Globular Protein Folding In Vitro and In Vivo

Martin Gruebele,[1,2,3] Kapil Dave,[1] and Shahar Sukenik[2]

[1]Center for Biophysics and Quantitative Biology, University of Illinois, Urbana, Illinois 61801; email: mgruebel@illinois.edu, kapildave91@gmail.com

[2]Department of Chemistry, University of Illinois, Urbana, Illinois 61801; email: shaharsu@gmail.com

[3]Department of Physics, University of Illinois, Urbana, Illinois 61801

Annu. Rev. Biophys. 2016. 45:233–51

The *Annual Review of Biophysics* is online at biophys.annualreviews.org

This article's doi: 10.1146/annurev-biophys-062215-011236

Keywords

energy landscape, lattice model, single molecule, folding kinetics, folding thermodynamics, free energy, protein function, protein evolution, in-cell folding

Abstract

In vitro, computational, and theoretical studies of protein folding have converged to paint a rich and complex energy landscape. This landscape is sensitively modulated by environmental conditions and subject to evolutionary pressure on protein function. Of these environments, none is more complex than the cell itself, where proteins function in the cytosol, in membranes, and in different compartments. A wide variety of kinetic and thermodynamics experiments, ranging from single-molecule studies to jump kinetics and from nuclear magnetic resonance to imaging on the microscope, have elucidated how protein energy landscapes facilitate folding and how they are subject to evolutionary constraints and environmental perturbation. Here we review some recent developments in the field and refer the reader to some original work and additional reviews that cover this broad topic in protein science.

Contents

INTRODUCTION

Protein folding produces much of the cell's signaling, structural, and catalytic machinery. It happens first upon ribosomal synthesis (21), often with membrane insertion via the translocon (63, 131), but also later on in the cell. One of the most important things learned from in vitro folding experiments is that even cytosolic globular proteins have fairly small folding equilibrium constants. Therefore, proteins will unfold and refold many times during their life cycle (76). With the exception of a few extraordinarily stable proteins (165), relatively low stability goes hand in hand with the flexibility required for protein function. Some proteins even remain unstructured after translation and fold upon binding to specific targets (146, 156).

In vitro studies also taught us that folding is a very fast chemical reaction (microseconds to hours at room temperature). Its free-energy barriers, $\Delta G^{\dagger}$, must be quite small, in some cases on the order of the thermal energy $k_B T$ (5, 20). Thanks to small folding free energies and small activation barriers, one might expect that the complex solvation environment of the cell can control protein stability and kinetics, and indeed it can (78). In addition, a network of chaperones can hold misfolded proteins, direct them toward degradation pathways, or unfold them, giving proteins inside cells another chance to fold autonomously, as most proteins do in vitro at low concentration when aggregation is unlikely (30, 153).

The plausibility of in vivo effects on folding is apparent from in vitro studies: Slight temperature changes, addition of small molecules, or crowding by large molecules can shift protein equilibria between unfolded and folded ensembles (75). Such shifts are often "cooperative," by which we mean that they occur over a narrow range of conditions (2). Although the cell modulates the folding free-energy landscape, it does not appear to fundamentally alter the way proteins are observed to fold in vitro (77).

True understanding requires that one should be able to put something back together after taking it apart. Protein scientists have gone through this process in a variety of ways. Although de novo design of active proteins is still not routine, much progress has been made in that field (40, 106). Likewise, model building has gone well. On the energetic side, the energy landscape model has explained many of the general (43, 121) and specific (127) features of folding. On the

structural side, models have advanced from beads on lattices (94) to all-atom simulations based on empirical force fields (19, 45, 145). The last 10 years have seen a remarkable confluence of protein (un)folding experiments, protein design, protein landscape models, and simulations of folding. The state of the art is proteins of ≈100 residues, folding faster than a few milliseconds if a direct comparison of simulations and experiments is to be made (107). Most domains of larger proteins are <150 amino acids long, and such domains usually fold relatively independently from one another (75). We are thus not far off from the holy grail where folds can be reliably computed, just as the structure of small organic molecules can be computed readily with quantum chemistry packages (36).

Many interesting problems remain to be solved. Although computation can predict the fold of some small proteins, how accurate the predicted mechanisms are is not yet clear. This is partly the fault of experiments, which have difficulty providing structural information on the timescale of the actual reaction (barrier crossing) events. An important question is, How detailed do we really need to be to have useful predictions? Although folding reactions can be described adequately by simple mechanisms along one or two reaction coordinates (29), considerable complexity lurks below this apparent simplicity (**Figure 1**) (11, 119). In particular, the unfolded ensemble has more structure and interesting dynamics than it is often given credit for (50, 157). And of course there is the question of how cells productively fine-tune the energy landscape of their proteins to enhance survival (114). Finally, other interesting problems, such as the effect of applied force on energy landscapes (164) or misfolding and amyloids (34), have been reviewed recently in volumes of the *Annual Review of Biophysics* and are only briefly referenced here.

IN VITRO PROTEIN FOLDING

Structural and Energetic Models for Folding

Perhaps the most basic model used to understand protein folding is the hydrophobic-hydrophilic (HP) residue model (42). The HP model accounts for hydrophobicity as the major driving force to make compact, desolvated structures (26, 42, 102), while also allowing local secondary structure formation. Hydrophobicity has been reviewed extensively (27, 149), and although hydrophobicity is not purely an entropic effect, water molecules avoiding ordered interaction with sidechains buried in the protein's core play an important role. Several key ideas emerged from such simple models, including that certain sequences are more likely to avoid kinetic traps and produce robust folded states that are rapidly accessible from unfolded conformations. Even though no solvent is included explicitly in HP models, hydrophobicity highlights that the solvent folds, too, when a protein folds. Although 2–amino acid alphabets do not fold proteins in practice, alphabets with as few as 5 residues have been successful (133), and disordered proteins are an excellent example of proteins with reduced alphabets (159). Of course, a larger alphabet of 20 different amino acid residues still leads to better-packed structures that are more fine-tuned by evolution for function (158). For a more detailed discussion of theoretical protein models, see the review in this volume by Schuler et al. (141).

Ideas such as local secondary structure formation or hydrophobicity involve a successive reduction of the search space as the search for the native state goes on. For example, hydrophobicity partitions residues into a more likely inside group and a more likely outside group. Gō (73) realized that proteins are evolved to have consistent interactions, and Frauenfelder et al. (64) proposed hierarchical energy landscapes of native proteins. Such concepts led to a quantitative energy landscape theory of folding (20, 42). In energy landscape theory, it has been long understood that when the bond energy scale approaches k_BT, enthalpy and configurational entropy effects can easily

