Published in final edited form as:
Rep Prog Phys. 2012 July ; 75(7): 076601. doi:10.1088/0034-4885/75/7/076601.

Biomolecular Dynamics: Order-Disorder Transitions and Energy Landscapes

Paul C. Whitford,
Center for Theoretical Biological Physics and Department of Physics, Rice University 6100 Main, Houston, TX 77005-1827, USA

Karissa Y. Sanbonmatsu, and
Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, MS K710, Los Alamos, NM 87545, USA

José N. Onuchic
Center for Theoretical Biological Physics and Department of Physics, Rice University, 6100 Main, Houston, TX 77005-1827, USA

Abstract

While the energy landscape theory of protein folding is now a widely accepted view for understanding how relatively-weak molecular interactions lead to rapid and cooperative protein folding, such a framework must be extended to describe the large-scale functional motions observed in molecular machines. In this review, we discuss 1) the development of the energy landscape theory of biomolecular folding, 2) recent advances towards establishing a consistent understanding of folding and function, and 3) emerging themes in the functional motions of enzymes, biomolecular motors, and other biomolecular machines. Recent theoretical, computational, and experimental lines of investigation are providing a very dynamic picture of biomolecular motion. In contrast to earlier ideas, where molecular machines were thought to function similarly to macroscopic machines, with rigid components that move along a few degrees of freedom in a deterministic fashion, biomolecular complexes are only marginally stable. Since the stabilizing contribution of each atomic interaction is on the order of the thermal fluctuations in solution, the rigid body description of molecular function must be revisited. An emerging theme is that functional motions encompass order-disorder transitions and structural flexibility provide significant contributions to the free-energy. In this review, we describe the biological importance of order-disorder transitions and discuss the statistical-mechanical foundation of theoretical approaches that can characterize such transitions.

1. Introduction: The Biological Context for Physical Questions

The biomolecular origins of cellular dynamics occur over a wide range of length and time scales. Each biomolecule has specific functional properties that are governed by its energetics. Two well-studied classes of biomolecules (at the functional level) are proteins and nucleic acids. Both are linear polymer chains composed of specific sequences of monomers called “residues.” Proteins are composed of amino acid residues (Figure 1A), of which there are twenty naturally-occurring types. Ribonucleic acid (RNA, Figure 1B) chains and deoxyribonucleic acid chains (DNA) are composed of sugar-phosphate backbones and four types of nucleic acid side chains (A,U,C,G in RNA and A,T,C,G in DNA). While the primary role of DNA is to carry genetic information, RNA molecules can adopt well-ordered three-dimensional configurations, perform chemistry (1, 2) and function as

molecular machines (3–5). In fact, in recent years it has been shown that the majority of the human genome is transcribed into RNA but not translated in protein, suggesting that functional noncoding RNAs may be equally important to cellular dynamics as proteins. Due to the functional importance of protein and RNA structural dynamics, this review will focus on their dynamics, though the concepts presented may be applicable to other biomolecular systems.

Gene expression in the cell includes two fundamental steps: transcription and translation. Transcription is the process by which DNA is read and a complementary messenger RNA (mRNA) molecule is generated (Figure 2). In prokaryotes (single-celled organisms), mRNA is directly translated by ribosomes to produce proteins. In eukaryotes (multicellular organisms), precursor-mRNA (pre-mRNA) is produced during transcription. In eukaryotes, pre-mRNA is further modified by the cell (e.g. splicing (6)), which yields mature mRNA that is translated by ribosomes into protein sequences. In addition to this central dogma for the flow of genetic information, many feedback channels exist in the cell. For example, proteins and RNAs often regulate transcription and translation (7–13), as well as aid in heritable changes in gene expression by coordinating the structural reorganization of biomolecular assemblies (14, 15).

During transcription and translation the resulting linear polymers initially adopt unstructured (or partially-structured) ensembles of configurations. This is followed by a conformational search for the lowest-energy configuration. Since these polymers can have hundreds (even thousands) of residues (where each residue has several degrees of freedom) the accessible phase space of these molecules is large. The vastness of the available phase space led to the proposal of the so-called “Levinthal's Paradox.” (16) That is, if a polymer chain has N residues and each residue has M available conformations, the number of possible configurations of the polymer chain will be M^N. If each residue has only two possible conformations, then a 100 residue chain would have $2^{100} \approx 10^{30}$ possible configurations. However, functional proteins and RNA molecules rapidly find a specific “native” three-dimensional structure (often on sub-second timescales). If the search process were completely random, and a configuration could be sampled every picosecond, then the average folding time would be $\frac{1}{2}10^{30} \times 10^{-12}$ seconds $\approx 10^{18}$ seconds $\approx 10^{10}$ years. While considering the excluded volume of a biomolecule, or by assuming that proteins are initially collapsed (17), may reduce this counting problem, these arguments alone cannot account for the apparent eighteen orders of magnitude discrepancy in timescales. Accordingly, the search process must not be random. Identifying how proteins have evolved to fold to specific structures, with rapid rates, is often referred to as the “protein folding problem.” While the theoretical biological physics community has focused largely on protein folding, similar considerations apply for the folding of RNA (18).

To rationalize the kinetic disparities presented by Levinthal and biological timescales, the “funneled” energy landscape concept was introduced (19). Combined with The Principle of Minimal Frustration (20–22), this theoretical framework for the energy landscape of biomolecular dynamics has demonstrated that in order for proteins to be foldable, there must exist an ensemble of routes by which the molecule may reach the folded configuration. Folding then proceeds as a diffusive process over a relatively-smooth energy landscape (i.e. one that lacks large traps), and the ensemble of possible configurations continuously narrows as the native configuration is approached.

Since folding and functional dynamics occur on the same energy landscape, the principles developed to understand protein folding must extend to, and account for, biomolecular function. Many multi-domain proteins, such as Adenylate Kinase and kinesin, undergo

conformational rearrangements during their function (23–27)). Accordingly, the energy landscapes of these molecules facilitate both, rapid folding and functional dynamics. Since the energy landscape (i.e. the Hamiltonian[‡]) is determined by the chemical composition of the molecule, polymer sequences encode the energetics of the native configurations and large-scale functionally-relevant rearrangements, making it desirable to have a consistent theoretical framework for understanding both.

In this review, we provide a brief historical perspective (1970s-2000) on the development of the energy landscape theory of biomolecular motion, as developed in the context of glass-forming systems and protein folding. Next, we discuss recent work (2000-2010) to extend protein folding concepts and methods to describe functional dynamics of biomolecules. The focus of this discussion is on the use of theoretical models of energy landscapes that allow us to characterize the relationship between structure, folding and function. We will finish by discussing the role of order-disorder transitions in biomolecular functional dynamics.

2. Development of Energy Landscape Theory for Protein Folding

Grounded in statistical mechanics, the energy landscape theory of biomolecular dynamics (20–22, 28) has provided a general framework for understanding the folding and functional properties of biomolecules. For the Levinthal's Paradox to manifest, the energy landscape of a biomolecule must be relatively random. That is, if the landscape is "rough" [§] (Figure 3), the Levinthal search problem could arise. When the stabilizing energy gap δE between the folded and unfolded ensembles is large, relative to the energetic roughness ΔE, the landscape can be described as an energetic "funnel." (19) In a funneled energy landscape, configurations are locally connected (i.e. are kinetically accessible through simple moves) to states that are slightly more native and to those that are slightly less native. The degree of native content is also correlated with the stabilizing energy, resulting in a stabilizing energy gap that is much larger than the energetic barriers separating locally-connected states. In such a landscape, the diffusive walk is guided by the stabilizing energy and connectivity of states, and the Levinthal's paradox is averted.

In many proteins, folding may be described as a pseudo-first order phase transition (22, 29) where the system goes from a disordered ensemble to an ordered ensemble by transitioning through partially-folded configurations that are not minima on the free-energy surface. The ensemble of configurations that separate the endpoints is known as the Transition State Ensemble (TSE). Phase transitions manifest throughout nature, suggesting that analogies between proteins and other physical system (such as spin glasses) may be identified. In spin systems, when the energetic roughness is small, the systems go through phase transitions to ordered states. Similar to glass-forming systems, when the energetic roughness in a protein is large (relative to the energy gap), the system is described as "frustrated" (22, 30–34). When there is a large degree of frustration, the fluctuations in energy that arise from non-zero temperatures are insufficient to escape from local minima and the timescales of the system's dynamics diverge.

2.1. Spin Glasses

To introduce the relationship between order-disorder transitions, glassy dynamics and energetic frustration, which are fundamental aspects of the energy landscape framework for biomolecular dynamics, we will discuss these properties in the context of spin glasses (35–37). This is not intended to be an exhaustive exploration of spin glasses and their dynamics.

[‡]In this review, "Hamiltonian" will refer to the potential energy contribution only.
[§]i.e. the energetic barriers separating locally-connected configurations are large, relative to the overall energy gap between the folded and unfolded ensembles.

Rather, we will focus on features of order-disorder systems that will assist in our understanding of the energy landscapes of biomolecules.

The textbook example for discussing the statistical mechanics of phase transitions is the Ising model. (38) The Ising model represents a ferromagnetic material as being composed of a lattice of discrete sites (in any number of dimensions), where each site adopts a spin value of +1, or -1 (Figure 4). In its original formulation (39), each spin is coupled to the adjacent spins and the energy for an arbitrary-dimensional system is given by:

$$H = -\sum_{\{ij\}} J_{ij} s_i s_j, \quad (1)$$

where $\Sigma_{\{ij\}}$ denotes the sum over all adjacent spins (s_i, s_j) and J_{ij} is the magnetic interaction energy of spins s_i and s_j. When $J_{ij} > 0$ (ferromagnetic interaction) for every pair, the lowest energy configuration of the system corresponds to all spins being aligned. If $J_{ij} > 0$, the system is described as “unfrustrated” since the interactions are consistent. That is, every microscopic interaction may be satisfied in the lowest energy configuration of the system and, at low temperatures, the magnitude of the average magnetization $\langle m \rangle$ approaches 1.

Ising systems of two or more dimensions (and $J_{ij} > 0$) undergo second-order phase transitions between disordered ($\langle m \rangle = 0$) and ordered states ($|\langle m \rangle| \approx 1$) (38) at the so-called critical temperature T_C. For $T < T_C$, the thermal fluctuations are sufficiently small that the stabilizing interactions are collectively satisfied and the majority of the spins align in a single direction. For $T_C > T > 0$ the system adopts a ferromagnetic state, though transient anti-ferromagnetic configurations are sampled. If J_{ij} is not constant for all pairs, then the probability of anti-ferromagnetic orientations are heterogeneously distributed across the system. Due to the Boltzmann weighting of configurations, this feature is obvious, though it is noted since biologically-relevant parallels will be drawn later.

In 1975, Sherrington and Kirkpatrick (SK) introduced an extended and exactly solvable Ising model that exhibits a glass phase (35). Rather than have $J_{ij} \neq 0$ for only the adjacent pairs, the J_{ij} coupling constants are defined for all spin pairs (i.e. infinite-range interactions), and the probability distribution of the constants is given by:

$$p(J_{ij}) = [(2\pi)^{1/2} J]^{-1} \exp[-(J_{ij} - J_0)^2 / 2J^2]. \quad (2)$$

In the limit $J \rightarrow 0$, the SK spin system reduces to the unfrustrated Ising model, and $J_{ij} = J_0$ for all pairs. As $J \rightarrow \infty$ the interactions are completely random and the system does not exhibit long-range order ($\langle m \rangle \rightarrow 0$). For large J (relative to J_0), the system is energetically frustrated, and not all interactions can be satisfied simultaneously.

There are several general features of the SK Ising model that are most relevant when describing the energy landscapes of biomolecules. Figure 4 shows the phase diagram for the SK model, in terms of the normalized quantities $\tilde{J}_0 = NJ_0$ and $\tilde{J} = N^{1/2}J$, where N is the number of spins in the system. The axes correspond to ratios of competing energetic parameters: $k_BT/\tilde{J}$ and $\tilde{J}_0/\tilde{J}$. $k_BT/\tilde{J}$ is the ratio between thermally-accessible energetic fluctuations and the heterogeneity of the interaction weights. For large T, individual spins have sufficient thermal energy to take any value and the system enters a paramagnetic state ($\langle m \rangle \rightarrow 0$). At low T, the system can enter a ferromagnetic state, or a spin-glass state. In the ferromagnetic state the individual spins are aligned ($|\langle m \rangle| \rightarrow 1$). In the spin-glass state, individual spins may align locally, but there is no long-range ordering. For low values of

$k_BT/\tilde{J}$, J_0 ~$/\tilde{J}$ determines whether the ferromagnetic or glass phase is adopted, where $\tilde{J}_0$ is the average coupling and $\tilde{J}$ is the dispersion of the coupling constants. When $\tilde{J}_0/\tilde{J} > 1$, the system may adopt a ferromagnetic or paramagnetic state. If $\tilde{J}_0$ is large relative to $\tilde{J}$, then the majority of the system's interactions will be ferromagnetic, and every microscopic ferromagnetic interaction may be satisfied in a global ferromagnetic state. If the dispersion is large, relative to the average coupling (i.e. $\tilde{J}_0/\tilde{J} < 1$), the system is composed of a mixture of ferromagnetic and anti-ferromagnetic interactions, and it is described as "energetically frustrated." In frustrated systems, not all interactions can be satisfied simultaneously (i.e. where each microscopic interaction is a minimum). Accordingly, if $\tilde{J}_0/\tilde{J}$ and $k_BT/\tilde{J}$ are small, the system will enter a glassy state. Hallmark features of glassy systems are a lack of long-range ordering ($\langle m \rangle \rightarrow 0$) and diverging relaxation times. Diverging relaxation times are due to interactions being satisfied locally, though the low-energy configurations are highly degenerate, which leads to competition of local structure formation. For a more detailed discussion on the thermodynamic properties of the SK model, we suggest the original manuscript (35).

Glass transitions are not isolated to Ising models. They are ubiquitous in nature and have been investigated computationally (40–42), theoretically (43–46) and experimentally (47–49) for a wide variety of systems. Here, we have limited the discussion to the general relationships between energetics, order-disorder transitions and glass dynamics, in order to guide our description of biomolecular dynamics.

2.2. Some Parallels Between Glasses and Protein Folding

While the energetics of biomolecules are not simply composed of spin-spin coupling interactions, analogous behaviors between biomolecular folding and glass transitions may be identified (49). Most proteins spontaneously fold under physiological conditions, and they do so in relatively short timescales (milliseconds). Similarly, in the Ising model, if the dispersion in the interactions is small, then a paramagnetic-to-ferromagnetic transition occurs without the formation of long-lived kinetic traps. In the SK spin model, when $1 < \tilde{J}_0/\tilde{J} < 1.25$, there are three accessible states. Consistent with this, in proteins, the adopted state is determined by the temperature-to-dispersion ratio $k_BT/\tilde{J}$. The paramagnetic:ferromagnetic:glass states thus correspond to unfolded:folded:misfolded ensembles. In polymers, glassy dynamics may emerge at low temperatures (50), where the glassy "state" is composed of many competing configurations and interconversion times between them is slow. Finally, for a given value of $\tilde{J}_0/\tilde{J}$ (between 1.0 and 1.25), these systems have well-defined order-disorder transition temperatures (paramagnetic-ferromagnetic transition, or folding-unfolding transition) and a glass transition temperatures (ferromagnetic-glass, or folded to glass/misfolded).

2.3. Energy Landscape Theory and The Principle of Minimal Frustration

The similarities between glass-forming systems and protein folding led to the proposal of The Principle of Minimal Frustration (19, 51, 52). That is, naturally occurring protein sequences are able to fold on short timescales, which indicates these sequences encode energetics for which the folding temperature T_f (i.e. the temperature at which the unfolded and folded ensembles are equally probable) is larger than the glass transition temperature T_g. Further, the principle states that nature has selected for protein sequences that maximize the ratio between T_f and T_g, which allows proteins to fold rapidly in vivo (i.e. $T_g < T_{cell} < T_f$). T_f/T_g in proteins parallels $\tilde{J}_0/\tilde{J}$ in the SK spin-glass model. $\tilde{J}_0$ is the average coupling strength of the spins, which is analogous to the average stabilizing energetic contribution of each native interaction in a protein. In the words of Gô (53), $\tilde{J}_0$ is similar to the degree of "consistency" between the native interactions in a protein. These consistent interactions lead to an energy gap δE between the disordered and ordered states (Figure 3). $\tilde{J}$ is the dispersion

in the strength of the spin interactions and it is correlated with the degree of frustration in the system. In proteins, frustration manifests as stabilizing non-native interactions (34). According to these corresponding features of proteins and glasses, The Principle of Minimal Frustration states that by maximizing the ratio between the stabilizing energy gap δE and the energetic roughness ΔE (which determines the ratio T_f/T_g), nature has selected for protein sequences that are able to fold rapidly.

3. Modeling Approaches for Protein Folding

Over the last 20 years, many theoretical and computational studies have investigated the protein folding problem. Of particular relevance to this review are studies that have extended the energy landscape framework for biomolecular folding and function. We will focus on advances that have employed structure-based models (SBMs) (54–56), since these models are founded on The Principle of Minimal Frustration. A common theme throughout these investigations is that by accounting for simple geometrical considerations, such as excluded volume and chain connectivity, dynamic features of protein folding and function may be accurately described. That is, once a protein has evolved to have a dominant energetic basin (or several basins), the geometry of the protein is the primary determinant of the rate/mechanism.

Energy Landscape Theory provides a general physical framework for understanding the energetics of protein folding, which allows for many classes of theoretical models to be employed. Early work used lattice models of proteins, where each residue is described as a single bead that moves between lattice vertices. With such models, the interplay between global collapse, connectivity of states and energetics was elucidated. Subsequent work employed more easily transferable coarse-grained off-lattice models. Recently, we have developed models that combine all-atom resolution with the energy landscape framework. To clarify the utility of each approach, we will provide an overview of how each class of models has advanced our understanding of protein folding. By considering the advantages and limitations of each approach, future investigations may strategically utilize the models that are best suited for each question.

3.1. Lattice models of protein folding

Inspired by the glass-theoretical roots of The Principle of Minimal Frustration, early simulation models for protein folding also used lattice representations of polymer chains to explore the relationship between energetic frustration, kinetics and structural mechanisms (19, 57, 58). In these models, the polymer is treated as a connected chain, where each amino acid residue is a link that transits across a discrete set of lattice positions (Figure 5). The functional form of the Hamiltonian used in many studies is similar to an Ising spin model:

$$E = -\sum_{\{ij\}} \epsilon(s_i, s_j)\Delta(i, j) + \sum_k \delta(|\vec{r}_k - \vec{r}_{k+1}|). \quad (3)$$

The chemical identity of the primary residue sequence S is defined as $s_1,\ldots, s_N$, $\epsilon(s_i, s_j)$ is the coupling constant between residue types, $\Delta(i, j) = 1$ if the residues are on adjacent sites and $\Delta(i, j) = 0$ otherwise. The summation over $\{ij\}$ is the sum over all residue pairs (not a sum over all lattice sites). The second term sums over all residues that are adjacent in sequence, ensuring the connectivity of the backbone (grey lines in Figure 5) is maintained. $\delta = 0$ if the residues are on adjacent sites and $\delta = \infty$ if they are not. Changing notation reveals the similarities between this model and the Ising model:

$$E = -\sum_{\{ij\}} J(R(i), R(j)) O_i O_j + \sum_{k=1}^{N-1} (|\vec{r}_k - \vec{r}_{k+1}|) \quad (4)$$

where $\{ij\}$ is now defined for *all adjacent* lattice sites. O_i is the occupancy of each lattice site (1 if occupied and 0 if unoccupied), $J(R(i), R(j))$ defines the coupling between the types of residues in the sites, and $R(i)$ is the type of residue at site i. $R(i)$ carries the same type of information as s_i. However, $R(i)$ is the type of residue at site i and s_i is the residue type in sequence position i. Accordingly, $R(i)$ is not defined if site i is unoccupied, in which case O_i is also equal to zero. In this specific formulation, there are no long-range interactions, though alternate formulations have included a combination of long-range and short-range interactions (59, 60).

The only differences between the Ising spin Hamiltonian and protein lattice Hamiltonian are the connectivity restriction (δ function in Equations 3 and 4) and the intrinsic time dependence of the site-site coupling constants $J(R(i), R(j))$ [‖]. While these complications have impeded efforts to find an analytic solution for the model, the relatively small number of discrete configurations on the lattice allows all compact configurations to be explicitly enumerated for small system (i.e. a 27-mer) (29, 30, 61–63).

3.1.1. Utility of Lattice models—While lattice models employ a simplified representation of the atomic structure and energetics, they possess several features that allow for a general understanding of polymer folding to emerge. First, there is specificity in the sequence of the monomers. Not only do individual beads (residues) have unique energetic content, but adjacent residues (in sequence) are constrained to be on adjacent vertices (i.e. covalent geometry is maintained). Another physically valid constraint in the model is that monomers can not pass through one another. In an Ising system, it is possible for adjacent lattice sites to simultaneously exchange values (e.g. position i goes from spin -1 to +1 when position i +1 goes from +1 to −1). In lattice models of proteins (and in real proteins) the repulsive nature of atomic nuclei lead to an excluded volume that prevents this from occurring.

The low computational cost of Ising-like lattice models for protein folding has allowed for extensive exploration of the balance between energetic stability, entropic contributions and energetic frustration in protein folding (19, 64–68), and has elucidated new strategies for protein design (69). One of the most significant results stemming from lattice model studies was the folding funnel paradigm for biomolecular folding. This concept emerged from investigation into the energetic, thermodynamic and kinetic differences between “foldable” and “unfoldable” polypeptide sequences. To rationalize the apparent paradox implied by the arguments of Levinthal, Leopold et al. (19) used a lattice model to simulate folding for two protein sequences. While the beads were not assigned energetic values based on a particular polypeptide sequence, the energetic parameters were designed, such that one sequence was foldable. That is, the sequence was designed to meet the following criteria: 1) 20 unique bead “flavors” were used, which reflects the 20 naturally-occurring residue types in proteins, 2) the overall strength of each bead-bead interaction was on the scale of a few $k_B T$ and 3) the potential energy had a dominant energetic basin. This small set of parametric requirements describe an “ideal” protein that has a unique native configuration. The non-foldable sequence was randomly assigned bead flavors, which represents a random heteropolymer that does not have a unique native configuration. While both sequences do have global energetic minima, it was shown that the folding time of the unfoldable sequence

[‖]Since residues move across the lattice, $R(i)$ changes with time.

was far greater than in real proteins. This is consistent with the arguments of Levinthal, where the search for the native configuration on a rough energy surface is also random, and the likelihood of finding the native configuration becomes vanishingly small. In contrast, for the foldable sequence, folding could be described as a diffusive process over an initially broad ensemble of configurations (i.e. the unfolded ensemble), and the kinetically connected sub-ensembles decrease in size as the native basin is approached. Essentially, the energy landscape resembled a "funnel," where the native ensemble resides at the bottom (Figure 3).

