Published in final edited form as:

Methods Enzymol. 2011 ; 492: 253–282. doi:10.1016/B978-0-12-381268-1.00020-3.

Strategies for the Thermodynamic Characterization of Linked Binding/Local Folding Reactions Within the Native State: Application to the LID Domain of Adenylate Kinase from *Escherichia coli*

Travis P. Schrank^{*}, W. Austin Elam[†], Jing Li[†], and Vincent J. Hilser^{†,‡}

^{*}Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, Texas, USA

[†]T.C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, Maryland, USA

[‡]Department of Biology, Johns Hopkins University, Baltimore, Maryland, USA

Abstract

Conformational fluctuations in proteins have emerged as an important aspect of biological function, having been linked to processes ranging from molecular recognition and catalysis to allostery and signal transduction. In spite of the realization of their importance, however, the connections between fluctuations and function have largely been empirical, even when they have been quantitative. Part of the problem in understanding the role of fluctuations in function is the fact that the mere existence of fluctuations complicates the interpretation of classic mutagenesis approaches. Namely, mutagenesis, which is typically targeted to an internal position (to elicit an effect), will change the fluctuations as well as the structure of the native state. Decoupling these effects is essential to an unambiguous understanding of the role of fluctuations in function. Here, we use a mutation strategy that targets surface-exposed sites in flexible parts of the molecule for mutation to glycine. Such mutations leave the ground-state structure unaffected. As a result, we can assess the nature of the fluctuations, develop a quantitative model relating fluctuations to function (in this case, molecular recognition), and unambiguously resolve the probabilities of the fluctuating states. We show that when this approach is applied to *Escherichia coli* adenylate kinase (AK), unique thermodynamic and structural insights are obtained, even when classic mutagenesis approaches targeted to the same region yield ambiguous results.

1. Introduction

It is now widely accepted that the native state of proteins is a dynamic and heterogeneous collection of conformations (i.e., an ensemble) that surround the average structure of the molecule, usually typified by the crystallographic structure. Dynamic excursions of proteins to excited (minor) states have been associated with biochemically important processes such as protein folding, catalysis, and binding (Korzhnev et al., 2006; Manson et al., 2009; Wolf-Watz et al., 2004). However, relatively few studies give insight into the physical/conformational character of the states that make up these functionally important dynamics (Vallurupalli et al., 2007). One particularly important experimental approach that has emerged over the past two decades has been the use of hydrogen–deuterium exchange to

investigate exchange-competent states in the native state ensembles of proteins (Bai, 2006; Krishna et al., 2004; Rogero et al., 1986). Although several models for exchange have been forwarded (Miller and Dill, 1995), with the precise origin of exchange for all positions in the protein unresolved (LeMaster et al., 2009), denaturant-dependent exchange experiments strongly suggest that exchange occurs as a result of local unfolding reactions that expose buried or hydrogen bonded amides to solvent (Englander, 1975).

Consistent with this view, a computational model of native state fluctuations that is based on local unfolding (i.e., the COREX algorithm) has also had success reconciling many other aspects of protein behavior, including allosteric effects of ligand binding, noncooperative protein unfolding at low temperatures, and the interaction of the effects of pH and denaturants in determining the apparent m -value of unfolding (Babu et al., 2004; Liu et al., 2007; Pan et al., 2000; Whitten et al., 2001). These results, taken in total, suggest that local unfolding to highly disordered states may be a ubiquitous property of the native state ensembles of most, if not all, proteins. Nevertheless, the agreement between COREX and experiment, although highly suggestive, cannot be taken as unambiguous proof of local unfolding. Indeed, there can be numerous models that are consistent with the data (Miller and Dill, 1995). To establish that local unfolding is present in the ensembles of proteins and plays a functionally relevant role, experimental strategies are needed that will directly challenge local unfolding relative to other types of conformational excursions.