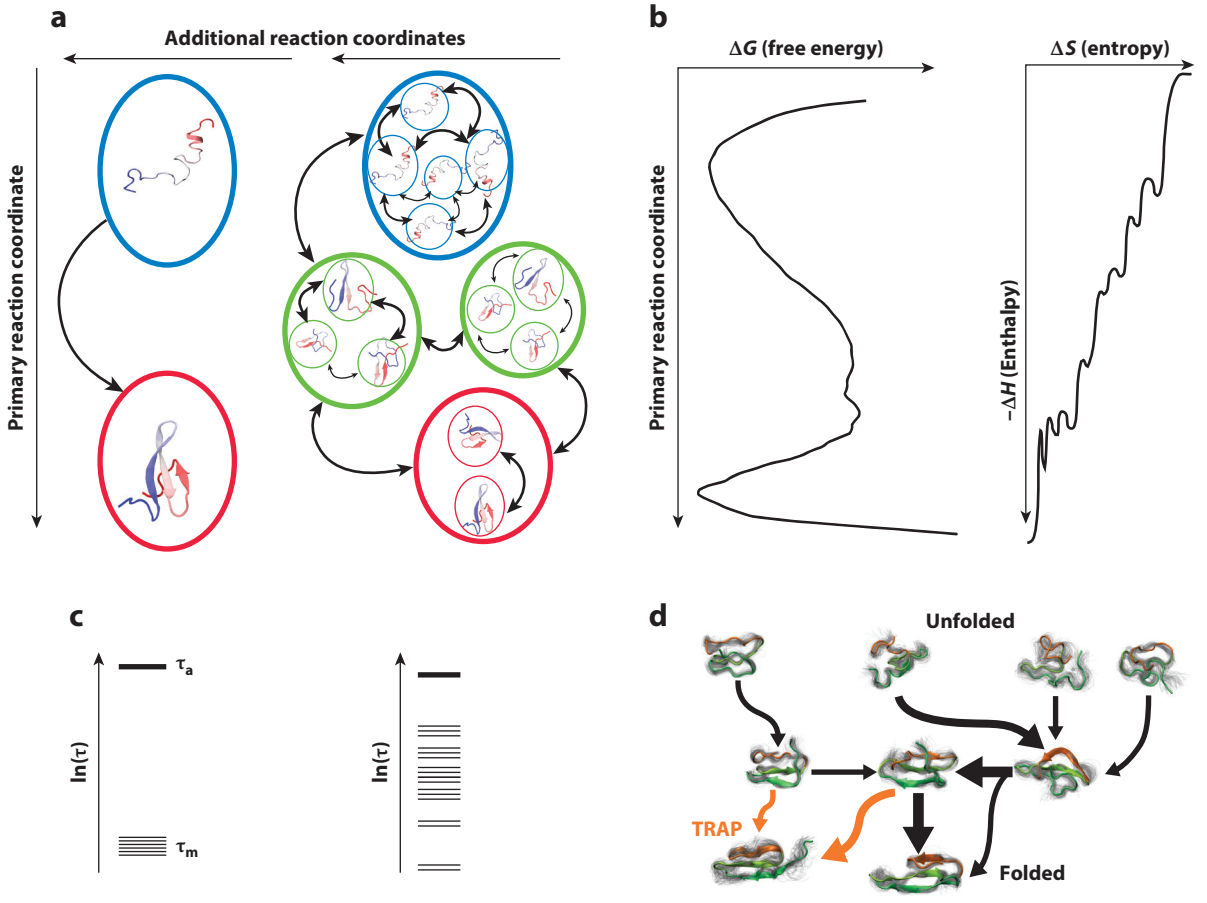

Figure 1

Folding simplicity and complexity: the structural and energetic view. (*a*) In the most coarse-grained structural picture on the very left, only two macrostates [unfolded (*top*) and folded (*bottom*)] play a role. The curved black arrow indicates interconversion through a transition state (whose probabilities of folding and unfolding are $p_F = p_U = 0.5$). In the next column of panel *a*, more detailed analysis reveals substates of folded, unfolded, and transitional ensembles, each containing many microstates (64). A rich kinetic network occasionally includes parallel paths (if their free energies are within a few k_BT, so one is not favored over the other), as shown by the curved arrows. In both cases, the main reaction coordinate coincides with the black downward arrow in panel *b*, whereas additional reaction coordinates correspond to the horizontal arrows at the top. (*b*) In the most coarse-grained energetic picture, the free energy $\Delta G(x)$ has unfolded and folded minima along just one primary reaction coordinate, x, and the folding enthalpy of a protein is well funneled as a function of the polypeptide configurational entropy. The lower-energy native state has lower configurational entropy and lower enthalpy. (*c*) For slow folders, the molecular time, τ_m (i.e., the ≈0.1–1 μs during which transition between states occurs), is well separated from the dwell time within states, τ_a. For fast folders, or proteins with many folding intermediates covering a wide range of barriers, the timescales overlap. (*d*) Simplified WW domain kinetic network, showing an actual calculated example of the less coarse-grained picture in panel *b* (119).

compensate (46, 160). Accordingly, in energy landscape theory, the Levinthal paradox (103) can be overcome because enthalpy loss ΔH and entropy loss ΔS_C are correlated as a protein folds, and such enthalpy-entropy compensation (31) overcomes unavoidable enthalpic noise or roughness. The correlation of enthalpy and entropy (funnel shape of ΔH as a function of ΔS_C in **Figure 1*b***) explains why proteins fold over low free-energy barriers (5). The combination of slope and roughness of the enthalpy funnel explains traps and intermediates when folding is frustrated, in analogy

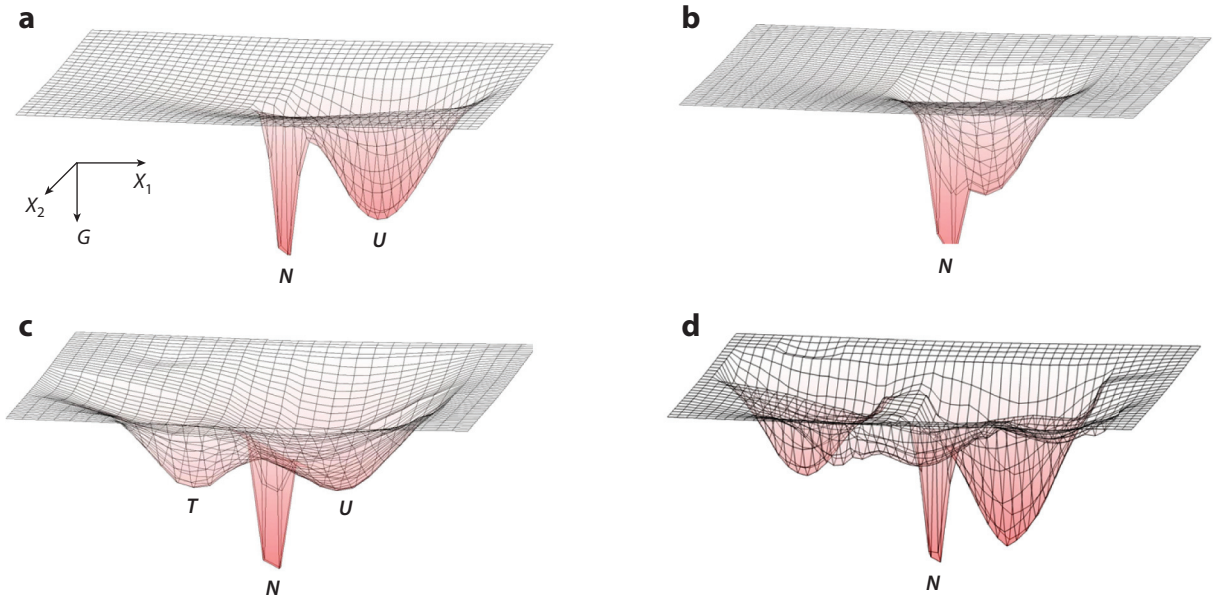

Figure 2

Free-energy landscapes of protein folding highlight several scenarios. (*a*) Scenario of two-state folding with well-defined native (labeled *N*) and unfolded (labeled *U*) well separated by a barrier. Axes illustrate the two folding reaction coordinates (x_1 and x_2) and folding free energy (*G*). (*b*) Downhill folding portrayed as native and unfolded well-separated by a low lying barrier ($\leq 3\ k_BT$). (*c*) The scenario of a free-energy trap (labeled *T*) is added to the two-state scenario. (*d*) The concept of multiple folding pathways available to the protein in case of change in environmental conditions or mutations is shown by the presence of various minima ending up in the native state of the protein.

to terminology used in dynamics of glasses. The funneled function ΔH (ΔS_C) should not be confused with the free energy $\Delta G(x) = \Delta H - T\Delta S$ as a function of reaction progress coordinate x (illustrated for several cases in **Figure 2**) (93). Although the funnel is downhill in enthalpy, the free energy is not necessarily downhill because ΔH and $-T\Delta S$ may not compensate for all values of x.

As computational power has grown, increasingly realistic computer models of folding have become possible (45, 107, 138, 145). Coarse graining is a powerful approach that dramatically decreases the computational demands of protein simulations (104, 137). In a coarse-grained model, clusters of atoms are modeled as a unit, interacting via an appropriately averaged force field. In parallel, implicit solvent models greatly reduce the number of atoms tracked in classical molecular dynamics simulations and yield interesting folding behavior such as a dominant but parallel pathway (23). The development of parallel simulation methods greatly improved sampling. Many parallel calculations can be sampled in search of a few successful folding events for comparison with experiment (148). Independent calculations can be stitched together (using Markov state models) to reveal short-lived or long-lived microstates (17). Replicas can exchange between simulations to provide rapid thermodynamic sampling (82, 175). For example, replica exchange has computationally revealed downhill free-energy surfaces for folding (125). Recent advances in computational power have made possible all-atom single-trajectory protein folding simulations, in which a single-domain protein unfolds and refolds many times in equilibrium (107). As with experiments, one of the greatest challenges of simulations is to find the most informative reaction coordinates (13, 44, 129).

