Review Article

Progress in Biological Sciences

Vol. 5, Number 1, Winter / Spring 2015/1-18

Cooperativity in biological systems

Received: October 15, 2014; Accepted: December 7, 2014

Reza Behrouzi *

Department of Cell Biology, Harvard Medical School, Boston, MA, USA

ABSTRACT

Living organisms can sense and respond to external and internal stimuli. Response is demonstrated in many forms including modulation of gene expression profiles, motility, secretion, cell death, etc. Nevertheless, all forms share a basic property: they depend on sensing small changes in the concentration of an effector molecule or subtle conformational changes in a protein and invoking the appropriate molecular response by the relevant signaling pathways. Sensing, transduction, and response to signals may be directly carried out by controlled changes in the conformation or the assembly of pre-existing components(1,2) or may involve changes in gene expression patterns (as in cell differentiation and development), which in turn is carried out by protein-nucleic acid interactions and complex formation. Hence, understanding conformational changes in proteins and nucleic acids, ligand binding, and complex formation play acentral role in advancing our knowledge of cellular dynamics. Large-scale interaction mapping projects continue to provide detailed (though approximate) interaction networks between pairs of proteins (3–6), but fall short of capturing the stability or dynamics of the interactions. Integration of these maps with thermodynamic and kinetic information about conformational changes and binding events in proteins and nucleic acids holds the promise of discovering simple universal mechanisms that explain and relate seemingly disparate biological phenomena at many levels of complexity. In this article, I will explore 'cooperativity', one of the most ubiquitous features in molecular biology and discuss how it impacts macromolecular folding, complex assembly, formation of biological networks, and eventually cellular function and pathology.

Keywords: higher-order assemblies, misfolding, self-assembly, two-state folding.

* behrouzi@hms.harvard.edu

Significance

Self-assembly of biological components and complexes may appear to be statistically impossible. Protein and RNA enzymes are made from linear chains that must fold to unique functional conformations. However, the conformational degrees of freedom for each residue of the chain seemingly creates an astronomical conformational space that renders the probability of finding a unique native fold in a global search essentially zero. Yet, many biological macromolecules achieve this seemingly impossible feat in milliseconds (7).

Upon folding, most proteins and nucleic acids become parts of multicomponent complexes that perform almost every cellular function (8). Complex assembly relies largely on weak interactions ($\Delta G^\circ \sim kT$) between limited numbers of modular domains that are embedded and reused in different proteins (1). The modularity of the binding domains allows proteins and nucleic acids to bind numerous targets in the cellular environment, so it is not uncommon for a protein to be part of more than one complex. Furthermore, notoriously prevalent nonspecific interactions between unrelated proteins and nucleic acids (9–11) could also overwhelm formation of defined complexes. However in normal cells, this seemingly improbable outcome is strongly favored over a vast space of possibilities.

Cooperative self-assembly is the remarkable property of biological systems that makes such apparently impossible search problems possible, in some cases so efficient that they approach the diffusion limit (12). A system or reaction is deemed cooperative when formation of one interaction favors (stabilizes) formation of subsequent ones. Through a bottom-up discussion of cooperativity and its implications in biological systems, this review strives to highlight a conserved biological phenomenon with diverse but thematically unified functions.

Cooperative interactions guide folding of biological macromolecules

RNA folding

Folding of RNA molecules with tertiary structures is largely, but not completely hierarchical (13–15), meaning that local structures are formed before long-distance contacts that stabilize the global fold of the RNA. Helices and loops are the building blocks of RNA secondary (*i.e.* Local) structure and have long been known to form cooperatively.

Cooperative formation of RNA double helices arises from nearest neighbor interactions between consecutive base pairs, which have been known since the early 1990s (13). Recent thermodynamic measurements provided higher resolution on how cooperativity changes with respect to the position of base pairs relative to the ends of the helix and showed large cooperativity between base pairs throughout the helix length of RNA (16). Interestingly, comparison of RNA and DNA hairpins showed that base pairing interactions in RNA are less cooperative than in DNA (17,18). The coupling between base pairs in RNA hairpins is direct, meaning that the coupling energy is insensitive to the disruption of other base pairs. In DNA hairpins however, coupling between each pair of base pairs depends on the other base pairs (18) (indirect cooperativity). As a result, individual mutations in RNA hairpins are less destabilizing than in DNA hairpins, leaving room for neutral genetic drift.

Local elements of the structure are thermodynamically more stable than long distance contact motifs, although global folding and tertiary contacts may be required for the formation of native local structure in some regions (19). As a result of stable structural elements, the folding energy landscapes of RNA molecules are rugged and contain multiple minima besides the native basin (20).

Until recently it was not clear whether global folding in large RNAs is cooperative or stepwise, consisting of cumulative formation of local structural elements. In fact, some long-distance (tertiary) interaction motifs have been shown to be non-cooperative (21,22) or only slightly cooperative (23). The first direct observation of cooperative tertiary folding came from single-molecule studies of an independently stable domain of *Tetrahymena* group I ribozyme (24) and revealed that two distant tertiary interaction motifs are strongly cooperative in stabilizing the global fold of the RNA. However, the simple topology of this domain led to speculations that global folding might not be cooperative in large structured RNAs that contain multiple stable domains and competing interactions (24).

Thermodynamic measurement of the coupling energies between tertiary interactions in the global folding transition of a large structured RNA was recently performed in the *Azoarcus* group I ribozyme (25). By making single-base-change mutations to slightly perturb the stability of a number of long-distance interactions in this ribozyme, the folding free energies of the native state and native-like intermediates of wild type and perturbed ribozymes were measured. The findings revealed that folding of this large RNA is globally cooperative, so that tertiary interactions that are spatially separated by more than 30 Å form simultaneously.

