Quinary protein structure and the consequences of crowding in living cells: Leaving the test-tube behind

Anna Jean Wirth[1)] *and Martin Gruebele*[1)2)3)*]

Although the importance of weak protein-protein interactions has been understood since the 1980s, scant attention has been paid to this quinary structure. The transient nature of quinary structure facilitates dynamic sub-cellular organization through loose grouping of proteins with multiple binding partners. Despite our growing appreciation of the quinary structure paradigm in cell biology, we do not yet understand how the many forces inside the cell – the excluded volume effect, the stickiness of the cytoplasm, and hydrodynamic interactions – perturb the weakest functional protein interactions. We discuss the unresolved problem of how the forces in the cell modulate quinary structure, and to what extent the cell has evolved to exert control over the weakest biomolecular interactions. We conclude by highlighting the new experimental and computational tools coming on-line for in vivo studies, which are a critical next step if we are to understand quinary structure in its native environment.

Keywords:
association; in vivo; protein folding; signaling

Introduction

The complexity of the cell – jam-packed with macromolecules, inherently heterogeneous, and always dynamic – has led to widespread recognition that in vitro studies of proteins cannot always provide physiologically relevant structural and functional information [1]. The momentum of protein science towards probing proteins in their native environments is leading to a departure from the canonical protein purification-to-characterization workflow towards more in vivo studies. But which aspects of protein structure, folding, and function are most likely to be modulated by the cell, by how much, and to what adaptive advantage, if any?

With the caveat that living cells are never in equilibrium, thermodynamics tells us "how much" modulation is significant: interaction of a protein with its environment must alter occupancy of different protein populations significantly. A disordered protein chain poised to fold from a state "U" (unfolded) to a state "F" (folded) upon binding a partner will switch from 80% U to 80% F if binding favors "F" over "U" by just 7 kJ/mol. By comparison, a single carbon-carbon bond in one amino acid of that protein requires 50 times more bond enthalpy to break. Teasing apart all of the interactions between a protein of interest and the cellular water, ions, macromolecules, and other co-solutes is difficult. Indeed, studying protein dynamics in vivo at all remains challenging.

The cell has evolved mechanisms to modulate proteins from the energy scale of primary structure (~370 kJ/mol amide bond enthalpy) to quaternary structure (20 kJ/mol hemoglobin-hemoglobin enthalpy increase in normal erythrocytes vs. sickle cells [2]). Biomolecule populations depend not just on enthalpy, but also on entropy: reactions producing greater disorder are favored. Thus, many biomolecular reactions have free energy changes (combining enthalpy and entropy effects) of just 10s of kJ/mol. For example, as strong as amide bonds are, they actually break spontaneously in acidic water, releasing about 10 kJ/mol of free energy [3]: the broken CN amide bond is partly made up for by forming NH and COH bonds, and on top of that, the two liberated peptides get to move about more randomly.

What about even weaker interactions – weaker than two protein surfaces binding or forming quaternary structure? Could such extremely weak interactions play a meaningful role in the cell? These transient protein-protein interactions

DOI 10.1002/bies.201300080

[1)] Department of Chemistry, University of Illinois, Urbana, IL, USA
[2)] Department of Physics, University of Illinois, Urbana, IL, USA
[3)] Center for Biophysics and Computational Biology, University of Illinois, Urbana, IL, USA

***Corresponding author:**
Martin Gruebele
E-mail: gruebele@scs.uiuc.edu

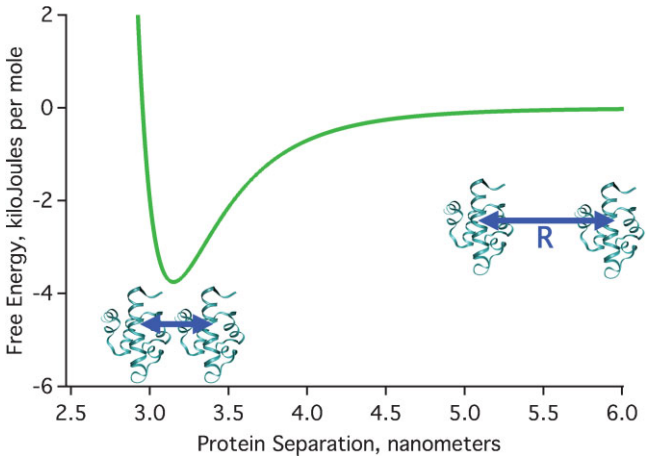

Figure 1. The interaction free energy for two lambda repressor fragment proteins, derived by us from small angle X-ray scattering measurements that look at how much proteins tend to cluster together as their concentration is increased. These proteins do not form quaternary structure or tightly bound dimers. Nonetheless, at a center-to-center separation R of 3.2 nanometers, the proteins "sticking" to each other with a small free energy release of 4 kJ/mol.

are collectively known as "quinary structure". Since the 1980s, quinary structure has been implicated in many cellular processes ranging from cellular metabolism [4] to protein translation [5]. In a few cases, illustrated in Fig. 1, the strength of quinary interactions has been measured quantitatively in vitro [6]. Their strength is difficult to quantify in vitro in ways that realistically mimic the interior of cells. Consequently, quinary structure is an area of research where in vivo study is warranted. Only recently have techniques become available to probe these weak transient interactions directly and quantitatively in the cell. With whole cell computational modeling becoming more feasible, the potential even exists to bridge computation and experiment.

The aim of this Problems & Paradigms essay is to provide a perspective on when and how quinary structure in the cell affects folding and function of biomolecules, and especially proteins. First, we will review forces at work in the cell. Then, we discuss quinary structure and examples where it may play a role. Finally, we will discuss the newest advances in in vivo experimental methodology and whole cell modeling and speculate on the future of the field.

The forces at work in the intracellular environment

The intracellular environment is highly crowded – macromolecular concentrations reach up to 400 g/L (or mg/mL) [1]. Such crowding by complex mixtures of biological molecules contributes only modestly to stabilization or destabilization of proteins, but when proteins interact very weakly to begin with, the consequences of crowding likely play an important role. We briefly review here how crowding, "stickiness" of the cytoplasm, hydrodynamic forces, and water dynamics manifest themselves from the macromolecular to the cellular level.

How does macromolecular crowding affect protein-protein interactions and protein folding through the excluded volume effect?

Rather than arising from attractive intermolecular interactions, the excluded volume effect depends on the size and concentration of molecules crowding the solution. In Fig. 1, at about 3.2 nm, the two proteins "touch": a third protein simply could not squeeze through between them. If they separate by less than the diameter of the third protein, it will still not fit through, and the volume between the pair is excluded. Only when they separate by enough (pair on the right) is the volume between the two proteins no longer excluded.

The main consequence of excluded volume is that the larger and less ordered form of a protein or protein complex – such as the unfolded state or a non-associated dimer – is entropically disfavored because of the lack of available space [7]. Figure 2A demonstrates how the more compact state of a protein is favored under crowded conditions. The excluded volume for a given protein in solution is defined as any region that protein's center of mass cannot occupy [8]. As the radius of gyration of a protein increases (such as would occur from a transition from the folded to unfolded state) the excluded volume increases, and the apparent protein concentration in the left-over accessible volume is greater than the actual concentration.