Fast Folding Proteins Unite Experiment and Computation

Some small protein domains fold/unfold in microseconds between just two macrostates (illustrated in **Figure 2*a***) or even downhill (**Figure 2*b***). Of course, disulfide bridges (35, 60), proline isomers (18), many types of intermediates (49), and domain interactions (66) can complicate the picture in general. Yet small, fast folders reveal the minimal requirements for folding and currently form the best link between experiment, theory, and simulation (129). **Figure 2*a*** illustrates the free-energy landscape two-state folding, with all highly populated conformations belonging to either the folded or the unfolded ensemble. These two ensembles are in local free-energy landscape minima, separated by a barrier that needs to be crossed by the protein in order for the protein to transition between the folded and unfolded ensembles. Experimentally, one hallmark of two-state folding is obtaining the same melting temperature (T_m) via different spectroscopic measurement techniques that each probe different parts of the energy landscape (86, 90).

Such behavior breaks down when intermediates are populated during the course of the folding (58, 116, 168) or in the scenario of downhill folding (**Figure 2*b***) (111). Downhill folding was predicted by energy landscape theory in the special case where decreasing enthalpy and entropy compensated throughout the whole reaction (163, 170). In addition to the thermodynamic observation of downhill folding (68), the gradual breakdown of timescale separation as downhill folding is approached (**Figure 1*d***) has also been seen kinetically (134, 168). Depending on initial conditions and protein stability modified by mutations, proteins can switch from downhill folding at low temperatures to two-state folding at temperatures close to T_m (110, 169) or from downhill folding to folding via intermediates (161). Two-state and downhill mechanisms are not common in large or multidomain globular proteins, where the so-called energy landscape roughness in the funnel (**Figures 1*c*** and ***2d***) is larger and where traps or intermediates occur in the free-energy landscape (10).

Friction and the Speed Limit of Folding

The scenario of downhill folding represents a case without kinetic barriers, when folding occurs at the "speed limit" (101). How fast can such proteins actually fold? The lack of a barrier between macrostates means that the rate of folding is limited by polypeptide chain diffusion and solvent friction (14, 25). Note that a process described as diffusion in a coarse-grained coordinate system can still involve many crossings over small barriers. For example, an individual backbone dihedral angle rotation occurs over a small barrier, but when many Ramachandran angles (130) are coarse grained into a few slower reaction coordinates, the fast motions can be treated as a friction-dependent prefactor in the Arrhenius equation $k_f \approx v^\dagger \exp[-\Delta G^\dagger/k_B T]$ (75). Here, $v^\dagger$ is the Arrhenius prefactor, $\Delta G^\dagger$ is the activation free energy, and k_f is the rate coefficient of the forward reaction. The exact scaling of the prefactor is still under debate (16, 62, 79).

Measurements on cytochrome c chain diffusion (28, 80) estimated a minimum time of ≈1 μs for the polypeptide chain to collapse. Secondary structure (helices, beta sheets) can form on a similar timescale (38, 39, 126). Studies using triplet energy transfer (62) have quantified chain length, location, and composition dependence for contact formation, ranging from 10 to 100 ns. Correspondingly large speed gains have been achieved for small proteins. An example of a protein mutated almost to the speed limit is the GTT variant of a WW domain (124), which was suggested by analysis of a long molecular dynamics trajectory. Another illustration is the three-helix bundle prb_{7-53}, in which the wild-type protein is mutated, replacing charged with hydrophobic residues. These computationally designed mutations again pushed folding down to ≈1 μs (172), close to the theoretical limit.

Trade-Off: Folding Versus Function

Globular proteins must attain a well-defined native structure in order to perform the requisite biological functions under specific environmental conditions (pH, temperature, solvent, salts). Even many disordered proteins fold upon binding (47, 150). The fast folders previously discussed show that evolution for function is an important factor that leads to frustrated folding. For example, when loop 1 of the Pin1 WW domain was truncated to speed up folding by an order of magnitude, the mutant lost the ability to bind to a phosphorylated target protein that is bound with high affinity (91). When the beta-bulge of interleukin-1β is replaced by a faster-folding β turn, protein function is again inhibited (74). Mutations that speed up folding—by making loops shorter, making helices more stable, or enlarging hydrophobic cores—are likely to eliminate charged residues needed for enzymatic function, to eliminate longer loops needed for binding, or to reduce flexibility needed for docking or substrate diffusion.

The possible explanation is that stabilizing mutations make the native structure too rigid compared with the wild-type, restricting it from sampling other conformational states that facilitate its binding. Low stability can even enhance function: According to the fly-casting mechanism (146), the unfolded form of the protein binds weakly at large distances, and folding and binding then go hand in hand. Functional proteins are an outcome of co-evolution between the need to fold and the need to perform functions (76).

The Diversity of Folding Pathways

There has been a long-standing discussion as to what extent proteins fold through sequential intermediates or parallel pathways; downhill or over obligatory barriers; and with or without traps (8, 48, 56, 154). The answer is all of the above! In vitro experiments, theory, and computation all agree that proteins have very shallow free-energy landscapes. Depths of valleys and heights of saddle points (barriers) are measured in tens of kJ/mol and not hundreds of kJ/mol as for chemical bond-making reactions. On such reaction surfaces, if their dimensionality is low but not necessarily equal to 1, many scenarios are possible. Nonetheless, a given mutant under a given solvent condition will almost always fold via a dominant pathway. In a typical experiment with a signal-to-noise ratio of 50:1, any additional pathways more than $\approx 4\ k_B T$ up in energy will simply be invisible (50). In simulations, such events will be rare and also hard to detect unless many folding/unfolding transitions can be sampled. For this reason, we have coined the phrase "apparent X-folder," where X is the mechanism of choice. Monitoring more reaction coordinates, going to higher free energy, or perturbing the system (e.g., temperature, pressure, solution conditions) will always reveal new paths and mechanisms (67, 97, 108, 109, 118, 169). A structural view complementary to the energetic view of the funnel picture builds on the original diffusion-collision model (95): Foldons (individual helices in the diffusion-collision model, later generalized to more complex cooperative folding units) can assemble sequentially, subject to off-pathway traps. Englander and colleagues (57) review these issues and the foldon picture in this volume in the context of nuclear magnetic resonance (NMR) hydrogen exchange experiments.

One case where alternative folding pathways become visible is for multirepeat proteins. Evidence of parallel folding pathways has been seen through the comparison of rates for symmetric consensus repeat proteins. Folding domains in parallel speeds up overall folding (1). The increase in folding rates with the chain length of the repeat protein stands in contrast to what is seen for globular proteins and is clear evidence of parallel folding. Perhaps not surprisingly, gene duplication is a key bootstrap for protein evolution.

Are these alternative paths important in general? This question deserves further investigation. Structure is more robust than mechanism, and this is why folded structure prediction is easier than

folding mechanism prediction (6). However, the very process of evolution that stabilizes native states versus higher-energy states while maximizing function may be the reason for alternative mechanisms and parallel paths. Evolution requires a certain flexibility, and digging too deep a funnel may reduce the evolvability of sequences (76). Appearance of new functionality upon mutation must eventually go hand in hand with a different folding mechanism and alternative folding pathways.

PROTEIN FOLDING IN COMPLEX SOLUTION ENVIRONMENTS

An important, though often overlooked, factor in protein folding is the surrounding environment. Above, we discussed the shallow free-energy landscape of proteins, which is easily perturbed by changes (e.g., temperature, pressure, mutations). Changes to the solvent in which proteins are contained can be thought of as another form of perturbation, which affects folding dynamics. Recall, for example, the importance of hydrophobic interactions in simplified models of protein folding (see the section titled Structural and Energetic Models for Folding). The interplay between the polypeptide chain and its surrounding water molecules creates the driving force for hydrophobic interactions. Perturbations to the properties of water (often called water chemical potential or activity) have a major impact on these interactions (27, 96). Other intra-protein interactions, including electrostatics and hydrogen bonds, are strongly affected by the solvent as well (7, 123, 143). In this section, we examine how changes to the aqueous environment owing to cosolute presence affect the protein folding landscape, linking in vitro studies to the complex solvation environment found in cells.

The one condition by which we define cosolutes is that they do not occupy a specific binding site on the protein of interest (as is the case, for example, of a substrate and an enzyme; see **Figure 3*a***). This broad category contains a plethora of molecules: from other proteins or nucleic-acid chains to electrolytes and small molecules. The hallmark of a cosolute effect is a linear dependence of folding free energy on cosolute concentration, as illustrated in **Figure 3*a***. For macromolecular cosolutes, the term quinary structure has been coined, and we return to it in our discussion of live cell folding in the section titled Protein Folding in the Cell. The importance of cosolutes and different solvents in industrial protein applications also cannot be overestimated (89).