3.1.2. How to describe folding: Reaction coordinates—Prior to the advent of lattice models for protein folding, kinetic modeling of protein folding involved a traditional biochemical description, where the unfolded and folded basins are separated by a single transition state. Since the free-energy had not been evaluated or measured directly at that time, the landscape was considered in terms of a generalized, ideal reaction coordinate, where approximations are necessary to connect the folding rates and free-energy barriers. Since lattice models provide pseudo-atomic resolution of the folding dynamics, they necessitated rigorous methods for charactering the collective motion of biomolecular processes. That is, the energy landscape of a system is defined in terms of $3N$ degrees of freedom, where N is the number of atoms. In these landscapes, the free-energy barrier of folding is described by the transition state ensemble (TSE) that separates the unfolded and folded ensembles. One of the objectives of statistical mechanics is to describe the collective dynamics of complex system along a low number of reaction coordinates, or order parameters. In the case of protein folding, the first questions to address were: Are there simple structural reaction coordinates that can capture the essential properties of a folding process? And, what are the appropriate coordinates?

There are several requirements for an appropriate reaction coordinate. First, for a given system configuration, the coordinate must be uniquely defined. Second, each basin and TSE must have distinct values of the coordinate. Third, local moves in Cartesian space must map continuously onto the reaction coordinate. If this requirement is not satisfied, then this coordinate can not describe diffusive motion, since discontinuities will result in the projection of ballistic-like tunneling events in that coordinate space. For a given coordinate, the potential of mean force and diffusion coefficient must also be consistent with the kinetics of interconversion between the local basins (See *Relationship between free-energy barriers and kinetics* below). Finally, by definition, a configuration in the TSE must be equally likely to proceed to either of the adjacent free-energy minima, and the coordinate must be able to identify a TSE with this property. Since lattice models provided the first opportunity to use atomistic-based reaction coordinates, two intuitive reaction coordinates were introduced: the radius of gyration R_g and the fraction of native contacts Q. Following the initial studies that employed these coordinates, many subsequent studies have investigated the relationships between diffusion, barrier-heights and kinetics along them, in both lattice and off-lattice models (70–78). In the context of off-lattice models, it has been shown that Q can satisfy the above requirements when describing folding of two-state proteins (i.e. two basins separated by a single barrier) (79). In contrast, R_g is able to distinguish between the folded and unfolded configurations, but it is less able to distinguish between native and near-native configurations. However, by considering the folding as a function of both coordinates, folding may be well described as collapse (decrease in R_g) followed by reorganization (increase in Q) (19, 29, 57, 64, 65, 80–83).

3.2. Off-Lattice Models of Protein Folding

3.2.1. C_α models—While lattice models enabled a general characterization of heteropolymer collapse, real proteins do not reside on lattices. Rather, the coordinates of proteins take continuous values in three-dimensional space. To provide a continuous

description of the phase space, simplified models that employ an off-lattice representation have been explored (31, 54, 84–92). Early off-lattice models employed a coarse-grained representation of the polymer, where each residue was represented as a single sphere (Figure 6C). With an off-lattice representation, it is possible to study specific proteins and not only model systems. That is, while it is possible to design lattice configurations that resemble the structures of specific proteins, the off-lattice description is transferable to any protein, which has enabled the exploration of the relationship between chain connectivity and excluded volume on the folding properties of many proteins.

The Principle of Minimal Frustration states that the energy gap between the unfolded ensemble and folded ensemble must be large, relative to the energetic roughness (i.e. frustration). Inspired by this, many off-lattice models employ forcefields that are completely specific to the native, folded configuration. These unfrustrated forcefields represent the limiting case of The Principle of Minimal Frustration. Similar to an ideal gas theory, these unfrustrated models represent a low-order approximation to a molecule's energetics, and thus provides a baseline framework upon which the role of energetic perturbations and roughness may be better understood (54, 81, 93–96). For example, this approach has been used to illuminate the effects of cellular crowding (97) and desolvation (81) on protein folding, folding under non-equilibrium conditions (87, 98–101), co-translational folding (102) and aggregation processes (103).

Clementi et al. introduced one of the most popular off-lattice models for protein folding, which is referred to as a Structure-Based Model (SBM), or a Go(x00302)-like model (54). In its initial formulation, the folded configuration is defined as the lowest-energy configuration. Each residue-residue interaction formed in the native configuration is assigned an attractive Lennard-Jones-like interaction, and the minimum energy corresponds to the residue-residue distances found in the native configuration. Adjacent residues are restrained by harmonic interactions, which account for the covalent backbone geometry of the system. The torsion angles formed by residues i, $i+1$, $i+2$, $i+3$ are given cosine functions, each with a minimum corresponding to the native configuration. The functional form of this potential is:

$$\begin{aligned} \mathrm{V} = & \sum_{i}^{\text{bonds}} \epsilon_r^i (r^i - r_o^i)^2 \\ & + \sum_{i}^{\text{angles}} \epsilon_\theta^i (\theta^i - \theta_o^i)^2 \\ & + \sum_{i}^{\text{dihedrals}} \epsilon_\varphi^i \{[1 - \cos(\varphi^i - \varphi_o^i)] + \tfrac{1}{2}[1 - \cos(3(\varphi^i - \varphi_o^i))]\} \\ & + \sum_{\text{ij}}^{\text{contacts}} \epsilon_C^{\text{ij}} \left[5\left(\frac{\sigma^{\text{ij}}}{r}\right)^{12} - 6\left(\frac{\sigma^{\text{ij}}}{r}\right)^{10} \right] + \sum_{\text{ij}}^{\text{non-contacts}} \epsilon_{\text{NN}}^{\text{ij}} \left(\frac{\sigma_{\text{NN}}^{\text{ij}}}{r}\right)^{12} \end{aligned} \tag{5}$$

Where ϵ_r^i=100, ϵ_θ^i=20, ϵ_φ^i=1, ϵ_C^{ij}=1, $\epsilon_{\text{NN}}^{\text{ij}}$=1, and $\sigma_{\text{NN}}^{\text{ij}}$=4.0 $\mathring{A}$.$r_o^i, \theta_o^i, \varphi_o^i$ and σ^{ij} are given the values found in the native structure. The bond and angle terms maintain the chain-like connectivity of the polymer. The 10-12 potential is used to model each native contact. While in this implementation the interactions are completely additive, other off-lattice models include non-additive (or, 3-body) interactions (104, 105), which is sometimes necessary to recover the proper degree of folding cooperatively.

In this model, since the solvent is not explicitly represented, the stabilizing 10-12 interactions are solvent-averaged effective interactions. That is, the native configuration has evolved to be the lowest free-energy configuration. Therefore, the native interactions are effectively stabilized, even though direct residue-residue interactions may be unfavorable.

For example, it is possible for a native configuration to have two positively charged residues near each other, even though the direct interaction between like charges is repulsive. However, since these residues are in close proximity in the lowest free-energy configuration, the potential of mean force between them has a minimum at the native distance. According to this logic, the native interactions in these forcefields are “effective” interactions that are averaged over the contributions of water and non-native interactions.

Structure-based models have forcefully demonstrated that chain connectivity and excluded volume significantly restrict the possible folding routes of biopolymers. This restriction of phase space can lead to a feature known as “Topological Frustration” (31, 106–111). Topological frustration may be characterized as the requirement of particular subunits to fold first, regardless of the relative stabilities. That is, it is possible that formation of particular native interactions will impose strong structural constraints on the system, leaving insufficient space for the chain to reorganize and form additional native interactions. Topological frustration is not energetic in the classical sense, where attractive interactions lead to large escape barriers from non-native configurations. Rather, similar to how excluded volume may lead to caged dynamics, where large barriers are associated with reorganization in collapsed polymers (112), in a topologically-frustrated fold the early formation of specific interactions can introduce energetic barriers. If the “incorrect” contacts form prematurely, the excluded volume of each atom will lead to high-energy barriers that impede folding. Therefore, it would more likely that the protein will return to the unfolded ensemble than continue folding. Topologically frustrated proteins must search for the appropriate initial interactions that are able to facilitate additional native-structure formation.

To illustrate ways in which topological frustration can manifest, let us compare the folding mechanisms of two small proteins: Protein A and SH3 (Figure 7). Protein A is a three-helix bundle protein and SH3 is composed of several beta sheets. In the folding of protein A, the transition state ensemble has rather homogeneous structural content (113). This homogeneity of the transition state ensemble suggests that the order of structure formation not crucial. Protein A is composed of α-helices that are formed by many interactions between residues that are close in sequence. Each turn of the helix results in interactions between residue i and i+4. In addition to helical interactions, the interface between the helices introduces non-local interactions. By inspecting the native structure, one can easily envision how individual turns of the helix may be relatively disordered while some interface interactions are maintained. In contrast, SH3 has a very polarized transition state ensemble, where some regions are highly ordered and others are largely disordered.

Two classes of proteins that are considered highly topologically frustrated are knotted proteins (114, 115) and the cytokine Interleukin family of proteins (108, 110, 116, 117) (Figure 8). A knotted structure may be defined by the following criteria: If the terminal ends of the chain are pulled in opposite directions, it will not reach a fully-extended structure. Rather, pulling results in a compact knot. Knots have been identified in a number of proteins and recent work has employed simplified models to characterize how complex native structures can be reached during folding (118, 119). Entropic considerations suggest that the most favorable way to fold would be to first form a loose loop and then thread the terminal end of the protein through the loop. Subsequent local rearrangements then lead to a completely folded configuration. If the loop were to close (i.e. fully form) early in the folding process, then there would be insufficient space for the threading process to occur, which highlights the topological difficulties associated with folding these proteins.

The excluded volume of individual atoms can also lead to backtracking in proteins (91, 107, 120, 121), including knotted ones (118, 119). Backtracking is defined as a phenomenon where native structure forms, breaks and reforms during forward progress of the global

folding process. In the case of IL-1β, when residues in the C-terminal trefoil region possess native structural content early, as folding continues to proceed, this region temporarily unfolds prior to it reaching a fully-folded configuration. As discussed above, one hallmark of topological frustration is that the protein returns to the unfolded ensemble if the "wrong" interactions form prematurely. In the case of IL-1β and backtracking, the unfolding is localized, where specific structural units reversibly unfold and refold as the forward global folding process takes place.

3.2.2. All-atom models—Following the progression from lattice to off-lattice C_α models, the next logical step was to extend biomolecular models to all-atom (AA) representations[¶]. In all-atom structure-based models (56, 122–128), every heavy atom (and, in some cases, hydrogen atoms) is explicitly represented. Similar to its C_α counterpart, every interaction in the native configuration is stabilizing, which ensures the native structure is the lowest-energy configuration. Adopting this representation has two advantages that are immediately evident. First, by including all atoms, the excluded volume of the atomic structure is explicitly accounted for, which allows one to characterize the relationship between the details of atomic packing and the functional/folding properties. Through comparison with a C_α model, it was shown that side-chain packing can significantly influence the folding of some proteins. Specifically, proteins that have spatially-symmetric backbone configurations can have two or more apparently equivalent folding pathways, when described by a C_α models. A second motivation for using an AA model is that non-specific atomic interactions (i.e. interactions that are not formed in the native configuration) may be explicitly included, where the native-centric SBM would serve as a baseline model. That is, one may partition the effects of individual energetic contributions on the folding process (or other molecular processes) by comparing the dynamics in an unfrustrated model and a model where non-specific attractive interactions are included.

While the cooperative transition associated with protein folding encompasses a large change in configurational entropy of the backbone, the contribution of side chain entropy is less clear. We recently extended the C_α SBM to an all-atom representation and detected a partial separation between backbone ordering and side-chain ordering (56). These models yield a peak in the specific heat curve, indicative of a phase transition between the disordered (unfolded) and ordered (folded) ensembles, which is coincident with backbone ordering. However, this transition is accompanied by only partial ordering of the side chains. Figure 9 shows a typical trajectory for a single protein, described by the radius of gyration R_g (which measures chain collapse), the fraction of all-atom native contacts formed Q_{AA} (which characterizes side-chain packing) and the fraction of $C_\alpha - C_\alpha$ contacts formed $Q_{C\alpha}$ (which measures tertiary structure formation). Both $\langle Q_{AA} \rangle$ and $\langle Q_{C\alpha} \rangle$ undergo sharp transitions at T_f (Figure 9B), but $\langle Q_{C\alpha} \rangle$ is larger than $\langle Q_{AA} \rangle$. This separation of backbone and side-chain ordering has also been noted in recent explicit-solvent simulations (129), suggesting it is a robust feature of protein folding. Figure 9D shows which residues exhibit ordered backbones with disordered side chains in the folded ensemble (highlighted in blue). It is striking that disorder in the side chains is not homogeneous between all residues. Even though the residue-residue contacts indicate that the tertiary structure of the molecule is formed, specific regions have more highly-disordered side chains, which is similar to anisotropies in ferromagnetic-phase spin-glass systems. The disorder accessible in a folded protein raises the question: Since folding does not necessitate the ordering of all side chains, does side-chain mobility fulfill a functional purpose? Loop residues are often involved in protein recognition processes (130–132). The predicted flexibility of side-chain

[¶] C_β models have also been studied, but there are more significant structural differences when moving to an all-atom representation, and we will focus attention there.

conformations in folded proteins is consistent with the notion that loop sequences may evolve in response to functional pressure that is independent of pressure to folding efficiently. Loop flexibility may also allow for conformational adaptability, which may enable proteins to participate in parallel signaling pathways. An alternative role of side-chain flexibility is that balancing order-disorder transitions in the side chains may be a way in which proteins modulate the free-energy barriers of interconversion between functional conformations (133–135).

All-atom models have highlighted modes by which side-chain dynamics may guide the folding process. For example, the experimentally-characterized disparate folding properties of protein L and protein G (two proteins that share the same global fold. Figure 7) could be accounted for by subtle differences in packing interactions of the side chains (122). Since these β-sheet structures are symmetric at a coarse (i.e. residue) level, the details of the atomic packing (between the α helix and β-sheet) significantly influence the predicted folding rates and mechanisms. Recently, it was also shown how orientational information contained in side-chain packing reduces the prevalence of malformed knots in proteins (119). Specifically, at the residue level, the energetic penalty of “crossing” mistakes is small, which can lead to the native configuration becoming kinetically inaccessible.

Similar to studies with C_α models, using all-atom models, one may ask: What dynamic properties are robust, and which may be more easily manipulated in order to gain precise functional control? Using an all-atom structure-based model, we investigated the effects of energetic perturbations on the folding properties of several well-studied proteins: CI2, SH3 and Protein A (56). To no surprise, the thermodynamic quantities were sensitive to the energetic details. Consistent with the theoretical work of Portman and Wolynes (136), these models demonstrated how the balance between chain persistence length and non-local stabilizing interactions impacts the free-energy barriers of folding. That is, as the persistence length is increased, the free-energy barriers also increase. Consistent with the findings from C_α models (137, 138), the folding mechanisms were largely robust, though there was a small degree of variation. Protein A, which is considered a simple fold, was most sensitive to the distribution of energetics, while SH3 was most robust. This is consistent with the SH3 fold being topologically frustrated (106). In topologically frustrated folds, the phase space that the protein may sample en-route to the native configuration is reduced, relative to a simple fold. The fact that the SH3 folding mechanism was only marginally perturbed by changes in energetics is consistent with this highly restricted phase space.

4. Theoretical Approaches to Investigate Conformational Rearrangements

Recent work has expanded the energy landscape approach beyond folding, in order to describe large-scale conformational transitions associated with function. Oftentimes, biomolecules do not possess a single well-defined native configuration. Rather, there are competing basins of attraction and the balance between the conformations governs the biological functions of the molecule (139–141). Some enzymes undergo conformational rearrangements upon substrate binding (142, 143), taking the molecule from an enzymatically inert state to a conformation that is competent for chemistry. These rearrangements can even be rate limiting in the catalytic cycles of enzymes (144). In molecular machines, processive functions require conformational cycling through functional configurations, such as the “walking” of kinesin on microtubules (145) and the elongation cycle of the ribosome (146–148). Molecular machines that are not processive may also utilize large-scale rearrangements during function (e.g. the spliceosome (6)).

At first glance, it may be tempting to characterize functional rearrangements as rigid-body movements about well-defined “hinge” regions, since macroscopic machines function in this

manner. While this approximation may be accurate in some cases (149), applying this framework to all biological systems can be misleading. In macroscopic machines, the rigid components are “infinitely” rigid (orders of magnitude stronger than the man-made hinges). The interactions that maintain macroscopic rigid bodies are extremely strong, relative to thermal energy, which allows macroscopic machines to maintain their structural integrity over a wide range of temperatures and conditions. Therefore, rigid macroscopic components undergo negligible changes in entropy during function. In contrast, biomolecules are only marginally stable, where a 10-20 % increase in temperature often results in high-entropy unfolding to occur. Further, in molecular machines, the weak interactions (i.e. on the scale of thermal fluctuations) that stabilize the folded configurations also govern the dynamics of functional motions. Since folding and function take place on a single energy landscape, we must expand our energy landscape approaches for folding to account for functional dynamics.

In this section, we will discuss current theoretical and computational approaches that are used to explore biomolecular functional energy landscapes. A challenge in these investigations is that in order for these molecules to interconvert between stable configurations, the native basin must not be uniquely defined. Since the folding and functional landscapes of a biomolecule are one and the same, the funneled energy landscape for folding (19) has to be extended to include competing basins. Similar to studies on protein folding, simple models allow for a low-order approximation to the energetics, which enables one to identify which features of the functional dynamics are most robust and which are most responsive to changes in energetics and cellular/experimental conditions. By partitioning the robust features, one may design new strategies to precisely target and regulate biomolecular dynamics.

4.1. Normal Mode Analysis

In normal mode analysis (NMA), the potential energy is expanded about a particular minimum to second order, and the local motions are described as a superposition of orthogonal oscillatory motions (150). In solution, biomolecules exhibit overdamped dynamics and their trajectories do not follow harmonic modes with precise frequencies. However, the energy local to the minimum may be expanded in terms of the modes:

$$E=\sum_n \frac{1}{2}\omega_n^2 q_n^2, \quad (6)$$

where the normal coordinate of mode n, q_n, has a characteristic frequency ω_n. Each eigenvector of the Hessian matrix (i.e. the normal mode) corresponds to a single degree of freedom, and the equipartition theorem dictates that each mode has an average energy of $1/2k_BT$. It then follows that $\langle q_n^2 \rangle \propto \omega_n^{-1}$, such that large-scale motions occur along the lowest-frequency modes. Interestingly, normal mode analysis of proteins has repeatedly demonstrated that low-frequency modes correspond to global deformations (151–153) that are in functionally-relevant directions (154).

A commonly-used potential energy function used for NMA of proteins is the Tirion potential (155). Similar to SBMs, the Tirion potential defines the native configurations as the lowest energy configuration and atoms that are near each other are assigned harmonic potentials, where the minima are assigned based on the initial configuration. Consistent with SBMs for folding, since the initial structure is typically a stable configuration, the Tirion potential provides a low-order approximation to the landscape about that minimum.

Additionally, this simple approximation has been shown to accurately reproduce the local fluctuations about stable protein configurations (133, 143, 156–158).

Low-frequency motions are largely determined by the global structure, and not specific energetic details (154). While individual interactions are important for the stability of the molecule, the connectivity of bonded and non-bonded interactions restricts the possible fluctuations, such that many Hamiltonians yield similar global motions about a chosen basin (159). These low-frequency motions in proteins and molecular complexes are also often correlated with functionally-relevant motions (160). In other words, the global structure determines the accessible local motions and these motions correlate with the large-scale changes required for function. Therefore, this provides a direct connection between molecular structure and function. The ribosome (Figure 10) is an excellent example that illustrates this relationship. Analysis of local-basin fluctuations have suggested that the movement of “stalk” regions is lower in energy than most other motions. While initial studies employed NMA with harmonic potentials (153, 153, 158, 161, 162), subsequent studies found similar results using all-atom structure-based simulations (163) and all-atom explicit-solvent simulations (164). These independent classes of theoretical models employ different assumptions and parameterizations. The fact that the fluctuations are so similar confirms that the dominant contributor to the local fluctuations is the overall structure of the molecular complex.

Since the local fluctuations described by NMA often correlate with functional motions, ongoing efforts are extending these methods to uncover the pathways between specified endpoints (133, 156, 165, 166). Miyashita et al. developed an iterative NMA approach to obtain a sequence of structures that represent a low-energy pathway between two functional configurations of the protein Adenylate Kinase (Figure 11). To do so, they first defined a Tirion potential for the initial configuration, calculated the normal modes and then displaced the structure along several low-frequency modes (where movement along each mode was proportional to the overlap with the displacements between endpoint configurations). After making a small displacement (such that the harmonic approximation was valid), the Tirion potential was re-defined for the new configuration, the modes were re-calculated and the system was displaced again. This process was repeated to generate a sequence of configurations along the modes that have the highest-overlap with the rearrangement. In this approach, only a few low-frequency modes were necessary to describe the majority of the global displacements. Following these initial studies, Adenylate Kinase has become the paradigm protein system for testing new concepts and methods in protein conformational transitions (133, 156, 165, 167–172). *A* common finding in these investigations is that the global motions are described well by low-frequency motions, though high-frequency motions are necessary to precisely organize catalytic sites.

