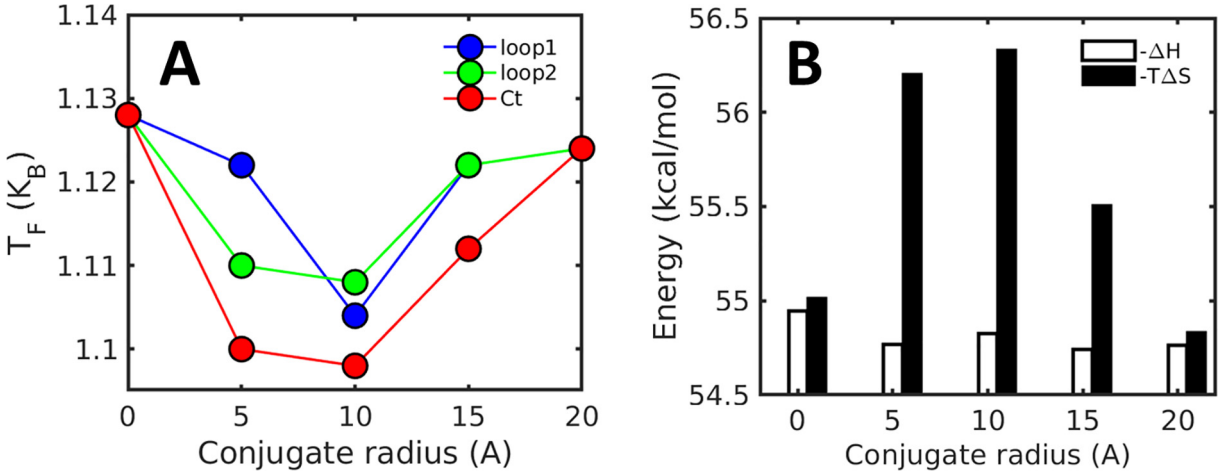

Fig. 3. Protein destabilization upon conjugation is size dependent and entropic in origin. (A) The size dependence of the transition folding temperature (T_F), determined as in Fig. 2, is shown for systems with a conjugate at loop 1, loop 2 and the C-terminus (color code indicated in the figure). The degree of destabilization is large for systems with a conjugate of size R = 5 or 10 Å, and decreases as the conjugate size increases. (B) Decomposition of the difference in free energy between the folded and unfolded state (ΔG) into entropic ($T\Delta S$) and enthalpic (ΔH) component, reveals that conjugation leads to a significant increase in $-T\Delta S$, and to a minor change in $-\Delta H$. The thermodynamic parameters of all systems were calculated from the PMF at the T_F of the protein without a conjugate. The size dependence of the changes in entropy are in concert with the trend observed in panel A for thermodynamic destabilization. The results shown here are for a protein with a conjugate at loop 1 and similar results were calculated for the other tethering positions (data not shown).

In light of the changes in the T_F and ΔG, we conclude that the conjugate leads to a thermodynamic destabilization of the protein, as reported before [5,6,10]. In addition, based on the large changes in $-T\Delta S$, we suggest that the destabilization is mostly entropic in origin. We then sought to determine whether this destabilization is caused by a change in the free energy of the unfolded or folded state. Analysis of the distribution of the radii of gyration (P(Rg)) of the protein reveals that P(Rg) of the folded state ($P(Rg_F)$) is not affected by a 5 Å conjugate, where the peak of the distribution and the mean Rg_F ($\langle Rg_F \rangle$) are ~10 Å with and without a conjugate (Fig. 4A, left y axis). By contrast, tethering leads to broadening of P (Rg_U) and to a shift of the peak in the unfolded state from ~14 Å to ~15 Å. Moreover, $\langle Rg_U \rangle$ increases from 15.7 Å to 16.2 Å (Fig. 4A, right y axis). Similar analysis of $\langle Rg_U \rangle$ and $\langle Rg_F \rangle$ for all the systems included in this study, shows that $\langle Rg_F \rangle$ is either not affected by conjugation (for conjugate at C terminus, Fig. 4B) or slightly

