Biochemistry

Perspective

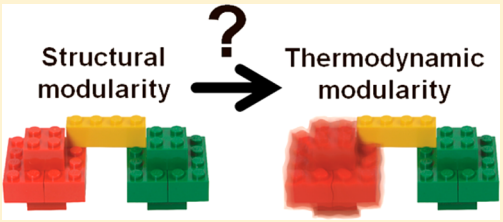
pubs.acs.org/biochemistry

Protein Assembly and Building Blocks: Beyond the Limits of the LEGO Brick Metaphor

Yaakov Levy*

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

ABSTRACT: Proteins, like other biomolecules, have a modular and hierarchical structure. Various building blocks are used to construct proteins of high structural complexity and diverse functionality. In multidomain proteins, for example, domains are fused to each other in different combinations to achieve different functions. Although the LEGO brick metaphor is justified as a means of simplifying the complexity of three-dimensional protein structures, several fundamental properties (such as allostery or the induced-fit mechanism) make deviation from it necessary to respect the plasticity, softness, and cross-talk that are essential to protein function. In this work, we illustrate recently reported protein behavior in multidomain proteins that deviates from the LEGO brick analogy. While earlier studies showed that a protein domain is often unaffected by being fused to another domain or becomes more stable following the formation of a new interface between the tethered domains, destabilization due to tethering has been reported for several systems. We illustrate that tethering may sometimes result in a multidomain protein behaving as "less than the sum of its parts". We survey these cases for which structure additivity does not guarantee thermodynamic additivity. Protein destabilization due to fusion to other domains may be linked in some cases to biological function and should be taken into account when designing large assemblies.



Many biological systems are described as being modular. Nature, like engineering, tends to construct large objects by combining smaller, independent parts or modules. Nature, thus, is opportunistic as it reuses instead of reinventing. The inclusion or exclusion of modules from different combinations enables new complex tasks to be fulfilled by previously approved modules whose reuse expedites the speed with which biological entities can evolve at many levels of cellular complexity.[1] Natural systems use modularity at many levels, ranging from the molecular level to the construction of multicellular organisms.[2]

MODULARITY IN PROTEINS

Proteins are often described as LEGO bricks (sometimes also as puzzle pieces). This analogy refers to the modular nature of proteins, which manifests at several levels. Proteins can be viewed as being constructed from at least four types of building blocks in a hierarchical manner. The smallest building blocks are the 20 amino acids, which are often organized as secondary structure elements (e.g., α helices, β sheets, and loops). Secondary structures can form "motifs" or a "supersecondary structure". The most frequent supersecondary structures are $\beta\beta$ hairpins, and $\beta\alpha\beta$ elements. Motifs assemble into larger subunits of structure called "domains". Although some proteins contain only a single domain, it is often the case that proteins are composed of multiple domains. Domains are the smallest cooperative and autonomous folding unit within a protein structure.

The domains are largely formed by a limited "vocabulary" of recurring supersecondary structural elements, often by repetition of the same element, and increasingly, elements similar in both structure and sequence are being discovered. Classification of protein domains (i.e., folds) also reveals a limited number of component piece self-assemblies. The number of folds, which is defined as the arrangement of secondary structural elements, is estimated by the CATH and by the SCOP to be <1400, and the rate of discovery of new folds appears to be low.[3] Although this number of folds is larger than the original estimate of a few hundred, there is general agreement that the number is bounded and is not larger than 10000 (the exact number may also depend on the classification used).[4] The small number of folds in comparison to the large number of proteins in the proteomes (>90000 proteins) or to the number of high-resolution three-dimensional (3D) structures in the Protein Data Bank (PDB; >127000 at the beginning of 2017) provides clear evidence of the modularity of proteins.

Proteins can alternatively be analyzed as being composed of small fragments. The fragment libraries that are created using nonhomologous proteins are found to be similar in sequence, structure, and sometimes even function, which points to a common evolutionary origin. The fragments, which can be viewed as structural alphabets of varying numbers of letters, were successfully used as building blocks in various computational approaches to analyze, predict, and design protein structures.[5−10] The usage of a small alphabet corresponding to a library of motifs or fragments to represent the complex and heterogeneous space of protein structures also illustrates their

Received: July 15, 2017
Revised: August 15, 2017
Published: August 15, 2017

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048

modular nature and the fact that modern proteins likely arose through the assembly and concatenation of smaller fragments.[11] Accordingly, proteins evolved by fusion and recombination from more ancient peptides and many of the core folds observed today may contain homologous building blocks. The self-assembly of the building blocks is facilitated by the principle of minimal frustration, which is biased toward sufficient energetic complementarity between the structural components.[7,12] A few consecutive building blocks may define an independently folding unit ("foldon").[7,13,14] The hierarchical nature of protein space is reflected in its clustering into a "periodic table".[15] Periodic tables were also drawn for coiled-coil protein structures[16] and for protein complexes.[17]

BUILDING BLOCKS IN SYNTHETIC BIOLOGY

Following the extensive evolutionary usage of peptides and proteins as building blocks for the generation of new proteins, a similar approach is used in synthetic biology. Peptides represent the most favorable building blocks for the design and synthesis of nanostructures. Peptides serve as attractive building blocks for bionanotechnology because of their ease of synthesis, small size, and relative stability, the great diversity of their chemical and physical properties, and their chemical and biological modifiability.[18] Frequently, very short peptides are used to form large structures. Although the peptide building blocks in synthetic proteinaceous nanostructures are not linked covalently as in proteins, the principles governing their self-assembly are expected to be identical. The self-organization properties of peptides have been exploited for the formation of bioinspired nanoassemblies, including nanotubes, nanospheres, nanofibrils, nanotapes, and hydrogels. The molecular self-assembly of peptides (i.e., the bottom-up approach) serves as an alternative paradigm for the *de novo* design of stable functional nanostructures and allows 3D structures with diverse shapes, sizes, and functions to be constructed from self-assembled molecules.[19,20] A bottom-up approach was used to engineer multidomain proteins from simpler domains that display the required properties for exploitation in nanotechnology.[21]

Although peptides naturally self-assemble and form nanostructures in irreversible amyloid fibrils, designing peptides as building blocks for supramolecular architectures is challenging. It is very difficult to rationally design, *a priori*, molecules that predictably self-assemble into and impart functionality to targeted 3D nanostructure materials, especially via non-commutative mechanisms. Many examples of peptides and proteins that are found to self-assemble are discovered serendipitously. The next step is to design molecules based on these rules that can assemble in a predictable manner into a targeted morphology.