4.2. Multi-basin Structure-based models

A physicist's approach to a new scientific question is to make the lowest-order approximation to the system's energetics, take that expansion to its limits of applicability and then expand to higher order. In this respect, the lowest-order approximation to a systems' dynamics is a static structure, which represents the average coordinates inside of a basin $\langle \vec{r} \rangle$. The first-order correction would be the isotropic dispersion in the coordinates $\left\langle \Delta \vec{r}_i^{\,2} \right\rangle$, which can be estimated from experimentally-accessible Debye-Waller factors (173). The dispersion may also be described in terms of anisotropic motions, where $\left\langle \Delta \vec{x}_i^{\,2} \right\rangle$, $\left\langle \Delta \vec{y}_i^{\,2} \right\rangle$ and $\left\langle \Delta \vec{z}_i^{\,2} \right\rangle$ are independent. The anisotropic dispersion characterizes the shape of the landscape about the average structure (173) and it can be obtained from normal mode

analysis or from x-ray crystallographic refinement methods (174, 175). While the correlation between anisotropic motions and functional rearrangements has allowed NMA to identify low-energy paths between basins of attraction, when atomic displacements involve the breaking and formation of interactions, the harmonic approximation becomes increasingly inaccurate. To expand the description of the landscape to include large-scale conformational rearrangements, recently-developed modeling approaches represent the potential energy landscape as having a defined number of experimentally-identified basins of attraction (176–178). While there is a range of molecular modeling approaches available to study conformational rearrangements, we will focus our attention on investigations that use multiple structure-based potentials to define a single multi-basin potential. For those interested in alternative approaches, such as the use of semi-empirical forcefields (179, 180), there are many reviews available that may be of interest (181–184).

Multi-basin structure-based models allow one to explore the robustness of functional dynamics and develop energetic descriptions of the process that are quantitatively consistent with available experimental knowledge. In contrast to the application of SBMs to protein folding, where there is a single dominant basin (i.e. the native ensemble), these approaches extend the description of the landscape to include multiple basins of attraction. When there are multiple well-characterized conformations associated with a particular biomolecule, the model must possess a minimum for each conformation (Figure 12). The general strategy when building these models is to construct a structure-based forcefield for each configuration and then systematically merge the energetic terms into a single potential energy function.

Multi-basin approaches may be partitioned into two broad categories, which we will refer to as “microscopic” and “macroscopic” mixing. To illustrate the differences, we will first consider the functional form of a single-basin structure-based forcefield (Equation 5). Each experimental conformation can be used to uniquely define a structure-based energy function V_X. A multi-basin Hamiltonian is then obtained by combining the information stored in these forcefields. Since the bonded functions (r_o, θ_o, χ_o) are based on the covalent geometry of the molecules, and the chemical composition does not change during the processes described here, we will assume that the bonded functions are identical for all conformations. For ease of discussion, we will also neglect the changes in dihedral angles (φ_o), which reduces the relevant contribution to the non-local contact energies of potential X, V_X^{contacts}.

In a microscopic mixing model, the individual energetic terms are combined term-by-term, which allows the combined potential energy function to be decomposed in the following way:

$$V_{\text{mixed}}^{\text{contacts}} = \sum_i f_i(r_i), \quad (7)$$

where $f_i(r_i)$ is a function that depends on a single degree of freedom (such as the distance between a specific atom-atom pair r_i). Let us consider an ideal protein that has two conformations and the only difference in these conformations is that atom pairs 1:20 and 30:60 (the numbers represent the sequence of the interacting atoms) are separated by different distances. In this example, the two degrees of freedom that determine the energetics of the system are the atom-atom distances $r_{1,20}$ and $r_{30,60}$. In a microscopic mixing model, the sum of the terms could be re-written as

$$V_{\text{mixed}}^{\text{contacts}} = f_{1,20}(r_{1,20}) + f_{30,60}(r_{30,60}) \quad (8)$$

In a macroscopically-mixed model, none of the energetic terms may be expressed as a function of a single degree of freedom. For our hypothetical system, the energy could only be expressed as

$$V_{\text{mixed}}^{\text{contacts}} = f(r_{1,20}, r_{30,60}). \quad (9)$$

In the macroscopic-mixing case, the energetic contribution of each degree of freedom is explicitly coupled to the energetic contribution of every other degree of freedom.

The microscopic and macroscopic mixing models represent the limiting ways by which to combine energetics from two forcefields. In principle, there is a continuum of partially-microscopically mixed models, though they are not as commonly used as other methods of mixing and they will not be discussed further. For a given functional transition, one must consider the differences between the endpoint structures when determining whether a microscopic or macroscopic mixing model is most appropriate. While there are no established rules that dictate the use of each approach, we will outline some general considerations and implications for each.

4.2.1. Microscopic Mixing Models—In microscopic mixing models, the non-bonded energetics of the system may be expressed as a sum of functions that each depend on only one atom-atom distance (Equation 7). While microscopic mixing of multiple SBMs has been implemented in a variety of ways (138, 177, 185), we will describe two classes of models that will illustrate the strengths and limitations of each. For this discussion, we will focus on Adenylate Kinase (Adk), a three-domain protein that is known to populate at least two different configurations during function: the apo and ligand-bound forms (Figure 11).

The first microscopic mixing protocol uses “mutually exclusive” contact interactions (168, 177). In a mutually exclusive potential, the native atomic interactions for each conformation are combined, such that the total interaction between atoms $i - j$ is identical to the potential in one of the single-basin potentials. The apo (ligand-free) and holo (ligand-associated) forms of Adk have 531 and 565 native contacts (177). For contacts that are common to both structures (i.e. same residue pairs at roughly the same distances), there is no need for mixing. For the unique contacts, one variant of a mutually-exclusively mixed forcefield would contain the union of the apo contacts Q_{apo} and the contacts unique to the holo form $Q_{\text{holo}}^{\text{unique}}$. One interpretation of this potential energy surface is that there is an energetic basin corresponding to the apo form, and “functional frustration” $\left(Q_{\text{holo}}^{\text{unique}}\right)$ allows the holo form to be significantly populated. Since the protein has evolved to populate both of these conformations (i.e. these conformations are basins on the landscape), this describes the system as having competing internal interactions that favor either the apo, or ligand-bound, conformation. Another interpretation of this forcefield is that Adk has evolved the apo configuration to be the global basin and ligand binding introduces additional minima between holo atom pairs. From this perspective, there is energetic competition between ligand binding and the evolutionarily-selected energetic minimum of the isolated protein. Computational work that has employed non-structure-based forcefields (171, 186, 187), single-molecule experimental investigations (188) and x-ray crystallographic studies (189) have shown that in the absence of a ligand, Adk exhibits large-scale fluctuations that span all the way to closed configuration. These findings, in addition to the framework provided

by the mixed-basin model, suggest that the apo basin is dominant and that residual stabilizing interactions of the holo atom pairs leads to partial sampling of the closed form.

In terms of the energy landscape of Adk, the above example suggests that minimally-frustrated proteins may contain residual frustration that contributes specifically to the functional dynamics. With advances in computing capabilities and algorithms (190–192), a question that should now be tractable is: Does the energetic roughness required for function influence folding mechanisms and rates, or are the folding properties robust? In other words, are function and folding governed by disparate features on the energy landscape? While we are unaware of a theoretical investigation that has explicitly probed this aspect of folding and function in Adk, ongoing work has aimed to identify regions of residual frustration, which may provide insights into the overlap of functional and folding energetics (34).

A second type of microscopic mixing has utility when contacts are formed in both conformations of the system, but the atom-atom distances of these pairs are significantly different (i.e. the differences can not be rationalized as local-basin fluctuations). For example, a contact between residue 10 and 35 may be formed in both conformations, but the atom pairs may be separated by 6 Å and 10 Å in the two conformations (Figure 12B). In these cases, it may be necessary to include a single potential energy function between the two atoms that has minima at both 6 and 10 Å (Figure 12D). This scenario is more frequently encountered in coarse-grained models (such as C_α models for proteins), where side-chain reorganization can lead to changes in the distance between C_α atoms of residues that are in contact. For example, it is possible for two Tryptophan residues (Figure 1) to be stacked, or be interacting in a extended orientation. At a coarse-grained level, the residues are considered "in-contact" in both conformations, though the distances between the C_α may change by ~10Å. If both configurations are stable, the potential of mean force between the residues must have two minima and this type of multi-basin atom-atom interaction protocol would be appropriate.

When microscopically mixing N Hamiltonians (each with M_i minima), the total number of minima in the mixed potential is not necessarily equal to $\sum_i^N M_i$. Since the functional form of a microscopically-mixed potential energy function can be rewritten as a sum of functions, each of which only depends on a single degree of freedom r_i, minima on the energy surface must only satisfy:

$$\frac{\partial}{\partial x_k} E_{\text{mixed}} = \frac{\partial}{\partial x_k} \sum_{\text{ij}} E_{\text{ij}}^{\text{contacts}}(r_{\text{ij}}) = \frac{\partial}{\partial x_k} \sum_{i} E_{\text{ik}}^{\text{contacts}}(r_{\text{ik}}) = 0 \quad (10)$$

for all x_k, where $E_{\text{ij}}^{\text{contacts}}(r_{\text{ij}})$ is the function that describes the total interaction potential between atom-pair *ij*. While, in principle, this condition may only be satisfied in the endpoint configurations, that is not necessarily the case. For Adk, microscopic mixing models have suggested at least two additional minima (177) that are not captured by a macroscopic mixing scheme (169). That is, two domains of Adk (LID and NMP) are known to interconvert between open and closed configurations (193, 194). Microscopic-mixing models suggest that two possible intermediate basins may also exist, where one domain is closed in each basin. Similar configurations have been identified through x-ray crystallographic analysis (195) and theoretical analysis indicates these partially-closed configurations may increase the fidelity (156) and rate of catalysis (171).

4.2.2. Macroscopic Mixing Models—In macroscopic mixing models the minima are unambiguously defined *a priori* and all energetic contributions of each potential are coupled

to one another. The simplest macroscopic mixing model would be defined by a simple "if"condition:

$$H_{\text{mixed}}^{\text{macro}}\left(\vec{x}\right)=\begin{cases} H_1\left(\vec{x}\right) & \text{if } H_1\left(\vec{x}\right)<H_2\left(\vec{x}\right) \\ H_2\left(\vec{x}\right) & \text{if } H_1\left(\vec{x}\right)>H_2\left(\vec{x}\right) \end{cases} \quad (11)$$

where $H_1(\vec{x})$ and $H_2(\vec{x})$ are the potential energy functions being mixed and the coordinates of the system are given by $\vec{x}$. By definition, the combined potential energy function takes on the value of the lowest energy surface at each point in configuration space. This approach may be generalized to N potential energy functions as $H(\vec{x}) = \min(H_1(\vec{x}), \ldots, H_N(\vec{x}))$. Since the molecules are on a single potential energy surface H_X at each point, the set of minima on this combined landscape is simply the union of the minima in the individual H_i.[+]

There are several considerations when using a macroscopically-mixed Hamiltonian. First, when integrating the equations of motion, the first-derivative must be defined for all possible configurations, which necessitates methods that continuously transition between the potentials. Second, if one seeks to identify the robust features of functional dynamics, the versatility of macroscopic mixing models can be of utility. For example, by using macroscopically-mixed multi-basin models that are capable of explicit modulation of the barrier heights and the depths of each basin, one may determine the extent to which the observed structural mechanisms are determined by the precise distribution of stabilizing energetic interactions. To illustrate the utility of these approaches, we will describe two commonly-implemented macroscopic-mixing algorithms: the "Boltzmann Weighting" and "Superposition" approaches.

In the Boltzmann Weighting approach (176), the mixed Hamiltonian H_{mixed} (constructed from Hamiltonians H_i) is defined according the following relationship:

$$\exp(-H_{\text{mixed}}\beta_{\text{mixing}})=\sum \exp(-(H_i - C_i)\beta_{\text{mixing}}). \quad (12)$$

$\beta_{mixing} = 1/(k_B T_{mixing})$ and C_i is an offset that increases/decreases the relative weight of H_i in the mixed-basin forcefield. Here, the barrier height and depth of each basin are governed by the parameters T_{mixing} and C_i. For a particular configuration, if $(H_1 - C_1) - (H_2 - C_2) = 10 k_B T_{mixing}$, then H_2 would be weighted by a factor of e^{10} relative to H_1. For example, in Adk, a particular experimental measurement may suggest that the open conformation is more stable by several $k_B T$, a condition that may be imposed by adjusting the values of C_i(169). T_{mixing} is the "mixing temperature," which is unrelated to the simulated temperature. T_{mixing} modulates the height of the barrier separating H_1 and H_2 by defining how sharp the transition is between potentials. For large values of T_{mixing}, the relative weight of each potential changes less with the difference between the potentials, and lower barriers are to be expected. Figure 12 shows an example of how two functions may be combined with the Boltzmann Weighting approach. For extensive discussion on specific applications of Boltzmann-weighted mixed potentials, we suggest additional reading (169, 176, 197).

[+]This assumes that for each set of coordinates that defines a minimum $\vec{x}_i^{\min}$, the following condition is satisfied: $\min(H_1(\vec{x}_i^{\min}), \ldots, H_N(\vec{x}_i^{\min})) = H_i(\vec{x}_i^{\min})$. In practice, this is normally imposed by shifting the energy of the input potentials appropriately. (176, 196)

In the Superposition approach (178, 196), a smooth potential, H^S_{mixed}, is defined as the eigenvalue of the characteristic equation:

$$\begin{pmatrix} H_1 & \Delta \\ \Delta & H_2+C_2 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}=H^S_{\text{mixed}}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (13)$$

where Δ is the coupling constant between H_1 and H_2, and ($c1$, $c2$) is the eigenvector. Δ governs the barrier height (similar to T_{mixing} in Boltzmann Mixing), and C_2 controls the relative weights of the input potentials. Solving for the secular equation yields the lowest-energy solution:

$$H^S_{\text{mixed}}=\frac{H_1+H_2+C_2}{2}-\sqrt{\left(\frac{H_1-H_2-C_2}{2}\right)^2+\Delta^2} \quad (14)$$

Similar to the Boltzmann Weighting approach, the Superposition approach ensures a smooth transition (i.e. defined first derivative) between potential energy surfaces and it allows the relative depths of each basin to be directly modulated.

4.2.3. Micro- vs. Macro-scopic mixing—Consideration must be given when deciding which mixing method best describes a particular transition. In the macroscopic mixing models, explicit declarations allow for straightforward calibration of some thermodynamic properties. Though, in these models, the “all-or-nothing” contribution of each potential is a restrictive constraint on the dynamics and may appear non-physical. In contrast, the flexibility of microscopic mixing approaches may necessitate a variety of qualitative decisions to be made, as well.

Each approach provides unique insights into the physics of the system. When we used a microscopic mixing model for Adenylate Kinase (177), we first asked whether or not a mixed set of contacts would provide sufficient information to encode the desired endpoints as energetic minima, while also exhibiting appropriate structural fluctuations about the endpoints. Since we were able to satisfy these conditions, the microscopic model provided an energetic framework to interpret the origins of Adk's functional transitions. However, with this approach, the free-energy barrier (potential of mean force as a function of inter-domain distance) was on the order of 1-2 k_BT, which may be a low estimate (described below). In contrast, in a macroscopically-mixed potential for Adk, it is possible to artificially increase the barrier height between the conformations. While the macroscopically-mixed potential has not yielded essential intermediate basins in earlier implementations (due to limitations on the available phase space in macroscopic mixing models), these approaches may be extended beyond the two-basin representation. Combing these approaches, one may use microscopic models to identify stable configurations that result from mixing and then use macroscopic models (with more than two minima) to tune the barrier heights further. Together, this layered approach may establish an overall picture of the energy landscape that is consistent with experimental knowledge that also provides an energetic/entropic characterization of the dynamics.

4.3. Relationship between free-energy barriers and kinetics

It is worth taking a moment to discuss the relationship between free-energy barriers and rates in biological molecules, since this is often where theoretical predictions and experimental kinetic measurements intersect. The barrier height of a conformational change may be related to the mean first passage time $\langle \tau \rangle$ via the relationship (20, 198):

$$\langle \tau \rangle = \int_{Q_{\text{initial}}}^{Q_{\text{final}}} dQ \int_{0}^{Q} dQ' \frac{\exp\left[(G(Q) - G(Q')) / k_B T \right]}{D(Q)}, \quad (15)$$

where $G(Q)$ is the free-energy as a function of the reaction coordinate Q, $D(Q)$ is the diffusion in Q-space, and Q_{final} and $Q_{initial}$ are the values of the coordinate for the endpoint configurations. When the diffusion coefficient in Q-space is not available, and the free-energy profile is dominated by a single barrier of height ΔG, then it is useful to approximate this relationship as (199)

$$\langle \tau \rangle = \frac{1}{C} \exp(\Delta G / k_B T). \quad (16)$$

C may be interpreted as the "attempt frequency" of barrier crossing. While we are not aware of an estimate of C for conformational transitions in Adk, estimates are available for protein folding (200, 201), RNA folding (18) and rearrangements in the ribosome (202). These independent theoretical, experimental, theoretical and computational studies suggest C is ~ $1 - 10 \mu s^{-1}$ for many folding and functional transitions. Since conformational transitions in Adk occur on a smaller length-scale than tRNA rearrangements inside of the ribosome, it would be reasonable to expect C to be larger for Adk's transitions. The corresponding $\langle \tau \rangle$ values would then represent an upper estimate (for a given ΔG). For example, free-energy barriers of 1-2 $k_B T$(177) and $C = 1 \mu s^{-1}$ correspond to $\langle \tau \rangle \sim 10 \mu s$. Kinetic measurements indicate that these conformational transitions are rate limiting and that they occur on timescales of hundreds of microseconds (144), which is slightly slower than the rates predicted by the microscopic mixing model. However, an order of magnitude in timescale only corresponds to a 2.8 $k_B T$ difference in barrier height and the observed deviation may be considered minor.

4.4. Computational considerations

Simulations allow one to enumerate the configuration space of a molecule, from which thermodynamic and kinetic properties may be calculated for a given model. With infinite computing time, one could exploit the ergodicity of these systems and simply wait for the full phase space to be sampled with a Boltzmann distribution. However, the energy landscapes of biomolecules are complex, and often possess many competing energetic basins. Therefore, when using conventional molecular dynamics (or, Monte Carlo) simulations, the majority of the calculation will be spent sampling the low-energy basins, and barrier-crossing events will be rare. Since the transition states between stable basins limit the overall functional dynamics, significant efforts have been dedicated to overcoming the sampling issues associated with high barriers. Some general strategies for characterizing transition states (or, transition state ensembles) are to enhance the overall sampling of the simulation (203–206), increase the sampling along particular coordinates (207), or to specifically search for the transition states (208–213).

In enhanced sampling approaches, system perturbations are introduced increase the sampling in particular regions of phase space and then statistical mechanical relationships are used to recover the physical quantities of the unperturbed system. For example, when using replica exchange methods, multiple copies of the system are simultaneously simulated, and each image is simulated with a unique Hamiltonian (H_i), or at a unique temperature (T_i). During the simulations, replicas are probabilistically interconverted (i.e. the copy at T_i goes to temperature T_{i+1} and the replica at T_{i+1} goes to T_i). The system at low temperature may then "jump" to the higher-energy configurations that are sampled at higher temperatures. Since these jumps are non-physical, the kinetics of the system are not directly observable.

However, if the jumps between temperatures satisfy detailed balance, a canonical ensemble will be sampled. While this method has been very successful for biomolecular folding (214–220), where the relevant transition is temperature induced, there are limitations to its applicability. For example, when studying the interconversion between competing energetic basins, performing simulations at high temperature can lead to global unfolding of the system. More generally, temperature changes can be associated with phase transitions that are not relevant to the process of interest (e.g. global folding, or evaporation of explicit water). When applied to appropriate rearrangements, replica exchange methods can significantly increase the statistics of a given process.

An alternative to calculating the full range of possible configurations is to use path sampling approaches (208–213). Rather than allow the system to freely search all phase space, these algorithms search for the lowest-energy pathway that connects the predefined endpoints. To search for these low-energy paths, the system is described by a set of images, where adjacent images are coupled though energetic restraints, or constraints. As the calculation proceeds, the set of images evolves with time, such that the set of images converge onto the lowest energy (or, lowest free-energy) pathway. These methods are highly effective for describing transitions that possess a dominant low-energy pathway. When there are large changes in entropy as a function of the progress coordinate, then sampling the high-entropy regions has the potential to be computationally challenging. In nudged elastic band methods (210–213, 221), the replicas are coupled to one another, which restricts the relative mobility of adjacent images. If the given rearrangement has a high-entropy substate, then the harmonic coupling between images can limit the characterization of that space. To overcome this issue, these methods often use harmonic approximations to the local density of states, in order to estimate the local contribution to the entropy. For the reader interested in the detailed use of these approaches, we suggest additional reading (222). If there are multiple competing paths, then these estimations of entropy will not be accurate, since the harmonic approximation to the density of states is no longer valid. In string methods (209–212, 223), rather than couple each image to the others, the images are only required to be equispaced along a predefined set of reaction coordinates. As the calculation proceeds, the images are evolved until the lowest-energy path is found (in the defined coordinate space). Since each image is allowed to fluctuate in directions perpendicular to the reaction coordinates, the local phase space may be fully sampled. Thus, string methods can provide a more complete characterization of the minimum free-energy pathways than elastic band methods.

5. Roles of Order-Disorder Transitions in Biomolecular Functional Dynamics

Largely inspired by the seminal works of Haldane and Pauling (224, 225), a conventional view of biomolecular functional dynamics is that there are rigid regions of residues and rearrangements are facilitated by highly flexible “hinge” regions (27, 149, 165, 226, 227). This description is often enticing due to its similarities with macroscopic machines. Though, in contrast to macroscopic machines, where the rigidity of components spans orders of magnitude (i.e. hinges are often lubricated, leading to “flat” potential energy functions, and moving components are composed of solid-phase metals), the weak interactions that maintain biomolecular stability are the same interactions that facilitate functional dynamics. Since the non-bonded energetic interactions that maintain biomolecular structure are normally on the scale of a few k_BT, there is no reason to assume that a hinge/rod description will account for the full range of biomolecular functional dynamics.

Here, we discuss a variety of recent findings that expand our picture of biomolecular functional dynamics beyond the simple hinge/rod picture. Of particular interest is the biological relevance of order-disorder transitions. While energetic interactions are essential

for dynamics, biomolecules may also exploit large changes in entropy through local, or global, folding/unfolding events. In doing so, biomolecules may tune their functional dynamics by evolving differential stabilities of particular subunits. To illustrate the biological role of heterogeneous biomolecular stability and flexibility, we will discuss several classes of order-disorder transitions that have been implicated in the functional dynamics of polypeptide and nucleic acid chains.