The low probability of excited states is a primary difficulty in the experimental investigation of native state fluctuations (Baldwin and Kay, 2009). Many spectroscopic techniques (such as fluorescence or circular dichroism) commonly used to investigate proteins report average properties of the molecule in solution, and therefore cannot easily detect minor populations. Significant progress has been made toward this end through the development of modern NMR relaxation–dispersion methods, although these methods too have a lower population limit of detectability (~0.5%; Baldwin and Kay, 2009). Given the experimental difficulties in detecting states with low probability, it would be useful to develop strategies that increase the population of such states, but doing so in a way that does not complicate the interpretation of experimental results.

To address these challenges, we have implemented a strategy that specifically targets mutational effects so that they are manifested as changes in the fluctuations around the ground-state structure and not as changes to the ground-state structure itself. Specifically, we utilize an entropy-enhancing mutation strategy that promotes conformational excursions to locally unfolded states, and we monitor the effects of the mutations to the thermodynamic properties of measurable binding reactions. Numerous studies have demonstrated that positions on the periphery of the binding sites are often among the most dynamic and unstable regions within the protein (Ferreon and Hilser, 2003a; Henzler-Wildman et al., 2007; Rundqvist et al., 2009). Not surprisingly, the dynamics of these unstable regions is also often found to be modulated by binding (Boehr et al., 2010; Ferreon and Hilser, 2003a). Because the approach we describe below specifically addresses the coupling between the fluctuations and binding, it provides a potential tool for probing the coupling between binding and fluctuations in many protein systems. Indeed, our laboratory has previously applied this strategy to successfully study the binding and fluctuations of the RT-loop of the

SEM-5 SH3 domain, as well as the LID domain of adenylate kinase (AK) from *Escherichia coli* (Ferreon et al., 2004; Manson et al., 2009; Schrank et al., 2009).

2. A Mutation Strategy to Amplify Locally Unfolded States

Although local unfolding appears to be a ubiquitous property of folded proteins, the probabilities of individual locally unfolded states are, in general, relatively small. As discussed above, mutational strategies that will selectively target such states are of great experimental utility. However, classic mutagenesis studies are targeted toward disruption of protein structure, a strategy that carries the risk of spurious changes that can greatly compromise the interpretation of results, especially when the point of the study is to clarify the structure and energetics of fluctuations. This point is highlighted in Fig. 9.1, which shows how typical mutation studies change the ground-state structure of a protein. In the example, a mutation which removes an interior group in the protein necessarily changes the ground-state structure and energy relative to the wild type (Fig. 9.1A). Such mutation may also change the number and energy of excited states, but deconvolution of the relative effects is clearly not straight-forward. An alternative ideal case would involve making mutations that do not change the ground-state structure of the molecule, but instead change the probability of fluctuations (Fig. 9.1B). Such an approach would greatly simplify the interpretation of results and would allow quantitative access to the energetics associated with the fluctuations. Although this ideal experiment is not possible, we utilize an approach that most closely approximates these conditions—the introduction of conformational entropy-enhancing mutations through substitution of surface-exposed residues with glycine (Gly).

To see how such a mutation strategy achieves this objective, we consider the following. In unfolded states, proteins and peptides have multiple conformational degrees of freedom due to the presence of the numerous rotatable bonds. For the majority of amino acids, which contain a β -carbon, the available conformational space is somewhat limited relative to Gly, which contains only an hydrogen (H). This limited conformational space arises because many (in fact most) conformations accessible to Gly are sterically occluded when the H of Gly is substituted with a more bulky side chain (D'Aquino et al., 1996). This principle is demonstrated in Fig. 9.2. A simulation was performed by sampling a uniform statistical distribution of ϕ and ψ angles for all residues in an AAAAA or AAGAA peptide, and constructing an ensemble of peptide conformations based on these angles and other well-known geometric constraints of polypeptides (Manson et al., 2009; Whitten et al., 2008). All conformers producing steric collisions were then discarded. The ϕ and ψ angles of the central residue in the sterically allowed space are displayed as a histogram in Fig. 9.2. Large differences between Ala and Gly can be appreciated, as previously demonstrated from the classical Ramachandran map (Ramakrishnan and Ramachandran, 1965). Using this simulation, the ratio of available conformational space for the Ala or Gly containing peptide can be calculated as (Leach et al., 1966),

$$\Omega_{\text{Ala-Gly}} = \frac{n_{\text{allowed,Ala}}/n_{\text{test,Ala}}}{n_{\text{allowed,Gly}}/n_{\text{test,Gly}}}, \quad (9.1)$$

where n_{allowed} is the number of conformers passing the steric check, and n_{test} is the total number of randomly sampled conformations. Accordingly, the expected change in entropy of an unfolded state affected by the mutation can then be estimated by using the Boltzmann entropy equation (Leach et al., 1966),

$$\Delta\Delta S_{\text{Ala-Gly}} = -R \ln(\Omega_{\text{Ala-Gly}}). \quad (9.2)$$