BUILDING BLOCKS IN *DE NOVO* PROTEIN DESIGN

The observation that natural proteins evolved in an opportunistic manner by reusing the same building blocks that are arranged into a relatively small set of 3D folds suggests that a large fraction of the sequence space and of the fold space was not sampled by evolution. These non-evolved proteins are variously named "never born proteins", the "dark matter of protein space", or fully *de novo* proteins.[3] In principle, some of these never born proteins could be evaluated by designing them *de novo* using the evolutionary approach, namely, through the assembly of building blocks (e.g., secondary or supersecondary structures). However, attempts to understand the dark matter of protein space via *de novo* protein design show this goal to be very difficult.[3] Although there have been some impressive successes in engineering complex protein structures, most *de novo* structures described so far are mimics of, or variations on, natural protein folds. Currently, there are few pure *de novo* designed proteins [TOP7 and CC-Hept (coiled coil)[3]].

The many efforts made to engineer proteins have yielded designed proteins of various structures, with different folds, and with different oligomeric states, including various $\alpha\beta$ proteins,[6] repeat proteins with different curvatures,[22] the symmetrical β propeller,[23] nanocages,[24] and α-helical barrels, bundles, and cages using coiled coils.[25−27] In many cases, this was achieved by fragment assembly and by applying a set of rules that relate secondary structural patterns to protein tertiary motifs (i.e., to the junctions between adjacent secondary structural elements). This has made possible the design of funnel-shaped protein folding energy landscapes leading into the target folded state.[28] For many years, building new structures using protein building blocks produced results inferior to those achieved using DNA building blocks. Because of the limited set of Watson−Crick base pairings, using DNA building blocks enabled the generation of extremely complicated shapes.[29] Nevertheless, the gap between the protein- and DNA-oriented approaches to structure design has narrowed. Often, successful protein design protocols result in ideal protein structures that are almost identical to those of the design model but are found to be considerably more stable than the natural proteins. Their ideality may stem from a lack of irregularities (such as kinked α helices and β strands) or the incorporation of minimal loops. However, it is still unclear whether the designed ideal proteins are suitable for biological functions. One may note that designed enzymes have not been able to compete with the functional precision of natural proteins.[6]

MODULARITY IN MULTIDOMAIN PROTEINS

Another manifestation of the modularity of protein structure is revealed by multidomain protein organization, in which the building blocks are not small motifs but structural domains. Modularity at the level of protein domains is widespread in both prokaryotes and eukaryotes, with approximately 40−65 and 65−80%, respectively, of their proteins containing multiple domains. Various data support the view that proteins evolve via domainwise rearrangements in which a limited repertoire of domains is used to form the diverse functional landscape of proteins.[30] Multidomain proteins likely evolved through simple linear concatenation of successive domains onto the polypeptide chains, by gene duplication, or through the insertion of one or more continuous sequences into the middle of another by gene fusion and fission. Various proteins are generated by the repetition of peptide segments (e.g., fibrous proteins, coiled-coils proteins, open-ended solenoid structures such as Trp, HEAT, Armadillo, and Ankyrin repeat proteins, or TIM barrel proteins) or even of a full protein domain (e.g., the giant muscle protein, titin, which consists of hundreds of immunoglobulin domains) most likely through internal duplications. The modularity of proteins and their hierarchical arrangements are illustrated *in vivo* by protein labeling (i.e., post-translational modifications) and in the lab by labeling with either polymers [e.g., the conjugation of polyethylene glycol (PEG) during PEGylation or of a polysaccharide during glycosylation] or proteins (e.g., conjugation of a ubiquitin chain during ubiquitination) (see Figure 1).

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048

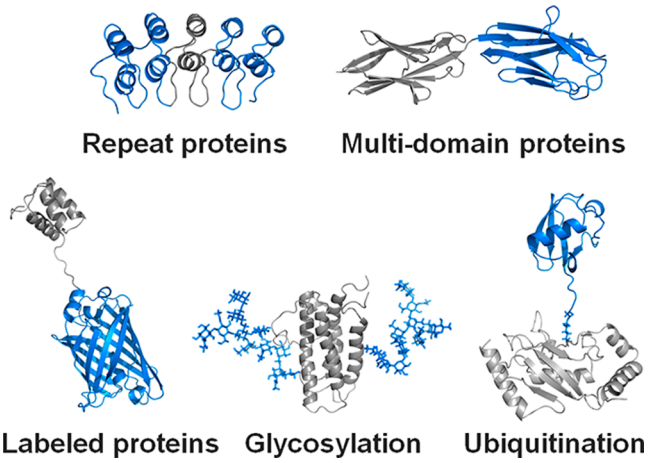


Figure 1. Gallery of conjugated proteins that are formed by different protein building blocks. The protein of interest is colored gray and the conjugation blue. The protein constructs shown are the repeat protein (ankyrin), a multidomain protein (FNfn9−FNfn10), a homeodomain protein labeled with GFP, a glycosylated erythropoietin, and a ubiquitinated Ubc7 protein (conjugated with monomeric ubiquitin at K94).

MULTIDOMAIN PROTEIN FOLDING: DOES STRUCTURAL MODULARITY GUARANTEE THERMODYNAMIC MODULARITY?

The folding of multidomain proteins is inherently more complex than that of single-domain proteins simply because the former are larger and contain several structural domains. The fusion of several domains may result in different compensation schemes for the enthalpy and entropy of folding. To investigate these effects, one must study domains both in isolation and within their multidomain architecture. While efforts in folding research have focused almost exclusively on studying single-domain proteins or small proteins in isolation, in recent years some studies have addressed questions regarding the possible effects of interdomain interactions on folding.

UNCOUPLED FOLDING THERMODYNAMICS IN MULTIDOMAIN PROTEINS: "SUM OF ITS PARTS" BEHAVIOR

Following the perspective that proteins behave like LEGO bricks, one may view each domain as being entirely independent of the other. In this case, the stability of each domain will be unaffected by its neighbor, and the folding and unfolding rate constants will be the same whether the domain is alone or in tandem. In this case, the stability of a two-domain protein, $\Delta G_{Tot}^{tethered}$, is simply the sum of the respective thermodynamic stabilities of the individual domains A and B, namely, ΔG_{A}^{free} and ΔG_{B}^{free}, respectively (Figure 2).

Indeed, for several multidomain proteins, it was shown that the constituent domains have similar thermodynamic and kinetic properties in the tethered and isolated forms. For example, the stability of each immunoglobulin domain in titin and in fibronectin and their folding and unfolding rate constants are unaffected by the neighboring domain and therefore are the same as those of each component of the multidomain protein in isolation (see Table 1). In these cases, the structural and biophysical characteristics of the individual domains are unaffected by the assembly. These tethered domains fold independently, and therefore, in this scenario, the properties of the multidomain protein are the sum of its parts,[31,32] it lacks cross-talk between constituent domains, and the interaction energy of a domain with its native (folded) or unfolded neighbor is zero. In known cases in which domains fold independently, the linkers between the domains are apparently short and flexible and the interfaces are small.