5.1. Conformational Rearrangements: Cracking

Early studies that employed iterative normal mode analysis to characterize the conformational transitions of Adenylate Kinase (133, 156, 165, 228) suggest that biomolecules may partially unfold and refold during conformational transitions. That is, it was predicted that internal "strain" energy accumulates during conformational rearrangements, and this strain is heterogeneously distributed across the protein. Not only was strain isolated to particular residues, but the level of strain energy would be sufficient to denature the protein (> 20 kCal/Mol). This high strain would also result in free-energy barriers would lead to biologically-irrelevant timescales (i.e. transitions would occur on hours, not milliseconds). To circumvent these high-strain states, it was proposed that these residues may "crack," or locally unfold and refold during conformational transitions (133). By cracking, residues would release stabilizing interactions, enter a higher-entropy unfolded state and then refold in the alternate configuration. The increase in entropy would reduce the free-energy barrier of rearrangement, and would increase the kinetic rates (Figure 13) (134). At first glance, the cracking mechanism may appear too complex to increase the rates of interconversion. Though, as discussed above, the kinetic prefactor for protein folding is ~ $1\mu s^{-1}$. Since cracking is a smaller-scale process than folding of an entire protein, one should expect the timescale associated with the cracking transition to be less than 1 μs (assuming that interconversion between order-disorder does not introduce an additional significant free-energy barrier). The ability to transition between order and disorder on such timescales makes it clear that cracking should be a kinetically-accessible mechanism by which conformational rearrangements may occur.

When the cracking mechanism is employed, the balance between order and disorder for individual residues will determine the functional dynamics. Therefore, protein functional dynamics may be tuned by selecting for amino-acid sequences that have appropriate local stabilities. Analogous to the SK Ising model, each residue is uniquely coupled to the rest of the protein, which allows residues to have differential propensities to spontaneously unfold, even under native conditions (229, 230).

Following the original prediction of cracking, many studies have aimed to use more complex modeling approaches to characterize the location and scale of cracking in Adk. With a microscopically-mixed multi-basin Hamiltonian, we found that the degree of cracking in Adk is largest in the highest-strain regions that were predicted by Miyashita et al. (177). In contrast, simulations of Adk that employed macroscopic mixing models have suggested a lower-degree of cracking (169). As discussed above, this attenuated level of cracking may have been the result of the stringent restrictions imposed by macroscopic mixing models (See *Macroscopic Mixing Models*). These studies have employed coarse-grained models that represent each residue as a single bead, located at the position of the C_α atom. Therefore, in these studies, cracking was measured in terms of large-scale changes in backbone dynamics. Models that explicitly include all-atoms have demonstrated that protein folding transition may be described as a superposition of partially-separable transitions: backbone collapse and side-chain rearrangements (56, 129, 231). This separation of backbone and side-chain ordering suggests that the details of cracking may be more subtle than originally envisioned. Specifically, cracking may also involve the local disordering of side chains, with a more minor contribution from backbone cracking (186). Together, these

studies suggest that multi-basin models that employ atomic resolution will reveal a broad range of cracking events that span from individual residue side chains, to regions of consecutive residues.

Cracking has also been identified in the motor protein kinesin. Kinesin molecules are able to "walk" on microtubules, which enables transport within the cell (26). During each "step," kinesin binds ATP, undergoes relative displacement of ≈ 80 Å of the domains (145), hydrolyzes ATP and releases ADP. Simulation studies that employed time-dependent Hamiltonians showed that the power stroke in the kinesin head domain may be facilitated by cracking (232, 233). Hyeon et al. showed that a simple model for substrate directed walking is that ATP binding shifts the effective potential energy surface of the leading domain, such that it prefers to place the trailing domain "in front." This leads to an internal conflict of domain energetics, which leads to partial disordering of the linker region that connects them. If flexibility of the linker were too high, then strain could not be transferred between the domains. Conversely, if the linker were rigid, it would not be capable of undergoing the conformational rearrangements associated with walking.

The potential importance of cracking has motivated the investigation of its role in other proteins' conformational dynamics, as well. Itoh and Sasai developed a dual-basin statistical mechanical model and found evidence of cracking in the C-terminal tail of Calmodulin (135). Similarly, Tripathi and Portman developed a variational model to describe conformational rearrangements in the N-terminal domain of Calmodulin and found that specific regions undergo significant changes in flexibility during rearrangements (234), which is consistent with the cracking paradigm. Using a coarse-grained model, Hyeon et al. also provided evidence of cracking during functional rearrangements in Protein Kinase A (235). Experimentally, biophysical measurements have implicated cracking during the functional dynamics of Adenylate Kinase (229, 236). Peng et al. used instantaneous normal mode (INM) calculations to guide conformational rearrangements in a simulation of Adk (237) and found that immediately preceding domain rearrangements, the INMs were dissimilar to those found in either of the end-state ensembles. This suggests that simple low-energy, rigid-body rearrangements are insufficient to account for Adk's functional dynamics and that cracking may occur. Finally, Okazaki et al. surveyed multiple proteins and showed that cracking appears to be a common feature in multiple protein rearrangements (196).

5.2. Biomolecular Association: Fly-Casting

The balance of many, often competing, kinetic processes is required for proper functional dynamics of the cell (238, 239). The highly non-equilibrium nature of cellular dynamics requires that signaling interactions be kinetically accessible and regulated. If bimolecular association events are simply 3-dimensional random walks, the only way to modulate the kinetics of association would be to evolve proteins of different shapes and sizes, thereby changing their diffusion coefficients. However, evolutionary pressure has led to nearly every aspect of the cell being tunable, including association kinetics.

To address the apparently limited tunability of bimolecular association, Shoemaker et al. proposed the "fly-casting" mechanism (240), where proteins partially unfold when isolated and the full structure forms upon substrate binding. Using a free-energy functional approach, they demonstrated that if part of the protein can unfold, and the unfolded form of the protein has a marginal affinity for the substrate, the exponential increase in phase space available for substrate capture can lead to increased rates of association. It was later demonstrated through molecular simulations (241, 242) that these changes in flexibility can impact the rates of bimolecular association events. Bioinformatic analysis of protein-RNA interactions have also suggested that the electrostatic composition of biomolecules have evolved to guide disordered regions during fly-casting events (243).

In addition to using fly-casting dynamics for bimolecular association, recent work has elaborated on the role of intrinsically-disordered regions during DNA "searching." That is, the molecular flexibility that enables fly-casting may also facilitate "sliding" motions (95), where the protein "scans" the DNA when looking for the lowest-energy binding site (244). During these searches, proteins may utilize a combination of sliding and "hopping" between sites. During hopping, the protein dissociates from the DNA and reattaches itself via the fly-casting mechanism. The balance between flexibility, disorder and atomic energetics determines the extent to which each mechanism is utilized, which ultimately determines if the genetic messages may be expressed at the appropriate time.

Similar to cracking, the fly-casting mechanism provides a means by which a balance of local stabilities dictates the functional dynamics of a protein. However, in contrast to cracking, where the endpoints are stable configurations, in fly casting the folded configuration only needs to be stable upon substrate binding.

5.3. Protein Recognition: Global Unfolding and Refolding

While fly-casting and cracking involve partial unfolding of substructures (Figure 3B), in some cases, global order-disorder transitions may be linked to functional dynamics. The Repressor of Primer (Rop) dimer* is formed by association of two alpha-helical protein chains that bind to "kissing" RNA hairpins and stabilize their interactions. Based on experimental kinetic data (245), it had been suggested that Rop and its mutants can adopt two symmetrically-related competing conformations during function (246). To study the dynamics of this process, a multi-basin potential energy function was employed to simulate the transitions between these competing configurations (247). Since a microscopically-mixed Hamiltonian was employed, it would have been possible for the protein to undergo this rearrangement through a sequence of local displacements, while the monomers remain in association with each other. Despite the energetic flexibility afforded by the microscopic model, these simulations indicate that the Rop dimer globally unfolds and then refolds during its conformational rearrangements. These predictions from simulations were later verified in single-molecule experiments (248). Consistent with cracking and fly-casting dynamics, this global denaturation is afforded by the marginal stability of proteins under physiological conditions. Global, or local increases in stability could make unfolding kinetically unfavorable, and functional regulation in Rop would then not be possible.

5.4. Riboswitch Signaling: Substrate-Assisted Ordering

Riboswitch aptamer domains are untranslated regions of mRNA that, in response to ligand binding, adopt well-defined three-dimensional configurations, whereas alternate ensembles of configurations are accessed in the absence of ligand. These conformational changes can even include large-scale unfolding and refolding of secondary structure, similar to the operation of the Rop dimer. The balance between these ligand-bound and ligand-free ensembles directly regulates transcription and translation (249, 250) (Figure 14A). Since the identification of riboswitches in bacteria (251), there has been vigorous investigation into the relationship between ligand binding, structure formation and gene expression. A frequently adopted framework for interpreting experimental data is that all secondary structure (i.e. helices), and some tertiary structure, is formed prior to substrate binding. For systems where this assumption holds, the role of ligand binding is relegated to facilitating minor rearrangements. That form of operating is reminiscent of allosteric protein function, where tertiary structure is usually fully formed and conformational transitions associated with ligand binding are limited to the repositioning of domains. However, bacterial riboswitches fold with the ligand in solution, suggesting that ligand binding may alter the

*Rop is also sometimes called Rom (RNA 1 Modulator).

dynamics of the global reorganization/folding processes, in addition to coordinating final atomic repositioning.

Simulations with unfrustrated models have shown that entropic bottlenecks during folding can be directly perturbed by ligand association. In the case of the SAM-1 riboswitch, entropic contributions disfavor the formation of the P1 helix (Figure 14B), which is the step associated with the largest free-energy barrier during folding (125). This large difference in configurational entropy can be attributed to the connectivity of the helices. The strands composing P2, P3 and P4 helices are each connected by a few-residue "turn," and the P1 helix is formed by the terminal residues in the domain (Figure 14A). While the P2/3/4 turns are composed of 4 residues, the P1 strands are connected by a linker of ~80 residues, hence the designation of "non-local helix." Thus, the configurational entropy of the unfolded P1 helix is much larger than the other helices. By simulating the process in the presence of the ligand, it was shown that this rate-limiting step is accelerated by ligand binding, suggesting that ligand binding can have a thermodynamic and kinetic influence on aptamer formation. Consistent with this prediction, recent chemical probing data has shown that in the absence of ligand, the P1 helix exhibits high mobility, and this mobility is attenuated by substrate binding (252).[#] Single-molecule experiments indicate that ligand binding most strongly affects the dynamics of the P1 helix (253), which highlights the potential of SAM to direct P1 folding.

Coarse-grained simulations of a purine riboswitch also found that the non-local helix is least stable and its formation is rate limiting (254). Similarly, simulations of the $preQ_1$ riboswitch showed that the folding mechanisms can be redirected by substrate binding (255). Single-molecule experiments have also revealed that ligand binding has a focused effect on the non-local helix in a purine riboswitch (256). Each of these studies demonstrate the role that ligand binding has on folding, suggesting that substrate recognition and folding are concomitant.

Similar to protein recognition, substrate-assisted folding in RNA may encompass fly-casting-like dynamics. As discussed above, the principal effect of the fly-casting mechanism is to accelerate substrate binding rates by increasing the capture radius for the substrate. Riboswitch studies suggest that there may be a symbiotic relationship between folding and ligand association: fly-casting increases association rates and ligand association increases the biomolecular folding rates. Similar to fly-casting proteins, if aptamer domain formation is too fast, or stable, then there would be a separation of timescales between ligand binding and riboswitch folding. In that scenario, ligand binding may increase the folding rate slightly, but the presence/absence of ligand would only represent a minor perturbation. Together, these studies corroborate a general theme in molecular biophysics: biomolecular complexes have evolved to be marginally stable, allowing them to access higher-entropy states and increase the available phase space for regulation.

5.5. Molecular Machines: Entropically-Guided Rearrangements

The final example of the role of entropy in biomolecular functional dynamics that we will discuss is found in the elongation cycle of the ribosome. During translation, mRNA is read by the ribosome through the use of transfer RNA (tRNA) molecules, which produces proteins with specific amino acid sequences. The decoding process involves individual aminoacyl-tRNA (aa-tRNA) molecules that form 3-base-pair interactions (i.e. codon-anticodon interactions) with the mRNA, where each base pair is composed of two (Adenine-

[#]It should be noted that this study was for an isolated aptamer, which lacked the expression platform. However, the dynamic effect of the ligand is anticipated to be similar in the context of the full riboswitch.

Uracil) or three (Guanine-Cytosine) hydrogen bonds. When accounting for competition with water, the hydrogen bonds of base pairs provide ~ $1 - 2k_BT$ of stabilizing energy, each. Since four types of nucleic acids compose these base-pair interactions, there are $4^3 = 64$ unique messages that may be encoded by a single mRNA codon. With only 64 possible ways to encode 20 different naturally-occurring amino acids, where all possible base-pair interactions are on comparable energetic scales, thermodynamic differences between cognate (correct tRNA-mRNA pairs) and nearcognate (1 base-pair mismatch, such as a U on the tRNA and a G on the mRNA) can only account for an error rate of about 1 in 100. Since the error rate of the ribosome is approximately 1 in $10^3 - 10^4$, additional proofreading mechanisms must be employed (257).

To explain the fidelity of ribosome dynamics, Hopfield proposed the “kinetic proofreading” mechanism (258). In this mechanism, by loading a marginally-stable intermediate via a driven process (such as GTP hydrolysis), the differences in kinetic accessibility of the initial and final configurations may be greater than the corresponding differences in thermodynamic stability. Kinetic proofreading was proposed as a general strategy by which molecular processes may increase the fidelity of a particular substrate selection step, though here we will only discuss it in context of ribosome fidelity (259). In the ribosome, elongation factor-Tu (EF-Tu) uses the energy stored in GTP to load the ribosome with aa-tRNA molecules (260–262) in an energetically-stained A/T configuration (163, 263). Upon EF-Tu release, cognate (i.e. correct) aa-tRNAs enter the ribosome (≈ 90 Å displacement) and adopt the “A/A” configuration, while noncognate aa-tRNAs are rejected. In the kinetic proofreading framework, the A/T configuration is loaded by EF-Tu, but both the A/A and dissociated configurations are more stable. The A/T-to-A/A transition (called “accommodation”) can then serve as the proofreading step in translation. Consistent with the kinetic proofreading mechanism, theoretical considerations have demonstrated that kinetics and fidelity impose competing limitations on ribosome dynamics (264). If the accommodation step were too fast (averaged over all tRNA species), large energetic factors would be necessary to remove the tRNA from the ribosome. However, no such factors are required for rejection. In contrast, if accommodation were too slow, then the dissociation of tRNA from the ribosome would be overly likely, even for cognate tRNAs.

Investigations into the balance of kinetics, energetics and fidelity of accommodation (265) have revealed atomic details of the endpoints (146, 261, 266), the transition state ensembles (260, 267, 268) and rates (265) of accommodation. These studies have provided many insights into the details of accommodation, including the role of disorder in the incoming aa-tRNA molecule. Using a modified SBM, we showed that the flexibility of the single-stranded 3′-CCA end (Figure 15) may lead to large changes in the configurational entropy during accommodation (Figure 15), which also facilitates a multi-step accommodation process (268). Upon dissociation of EF-Tu, the 3′-CCA end can adopt a disperse set of configurations within the A/T ensemble, which leads to an increase in configurational entropy. As the aa-tRNA enters the ribosome, the initial accommodation sub-step is aa-tRNA elbow movement towards an A/A-like configuration. This is followed by aa-tRNA arm movement and finally 3′-CCA end movement into the catalytic center of the ribosome (the peptidyltransferase center, PTC). During the elbow-association step, the accessible configurations of the 3′-CCA end appear to be constant, which indicates a negligible change in configurational entropy of the 3′-CCA end during this transition. As the tRNA arm moves into the A site, movement of the 3′-CCA end is partially restricted (Figure 15), suggesting that a decrease in configurational entropy resists such motions. Finally, as the 3′-CCA end enters the PTC, it becomes highly ordered, as required for the chemical step (peptide bond formation) to occur (3). This further decrease in configurational entropy represents an additional entropic penalty during accommodation.

This order-disorder-order transition in the 3′-CCA end of the aa-tRNA can have profound biological implications. Specifically, changes in configurational entropy of the universally conserved (across all tRNA species) 3′-CCA end may provide an entropic counterweight that ensures 3′-CCA entry into the PTC is the last step during accommodation. Since the incoming amino acid is attached to the 3′-CCA end, once it accommodates the amino acid may be added to the growing peptide chain. By favoring the late entry of the 3′-CCA end, this entropic contribution can facilitate the use of additionally proofreading "machinery" prior to incorporating a potential mistake.

While previous simulations showed that flexibility in the 3′-CCA end is necessary for it to enter the PTC (267), this was the first time that configurational entropy of the 3′-CCA end was explicitly identified as contributing to the global tRNA dynamics. There are many signatures of a dynamic 3′-CCA end in tRNA may be found in the literature. For example, the 3′-CCA end of tRNA molecules are often not resolved in crystallographic structures (269), or the dispersion in their coordinates is large (as measured by the Debye-Waller factors) (270). In cryogenic electron microscopy studies there is often a low density obtained for the 3′-CCA end (271), indicating that the 3′-CCA end structure is composed of a diverse set of configurations, relative to other molecular components. The 3′-CCA end also interacts with many distinct binding partners (272), including synthetases (273), EF-Tu (274) and the ribosome (275), which necessitates that its conformation be easily adaptable.

These studies highlight the balance between stabilizing energetic interactions (that favor accommodation) and destabilizing entropic contributions (that resist the process). If flexibility of the 3′-CCA end were increased, it would be possible for accommodation to become too slow, and correct tRNA molecules would be rejected. Though, if 3′-CCA flexibility were decreased, it may not be possible for the tRNA molecules to navigate through the narrow corridor of the ribosome (267, 276).

6. Closing Remarks

Grounded in statistical mechanics, the energy landscape framework has enabled the quantitative characterization of biomolecular dynamics at the atomic level. Here, we have described the historical context surrounding the development of this framework, which originated in protein folding. Since then, it has been repeatedly extended to address biomolecular functional dynamics of increasing complexity. An exciting consequence of the energy landscape description is that there is an interplay between order-disorder transitions and the functional dynamics of biomolecules. From a wide range of studies, there are many examples demonstrating how the marginal stability of biomolecules allows them to adjust their unfolding properties as a mode for tuning the functional dynamics. Accordingly, these deceptively-complex molecular assemblies possess an extremely large range of phase space available for functional dynamics. In the future, we anticipate that this theme will be echoed in assembly/disassembly processes, transient formation of excited states, and large-scale collective motions in the cell. Eventually, the molecular description will reach scales that overlap with systemslevel biology, which will represent one of the next major challenges for the biological physics community.

Acknowledgments

PCW would like to thank Professor Alexander Schug (Karlsruher Institut für Technologie) and Dr. Jeffrey Noel (UCSD) for critical feedback during manuscript preparation. This work was supported by the Center for Theoretical Biological Physics sponsored by the NSF(Grant PHY-0822283) and by NSF- MCB-1214457, in addition to support from NIH Grant R01-GM072686. JNO is a CPRIT Scholar in Cancer Research sponsored by the Cancer Prevention and Research Institute of Texas.

References

1. Cech TR. Self-splicing of group-I introns. Annu Rev Biochem. Jan.1990 59:543–568. [PubMed: 2197983]
2. Mcclain WH, Guerriertakada C, Altman S. Model substrates for an RNA enzyme. Science. Jan; 1987 238(4826):527–530. [PubMed: 2443980]
3. Yonath A. Large facilities and the evolving ribosome, the cellular machine for genetic-code translation. J Royal Soc Inter. Aug.2009 6:S575–S585.
4. Frank J, Gao H, Sengupta J, Gao N, Taylor DJ. The process of mRNA-tRNA translocation. Proc Nat Acad Sci USA. Dec; 2007 104(50):19671–8. [PubMed: 18003906]
5. Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH, Noller HF. Crystal structure of the ribosome at 5.5 a resolution. Science. May; 2001 92(5518):883–96. [PubMed: 11283358]
6. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009; 136(4):701–718. [PubMed: 19239890]
7. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet. Jan; 2009 10(2):94–108. [PubMed: 19148191]
8. Bartel DP. MicroRNAs: Target recognition and regulatory functions. Cell. Jan; 2009 136(2):215–233. [PubMed: 19167326]
9. Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem. Jan.2010 79:351–379. [PubMed: 20533884]
10. Fuxreiter M, Simon I, Bondos S. Dynamic protein-DNA recognition: beyond what can be seen. Trnd Biochem Sci. Jan; 2011 36(8):415–423.
11. Vallabhapurapu S, Karin M. Regulation and function of NF-κB transcription factors in the immune system. Annu Rev Immunol. Apr; 2009 27(1):693–733. [PubMed: 19302050]
12. Nieto MA. The snail superfamily of zinc-finger transcription factors. Nat Rev Mol Cell Biol. Jan; 2002 3(3):155–166. [PubMed: 11994736]
13. Baldwin AS. The NF-κB and I κB proteins: New discoveries and insights. Annu Rev Immunol. Jan.1996 14:649–683. [PubMed: 8717528]
14. Vignali M, Hassan AH, Neely KE, Workman JL. ATP-dependent chromatin-remodeling complexes. Mol and Cell Biol. Jan; 2000 20(6):1899–1910. [PubMed: 10688638]
15. Olave IA, Reck-Peterson SL, Crabtree GR. Nuclear actin and actin-related proteins in chromatin remodeling. Annu Rev Biochem. Jan.2002 71:755–781. [PubMed: 12045110]
16. Levinthal C. Mossbauer Spectroscopy in Biological Systems. 1969
17. Dill KA. Theory for the folding and stability of globular proteins. Biochemistry. 1985; 4:1501–1509. [PubMed: 3986190]
18. Thirumalai D, Hyeon C. RNA and protein folding: common themes and variations. Biochemistry. Apr; 2005 44(13):4957–70. [PubMed: 15794634]
19. Leopold PE, Montal M, Onuchic JN. Protein folding funnels - a kinetic approach to the sequence structure relationship. Proc Nat Acad Sci USA. Jan; 1992 89(18):8721–8725. [PubMed: 1528885]
20. Bryngelson D, Wolynes PG. Intermediates and barrier crossing in a random energy-model (with applications to protein folding). J Phys Chem-Us. Jan; 1989 93(19):6902–6915.
21. Bryngelson JD, Wolynes PG. A simple statistical field-theory of heteropolymer collapse with application to protein folding. Biopolymers. Jan; 1990 30(1-2):177–188.
22. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein-folding - a synthesis. Proteins. Jan; 1995 21(3):167–195. [PubMed: 7784423]
23. Svoboda K, Schmidt CF, Schnapp BJ, Block SM. Direct observation of kinesin stepping by optical trapping interferometry. Nature. Jan; 1993 365(6448):721–727. [PubMed: 8413650]
24. Bustamante C, Chemla YR, Forde NR, Izhaky D. Mechanical processes in biochemistry. Annu Rev Biochem. Jan.2004 73:705–748. [PubMed: 15189157]
25. Tan Y, Hanson J, Yang H. Direct Mg2+ binding activates adenylate kinase from escherichia coli. J Biol Chem. Nov.2008