Cosolutes Affect Protein Folding: Molecular Crowding, Osmolytes, and Salts

To consider a simplistic model, we can assume that cosolutes have no enthalpic interaction with our protein of interest. Instead, they are only sterically (entropically) repulsed from its surface and only take up space in the solution (3). This simple model has been extended to biology in what is commonly referred to as molecular crowding (55, 115, 171). Crowding acts to push proteins toward more compact states because these states take up less space, freeing up volume and increasing the configurational entropy for the entire system. Thus, upon addition of a crowding agent (often polymers such as Ficoll or dextran), the native state is stabilized. This scenario should not be confused with that of confinement, in which the system's volume is reduced in a fixed way (e.g., trapping the protein in a fixed volume) and has different consequences on protein structure (171). To quantify the effect of crowding, one needs simply to estimate the volumes of the species in the system and the volume changes upon folding to obtain a free-energy estimate (171). Molecular crowding is predictive mainly in cases where the cosolutes used are inert and relatively large (compared with the probe protein). The appearance of nonsteric interactions that are on the order of or larger than the steric repulsion effect causes the model to fail (98, 135). For example, recent evidence points to the problem of leaving out the solute (water) in crowding

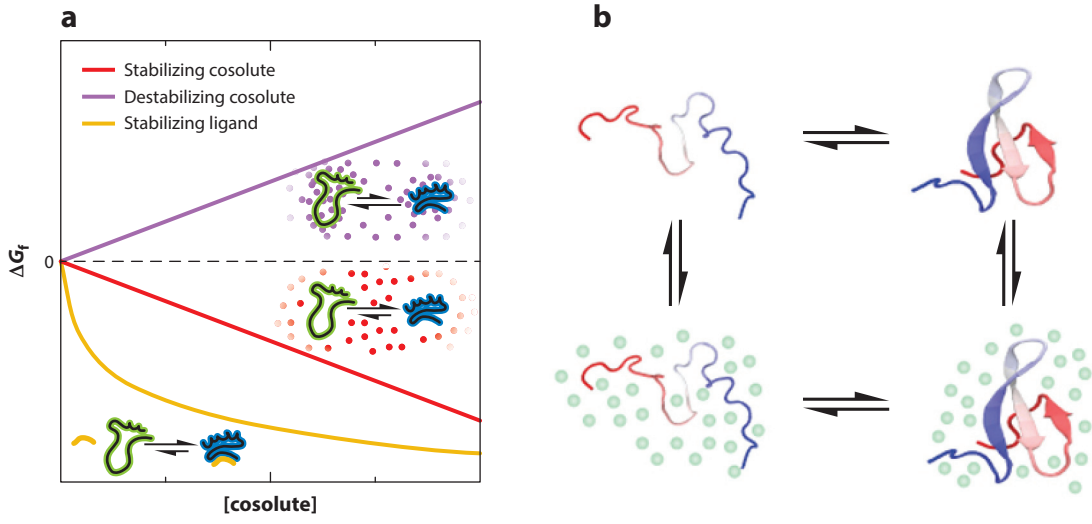

Figure 3

Protein folding in the presence of cosolutes. (*a*) Free energy of protein folding in the presence of destabilizing (denaturing) cosolute (*purple*), stabilizing cosolute (*red*), or ligand stabilizer (*orange*). The effect of cosolutes on folding free energy is linear in relation to cosolute concentration, even in concentrations that are orders of magnitude beyond those of the protein. This is because cosolutes do not have a single binding site—rather, they may bind the entire surface of the protein with weak affinity (for the case of denaturants) or be repelled from protein surface (for the case of stabilizing cosolutes). Ligands that have specific binding sites show a logarithmic dependence on concentration as a result of binding site saturation (37). (*b*) The Tanford transfer model. The free energy of solvation of the folded or unfolded ensembles in a cosolute solution can be determined from the folding free energy in the presence and absence of cosolutes, and vice versa.

calculations (144), although we believe that entropy-enthalpy compensation reduces the apparent problem, making the crowding process less sensitive to water. In any case, macromolecular crowding (volume exclusion) as a blanket description of any complex environment surrounding proteins is insufficient, and solvent effects and stickiness of the environment (via electrostatic and hydrophobic interactions), as well as hydrodynamics, have to be taken into account.

Osmolytes and salts are two naturally occurring cosolutes where molecular crowding models fail. Osmolytes—a family of molecules containing free amino acids, polyols, urea, and methylamine derivatives—exist ubiquitously in organisms spanning all kingdoms of life (167). They are extensively utilized by organisms that exist under high environmental stress (e.g., extreme temperature, pressure, osmolarity). In addition, high concentrations of osmolytes are found in cells that have high concentrations of denaturants such as urea (69). Osmolytes have recently been compared to inert crowders and have been found to affect protein folding through an enthalpically dominated mechanism [e.g., repulsive enthalpic interactions with side chain or backbone moieties (112)]. This is in contrast to inert polymers, which affect protein folding in an entropically dominated mechanism (e.g., steric repulsion) (136, 142, 152).

Small cations and anions are another ubiquitous component in the cell for which molecular crowding is nonpredictive. The presence of ions in the cellular milieu is highly regulated, and their intake is optimized by membrane channels (113). The effect of salts is compounded by the additive effect of electrostatic interactions and preferential interactions of the ion with the protein surface (132, 151). The effect of electrolytes also depends on the specific value and localization of charge in the folded and the unfolded ensemble. As concentrations increase, this effect diminishes owing to screening, and preferential interactions dominate salt effects on protein folding (99, 152).

Empirical and Theoretical Models of Cosolute Effects

The surface of globular proteins is often key for function (binding sites, substrate access). For protein folding, the effect of cosolute addition is linked to the change in solvent accessible surface area (SASA) between the folded and unfolded states. If we were to put this in a thermodynamic context, a convenient way to model the effect in terms of simple processes is given by the Tanford transfer model (120, 174), illustrated by the thermodynamic cycle in **Figure 3*b***. This cycle contains the interconversion between folded and unfolded protein in the presence and absence of cosolutes.

Yet this relatively simple picture is exacerbated by two factors. First, the surface area that changes upon protein folding is a complex entity. Amino acids contain many different moieties, which possess disparate chemical characteristics optimized for protein function at the surface, each of which has a different interaction with the solution components. Second, calculating the change in SASA requires a thermodynamic (by Boltzmann-weighted sampling) description of the conformations contained in the native and unfolded ensembles. To answer the first problem, several researchers have used model compounds to experimentally determine the free energy of transfer of different protein moieties from water under reference conditions (4, 99). The second problem can now be addressed owing to the advancement of molecular dynamics simulations. Extensive simulations can now give realistic representations of folded, unfolded, and perhaps even misfolded ensembles, facilitating calculations of the conformational free energy of interaction with respective solutions and subsequent thermodynamic cycles (22).

PROTEIN FOLDING IN THE CELL

To facilitate proper function, a cell must maintain an internal balance of metabolic, regulation, and transcriptional pathways. In addition, the cell must be able to maintain homeostasis in a changing environment. This is possible due to a complex network of regulation, which is carried out primarily by proteins, in response to internal and external signals. To this end, cells use a range of strategies to deal with deleterious environmental conditions: they contain specialized protein machines that store or unfold proteins for proper refolding (chaperones); they contain other protein machines that induce proteolysis of misfolded proteins (proteasomes and the ubiquitination mechanism) (139); and they take up or synthesize stabilizing osmolytes (discussed in the previous section) (166). Importantly, many factors in the cell, as well as in the cell's environment, will affect protein folding stability and speed, as illustrated in **Figure 4**.

Macromolecules are estimated to take up roughly 0.3–0.5 g per mL of cellular solution (173). Water content is roughly 70% of total cell mass (24). Proteins thus take up roughly half of the dry weight of the cell. DNA and lipids take up about a third of the same dry mass, and 10–15% belong to other molecules, mostly low-molecular mass species. The small percentage by mass of small molecules is deceptive because these molecules can exist at molar cellular concentrations.

As we zoom in on the local environment of a protein, surrounding solution components can vary dramatically owing to cytoskeleton-induced and organelle-induced local environments. From the complexities of their locale, it follows that proteins are surrounded by a staggeringly complex cellular environment (117). In an average mammalian cell, roughly 4 nm of water separate two adjacent proteins. This distance also contains other cosolutes, including electrolytes and metabolites. Because the diameter of a single hydration layer is on the order of the diameter of a water molecule, 4 nm contain ≈15 such layers—an exceedingly small number compared to many in vitro experiments. Water dynamics in those layers are modified by proteins anywhere from 0.2 to 2 nm from the protein surface, depending on the molecular property being examined (53, 81).