COUPLED THERMODYNAMICS LEADING TO INCREASED STABILITY UPON TETHERING

The folding of large protein constructs might be different from the folding of the individual protein components. An intuitive claim is that tethering may modify the energy landscape of the protein domains. Protein folding is a complex process that involves large ensembles of conformations whose population can, in principle, be affected by the self-assembly process, and therefore, the kinetics might deviate. Similarly, the thermodynamics of the large assembly might be different from that of the individual proteins because of the formation of an interface between the domains. In this case, the folding thermodynamics and kinetics of each domain may be affected by its neighbor, such that the thermodynamics of the assembly may not be the sum of its parts. Accordingly, $\Delta G_{Tot}^{tethered} = \Delta G_{A}^{free} + \Delta G_{B}^{free} + \Delta G_{INT}$, where ΔG_{INT} is the interaction energy between the domains (namely, $\Delta G_{A/B}^{tethered} > \Delta G_{A/B}^{free}$) (Figure 2). A similar thermodynamic scheme was applied to dissect the energetics of modular folding of repeat proteins.[33]

Many multidomain proteins show stronger stabilization because of the formation of a new interface between the neighboring domains. In some cases, the individual domains are very similar in isolation and in the multidomain constructs, such that the higher stability clearly arises from the new interface between the domains, which can be viewed as associating as LEGO bricks. In other cases, the domain may behave differently in its free and tethered forms. For example, some domains in the multidomain architecture share significant interfaces and are attached to each other by either flexible or structured linkers of various lengths[34] without an interface. Some domains are unable to fold on their own.[35] Experimentally studied systems usually show that the folding characteristics of the individual domains depend on whether they fold as part of a multidomain complex or in isolation. These variations can potentially be traced back to differences in cross-talk between tethered domains, which may also be reflected in their multidomain topologies. For example, it was suggested that an extensive domain-linking helix correlates with measurable cooperative folding, and large densely packed interfaces were shown to correlate with the stabilization of the tethered domains.[32]

The observed enhanced thermostability, which is imposed by interfacial interactions that stabilize the native state, was also associated with slower unfolding rates. Interactions in the non-fully folded protein can be exploited to catalyze the folding of an adjacent domain, resulting in more efficient folding, or, on the contrary, may significantly impede folding. A domain may also be affected by an unfolded neighboring domain (for more details, see ref 32). In light of evolutionary pressure, an extensive, densely packed interface confers more stability on tethered domains than a limited interface. See, for example, the increased stability of the VH domain in the scFV protein or of the crystalline-like domain in Protein S (Table 1). However, the energy of the interaction depends on the interface topology and on the specificity of the interactions between the adjacent

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048

Table 1. Multidomain Proteins Exemplifying the Range of Different Effects of Tethering Might Exert on Protein Stability

protein	multidomain construct	T_M^{free} (°C)	$T_M^{tethered}$ (°C)	ΔG^{free} (kcal/mol)	$\Delta G^{tethered}$ (kcal/mol)	ref	remarks
				Increased Stability			
Protein S	two γ-crystallin-like domains (all γΔ-Greek key)	50	65	–	–	47	
spectrin	R16 linked to R17	–	–	6.1	12.2	48	
	R15 linked to R16	–	–	6.8	8.7	49	
scFV	VH tethered to VL	–	–	3.9	7.4	50, 51	
titin	30th Ig domain linked to the 32nd domain	–	–	1.9	3.0	31	
				Decreased Stability			
human utrophin tandem calponin homology	C-terminal domain in comparison with the full length domain	71.5	56.9	10.48	6.49	46	Similar results obtained by Trp fluorescence and circular dichroism spectroscopy
phosphoglycerate kinase	AcGFP1 at the N-terminus	52	43	–	–	45	Similar destabilization was obtained when mCherry was used at the N- or C-terminus
ubiquitin	two covalently linked ubiquitin proteins	88	85	–	–	52	This study includes 12 variants of fused proteins: for most of them, decreased stability is observed
	β1 domain of streptococcal protein G attached at the N-terminus	88.5	80.2	–	–		
human FK506 binding protein (FKBP12)	ubiquitin attached	59.1	43.5	–	–	53	Additional variants of ubiquitinated FKBP12 also show destabilization
	polymer of six ubiquitins attached	59.1	41.8	–	–		
human fatty acid binding protein 4 (FABP4)	ubiquitin attached	63.1	55.8	–	–		Additional variants of ubiquitinated FKBP4 also show destabilization
	polymer of six ubiquitins attached	63.1	54.4	–	–		
Engrailed homeodomain (EnHD)	GFP tethered at the N-terminus	52.0	43.6	–	–	44	Similar results are seen for a linker of three or nine residues. A similar but smaller effect is seen when GFP is tethered at the C-terminus
staphylococcal nuclease K70C	linked dimer with a linker of zero to two residues	48.4	43.4	–	–	54	Similar destabilization was observed for eight additional mutants and for linkers of zero to three residues
				Independent Folding			
titin	28th Ig domain linked to the 27th domain	–	–	3.2	3.1	31	Similar behavior is observed when the 29th domain is fused
	29th domain linked to the 28th domain	–	–	5.1	5.1		Similar behavior is observed when the 30th domain is fused
	30th Ig domain linked to the 29th domain	–	–	5.0	5.0		Similar behavior is observed when the 31st or 32nd domain is fused
	32nd Ig domain linked to the 31st domain	–	–	6.2	6.3		
fibronectin	FNfn9 Ig domain linked to FNfn10	–	–	4.30	4.33	55	

domains. Several multidomain systems surveyed demonstrated a tendency toward largely hydrophobic interdomain interfaces,[36,37] which is expected to influence the compatibility and strength of the interactions.

The size and geometry of interfaces in a multidomain architecture may alter the folding pathways of tethered domains such that they become very different from those of the isolated domains.[38] For example, in the yeast phosphoglycerate kinase (PGK), domain−domain interactions were shown to direct the folding of the N domain along a pathway completely different from that of an isolated N domain.[39] The R16 immunoglobulin domain in spectrin is stabilized when fused to R17 because the linker adopts a helical conformation that interacts with both domains (Table 1). This suggests that the linkers, which connect between the constituent domains, must be considered more than simple covalent connectors.[40−42] The linkers perform the important task of establishing communication by directing the correlated movements of various domains.