The energetic consequences of protein folding are summarized in Fig. 2B. As more volume is excluded, the unfolded state has no place to fit, and becomes less stable, and the unfolded protein crosses the barrier more easily and thus folds more rapidly. This behavior is entirely driven by the entropy term in the free energy relation. Likewise, a compact protein complex is favored over the separated protein monomers – unless the leftover volumes between crowding molecules are so small that the complex just cannot fit.

The role of crowding on protein structure and stability has been studied extensively in vitro and in silico, and experiment verifies the predictions of theory [9–13]. The role that the excluded volume effect plays in biological reactions in general was recently reviewed [14]. In the heterogeneous environment of the cytoplasm, a given protein will be subject to a wide range of excluded volumes of different shapes.

An approximate free energy range for the effect can be assigned based on experimental and theoretical results. A computational study found that at high crowder concentration, the excluded volume effect increased the stability of the folded state by about 1–6 kJ/mol and protein-protein interactions by 2.5–8.5 kJ/mol [15]. Similarly, in a model of a bacterial cell, it was found that the stabilization of the folded state in the crowded cytoplasm due to the excluded volume effect was about 4–8 kJ/mol and for protein-protein interactions 4–5 kJ/mol [16]. Experimentally, the folded state of FRET-labeled yeast phosphoglycerate kinase was stabilized by 2 kJ/mol when 200 mg/mL Ficoll was added to the solution (Fig. 2C) [12].

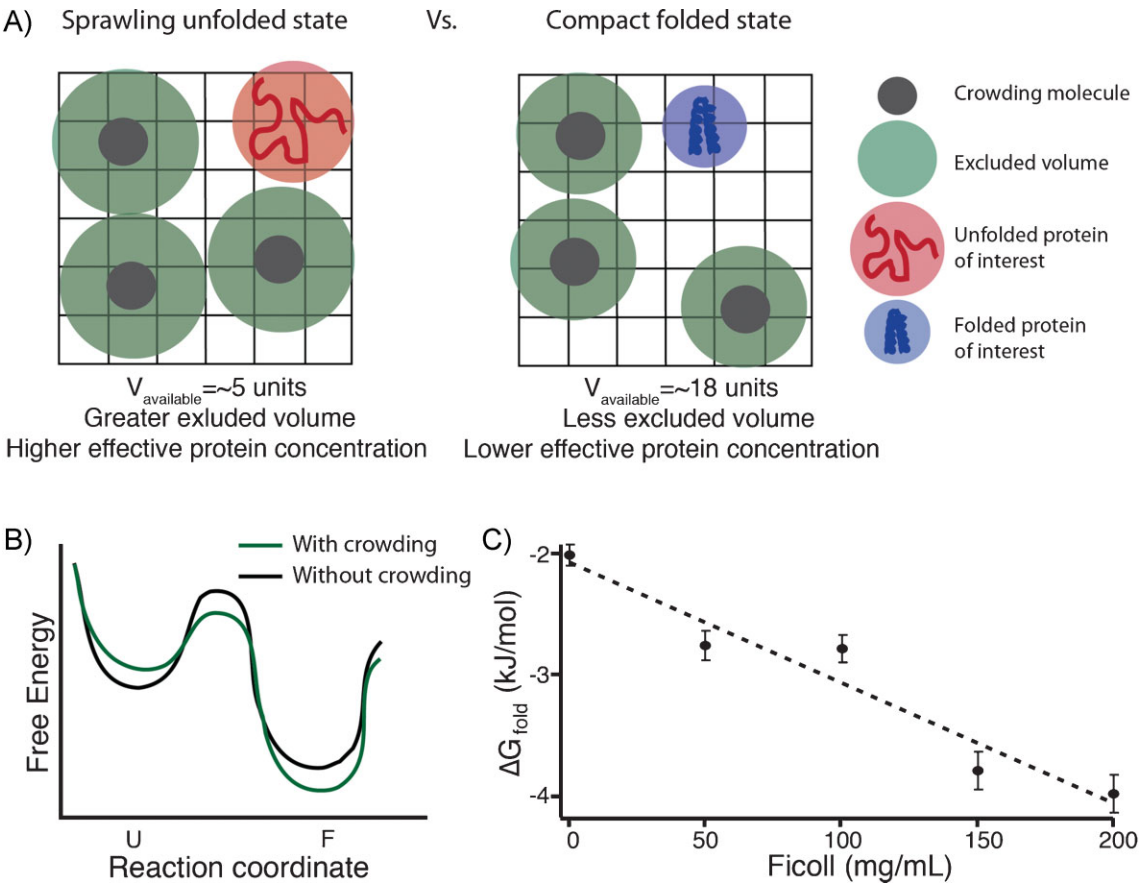

Figure 2. The excluded volume and its energetic consequences. **A:** The compact, folded state of a protein has less excluded volume than the less ordered unfolded state even though the size and concentration of crowders is the same. The excluded volume (green) is defined as any region of space that the protein's center of mass cannot occupy. The apparent protein concentration is greater for the unfolded state. **B:** Increasing crowding decreases the stability of the unfolded state, and decreases the activation barrier for folding. **C:** Experimentally measured free energy of folding ΔG_{fold} for FRET-labeled yeast phosphoglycerate kinase (PGK) as macromolecular crowding is increased. The protein is stabilized as crowding increases, illustrating the net stabilization of the folded state due to the excluded volume effect [12]. ΔG_{fold} is calculated at a reference temperature of 37°C.

Can the "stickiness" of the cytoplasm modulate protein-protein interaction and protein folding in a sequence dependent manner?

The excluded volume effect predicts size-specific effects, but there are also sequence-specific effects. Non-specific interactions between proteins in the crowded cellular environment, either through electrostatics (due to charges) or hydrophobicity (mainly the protein's effect on solvent water disorder), make proteins stick to each other or other macromolecules. This makes the apparent viscosity of the cytoplasm larger than bulk water's viscosity [17].

The diffusion coefficient depends inversely on solution viscosity. Consequently, most experimental efforts to measure cytoplasmic viscosity have relied on diffusion measurements [18]. This effort is hampered, however, by anomalous diffusion – basically a time-dependent diffusion coefficient – which abounds in the cell due to non-uniformity of the environment [19]. In one striking example that highlights the challenges of using translational diffusion to directly measure cytoplasmic viscosity, the translational diffusion of chromosomes has been shown to be sensitive to the metabolic activity level of the cell [20].

Rotational diffusion, on the other hand, has been shown to be almost exclusively sensitive to localized intermolecular interactions [21]. Nuclear magnetic resonance (NMR) spectroscopy is highly sensitive to the tumbling of molecules (and thus can extract the rotational diffusion coefficient). Wang et al. [21] found that rotational diffusion coefficients for a test protein varied widely between solutions containing inert crowders, which only increase the excluded volume, and those containing protein crowders, which can stick non-specifically to the protein of interest. Another group used a whole cell NMR approach to measure the rotational diffusion of three globular proteins in *Escherichia coli* cells and found that although the proteins studied had similar size and fold, they showed significantly different rotational diffusion [17]. This result suggests that stickiness of the cytoplasm is sequence-dependent.