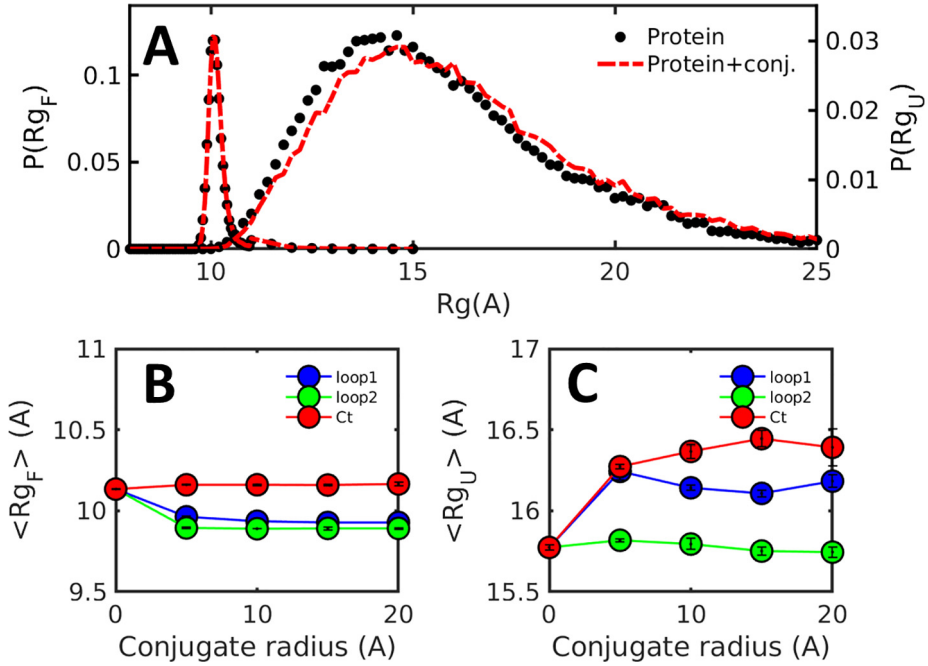

Fig. 4. Tethering leads to expansion of the unfolded state of the protein. (A) Left y axis: The probability distribution of the radius of gyration of the folded state, $P(Rg_F)$, of the protein alone (black circles) and with a 5 Å conjugate at the C-terminus (dashed red line). Right y axis: The probability distribution of the radius of gyration of the unfolded state, $P(Rg_U)$, of the protein alone (black circles) and with a conjugate (dashed red line). The Rg values were determined at the transition folding temperature of the protein without a conjugate. The size of the sphere was not included in the Rg calculations. The shift of the peak in the unfolded state toward higher Rg values in the case of a protein with a conjugate reflects the expanded conformations due to tethering. (B) The mean radius of gyration of the folded state ($\langle Rg_F \rangle$) is shown for all the studied systems. (C) The mean radius of gyration of the unfolded state ($\langle Rg_U \rangle$) is shown for the studied systems. The folded and unfolded states in panels B and C were determined as in panel A and the error bars are standard deviations from three independent simulations.

decreases (i.e., conjugate at loop 1 and 2, Fig. 4B). By contrast, <Rg_U> of most systems (conjugate at C terminus or loop 1) increases significantly (Fig. 4C). These results indicate that there is a change in the dimensions of the unfolded state, which suggest a change in the biophysical characteristics of the unfolded state.

3.3. Tethering leads to unwinding of contacts in the unfolded state

To decipher the effect tethering has on the biophysical properties of the unfolded state of the protein, we calculated the probability of contact formation, Q_i, for each *ith* contact in the unfolded state of the protein with and without a conjugate. The value of Q_i was calculated by $Q_i = \frac{N_{cont}}{N_{total}}$, where N_{cont} is the number of conformations in the unfolded state in which the *ith* contact is formed, and N_{total} is the total number of conformations in the sampled ensemble of the unfolded state. The values of Q_i can be plotted in a form of a contact map in which the intensity of each pixel represents the value of Q_i corresponding to an interaction between the two residues that form contact *i*. Fig. 5 shows the values of $\Delta Q_i = Q_i^{WT} - Q_i^{conj}$ where Q_i^{WT} and Q_i^{conj} are the values of Q_i of the protein without and with a conjugate, respectively. Hence, positive values of ΔQ_i reflect lower probability of contact formation after tethering. Tethering of a 5 Å conjugate at loop *1* leads to an increase in ΔQ_i across the whole contact map, which indicates that there is global unwinding of residual contacts in the unfolded state of the protein (Fig. 5A). Still, it is noteworthy that the degree of unwinding is not uniform for all contacts. Tethering of a 20 Å conjugate at the same position, maintains the signature of unwinding for most contacts (Fig. 5B). However, unwinding in the top left part of the contact map, corresponding to the interactions of residues ~15–25 with residues ~5–10, disappears. These contacts are in proximity to the tethering point. We therefore suggest that tethering leads to a global unwinding of residual contacts in the unfolded state ensemble, but that unwinding decreases locally upon tethering of larger conjugates. Unwinding of contacts is expected to directly affect the enthalpy of the unfolded state and consequently its entropy. A change in the thermodynamic stability of the conjugated proteins indicate that these two effects do not perfectly cancel out each other.

3.4. Effect of tethering on conformational dynamics of the protein

Our observations that tethering leads to entropy driven destabilization of the protein and that the dimensions of the unfolded state change significantly upon tethering may indicate that there is a major change in the entropy of the unfolded state due to tethering. In attempt to understand the effect tethering has on the entropy of the unfolded state, we sought to examine the effect tethering will have on the protein dynamics, which is represented here by the distribution of pairwise distances between all the protein residues in a specific ensemble. The dynamics of each pairwise distance is estimated by the width of its distribution, σ_{ij}. For clarity, we show the difference $\Delta\sigma_{i,j} = \sigma_{i,j}^{conj} - \sigma_{i,j}^{WT}$, where $\sigma_{i,j}^{conj}$ and $\sigma_{i,j}^{WT}$ are the standard deviation of the distributions of the pairwise distances in the protein with and without a conjugate, respectively. Tethering of a 5 Å conjugate at loop 1 has a minor effect on the dynamics of the folded state of the protein, as reflected by values of $\Delta\sigma_{ij} \sim 0$ (Fig. 6, top left panel). By contrast, the same tethering leads to a global increase in the dynamics of the unfolded state of the protein, as reflected by values of $\Delta\sigma_{ij} \sim$ 3–4 Å (Fig. 6, bottom left panel). Interestingly, the dynamics of residues ~15–25 is restricted as reflected by negative values of $\Delta\sigma_{ij}$ (Fig. 6, bottom left panel, dark blue). Since the restricted residues are in proximity to the tethering point of loop 1 (residue 14), we appoint this effect to local crowding, caused by the conjugate.

In order to understand the effect a larger conjugate has on protein dynamics, we calculated the difference $\Delta\sigma_{i,j} = \sigma_{i,j}^{R=20} - \sigma_{i,j}^{R=5}$, where $\sigma_{i,j}^{R=20}$ and $\sigma_{i,j}^{R=5}$ are the standard deviation of the pairwise distances in the protein with a conjugate of size R = 20 Å and $\underline{R}$ = 5 Å, respectively. The dynamics of the folded state remains globally similar to the dynamics of the folded state of the 5 Å conjugate, as reflected in values of $\Delta\sigma_{ij} \sim 0$ (Fig. 6, top right panel). By contrast, tethering of a large conjugate lead to a modest increase in dynamics in most regions of the protein in the unfolded state (Fig. 6, bottom