26. Vale RD, Milligan RA. The way things move: Looking under the hood of molecular motor proteins. Science. Jan; 2000 288(5463):88–95. [PubMed: 10753125]
27. Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, Kern D. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. Nature. Dec; 2007 450(7171):913–6. [PubMed: 18026087]
28. Frauenfelder H, Sligar SG, Wolynes PG. The energy landscapes and motions of proteins. Science. 1991; 254(5038):1598–1603. [PubMed: 1749933]
29. Socci ND, Onuchic J. Kinetic and thermodynamic analysis of protein-like heteropolymers - monte-carlo histogram technique. J Chem Phys. Jan; 1995 103(11):4732–4744.
30. Onuchic JN, Luthey-Schulten Z, Wolynes PG. Theory of protein folding: The energy landscape perspective. Annu Rev Phys Chem. Jan.1997 48:545–600. [PubMed: 9348663]
31. Shea JE, Onuchic JN, Brooks CL. Exploring the origins of topological frustration: Design of a minimally frustrated model of fragment B of protein A. Proc Nat Acad Sci USA. Jan; 1999 96(22): 12512–12517. [PubMed: 10535953]
32. Shea JE, Onuchic JN, Brooks CL. Energetic frustration and the nature of the transition state in protein folding. J Chem Phys. Jan; 2000 113(17):7663–7671.
33. Cho SS, Levy Y, Onuchic JN, Wolynes PG. Overcoming residual frustration in domain-swapping: the roles of disulfide bonds in dimerization and aggregation. Phys Biol. Jan; 2005 2(2):S44–S55. [PubMed: 16204848]
34. Ferreiro DU, Hegler JA, Komives EA, Wolynes PG. Localizing frustration in native proteins and protein assemblies. Proc Nat Acad Sci USA. Dec; 2007 104(50):19819–24. [PubMed: 18077414]
35. Sherrington D, Kirkpatrick S. Solvable model of a spin-glass. Phys Rev Lett. 1975; 35:1792–1796.
36. Edwards SF, Anderson PW. Theory of spin glasses. J Phys F- Met Phys. 1975; 5:9655–974.
37. Binder K, Young AP. Spin-glasses - experimental facts, theoretical concepts, and open questions. Rev Mod Phys. 1986; 58:801–976.
38. Onsager L. Crystal statistics. I. A two-dimensional model with an order-disorder transition. Physical Review. 1944; 65(3-4):117.
39. Ising E. Zeits f Physik. 1925; 31:253.
40. Glotzer SC. Spatially heterogeneous dynamics in liquids: insights from simulation. Journal Of Non-Crystalline Solids. Sep; 2000 274(1-3):342–355.
41. Sastry S. The relationship between fragility, configurational entropy and the potential energy landscape of glass-forming liquids. Nature. Jan 11; 2001 409(6817):164–167. [PubMed: 11196634]
42. Angell CA. Formation of glasses from liquids and biopolymers. Science. 1995; 267(5206):1924. [PubMed: 17770101]
43. Stevenson JD, Schmalian J, Wolynes PG. The shapes of cooperatively rearranging regions in glass-forming liquids. Nature Physics. Apr; 2006 2(4):268–274.
44. Xia XY, Wolynes PG. Microscopic theory of heterogeneity and nonexponential relaxations in supercooled liquids. Phys Rev Lett. Jun 11; 2001 86(24):5526–5529. [PubMed: 11415292]
45. Kirkpatrick TR, Thirumalai D, Wolynes PG. Scaling concepts for the dynamics of viscous-liquids near an ideal glassy state. Physical Review A. Jul 15; 1989 40(2):1045–1054. [PubMed: 9902230]
46. Garrahan JP, Chandler D. Geometrical explanation and scaling of dynamical heterogeneities in glass forming systems. Phys Rev Lett. Jul 15.2002 89(3)
47. Ediger MD, Angell CA, Nagel SR. Supercooled liquids and glasses. J Phys Chem. 1996; 100(31): 13200–13212.
48. Inoue A. Stabilization of metallic supercooled liquid and bulk amorphous alloys. Acta Mater. Jan; 2000 48(1):279–306.
49. Frauenfelder H. Proteins, supercooled liquids, and glasses: A micro-review. Physica E-Low-Dimensional Systems & Nanostructures. Jan; 2010 42(3):662–665.
50. Iben I, Braunstein D, Doster W, Frauenfelder H, Hong MK, Johnson JB, Luck S, Ormos P, Schulte A, Steinbach PJ, Xie AH, Young RD. Glassy behavior of a protein. Phys Rev Lett. Apr 17; 1989 62(16):1916–1919. [PubMed: 10039803]

51. Bryngelson JD, Wolynes PG. Spin glasses and the statistical mechanics of protein folding. Proc Nat Acad Sci USA. 987; 84(21):7524.
52. Goldstein RA, Luthey-Schulten ZA, Wolynes PG. Protein tertiary structure recognition using optimized hamiltonians with local interactions. Proc Nat Acad Sci USA. Oct 1; 1992 89(19):9029–9033. [PubMed: 1409599]
53. Gō N. Theoretical studies of protein folding. Ann rev biophys and bioeng. 1983; 12(1):183–210. [PubMed: 6347038]
54. Clementi C, Nymeyer H, Onuchic JN. Topological and energetic factors: What determines the structural details of the transition state ensemble and “en-route” intermediates for protein folding? An investigation for small globular proteins. J Mol Biol. Jan; 2000 298(5):937–953. [PubMed: 10801360]
55. Cheung MS, Chavez LL, Onuchic JN. The energy landscape for protein folding and possible connections to function. Polymer. Jan; 2004 45(2):547–555.
56. Whitford PC, Noel JK, Gosavi S, Schug A, Sanbonmatsu KY, Onuchic JN. An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields. Proteins. May; 2009 75(2):430–41. [PubMed: 18837035]
57. Dill KA, Bromberg S, Yue KZ, Fiebig KM, Yee DP, Thomas PD, Chan HS. Principles of protein-folding - a perspective from simple exact models. Protein Sci. Jan; 1995 4(4):561–602. [PubMed: 7613459]
58. Sali A, Shakhnovich E, Karplus M. Kinetics of protein-folding - a lattice model study of the requirements for folding to the native-state. J Mol Biol. Feb 4; 1994 235(5):1614–1636. [PubMed: 8107095]
59. Gō N, Taketomi H. Respective roles of short- and long-range interactions in protein folding. Proc Nat Acad Sci USA. Mar.1978 75:1–5. [PubMed: 16592476]
60. Socci ND, Onuchic JN. Folding kinetics of protein-like heteropolymers. J Chem Phys. Jan; 1994 101(2):1519–1528.
61. Shrivastava I, Vishveshwara S, Cieplak M, Maritan A, Banavar JR. Lattice model for rapidly folding protein-like heteropolymers. Proc Nat Acad Sci USA. Sep 26; 1995 92(20):9206–9209. [PubMed: 7568102]
62. Socci ND, Onuchic JN, Wolynes PG. Protein folding mechanisms and the multidimensional folding funnel. Proteins. Jan; 1998 32(2):136–158. [PubMed: 9714155]
63. Onuchic JN, Nymeyer H, Garcia AE, Chahine J, Socci ND. The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios. Adv Protein Chem. Jan.2000 53:87–152. [PubMed: 10751944]
64. Camacho CJ, Thirumalai D. Kinetics and thermodynamics of folding in model proteins. Proc Nat Acad Sci USA. Jul 1; 1993 90(13):6369–6372. [PubMed: 8327519]
65. Klimovn DK, Thirumalai D. Factors governing the foldability of proteins. Proteins-Structure Function And Bioinformatics. Dec; 1996 26(4):411–441.
66. Sorenson JM, Head-Gordon T. The importance of hydration for the kinetics and thermodynamics of protein folding: simplified lattice models. Folding & Design. 1998; 3(6):523–534. [PubMed: 9889163]
67. Shakhnovich E, Farztdinov G, Gutin AM, Karplus M. Protein folding bottlenecks: A lattice monte carlo simulation. Phys Rev Lett. Sep.1991 67:1665–1668. [PubMed: 10044213]
68. Dinner A, Sali A, Karplus M, Shakhnovich E. Phase-diagram of a model protein-derived by exhaustive enumeration of the conformations. J Chem Phys. Jul 15; 1994 101(2):1444–1451.
69. Saven JG, Wolynes PG. Statistical mechanics of the combinatorial synthesis and analysis of folding macromolecules. Phys Chem B. Oct 9; 1997 101(41):8375–8389.
70. Yang S, Onuchic JN, Levine H. Effective stochastic dynamics on a protein folding energy landscape. J Chem Phys. Jan.2006 125(5):054910. [PubMed: 16942260]
71. Yang S, Onuchic JN, Garcia AE, Levine H. Folding time predictions from all-atom replica exchange simulations. J Mol Biol. Jan; 2007 372(3):756–763. [PubMed: 17681536]
72. Chavez LL, Onuchic JN, Clementi C. Quantifying the roughness on the free energy landscape: Entropic bottlenecks and protein folding rates. J Am Chem Soc. Jan; 2004 126(27):8426–8432. [PubMed: 15237999]

73. Chahine J, Oliveira RJ, Leite VBP, Wang J. Configuration-dependent diffusion can shift the kinetic transition state and barrier height of protein folding. Proc Nat Acad Sci USA. Sep; 2007 104(37):14646–51. [PubMed: 17804812]
74. Hummer G. Position-dependent diffusion coefficients and free energies from bayesian analysis of equilibrium and replica molecular dynamics simulations. New J Phys. Feb.2005 7:34–34.
75. Best R, Hummer G. Diffusive model of protein folding dynamics with Kramers turnover in rate. Phys Rev Lett. Jan.2006 96:228104. [PubMed: 16803349]
76. Das P, Moll M, Stamati H, Kavraki LE, Clementi C. Low-dimensional, free-energy landscapes of protein-folding reactions by nonlinear dimensionality reduction. Proc Nat Acad Sci USA. Jan; 2006 103(26):9885–9890. [PubMed: 16785435]
77. Oliveira RJ, Whitford PC, Chahine J, Leite VBP, Wang J. Coordinate and time-dependent diffusion dynamics in protein folding. Methods. May.2010 52:91–98. [PubMed: 20438841]
78. Oliveira RJ, Whitford PC, Chahine J, Wang J, Onuchic JN, Leite VBP. The origin of nonmonotonic complex behavior and the effects of nonnative interactions on the diffusive properties of protein folding. Biophys J. Jul; 2010 99(2):600–8. [PubMed: 20643080]
79. Cho SS, Levy Y, Wolynes PG. P versus Q: structural reaction coordinates capture protein folding on smooth landscapes. Proc Nat Acad Sci USA. Jan; 2006 103(3):586–91. [PubMed: 16407126]
80. Guo Z, Brooks CL. Thermodynamics of protein folding: a statistical mechanical study of a small all-β protein. Biopolymers. Dec; 1997 42(7):745–57. [PubMed: 10904547]
81. Cheung MS, Garcia AE, Onuchic JN. Protein folding mediated by solvation: Water expulsion and formation of the hydrophobic core occur after the structural collapse. Proc Nat Acad Sci USA. Jan; 2002 99(2):685–690. [PubMed: 11805324]
82. Levy Y, Onuchic JN. Water mediation in protein folding and molecular recognition. Annu Rev Bioph Biom. Jan.2006 35:389–415.
83. Oliveira LC, Silva RTH, Leite VBP, Chahine J. Frustration and hydrophobicity interplay in protein folding and protein evolution. J Chem Phys. 2006; 125:084904. [PubMed: 16965054]
84. Hillson N, Onuchic JN, Garcia AE. Pressure-induced protein-folding/unfolding kinetics. Proc Nat Acad Sci USA. Jan; 1999 96(26):14848–14853. [PubMed: 10611301]
85. Wales DJ, Dewsbury PEJ. Effect of salt bridges on the energy landscape of a model protein. J Chem Phys. Jan; 2004 121(20):10284–10290. [PubMed: 15549905]
86. Karanicolas J, Brooks CL. The importance of explicit chain representation in protein folding models: an examination of ising-like models. Proteins. Nov; 2003 53(3):740–7. [PubMed: 14579364]
87. Pincus, DL.; Cho, SS.; Hyeon, C.; Thirumalai, D. Minimal models for proteins and RNA: From folding to function. In: Conn, PM., editor. Molecular Biology Of Protein Folding, Pt B, volume 84 of Progress in Molecular Biology and Translational Science, pages 203+. Elsevier Academic Press Inc; 525 B Street, Suite 1900, San Diego, Ca 92101-4495 Usa: 2008.
88. Kouza M, Li MS, O'Brien EP, Hu CY, Thirumalai D. Effect of finite size on cooperativity and rates of protein folding. J Phys Chem A. Jan; 2006 110(2):671–6. [PubMed: 16405339]
89. Komatsuzaki T, Hoshino K, Matsunaga Y, Rylance GJ, Johnston RL, Wales DJ. How many dimensions are required to approximate the potential energy landscape of a model protein? J Chem Phys. Jan.2005 122(8):084714.
90. Chan HS, Zhang Z, Wallin S, Liu Z. Cooperativity, local-nonlocal coupling, and nonnative interactions: Principles of protein folding from coarse-grained models. 2011; 62:301–326.
91. Hills RD, Brooks CL. Insights from coarse-grained Gō models for protein folding and dynamics. Int J Mol Sci. Mar; 2009 10(3):889–905. [PubMed: 19399227]
92. Hyeon C, Thirumalai D. Capturing the essence of folding and functions of biomolecules using coarse-grained models. Nat Comms. Jan.2011 2:487.
93. Plotkin SS, Onuchic JN. Understanding protein folding with energy landscape theory - part I: Basic concepts. Quart Rev Biophys. Jan; 2002 35(2):111–167.
94. Plotkin SS. Speeding protein folding beyond the Gō model: How a little frustration sometimes helps. Proteins. Dec; 2001 45(4):337–45. [PubMed: 11746681]

95. Givaty O, Levy Y. Protein sliding along DNA: Dynamics and structural characterization. J Mol Biol. Nov.2009 385:1087–1097. [PubMed: 19059266]

96. Das P, Matysiak Silvina, Clementi C. Balancing energy and entropy: a minimalist model for the characterization of protein folding landscapes. Proc Nat Acad Sci USA. Jul; 2005 102(29):10141–6. [PubMed: 16006532]

97. Samiotakis A, Cheung MS. Folding dynamics of trp-cage in the presence of chemical interference and macromolecular crowding. I. J Chem Phys. Nov.2011 135(17):175101. [PubMed: 22070323]

98. Sulkowska JI, Cieplak M. Mechanical stretching of proteins—a theoretical survey of the protein data bank. J Phys Condens Matter. Jun.2007 19(28):283201.

99. Sulkowska JI, Cieplak M. Selection of optimal variants of Gō-like models of proteins through studies of stretching. Biophys J. Oct; 2008 95(7):3174–91. [PubMed: 18567634]

100. Hyeon C, Morrison G, Pincus Dl, Thirumalai D. Refolding dynamics of stretched biopolymers upon force quench. Proc Nat Acad Sci USA. Dec; 2009 106(48):20288–93. [PubMed: 19915145]

101. Cieplak M, Sulkowska JI. Structure-based models of biomolecules: Stretching of proteins, dynamics of knots, hydrodynamic effects, and indentation of virus capsids. Multiscale Approaches to Protein Modeling. 2011:179–208.

102. O'Brien EP, Christodoulou J, Vendruscolo M, Dobson CM. New scenarios of protein folding can occur on the ribosome. J A m Chem Soc. Jan.2011 133:516–516.

103. Bellesia G, Shea JE. Diversity of kinetic pathways in amyloid fibril formation. J Chem Phys. 2009; 131:111102. [PubMed: 19778093]

104. Eastwood MP, Wolynes PG. Role of explicitly cooperative interactions in protein folding funnels: A simulation study. J Chem Phys. Mar 8; 2001 114(10):4702–4716.

105. Plotkin SS, Wang J, Wolynes PG. Statistical mechanics of a correlated energy landscape model for protein folding funnels. J Chem Phys. Feb 15;1997 106(7):2932–2948.

106. Clementi C, Jennings PA, Onuchic JN. Prediction of folding mechanism for circular-permuted proteins. J Mol Biol. Jan; 2001 311(4):879–890. [PubMed: 11518537]

107. Gosavi S, Chavez LL, Jennings PA, Onuchic JN. Topological frustration and the folding of interleukin-1β. J Mol Biol. Jan; 2006 357(3):986–996. [PubMed: 16469330]

108. Chavez LL, Gosavi S, Jennings Pa, Onuchic JN. Multiple routes lead to the native state in the energy landscape of the β-trefoil family. Proc Nat Acad Sci USA. Jan; 2006 103(27):10254–10258. [PubMed: 16801558]

109. Andrews BT, Gosavi S, Finke JM, Onuchic JN, Jennings PA. The dual-basin landscape in GFP folding. Proc Nat Acad Sci USA. Jan; 2008 105(34):12283–12288. [PubMed: 18713871]

110. Gosavi S, Whitford PC, Jennings PA, Onuchic JN. Extracting function from a β-trefoil folding motif. Proc Nat Acad Sci USA. Jan; 2008 105(30):10384–10389. [PubMed: 18650393]

111. Norcross TS, Yeates TO. A framework for describing topological frustration in models of protein folding. J Mol Biol. 2006; 362(3):605–621. [PubMed: 16930616]

112. McCammon JA, Wolynes PG, Karplus M. Picosecond dynamics of tyrosine side chains in proteins. Biochemistry. 1979; 18(6):927–942. [PubMed: 427100]

113. Garcia AE, Onuchic JN. Folding a protein in a computer: An atomic description of the folding/unfolding of protein A. Proc Nat Acad Sci USA. Jan; 2003 100(24):13898–13903. [PubMed: 14623983]

114. Yeates TO, Norcross TS, King NP. Knotted and topologically complex proteins as models for studying folding and stability. Curr Opin Chem Biol. 2007; 11(6):595–603. [PubMed: 17967433]

115. King NP, Jacobitz AW, Sawaya MR, Goldschmidt L, Yeates TO. Structure and folding of a designed knotted protein. Proc Nat Acad Sci USA. 2010; 107(48):20732. [PubMed: 21068371]

116. Roy M, Chavez LL, Finke JM, Heidary DK, Onuchic JN, Jennings PA. The native energy landscape for interleukin-1β. modulation of the population ensemble through native-state topology. J Mol Biol. Jan; 2005 348(2):335–347. [PubMed: 15811372]

117. Heidary DK, Roy M, Daumy GO, Cong Y, Jennings PA. Long-range coupling between separate docking sites in interleukin-1β. J Mol Biol. Nov; 2005 353(5):1187–98. [PubMed: 16216268]

118. Sulkowska JI, Sulkowski P, Szymczak P, Cieplak M. Stabilizing effect of knots on proteins. Proc Nat Acad Sci USA. Dec; 2008 105(50):19714–9. [PubMed: 19064918]

119. Noel JK, Sulkowska JI, Onuchic JN. Slipknotting upon native-like loop formation in a trefoil knot protein. Proc Nat Acad Sci USA. Aug; 2010 107(35):15403–8. [PubMed: 20702769]
120. Capraro DT, Roy M, Onuchic JN, Jennings PA. Backtracking on the folding landscape of the β-trefoil protein interleukin-1β? Proc Nat Acad Sci USA. Sep 30; 2008 105(39):14844–14848. [PubMed: 18806223]
121. Baxter EL, Jennings Pa, Onuchic JN. Interdomain communication revealed in the diabetes drug target mitoneet. Proc Nat Acad Sci USA. Mar.2011 108:5266–5271. [PubMed: 21402934]
122. Clementi C, Garcia AE, Onuchic JN. Interplay among tertiary contacts, secondary structure formation and side-chain packing in the protein folding mechanism: All-atom representation study of protein L. J Mol Biol. Jan; 2003 326(3):933–954. [PubMed: 12581651]
123. Zhou Y, Linhananta A. Role of hydrophilic and hydrophobic contacts in folding of the second hairpin fragment of protein G: Molecular dynamics simulation studies of an all-atom model. Prot Struct Func Bioinfo. 2002; 47(2):154–162.
124. Luo Z, Ding J, Zhou Y. Folding mechanisms of individual β-hairpins in a Go model of Pin1 WW domain by all-atom molecular dynamics simulations. J Chem Phys. 2008; 128:225103. [PubMed: 18554060]
125. Whitford PC, Schug A, Saunders J, Hennelly SP, Onuchic JN, Sanbonmatsu KY. Nonlocal helix formation is key to understanding S-adenosylmethionine-1 riboswitch function. Biophys J. Jan; 2009 96(2):L7–L9. [PubMed: 19167285]
126. Shimada J, Kussell EL, Shakhnovich EI. The folding thermodynamics and kinetics of crambin using an all-atom monte carlo simulation. J Mol Biol. Apr; 2001 308(1):79–95. [PubMed: 11302709]
127. Sorin EJ, Nakatani BJ, Rhee YM, Jayachandran G, Vishal V, Pande VS. Does native state topology determine the RNA folding mechanism? J Mol Biol. Apr; 2004 337(4):789–97. [PubMed: 15033351]
128. Chen K, Eargle J, Lai J, Kim H, Abeysirigunawardena S, Mayerle M, Woodson SA, Ha T, Schulten Z. Assembly of the five-way junction in the ribosomal small subunit using hybrid MD-Ĝ simulations. J Phys Chem B. Mar.2012 in press.
129. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W. Atomic-level characterization of the structural dynamics of proteins. Science. Oct; 2010 330(6002):341–6. [PubMed: 20947758]
130. Greenfeder SA, Varnell T, Powers G, Lombardgillooly K, Shuster D, Mcintyre KW, Ryan DE, Levin W, Madison V, Ju G. Insertion of a structural domain of interleukin (IL)-1β confers agonist activity to the Il-1 receptor antagonist - implications for IL-1 bioactivity. J Biol Chem. Sep 22; 1995 270(38):22460–22466. [PubMed: 7673234]
131. Vigers GPA, Anderson LJ, Caffes P, Brandhuber BJ. Crystal structure of the type-I interleukin-1 receptor complexed with interleukin-1β. Nature. Mar 13; 1997 386(6621):190–194. [PubMed: 9062193]
132. Koussounadis AI, Ritchie DW, Kemp GJL, Secombes CJ. Analysis of fish IL-1β and derived peptide sequences indicates conserved structures with species-specific IL-1 receptor binding: Implications for pharmacological design. Current Pharmaceutical Design. 2004; 10(31):3857–3871. [PubMed: 15579075]
133. Miyashita O, Onuchic JN, Wolynes PG. Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. Proc Nat Acad Sci USA. Jan; 2003 100(22): 12570–12575. [PubMed: 14566052]
134. Whitford PC, Onuchic JN, Wolynes PG. Energy landscape along an enzymatic reaction trajectory: hinges or cracks? Hfsp J. Jan; 2008 2(2):61–64. [PubMed: 19404472]
135. Itoh K, Sasai M. Statistical mechanics of protein allostery: roles of backbone and side-chain structural fluctuations. J Chem Phys. Mar.2011 134(12):125102. [PubMed: 21456702]
136. Portman JJ, Takada S, Wolynes PG. Microscopic theory of protein folding rates. I. Fine structure of the free energy profile and folding routes from a variational approach. J Chem Phys. Mar 15; 2001 114(11):5069–5081.
137. Sulkowska JI, Cieplak M. Selection of optimal variants of Go-like models of proteins through studies of stretching. Biophys J. Oct 1; 2008 95(7):3174–3191. [PubMed: 18567634]