COUPLED THERMODYNAMICS LEADING TO DECREASED STABILITY UPON TETHERING

Although there have been several reports of multidomain proteins possessing thermodynamic stability similar to or greater than that of their isolated constituent domains, destabilization because of tethering was not reported until very recently. A recent theoretical study has shown that tethering two domains may have an intrinsically destabilizing effect.[43] This effect originates in the larger entropy of the unfolded state of a given subunit because some of the residual structures located close to the tethering site are less likely to form in the tethered state than in the isolated state. Therefore, the domain is more stable in its unfolded, tethered form than in its unfolded, isolated form. The free energy for folding is therefore more favorable for this domain in the isolated (rather than tethered) state (Figure 2).

It was illustrated that this destabilization as a result of tethering exists even for two-domain proteins whose behavior is

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048

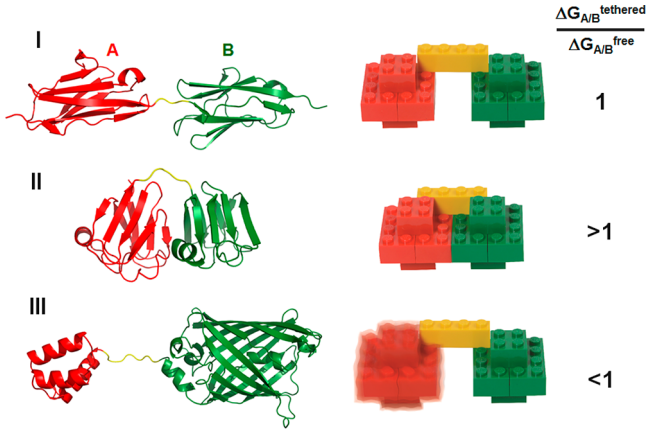


Figure 2. Folding thermodynamics of multidomain proteins. $\Delta G_{A/B}^{free}$ and $\Delta G_{A/B}^{tethered}$ correspond to the folding free energy of the isolated and tethered A or B domains, respectively. The two domains are colored red and green, and the linker that connects them is colored yellow. Illustrated are three cases in which the stability of at least one of the domains is unaffected (I), increased (II), or decreased (III) by the fusion to the other domain. The proteins are Fnfn9−Fnfn10 (I), Protein S (II), and homeodomains fused to GFP (III), where the softness of the homeodomains is indicated by an out-of-focus LEGO brick. The experimental thermodynamic stabilities of these proteins are listed in Table 1. The increased stability in II is due to a change in the folded state, but the decreased stability in III is mostly from a change in the unfolded state. While the multidomain proteins described in cases I and II can be viewed as being consistent with the LEGO perspective, the proteins that correspond to case III deviate from that as structural and thermodynamic additivities do not coincide.

the sum of their parts; however, several competing effects may come into play to determine the thermostability of the constituent domains.[43] The existence of a small interface between the tethered domains or a difference in the folding thermostability of the constituent domains may compensate for the destabilization effect of the tethering itself. The various determinants may combine to influence the cross-talk between domains in numerous ways and to various extents, depending on the specific characteristics of the inspected multidomain constructs. As a result, considerable diversity is expected. Therefore, identifying correlations between topology-driven cross-talk and measurable folding observables, such as thermostability and folding rates, is of major importance if one wishes to understand better the folding process in relation to multidomain proteins.

Overall, the experimental "tethered yet independent" perspective (i.e., two-domain constructs whose overall behavior is the sum of their parts) is the combined result of several folding determinants, which may mask each other's effects. Uncoupled folding behavior, with $\Delta G_{INT} = 0$,[32] therefore involves a complex microscopic mechanism in which the thermodynamics of each domain may be coupled, and it should not be assumed that the tethering has no effect.

Since the original theoretical study predicting that tethering could lead to destabilization in multidomain proteins,[43] several experimental studies have reported destabilization in multidomain proteins. A thermodynamic study showed that the Engrailed homeodomain (EnHD) undergoes significant destabilization upon its fusion it to green fluorescent protein (GFP) regardless of the linker length used and whether the tethering is to its N- or C-terminus (Table 1). The destabilization of EnHD caused by tethering it to GFP is reflected in melting temperatures reduced by up to 9 °C.[44] A similar destabilization effect following tethering of the protein to GFP was observed for phosphoglycerate kinase (PGK). Thermal destabilization was observed for PGK when GFP was linked at the N- or C-terminus. While destabilization was observed when GFP was linked at the N- and C-termini simultaneously, the effect was not additive, which suggests that cross-talk between EnHD and GFP may involve a crowding effect that decreases the magnitude of the destabilization effect.[45] Destabilization was also reported for the C domain of the natural protein utrophin tandem calponin homology, which is less stable in its full length when it is tethered to the N domain.[46]

BIOLOGICAL IMPLICATIONS OF DESTABILIZATION FOLLOWING PROTEIN FUSION

Destabilization as a result of conjugation may have biological significance. It was hypothesized that the connection of a ubiquitin polymer to a protein substrate (ubiquitination) might be linked to the function performed by the modification. In particular, the thermodynamic destabilization of a protein after ubiquitination by the ubiquitin chain (which is often associated with proteasomal degradation) might be advantageous as it may assist the degradation process, which itself requires protein unfolding. This idea, which was raised approximately 25 years ago,[56] was illustrated in coarse-grained molecular dynamics simulations for the Ubc7 protein.[57] Tethering the Ubc7 enzyme to a K48-tetrameric Ub resulted in destabilization when the conjugation occurred at the *in vivo* sites for degradation. A much smaller effect was observed when conjugation was performed using other forms of ubiquitin or at alternative sites (i.e., other lysine residues that might be used for ubiquitination). The destabilization effect originates from a disruption of the residual structure of the unfolded state that increases its configurational entropy and therefore decreases its thermodynamic stability.