Besides direct measurements of coupling free energies, other studies present strong evidence for the generality of cooperativity in large structured RNAs. Hydroxyl radical foot printing assays have been used in combination with urea titration experiments to show that all tertiary motifs of a large RNA, *Tetrahymena thermophila* group I ribozyme, have comparable stabilities (26,27). This is consistent with a two-state unfolding process indicative of high coupling between tertiary interactions. Furthermore, rapid and two state transitions from extended to collapsed and compact structures that have been observed for several large RNAs support the hypothesis that large RNA structures fold cooperatively, despite a rugged energy landscape (28,29).

Nevertheless, the global folding transition in some RNAs may not be fully cooperative when measured in vitro. Imbalance of stability in different subdomains of RNA structures is not consistent with cooperative folding. This has been experimentally demonstrated in the *Tetrahymena* group I ribozyme. Engineered destabilization of an independently stable domain suppresses folding intermediates and makes folding more cooperative (30).

In general, global folding of large RNA structures may be less robust compared to proteins of the same size, as strong non-native contacts are more likely to form in RNA. Given the small number of specific native interactions, a few non-native interactions can significantly reduce the efficiency of RNA folding (31). On the other hand, addition of individual tertiary interactions may stabilize both native and non-native intermediates and reduce folding specificity, despite increasing native state

stability (32,33). This fundamental difference in the folding robustness of RNA and protein polymers may have had a role in the transfer of enzymatic roles from RNA to proteins during the evolution.

Protein folding

Folding cooperativity implies that under various conditions (temperature, pH, denaturant concentration, etc.), the unfolded and native states are the only conformations that are significantly populated. Hence, two-state and cooperative are often used interchangeably. Two-state folding is a general property of almost all natural single-domain peptides and many larger proteins (34,35). Cooperative folding is prominent in formation of local (secondary structure elements) and global (long range tertiary interactions) structures (36).

The mechanisms of cooperative folding in polypeptides have been investigated extensively, originating in the helix-coil theory of Schellman (37) and Zimm and Bragg (38).This model views the folding of a polypeptide chain as the increasing growth of local helical elements and is most suited to the formation of regular local structures, such as an alpha helix. In the early 90s, the hydrophobic zipper model was proposed by Dill and coworkers (12,39) to address the shortcomings of this model to account for cooperative global folding of proteins. Unlike helix-coil transitions, the hydrophobic zipper model posits that global folding is driven by nonlocal hydrophobic interactions between sequence-distant residues. Since formation of each interaction acts as a constraint on the freedom of the polypeptide chain, it makes it easier for sequential formation of other interactions. This review will not address concepts and models of folding in small proteins, as these ideas have been extensively discussed elsewhere (36).

Our understanding of protein folding cooperativity is largely rooted in studies of small single domain peptides and much less is known about the folding of large multi-domain proteins. Although, they sometimes exhibit stepwise (as opposed to two-state) folding (40–42), cooperative interactions among multiple domains are important for the function of many large proteins as well (43). Since large proteins consist of modular, independently stable domains (44) cooperative folding must rely on some form of communication among the domains to achieve a concerted transition from unfolded to the fully folded structure. Otherwise, folding of each domain would be determined solely by its own thermodynamic stability and kinetic properties, independent of the state of the other domains. Compared to the folding of small proteins, the mechanisms of inter-domain communication have been less explored. However, studies on small protein structures that include independently stable subdomains, as well as repeat proteins have paved the way to a better understanding of cooperative folding.

T4 lysozyme is a small dumbbell shaped protein with two subdomains that has been studied extensively as a model system for protein folding studies. Despite containing distinct structural domains, it folds in a two-state cooperative transition (45). Hence, it is a great model for studying how interdomain interactions affect folding cooperativity. Changing the order and topology of the subdomains while maintaining the original sequence (circular permutation) strongly destabilized the protein, but the protein still demonstrated enzymatic activity and a cooperative unfolding transitions in denaturant solution (45). Therefore, it was concluded that the chain topology does not

affect the native protein fold. However, recent single molecule force measurements combined with non-equilibrium thermodynamic analyses unraveled finer details that changed this picture. Comparison of force-induced unfolding free energies, when the force was applied to the whole protein or to one domain alone, revealed that circular permutation diminishes cooperativity (46). If the natural domain architecture of the protein is broken, folding generally proceeds through partially folded intermediates.

Repeat proteins constitute a special class of large proteins and are involved in important cell signaling events, such as the Notch signaling pathway (35). The structure of these large proteins consists of tandem repeats of a relatively small domain, connected by flexible loops. Despite their linear organization, most of these proteins fold cooperatively in two-state transitions (35,47), though there are instances of repeat proteins that form partially folded structures (48,49). Despite their large size, these proteins are amenable to biophysical analysis of the mechanism of folding because of their linear structure and dissectible chain topology (35,47). For example, the folding energy landscape has been determined for several natural or consensus repeat proteins by measuring internal and inter-domain free energies using systematic deletion and folding measurements (50,51).

A simple but important conclusion from the folding of repeat proteins is that cooperative folding requires homogenous stabilities of individual repeating units. Proteins in which repeats have varying stabilities fold in multiple steps. Furthermore, direct manipulation of inter-domain interactions by engineered mutations shows that ‘energetic coupling’ between repeating units is required for cooperative global folding (35,47). Such observations are likely to be relevant to globular proteins as well. High-resolution experimental techniques such as hydrogen/deuterium exchange measurements demonstrate that different regions of proteins may have different stabilities and that imbalance of local stabilities may undermine the cooperative nature of global folding (52–54). However, the difficulties in defining and measuring individual behavior of subdomains in a globular protein makes it hard to directly test these ideas in globular proteins.