Computation is not far behind. McGufee and Elcock [16] used Brownian dynamics simulations to show that specific protein-protein binding interactions were destabilized consistently by cytoplasmic stickiness, while its effect on protein

folding varied significantly from about 10 kJ/mol of destabilization to the same amount of stabilization. MD simulations of proteins in crowded, "sticky", solutions have shown that the enthalpic consequences of many protein-protein interactions may destabilize the fold of a protein and cause accumulation of non-native and partially denatured states [22]. The effect of specific interactions, on the other hand, are thought to impart general stabilization to a protein fold [23].

What happens when water and other molecules get dragged around by proteins inside the cell?

Hydrodynamic interactions arise from collisions between macromolecules and solvent molecules and affect the kinetics of protein-protein binding. At physiological temperature, collisions with water molecules impart energy fluctuations of ± 0.04 kJ/mol to an average-size protein. This may seem like a small effect, and when these collisions between solute and solvent are completely random, they give rise to the familiar Brownian motion.

In a crowded environment, however, motions of particles are correlated, and collisions between protein and solute can result in non-random motion. Figure 3 illustrates how hydrodynamic interactions influence the binding of two proteins. As a protein moves towards its partner, the water molecules between the two proteins must be pushed out of the interfacial region, carrying away energy. Some of those displaced water molecules will also bump into the partner protein, pushing it away. Together, these two hydrodynamic effects slow down the kinetics of protein-protein binding [24].

The energetic magnitude of hydrodynamic interactions on proteins – i.e. the reduction or increase of the transition state barrier due to hydrodynamic interactions – is again relatively small. Theory predicts that rates of association should decrease by $<30\%$ depending on the surface charge of the molecules in question [25]. Simulations also have explored the role that hydrodynamic interactions can play in protein folding and protein-protein association kinetics. Frembgen-Kesner and Elcock [26] showed that the association rate of proteins decreased by 30–80% in the presence of hydrodynamic interactions. The same researchers also found that inclusion of hydrodynamics increases the folding rate of proteins 1.5- to 3-fold, corresponding to at most 3 kJ/mol [27]. Moving towards simulating hydrodynamic effects in cytoplasmic models is extremely challenging. However, Ando et al. [28] found that inclusion of hydrodynamic interactions in a model *E. coli* cytoplasm along with the excluded volume effect replicated the experimentally measured translational and rotational diffusion of GFP.

Is there any bulk water left inside the cell?

Water often mediates protein-protein interactions, or can be thought of as part of the protein structure. Water molecules inside protein cavities form on average three hydrogen bonds, each providing a stabilization of 2.5 kJ/mol [29]. Of more interest for water-mediated protein interactions is the intermediate regime between bulk solvent water and the "frozen", long-resident time structural water molecules. This intermediate regime is often called the solvent-shell of the protein and consists of water molecules that are transiently bound to the protein surface [30], as well as water molecules up to several nanometers away whose hydrogen bonds rearrange more slowly than in the bulk because of the influence of the protein [31]. Although NMR experiments indicate some presence of bulk-like water [32], THz experiments that probe further from protein surfaces indicate that the hydrogen bond rearrangement time is retarded up to 2 nm away from protein surfaces [33]. At typical packing densities in the cell, the separation between macromolecules is only a few nanometers, leaving little room for any bulk water inside the cell by the THz criterion.

Figure 4 shows an example of a protein-protein interaction that is mediated by water. Transiently bound water molecules have been proposed to mediate recognition during protein-protein binding and to contribute to the stability of the interaction [34]. Experiments have used double mutant cycles, which compare the overall loss of stability from binding site mutations that remove hydrogen bond donors or acceptors, to elucidate the energetic contribution of water-mediated protein interactions to the overall stability of protein complexes. Stabilization, if any, ranges from 0 to about 4 kJ/mol [35–37].

Quinary structure

Quinary structure is the fifth level of protein structural organization

The hierarchy of protein structure and its relation to energy and free energy is shown in Fig. 5. Quinary structure is the most fleeting, and thus the most likely to be affected by the weak forces acting at the macromolecule to cell level.

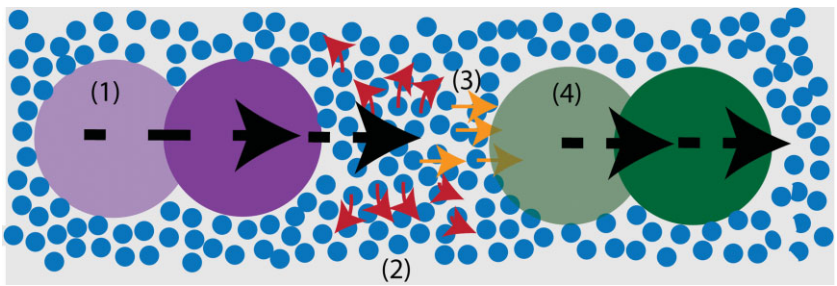

Figure 3. Cartoon depiction of the effect of hydrodynamic interactions on the kinetics of bimolecular association. (1) As one protein (purple) moves towards its binding partner (green), it displaces the water molecules between the two proteins. (2) The water molecules must move from the region between the proteins and reorganize (red arrows). (3) Some water molecules are displaced towards the binding partner (orange arrows) and push the binding partner away. (4) This hydrodynamic repulsion further slows the kinetics of association.

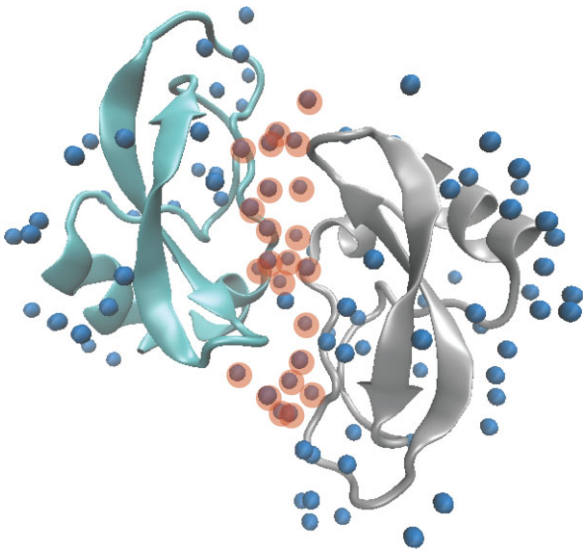

Figure 4. Structure of protease inhibitor domain of the amyloid beta-protein precursor with associated waters (pdb 1AAP [77]). The density of water molecules at the dimer interface (highlighted in orange) illustrates a water-mediated protein-protein interaction.

The term quinary structure was introduced by McConkey [5] to define a fifth level of protein structural organization: protein-protein interactions that he described as "inherently transient". Quinary structure, he argued, could be disrupted by the isolation methods of the time, and consequently had not received the recognition it deserved. We discuss quinary structure in terms of three criteria: functionality, thermodynamics, and kinetics.