Many processes in the cell occur concomitantly with cellular shape change (122). These include the obvious cell cycle changes, which cause dramatic changes in cell composition and shape (162),

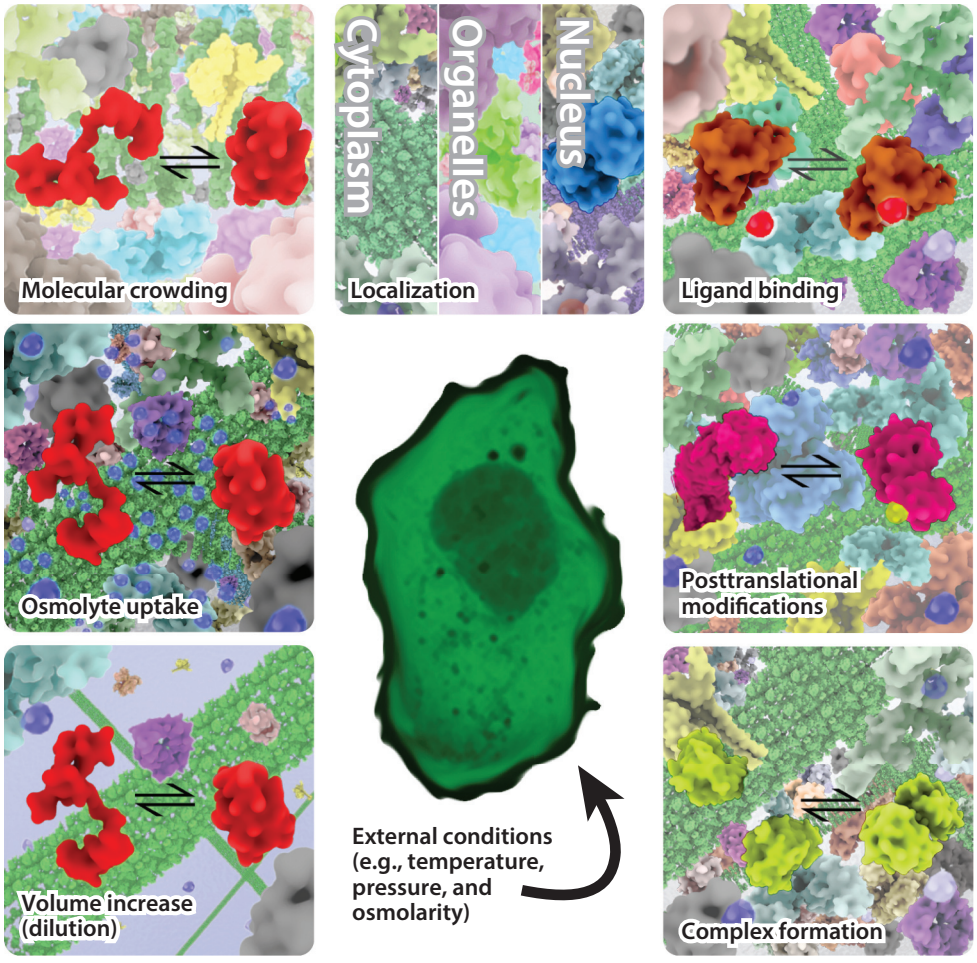

Figure 4

Factors that affect protein folding in the cell. The panels illustrate processes and effects that occur within a living cell and that affect protein function as well as folding stability and kinetics. Panels on the left depict cellular processes that affect folding, whereas panels on the right show protein reactions that occur in the cell and that affect the protein's folding equilibrium.

but other processes, such as migration, also cause cellular shape change (85). Such changes trickle down to local solution composition. In terms of kinetics, thermodynamic stability, and protein structure, these changes need not be dramatic to have an effect. As we saw in the previous section, a few k_BT suffice because many important proteins in the cell [including intrinsically disordered proteins (155) and transcription factors such as p53 (92)] are only marginally stable and because kinetic barriers for folding are small. Hence, changes to local solvent composition can have a very real effect on cellular function, in both normal and stress conditions.

Cellular Effects on Protein Folding and Interactions

How does folding happen in the cell? To a first approximation, it happens as it does in vitro. There are still cooperative folding curves (87) and similar kinetics (51). But to understand how a

cell modulates folding, adopting the viewpoint of a protein diffusing in the cytoplasm may be a useful exercise. Surrounding the protein are perturbed layers of water, interspersed with abundant dissolved ions, metabolites, sugars, signaling molecules, and short nucleic acids. Potassium, for example, exists in concentrations of ≈140 mM in the cytoplasm, making it very abundant in the cellular environment. At a distance roughly 10 water layers away are larger biomolecules such as other proteins at a high abundance. In this crowded environment, a protein must remain relatively inert to most solution components. Indeed, bioinformatics studies show a tendency for proteins to use less reactive amino acids to coat their surface in the native conformation (105). Evolution for specific interactions (e.g., signaling) must also evolve against the many nonspecific interactions that compete with a protein's interaction network.

The stability of the folded state of a protein in a mammalian cell milieu varies widely (41, 70, 147). At the lower end of this stability range (≈8 kJ/mol), over 5% of that protein's population at any given time is unfolded and subjected to misfolding that occurs during folding. This goes on for the lifetime of the protein, estimated to be between half an hour to several days (170), until it is sent to degradation. Cells had to develop complex machinery to monitor initial folding (33, 61), fold unfolded proteins [e.g., Hsp70 chaperones (83)], and degrade misfolded proteins (71).

The regulatory pathways tied to this machinery, termed collectively the unfolded protein response (UPR), are able to detect protein misfolding stress and act accordingly: slowing down or halting protein synthesis, increasing the specific synthesis of chaperone proteins, or uptaking or synthesizing osmolytes. In extreme cases, the UPR can initiate apoptosis, the self-destruction of the cell. Importantly, this machinery is initiated not only at times of duress but also during protein synthesis, as the nascent chain emerges from the ribosome (59, 128).

Experimental Techniques to Monitor Protein Folding In Situ

For decades, protein folding in the cell has been quantified using biochemical methods such as cross-linking and enzymatic digestion, followed by lysis and assaying. This type of methodology is invasive, is low in resolution, and cannot observe proteins in their natural environment. Today, new techniques enable minimally invasive observation inside cells of protein structure (147), of folding in real time (52) and with subcellular resolution (41), of protein stability (88), and even of single protein molecules (100).

One of the first methods to answer in-cell protein folding questions is the use of live cell NMR (147). Although this technique is technically challenging owing to the high concentrations of protein it requires, it has yielded interesting results ranging from gain of structure to decreased stability inside cells (65).

The explosion of fluorescence microscopy techniques has led to the most sensitive probes of protein dynamics in cells. Förster resonance energy transfer (FRET) (54) is utilized extensively today, from single-molecule protein folding experiments (9, 84, 140) to measurements of protein folding in living cells (51, 88). Robust and red probes that avoid auto-fluorescence of the cell have enabled even single-molecule FRET in living cells (100). These studies, together with the development of new fluorescent probes (15, 72), reveal a protein folding environment in the cell that is far from homogeneous (52). Folding thermodynamics and kinetics are affected not only by spatial localization (41) but also by temporal changes in cell cycle (162). In addition, protein identity plays a major role in determining whether it is stabilized or destabilized in the cell (12, 70, 78). The picture that emerges from these studies is that of a complex, heterogeneous, and dynamic system, where the solvent environment of a protein in the cell can control folding and activity.

SUMMARY AND OUTLOOK

With a firm basis of in vitro and computational studies now established, folding science can focus on questions such as the effect of residual structure in unfolded states and the effect of complex environments, including in the cell, on folding. The marginal stability of most proteins opens up control of folding in situ as a new area of study. The increased cross talk between protein science, computation, and cell biology will lead to a better understanding of how folding, function, and protein evolution are connected.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work is supported by National Institutes of Health (NIH) grant 2R01 GM093318-06 and National Science Foundation (NSF) grant NSF MCB 1413256 to M.G.