According to this simulation, the entropy difference of unfolding an Ala versus a Gly residue is 2.2 cal/mol K (Table 9.1), similar to values measured from calorimetric protein unfolding experiments using Ala/Gly mutant pairs (2.3–2.7 cal/mol; D'Aquino et al., 1996) and other computational methods (D'Aquino et al., 1996; Leach et al., 1966). Therefore, based on the constraints imposed by molecular geometry, Ala to Gly mutations should invariably decrease the stability through enhancement of the entropy of the unfolded state. Indeed, several authors have used this robust feature of protein folding as a tool for protein destabilization (Huyghues-Despointes et al., 1999; Maity et al., 2003).

Thus, the mutation strategy is expected to promote conformational excursions to highly dissimilar, that is disordered conformational states. To illustrate this point, the above simulation was again used to test the relative availability of conformational space within a given window of ϕ/ψ space. Interestingly, such a calculation reveals that the density of available (steric) conformations is similar for an Ala or Gly residing in a ϕ/ψ space accessible to both amino acid types (e.g., *Case-1*; Fig. 9.2; Table 9.1). If the conformation of a mutated position is accessible to both Ala and Gly, the entropy of the folded ground state (a range of available $\phi-\psi$ space compatible with the compact fold) should be minimally affected (see *Case-1*; Table 9.1). A mutation of a residue in a β -sheet (see position of Valine 142, Fig. 9.2), for instance, is expected to provide a relatively isolated entropic stabilization to states that can access distant regions of ϕ/ψ space (in this case, greater than 90 ° in rotation around one or both dihedral angles). Others have noted that the entropy of an Ala or Gly confined to an α -helix should be similar in conformational entropy (D'Aquino et al., 1996). However, mutation of residues that reside near an overlapping boundary of the available Ala and Gly spaces (*Case-2*, Fig. 9.2, Table 9.1) might allow access to new states similar in conformation to the folded (i.e., crystallographic) state. The magnitude of this effect in terms of entropy will be highly dependent on the conformational envelope of the state in question. In short, considerable data support the notion that Ala (or Val) to Gly mutations will robustly effect an increase in the conformational entropy of unfolded or highly disordered states, thus destabilizing the conformationally constrained folded state. Since the relevant changes are intrinsic to the unfolded state itself, the mutational effects should be realized for any mutation site, independent of structural context.

Based on such considerations, we have designed (Val to Gly) mutations to promote disordered states in highly dynamic structures, in an effort to increase the population of highly disordered or unfolded states involving these structures. To ensure that the mutations closely resembled the ideal strategy described in Fig. 9.1, the selected Val residues were surface exposed, with the side chain involved in few to no molecular contacts. This strategy

allows for destabilization without perturbation to the conformation of the folded state, a condition that minimizes undesirable/confounding changes in structure.

To illustrate the extensive thermodynamic information that can be obtained from such an approach, we will summarize the results of a recent ITC analysis of three mutations of AK from *E. coli* (v142g, v135g, v148g). The highly exposed location of the mutation sites, as well as the structural relationship to the binding cleft, is depicted in Fig. 9.3. The positions of the mutated residues in ϕ/ψ space are also indicated in Fig. 9.2. As a counter example, we also present data gathered for a mutation of the same region that is designed to destabilize the *folded* conformation by eliminating a native state hydrogen bond (s129a), a common strategy used to probe mutational effects.