The ability of ubiquitin chains to affect the biophysical characteristics of a protein substrate is highly dependent on the type of ubiquitin chain and conjugation site. For example, K48-linked polyUb induces destabilization much more effectively than K63-linked polyUb or monomeric Ub does (Figure 3). The effect of Ub on substrate biophysics suggests that Ub acts as a "conformation-perturbing device"[56] and that Ub does not serve just as a signaling tag.[57] It was estimated that ubiquitin's active role in assisting degradation by inducing local unfolding that introduces the required disordered regions is relevant for ∼25% of proteins, whereas the rest already have a sufficiently long intrinsic disordered region.[58] The ubiquitin chains may affect not only the unfolded ensemble but also the folded state ensemble. It was shown that, despite a lack of specificity in the interaction between ubiquitin and its substrate, the ubiquitin can affect the conformational ensemble of particular structural regions of the substrate, and this may also contribute to reduced stability and affect function.[59,60]

Recently, the effect of ubiquitination on protein stability was measured experimentally for two proteins: the human FK506 binding protein (FKBP12) and human fatty acid binding protein 4 (FABP4). FKBP12 and FABP4 were conjugated with two types of ubiquitin chains (monomeric and hexameric ubiquitin chains) at different ubiquitination sites.[53] The 10 ubiquitinated variants of FKBP12 and FABP4 that were studied exhibited thermal destabilization. The largest destabilization

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048

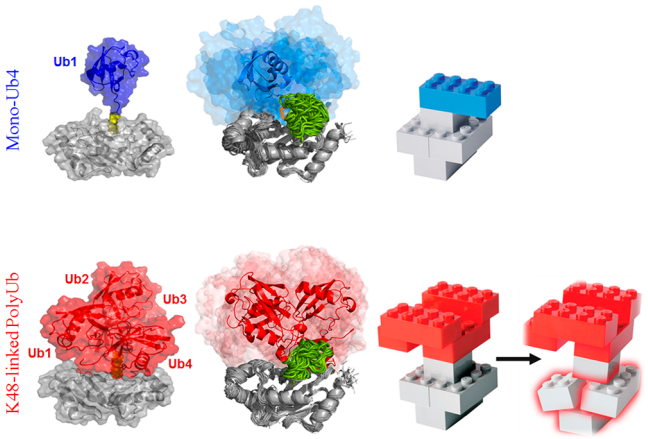


Figure 3. Effect of ubiquitination on protein biophysics. Conjugation of ubiquitin may change protein stability depending on the polymeric nature of ubiquitin. While monomeric ubiquitin may have an only minor effect on substrate stability, a K48-linked polyubiquitin may induce destabilization, which suggests that ubiquitin may not serve only as a signaling tag. The substrate protein is illustrated by the Ubc7 protein and is depicted in grey and its loop residues 95−105 are depicted in green (see ref 57 and 59). Monomeric ubiquitin and K48-linked polyubiquitin are shown in blue and red, respectively.

effect was obtained when conjugating with hexameric ubiquitin; this modification reduced the melting temperatures of FKBP12 and FABP4 by 18 and 9 °C, respectively. Similarly, cross-linking two staphylococcal nuclease proteins resulted in thermodynamic destabilization whose magnitude was dependent on the conjugation site.[54] Reduced thermal stability was also reported experimentally due to ubiquitination of the β1 domain of streptococcal protein G (Gβ1) and the B domain of staphylococcal protein A (SpAB). Several variants of the Gβ1 and SpA_B proteins were linked to a monomeric ubiquitin at the N-terminus.[52] Most of the variants showed lower melting temperatures with an average difference between the free Gb1 (SpA_B) variant and the Gb1 (SpA_B) variant fused to Ub of 7.2 ± 3.7 °C (1.8 ± 1.5 °C).

OTHER BIOPHYSICAL MANIFESTATIONS FOR CROSS-TALK WITHIN AND BETWEEN PROTEIN DOMAINS

Cross-talk between protein domains that are linked covalently or that interact via intermolecular forces is essential and does not have to be associated solely with a thermodynamic effect. The "induced-fit" mechanism is well accepted as being relevant to a variety of biomolecular recognitions, whereas the old "lock-and-key" mechanism, which ignores cross-talk, may be relevant in much fewer cases.[61] The plasticity of proteins and the existence of a conformational ensemble associated with the native state basin of the energy landscape are the subject of ongoing research to quantify their magnitude, directionality, and determinants. Allostery, in which a change in substrate binding affinity is mediated by a change in the free energy landscape of a protein conformation (or a change in protein dynamics), is an important mechanism of communication within a single domain or between domains and is vital for function.[62] Another illustration is the "cracking" mechanism, in which the system makes excursions along a low-frequency normal mode (which is not costly in energy), during which it locally unfolds (which is also not too costly in free energy) and then refolds.[63,64] Another manifestation consistent with the view that proteins are evolutionarily selected to adopt different states and to locally unfold for function is the existence of frustrated sites[65] in which interactions are not necessarily optimized energetically. These types of energetic conflicts may trigger conformational changes and therefore functional transitions.[66]

A protein domain response to conjugation was also reported for nonproteinaceous conjugating molecules. Tethering a protein to surfaces was found to affect the protein folding mechanism and protein stability.[67,68] The magnitude of the effects depends on the protein sites selected for tethering and the hydrophobicity of the surface. Glycosylation and PEGylation can also change protein biophysics, and their effect is less obvious than that of other post-translational modifications, such as phosphorylation or acetylation, that involve changes in charge and thus change the chemical environment. Glycosylation and PEGylation, like ubiquitination, might be expected to have minor effects on protein stability because these conjugates are highly polar and exposed to the solvent. However, several studies have shown large effects caused by these modifications. Computational studies showed that the degree of stabilization mostly depends on the degree of glycosylation,[69] but destabilization by glycosylation is also not uncommon.[70,71] In particular, a series of studies on the pinWW protein have shown glycosylation and PEGylation to exert complex effects on the stability and folding rate of pinWW that are very sensitive to the conjugation sites,[72] the type of conjugating polymers, and the type of linker.[73] Accordingly, conjugating pinWW with either polysaccharides or PEG cannot be described as "sum of its parts" thermodynamics. These polymers can have an either stabilizing or destabilizing effect depending on the modified sites. Sometimes they may exert opposite effects despite being conjugated to the same protein site.[74,75] Thus, predicting the effect of glycosylation and PEGylation on protein stability is not trivial.[76]

SUMMARY

Metaphors are widely used in science. They are essential tools for understanding, to generate novel insight, and to facilitate the communication of an idea,[77,78] especially for complicated molecular processes. However, some metaphors are so commonplace that it is easy to mistake them for literal descriptions of the molecular world. In this opinion, we discuss the consequences of describing proteins as LEGO bricks. Certainly, the modularity of protein structures and the hierarchy of their construction justify this perspective. Proteins are built from a limited set of building blocks of different sizes that are reused to build a small protein domain as well as large proteins that comprise many domains. Modern proteins likely arose through the assembly and concatenation of smaller fragments via duplication, mutation, and shuffling from existing domains. This modularity lies at the heart of protein design approaches in which novel proteins are constructed, for example, to optimize their properties, to perform new biological functions, or even to explore regions in the sequences and structure spaces that were not sampled by evolution. Protein building blocks are often used in the protocol of *ab initio* structure prediction for proteins. The modularity of proteins is also fundamental to synthetic biology where various nano-assemblies are designed for a range of biotechnological applications.