138. Lammert H, Schug A, Onuchic JN. Robustness and generalization of structure-based models for protein folding and function. Proteins. Dec; 2009 77(4):881–91. [PubMed: 19626713]
139. Thirumalai D, Lorimer GH. Chaperonin-mediated protein folding. Ann Rev Biophys Biomol Struct. 2001; 30:245–269. [PubMed: 11340060]
140. Huse M, Kuriyan J. The conformational plasticity of protein kinases. Cell. May; 2002 109(3): 275–82. [PubMed: 12015977]
141. Munro JB, Sanbonmatsu KY, Spahn CMT, Blanchard SC. Navigating the ribosome's metastable energy landscape. Trends Biochem Sci. Jul.2009 34:390–400. [PubMed: 19647434]
142. Ogawa A, Takayama Y, Sakai H, Chong KT, Takeuchi S, Nakagawa A, Nada S, Okada M, Tsukihara T. Structure of the carboxyl-terminal src kinase, Csk. J Biol Chem. Apr; 2002 277(17): 14351–4. [PubMed: 11884384]
143. Wong L, Lieser SA, Miyashita O, Miller M, Tasken K, Onuchic JN, Adams JA, Woods VL, Jennings PA. Coupled motions in the SH2 and kinase domains of Csk control Src phosphorylation. J Mol Biol. Jan; 2005 351(1):131–143. [PubMed: 16002086]
144. Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, Kern D. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. Nat Struct Mol Biol. Oct; 2004 11(10):945–9. [PubMed: 15334070]
145. Block SM. Kinesin motor mechanics: Binding, stepping, tracking, gating, and limping. Biophys J. May 1; 2007 92(9):2986–2995. [PubMed: 17325011]
146. Korostelev A, Ermolenko Dn, Noller HF. Structural dynamics of the ribosome. Curr Opin Chem Biol. Dec; 2008 12(6):674–83. [PubMed: 18848900]
147. Marshall RA, Aitken CE, Dorywalska M, Puglisi JD. Translation at the single-molecule level. Annu Rev Biochem. Jan.2008 77:177–203. [PubMed: 18518820]
148. Blanchard S. Single-molecule observations of ribosome function. Curr Op Struct Biol. Feb.2009 19:103–109.
149. Kumar S, Ma B, Tsai CJ, Wolfson H, Nussinov R. Folding funnels and conformational transitions via hinge-bending motions. Cell Biochem Biophys. Jan; 1999 31(2):141–64. [PubMed: 10593256]
150. Goldstein, H. Classical Mechanics. 3rd. Addison-Wesley Publishing Company; Reading, MA: 2001.
151. Bahar I, Chennubhotla C, Tobi D. Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. Curr Op Struct Biol. 2007; 17(6):633–640.
152. Chennubhotla, Chakra; Rader, AJ.; Yang, Lee-Wei; Bahar, Ivet. Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies. Phys Biol. Nov; 2005 2(4):S173–80. [PubMed: 16280623]
153. Wang Y, Rader AJ, Bahar I, Jernigan RL. Global ribosome motions revealed with elastic network model. Journal of Structural Biology. Sep; 2004 147(3):302–14. [PubMed: 15450299]
154. Tama F, Sanejouand YH. Conformational change of proteins arising from normal mode calculations. Protein Eng. Jan; 2001 14(1):1–6. [PubMed: 11287673]
155. Tirion MM. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys Rev Lett. Aug 26; 1996 77(9):1905–1908. [PubMed: 10063201]
156. Whitford PC, Gosavi S, Onuchic JN. Conformational transitions in adenylate kinase - allosteric communication reduces misligation. J Biol Chem. Jan; 2008 283(4):2042–2048. [PubMed: 17998210]
157. Kurkcuoglu O, Doruker P, Sen TZ, Kloczkowski A, Jernigan RL. The ribosome structure controls and directs mRNA entry, translocation and exit dynamics. Phys Biol. Jan.2008 5(4):46005.
158. Kurkcuoglu O, Kurkcuoglu Z, Doruker P, Jernigan RL. Collective dynamics of the ribosomal tunnel revealed by elastic network modeling. Prot Struct Func Bioinfo. 2009; 75(4):837–845.
159. Sen TZ, Feng Y, Garcia JV, Kloczkowski A, Jernigan RL. The extent of cooperativity of protein motions observed with elastic network models is similar for atomic and coarser-grained models. J Chem Theory Comp. 2006; 2(3):696–704.

160. Jiang J, Shrivastava IH, Watts SD, Bahar I, Amara SG. Large collective motions regulate the functional properties of glutamate transporter trimers. Proc Nat Acad Sci USA. Sep; 2011 108(37):15141–6. [PubMed: 21876140]

161. Tama F, Valle Mikel, Frank J, Brooks CL. Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. Proc Nat Acad Sci USA. Aug; 2003 100(16):9319–23. [PubMed: 12878726]

162. Trylska J, Tozzini V, McCammon JA. Exploring global motions and correlations in the ribosome. Biophys J. Sep; 2005 89(3):1455–63. [PubMed: 15951386]

163. Whitford, PC.; Altman, RB.; Geggier, P.; Terry, D.; Munro, JB.; Onuchic, JN.; Spahn, CMT.; Sanbonmatsu, KY.; Blanchard, SC. Ribosomes: Structure, Function and Dynamics. Vol. chapter 24. Springer; 2011.

164. Trabuco LG, Schreiner E, Eargle J, Cornish P, Ha T, Luthey-Schulten Z, Schulten K. The role of L1 stalk-tRNA interaction in the ribosome elongation cycle. J Mol Biol. Oct; 2010 402(4):741–60. [PubMed: 20691699]

165. Maragakis, P.; Karplus, M. Large amplitude conformational change in proteins explored with a plastic network model: Adenylate Kinase. J Mol Biol; Workshop For Multiscale Modelling And Simulation; Lugano, Switzerland. AUG, 2003; Sep 30. 2005 p. 807-822.

166. Schuyler AD, Jernigan RL, Qasba PK, Ramakrishnan B, Chirikjian GS. Iterative cluster-NMA: A tool for generating conformational transitions in proteins. Proteins. Feb; 2009 74(3):760–76. [PubMed: 18712827]

167. Kubitzki MB, de Groot BL. The atomistic mechanism of conformational transition in adenylate kinase: a TEE-REX molecular dynamics study. Structure. Aug; 2008 16(8):1175–82. [PubMed: 18682219]

168. Lu Q, Wang J. Kinetics and statistical distributions of single-molecule conformational dynamics. J Phys Chem B. Jan.2009 113:1517–1521. [PubMed: 19140753]

169. Daily, Md; Phillips, GN.; Cui, Q. Many local motions cooperate to produce the adenylate kinase conformational transition. J Mol Biol. Jul; 2010 400(3):618–31. [PubMed: 20471396]

170. Korkut A, Hendrickson W. Computation of conformational transitions in proteins by virtual atom molecular mechanics as validated in application to adenylate kinase. Proc Nat Acad Sci USA. Aug.2009 106:15673–15678. [PubMed: 19706894]

171. Jana B, Adkar BV, Biswas R, Bagchi B. Dynamic coupling between the LID and NMP domain motions in the catalytic conversion of ATP and AMP to ADP by adenylate kinase. J Chem Phys. Jan 21;2011 134(3):035101. [PubMed: 21261390]

172. Kantarci-Carsibasi N, Haliloglu T, Doruker P. Conformational transition pathways explored by Monte Carlo simulation integrated with collective modes. Biophys J. Aug.2008 95:5862–5873. [PubMed: 18676657]

173. García AE, Krumhansl JA, Frauenfelder H. Variations on a theme by Debye and Waller: from simple crystals to proteins. Proteins. Oct; 1997 29(2):153–60. [PubMed: 9329080]

174. Korostelev A, Noller HF. Analysis of structural dynamics in the ribosome by TLS crystallographic refinement. J Mol Biol. Nov; 2007 373(4):1058–70. [PubMed: 17897673]

175. Poon BK, Chen X, Lu M, Vyas NK, Quiocho FA, Wang Q, Ma Jianpeng. Normal mode refinement of anisotropic thermal parameters for a supramolecular complex at 3.42-Å crystallographic resolution. Proc Nat Acad Sci USA. May; 2007 104(19):7869–74. [PubMed: 17470791]

176. Best, Rb; Chen, YG.; Hummer, G. Slow protein conformational dynamics from multiple experimental structures: the helix/sheet transition of arc repressor. Structure (London England: 1993). Dec; 2005 13(12):1755–63.

177. Whitford PC, Miyashita O, Levy Y, Onuchic JN. Conformational transitions of adenylate kinase: Switching by cracking. J Mol Biol. Jan; 2007 366(5):1661–1671. [PubMed: 17217965]

178. Okazaki K, Takada S. Dynamic energy landscape view of coupled binding and protein conformational change: Induced-fit versus population-shift mechanisms. Proc Nat Acad Sci USA. Jan; 2008 105(32):11182–11187. [PubMed: 18678900]

179. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. J Comp Chem. 2005; 26(16):1668. [PubMed: 16200636]
180. Brooks B, Brooks C, Mackerell A, Nilsson L, Petrella R, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner A, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor R, Post C, Pu J, Schaefer M, Tidor B, Venable R, Woodcock H, Wu X, Yang W, York D, Karplus M. CHARMM: The biomolecular simulation program. J Comp Chem. May.2009 30:1545–1614. [PubMed: 19444816]
181. Chen J, Brooks CL. Implicit modeling of nonpolar solvation for simulating protein folding and conformational transitions. Phys Chem Chem Phys. Jan; 2008 10(4):471–81. [PubMed: 18183310]
182. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nat Struct Mol Biol. 2002; 9(9):646–652.
183. Grant BJ, Gorfe AA, McCammon JA. Large conformational changes in proteins: signaling and other functions. Curr Opin Struct Biol. 2010; 20(2):142–147. [PubMed: 20060708]
184. Klepeis JL, Lindorff-Larsen K, Dror RO, Shaw DE. Long-timescale molecular dynamics simulations of protein structure and function. Curr Opin Struct Biol. Apr; 2009 19(2):120–7. [PubMed: 19361980]
185. Zuckerman DM. Simulation of an ensemble of conformational transitions in a united-residue model of calmodulin. J Phys Chem B. Apr 22; 2004 108(16):5127–5137.
186. Brokaw JB, Chu JW. On the Roles of Substrate Binding and Hinge Unfolding in Conformational Changes of Adenylate Kinase. Biophys J. Nov 17; 2010 99(10):3420–3429. [PubMed: 21081091]
187. Arora K, Brooks CL. Large-scale allosteric conformational transitions of adenylate kinase appear to involve a population-shift mechanism. Proc Nat Acad Sci USA. Nov; 2007 104(47):18496–501. [PubMed: 18000050]
188. Hanson JA, Duderstadt K, Watkins LP, Bhattacharyya S, Brokaw J, Chu JW, Yang H. Illuminating the mechanistic roles of enzyme conformational dynamics. Proc Nat Acad Sci USA. Nov; 2007 104(46):18055–60. [PubMed: 17989222]
189. Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, Huebner CG, Kern D. Intrinsic motions along an enzymatic reaction trajectory. Nature. Jan; 2007 450(7171):838–U13. [PubMed: 18026086]
190. Vendruscolo M, Dobson CM. Protein Dynamics: Moore's Law in Molecular Biology. Current Biology. Jan 25; 2011 21(2):R68–R70. [PubMed: 21256436]
191. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How fast-folding proteins fold. Science. Oct; 2011 334(6055):517–520. [PubMed: 22034434]
192. Schulz R, Lindner B, Petridis L, Smith JC. Scaling of multimillion-atom biological molecular dynamics simulation on a petascale supercomputer. J Chem Theory Comput. 2009; 5(10):2798–2808.
193. Müller CW, Schulz GE. Structure of the complex between adenylate kinase from escherichia coli and the inhibitor Ap5A refined at 1.9 Å resolution: A model for a catalytic transition state. J Mol Biol. 1992; 224(1):159–177. [PubMed: 1548697]
194. Müller CW, Schlauderer GJ, Reinstein J, Schulz GE. Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. Structure. Feb; 1996 4(2):147–56. [PubMed: 8805521]
195. Schlauderer GJ, Proba K, Schulz GE. Structure of a mutant adenylate kinase ligated with an ATP-analogue showing domain closure over ATP. J Mol Biol. Feb; 1996 256(2):223–7. [PubMed: 8594191]
196. Okazaki K, Koga N, Takada S, Onuchic JN, Wolynes PG. Multiple-basin energy landscapes for large-amplitude conformational motions of proteins: Structure-based molecular dynamics simulations. Proc Nat Acad Sci USA. Jan; 2006 103(32):11844–11849. [PubMed: 16877541]
197. Yang S, Roux Benoît. Src kinase conformational activation: thermodynamics, pathways, and mechanisms. PLoS Comput Biol. Mar.2008 4(3):e1000047. [PubMed: 18369437]

198. Zwanzig R. Diffusion in a rough potential. Proc Nat Acad Sci USA. Apr; 1988 85(7):2029–30. [PubMed: 3353365]

199. Kramers H. Brownian motion in a field of force and the diffusion model of chemical reactions. Physica. Jan.1940 7:284–304.

200. Kubelka J, Hofrichter J, Eaton WA. The protein folding 'speed limit'. Curr Op Struct Biol. Feb; 2004 14(1):76–88.

201. Tang J, Kang SG, Saven JG, Gai F. Characterization of the cofactor-induced folding mechanism of a zinc-binding peptide using computationally designed mutants. J Mol Biol. May; 2009 389(1):90–102. [PubMed: 19361525]

202. Whitford PC, Onuchic JN, Sanbonmatsu KY. Connecting energy landscapes with experimental rates for aminoacyl-tRNA accommodation in the ribosome. J Am Chem Soc. Aug.2010 132:13170–13171. [PubMed: 20806913]

203. Okamoto Y. Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. J Mol Graph Model. Jan; 2004 22(5):425–439. [PubMed: 15099838]

204. Nymeyer H, Gnanakaran S, Garcia AE. Atomic simulations of protein folding, using the replica exchange algorithm. Method Enzymol. Jan.2004 383:119–149.

205. Rodinger T, Pomes R. Enhancing the accuracy, the efficiency and the scope of free energy simulations. Curr Opin Struc Biol. Jan; 2005 15(2):164–170.

206. Gnanakaran S, Nymeyer H, Portman J, Sanbonmatsu KY, Garcia AE. Peptide folding simulations. Curr Opin Struc Biol. Jan; 2003 13(2):168–174.

207. Roux B. The calculation of the potential of mean force using computer-simulations. Comput Phys Commun. Jan; 1995 91(1-3):275–282.

208. Carr JM, Trygubenko SA, Wales DJ. Finding pathways between distant local minima. J Chem Phys. Jan.2005 122(23):234903. [PubMed: 16008483]

209. Maragliano L, Fischer A, Vanden-Eijnden E, Ciccotti G. String method in collective variables: Minimum free energy paths and isocommittor surfaces. J Chem Phys. Jan.2006 125(2):024106.

210. Sheppard D, Terrell R, Henkelman G. Optimization methods for finding minimum energy paths. J Chem Phys. Jan.2008 128(13):134106. [PubMed: 18397052]

211. Koslover EF, Wales DJ. Comparison of double-ended transition state search methods. J Chem Phys. Jan.2007 127(13):134102. [PubMed: 17919006]

212. Peters B, Heyden A, Bell AT, Chakraborty A. A growing string method for determining transition states: Comparison to the nudged elastic band and string methods. J Chem Phys. Jan; 2004 120(17):7877–7886. [PubMed: 15267702]

213. Henkelman G, Uberuaga BP, Jonsson H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. J Chem Phys. Jan; 2000 113(22):9901–9904.

214. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett. Jan; 1999 314(1-2):141–151.

215. Garcia AE, Sanbonmatsu KY. Exploring the energy landscape of a β hairpin in explicit solvent. Proteins. Jan; 2001 42(3):345–354. [PubMed: 11151006]

216. Zhou RH, Berne BJ, Germain R. The free energy landscape for β hairpin folding in explicit water. P Nat Acad Sci USA. Jan; 2001 98(26):14931–14936.

217. Zhou RH. Free energy landscape of protein folding in water: Explicit vs. implicit. solvent Proteins. Jan; 2003 53(2):148–161.

218. Rao F, Caflisch A. Replica exchange molecular dynamics simulations of reversible folding. J Chem Phys. Jan; 2003 119(7):4035–4042.

219. Baumketner A, Shea JE. The structure of the alzheimer amyloid β 10-35 peptide probed through replica-exchange molecular dynamics simulations in explicit solvent. J Mol Biol. Jan; 2007 366(1):275–285. [PubMed: 17166516]

220. Scheraga HA, Khalili M, Liwo A. Protein-folding dynamics: Overview of molecular simulation techniques. Annu Rev Phys Chem. Jan.2007 58:57–83. [PubMed: 17034338]

221. Mathews DH, Case DA. Nudged elastic band calculation of minimal energy paths for the conformational change of a gg non-canonical pair. J Mol Biol. Apr; 2006 357(5):1683–93. [PubMed: 16487974]

222. Wales, DJ. Energy Landscapes: Applications to Clusters, Biomolecules and Glasses. 1st. Cambridge University Press; Cambridge, UK: 2004.

223. Miller TF, Vanden-Eijnden E, Chandler D. Solvent coarse-graining and the string method applied to the hydrophobic collapse of a hydrated chain. Proc Nat Acad Sci USA. Jan; 2007 104(37): 14559–14564. [PubMed: 17726097]

224. Haldane, JSB. Enzymes. Longmans Green; New York, NY: 1930.

225. Pauling L. Nature of forces between large molecules of biological interest. Nature. Jan; 1948 161(4097):707–709. [PubMed: 18860270]

226. Horiuchi T, Gō N. Projection of monte-carlo and molecular-dynamics trajectories onto the normal mode axes - human lysozyme. Proteins. Jan; 1991 10(2):106–116. [PubMed: 1896424]

227. Bae E, Phillips GN. Roles of static and dynamic domains in stability and catalysis of adenylate kinase. Proc Nat Acad Sci USA. Jan; 2006 103(7):2132–2137. [PubMed: 16452168]

228. Miyashita O, Wolynes PG, Onuchic JN. Simple energy landscape model for the kinetics of functional transitions in proteins. J Phys Chem B. Jan; 2005 109(5):1959–1969. [PubMed: 16851180]

229. Rundqvist L, Adén J, Sparrman T, Wallgren M, Olsson U, Wolf-Watz M. Noncooperative folding of subdomains in adenylate kinase. Biochemistry. Mar; 2009 48(9):1911–27. [PubMed: 19219996]

230. Schrank T, Elam W, Li J, Hilser VJ. Strategies for the thermodynamic characterization of linked binding/local folding reactions within the native state application to the LID domain of adenylate. Methods in enzymology. Jan.2011

231. Lammert H, Wolynes Pg, Onuchic JN. The role of atomic level steric effects and attractive forces in protein folding. Proteins. Sep.2012 80:362–373.

232. Hyeon C, Onuchic JN. Internal strain regulates the nucleotide binding site of the kinesin leading head. Proc Nat Acad Sci USA. Jan; 2007 104(7):2175–2180. [PubMed: 17287347]

233. Hyeon C, Onuchic JN. Mechanical control of the directional stepping dynamics of the kinesin motor. Proc Nat Acad Sci USA. Jan; 2007 104(44):17382–17387. [PubMed: 17959770]

234. Tripathi S, Portman JJ. Inherent flexibility determines the transition mechanisms of the EF-hands of calmodulin. Proc Nat Acad Sci USA. Feb; 2009 106(7):2104–9. [PubMed: 19190183]

235. Hyeon C, Jennings PA, Adams JA, Onuchic JN. Ligand-induced global transitions in the catalytic domain of protein kinase A. Proc Nat Acad Sci USA. Mar; 2009 106(9):3023–8. [PubMed: 19204278]

236. Olsson U, Wolf-Watz M. Overlap between folding and functional energy landscapes for adenylate kinase conformational change. Nat Comms. Nov.2010 1(8):111.