Emergence of cooperative folding in biological macromolecules

The role of cooperativity may best be understood by considering how it shapes the folding energy landscape of macromolecules. The folding of macromolecules is accompanied by loss of configurational entropy, which has to be energetically compensated by formation of stabilizing interactions. In cooperatively folding biopolymers, the shape of energy landscape resembles a funnel with a unique global minimum (native state) (55). Nuclei of tertiary interactions can form in parallel in different regions of the structure, and grow to form the native fold. Cooperativity among interactions means that formation of any set of tertiary interactions is consistent with other interactions. Hence, trapped intermediates do not limit the folding process and folding proceeds quickly to structures with a significant degree of native structure (55,56). In contrast, the energy landscape of polymers with competing, but mutually inconsistent nucleation sites contain other funnels beside the global energy minimum as formation of some tertiary interactions frustrates the other interactions (57).

It is remarkable that both major classes of

functional biopolymers, RNA and polypeptides, fold specifically and form unique conformations. The nature of physical interactions that stabilizes these biopolymers and drives their folding is vastly different. Strong electrostatic interactions with cations, stacking interactions between bases, and hydrogen bonds are the main stabilizers of RNA, whilethe global structure of polypeptides is largely stabilized by entropically driven hydrophobic interactions.

The specificity of folding is not due to intrinsic chemical properties of these polymers. Randomized RNA sequences and rationally designed polypeptides (lacking detectable homology to natural ones) generally demonstrate a multitude of equilibrium and kinetically trapped intermediates (56,58–60). Therefore, folding cooperativity is likely a product of evolution.

Altogether, natural selection favors chain topologies that are conducive to cooperative folding to ensure folding specificity. Interestingly, domain topology is dispensable for enzyme function and stability (46). Consequently, it has been possible to create new proteins that are more stable than their natural counterparts by de novo design or by construction of consensus sequence of a family of homologous proteins. However, these proteins, lacking the evolutionary history of natural proteins, lack folding cooperativity and demonstrate stable folding intermediates (35,59,60). Similarly, concerted changes in more than one helical domain are necessary to increase the folding cooperativity and specificity of RNA structure (61). Therefore, explicit integration of topological principles is necessary to design macromolecules with unique stable folds.

Cooperative interactions guide macromolecular complex assembly

Self-assembly of complexes

Proteins and ribonucleoprotein complexes perform the majority of cellular functions. Complexes bring related enzymatic functions to close proximity to increase the efficiency of multistep reactions and decrease inadvertent escape and accumulation of reaction intermediates. Simultaneously, catalytic processes remain dynamically regulated through modulation of the complex assembly and disassembly rates (8). These rates can be modulated by changes in concentration or post-translational modifications of one or more of the components.

Cellsuse a limited number of modular domains to facilitate interactions between proteins and nucleic acids (1). Many proteins, especially those involved in signaling pathways or complexes, contain a number of different domains allowing them to bind to different targets simultaneously. Additionally, many interaction domains are versatile as they can recognize multiple ligands in different contexts (1,62).

Complex assembly relies on the specific binding interactions among components. However, nonspecific interactions between unrelated proteins and nucleic acids are the rule rather than the exception (1,10,63). Although off-target binding events are often weaker than binding of a domain to its specific target, the number of possible nonspecific binding targets greatly outnumbers the specific binding loci. Hence, the differential stability of specific and nonspecific binding loci is not enough to guarantee specific recruitment. Increasing the binding strength of specific partners by forming larger contact surfaces increases the

strength of nonspecific binding events as well (for examples, see Pawson and Nash (1). Hence, cumulative evolutionary improvements in binding affinity do not provide a solution for specific recruitment and assembly of components of a complex.

Instead, cooperative self-assembly has evolved to ensure efficient and specific formation of complexes and is widespread in modern macromolecular complexes. Ribosome, proteasome, and viral capsids are only a few examples of complex structures that assemble cooperatively from their components. DNA and RNA bound complexes, such as transcription machinery (64,65), various transcriptional silencing (66) and other chromatin-bound complexes (67), as well as mRNA-bound regulatory complexes (62) rely on cooperative interactions between their components to assemble and function. Another important venue for cooperative assembly involves organization of cell signaling proteins that translate external stimuli to intracellular responses (68–70).

Two mechanisms underlie the cooperativity observed in the assembly of different complexes from the same group of proteins. The first mechanism relies on the 'multivalent' structure of many proteins, meaning that there are two or more different interaction domains in a given protein that allow it to bind multiple targets simultaneously. For example, assume a trimeric complex ABC in which A and C both bind the common partner B. If A has an additional binding pocket for C, then C is more likely to bind AB than B, because of the extra binding energy gained from binding to A. Combinatorial post-translational modifications on histone proteins are proposed to act in this manner to induce specificity in binding of chromatin reader proteins (71,72).

The second mechanism is dependent on the dynamic nature of protein and nucleic acid structures. In this case, cooperative binding to a common partner does not require direct interaction between the binders. Instead, one binding event causes a change in the conformation of the common partner that increases its affinity for the second. This mechanism is analogous to the allosteric effects observed in ligand binding and is often harder to discern (73–76).

Cooperative assembly decreases the possibility of partially assembled or unproductive (dead end) intermediates (75) and thus guards against wasteful diversion of a cell's resources. Consider the assembly of ribosomal subunits, consisting of large noncoding RNAs and over 20 proteins. In an exponentially growing *Escherichia coli* culture, doubling every 20 minutes, ribosomes make up about half of the dry cell weight (77). About sixty percent of transcription is devoted to ribosomal RNA transcription (78) and about 40% of total energy is devoted to protein synthesis (79). It is easy to see how inefficient assembly of ribosomes would lead to toxic accumulation of nonfunctional intermediates and aggregates, resulting in growth arrest and cell death.

In addition, for complexes that are assembled on a specifically localized nucleating element, such as transcription initiation and silencing complexes, cooperative interaction among the components ensures the specificity of recruitment at the target loci, because the free energy of thermodynamic coupling outweighs nonspecific binding energies of each component. Multivalent proteins that can bind several motifs simultaneously play a critical role in the cooperative assembly of complexes at their target loci (66,67,70,71).