However, the LEGO brick metaphor should not be stretched to the extent that proteins are mistakenly interpreted as rigid

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048

independent objects that can simply be fitted together to form larger proteins that behave as the sum of their parts. The LEGO brick metaphor may suggest that structural additivity guarantees thermodynamic additivity. Obviously, proteins are soft, and their plasticity and local unfolding are widely acknowledged as contributing to protein function via various mechanisms, such as induced fit, allostery, and cracking. We argue that even in cases in which the proteins seem to behave like LEGO bricks, the energy landscape of the protein might be affected by the LEGO-like assembly mechanism. The monomers in multidomain proteins do not necessarily behave as they do in isolation. Furthermore, recent studies show that protein tethering not only may produce stabilization but also can create destabilization, which is somewhat surprising. The destabilization due to tethering reveals the potential complexity of multidomain proteins simply by flexible linkers. Destabilization, in principle, can be caused not only by tethering one protein to another but also via post-translational modification; therefore, it might be linked to function. This mechanism was proposed to explain the induced thermodynamic destabilization by ubiquitination that may indirectly assist proteasomal degradation.[53,57] The destabilization of a tethered protein compared with a protein in isolation may explain the lower thermal stability that was recently measured *in vivo* for longer proteins.[79]

Protein design approaches often incorporate ideal protein modules (e.g., optimal helices and sheets) that result in highly stable proteins.[6,27] In light of these hyperstable designed proteins together with the limited success of *de novo* protein design, we speculate that incorporating soft LEGO brick fragments that deviate from ideality may improve protein design protocols and enable the design of proteins whose underlying energy landscape includes multiple distinct energy minima and not a single deep funnel. Natural proteins provide numerous examples of rich functionality, including allostery and signaling, that can emerge in protein systems with multiple low-energy states and moving parts that can be toggled by external stimuli. Designing proteins using soft building blocks that deviate from ideality and that do not assemble perfectly (see Figure 2, III) may result in robust molecules that achieve their intended design function.

AUTHOR INFORMATION

Corresponding Author

*Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel. E-mail: Koby.Levy@weizmann.ac.il. Telephone: 972-8-9344587.

ORCID

Yaakov Levy: 0000-0002-9929-973X

Funding

This work was supported by the Benoziyo Fund for the Advancement of Science and by the Kimmelman Center for Macromolecular Assemblies. Y.L. is The Morton and Gladys Pickman Professional Chair in Structural Biology.

Notes

The author declares no competing financial interest.

REFERENCES

(1) Moore, A. D., Bjorklund, A. K., Ekman, D., Bornberg-Bauer, E., and Elofsson, A. (2008) Arrangements in the modular evolution of proteins. *Trends Biochem. Sci. 33*, 444−451.

(2) Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. (1999) From molecular to modular cell biology. *Nature 402*, C47−C52.

(3) Woolfson, D. N., Bartlett, G. J., Burton, A. J., Heal, J. W., Niitsu, A., Thomson, A. R., and Wood, C. W. (2015) De novo protein design: how do we expand into the universe of possible protein structures? *Curr. Opin. Struct. Biol. 33*, 16−26.

(4) Kolodny, R., Pereyaslavets, L., Samson, A. O., and Levitt, M. (2013) On the Universe of Protein Folds. *Annu. Rev. Biophys. 42*, 559−582.

(5) Koga, N., Tatsumi-Koga, R., Liu, G. H., Xiao, R., Acton, T. B., Montelione, G. T., and Baker, D. (2012) Principles for designing ideal protein structures. *Nature 491*, 222−227.

(6) Huang, P. S., Boyken, S. E., and Baker, D. (2016) The coming of age of de novo protein design. *Nature 537*, 320−327.

(7) Panchenko, A. R., LutheySchulten, Z., and Wolynes, P. G. (1996) Foldons, protein structural modules, and exons. *Proc. Natl. Acad. Sci. U. S. A. 93*, 2008−2013.

(8) Le, Q., Pollastri, G., and Koehl, P. (2009) Structural Alphabets for Protein Structure Classification: A Comparison Study. *J. Mol. Biol. 387*, 431−450.

(9) Tyagi, M., Sharma, P., Swamy, C. S., Cadet, F., Srinivasan, N., de Brevern, A. G., and Offmann, B. (2006) Protein Block Expert (PBE): a web-based protein structure analysis server using a structural alphabet. *Nucleic Acids Res. 34*, W119−W123.

(10) Vanhee, P., Verschueren, E., Baeten, L., Stricher, F., Serrano, L., Rousseau, F., and Schymkowitz, J. (2011) BriX: a database of protein building blocks for structural analysis, modeling and design. *Nucleic Acids Res. 39*, D435−D442.

(11) Remmert, M., Biegert, A., Linke, D., Lupas, A. N., and Soding, J. (2010) Evolution of Outer Membrane beta-Barrels from an Ancestral beta beta Hairpin. *Mol. Biol. Evol. 27*, 1348−1358.

(12) Fitzkee, N. C., Fleming, P. J., Gong, H. P., Panasik, N., Street, T. O., and Rose, G. D. (2005) Are proteins made from a limited parts list? *Trends Biochem. Sci. 30*, 73−80.

(13) Hu, W. B., Kan, Z. Y., Mayne, L., and Englander, S. W. (2016) Cytochrome c folds through foldon-dependent native-like intermediates in an ordered pathway. *Proc. Natl. Acad. Sci. U. S. A. 113*, 3809−3814.

(14) Weinkam, P., Zimmermann, J., Romesberg, F. E., and Wolynes, P. G. (2010) The Folding Energy Landscape and Free Energy Excitations of Cytochrome c. *Acc. Chem. Res. 43*, 652−660.

(15) Taylor, W. R. (2002) A 'periodic table' for protein structures. *Nature 416*, 657−660.

(16) Moutevelis, E., and Woolfson, D. N. (2009) A Periodic Table of Coiled-Coil Protein Structures. *J. Mol. Biol. 385*, 726−732.

(17) Ahnert, S. E., Marsh, J. A., Hernandez, H., Robinson, C. V., and Teichmann, S. A. (2015) Principles of assembly reveal a periodic table of protein complexes. *Science 350*, aaa2245.

(18) Lakshmanan, A., Zhang, S. G., and Hauser, C. A. E. (2012) Short self-assembling peptides as building blocks for modern nanodevices. *Trends Biotechnol. 30*, 155−165.

(19) Gazit, E. (2007) Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chem. Soc. Rev. 36*, 1263−1269.

(20) Rajagopal, K., and Schneider, J. P. (2004) Self-assembling peptides and proteins for nanotechnological applications. *Curr. Opin. Struct. Biol. 14*, 480−486.

(21) Gilardi, G., Meharenna, Y. T., Tsotsou, G. E., Sadeghi, S. J., Fairhead, M., and Giannini, S. (2002) Molecular Lego: design of molecular assemblies of P450 enzymes for nanobiotechnology. *Biosens. Bioelectron. 17*, 133−145.