LITERATURE CITED

1. Aksel T, Barrick D. 2014. Direct observation of parallel folding pathways revealed using a symmetric repeat protein system. *Biophys. J.* 107:220–32
2. Anfinsen CB. 1973. Principles that govern the folding of protein chains. *Science* 181:223–30
3. Asakura S, Oosawa F. 1954. On interaction between two bodies immersed in a solution of macromolecules. *J. Chem. Phys.* 22:1255–56
4. Auton M, Bolen DW. 2005. Predicting the energetics of osmolyte-induced protein folding/unfolding. *PNAS* 102:15065–68
5. Baker D. 1998. Metastable states and folding free energy barriers. *Nat. Struct. Biol.* 5:1021–24
6. Baker D. 2000. A surprising simplicity to protein folding. *Nature* 405:39–42
7. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *PNAS* 98:10037–41
8. Baldwin RL. 1995. The nature of protein folding pathways: the classical versus the new view. *J. Biomol. NMR* 5:103–9
9. Banerjee PR, Deniz AA. 2014. Shedding light on protein folding landscapes by single-molecule fluorescence. *Chem. Soc. Rev.* 43:1172–88
10. Barrick D. 2009. What have we learned from the studies of two-state folders, and what are the unanswered questions about two-state protein folding? *Phys. Biol.* 6:015001
11. Beauchamp KA, McGibbon R, Lin Y-S, Pande VS. 2012. Simple few-state models reveal hidden complexity in protein folding. *PNAS* 109:17807–13
12. Benton LA, Smith AE, Young GB, Pielak GJ. 2012. Unexpected effects of macromolecular crowding on protein stability. *Biochemistry* 51:9773–75
13. Best RB, Hummer G. 2005. Reaction coordinates and rates from transition paths. *PNAS* 102:6732–37
14. Best RB, Hummer G. 2010. Coordinate-dependent diffusion in protein folding. *PNAS* 107:1088–93
15. Boersma AJ, Zuhorn IS, Poolman B. 2015. A sensor for quantification of macromolecular crowding in living cells. *Nat. Meth.* 12:227–29
16. Borgia A, Wensley BG, Soranno A, Nettels D, Borgia MB, et al. 2012. Localizing internal friction along the reaction coordinate of protein folding by combining ensemble and single-molecule fluorescence spectroscopy. *Nat. Commun.* 3:1195
17. Bowman GR, Voelz VA, Pande VS. 2011. Atomistic folding simulations of the five helix bundle protein λ_{6-85}. *J. Am. Chem. Soc.* 133:664–67

18. Brandts JF, Halvorson HR, Brennan M. 1975. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* 14:4953–63
19. Brooks BR, Brooks CL, MacKerell AD, Nilsson L, Petrella RJ, et al. 2009. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30:1545–614
20. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. 1995. Funnels, pathways, and the energy landscape of protein-folding: a synthesis. *Proteins* 21:167–95
21. Cabrita LD, Dobson CM, Christodoulou J. 2010. Protein folding on the ribosome. *Curr. Opin. Struct. Biol.* 20:33–45
22. Camilloni C, Rocco AG, Eberini I, Gianazza E, Broglia RA, Tiana G. 2008. Urea and guanidinium chloride denature protein L in different ways in molecular dynamics simulations. *Biophys. J.* 94:4654–61
23. Cavalli A, Haberthür U, Paci E, Caflisch A. 2003. Fast protein folding on downhill energy landscape. *Protein Sci.* 12:1801–3
24. Cayley S, Lewis BA, Guttman HJ, Record MT. 1991. Characterization of the cytoplasm of *Escherichia coli* K-12 as a function of external osmolarity. Implications for protein-DNA interactions in vivo. *J. Mol. Biol.* 222:281–300
25. Chahine J, Oliveira RJ, Leite VB, Wang J. 2007. Configuration-dependent diffusion can shift the kinetic transition state and barrier height of protein folding. *PNAS* 104:14646–51
26. Chan HS, Dill KA. 1991. Polymer principles in protein structure and stability. *Annu. Rev. Biophys. Biophys. Chem.* 20:447–90
27. Chandler D. 2005. Interfaces and the driving force of hydrophobic assembly. *Nature* 437:640–7
28. Chang AC, Chuang SS, Gray M, Soong Y. 2003. In-situ infrared study of CO_2 adsorption on SBA-15 grafted with γ-(aminopropyl)triethoxysilane. *Energy Fuels* 17:468–73
29. Cho SS, Levy Y, Wolynes PG. 2006. P versus Q: Structural reaction coordinates capture protein folding on smooth landscapes. *PNAS* 103:586–91
30. Cho Y, Zhang X, Pobre KF, Liu Y, Powers DL, et al. 2015. Individual and collective contributions of chaperoning and degradation to protein homeostasis in *E. coli*. *Cell Rep.* 11:321–33
31. Chodera JD, Mobley DL. 2013. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.* 42:121–42
32. Chung HS, McHale K, Louis JM, Eaton WA. 2012. Single-molecule fluorescence experiments determine protein folding transition path times. *Science* 335:981–84
33. Clark PL. 2004. Protein folding in the cell: reshaping the folding funnel. *Trends Biochem. Sci.* 29:527–34
34. Comellas G, Rienstra CM. 2013. Protein structure determination by magic-angle spinning solid-state NMR, and insights into the formation, structure, and stability of amyloid fibrils. *Annu. Rev. Biophys.* 42:515–36
35. Creighton T, Goldenberg D. 1984. Kinetic role of a meta-stable native-like two-disulphide species in the folding transition of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* 179:497–526
36. Curtiss LA, Raghavachari K, Redfern PC, Pople JA. 1997. Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation. *J. Chem. Phys.* 106:1063–79
37. Daniels KG, Tonthat NK, McClure DR, Chang YC, Liu X, et al. 2014. Ligand concentration regulates the pathways of coupled protein folding and binding. *J. Am. Chem. Soc.* 136:822–25
38. Davis CM, Xiao S, Raleigh DP, Dyer RB. 2012. Raising the speed limit for β-hairpin formation. *J. Am. Chem. Soc.* 134:14476–82
39. De Sancho D, Best RB. 2011. What is the time scale for α-helix nucleation? *J. Am. Chem. Soc.* 133:6809–16
40. DeGrado WF, Summa CM, Pavone V, Nastri F, Lombardi A. 1999. De novo design and structural characterization of proteins and metalloproteins. *Annu. Rev. Biochem.* 68:779–819
41. Dhar A, Girdhar K, Singh D, Gelman H, Ebbinghaus S, Gruebele M. 2011. Protein stability and folding kinetics in the nucleus and endoplasmic reticulum of eucaryotic cells. *Biophys. J.* 101:421–30
42. Dill KA. 1985. Theory for the folding and stability of globular proteins. *Biochemistry* 24:1501–9
43. Dill KA, MacCallum JL. 2012. The protein-folding problem, 50 years on. *Science* 338:1042–46
44. Dror RO, Dirks RM, Grossman JP, Xu H, Shaw DE. 2012. Biomolecular simulation: a computational microscope for molecular biology. *Annu. Rev. Biophys.* 41:429–52

45. Duan Y, Kollman PA. 1998. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 282:740–44
46. Dunitz JD. 1995. Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem. Biol.* 2:709–12
47. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, et al. 2001. Intrinsically disordered protein. *J. Mol. Graph. Model.* 19:26–59
48. Dyer RB. 2007. Ultrafast and downhill protein folding. *Curr. Opin. Struct. Biol.* 17:38–47
49. Dyson HJ, Rance M, Houghten RA, Lerner RA, Wright PE. 1988. Folding of immunogenic peptide fragments of proteins in water solution: I. Sequence requirements for the formation of a reverse turn. *J. Mol. Biol.* 201:161–200
50. Dyson HJ, Wright PE. 2002. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Adv. Protein Chem.* 62:311–40
51. Ebbinghaus S, Dhar A, McDonald D, Gruebele M. 2010. Protein folding stability and dynamics imaged in a living cell. *Nat. Methods* 7:319–23
52. Ebbinghaus S, Gruebele M. 2011. Protein folding landscapes in the living cell. *J. Phys. Chem. Lett.* 2:314–19
53. Ebbinghaus S, Kim SJ, Heyden M, Yu X, Heugen U, et al. 2007. An extended dynamical hydration shell around proteins. *PNAS* 104:20749–52
54. Edelhoch H, Brand L, Wilchek M. 1967. Fluorescence studies with tryptophyl peptides. *Biochemistry* 6:547–59
55. Ellis RJ, Minton AP. 2003. Cell biology: Join the crowd. *Nature* 425:27–28
56. Englander SW. 2000. Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.* 29:213–38
57. Englander SW, Mayne L, Kan Z-Y, Hu W. 2016. Protein folding—how and why: by hydrogen exchange, fragment separation, and mass spectrometry. *Annu. Rev. Biophys.* 45:135–52
58. Englander SW, Sosnick TR, Mayne LC, Shtilerman M, Qi PX, Bai Y. 1998. Fast and slow folding in cytochrome c. *Acc. Chem. Res.* 31:737–44
59. Evans MS, Clark TF, Clark PL. 2005. Conformations of co-translational folding intermediates. *Protein Pept. Lett.* 12:189–95
60. Fass D. 2012. Disulfide bonding in protein biophysics. *Annu. Rev. Biophys.* 41:63–79
61. Fedyukina DV, Cavagnero S. 2011. Protein folding at the exit tunnel. *Annu. Rev. Biophys.* 40:337–59
62. Fierz B, Kiefhaber T. 2007. End-to-end versus interior loop formation kinetics in unfolded polypeptide chains. *J. Am. Chem. Soc.* 129:672–79
63. Fleming KG. 2014. Energetics of membrane protein folding. *Annu. Rev. Biophys.* 43:233–55
64. Frauenfelder H, Sligar S, Wolynes P. 1991. The energy landscapes and motions of proteins. *Science* 254:1598–603
65. Freedberg DI, Selenko P. 2014. Live cell NMR. *Annu. Rev. Biophys.* 43:171–92
66. Freire E, Murphy KP, Sanchez-Ruiz JM, Galisteo ML, Privalov PL. 1992. The molecular basis of cooperativity in protein folding. Thermodynamic dissection of interdomain interactions in phosphoglycerate kinase. *Biochemistry* 31:250–56
67. Fuentes EJ, Wand AJ. 1998. Local stability and dynamics of apocytochrome b_{562} examined by the dependence of hydrogen exchange on hydrostatic pressure. *Biochemistry* 37:9877–83
68. Garcia-Mira MM, Sadqi M, Fischer N, Sanchez-Ruiz JM, Muñoz V. 2002. Experimental identification of downhill protein folding. *Science* 298:2191–95
69. Garcia-Perez A, Burg MB. 1991. Renal medullary organic osmolytes. *Physiol. Rev.* 71:1081–115
70. Ghaemmaghami S, Oas TG. 2001. Quantitative protein stability measurement in vivo. *Nat. Struct. Mol. Biol.* 8:879–82
71. Glickman MH, Ciechanover A. 2002. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* 82:373–428
72. Gnutt D, Gao M, Brylski O, Heyden M, Ebbinghaus S. 2014. Excluded-volume effects in living cells. *Angew. Chem. Int. Ed.* 54:2548–51
73. Gō N. 1983. Theoretical studies of protein folding. *Annu. Rev. Biophys. Bioeng.* 12:183–210