Higher-order assemblies and macroscopic transitions

Several cellular functions involve association and polymerization of many copies of one or more macromolecular complexes. For example, nucleosomes in transcriptionally silent domains of chromatin that extend several kilobasepairs are bound by silencing complexes or heterochromatin protein HP1 analogues (80,81). Tumor necrosis factor (TNF) receptor (TNFR) and the Toll-like receptor/interleukin-1receptor (TLR/IL-1R) super families are instances of signal transduction systems that self-associate to form polymeric higher order assemblies upon activation by external signals (see, Wu (2)). Binding of carbohydrates and lectin, forming lattices and other higher order structures on the cell surface, is yet another example of this type of system (82).

Formation of these higher order complexes relies on 'multivalent' monomers, i.e. monomers that have more than one interacting domain to engage with the other building blocks. For example, in chromatin-templated assembly and spreading of silencing complexes, at least two types of binding domains (valencies) are required: a modified-histone binding domain that recognizes histone post-translational modifications specific to heterochromatin, and a dimerization domain in one component of the complex (66). Consolidation of multiple interaction domains into one protein or a multicomponent complex couples the binding of each domain to its respective target with the other domains, thus making the polymerization reaction cooperative. Domain-deletion studies in yeast for instance have shown that heterochromatin domains cannot be established and maintained when either dimerization or histone-binding domains are absent (83). This is also true for proteins with multiple copies of a binding domain. In fact, in vivo measurements of higher order assembly of SH3 and PRM domains in engineered peptides with different numbers of repeats established that the formation of these assemblies requires more than one instance of the domains. Furthermore, increasing the number of domains in the peptides reduces the minimum concentration of SH3 and PRM repeats that are needed for higher order assembly formation (84). This work demonstrated for the first time that the binding of multivalent signaling proteins results in a sharp (cooperative) macroscopic phase transition leading to liquid-liquid demixing of the higher-order assembly in cells. Importantly, the phase transition boundaries can be regulated by the phosphorylation state of the binding partners. Such phase transitions are emerging as a new paradigm in signaling pathways (2) and may be more important than previously appreciated.

Cooperative interactions and construction of gene regulatory networks

Many biological regulatory systems are bistable; they can stably exist in two alternative states (ON/OFF) depending on the level of the input stimulus. Theoretical investigations have discovered a number of key shared features in such systems, including ultra-sensitivity and positive feedback (85). Cooperativity among biomolecules is one of the mechanisms used by cells to create ultra-sensitivity.

For instance, epigenetic states of gene expression that depend on chromatin structure are thought to be established and maintained by cooperative interactions among specific targeting proteins, general silencing complexes, and modified histones (66). Computational investigations of these

systems show that bi-stability in the specialized chromatin regions is impossible without cooperative interactions among nucleosomes and chromatin-modifying complexes (86,87). These epigenetic programs are established early during development and faithfully inherited through many cell divisions (88). Similarly, cooperative spatial organization of components of certain signaling pathways, such as mitogen-activated protein kinase (MAPK), plays a key role in bistable signaling and suppression of cross-talk among different pathways (68,69).

However, it must be noted that other mechanisms besides binding cooperativity may underlie ultra-sensitivity in other bistable systems. For example, positive and double-negative feedback circuits among pairs of gene products are involved in many gene regulatory systems(89), but fall outside the scope of this review.

Conclusion

In the context of macromolecular folding and complex assembly, cooperative interactions provide important advantages that underlie their natural selection. As might be expected, loss of cooperativity leads to serious pathological states.

Cooperativity brings specificity to biological reactions

Cooperative interactions among the components of a system decrease the abundance of reaction intermediates. That is why a fully cooperative transition occurs in one step from the initial to the final state. Non-cooperative interactions, on the other hand, proceed in two or more steps and may involve on- or off-pathway intermediates.

In the folding of protein and RNA molecules, intermediates are partially folded chains that may be native-like (on pathway) or non-native (off pathway) in their structure. In the former type, sometimes observed in the folding of large or repeat proteins and large noncoding RNA molecules *in vitro*(see above), part of the polymer adopts its native tertiary structure, while the rest still remains unfolded. Subsequent steps consist of folding the remaining regions and accumulation of the native structure. In contrast, in non-native intermediates, some or part of the polymer chain adopts a tertiary structure unlike the native fold. Therefore, a partial or complete unfolding transition is required to reset the polymer chain and the possibility of folding to the native conformation. In macromolecular assembly reactions, intermediates may include incomplete on-pathway assembly products, or trapped sub-complexes that cannot proceed toward the native complex without the dissociation of some of the components.

In reactions, where non-native intermediates are thermodynamically possible, cooperative transitions significantly increase the reaction yield, as fewer molecules or complexes are trapped in “dead-end” off-pathway intermediates. On the other hand, if all intermediates in a reaction pathway are native-like (on pathway), it’s hard to imagine any intrinsic advantage for a cooperative mechanism over stepwise transitions. However, due to some idiosyncratic properties of biopolymers and complexes, even on pathway intermediates at high-enough concentrations may lead to detrimental biological side effects.

The suppression of reaction intermediates makes cooperative transitions specific, meaning that they mainly produce the intended final state (folded conformation or multicomponent complex) with little or no

side products. This is arguably the most remarkable feature of cooperative systems. As mentioned before, proteins and nucleic acids are generally capable of binding to a wide range of partners specifically and nonspecifically. Without cooperative interactions between components formation of a specific complex would be highly improbable.