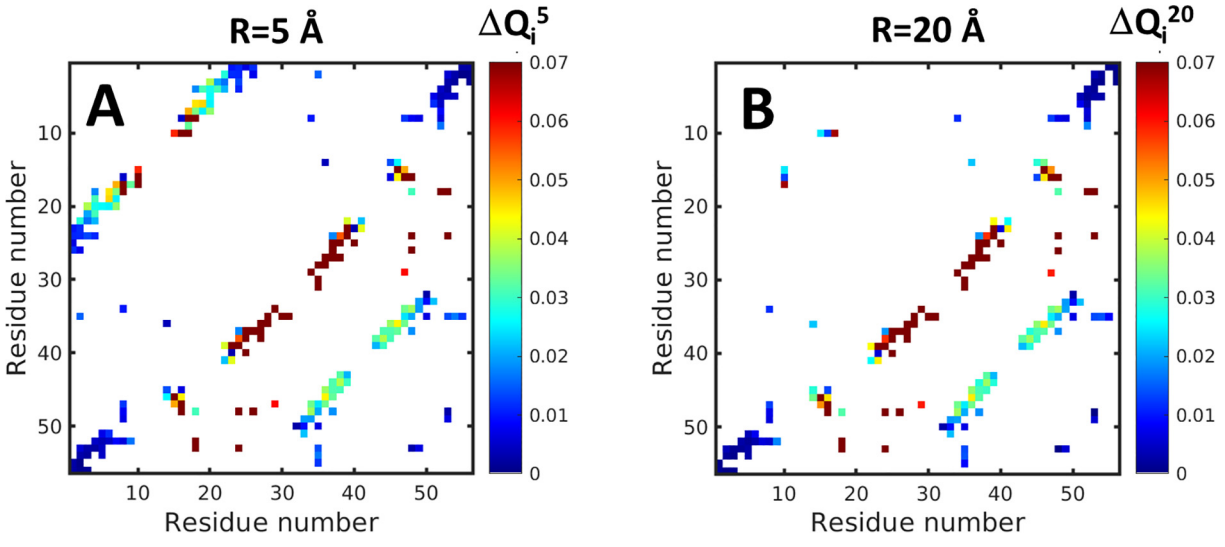

Fig. 5. Tethering leads to unwinding of contacts in the unfolded state of the protein. Contact maps showing the probability of formation of native contacts (Q_i, see *Methods* for details) in the unfolded state of the protein. Each pixel in the maps represents the difference in the probability of the formation of a specific contact $\Delta Q_i^X = Q_i^{WT} - Q_i^X$, where Q_i^{WT} is the probability of contact formation of the protein without a conjugate and Q_i^X is the probability of contact formation of the protein with a conjugate of size X. Hence, a positive value of ΔQ_i indicates a lower probability of contact formation. The unfolded state is defined as in Fig. 4. (A) Contact map of a protein with a conjugate of size R = 5 Å at loop 1. The positive values of ΔQ_i throughout the contact map indicate unwinding of residual contacts in the unfolded state due to conjugation. (B) Contact map of a protein with a conjugate of size R = 20 Å at loop 1. The intensity of ΔQ_i in the upper left part of the map (i.e., the interaction between residues ~15–25 with residues ~5–10) is lower than for the protein with a conjugate of size R = 5 Å (panel A), and is in proximity with the position of tethering in loop 1 (residue 14). Analysis preformed for other tethering positions showed similar results, although the effect reported here was not as significant for tethering at the C-terminus of the protein.