The quinary structure terminology is intermittently used in the literature, but in our – and others [1] – opinion it deserves to be explicitly recognized. Any individual quinary interaction may not confer a large functional advantage to the cell. Yet even small differences can serve as the basis for evolutionary selection, and a large number of them could add up to a significant advantage. Indeed, many interactions that eventually evolve to tight binding may start out highly transient, while others remain weak to be optimal for the cell.

Very low thermodynamic stability has been used as a criterion for quinary structure (i.e. K_d values $>1\,\mu M$ for complex formation [17]). However, some interactions have variable affinity. They can be extremely weak unless the proper ligand assists the interaction [38]. Such "three body" effects abound in biology. One particularly important example is the binding of the molecular chaperone DnaK to an unfolded protein at low affinity when ATP is bound, but with high affinity following hydrolysis of ATP to ADP [39]. McConkey does not exclude such variable protein-protein interactions from his original definition. One of the examples he gives of quinary structure – calmodulin binding to actin fibers in interphase cells [40] – is a variable affinity protein-protein interaction modulated by calcium binding. Thus, in our definition here, quinary structure includes constant low affinity protein-protein interactions, as well as the low affinity state of variable affinity protein-protein interactions.

McConkey's definition of "inherently transient" implies rapid kinetics in addition to low stability. Only when both criteria are fulfilled is an interaction "inherently transient." Consider two states: protein dimer versus two monomers. If these two states have similar free energy, so neither is very stable with respect to the other, their connecting barrier could still be high, trapping the system for a long time in one state or the other. If the two states are connected by a small barrier, but one state is much lower in free energy than the other, the system will simply end up in the lower state, and no function associated with switching states can occur. Only when stability and barriers are both low, is the system inherently transient.

We can use the thermodynamic, kinetic, and functional criteria to define quinary structure as inherently transient (none of the states is thermodynamically favored, nor are any two states separated by a large barrier that traps population), yet having some functional specificity subject to evolutionary selection. Thus, quinary interactions between proteins or other biomolecules must be on the order of a few kJ/mol on the free energy scale: larger free energy differences would lock in the lowest energy population, and larger barriers would make each state, once formed, rather long lived. This low free energy scale makes quinary interaction subject to modulation by the forces in the cell acting on the macromolecular to whole-cell scale, as discussed in the previous section.

What is not a quinary structure?

Interactions between water molecules are fleeting; they may even be functionally important. However, there is no functional specificity that can be selected by evolution. Water is just water and cannot evolve like a protein. Obligatory protein-protein interactions form strongly bound thermodynamically stable structures. An example is the arc repressor dimer, which as a monomer is unstable, but as a dimer very stable [38, 41]. Another example is the intrinsically disordered protein P27 that gains structure when it tightly binds and inhibits various Cyclin/Cdk complexes throughout the cell cycle [42]. These types of interactions will not be classified as quinary because they result in tight binding and/or can be assigned to lower levels of protein structure: quaternary for the dimerization of the arc repressor and tertiary for the folding of p27.

What is quinary structure and what does it do?

Quinary structure has been implicated in a number of cellular processes from cell signaling to metabolomics. Its inherent transience allows quinary structure to facilitate dynamic spatial organization of macromolecules in the cytoplasm. It can facilitate loose groupings of proteins when they are working together, but not otherwise. In other examples, quinary structure can enable a single protein to serve multiple functions depending on its binding partner or serve as a dose-dependent molecular switches. Very recent work has presented empirical and theoretical evidence that a protein in a specific interaction stabilizes its partner's fold and hypothesizes that the protein interaction network is of critical

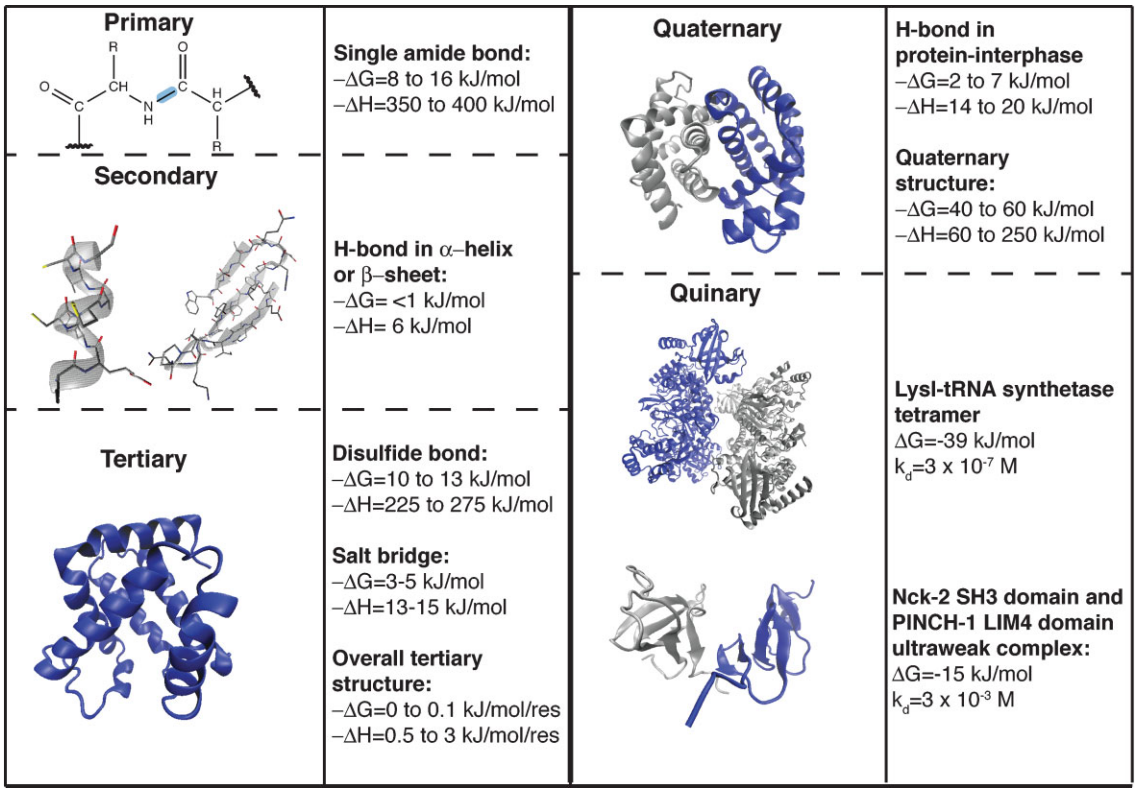

Figure 5. Structural and energy scales (absolute values around 37°C) in the cell in terms of enthalpy change ΔH and free energy change ΔG. Enthalpy describes the strength of interactions, while free energy describes which interactions are most likely by also including the effect of disorder. For example, amide bonds are very strong, but nonetheless hydrolyze in water because new NH and OH bonds are formed and because the two peptide fragments have higher disorder. Primary structure [3, 78], secondary structure [79, 80], tertiary structure [12, 81–84], quaternary structure [85, 86], and two examples of quinary structure [47, 87] are shown. PDB structures (in order of appearance): 1EG3 [88], 1HHO [89], 1U55 [90], 3BJU [47].

importance to the overall thermodynamic stability of the proteome [23].