An important, but often overlooked corollary is that specific complex assembly requires component concentrations to be tightly regulated (90). This is because the differential stability of independent and cooperative binding is dependent on the concentration of the free components. When free concentrations are too high or too low relative to the binding affinity of the components, cooperative stabilization effects will not be observed. Hence, cooperativity as a mechanism of guiding complex assembly to specific outcomes can only function when component concentrations are tightly regulated. As expected, deregulation of concentrations, for example by overexpression in recombinant systems, results in the loss of specificity in complex assembly. The assembly of transcriptional silencing complexes and the formation of heterochromatin regions provide excellent examples of such effects. Artificial overexpression of a protein required for transcriptional silencing often results in partial or complete loss of heterochromatin formation because instead of a limited number of specifically localized and fully associated silencing complexes, components spread throughout the genome as partially assembled sub-complexes.

Kinetic effects of cooperative interactions

Cooperativity by definition is a thermodynamic property and, in principle, it does not provide any information about the rate of reaching the final state of the transition. Thus, in contrast to the general relationship between cooperativity and specificity discussed in the previous section, a generalized relationship between cooperativity and kinetics does not exist. However, as discussed below, the geometric features of protein and RNA chains strongly affect folding kinetics and cooperativity in similar ways, making it possible to form a general picture of how cooperativity might affect kinetics.

Nearly three decades of computational and experimental research on protein folding rates and mechanisms has established that the rate of folding of proteins is largely determined by the native topology of the polypeptide chain. The average sequence separation of residues that interact in the tertiary structure is called contact order and correlates with protein folding rate[87].This correlation holds for over 6 orders of magnitude in folding rates and implies that geometric features of the protein, not the details of side chain interactions determined by the sequence, have the strongest role in determining the mechanism and rate of folding (91).

Polypeptides with low contact order generally fold faster because the residues that make the initiallong-range interactions in the folding process are on average closer in sequence space. Hence, relatively smaller configurational entropy is lost by the formation of these native contacts and a smaller free energy barrier has to be overcome to reach the transition state. Transition state energy is directly proportional to the reaction rate, thus the faster folding rates. Additionally, once the first long-range interactions are formed, formation of the remaining native interactions occurs at a significantly smaller entropic cost.

Thus, in effect, these interactions are significantly stabilized by the previous interactions (cooperativity). Most small single domain proteins are thought to fold in this manner, demonstrating fast kinetics and two-state cooperative equilibrium transitions (see above).

The above folding pathway and the ensuing correlation between folding rate and contact order is possible only when the native fold is the unique thermodynamically favorable folded conformation (energy minimum) available to the peptide chain. Otherwise, formation of the initial long-range interactions may be followed by non-native interactions that lead to misfolded or kinetically trapped intermediates. In such cases, some interactions that form quickly trap the protein conformation in a dead-end state that cannot proceed toward the native state, unless some interactions are unfolded first. The observation of such folding pathways indicates that the folding transition is not fully cooperative, since the formation of some interactions opposes the formation of other native interactions. This is the case for some large RNA molecules folding in vitro as well as protein structures with the possibility of non-native interactions.

Finally, it is possible to imagine another folding pathway for a multi-domain protein or RNA molecule, in which the native state is the only thermodynamically stable state, but the global folding of the protein is not cooperative as the domains of the macromolecule fold independently. Here, the kinetics of folding will depend on the individual domains and could be either slow or fast. Under such conditions, no correlation between global cooperativity and kinetics is expected.

In summary, the relationship between cooperative equilibrium transitions and kinetic folding pathways are not necessarily direct and simple. This is highlighted by inconsistencies between theoretical expectations and experimental measures of the degree of folding of residues in the transition states of many proteins (35). Nevertheless, it remains true that cooperativity among native interactions makes the folding energy surface ‘smoother’ by suppressing the formation of non-native interactions. This has a generally favorable effect on folding kinetics, as direct pathways from unfolded to folded state are more likely to be travelled on a smooth energy landscape.

The same general relationship and complications between cooperativity and kinetics also hold true for macromolecular complex assembly reactions. Generally, when dead-end assembly intermediates are not possible, the binding cooperativity of components increases the binding kinetics. However, in many complexes, off-pathway assembly intermediates are possible, which give rise to complexities in the kinetic folding pathway.

Loss of cooperativity in disease and pathology

The discussions above make it clear that even though the formation of the native state of biological macromolecules and complexes might be driven thermodynamically, folding or self-assembly processes may be frustrated by the formation of nonproductive (kinetically trapped) intermediates. Besides wasting metabolic resources and energy, the accumulation of partial folds and assemblies leads to other detrimental consequences in the crowded environment of the cell.

A pathologically familiar instance of such consequences is the case of amyloid diseases. These are accompanied by the unusual aggregation of specific proteins in the form of

fibrils or plaques that are deposited in various tissues and organs (43). For example, point mutations in the human lysozyme gene cause hereditary non-neuropathic systemic amyloidosis (92). Amyloidogenic variants have native folds and are catalytically active, but they are less stable and occasionally form partially folded intermediates, in which only one of the two domains of the protein is folded. The point mutations destabilize part of the protein structure, thus folding of the protein can no longer be a cooperative transition accompanied by structure formation in both domains. Consequently, unfolded parts of the protein provide extensive interaction surfaces that underlie the formation of higher order aggregates and insoluble fibrils (93,94).

Conversely, one may infer that recovering the folding cooperativity by any mechanism would circumvent the problem and thus prevent fibril formation. Remarkably, an antibody that was discovered to inhibit fibril formation in vitro, achieves this effect by restoring folding cooperativity to the protein (95). Interestingly, the antibody makes minimum contacts to the domain affected by the mutation (and not the site of point mutation), but seems to exert its stabilizing effect by long-range conformation effects at the interface of the two domains (95).