237. Peng C, Zhang L, Head-Gordon T. Instantaneous normal modes as an unforced reaction coordinate for protein conformational transitions. Biophys J. May; 2010 98(10):2356–64. [PubMed: 20483345]

238. Walczak AM, Onuchic JN, Wolynes PG. Absolute rate theories of epigenetic stability. Proc Nat Acad Sci USA. Jan; 2005 102(52):18926–18931. [PubMed: 16361441]

239. Schultz D, Jacob E Ben, Onuchic JN, Wolynes PG. Molecular level stochastic model for competence cycles in bacillus subtilis. Proc Nat Acad Sci USA. Jan; 2007 104(45):17582–17587. [PubMed: 17962411]

240. Shoemaker BA, Portman JJ, Wolynes PG. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc Nat Acad Sci USA. Aug; 2000 97(16):8868–73. [PubMed: 10908673]

241. Levy Y, Cho SS, Onuchic JN, Wolynes PG. A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. J Mol Biol. Jan; 2005 346(4):1121–1145. [PubMed: 15701522]

242. Levy Y, Onuchic JN, Wolynes PG. Fly-casting in protein-DNA binding: Frustration between protein folding and electrostatics facilitates target recognition. J Am Chem Soc. Jan; 2007 129(4):738–739. [PubMed: 17243791]

243. Chen K, Eargle J, Sarkar K, Gruebele M, Luthey-Schulten Z. Functional role of ribosomal signatures. Biophys J. Jan.2010 99:3930–3940. [PubMed: 21156135]

244. Vuzman D, Levy Y. DNA search efficiency is modulated by charge composition and distribution in the intrinsically disordered tail. Proc Nat Acad Sci USA. 2010; 107(49):21004. [PubMed: 21078959]

245. Munson M, Obrien R, Sturtevant JM, Regan L. Redesigning the hydrophobic core of a 4-helix-bundle protein. Protein Science. Nov; 1994 3(11):2015–2022. [PubMed: 7535612]

246. Levy Y, Cho SS, Shen T, Onuchic JN, Wolynes PG. Symmetry and frustration in protein energy landscapes: a near degeneracy resolves the Rop dimer-folding mystery. Proc Nat Acad Sci USA. Feb; 2005 102(7):2373–8. [PubMed: 15701699]

247. Schug A, Whitford PC, Levy Y, Onuchic JN. Mutations as trapdoors to two competing native conformations of the Rop-dimer. Proc Nat Acad Sci USA. Jan; 2007 104(45):17674–17679. [PubMed: 17968016]

248. Gambin Y, Schug A, Lemke Ea, Lavinder JJ, Ferreon ACM, Magliery TJ, Onuchic JN, Deniz AA. Direct single-molecule observation of a protein living in two opposed native structures. Proc Nat Acad Sci USA. Jun; 2009 106(25):10153–8. [PubMed: 19506258]

249. Breaker RR. Prospects for riboswitch discovery and analysis. Molecular Cell. Sep; 2011 43(6): 867–79. [PubMed: 21925376]

250. Blouin S, Mulhbacher J, Penedo JC, Lafontaine DA. Riboswitches: ancient and promising genetic regulators. ChemBioChem. Feb; 2009 10(3):400–16. [PubMed: 19101979]

251. Winkler W, Nahvi A, Breaker RR. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature. Oct; 2002 419(6910):952–6. [PubMed: 12410317]

252. Hennelly SP, Sanbonmatsu KY. Tertiary contacts control switching of the SAM-I riboswitch. Nucleic Acids Res. Mar; 2011 39(6):2416–31. [PubMed: 21097777]

253. Heppell B, Blouin S, Dussault AM, Mulhbacher J, Ennifar E, Penedo JC, Lafontaine DA. Molecular insights into the ligand-controlled organization of the SAM-I riboswitch. Nature Chemical Biology. May.2011

254. Lin JC, Thirumalai D. Relative stability of helices determines the folding landscape of adenine riboswitch aptamers. J Am Chem Soc. Oct; 2008 130(43):14080–1. [PubMed: 18828635]

255. Feng J, Walter NG, Brooks CL. Cooperative and directional folding of the preQ1 riboswitch aptamer domain. J Am Chem Soc. Mar; 2011 133(12):4196–9. [PubMed: 21375305]

256. Greenleaf WJ, Frieda Kl, Foster DAN, Woodside MT, Block SM. Direct observation of hierarchical folding in single riboswitch aptamers. Science. Feb; 2008 319(5863):630–3. [PubMed: 18174398]

257. Green R, Noller HF. Ribosomes and translation. Annu Rev Biochem. Jan.1997 66:679–716. [PubMed: 9242921]

258. Hopfield JJ. Kinetic proofreading - new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc Nat Acad Sci USA. Jan; 1974 71(10):4135–4139. [PubMed: 4530290]

259. Rodnina MV, Wintermeyer W. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. Annu Rev Biochem. Jan.2001 70:415–35. [PubMed: 11395413]

260. Blanchard SC, Gonzalez RL, Kim HD, Chu S, Puglisi JD. tRNA selection and kinetic proofreading in translation. Nat Struct Mol Biol. Oct; 2004 11(10):1008–14. [PubMed: 15448679]

261. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy FV, Weir JR, Ramakrishnan V. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science. Oct; 2009 326(5953):688–94. [PubMed: 19833920]

262. Geggier P, Dave R, Feldman Mb, Terry Ds, Altman Rb, Munro JB, Blanchard SC. Conformational sampling of aminoacyl-tRNA during selection on the bacterial ribosome. J Mol Biol. Apr.2010

263. Frank J, Sengupta J, Gao H, Li W, Valle M, Zavialov A, Ehrenberg M. The role of tRNA as a molecular spring in decoding, accommodation, and peptidyl transfer. FEBS Lett. Feb; 2005 579(4):959–62. [PubMed: 15680982]

264. Johansson M, Lovmar M, Ehrenberg M. Rate and accuracy of bacterial protein synthesis revisited. Curr Op Microbio. 2008; 11(2):141–147.

265. Wohlgemuth I, Pohl C, Mittelstaet J, Konevega Al, Rodnina MV. Evolutionary optimization of speed and accuracy of decoding on the ribosome. Phil Trans Royal Soc B: Biol Sci. Oct; 2011 366(1580):2979–86.

266. Frank J, Spahn CMT. The ribosome and the mechanism of protein synthesis. Reports On Progress In Physics. Jan; 2006 69(5):1383–1417.

267. Sanbonmatsu KY, Joseph S, Tung CS. Simulating movement of tRNA into the ribosome during decoding. Proc Nat Acad Sci USA. Nov; 2005 102(44):15854–9. [PubMed: 16249344]

268. Whitford PC, Geggier P, Altman RB, Blanchard SC, Onuchic JN, Sanbonmatsu KY. Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. RNA. Apr.2010 16:1196–1204. [PubMed: 20427512]

269. Byrne RT, Konevega AL, Rodnina MV, Antson AA. The crystal structure of unmodified tRNA-Phe from Escherichia coli. Nucleic Acids Res. Jul; 2010 38(12):4154–62. [PubMed: 20203084]

270. Shi H, Moore PB. The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. RNA. Aug; 2000 6(8):1091–105. [PubMed: 10943889]

271. Fischer N, Konevega AL, Wintermeyer W, Rodnina MV, Stark H. Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. Nature. Jul; 2010 466(7304):329–33. [PubMed: 20631791]

272. Chladek S, Sprinzl M. The 3′-end of transfer-RNA and its role in protein-biosynthesis. Angew Chem Int Edit. Dec; 1985 24(5):371–391.

273. Rebecca WA, Eargle J, Luthey-Schulten Z. Experimental and computational determination of tRNA dynamics. FEBS Lett. Jan; 2010 584(2):376–86. [PubMed: 19932098]

274. Eargle J, Black AA, Sethi A, Trabuco LG, Luthey-Schulten Z. Dynamics of recognition between tRNA and elongation factor Tu. J Mol Biol. Apr; 2008 377(5):1382–405. [PubMed: 18336835]

275. Bashan A, Agmon I, Zarivach R, Schluenzen F, Harms J, Berisio R, Bartels H, Franceschi F, Auerbach T, Hansen HAS, Kossoy E, Kessler M, Yonath A. Structural basis of the ribosomal machinery for peptide bond formation, translocation, and nascent chain progression. Molecular Cell. Jan; 2003 11(1):91–102. [PubMed: 12535524]

276. Sanbonmatsu KY. Alignment/misalignment hypothesis for tRNA selection by the ribosome. Biochimie. Aug; 2006 88(8):1075–89. [PubMed: 16890341]

277. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. Feb; 1996 14(1):33–38. [PubMed: 8744570]

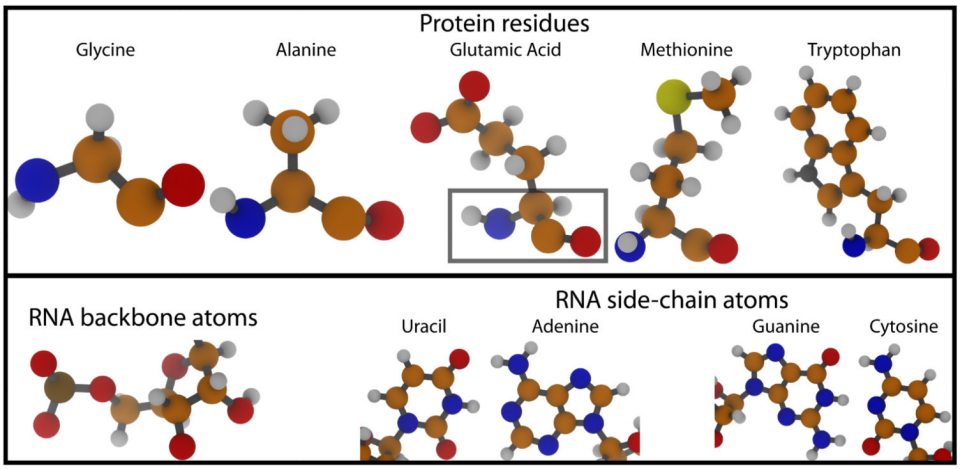

Figure 1. The diversity of protein and RNA residues (top)

Protein chains are composed of amino acid residues, which contain carbon (orange spheres), nitrogen (blue), oxygen (red), sulfur (yellow) and hydrogen (white) atoms. Each amino acid has a common set of C, N, O and H atoms (boxed). The range of chemical compositions of residues is depicted by the atomic structures of 5 of the 20 naturally-occurring) different amino acids. (bottom) RNA chains are composed of 4 type of nucleic acids. Each nucleic acid has a common set of backbone atoms (left). Canonical Watson-Crick base pairs are formed between A-U and C-G pairs (right). Structural figures were prepared with VMD (277).

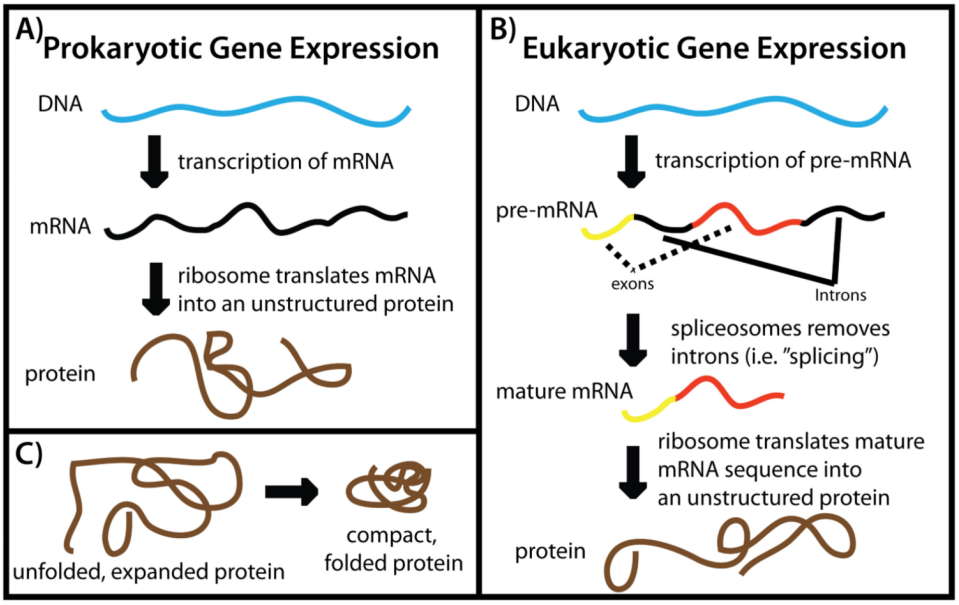

Figure 2. Gene expression in the cell

The two fundamental steps of gene expression are transcription and translation. A) In prokaryotes, messenger RNA (mRNA) is transcribed and then read by ribosomes to produce proteins. B) In eukaryotes, precursor-mRNA is produced during transcription. Pre-mRNA is modified and transported in the cell and then mature mRNA is read by ribosomes. There are many types of post-transcriptional modifications that may be performed on pre-mRNA, though only splicing is depicted. During splicing, specific sequence of mRNA (introns) are removed from the pre-mRNA. C) Once a protein is synthesized by the ribosome, it folds to the lowest-energy "native" ensemble.

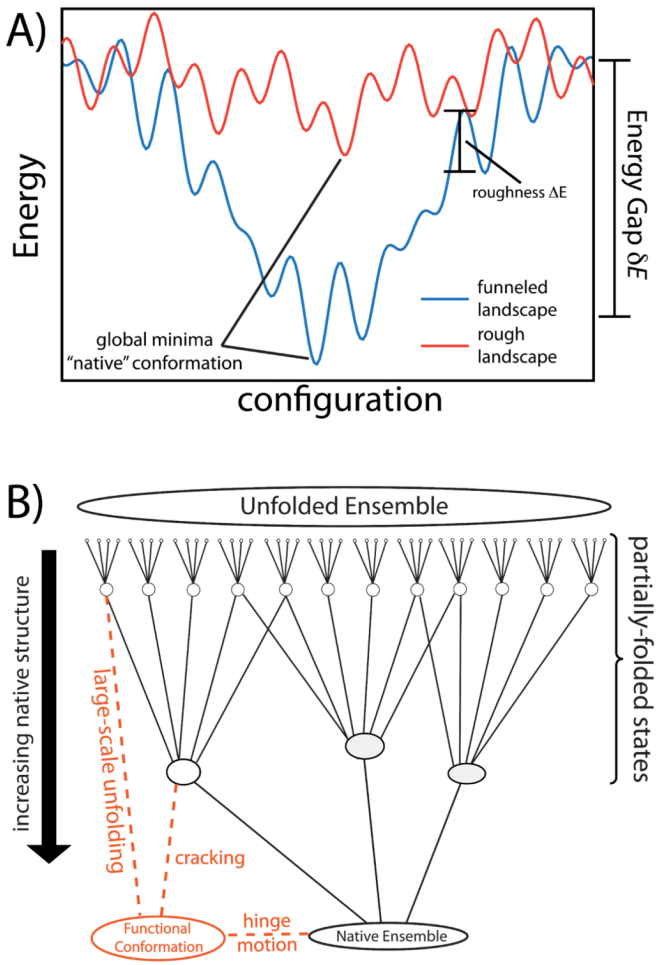

Figure 3. Energy landscapes of biomolecular folding

A) When the energy landscape of a system is rough red), the system must search randomly for the lowest-energy configuration and glass-like dynamics may arise. Biomolecules have evolved to have a large energy gap between the unfolded and folded ensembles δE, relative to the energetic roughness ΔE. "Funneled" energy landscapes (blue) enable rapid folding of biomolecules *in vivo*. B) An alternate representation of a funneled landscape that is extended to account for functional rearrangements. Functional biomolecules often possess multiple basins of attraction and there are many mechanisms by which functional transitions may occur. Cracking, hinge-motion and large-scale unfolding orange) encompass different degrees of molecular disorder and they result in differential forms of connectivity between the functional and native ensembles.

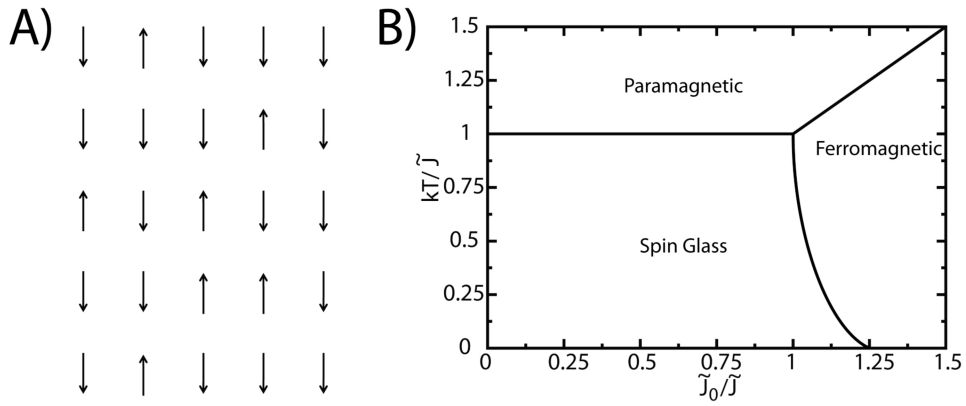

Figure 4. Spin systems and glass transitions

A) A two-dimensional Ising spin model. B) Schematic of the phase diagram for the SK spin-glass model. In the SK spin-glass, three states can exist: paramagnetic, ferromagnetic, and glass. The balance of energetic roughness ($\tilde{J}$), ferromagnetic coupling ($\tilde{J}_0$) and temperature T determines which state the system adopts.

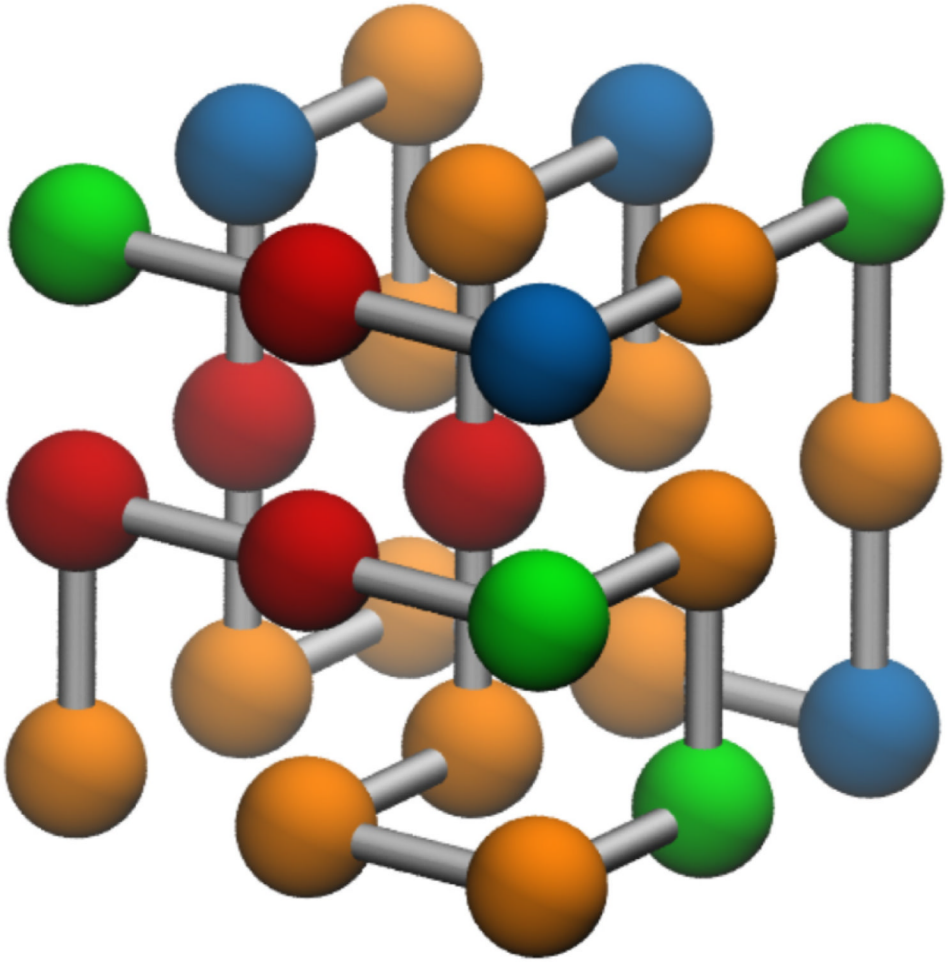

Figure 5. Lattice model for protein dynamics

In lattice models of proteins, the residues (spheres) move between adjacent lattice positions. Similar to proteins in solution, backbone connectivity (grey) leads to sequential residues being adjacent in space. In this model, the degree of hydrophobicity is indicated by the sphere colors. The sequence shown is taken from Leopold et al., and it represents a maximally-compact configuration (19).

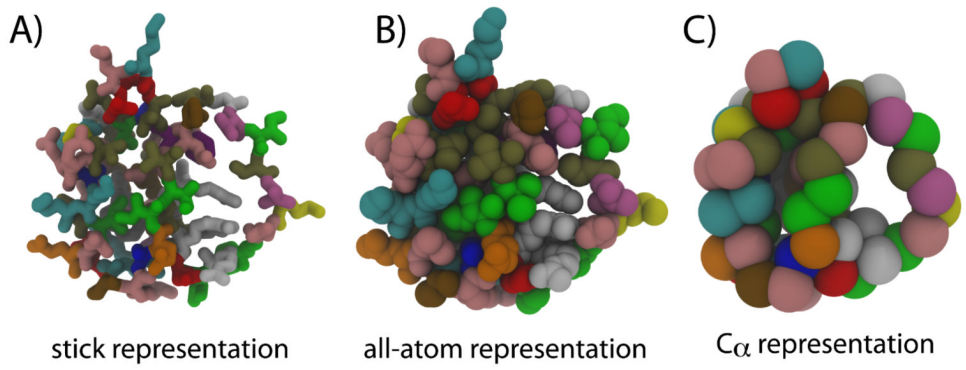

Figure 6. Multiple representations of proteins

The protein Chymotrypsin Inhibitor 2 (CI2) pdb entry: 2ci2) shown in A) stick representation, B) all-atom representation excluding hydrogens) that depicts the excluded volume of the atoms, and C) C_a representation that is often used in molecule simulations. Each residue type is give a unique color, highlighting the heterogeneity of protein sequences.