74. Gosavi S, Whitford PC, Jennings PA, Onuchic JN. 2008. Extracting function from a β-trefoil folding motif. *PNAS* 105:10384–89
75. Gruebele M. 1999. The fast protein folding problem. *Annu. Rev. Phys. Chem.* 50:485–516
76. Gruebele M. 2005. Downhill protein folding: Evolution meets physics. *C. R. Biol.* 328:701–12
77. Guo M, Xu Y, Gruebele M. 2012. Temperature dependence of protein folding kinetics in living cells. *PNAS* 109:17863–67
78. Guzman I, Gelman H, Tai J, Gruebele M. 2014. The extracellular protein VlsE is destabilized inside cells. *J. Mol. Biol.* 426:11–20
79. Hagen SJ. 2010. Solvent viscosity and friction in protein folding dynamics. *Curr. Protein Pept. Sci.* 11:385–95
80. Hagen SJ, Hofrichter J, Szabo A, Eaton WA. 1996. Diffusion-limited contact formation in unfolded cytochrome c: estimating the maximum rate of protein folding. *PNAS* 93:11615–17
81. Halle B. 2004. Protein hydration dynamics in solution: a critical survey. *Philos. Trans. R. Soc. Lond. Ser. B.* 359:1207–23
82. Hansmann UH. 1997. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* 281:140–50
83. Hartl FU, Hayer-Hartl M. 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295:1852–58
84. Hoffmann A, Nettels D, Clark J, Borgia A, Radford SE, et al. 2011. Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Phys. Chem. Chem. Phys.* 13:1857–71
85. Hoffmann EK, Lambert IH, Pedersen SF. 2009. Physiology of cell volume regulation in vertebrates. *Physiol. Rev.* 89:193–277
86. Huang GS, Oas TG. 1995. Structure and stability of monomeric lambda repressor: NMR evidence for two-state folding. *Biochemistry* 34:3884–92
87. Ignatova Z, Gierasch LM. 2004. Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling. *PNAS* 101:523–28
88. Ignatova Z, Gierasch LM. 2006. Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant. *PNAS* 103:13357–61
89. Iyer PV, Ananthanarayan L. 2008. Enzyme stability and stabilization—aqueous and non-aqueous environment. *Process Biochem.* 43:1019–32
90. Jackson SE, Fersht AR. 1991. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* 30:10428–35
91. Jäger M, Zhang J, Bieschke J, Nguyen H, Dendle G, et al. 2006. The structure-function-folding relationship in a WW domain. *PNAS* 108:10648–53
92. Johnson CR, Morin PE, Arrowsmith CH, Freire E. 1995. Thermodynamic analysis of the structural stability of the tetrameric oligomerization domain of p53 tumor suppressor. *Biochemistry* 34:5309–16
93. Karplus, M. 2011. Behind the folding funnel diagram. *Nat. Chem. Biol.* 7:401–4
94. Karplus M, Sali A, Shakhnovich E. 1995. Kinetics of protein folding. *Nature* 373:664–65
95. Karplus M, Weaver DL. 1979. Diffusion–collision model for protein folding. *Biopolymers* 18:1421–37
96. Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* 14:1–63
97. Kim SJ, Matsumura Y, Dumont C, Kihara H, Gruebele M. 2009. Slowing down downhill folding: a three-probe study. *Biophys. J.* 97:295–302
98. Kim YC, Mittal J. 2013. Crowding induced entropy-enthalpy compensation in protein association equilibria. *Phys. Rev. Lett.* 110:208102
99. Knowles DB, Lacroix AS, Deines NF, Shkel I, Record MT Jr. 2011. Separation of preferential interaction and excluded volume effects on DNA duplex and hairpin stability. *PNAS* 108:12699–704
100. König I, Zarrine-Afsar A, Aznauryan M, Soranno A, Wunderlich B, et al. 2015. Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat. Methods* 12:773–79
101. Kubelka J, Hofrichter J, Eaton WA. 2004. The protein folding 'speed limit.' *Curr. Opin. Struct. Biol.* 14:76–88

102. Lau KF, Dill KA. 1989. A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules* 22:3986–97
103. Levinthal C. 1968. Are there pathways for protein folding? *J. Chim. Phys.* 65:44
104. Levitt M, Warshel A. 1975. Computer simulation of protein folding. *Nature* 253:694–98
105. Levy ED, De S, Teichmann SA. 2012. Cellular crowding imposes global constraints on the chemistry and evolution of proteomes. *PNAS* 109:20461–66
106. Li Z, Yang Y, Zhan J, Dai L, Zhou Y. 2013. Energy functions in de novo protein design: current challenges and future prospects. *Annu. Rev. Biophys.* 42:315–35
107. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. 2011. How fast-folding proteins fold. *Science* 334:517–20
108. Liu F, Du D, Fuller AA, Davoren JE, Wipf P, et al. 2008. An experimental survey of the transition between two-state and downhill protein folding scenarios. *PNAS* 105:2369–74
109. Liu F, Gao YG, Gruebele M. 2010. A survey of λ repressor fragments from two-state to downhill folding. *J. Mol. Biol.* 397:789–98
110. Liu F, Gruebele M. 2007. Tuning $\lambda_{6\text{-}85}$ towards downhill folding at its melting temperature. *J. Mol. Biol.* 370:574–84
111. Liu J, Campos LA, Cerminara M, Wang X, Ramanathan R, et al. 2012. Exploring one-state downhill protein folding in single molecules. *PNAS* 109:179–84
112. Liu Y, Bolen D. 1995. The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry* 34:12884–91
113. MacKinnon R. 2003. Potassium channels. *FEBS Lett.* 555:62–65
114. McConkey EH. 1982. Molecular evolution, intracellular organization, and the quinary structure of proteins. *PNAS* 79:3236–40
115. Minton AP, Wilf J. 1981. Effect of macromolecular crowding upon the structure and function of an enzyme: glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* 20:4821–26
116. Muñoz V. 2002. Folding plasticity. *Nat. Struct. Biol.* 9:792–94
117. Nelson DR. 2012. Biophysical dynamics in disorderly environments. *Annu. Rev. Biophys.* 41:371–402
118. Nguyen H, Jäger M, Moretto A, Gruebele M, Kelly JW. 2003. Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. *PNAS* 100:3948–53
119. Noé F, Schütte C, Vanden-Eijnden E, Reich L, Weikl TR. 2009. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *PNAS* 106:19011–16
120. O'Brien EP, Ziv G, Haran G, Brooks BR, Thirumalai D. 2008. Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model. *PNAS* 105:13403–8
121. Onuchic JN, Wolynes PG, Luthey-Schulten Z, Socci ND. 1995. Toward an outline of the topography of a realistic protein-folding funnel. *PNAS* 92:3626–30
122. Pedersen SF, Hoffmann EK, Mills JW. 2001. The cytoskeleton and cell volume regulation. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 130:385–99
123. Pegram LM, Record MT Jr. 2008. Thermodynamic origin of Hofmeister ion effects. *J. Phys. Chem. B* 112:9428–36
124. Piana S, Sarkar K, Lindorff-Larsen K, Guo M, Gruebele M, Shaw DE. 2011. Computational design and experimental testing of the fastest-folding beta-sheet protein. *J. Mol. Biol.* 405:43–48
125. Pitera JW, Swope WC, Abraham FF. 2008. Observation of noncooperative folding thermodynamics in simulations of 1BBL. *Biophys. J.* 94:4837–46
126. Poland D, Scheraga HA. 1966. Phase transitions in one dimension and the helix-coil transition in polyamino acids. *J. Chem. Phys.* 45:1456–63
127. Portman JJ, Takada S, Wolynes PG. 1998. Variational theory for site resolved protein folding free energy surfaces. *Phys. Rev. Lett.* 81:5237–40
128. Powers ET, Morimoto RI, Dillin A, Kelly JW, Balch WE. 2009. Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem.* 78:959–91
129. Prigozhin MB, Gruebele M. 2013. Microsecond folding experiments and simulations: A match is made. *Phys. Chem. Chem. Phys.* 15:3372–88
130. Ramakrishnan C, Ramachandran G. 1965. Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. *Biophys. J.* 5:909–33