Chaperone proteinsplay an analogous role to this antibody. Binding of chaperones to the polypeptide chains as they emerge from the ribosome may prevent partial folding of proteins before they are completely synthesized. Other chaperones use the energy of ATP hydrolysis to unfold partially folded intermediates and allow them to re-fold. If the native state is the most stable conformation, this action will shift the balance of conformations toward the native state. In effect, both chaperone mechanisms will shift a kinetically controlled conformational population to a thermodynamically controlled one, favoring the native state.

Concluding remarks

As progress in experimental and computational methodologies enables quantitative investigation of larger subsets of cellular interaction networks, new examples of cooperative interactions among proteins and nucleic acids are discovered. Cooperative interactions enable cells to construct specific complexes from modular building blocks, localize reactions to specific recruiting loci, and create stable states that are resistant to fluctuations in their components. Understanding the basis and implications of cooperativity is necessary for proper understanding of molecular interactions and cellular regulatory circuits.

Acknowledgements

The author greatly appreciates helpful suggestions and critical reading of this manuscript by Rebecca Mathews (Harvard Medical School) and Yasaman Etminan (University of Maryland, Baltimore County).

.

REFERENCES

1. Pawson T, Nash P. Assembly of cell regulatory systems through protein interaction domains. Science . 2003;300(5618):445–52.
2. Wu H. Higher-order assemblies in a new paradigm of signal transduction. Cell . 2013;153(2):287–92.
3. Arabidopsis Interactome Mapping Consortium, Dreze M, Carvunis A-R, Charloteaux B, Galli M, Pevzner SJ, et al. Evidence for network evolution in an Arabidopsis interactome map. Science . 2011;333:601–7.
4. Rual J-F, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, et al. Towards a proteome-scale map of the human protein-protein interaction network. Nature . 2005;437:1173–8.
5. Yu H, Braun P, Yildirim M a, Lemmens I, Venkatesan K, Sahalie J, et al. High-quality binary protein interaction map of the yeast interactome network. Science . 2008;322:104–10.
6. Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, et al. A map of the interactome network of the metazoan C. elegans. Science . 2004;303(5657):540–3.
7. Zwanzig R, Szabo A, Bagchi B. Levinthal's paradox. Proc Natl Acad Sci U S A. 1992;89:20–2.
8. Perica T, Marsh JA, Sousa FL, Natan E, Colwell LJ, Ahnert SE, et al. The emergence of protein complexes: quaternary structure, dynamics and allostery. Biochem Soc Trans . 2012;40:475–91.
9. Bonetta L. Protein-protein interactions: Interactome under construction. Nature . 2010;468(7325):851–4.
10. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. Nat Struct Mol Biol. 2013;20(11):1250–7.
11. Van Leeuwen F, Gottschling DE. Genome-wide histone modifications: Gaining specificity by preventing promiscuity. Curr Opin Cell Biol . 2002;14(6):756–62.
12. Dill K a, Fiebig KM, Chan HS. Cooperativity in protein-folding kinetics. Proc Natl Acad Sci U S A . 1993;90(5):1942–6.
13. Tinoco I, Bustamante C. How RNA folds. J Mol Biol . 1999;293(2):271–81.
14. Pan J, Woodson S a. The effect of long-range loop-loop interactions on folding of the Tetrahymena self-splicing RNA. J Mol Biol . 1999;294(4):955–65.
15. Woodson S a. Recent insights on RNA folding mechanisms from catalytic RNA. Cell Mol Life Sci . 2000;57(5):796–808.
16. Siegfried N a., Bevilacqua PC. Thinking Inside the Box. Designing, Implementing, and Interpreting Thermodynamic Cycles to Dissect Cooperativity in RNA and DNA Folding . Methods in Enzymology. 2009. 365–93.
17. Moody EM, Feerrar JC, Bevilacqua PC. Evidence that folding of an RNA tetraloop hairpin is less cooperative than its DNA counterpart. Biochemistry . 2004;43(25):7992–8.
18. Moody EM, Bevilacqua PC. Structural and energetic consequences of expanding a highly cooperative stable DNA hairpin loop. J Am Chem Soc . 2004;126(31):9570–7.

19. Rangan P, Masquida B, Westhof E, Woodson S a. Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. Proc Natl Acad Sci U S A . 2003;100(4):1574–9.

20. Thirumalai D, Woodson S a. Kinetics of folding of proteins and RNA. Acc Chem Res . 1996;29(9):433–9.

21. Silverman SK, Cech TR. Energetics and cooperativity of tertiary hydrogen bonds in RNA structure. Biochemistry . 1999];38(27):8691–702.

22. Silverman SK, Zheng M, Wu M, Tinoco I, Cech TR. Quantifying the energetic interplay of RNA tertiary and secondary structure interactions. RNA . 1999;5(12):1665–74.

23. Klostermeier D, Millar DP. Time-resolved fluorescence resonance energy transfer: a versatile tool for the analysis of nucleic acids. Biopolymers . 2002;61(3):159–79.

24. Sattin BD, Zhao W, Travers K, Chu S, Herschlag D. Direct measurement of tertiary contact cooperativity in RNA folding. J Am Chem Soc . 2008;130(19):6085–7.

25. Behrouzi R, Roh JH, Kilburn D, Briber RM, Woodson S a. Cooperative tertiary interaction network guides RNA folding. Cell . 2012;149(2):348–57.

26. Deras ML, Brenowitz M, Ralston CY, Chance MR, Woodson S a. Folding mechanism of the Tetrahymena ribozyme P4-P6 domain. Biochemistry . 2000 12;39(36):10975–85.

27. Ralston CY, He Q, Brenowitz M, Chance MR. Stability and cooperativity of individual tertiary contacts in RNA revealed through chemical denaturation. Nat Struct Biol . 2000;7(5):371–4.

28. Koculi E, Cho SS, Desai R, Thirumalai D, Woodson S a. Folding path of P5abc RNA involves direct coupling of secondary and tertiary structures. Nucleic Acids Res . 2012;40:8011–20.