The most famous and well-studied example of quinary structure is the metabolon, a multi-enzyme complex associated with multi-step metabolic pathways. These pathways produce numerous intermediates, ~80% of which have no use in the cell other than to be fed to the next processing enzyme. It would be entropically very unfavorable for a cell to support a random distribution of these molecules and their processing enzyme throughout the entire cytoplasm [4]. Metabolons alleviate this problem by "substrate channeling": intermediates along a metabolic pathway are kept from dispersing throughout the cytoplasm because they are passed between enzymes associated into quinary structures (thus effectively localizing reactions [43].

The protein biosynthesis pathway is one metabolic process that features such quinary structure and is an example of quinary structure providing both spatial organization and enabling protein multifunctionality. Aminoacyl-tRNA synthetases, the family of enzymes which aminacylate tRNA molecules, were first observed in the 1980s to form high molecular weight complexes of then-unknown function [44]. The multisynthetase complexes appear to play multiple roles depending on cellular conditions and since their original discovery have been associated with several processes from substrate channeling in protein biosynthesis to cytokine release in the inflammation response [45]. The exact mode of assembly of these complexes is debated, but many of the protein-protein interactions in the complex are thought to be quite transient [46].

One component of the complex that is involved in quinary structure is the lysl-tRNA synthetase (LysRS). Guo et al. [47] showed that the normally dimeric LysRS forms a weakly associated ($K_d \sim 1\,\mu M$) homotetramer that competes with the formation of the larger multi-synthetase complex. Because release of LysRS from the complex is thought to be a trigger for the proinflammatory response, competition between these two equilibria could allow for LysRS to shuttle between two functions. Thus, in this case, quinary structure allows for protein multi-functionality.

One of the exciting new frontiers of quinary structure is the intracellular formation of membrane-free macromolecular structures that arise because of phase separation [48]. Weak association between proteins at sufficiently high concentration has been shown to drive phase separation of proteins into a protein-rich and protein-poor liquid phase in vitro [49]. Such protein-rich droplets behave as a separate liquid phase – that is, they are embedded in but distinct from the surrounding cytoplasm, and they can merge with each other to form larger droplets. These phase-separated regions without membranes have been observed recently in vivo, first as the P-granules in germ-line cells, and most recently with signalling related

proteins [50, 51]. P-granules are an example of quinary structure where interactions driving the phase transition derive from transient RNA-protein binding [50].

The recent observation of phase transitions in actin-signaling associated proteins, is another example that relies on transient protein-protein interactions only. Li et al. [51] demonstrated that the SH3 domain and its ligand, the PRM domain, form liquid droplets in vitro and in vivo and found that the ability to form multivalent associations is necessary for phase separation. Moving to the nephrin-NCK-N-WASP system, a signaling pathway associated with actin stimulation, they showed that this complex forms liquid droplets in vitro when nephrin is phosphorylated and that the downstream signaling activity sharply increases as phase separation occurs. This example of quinary structure is remarkable because it could provide another degree of cytoplasmic structure – one that is far more spatially and temporally dynamic than membrane bound organelles.

When the phase transition is tied to a downstream response, quinary structure enables a dose-dependent, nearly binary switch for the signaling pathway. Only when a sufficient concentration of nephrin is phosphorylated will the phase transition occur, but when the threshold is reached the transition occurs rapidly thus "turning on" the downstream signaling pathway. Such a dose dependent response mediated by quinary structure has also been seen in signaling pathways associated with the immune system [52].

In addition to metabolism and signaling, quinary structure can also affect the folding stability and kinetics of proteins. For example, Denos et al. [53] investigated the folding of low concentrations of lambda repressor embedded in a crowding matrix of SubL, a thermophilic protein that remains stable while lambda repressor is thermally unfolded. Protein-protein interaction raises the melting temperature of lambda repressor by 13°C. This change is large enough that lambda repressor could switch from 80% unfolded to 80% folded simply by adding the crowder. The folding time of lambda repressor is $<100\,\mu s$, so this is a highly transient equilibrium. Some proteins that are intrinsically disordered in vitro (IDPs) may undergo a similar equilibrium shift inside cells, promoted by only transient interactions [54], rather than the tight binding interactions that we exclude (like p27).

The key take-home message is that quinary structure abounds in cells, and produces far more fine-grained structure inside cells than was originally suspected. Cells are very far from random mixtures of macromolecules. Even in regions that are not bounded by membranes, compartmentalization exists.

Leaving the test-tube behind

Experimental advances for detecting quinary structure in vivo are beginning to yield quantitative information on transient protein-protein interaction in its native environment

Numerous advances in the past decade have allowed for measurement of protein-protein interactions inside cells; these are thoroughly reviewed elsewhere [55]. A few standouts have great potential to yield useful data on the weakest and more transient protein-protein interactions.

For example, a recent solid-state NMR study of frozen cells produced atomic resolution protein structural information, and as the technique further develops, protein-protein interactions and structural changes of proteins in vivo may be characterized in great detail [56]. With improving isotopic labeling schemes and pulse sequences, whole cell NMR of living cells also promises to soon be applicable to eukaryotic cell types [57]. Similarly, mass spectrometry has been used to probe how the interior of bacterial cells changes protein stability [58], and could be expanded to eukaryotic cells.

Super-resolution imaging techniques, which provide imaging capabilities below the diffraction limit, have already provided rich information on the co-localization of proteins [59]. One of the more recent and spectacular advances in this family of techniques is improved temporal and spatial resolution in living cells using STORM in both 2 and 3 dimensions [60]. As time resolution improves even further, such imaging capabilities will undoubtedly allow direct detection of even transiently bound protein complexes.

Spatial and temporal patterns of protein stability and folding kinetics modulated by quinary structure can now be imaged inside living cells using Fast Relaxation Imaging (FReI). Ebbinghaus and Gruebele [61] review micro-patterned stability of the enzyme phosphoglycerate kinase inside mammalian cells. Similar to in vitro observations [12], crowding in the cell increases the average stability of the enzyme, but in addition, microenvironments that persist for at least minutes differ in protein stability and folding rate, as shown in Fig. 6. The folding free energy landscape of the protein is modulated by several kJ/mol inside the cell. It remains to be seen whether such modulation has functional significance.

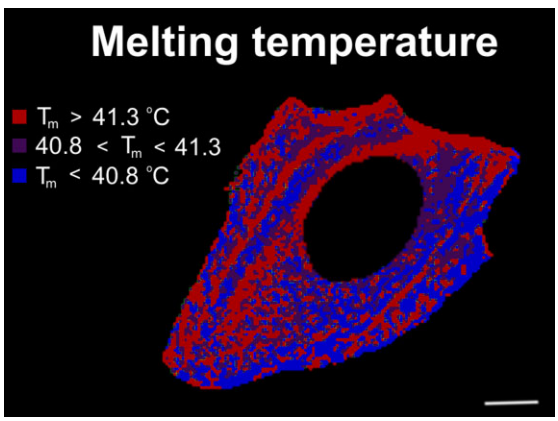

Figure 6. Microenvironments with different protein stabilities inside a U2OS cell (thermodynamic data from [61], Fig. 2). Stability is characterized by the melting temperature T_m, at which 50% of the protein is folded, 50% unfolded. Stability patterns extend from the resolution limit (2 μm) to 10s of μm in size. The scale bar is 10 μm in length.