131. Rath A, Deber CM. 2012. Protein structure in membrane domains. *Annu. Rev. Biophys.* 41:135–55
132. Record MT Jr., Anderson CF, Lohman TM. 1978. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* 11:103–78
133. Roy S, Hecht MH. 2000. Cooperative thermal denaturation of proteins designed by binary patterning of polar and nonpolar amino acids. *Biochemistry* 39:4603–7
134. Sabelko J, Ervin J, Gruebele M. 1999. Observation of strange kinetics in protein folding. *PNAS* 96:6031–36
135. Sapir L, Harries D. 2015. Is the depletion force entropic? Molecular crowding beyond steric interactions. *Curr. Opin. Colloid Interface Sci.* 20:3–10
136. Saunders AJ, Davis-Searles PR, Allen DL, Pielak GJ, Erie DA. 2000. Osmolyte-induced changes in protein conformational equilibria. *Biopolymers* 53:293–307
137. Saunders MG, Voth GA. 2013. Coarse-graining methods for computational biology. *Annu. Rev. Biophys.* 42:73–93
138. Scheraga HA, Khalili M, Liwo A. 2007. Protein-folding dynamics: overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.* 58:57–83
139. Schroder M, Kaufman RJ. 2005. The mammalian unfolded protein response. *Annu. Rev. Biochem.* 74:739–89
140. Schuler B, Eaton WA. 2008. Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* 18:16–26
141. Schuler B, Sorrano A, Hofmann H, Nettels D. 2016. Single-molecule FRET spectroscopy and the polymer physics of unfolded and intrinsically disordered proteins. *Annu. Rev. Biophys.* 45:207–31
142. Senske M, Törk L, Born B, Havenith M, Herrmann C, Ebbinghaus S. 2014. Protein stabilization by macromolecular crowding through enthalpy rather than entropy. *J. Am. Chem. Soc.* 136:9036–41
143. Sharp KA. 1995. Polyelectrolyte electrostatics: salt dependence, entropic, and enthalpic contributions to free energy in the nonlinear Poisson-Boltzmann model. *Macromolecules* 36:227–43
144. Sharp, KA. 2015. Analysis of the size dependence of macromolecular crowding shows that smaller is better. *PNAS* 112:7990–95
145. Shea JE, Brooks CL III. 2001. From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys. Chem.* 52:499–535
146. Shoemaker BA, Portman JJ, Wolynes PG. 2000. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *PNAS* 97:8868–73
147. Smith AE, Zhang Z, Pielak GJ, Li C. 2015. NMR studies of protein folding and binding in cells and cell-like environments. *Curr. Opin. Struct. Biol.* 30:7–16
148. Snow CD, Nguyen H, Pande VS, Gruebele M. 2002. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* 420:102–6
149. Southall NT, Dill KA, Haymet ADJ. 2002. A view of the hydrophobic effect. *J. Phys. Chem. B* 106:521–33
150. Sugase K, Dyson HJ, Wright PE. 2007. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447:1021–25
151. Sukenik S, Boyarski Y, Harries D. 2014. Effect of salt on the formation of salt-bridges in β-hairpin peptides. *Chem. Comm.* 50:8193–96
152. Sukenik S, Sapir L, Gilman-Politi R, Harries D. 2013. Diversity in the mechanisms of cosolute action on biomolecular processes. *Faraday Discuss.* 160:225
153. Thirumalai D, Lorimer GH. 2001. Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 30:245–69
154. Udgaonkar JB. 2008. Multiple routes and structural heterogeneity in protein folding. *Annu. Rev. Biophys.* 37:489–510
155. Uversky VN. 2009. Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J.* 28:305–25
156. Uversky VN, Oldfield CJ, Dunker AK. 2008. Intrinsically disordered proteins in human diseases: introducing the D^2 concept. *Annu. Rev. Biophys.* 37:215–46

157. Voelz VA, Singh VR, Wedemeyer WJ, Lapidus LJ, Pande VS. 2010. Unfolded state dynamics and structure of protein L characterized by simulation and experiment. *J. Am. Chem. Soc.* 132:4702–9
158. Wagner A. 2005. *Robustness and Evolvability in Living Systems.* Princeton, NJ: Princeton Univ. Press
159. Weathers EA, Paulaitis ME, Woolf TB, Hoh JH. 2004. Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein. *FEBS Lett.* 576:348–52
160. Weber G. 1995. van't Hoff revisited: enthalpy of association of protein subunits. *J. Phys. Chem.* 99:1052–59
161. Wirth AJ, Liu Y, Prigozhin MB, Schulten K, Gruebele M. 2015. Comparing fast pressure jump and temperature jump protein folding experiments and simulations. *J. Am. Chem. Soc.* 137:7152–59
162. Wirth AJ, Platkov M, Gruebele M. 2013. Temporal variation of a protein folding energy landscape in the cell. *J. Am. Chem. Soc.* 135:19215–21
163. Wolynes PG, Luthey-Schulten Z, Onuchic JM. 1996. Fast-folding experiments and the topography of protein folding energy landscapes. *Chem. Biol.* 3:425–32
164. Woodside MT, Block SM. 2014. Reconstructing folding energy landscapes by single-molecule force spectroscopy. *Annu. Rev. Biophys.* 43:19–39
165. Xia K, Zhang S, Bathrick B, Liu S, Garcia Y, Colon W. 2012. Quantifying the kinetic stability of hyperstable proteins via time-dependent SDS trapping. *Biochemistry* 51:100–7
166. Yancey PH. 2001. Water stress, osmolytes and proteins. *Integr. Comp. Biol.* 41:699–709
167. Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN. 1982. Living with water stress: evolution of osmolyte systems. *Science* 217:1214–22
168. Yang WY, Gruebele M. 2003. Folding at the speed limit. *Nature* 423:193–97
169. Yang WY, Gruebele M. 2004. Folding λ-repressor at its speed limit. *Biophys. J.* 87:596–608
170. Yen H-CS, Xu Q, Chou DM, Zhao Z, Elledge SJ. 2008. Global protein stability profiling in mammalian cells. *Science* 322:918–23
171. Zhou HX, Rivas G, Minton AP. 2008. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* 37:375–97
172. Zhu Y, Fu X, Wang T, Tamura A, Takada S, et al. 2004. Guiding the search for a protein's maximum rate of folding. *Chem. Phys.* 307:99–109
173. Zimmerman SB, Trach SO. 1991. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J. Mol. Biol.* 222:599–620
174. Ziv G, Haran G. 2009. Protein folding, protein collapse, and Tanford's transfer model: lessons from single-molecule FRET. *J. Am. Chem. Soc.* 131:2942–47
175. Zuckerman DM. 2011. Equilibrium sampling in biomolecular simulation. *Annu. Rev. Biophys.* 40:41–62

Annual Review of Biophysics
Volume 45, 2016

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Biophysics* articles may be found at http://www.annualreviews.org/errata/biophys