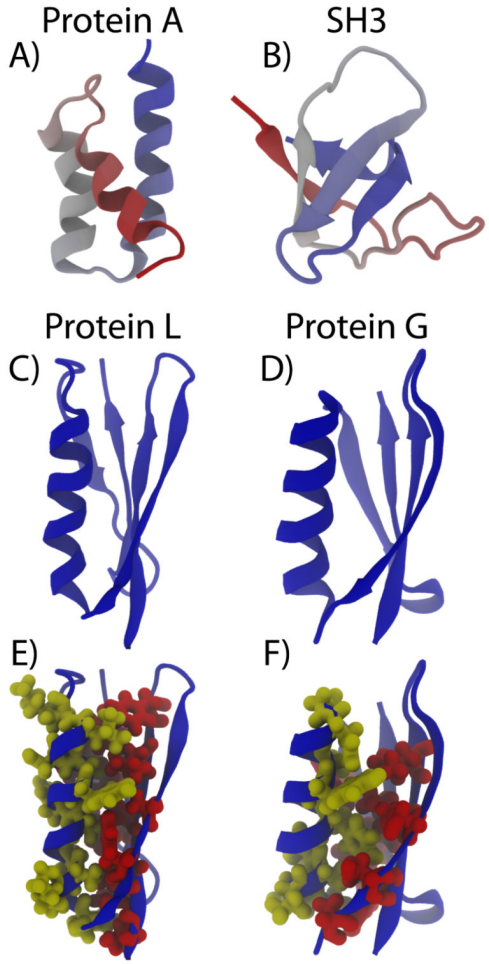

Figure 7. Protein structures vary in complexity

A) Protein A pdb: 1BDD) is composed of three α-helices and is considered a simple fold. B) The SH3 domain of Src (pdb: 1FMK) is composed of two adjacent β-hairpins, which leads to more-complex folding dynamics. Protein L pdb: 2ptl) (panel C) and G (panel D) (pdb: 3gp1) share a common overall fold, but they exhibit different folding dynamics in solution. Inspection of the packing of residues in the core of the protein (E, F) shows that the fine details of these interactions are different. Side chains are shown explicitly for residues that are in contact between the helix and the sheet. These packing effects lead to unique folding dynamics of the proteins (122).

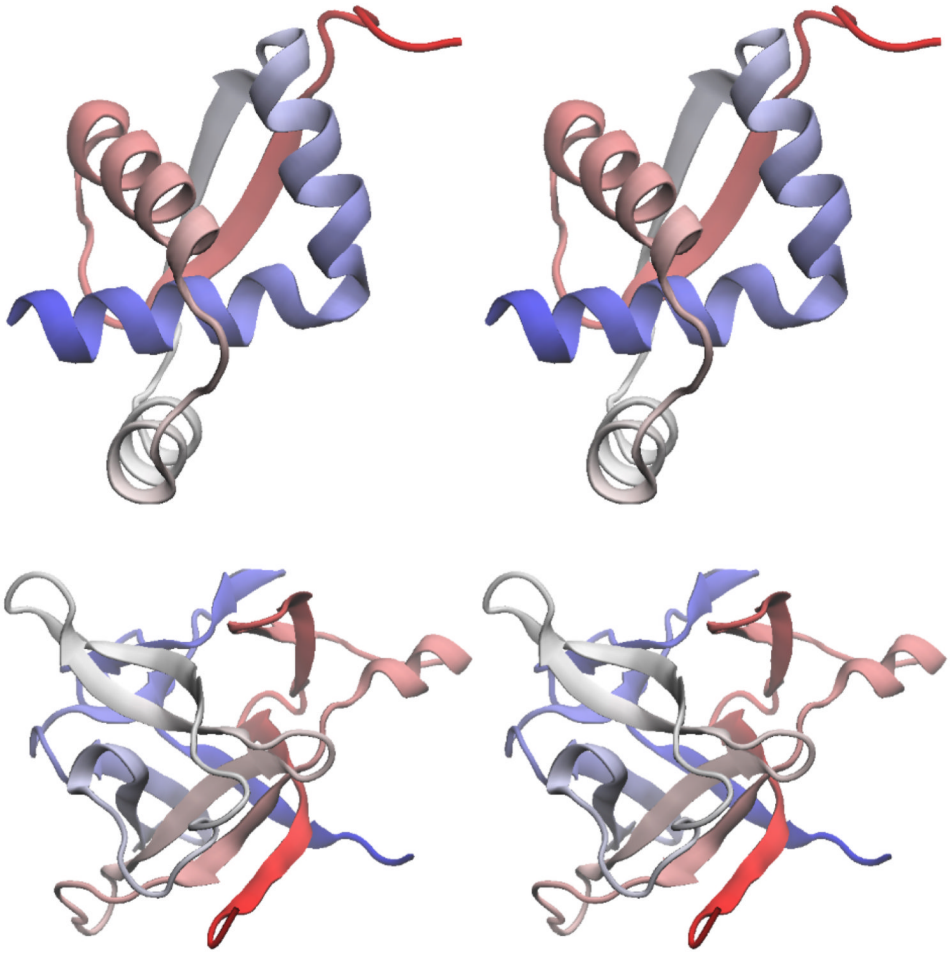

Figure 8. Topologically frustrated structures

Two classes of proteins that are considered to be of “complex” folds are knotted proteins (top, stereo view. PDB: 2EFV) and the Interleukin family of proteins (bottom, stereo view. PDB: 6I1B). In the knotted protein, folding involves the formation of a loop and then threading of the terminal end through the loop. If the loop closes prematurely, then the threading process will be inhibited by excluded volume effects. In the Interleukin family, there is structural competition between regions during folding. If the C-terminal trefoil folds early, the protein “backtracks” and partially unfolds before full folding occurs.

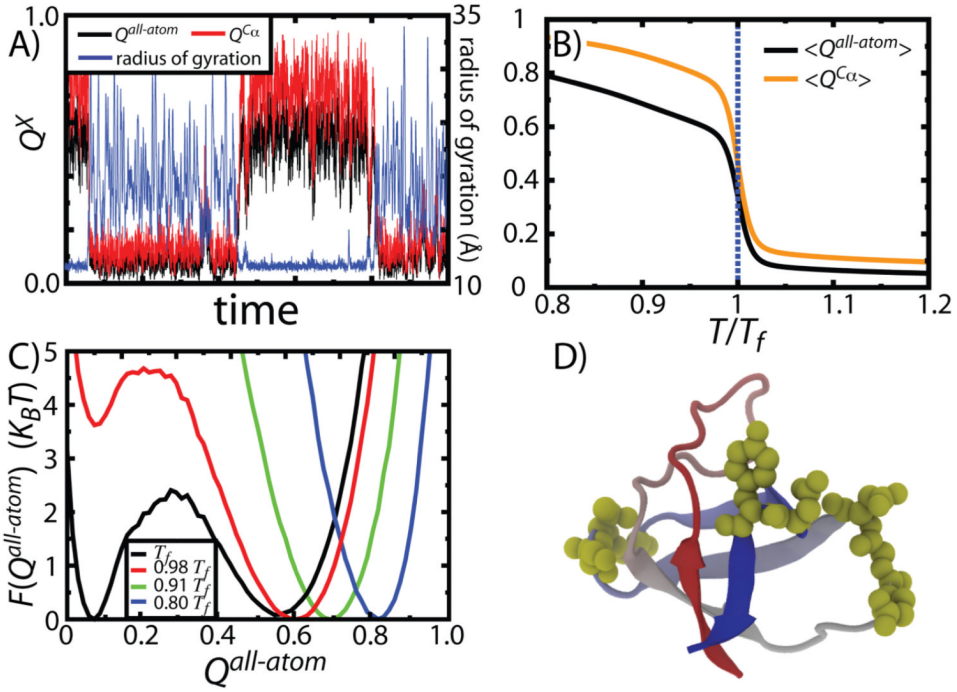

Figure 9. Side-chain ordering and backbone collapse during protein folding

A) A simulated folding trajectory of CI2, shown as a function of the radius of gyration R_g (blue), fraction of native all-atom contacts formed $Q^{all\text{-}atom}$ (black) and fraction of native C_α contacts formed $Q^{C\alpha}$. R_g measures backbone collapse, which decreases when $Q^{all\text{-}atom}$ and $Q^{C\alpha}$ increase. $Q^{C\alpha}$ reaches larger values than $Q^{all\text{-}atom}$ indicating that the backbone is ordered, while the side chains are not in their native configurations. B) $\langle Q^{all\text{-}atom} \rangle$ and $\langle Q^{C\alpha} \rangle$ as function of T indicate that both coordinates capture the global folding transition. C) As temperature is reduced below T_f, there is continuous movement of the native basin to higher values of $Q^{all\text{-}atom}$. D) Residues in SH3 that have a larger fraction of C_α contacts formed than all-atom contacts are depicted by yellow spheres.

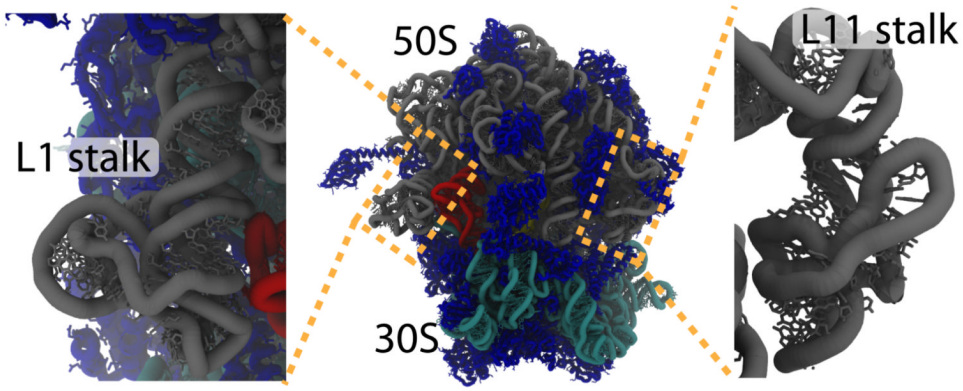

Figure 10. The Ribosome Center)

Atomic structure of an in-tact 70S ribosome (~150,000 non-hydrogen atoms). Transfer (RNA tRNA, P-site tRNA shown in red) molecules read messenger RNA (mRNA) through interactions on the “small” (30S) subunit and add the incoming amino acid to the growing protein chain located in the “large” (50S) subunit. An additional tRNA i.e. the A-site tRNA) is buried deep inside of the ribosome. During protein synthesis, the 50S “stalks” assist movement of tRNA molecules. The L11 stalk (right) aids the entry of incoming tRNA molecules and the L1 stalk (left) facilitates the disassociation of tRNA molecules. NMA, simulations with SBMs and simulations with explicit solvent suggest that the stalks are very flexible (i.e. have low-energy displacements), relative to other components of the ribosome.

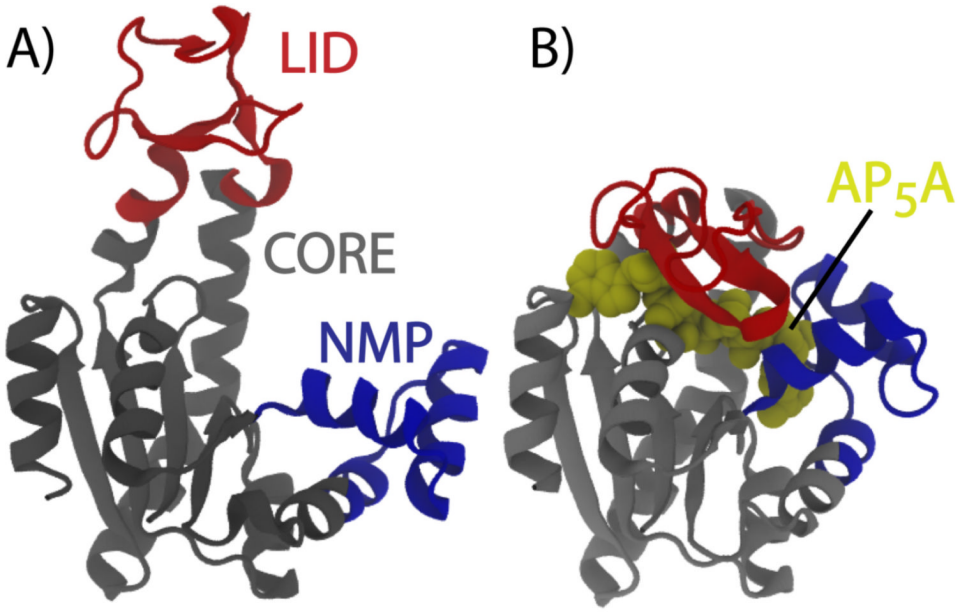

Figure 11. Functional configurations of Adenylate Kinase

Adk is composed of three domains: LID (red), NMP (blue) and core (grey). A) When a ligand is not present, the dominant configuration has the domains arranged in a more extended fashion. B) Upon ligand binding, the molecule adopts a compact conformation, where the ligand is isolated and chemistry is permitted. AP_5A is a bi-substrate analogue that mimics the natural substrates of Adk: ATP and AMP.

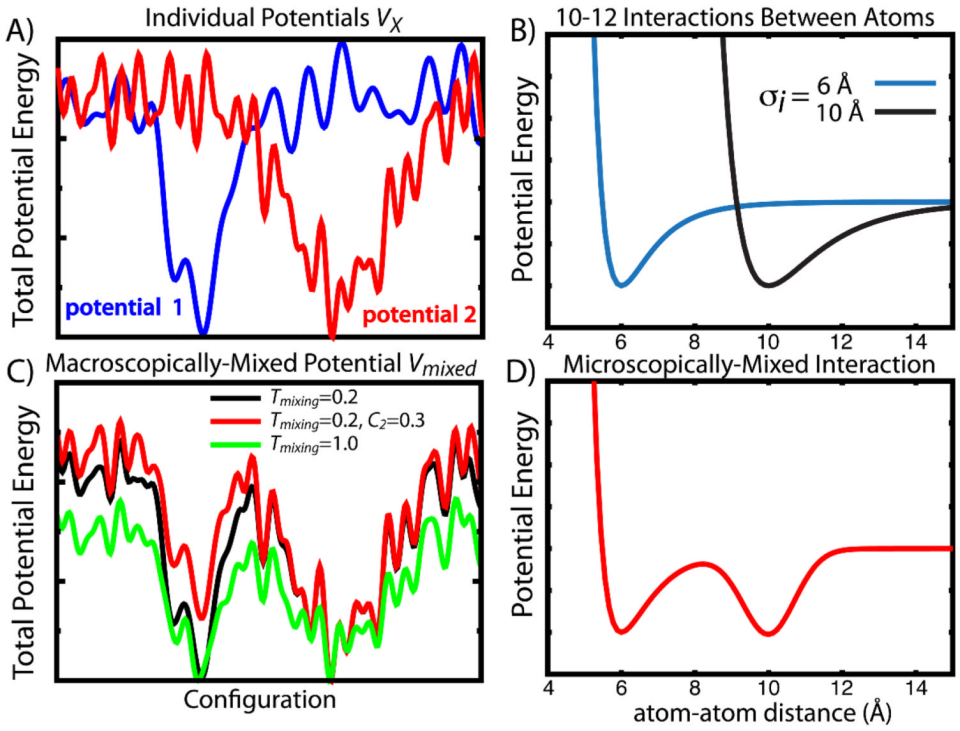

Figure 12. Macroscopic and Microscopic Mixing

When combing two potential energy functions, one may combine them term-by-term, or in a global fashion. A) Two hypothetical potential energy surfaces, where each curve depicts the total energy in each potential, given in terms of an arbitrary coordinate. B) Two 10-12 interactions that are defined for the same atom-atom pair. This depiction represents a pair that is 6Å in one configuration and 10Å in the second. C) A macroscopically-mixed (Boltzmann) Hamiltonian, using the two potentials in (A) as input. Different values of the mixing temperature T_{mixing} and weighting factor C_i are shown. D) A microscopically-mixed atom-atom potential energy function that includes minima at both 6 and 10 Å. This representation uses a Gaussian function for the minimum at 10 Å.

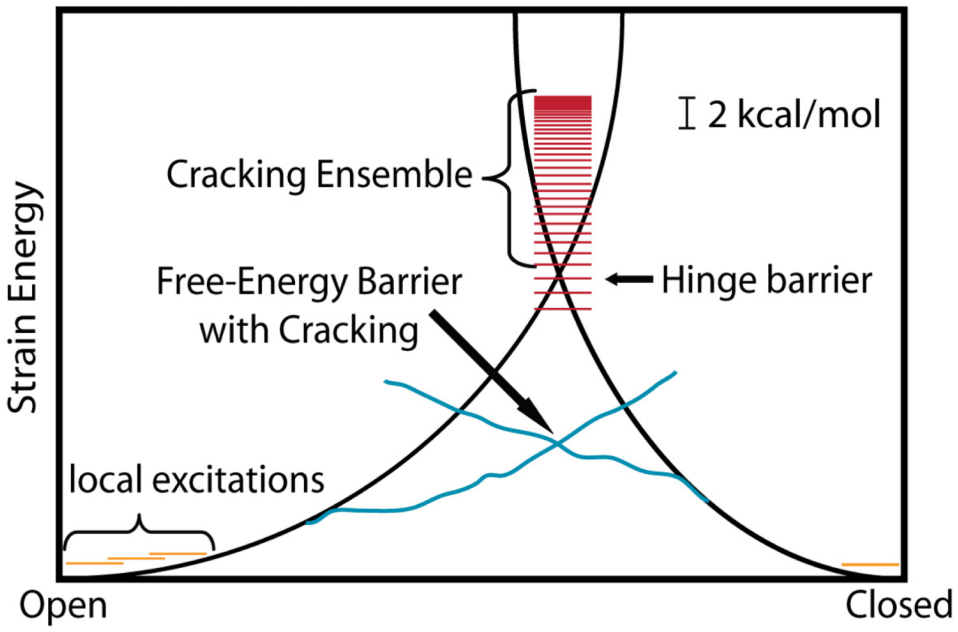

Figure 13. Cracking during functional rearrangements

Iterative normal mode analysis suggests large levels of strain energy accumulate (black) during conformational rearrangements. To relieve strain that is localized to specific residues, biomolecules may crack, or partially unfold and refold. By unfolding, a large number of unfolded configurations are accessible (red), which increases the entropy of the transition state ensemble and reduces the free-energy barrier associated with the transition (blue). X-ray crystallographic data and NMR measurements (27) probe the fluctuations local to a particular energetic basin (yellow lines), and the local excitations can have dynamics that are correlated with large-scale rearrangements (134).

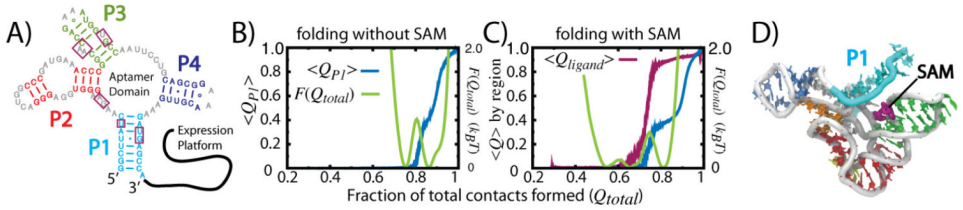

Figure 14. Riboswitch folding and ligand recognition

A) Secondary structure of the SAM-1 riboswitch. The P1 helix is formed by the terminal residues, making it a non-local helix (125). B) Average fraction of native P1 contacts formed ($\langle Q_{P1} \rangle$) as a function of the fraction of all native contacts being formed (Q_{total}) is shown in blue. Potential of mean force shown in green. The largest barrier in the pmf corresponds to folding of the P1 helix. C) Same as (B), but in the presence of the SAM ligand. The average fraction of ligand interactions $\langle Q_{ligand} \rangle$ is shown in magenta. Ligand binding induces earlier folding of P1 and it reduces the associated barrier. D) Structural representation of the final step of folding, when the SAM ligandis present. All secondary structure, along with some tertiary structure is formed, and the ligand binds prior to P1 structure formation.

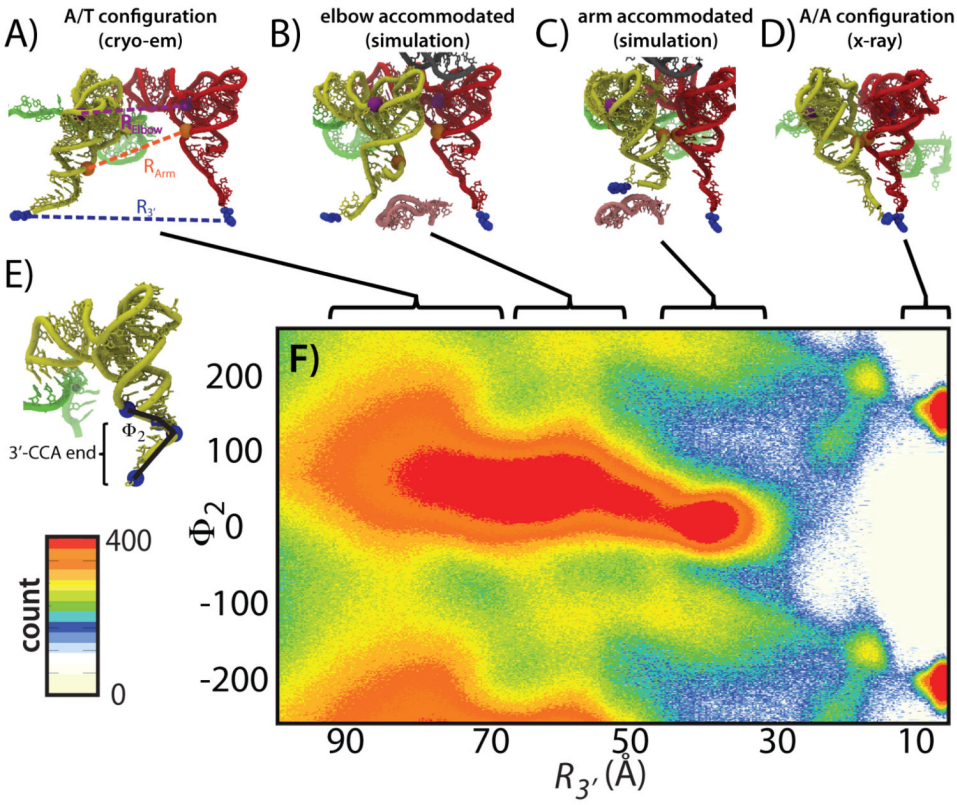

Figure 15. Entropy changes in the 3′-CCA end of aa-tRNA during accommodation

Atomic models of tRNA (A-site tRNA in yellow, P-site tRNA in red) and mRNA (green) at different points during accommodation. Amino acids attached to the ends of tRNA are shown in blue. A) Atomic model of the A/A configuration of tRNA, based on cryo-EM densities. Three reaction coordinates for describing accommodation (R_{Elbow}, R_{Arm}, $R_{3'}$) are indicated. B/C) Intermediate populations predicted from simulations (268) suggest that transient interactions are formed with the A loop (pink) and H89 (gray) of the large subunit. D) Endpoint of accommodation, as determined from x-ray crystallography. E) Pseudo-dihedral angle Φ_2 describes the configuration of the 3′-CCA end as it enters the center of the ribosome. F) During accommodation, the range of accessible configurations of the 3′-CCA end decreases, indicative of an entropic penalty for accommodation.