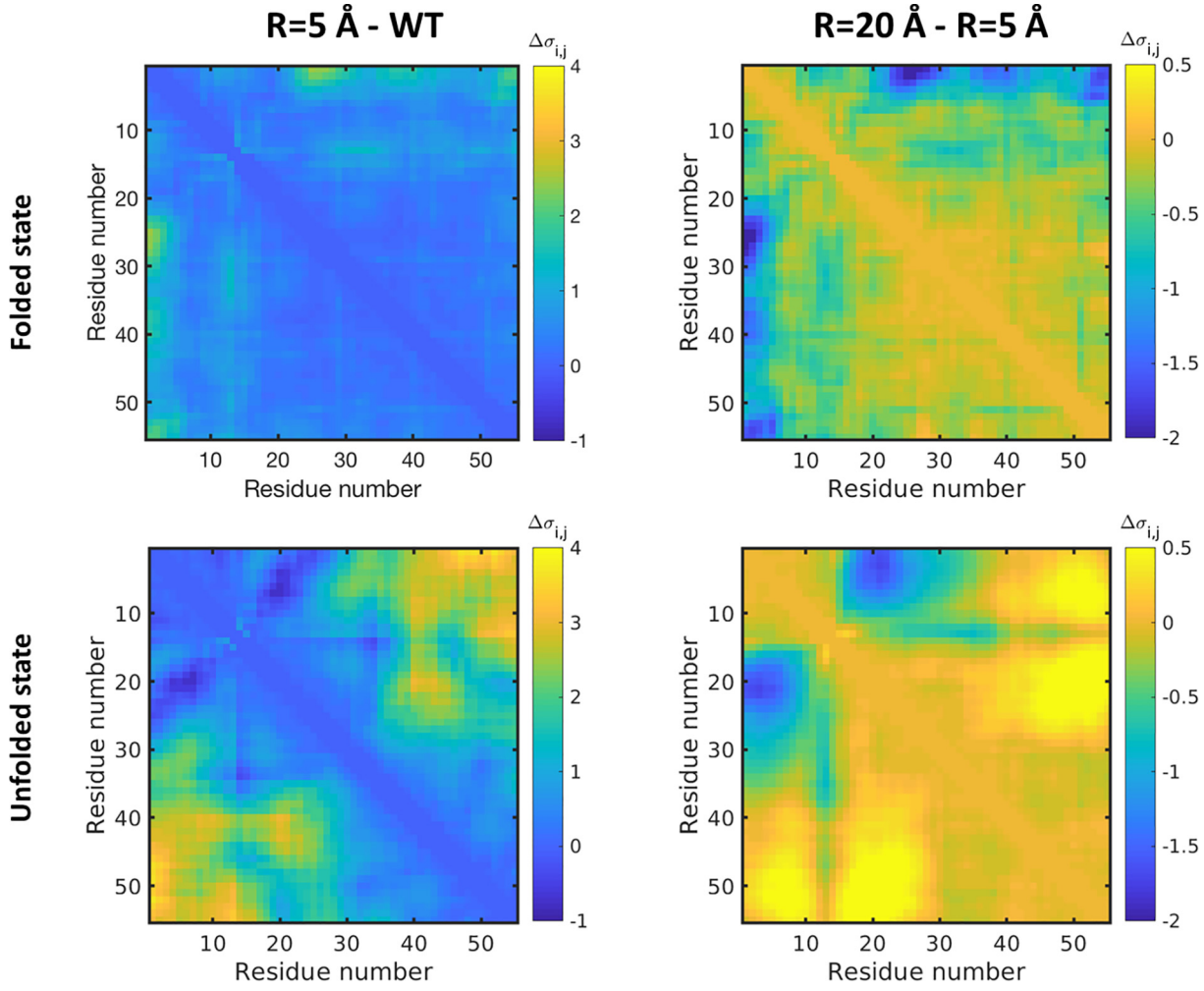

Fig. 6. Effect of tethering on conformational dynamics of the protein. Conformational dynamics are represented by the standard deviation of the distribution of each pairwise distances in the protein ($\sigma_{i,j}$) in a specific ensemble. Top left: Conformational dynamics of the folded state of the protein with a conjugate of size R = 5 Å tethered at loop 1. The colorbar represents the difference $\Delta\sigma_{i,j} = \sigma_{i,j}^{conj} - \sigma_{i,j}^{WT}$, where $\sigma_{i,j}^{conj}$ and $\sigma_{i,j}^{WT}$ are the standard deviation of the pairwise distances in the protein with and without a conjugate, respectively. The dominance of the blue color throughout the standard deviation matrix, indicates that the dynamics of the folded state is not affected much by tethering. Bottom left: Same as in top left, except that the standard deviation matrix is shown for the unfolded state. In the unfolded state a large fraction of the protein gains higher dynamics due to tethering, although a local crowding effect (e.g., restriction of dynamics, indicated by negative values of $\Delta\sigma_{i,j}$) can also be observed. Top right: Conformational dynamics of the folded state of the protein with a conjugate of size R = 20 Å tethered at loop 1. The colorbar represents the difference $\Delta\sigma_{i,j} = \sigma_{i,j}^{R=20} - \sigma_{i,j}^{R=5}$, where $\sigma_{i,j}^{R=20}$ and $\sigma_{i,j}^{R=5}$ are the standard deviation of the pair wise distances in the protein with a conjugate of size R = 20 Å and R = 5 Å, respectively. The dominance of the orange color throughout the matrix, reflects that tethering of a larger conjugate overall does not affect the dynamics of the folded state. Still, a local effect of crowding can be observed. Bottom right: Same as top right, except that here we show the unfolded state. Tethering of a larger conjugate leads to increased dynamics in some regions of the protein, and to a larger extent of crowding in others (i.e., the deviation of distances between residues ~15–25 and residues ~5–10). The area of restricted dynamics is in proximity of the position of tethering in loop 1 (residue 14). Analysis preformed for other tethering positions showed similar results, although the magnitude of the effect of conjugation on dynamics may vary depending on the conjugation site.

right panel), as reflected in positive values of $\Delta\sigma_{ij}$. This increase in dynamics is coupled with a more profound local crowding effect, as reflected by negative values of $\Delta\sigma_{ij}$ (Fig. 6, bottom right panel). Hence, we conclude that a larger conjugate may lead to a global increase in dynamics, coupled with a locally more profound crowding effect.

3.5. Dual effect of tethering on the entropy of the unfolded state

Our results suggest that tethering may involve several effects on the unfolded state of the protein, some of which might have opposite effects on stability. Tethering leads to a global unwinding of contacts and thus an increase in enthalpy. This change in enthalpy might be coupled to change in entropy via an increase in dynamics of the unfolded state. At the same time, at the proximity of the tethering point, the degree of unwinding decreased when larger conjugates were tethered, and a significant crowding effect was observed. It is expected, that a competition between the opposing effects will determine the precise thermodynamic properties of the unfolded state of the protein. Therefore, we sought to calculate directly the effect tethering will have on the entropy of the unfolded state of the protein.

For that purpose, we simulated the protein at high temperatures (k_BT = 1.25), so ensuring that the average number of contacts in the ensemble comprised ~10% of the native protein structure. Then, we calculated the entropy of the unfolded state based on the distributions of all the dihedral angles (see *Methods* for details). Our calculations show that tethering leads to an increase in the entropy of the unfolded state, independent of the size or the position of the conjugate (Fig. 7A). This increase is unexpected. The conjugate is represented solely by its excluded volume, and so it is expected that tethering the conjugate will lead to a decrease in the entropy by acting as an intrinsic crowder that restricts the conformational flexibility of the protein, particularly in the unfolded