(22) Park, K., Shen, B. W., Parmeggiani, F., Huang, P. S., Stoddard, B. L., and Baker, D. (2015) Control of repeat-protein curvature by computational protein design. *Nat. Struct. Mol. Biol. 22*, 167−174.

(23) Voet, A. R. D., Noguchi, H., Addy, C., Simoncini, D., Terada, D., Unzai, S., Park, S. Y., Zhang, K. Y. J., and Tame, J. R. H. (2014) Computational design of a self-assembling symmetrical beta-propeller protein. *Proc. Natl. Acad. Sci. U. S. A. 111*, 15102−15107.

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048

(24) King, N. P., Bale, J. B., Sheffler, W., McNamara, D. E., Gonen, S., Gonen, T., Yeates, T. O., and Baker, D. (2014) Accurate design of co-assembling multi-component protein nanomaterials. *Nature 510*, 103.
(25) Thomson, A. R., Wood, C. W., Burton, A. J., Bartlett, G. J., Sessions, R. B., Brady, R. L., and Woolfson, D. N. (2014) Computational design of water-soluble alpha-helical barrels. *Science 346*, 485−488.
(26) Fletcher, J. M., Harniman, R. L., Barnes, F. R. H., Boyle, A. L., Collins, A., Mantell, J., Sharp, T. H., Antognozzi, M., Booth, P. J., Linden, N., Miles, M. J., Sessions, R. B., Verkade, P., and Woolfson, D. N. (2013) Self-Assembling Cages from Coiled-Coil Peptide Modules. *Science 340*, 595−599.
(27) Huang, P. S., Oberdorfer, G., Xu, C. F., Pei, X. Y., Nannenga, B. L., Rogers, J. M., DiMaio, F., Gonen, T., Luisi, B., and Baker, D. (2014) High thermodynamic stability of parametrically designed helical bundles. *Science 346*, 481−485.
(28) Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnels, Pathways, and the Energy Landscape of Protein-Folding - a Synthesis. *Proteins: Struct., Funct., Genet. 21*, 167−195.
(29) Rothemund, P. W. K. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature 440*, 297−302.
(30) Bornberg-Bauer, E., Huylmans, A. K., and Sikosek, T. (2010) How do new proteins arise? *Curr. Opin. Struct. Biol. 20*, 390−396.
(31) Scott, K. A., Steward, A., Fowler, S. B., and Clarke, J. (2002) Titin; a multidomain protein that behaves as the sum of its parts. *J. Mol. Biol. 315*, 819−829.
(32) Batey, S., Nickson, A. A., and Clarke, J. (2008) Studying the folding of multidomain proteins. *HFSP J. 2*, 365−377.
(33) Aksel, T., and Barrick, D. (2009) Analysis of Repeat-Protein Folding Using Nearest-Neighbor Statistical Mechanical Models. *Methods Enzymol. 455*, 95−125.
(34) Yuwen, T., Post, C. B., and Skrynnikov, N. R. (2011) Domain cooperativity in multidomain proteins: what can we learn from molecular alignment in anisotropic media? *J. Biomol. NMR 51*, 131−150.
(35) Bhaskara, R. M., and Srinivasan, N. (2011) Stability of domain structures in multi-domain proteins. *Sci. Rep. 1*, 00040.
(36) Han, J. H., Batey, S., Nickson, A. A., Teichmann, S. A., and Clarke, J. (2007) The folding and evolution of multidomain proteins. *Nat. Rev. Mol. Cell Biol. 8*, 319−330.
(37) Hua, L., Huang, X., Liu, P., Zhou, R., and Berne, B. J. (2007) Nanoscale dewetting transition in protein complex folding. *J. Phys. Chem. B 111*, 9069−9077.
(38) Sikora, M., and Cieplak, M. (2011) Mechanical stability of multidomain proteins and novel mechanical clamps. *Proteins: Struct., Funct., Genet. 79*, 1786−1799.
(39) Osvath, S., Kohler, G., Zavodszky, P., and Fidy, J. (2005) Asymmetric effect of domain interactions on the kinetics of folding in yeast phosphoglycerate kinase. *Protein Sci. 14*, 1609−1616.
(40) Gokhale, R. S., and Khosla, C. (2000) Role of linkers in communication between protein modules. *Curr. Opin. Chem. Biol. 4*, 22−27.
(41) Michnick, S. W., Ear, P. H., Landry, C., Malleshaiah, M. K., and Messier, V. (2011) Protein-fragment complementation assays for large-scale analysis, functional dissection and dynamic studies of protein-protein interactions in living cells. *Methods Mol. Biol. 756*, 395−425.
(42) Papaleo, E., Saladino, G., Lambrughi, M., Lindorff-Larsen, K., Gervasio, F. L., and Nussinov, R. (2016) The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem. Rev. 116*, 6391−6423.
(43) Arviv, O., and Levy, Y. (2012) Folding of multidomain proteins: biophysical consequences of tethering even in apparently independent folding. *Proteins: Struct., Funct., Genet. 80*, 2780−2798.
(44) Sokolovski, M., Bhattacherjee, A., Kessler, N., Levy, Y., and Horovitz, A. (2015) Thermodynamic Protein Destabilization by GFP Tagging: A Case of Interdomain Allostery. *Biophys. J. 109*, 1157−1162.
(45) Dave, K., Gelman, H., Thu, C. T. H., Guin, D., and Gruebele, M. (2016) The Effect of Fluorescent Protein Tags on Phosphoglycerate Kinase Stability Is Nonadditive. *J. Phys. Chem. B 120*, 2878−2885.
(46) Bandi, S., Singh, S. M., and Mallela, K. M. (2014) The C-terminal domain of the utrophin tandem calponin-homology domain appears to be thermodynamically and kinetically more stable than the full-length protein. *Biochemistry 53*, 2209−2211.
(47) Wenk, M., Jaenicke, R., and Mayr, E. M. (1998) Kinetic stabilisation of a modular protein by domain interactions. *FEBS Lett. 438*, 127−130.
(48) Batey, S., Randles, L. G., Steward, A., and Clarke, J. (2005) Cooperative folding in a multi-domain protein. *J. Mol. Biol. 349*, 1045−1059.
(49) Batey, S., and Clarke, J. (2006) Apparent cooperativity in the folding of multidomain proteins depends on the relative rates of folding of the constituent domains. *Proc. Natl. Acad. Sci. U. S. A. 103*, 18113−18118.
(50) Rothlisberger, D., Honegger, A., and Pluckthun, A. (2005) Domain interactions in the Fab fragment: A comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *J. Mol. Biol. 347*, 773−789.
(51) Jager, M., Gehrig, P., and Pluckthun, A. (2001) The scFv fragment of the antibody hu4D5−8: Evidence for early premature domain interaction in refolding. *J. Mol. Biol. 305*, 1111−1129.
(52) Navarro, M. F., Carmody, L., Romo-Fewell, O., Lokensgard, M. E., and Love, J. J. (2014) Characterizing Substrate Selectivity of Ubiquitin C-Terminal Hydrolase-L3 Using Engineered alpha-Linked Ubiquitin Substrates. *Biochemistry 53*, 8031−8042.
(53) Morimoto, D., Walinda, E., Fukada, H., Sugase, K., and Shirakawa, M. (2016) Ubiquitylation Directly Induces Fold Destabilization of Proteins. *Sci. Rep. 6*, 39453.
(54) Kim, Y. H., and Stites, W. E. (2008) Effects of excluded volume upon protein stability in covalently cross-linked proteins with variable linker lengths. *Biochemistry 47*, 8804−8814.
(55) Steward, A., Adhya, S., and Clarke, J. (2002) Sequence conservation in Ig-like domains: the role of highly conserved proline residues in the fibronectin type III superfamily. *J. Mol. Biol. 318*, 935−940.
(56) Varshavsky, A. (1992) The N-end rule. *Cell 69*, 725−735.
(57) Hagai, T., and Levy, Y. (2010) Ubiquitin not only serves as a tag but also assists degradation by inducing protein unfolding. *Proc. Natl. Acad. Sci. U. S. A. 107*, 2001−2006.
(58) Hagai, T., Azia, A., Toth-Petroczy, A., and Levy, Y. (2011) Intrinsic disorder in ubiquitination substrates. *J. Mol. Biol. 412*, 319−324.
(59) Gavrilov, Y., Hagai, T., and Levy, Y. (2015) Nonspecific yet decisive: Ubiquitination can affect the native-state dynamics of the modified protein. *Protein Sci. 24*, 1580−1592.
(60) Ball, K. A., Johnson, J. R., Lewinski, M. K., Guatelli, J., Verschueren, E., Krogan, N. J., and Jacobson, M. P. (2016) Non-degradative Ubiquitination of Protein Kinases. *PLoS Comput. Biol. 12*, e1004898.
(61) Boehr, D. D., Nussinov, R., and Wright, P. E. (2009) The role of dynamic conformational ensembles in biomolecular recognition (vol 5, pg 789, 2009). *Nat. Chem. Biol. 5*, 954−954.
(62) Guo, J. J., and Zhou, H. X. (2016) Protein Allostery and Conformational Dynamics. *Chem. Rev. 116*, 6503−6515.
(63) Miyashita, O., Onuchic, J. N., and Wolynes, P. G. (2003) Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci. U. S. A. 100*, 12570−12575.
(64) Whitford, P. C., Miyashita, O., Levy, Y., and Onuchic, J. N. (2007) Conformational transitions of adenylate kinase: switching by cracking. *J. Mol. Biol. 366*, 1661−1671.
(65) Ferreiro, D. U., Hegler, J. A., Komives, E. A., and Wolynes, P. G. (2007) Localizing frustration in native proteins and protein assemblies. *Proc. Natl. Acad. Sci. U. S. A. 104*, 19819−19824.
(66) Ferreiro, D. U., Komives, E. A., and Wolynes, P. G. (2014) Frustration in biomolecules. *Q. Rev. Biophys. 47*, 285−363.