29. Woodson S a. Compact intermediates in RNA folding. Annu Rev Biophys . 2010;39:61–77.

30. Pan J, Woodson S a. Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. J Mol Biol . 1998;280(4):597–609.

31. Chauhan S, Behrouzi R, Rangan P, Woodson S a. Structural Rearrangements Linked to Global Folding Pathways of the Azoarcus Group I Ribozyme. J Mol Biol . 2009;386(4):1167–78.

32. Treiber DK, Rook MS, Zarrinkar PP, Williamson JR. Kinetic intermediates trapped by native interactions in RNA folding. Science . 1998;279(5358):1943–6.

33. Wan Y, Russell R. Enhanced specificity against misfolding in a thermostable mutant of the tetrahymena ribozyme. Biochemistry . 2011;50(5):864–74.

34. Jackson SE. How do small single-domain proteins fold? Fold Des.; 1998;3(4):R81–91.

35. Kloss E, Courtemanche N, Barrick D. Repeat-protein folding: New insights into origins of cooperativity, stability, and topology. Arch Biochem Biophys . 2008;469(1):83–99.

36. Dill K a. Polymer principles and protein folding. Protein Sci . 1999;8(6):1166–80.

37. Schellman JA. The Factors Affecting the Stability of Hydrogen-bonded Polypeptide Structures in Solution. J Phys Chem . 1958;62:1485–94.

38. Zimm BH, Bragg JK. Theory of the Phase Transition between Helix and Random Coil in Polypeptide Chains. J Chem Phys . 1959;31(2):526.

39. Dill K a, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, et al. Principles of protein folding-a perspective from simple exact models. Protein Sci . 1995 ;4(4):561–602.

40. Jaenicke R. Protein folding: local structures, domains, subunits, and assemblies. Biochemistry. 1991;30:3147–61.

41. Wetlaufer DB. Nucleation, rapid folding, and globular intrachain regions in proteins. Proc Natl Acad Sci U S A. 1973;70:697–701.

42. Rose GD. Hierarchic organization of domains in globular proteins. J Mol Biol. 1979;134:447–70.

43. Dobson CM. The structural basis of protein folding and its links with human disease. Biochem Soc Symp . 2001;1–26.

44. Porter LL, Rose GD. A thermodynamic definition of protein domains. Proc Natl Acad Sci . 2012;109(24):9420–5.

45. Llinás M, Marqusee S. Subdomain interactions as a determinant in the folding and stability of T4 lysozyme. Protein Sci . 1998;7(1):96–104.

46. Shank E a, Cecconi C, Dill JW, Marqusee S, Bustamante C. The folding cooperativity of a protein is controlled by its chain topology. Nature . 2010;465(7298):637–40.

47. Tsytlonok M, Itzhaki LS. The how's and why's of protein folding intermediates. Arch Biochem Biophys . 2013;531:14–23.

48. Croy CH, Bergqvist S, Huxford T, Ghosh G, Komives EA. Biophysical characterization of the free IkappaBalpha ankyrin repeat domain in solution. Protein Sci. 2004;13:1767–77.

49. Nam Y, Sliz P, Song L, Aster JC, Blacklow SC. Structural basis for cooperativity in recruitment of MAML coactivators to Notch transcription complexes. Cell. 2006;124:973–83.

50. Aksel T, Barrick D. Analysis of repeat-protein folding using nearest-neighbor statistical mechanical models. Methods Enzymol . 2008;455:95–9125.

51. Aksel T, Majumdar A, Barrick D. The contribution of entropy, enthalpy, and hydrophobic desolvation to cooperativity in repeat-protein folding. Struct. 2011;19(3):349–60.

52. Spudich G, Marqusee S. A change in the apparent m value reveals a populated intermediate under equilibrium conditions in Escherichia coli ribonuclease HI. Biochemistry. 2000;39(38):11677–83.

53. Englander SW, Mayne L, Rumbley JN. Submolecular cooperativity produces multi-state protein unfolding and refolding. Biophys Chem. 2002;101-102:57–65.

54. Krantz BA, Mayne L, Rumbley J, Englander SW, Sosnick TR. Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. J Mol Biol. 2002;324:359–71.

55. Wolynes PG, Onuchic JN, Thirumalai D. Navigating the folding routes. Science. 1995;267(5204):1619–20.

56. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: A synthesis. Proteins Struct Funct Genet . 1995;21(3):167–95.

57. Lindberg MO, Oliveberg M. Malleability of protein folding pathways: a simple reason for complex behaviour. Curr Opin Struct Biol . 2007;17(1):21–9.

58. Schultes E a, Spasic A, Mohanty U, Bartel DP. Compact and ordered collapse of randomly generated RNA sequences. Nat Struct Mol Biol . 2005;12(12):1130–6.

59. Scalley-Kim M, Baker D. Characterization of the folding energy landscapes of computer generated proteins suggests high folding free energy barriers and cooperativity may be consequences of natural selection. J Mol Biol . 2004;338(3):573–83.

60. Watters AL, Deka P, Corrent C, Callender D, Varani G, Sosnick T, et al. The Highly Cooperative Folding of Small Naturally Occurring Proteins Is Likely the Result of Natural Selection. Cell . 2007;128(3):613–24.

61. Baird NJ, Srividya N, Krasilnikov AS, Mondragón A, Sosnick TR, Pan T. Structural basis for altering the stability of homologous RNAs from a mesophilic and a thermophilic bacterium. RNA . 2006;12(4):598–606.

62. Campbell ZT, Bhimsaria D, Valley CT, Rodriguez-Martinez J a., Menichelli E, Williamson JR, et al. Cooperativity in RNA-Protein Interactions: Global Analysis of RNA Binding Specificity. Cell Rep . 2012;1(5):570–81.