NMR, mass spectrometry, and imaging techniques are not the only areas in which progress has been made. Biochemistry has also offered new solutions for detecting quinary structure in vivo. M-track uses a "bait" and "prey" methylation-based approach to detect very short-lived interactions between proteins in yeast [62]. Another technique, PICT, uses a "bait" and "prey" approach paired with fluorescence live cell spectroscopy to detect protein-protein interactions in living cells [63].

How can computation and experiment work together to study quinary structure?

Only now are simulations of large protein complexes and even whole cells becoming feasible, allowing comparison with the new array of experiments probing transient structures inside the cell [16, 64, 65]. By combining computation, theory, and experiment we can attempt to dissect how each of the various forces at play in the intracellular environment affects a specific protein of interest. For the case of excluded volume and cytoplasmic stickiness, theory and experiment are beginning to paint a clear picture of how these two effects compete against one another in the cell. Several theoretical studies have predicted that attractive interactions between crowders and proteins forming a complex (an enthalpic effect) can counteract entropic stabilization of the complex by the excluded volume effect [66, 67]. Jiao et al. [68] demonstrated this principle for protein association in mixtures of attractive polymers and found that they could tune the stability of a protein complex by using polymers with stronger and weaker attractions to the protein of interest. Simulation studies have shown that moderate attraction between proteins and crowding molecules, on the order of 1 kJ/mol, can counteract the excluded volume effect [69].

The competition of stabilizing and destabilizing effects can also be seen in protein folding inside cells. Dhar and coworkers [70] directly measured both protein folding thermodynamics and kinetics in U2OS cells. Using a FReI, they found that the folded state of yeast phosphoglycerate kinase was stabilized by 6.3 kJ/mol in vivo compared to in vitro. On the other hand, Ignatova and Gierash [71] monitored the unfolding of CRABP I in living bacterial cells and found a net destabilization of 0.8 kJ/mol compared to in vitro. In an update to their original work, the same group found that although thermodynamic stability between in vitro and in vivo is not significant, the unfolding rate of CRABP I was significantly faster in vivo [72]. Philip et al. [73] used a similar FRET construct to directly measure the binding dynamics between β-lactamase to its inhibitor in HeLa cells and found that association was 0.8–1.5 kJ/mol destabilized in vivo compared to in vitro. These studies highlight a wide range of behaviors. The cell can both stabilize and destabilize proteins in a sequence-specific manner. Because of this sequence dependence, proteins that are energetically similar in vitro might be quite different in vivo. The holy grail for simulation will be to predict from protein sequence the effect that the intracellular environment will exert on specific protein-protein interactions and protein folding reactions. In particular, as the field of whole cell modeling continues to improve, it may be possible to predict the effect that a changing intracellular environment has on complex protein-protein interaction networks.

Conclusions and prospects

The energy and distance scales of the forces that arise in the crowded intracellular environment are such that we can expect them to influence modestly protein tertiary and quaternary structure and to affect significantly quinary structure. The role that quinary structure plays in cell signaling, subcellular localization of proteins, metabolomics, and other systems is clearly significant, and studying these weak protein interactions in their native environment will become increasingly important if we are to fully understand their role in biology. Experimental and computational advances will likely soon make these types of studies more routine and allow for careful consideration of how the interplay between intracellular forces and protein structure fits into the biology of the cell.

One intriguing possibility is to consider that structure could be dynamically tuned by the changing microenvironments inside the cell, in a yet unexplored mode of protein regulation. Is it conceivable that the physical characteristics of the cytoplasm could change substantially enough to make quinary structure an important regulatory mechanism inside the cell?

For one cellular process, the cell cycle, it certainly is. During the procession of events leading to division into two genetically identical daughter cells, the nucleus dissolves during mitosis and dumps its highly charged chromosomes into the cytoplasm. Furthermore, the dry mass, or total biomolecular content of a cell, does not increase at the same rate throughout the cell cycle as the overall cell volume, so the intracellular protein density may also fluctuate in a cell cycle-dependent fashion [74, 75]. Finally, Hinde et al. [76] measured diffusion across the nucleus and showed that diffusion is limited during certain stages of the cell cycle by changing degrees of chromatin compaction. This result implies that the cell-cycle environment is extremely dynamic.

Such a strikingly changing environment during the cell cycle invites speculation that the interplay between the physical forces in the cell and quinary structure could be actively exploited by the cell for the purposes of process regulation. Indeed, the tightly regulated series of events that comprise the cell cycle are controlled by an incredibly complex, and still not fully elucidated, network of protein-protein interactions – some of which may be sensitive to environmental fluctuation. As more biophysical studies emerge that examine protein structure and protein-protein interactions across dynamic cellular processes, exploration of potentially new types of protein regulation will become tractable, and the in vivo structure-function relationship more completely understood.

Acknowledgements

AJW is supported by an NSF Graduate Research Fellowship under grant number DGE-1144245. MG thanks the Eiszner

family for support of this research. We thank Dr. S. J. Kim for the calculation in Figure 1, based on reference 6. The research was funded by National Science Foundation grant MCB 1019958.