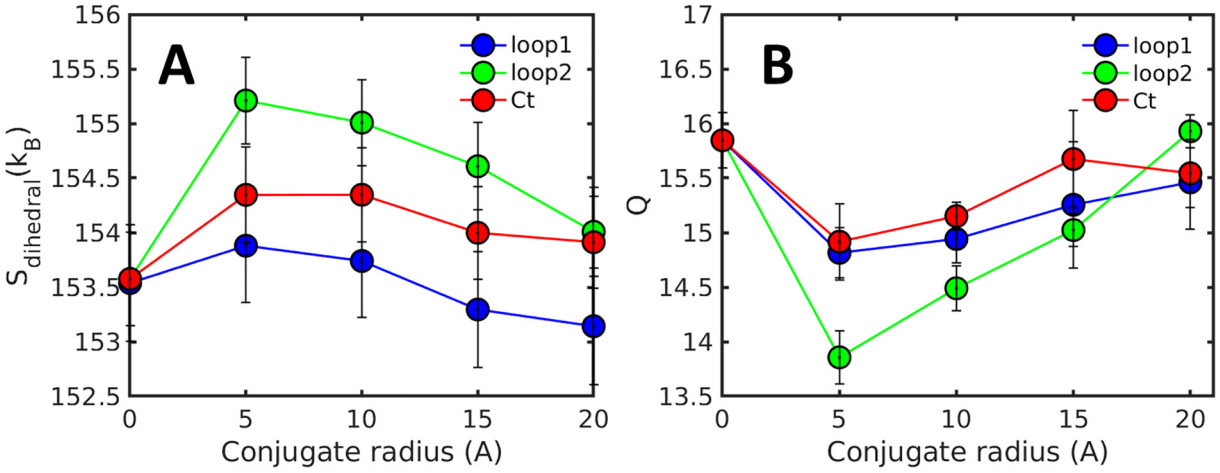

Fig. 7. Tethering increases the configurational entropy of the unfolded state by unwinding residual contacts. (A) The entropy of the unfolded state for conjugates of different sizes (conjugate radius ranges between 5 and 20 Å) and different positions (color code indicated on the figure). The entropy of the untethered protein is shown for $R = 0$. Tethering of a conjugate at all positions examined in this study leads to an increase in the entropy. However, increasing the size of the conjugate leads to a reduced effect of the conjugate on the entropy of the unfolded state. The entropy was calculated based on the distributions of the dihedral angle described in the *Methods*. The values and the error bars are the average and standard deviation from five independent simulations. (B) The average number of native contacts (Q) versus the conjugate size for the protein alone, and for the protein with a conjugate at different positions (color code indicated on the figure). Tethering leads to a loss of residual contacts in the unfolded state. Increasing the size of the conjugate leads to the loss of fewer contacts. Note that the extent of lost contacts correlates with the increase in entropy. The error bars are standard deviations from five independent simulations.

state. However, the increase in entropy is in concert with our observations that tethering leads to increased dynamics of the protein. Since the increase in entropy is correlated with a decrease in the probability of the formation of residual contacts (Fig. 7B), we suggest that the entropic gain is due to unwinding of some residual structure in the unfolded state following breakage of some contacts. This suggestion is strongly supported by our observation that tethering leads to global unwinding of contacts (Fig. 5).

Interestingly, although higher entropy was observed for all modified unfolded states, the magnitude of this increase depends on the radius of the conjugate. While small conjugates lead to a large entropic gain, increasing the size of the conjugate leads to a smaller entropic gain. This trend is similar to what we observed for the size dependent destabilization of the protein (Fig. 3A), and the size dependent increase in $-T\Delta S$ (Fig. 3B). The decrease in entropy as the radius of the tethered sphere increases can be explained by the so-called "crowding effect" [45] whereby conformational freedom is restricted due to the interaction of the protein with the crowding molecules. This explanation is strongly supported by our observation that large conjugates induce a more profound crowding effect at the proximity of the tethering position (Fig. 6). Taken together, our analysis of the entropy of the unfolded state reveals two opposing entropic effects: unwinding residual native contacts leads to an increase in entropy, while concomitantly the excluded volume of the conjugate leads to a decrease in entropy.

3.6. Entropy–enthalpy compensation?

The changes in entropy of the simulated unfolded state are coupled with similar changes in its fractions of formed contacts (Fig. 7), which have an opposite effect on the overall stability of the unfolded state of the protein. So the question of which term leads in the competition between entropy and enthalpy remains elusive.

Our results show that tethering leads to thermodynamic destabilization of a protein (Figs. 2 and 3A) which is entropic in origin (Fig. 3B). In addition, we demonstrate that the folded state of the protein is scarcely affected by tethering (Fig. 4). Consequently, stabilization of the unfolded state is the origin of protein destabilization upon tethering. Our observation that tethering leads to unwinding of contacts and to an increase in the dynamics of the unfolded, coincides with the observation that the entropy of the unfolded state increases upon tethering in a size dependent manner (Fig. 7A). These results can explain the stabilization of the unfolded state and the overall destabilization of the protein. But why does the increase in entropy overcome the increase in enthalpy, with both caused by the unwinding of residual contacts in the unfolded state?

While the enthalpy is determined by the number of contacts in a specific state (Q), the entropy is affected also by the distribution of the contacts; namely the formation probability of each contact. Therefore, a complete breakage of a single contact (*i.e.*, a change in its formation probability from 1 to 0), will have the same effect on enthalpy as a 10% reduction in the formation probability of ten different contacts. By contrast, since entropy is affected also by the distribution of contacts, the latter scenario will result in a larger change in entropy than the former. The contacts maps (Fig. 5) clearly show that tethering leads to a global unwinding in the unfolded state.

To quantitatively compare the changes in entropy and enthalpy, we used an analytical model, developed by Plotkin and coworkers [39–41], that takes into account mainly geometrical features, to calculate the thermodynamic properties of the protein. In Plotkin's model, the entropy is calculated based on the distribution of contacts in a specific state along the folding pathway, and on the distance in sequence between two residues that form a native contact (see *Methods* for details). The enthalpy is calculated based on the sum of contacts in a specific state. In our study, we used a modified version of Plotkin's model [39,40,44,46] that was tailored to fit the results from native-topology based simulations.

Using this model, we estimated the entropy and enthalpy of an ensemble of conformations of a given Q. We first calculated the entropy and enthalpy of SH3 without modifications. To mimic the effect of tethering on the unfolded state, we introduced unwinding of contacts by artificially decreasing the probability of contact formation of a selected set of contacts at varying degrees, by a weight w_i in the range 0.1–1 (e.g., if $w_i = 0.1$ the probability of contact i, Q_i, is reduced by 90%). Accordingly, the unfolded state of the modified protein is characterized by $Q = \sum w_i Q_i / M$. Then, we