63. Swygert SG, Manning BJ, Senapati S, Kaur P, Lindsay S, Demeler B, et al. Solution-state conformation and stoichiometry of yeast Sir3 heterochromatin fibres. Nat Commun . 2013;5.

64. Kuriyan J, Eisenberg D. The origin of protein interactions and allostery in colocalization. Nature . 2007;450(7172):983–90.

65. Ptashne M, Ptashne M, Gann a, Gann a. Transcription initiation: imposing specificity by localization. Essays Biochem. 2001;37:1–15.

66. Moazed D. Mechanisms for the inheritance of chromatin states. Cell . 2011;146(4):510–8.

67. Hennig J, Militti C, Popowicz GM, Wang I, Sonntag M, Geerlof A, et al. Structural basis for the assembly of the Sxl–Unr translation regulatory complex. Nature. 2014;515(7526):287–90.

68. Levchenko a, Bruck J, Sternberg PW. Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. Proc Natl Acad Sci U S A . 2000;97(11):5818–23.

69. Kim AH, Yano H, Cho H, Meyer D, Monks B, Margolis B, et al. Akt1 regulates a JNK scaffold during excitotoxic apoptosis. Neuron . 2002;35(4):697–709.

70. Seeliger M a, Breward SE, Friedler A, Schon O, Itzhaki LS. Cooperative organization in a macromolecular complex. Nat Struct Biol . 2003;10(9):718–24.

71. Ruthenburg AJ, Li H, Patel DJ, Allis CD. Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol . 2007;8(12):983–94.

72. Ruthenburg AJ, Li H, Milne T a., Dewell S, McGinty RK, Yuen M, et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. Cell . 2011;145(5):692–706.

73. Kim H, Abeysirigunawarden SC, Chen K, Mayerle M, Ragunathan K, Luthey-Schulten Z, et al. Protein-guided RNA dynamics during early ribosome assembly. Nature . 2014;506(7488):334–8.

74. Woodson S a. RNA folding pathways and the self-assembly of ribosomes. Acc Chem Res . 2011;44(12):1312–9.

75. Williamson JR. Cooperativity in macromolecular assembly. Nat Chem Biol . 2008;4(8):458–65.

76. Ramaswamy P, Woodson S a. S16 throws a conformational switch during assembly of 30S 5' domain. Nat Struct Mol Biol . 2009 (4):438–45.

77. Ridgway D, Broderick G, Lopez-Campistrous A, Ru'aini M, Winter P, Hamilton M, et al. Coarse-grained molecular simulation of diffusion and reaction kinetics in a crowded virtual cytoplasm. Biophys J . 2008;94(10):3748–59.

78. Warner JR. The economics of ribosome biosynthesis in yeast. Trends Biochem Sci. 1999;24(11):437–40.

79. Wilson DN, Nierhaus KH. The weird and wonderful world of bacterial ribosome regulation. Crit Rev Biochem Mol Biol . 2006;42(3):187–219.

80. Kim C a, Gingery M, Pilpa RM, Bowie JU. The SAM domain of polyhomeotic forms a helical polymer. Nat Struct Biol . 2002;9(6):453–7.

81. Moazed D. Common themes in mechanisms of gene silencing. Mol Cell . 2001;8(3):489–98.

82. Brewer CF, Miceli MC, Baum LG. Clusters, bundles, arrays and lattices: Novel mechanisms for lectin-saccharide-mediated cellular interactions. Curr Opin Struct Biol . 2002;12(5):616–23.

83. Oppikofer M, Kueng S, Keusch JJ, Hassler M, Ladurner AG, Gut H, et al. Dimerization of Sir3 via its C-terminal winged helix domain is essential for yeast heterochromatin formation. EMBO J . 2013;32:437–49.

84. Li P, Banjade S, Cheng H-C, Kim S, Chen B, Guo L, et al. Phase transitions in the assembly of multivalent signalling proteins. Nature. 2012;483(7389):336–40.

85. Ferrell JE. Self-perpetuating states in signal transduction: Positive feedback, double-negative feedback and bistability . Current Opinion in Cell Biology. 2002. 140–8.

86. Dodd IB, Micheelsen M a., Sneppen K, Thon G. Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification. Cell. 2007;129:813–22.

87. Dodd IB, Sneppen K. Barriers and silencers: a theoretical toolkit for control and containment of nucleosome-based epigenetic states. J Mol Biol . 2011;414(4):624–37.

88. Grewal SIS, Moazed D. Heterochromatin and epigenetic control of gene expression. Allis CD, Jenuwein T, Reinberg D, editors. Science . Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2003;301(5634):798–802.

89. Qian H. Cooperativity in Cellular Biochemical Processes: Noise-Enhanced Sensitivity, Fluctuating Enzyme, Bistability with Nonlinear Feedback, and Other Mechanisms for Sigmoidal Responses. Annu Rev Biophys . 2012;41:179–204.

90. Chen L, Widom J. Molecular basis of transcriptional silencing in budding yeast. Biochem Cell Biol . 2004;82(4):413–8.

91. Baker D. A surprising simplicity to protein folding. Nature . 2000;405(6782):39–42.

92. Pepys MB, Hawkins PN, Booth DR, Vigushin DM, Tennent G a, Soutar a K, et al. Human lysozyme gene mutations cause hereditary systemic amyloidosis. Nature . 1993;362(6420):553–7.

93. Booth DR, Sunde M, Bellotti V, Robinson C V, Hutchinson WL, Fraser PE, et al. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. Nature . 1997;385(6619):787–93.

94. Canet D, Last AM, Tito P, Sunde M, Spencer A, Archer DB, et al. Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme. Nat Struct Biol . 2002;9(4):308–15.

95. Dumoulin M, Last AM, Desmyter A, Decanniere K, Canet D, Larsson G, et al. A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. Nature . 2003;424(6950):783–8.