References

1. **Gierasch LM, Gershenson A**. 2009. Post-reductionist protein science, or putting Humpty Dumpty back together again. *Nat Chem Biol* **5**: 774–7.
2. **Kuczera K, Gao J, Tidor B, Karplus M**. 1990. Free energy of sickling. A simulation analysis. *Proc Natl Acad Sci USA* **87**: 8481–5.
3. **Martin RB**. 1998. Free energies and equilibria of peptide bond hydrolysis and formation. *Biopolymers* **45**: 351–3.
4. **Srere PA**. 1987. Complexes of sequential metabolic enzymes. *Annu Rev Biochem* **56**: 89–124.
5. **McConkey EH**. 1982. Molecular evolution, intracellular organization, and the quinary structure of proteins. *Proc Natl Acad Sci USA* **79**: 3236–40.
6. **Kim SJ, Dumont C, Gruebele M**. 2008. Simulation-based fitting of protein-protein interaction potentials to SAXS experiments. *Biophys J* **94**: 4924–31.
7. **Minton AP**. 1981. Excluded volume as a determinant of macromolecular structure and reactivity. *Biopolymers* **20**: 2093–120.
8. **Minton AP**. 2001. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J Biol Chem* **276**: 10577–80.
9. **Cheung MS, Klimov D, Thirumalai D**. 2005. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc Natl Acad Sci USA* **102**: 4753–8.
10. **Perham M, Stagg L, Wittung-Stafshede P**. 2007. Macromolecular crowding increases structural content of folded proteins. *FEBS Lett* **581**: 5065–9.
11. **Minton AP**. 2005. Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: macromolecular crowding and protein stability revisited. *Biophys J* **88**: 971–85.
12. **Dhar A, Samiotakis A, Ebbinghaus S, Nienhaus L**, et al. 2010. Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. *Proc Natl Acad Sci USA* **107**: 17586–91.
13. **Hong JA, Gierasch LM**. 2010. Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J Am Chem Soc* **132**: 10445–52.
14. **Zhou H-X, Rivas G, Minton AP**. 2008. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* **37**: 375–97.
15. **Qin S, Zhou H-X**. 2009. Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability. *Biophys J* **97**: 12–9.
16. **McGuffee SR, Elcock AH**. 2010. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput Biol* **6**: e1000694.
17. **Wang Q, Zhuravleva A, Gierasch LM**. 2011. Exploring weak, transient protein–protein interactions in crowded in vivo environments by in-cell nuclear magnetic resonance spectroscopy. *Biochemistry* **50**: 9225–36.
18. **Verkman AS**. 2002. Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem Sci* **27**: 27–33.
19. **Dix JA, Verkman AS**. 2008. Crowding effects on diffusion in solutions and cells. *Annu Rev Biophys* **37**: 247–63.
20. **Weber SC, Spakowitz AJ, Theriot JA**. 2012. Nonthermal ATP-dependent fluctuations contribute to the in vivo motion of chromosomal loci. *Proc Natl Acad Sci USA* **109**: 7338–43.
21. **Wang Y, Li C, Pielak GJ**. 2010. Effects of proteins on protein diffusion. *J Am Chem Soc* **132**: 9392–7.
22. **Harada R, Tochio N, Kigawa T, Sugita Y**, et al. 2013. Reduced native state stability in crowded cellular environment due to protein-protein interactions. *J Am Chem Soc* **135**: 3696–701.
23. **Dixit PD, Maslov S**. 2013. Evolutionary capacitance and control of protein stability in protein-protein interaction networks. *PLoS Comput Biol* **9**: e1003023.
24. **Bamford C, Tipper C, Compton R**. 1985. *Diffusion-Limited Reactions*. Elsevier. p. 213–53.
25. **Wolynes PG**. 1976. Slip boundary conditions and the hydrodynamic effect on diffusion controlled reactions. *J Chem Phys* **65**: 450.
26. **Frembgen-Kesner T, Elcock AH**. 2010. Absolute protein-protein association rate constants from flexible, coarse-grained Brownian dynamics simulations: the role of intermolecular hydrodynamic interactions in Barnase-Barstar association. *Biophys J* **99**: L75–7.
27. **Frembgen-Kesner T, Elcock AH**. 2009. Striking effects of hydrodynamic interactions on the simulated diffusion and folding of proteins. *J Chem Theory Comput* **5**: 242–56.
28. **Ando T, Skolnick J**. 2013. On the importance of hydrodynamic interactions in lipid membrane formation. *Biophys J* **104**: 96–105.
29. **Williams MA, Goodfellow JM, Thornton JM**. 1994. Buried waters and internal cavities in monomeric proteins. *Protein Sci* **3**: 1224–35.
30. **Timasheff SN**. 2002. Protein hydration, thermodynamic binding, and preferential hydration. *Biochemistry* **41**: 13473–82.
31. **Ebbinghaus S, Kim SJ, Heyden M, Yu X**, et al. 2007. An extended dynamical hydration shell around proteins. *Proc Natl Acad Sci USA* **104**: 20749–52.
32. **Persson E, Halle B**. 2008. Cell water dynamics on multiple time scales. *Proc Natl Acad Sci USA* **105**: 6266–71.
33. **Meister K, Ebbinghaus S, Xu Y, Duman JG**, et al. 2013. Long-range protein-water dynamics in hyperactive insect antifreeze proteins. *Proc Natl Acad Sci USA* **110**: 1617–22.
34. **Papoian GA, Ulander J, Wolynes PG**. 2003. Role of water mediated interactions in protein–protein recognition landscapes. *J Am Chem Soc* **125**: 9170–8.
35. **Langhorst U, Backmann J, Loris R, Steyaert J**. 2000. Analysis of a water mediated protein–protein interactions within RNase T1. *Biochemistry* **39**: 6586–93.
36. **Reichmann D, Phillip Y, Carmi A, Schreiber G**. 2008. On the contribution of water-mediated interactions to protein-complex stability. *Biochemistry* **47**: 1051–60.
37. **Ikura T, Urakubo Y, Ito N**. 2004. Water-mediated interaction at a protein-protein interface. *Chem Phys* **307**: 111–9.
38. **Nooren I, Thornton JM**. 2003. Diversity of protein-protein interactions. *EMBO J* **22**: 3486–92.
39. **Sekhar A, Lam HN, Cavagnero S**. 2012. Protein folding rates and thermodynamic stability are key determinants for interaction with the Hsp70 chaperone system. *Protein Sci* **21**: 1489–502.
40. **Welsh MJ, Dedman JR, Brinkley BR, Means AR**. 1978. Calcium-dependent regulator protein: localization in mitotic apparatus of eukaryotic cells. *Proc Natl Acad Sci USA* **75**: 1867–71.
41. **Bowie JU, Sauer RT**. 1989. Equilibrium dissociation and unfolding of the arc repressor dimer. *Biochemistry* **28**: 7139–43.
42. **Ou L, Ferreira AM, Otieno S, Xiao L**, et al. 2011. Incomplete folding upon binding mediates Cdk4/Cyclin D complex activation by tyrosine phosphorylation of inhibitor p27 protein. *J Biol Chem* **286**: 30142–51.
43. **Ovádi J, Srere PA**. 2000. Macromolecular compartmentation and channeling. *Int Rev Cytol* **192**: 255–80.
44. **Dang CV, Ferguson B, Burke DJ, Garcia V**, et al. 1985. Interactions of aminoacyl-tRNA synthetases in high-molecular-weight multienzyme complexes from rat liver. *Biochim Biophys Acta* **829**: 319–26.
45. **Lee SW, Cho BH, Park SG, Kim S**. 2004. Aminoacyl-tRNA synthetase complexes: beyond translation. *J Cell Sci* **117**: 3725–34.
46. **Hausmann CD, Ibba M**. 2008. Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. *FEMS Microbiol Rev* **32**: 705–21.
47. **Guo M, Ignatov M, Musier-Forsyth K, Schimmel P**, et al. 2008. Crystal structure of tetrameric form of human lysyl-tRNA synthetase: implications for multisynthetase complex formation. *Proc Natl Acad Sci USA* **105**: 2331–6.
48. **Weber SC, Brangwynne CP**. 2012. Getting RNA and protein in phase. *Cell* **149**: 1188–91.
49. **Muschol M, Rosenberger F**. 1997. Liquid–liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *J Chem Phys* **107**: 1953.
50. **Brangwynne CP, Eckmann CR, Courson DS, Rybarska A**, et al. 2009. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**: 1729–32.
51. **Li P, Banjade S, Cheng H-C, Kim S**, et al. 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**: 336–40.
52. **Wu H**. 2013. Higher-order assemblies in a new paradigm of signal transduction. *Cell* **153**: 287–92.
53. **Denos S, Dhar A, Gruebele M**. 2012. Crowding effects on the small, fast-folding protein lambda(6-85). *Faraday Discuss* **157**: 451–62.
54. **Cino EA, Karttunen M, Choy WY**. 2012. Effects of molecular crowding on the dynamics of intrinsically disordered proteins. *PLoS ONE* **7**: e49876.
55. Cai W, Hong H, (eds). 2012. *Protein–protein interactions – computational and experimental tools*. InTech, New York, DOI: 10.5772/2679