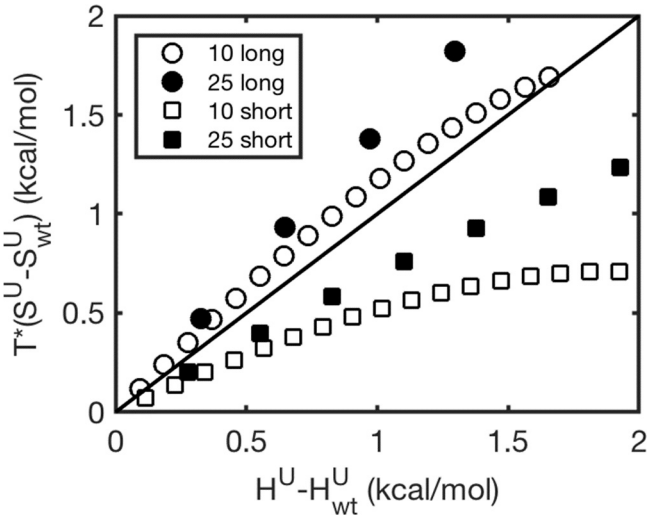

Fig. 8. Entropy-enthalpy compensation of the unfolded state of conjugated proteins. The thermodynamic stability of the unfolded state of SH3 was calculated using an analytical model that takes into account geometrical features of the proteins such as the formation probability of each contact, Q_i and its loop length l (see main text and *Methods* for more details). Q is the mean of Q_i at a given protein conformation during the folding reaction. Based on the analytical model, we examined the effect of unwinding of residual contacts in the unfolded state on the configurational entropy and effective enthalpy of the protein. We focused on an unfolded state at which $Q = 0.2$ (i.e., 20% of the native contacts are formed), and gradually decreased the probability of contact formation Q_i, of a specific set of contacts by a weight w ranging from 1 to 0.1. We tested the effect of unwinding 10 or 25 contacts (white and black symbols, respectively) that are either close in sequence ($5 < l < 15$, squares) or distant in sequence ($15 < l < 35$, circles). The effect different distributions of unwinding had on enthalpy and entropy was calculated as described in detail in the main text and in the *Methods*. Unwinding of short contacts increases the enthalpy more than it contributes to entropy, as reflected by squares all being below the diagonal that represents full entropy-enthalpy compensation. Unwinding of 25 short-range contacts results in a larger increase in entropy than unwinding of 10 contacts. By contrast, unwinding of long-range contacts leads to gain of entropy that is larger than the increase in enthalpy due to unwinding. This is reflected by circles being above the diagonal of full compensation. Note, that also unwinding of 25 long-range contacts leads to a larger entropic gain than unwinding of 10 long-range contacts.

used Plotkin's model to calculate the effect unwinding may have on the entropy and the enthalpy of the specific state ($Q = 0.2$ was selected to represent the unfolded state). To determine the effect unwinding has on different types of contacts, we classified the native contacts according to the sequence distance between the two interacting residues, represented by l (for "loop length").

We found that unwinding short-range contacts ($5 < l < 15$) leads to an increase in enthalpy that is not compensated by entropy (Fig. 8, squares), and to an overall destabilization of the unfolded state. This is reflected in data points below the diagonal in Fig. 8 (the diagonal represents full compensation between entropy and enthalpy). By contrast, unwinding long-range contacts (i.e., $15 < l < 35$) leads to an increase in entropy that is higher than the increase in enthalpy, and to an overall stabilization of the unfolded state (Fig. 8, circles). This is reflected in data points above the diagonal in Fig. 8.

Remarkably, for both long and short-range contacts, unwinding 25 contacts resulted (full symbols) in a larger entropic effect than unwinding 10 contacts (empty symbols). Hence, suggesting that unwinding long-range contacts is more likely to lead to an entropic stabilization of the unfolded state, though the precise definition of long and short contacts is probably case sensitive. More importantly, we conclude that a broad distribution of contact unwinding leads to a larger increase in entropy than when the unwinding is local, that can lead, in some cases, to entropic stabilization of the unfolded state.

To further demonstrate that the larger effect of unwinding on entropy than on enthalpy is linked by the number of contacts whose formation probability is affected, we represent the microstates of a system using a Gaussian distribution of the contact formation probabilities, p. The entropy in this system is estimated by $-k_BT\cdot\sum p\log(p)$, where we take $k_BT = 0.6$ kcal/mol and p is the change in probability of contact formation due to the tethering of the spherical conjugate. The enthalpy of this system is varied artificially from 0 to 5 kcal/mol. Each change in enthalpy can be realized by different situations depending on the change of contact probabilities. Obviously, while the same enthalpic change can be achieved by breaking a single contact or reducing by n% the forma-