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048

(67) Friedel, M., Baumketner, A., and Shea, J. E. (2007) Stability of a protein tethered to a surface. *J. Chem. Phys. 126*, 095101.
(68) Zhuang, Z., Jewett, A. I., Soto, P., and Shea, J. E. (2009) The effect of surface tethering on the folding of the src-SH3 protein domain. *Phys. Biol. 6*, 015004.
(69) Shental-Bechor, D., and Levy, Y. (2008) Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U. S. A. 105*, 8256−8261.
(70) Shental-Bechor, D., and Levy, Y. (2009) Folding of glycoproteins: toward understanding the biophysics of the glycosylation code. *Curr. Opin. Struct. Biol. 19*, 524−533.
(71) Gavrilov, Y., Shental-Bechor, D., Greenblatt, H. M., and Levy, Y. (2015) Glycosylation May Reduce Protein Thermodynamic Stability by Inducing a Conformational Distortion. *J. Phys. Chem. Lett. 6*, 3572−3577.
(72) Price, J. L., Shental-Bechor, D., Dhar, A., Turner, M. J., Powers, E. T., Gruebele, M., Levy, Y., and Kelly, J. W. (2010) Context-dependent effects of asparagine glycosylation on Pin WW folding kinetics and thermodynamics. *J. Am. Chem. Soc. 132*, 15359−15367.
(73) Lawrence, P. B., Billings, W. M., Miller, M. B., Pandey, B. K., Stephens, A. R., Langlois, M. I., and Price, J. L. (2016) Conjugation Strategy Strongly Impacts the Conformational Stability of a PEG-Protein Conjugate. *ACS Chem. Biol. 11*, 1805−1809.
(74) Price, J. L., Powers, E. T., and Kelly, J. W. (2011) N-PEGylation of a Reverse Turn Is Stabilizing in Multiple Sequence Contexts, unlike N-GlcNAcylation. *ACS Chem. Biol. 6*, 1188−1192.
(75) Lawrence, P. B., and Price, J. L. (2016) How PEGylation influences protein conformational stability. *Curr. Opin. Chem. Biol. 34*, 88−94.
(76) Lawrence, P. B., Gavrilov, Y., Matthews, S. S., Langlois, M. I., Shental-Bechor, D., Greenblatt, H. M., Pandey, B. K., Smith, M. S., Paxman, R., Torgerson, C. D., Merrell, J. P., Ritz, C. C., Prigozhin, M. B., Levy, Y., and Price, J. L. (2014) Criteria for Selecting PEGylation Sites on Proteins for Higher Thermodynamic and Proteolytic Stability. *J. Am. Chem. Soc. 136*, 17547−17560.
(77) Brown, T. L. (2003) *Making Truth: Metaphor in Science*, University of Illinois, Champaign, IL.
(78) Chew, M. K., and Laubichler, M. D. (2003) Natural enemies - Metaphor or misconception? *Science 301*, 52−53.
(79) Leuenberger, P., Ganscha, S., Kahraman, A., Cappelletti, V., Boersema, P. J., von Mering, C., Claassen, M., and Picotti, P. (2017) Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science 355*, eaai7825.

DOI: 10.1021/acs.biochem.7b00666
Biochemistry 2017, 56, 5040−5048