56. **Renault M, Tommassen-van Boxtel R, Bos MP, Post JA**, et al. 2012. Cellular solid-state nuclear magnetic resonance spectroscopy. *Proc Natl Acad Sci USA* **109**: 4863–8.
57. **Selenko P, Wagner G**. 2007. Looking into live cells with in-cell NMR spectroscopy. *J Struct Biol* **158**: 244–53.
58. **Ghaemmaghami S, Oas TG**. 2001. Quantitative protein stability measurement in vivo. *Nat Struct Biol* **8**: 879–82.
59. **Xu K, Zhong G, Zhuang X**. 2013. Actin, spectrin, and associated proteins form a periodic cytoskeletal structure in axons. *Science* **339**: 452–6.
60. **Jones SA, Shim S-H, He J, Zhuang X**. 2011. Fast, three-dimensional super-resolution imaging of live cells. *Nat Methods* **8**: 499–505.
61. **Ebbinghaus S, Gruebele M**. 2011. Protein folding landscapes in the living cell. *J Phys Chem Lett* **2**: 314–9.
62. **Zuzuarregui A, Kupka T, Bhatt B, Dohnal I**, et al. 2012. M-Track: detecting short-lived protein-protein interactions in vivo. *Nat Methods* **9**: 594–6.
63. **Gallego O, Specht T, Brach T, Kumar A**, et al. 2013. Detection and characterization of protein interactions in vivo by a simple live-cell imaging method. *PLoS ONE* **8**: e62195.
64. **Wang Q, Cheung MS**. 2012. A physics-based approach of coarse-graining the cytoplasm of *Escherichia coli* (CGCYTO). *Biophys J* **102**: 2353–61.
65. **Roberts E, Stone JE, Luthey-Schulten Z**. 2013. Lattice microbes: high-performance stochastic simulation method for the reaction-diffusion master equation. *J Comp Chem* **34**: 245–55.
66. **Douglas J, Dudowicz J, Freed K**. 2009. Crowding induced self-assembly and enthalpy-entropy compensation. *Phys Rev Lett* **103**: 135701.
67. **Kim YC, Mittal J**. 2012. Crowding induced entropy-enthalpy compensation in protein association equilibria. *Phys Rev Lett* **110**: 208102.
68. **Jiao M, Li H-T, Chen J, Minton AP**, et al. 2010. Attractive protein-polymer interactions markedly alter the effect of macromolecular crowding on protein association equilibria. *Biophys J* **99**: 914–23.
69. **Rosen J, Kim YC, Mittal J**. 2011. Modest protein–crowder attractive interactions can counteract enhancement of protein association by intermolecular excluded volume interactions. *J Phys Chem B* **115**: 2683–9.
70. **Ebbinghaus S, Dhar A, McDonald D, Gruebele M**. 2010. Protein folding stability and dynamics imaged in a living cell. *Nat Methods* **7**: 319–23.
71. **Ignatova Z, Gierasch LM**. 2004. Monitoring protein stability and aggregation in vivo by real-time flourescent labeling. *Proc Natl Acad Sci USA* **101**: 523–8.
72. **Ignatova Z, Krishnan B, Bombardier JP, Marcelino AMC**, et al. 2007. From the test tube to the cell: exploring the floding and aggregation of a beta-clam protein. *Biopolymers* **88**: 157–63.
73. **Phillip Y, Kiss V, Schreiber G**. 2012. Protein-binding dynamics imaged in a living cell. *Proc Natl Acad Sci USA* **109**: 1461–6.
74. **Mir M, Wang Z, Shen Z, Bednarz M**, et al. 2011. Optical measurement of cycle-dependent cell growth. *Proc Nat Acad Sci USA* **108**: 13124–9.
75. **Tzur A, Kafri R, LeBleu VS, Lahav G**, et al. 2009. Cell growth and size homeostasis in proliferating animal cells. *Science* **325**: 167.
76. **Hinde E, Cardarelli F, Digman MA, Gratton E**. 2012. Changes in chromatin compaction during the cell cycle revealed by micrometer-scale measurement of molecular flow in the nucleus. *Biophys J* **102**: 691–7.
77. **Hynes TR, Randal M, Kennedy LA, Eigenbrot C**, et al. 1990. X-ray crystal-structure of the protease inhibitor domain of alzheimers amyloid beta-protein precursor. *Biochemistry* **29**: 10018–22.
78. **Buczek O, Krowarsch D, Otlewski J**. 2002. Thermodynamics of single peptide bond cleavage in bovine pancreatic trypsin inhibitor (BPTI). *Protein Sci* **11**: 924–32.
79. **Scholtz JM, Baldwin RL**. 1992. The mechanism of alpha-helix formation by peptides. *Annu Rev Biophys Biomol Struct* **21**: 95–118.
80. **Wieczorek R, Dannenberg JJ**. 2005. Enthalpies of hydrogen-bonds in α-helical peptides. An ONIOM DFT/AM1 study. *J Am Chem Soc* **127**: 14534–5.
81. **Doig AJ, Williams DH**. 1991. Is the hydrophobic effect stabilizing or destabilizing in proteins? *Exp Cell Res* **217**: 389–98.
82. **Jones K, Wittung-Stafshede P**. 2003. The largest protein observed to fold by two-state kinetic mechanism does not obey contact-order correlation. *J Am Chem Soc* **125**: 9606–7.
83. **Nguyen H, Jäger M, Kelly J, Gruebele M**. 2005. Engineering beta-sheet protein toward the folding speed limit. *J Phys Chem B* **109**: 15182–6.
84. **Shiroishi M, Yokota A, Tsumoto K, Kondo H**, et al. 2001. Structural evidence for entropic contribution of salt bridge formation to a protein antigen-antibody interaction. The case of hen lysozyme-HyHEL-10 Fv complex. *J Biol Chem* **276**: 23042–50.
85. **Fersht AR**. 1987. The hydrogen bond in molecular recognition. *Trends Biochem Sci* **12**: 301–4.
86. **Klotz IM, Langebman NR, Dahnall DW**. 1970. Quaternary structure of proteins. *Annu Rev Biochem* **39**: 25–62.
87. **Vaynberg J, Fukuda T, Chen K, Vinogradova O**, et al. 2005. Structure of an ultraweak protein-protein complex and its crucial role in regulation of cell morphology and motility. *Mol Cell* **17**: 513–23.
88. **Huang X, Poy F, Zhang RG, Joachimiak A**, et al. 2000. Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat Struct Biol* **7**: 634–8.
89. **Shaanan B**. 1983. Structure of human oxyhemoglobin at 2.1 Å resolution. *J Mol Biol* **171**: 31–59.
90. **Pellicena P, Karow DS, Boon EM, Marletta MA**, et al. 2004. Crystal structure of an oxygen-binding heme domain related to soluble guanylate cyclases. *Proc Natl Acad Sci USA* **101**: 12854–9.