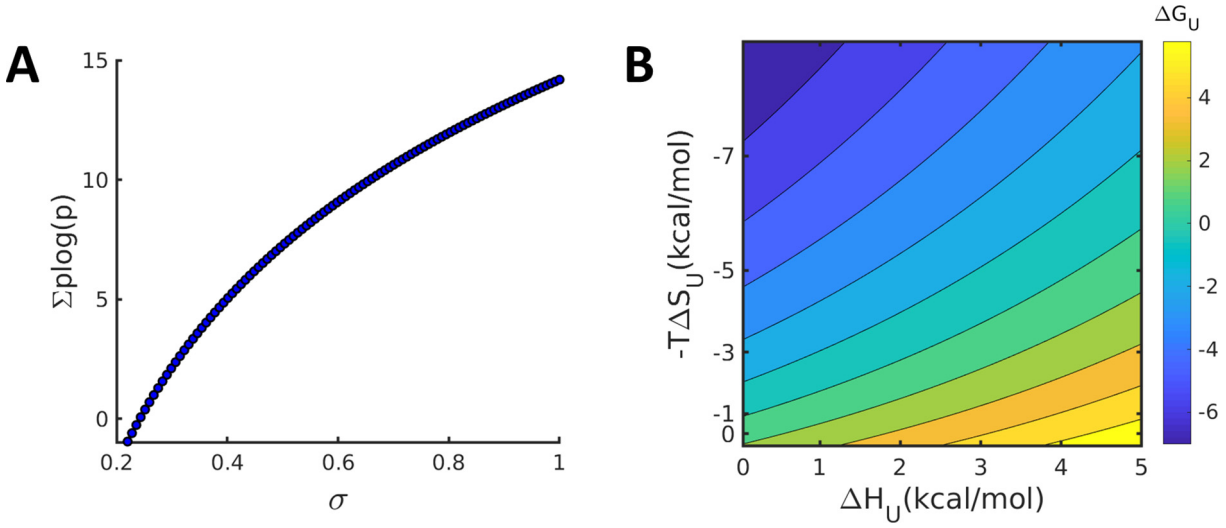

Fig. 9. Distribution of microstates affects the extent of entropy–enthalpy compensation. (A) The change in entropy is proportional to the number of states that are affected by the perturbation. This number is estimated by $-k_BT\sum p\log(p)$, where p is the change in probability of contact formation due to the tethering of the spherical conjugate. This effect is illustrated by estimating the entropy for different probability distributions, which are represented by different standard deviations (σ) of an arbitrary Gaussian distribution. Increasing σ leads to an increase in $-k_BT\sum p\log(p)$. (B) Contour plot of the change in free energy (ΔG_U) for corresponding values of change in enthalpic (ΔH_U) and entropic ($-T\Delta S_U$) components of the free energy term. For a value of $\Delta H_U > 0$, which resembles loss of contacts in the unfolded state, the entropy increases dramatically as the standard deviation of the distribution increases. The competition between entropy and enthalpy can lead to overall destabilization of the unfolded state ($\Delta G_U > 0$, yellow-orange colors on the contour plot) originating from the loss of contacts. Alternatively, if the distribution of the lost contacts is broad enough, the entropy gain will dominate over the contact loss and will result in overall stabilization of the unfolded state ($\Delta G_U < 0$, blue colors on the contour plot). For a specific set of combinations of ΔH_U and $-T\Delta S_U$ there is full entropy–enthalpy compensation (green colors on the contour plot).

tion probability of n different contacts, the later scenario will have a greater effect on the entropy. Indeed, increasing the standard deviation of the Gaussian distribution, σ, leads to an increase in the entropy (Fig. 9A). Fig. 9B shows the different schemes for contributions of the enthalpy and entropy to the free energy for different values of enthalpy loss and entropy gain (the entropy is modulated by the standard deviation, σ, of the distribution of the probabilities of breaking contacts). The competition between the entropy and enthalpy is shown in a contour plot of the change in free energy, ΔG_U. For a given enthalpy, the entropy can vary dramatically, and lead to either stabilization (Fig. 9B, dark blue) or destabilization (Fig. 9B, yellow-orange) of the system.

In the context of the unfolded state of the protein, it seems that although tethering leads to loss of contacts and increased enthalpy ($\Delta H_U > 0$ on the x axis in Fig. 9B), the entropic gain ($-T\Delta S_U < 0$) is more dominant, because of a broad distribution of the lost contacts along the protein. Therefore, we suggest that tethering leads to entropic stabilization of the unfolded state of the protein, as represented by the blue areas in the contour plot of Fig. 9B.

4. Conclusions

Several studies have shown that the tethering of a conjugate, such as GFP or ubiquitin, to a protein leads to destabilization of the protein [5]. This destabilization was suggested to be related to conformational changes in the unfolded state of the protein [7], but to date no direct thermodynamic analysis of the unfolded state of a protein upon tethering was reported. Computational approaches are powerful to quantify the effect of tethering on the unfolded state and consequently on protein biophysics. Here, we used coarse-grained MD simulations to study the effect of a conjugate on the thermodynamic stability of the unfolded state of a protein.

Our results show that tethering leads to an overall destabilization of the protein, which is entropic in origin. In addition, we found that while the unfolded state of the protein samples more expanded conformations in the presence of a conjugate, the folded state remains mostly unaffected. These findings suggest that the destabilization of the protein is caused by stabilization of the unfolded state of the protein. Analysis of the entropy of the unfolded state of the protein revealed that tethering leads to a surprising increase in its entropy. This increase in entropy is correlated with a loss in the residual contacts of the unfolded state. Hence, we suggest that unwinding of residual contacts leads to this increase in entropy which is larger than the expected reduction of entropy associated with the crowding introduced by the large conjugate. Since we observed an overall destabilization of the protein, and no apparent change in the properties of the folded state, it appears that the increase in entropy of the unfolded state is not compensated for by an increase in enthalpy. We further suggest, in light of the analysis of contact maps of the protein, that the reason for this lack of enthalpic compensation is that the distribution of the lost residual contacts is broad and that breaking of long-range contacts has larger contribution to the configurational entropy than on the effective enthalpy. This suggestion is supported by the direct comparison of changes in enthalpy and entropy using an analytical model. We conclude that an unexpected increase in the entropy of the unfolded state leads to stabilization of the unfolded state, and overall destabilization of the protein.

It is suggested that this destabilization is common to all conjugated proteins, yet the fine competition between entropy and enthalpy can be further tuned by other biophysical effects [6], global or local, that may dictate the net balance between them and the overall thermodynamic stability. For example, the existence of soft interactions between the protein and the sphere may affect the magnitude of the effect. The balance between entropy and enthalpy due to the unwinding and crowding effects of the unfolded state may depend on the studied proteins, the dimension of the conjugates and the conjugation sites. We found that small conjugate will destabilize the protein and this effect might be smaller for larger conjugates. It is not trivial to predict the exact effect of large conjugates on protein stability and indeed several studies reported diverse effects of protein tethering to surfaces [47–49].

Acknowledgments

This work was supported by Benoziyo Fund for the Advancement of Science and by the Kimmelman Center for Macromolecular Assemblies. YL is The Morton and Gladys Pickman professional chair in Structural Biology.

References

[1] D. Ekman, A.K. Bjorklund, J. Frey-Skott, A. Elofsson, J. Mol. Biol. 348 (2005) 231–243.
[2] H. Lis, N. Sharon, Eur. J. Biochem. 218 (1993) 1–27.
[3] D. Komander, M. Rape, Annu. Rev. Biochem. 81 (2012) 203–229.
[4] E. Bornberg-Bauer, A.K. Huylmans, T. Sikosek, Curr. Opin. Struc. Biol. 20 (2010) 390–396.
[5] Y. Levy, Biochemistry-Us 56 (2017) 5040–5048.
[6] O. Arviv, Y. Levy, Proteins 80 (2012) 2780–2798.
[7] M. Sokolovski, A. Bhattacherjee, N. Kessler, Y. Levy, A. Horovitz, Biophys. J. 109 (2015) 1157–1162.
[8] T. Hagai, Y. Levy, Proc. Natl. Acad. Sci. USA 107 (2010) 2001–2006.
[9] M.F. Navarro, L. Carmody, O. Romo-Fewell, M.E. Lokensgard, J.J. Love, Biochemistry-Us (2014).
[10] D. Morimoto, E. Walinda, H. Fukada, K. Sugase, M. Shirakawa, Sci. Rep.-Uk 6 (2016).
[11] J.H. Han, S. Batey, A.A. Nickson, S.A. Teichmann, J. Clarke, Nat. Rev. Mol. Cell Biol. 8 (2007) 319–330.
[12] K. Dave, H. Gelman, C.T.H. Thu, D. Guin, M. Gruebele, J. Phys. Chem. B 120 (2016) 2878–2885.
[13] L. Liu, C. Yang, Q.X. Guo, Biophys. Chem. 84 (2000) 239–251.
[14] A.I. Dragan, C.M. Read, C. Crane-Robinson, Eur. Biophys. J. Biophy. 46 (2017) 301–308.
[15] J.D. Chodera, D.L. Mobley, Annu. Rev. Biophys. 42 (42) (2013) 121–142.
[16] K. Sharp, Protein Sci. 10 (2001) 661–667.
[17] S. Dagan, T. Hagai, Y. Gavrilov, R. Kapon, Y. Levy, Z. Reich, Proc. Natl. Acad. Sci. USA 110 (2013) 10628–10633.
[18] Y. Gavrilov, S. Dagan, Y. Levy, Proteins 83 (2015) 2137–2146.
[19] C. Clementi, H. Nymeyer, J.N. Onuchic, J. Mol. Biol. 298 (2000) 937–953.
[20] J.K. Noel, P.C. Whitford, K.Y. Sanbonmatsu, J.N. Onuchic, Nucleic Acids Res. 38 (2010) W657–W661.
[21] V.V.H.G. Rao, S. Gosavi, Plos Comput. Biol. 10 (2014).
[22] J.K. Noel, M. Levi, M. Raghunathan, H. Lammert, R.L. Hayes, J.N. Onuchic, P.C. Whitford, Plos Comput. Biol. 12 (2016).
[23] P.E. Leopold, M. Montal, J.N. Onuchic, Proc. Natl. Acad. Sci. USA 89 (1992) 8721–8725.
[24] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Proteins, Struct. Func. Gene. 21 (1995) 167–195.
[25] J.N. Onuchic, N.D. Socci, Z. Luthey-Schulten, P.G. Wolynes, Fold. Des. 1 (1996) 441–450.
[26] J.N. Onuchic, Z. LutheySchulten, P.G. Wolynes, Annu. Rev. Phys. Chem. 48 (1997) 545–600.
[27] K. Nguyen, P.C. Whitford, Nat. Commun. 7 (2016).
[28] A.B. Oliveira, F.M. Fatore, F.V. Paulovich, O.N. Oliveira, V.B.P. Leite, Plos One 9 (2014).
[29] J. Wang, R.J. Oliveira, X.K. Chu, P.C. Whitford, J. Chahine, W. Han, E.K. Wang, J.N. Onuchic, V.B.P. Leite, P. Natl. Acad. Sci. USA 109 (2012) 15763–15768.
[30] A. Azia, Y. Levy, J. Mol. Biol. 393 (2009) 527–542.
[31] D. Shental-Bechor, Y. Levy, Proc. Natl. Acad. Sci. USA 105 (2008) 8256–8261.
[32] D. Shental-Bechor, Y. Levy, Curr. Opin. Struc. Biol. 19 (2009) 524–533.
[33] D. Shental, O. Arviv, T. Hagai, L. Y., Annu Rep, Comput. Chem. 6 (2010) 263–277.
[34] Y. Gavrilov, T. Hagai, Y. Levy, Protein Sci. 24 (2015) 1580–1592.
[35] Y. Gavrilov, D. Shental-Bechor, H.M. Greenblatt, Y. Levy, J. Phys. Chem. Lett. 6 (2015) 3572–3577.
[36] O. Givaty, Y. Levy, J. Mol. Biol. 385 (2009) 1087–1097.
[37] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, J. Comput. Chem. 13 (1992) 1011–1021.
[38] S.S. Cho, Y. Levy, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 103 (2006) 586–591.
[39] S.S. Plotkin, J.N. Onuchic, Q. Rev. Biophys. 35 (2002) 205–286.
[40] S.S. Plotkin, J.N. Onuchic, Q. Rev. Biophys. 35 (2002) 111–167.

[41] S.S. Plotkin, J.N. Onuchic, P. Natl. Acad. Sci. USA 97 (2000) 6509–6514.
[42] Y. Suzuki, J.K. Noel, J.N. Onuchic, J. Chem. Phys. 134 (2011).
[43] Y. Suzuki, J.K. Noel, J.N. Onuchic, J. Chem. Phys. 128 (2008).
[44] Y. Suzuki, J.N. Onuchic, J. Phys. Chem. B 109 (2005) 16503–16510.
[45] H.X. Zhou, G.N. Rivas, A.P. Minton, Annu. Rev. Biophys. 37 (2008) 375–397.
[46] C. Clementi, S.S. Plotkin, Protein Sci. 13 (2004) 1750–1766.
[47] M. Friedel, A. Baumketner, J.E. Shea, P. Natl. Acad. Sci. USA 103 (2006) 8396–8401.
[48] Z. Zhuang, A.I. Jewett, P. Soto, J.E. Shea, Phys. Biol. 6 (2009) 15004.
[49] T.A. Knotts, N. Rathore, J.J. de Pabloz, Biophys. J. 94 (2008) 4473–4483.