3. Thermodynamic Properties of Linked Folding and Binding Reactions

Thermodynamically linked folding (local or global) and binding has been found to be a property of many proteins (Cliff et al., 2005; Wright and Dyson, 2009). In such cases, ITC binding experiments become a powerful tool to investigate the folding reaction. The simplest conformational model that can be used to describe any linked folding and binding reaction, is the model where the (locally) unfolded state has little or no affinity for the ligand, or is *binding incompetent* (BI), and is in equilibrium with a folded, high affinity or *binding competent* (BC) state. Eftink et al. term this model a “mandatory coupling model,” to emphasize that binding occurs with only one of the two possible states (Eftink et al., 1983). In this case, the addition of ligand will promote the population of the BC state through mass action. The impact of such a model on direct (ΔH and ΔG) and indirect (ΔC_p) calorimetric observables of binding has been published (Eftink et al., 1983).

According to this model, the apparent (measured) binding constant will have the form

$$K_{\text{app}}(T) = \frac{[\text{BCX}]}{[\text{BC} + \text{BI}][\text{X}]} \quad (9.3)$$

Since the equilibrium between the BC and BI states can be written as $K_{\text{conf}}(T) = [\text{BI}]/[\text{BC}]$, Eq. (9.3) reduces to

$$K_{\text{app}}(T) = \frac{K_0(T)}{(1 + K_{\text{conf}}(T))}, \quad (9.4)$$

where $K_0(T)$ is the intrinsic association constant between the BC state and the ligand.

$$K_0(T) = \frac{[\text{BCX}]}{[\text{BC}][\text{X}]} \quad (9.5)$$

Considering Eqs. (9.3)–(9.5), it is clear that an increase in the probability of the BI should decrease the measured binding affinity of the system. The free energy of binding will consist of two terms,

$$\Delta G_{\text{app}}(T) = -RT \ln K_0(T) + RT \ln(1 + K_{\text{conf}}(T)), \quad (9.6)$$

$$\Delta G_{\text{app}}(T) = \Delta G_0(T) - \Delta G_{\text{conf,app}}(T), \quad (9.7)$$

where the first term is the intrinsic free energy of interaction between the BC state and the ligand, and the second term is the apparent conformational free energy contributed by the equilibrium between the BI and BC states (Eftink et al., 1983). We note that in Eq. (9.6), the term

$$K_{\text{conf}}(T) = \exp(\Delta G_{\text{conf}}/RT), \quad (9.8)$$

where

$$\Delta G_{\text{conf}}(T) = \Delta H_{\text{conf}} \left(\frac{T_{\text{m,conf}} - T}{T_{\text{m,conf}}} \right) - (T_{\text{m,conf}} - T) \Delta C_{\text{p,conf}} + T \Delta C_{\text{p,conf}} \ln \left(\frac{T_{\text{m,conf}}}{T} \right), \quad (9.9)$$

and H_{conf} , $T_{\text{m,conf}}$, and $C_{\text{p,conf}}$ are the enthalpy, transition midpoint temperature, and heat capacity difference between the BI and BC states.

The apparent enthalpy for the binding process can be obtained through the derivative of Eq. (9.4) with respect to $1/T$,

$$\Delta H_{\text{app}}(T) = -R \frac{d \ln K_{\text{app}}(T)}{d(1/T)}, \quad (9.10)$$

to yield

$$\Delta H_{\text{app}}(T) = \Delta H_0(T) + \frac{K_{\text{conf}}}{(1 + K_{\text{conf}})} \Delta H_{\text{conf}}(T) \quad (9.11)$$

where, the first term corresponds to the enthalpy of the intrinsic association process and the second term corresponds to the enthalpy difference between the BI and BC state, weighted according to the population of molecules in the BI state.

$$\Delta H_{\text{app}}(T) = \Delta H_0(T) + P_{\text{BI}} \Delta H_{\text{conf}}(T) \quad (9.12)$$

$$\Delta H_{\text{app}}(T) = \Delta H_0(T) + \Delta H_{\text{conf,app}}(T) \quad (9.13)$$

Similarly, two terms are expected to contribute to the apparent C_p of binding, one corresponding to the intrinsic and one to the conformational reaction. Formally,

$$\Delta C_{p,\text{app}} = \Delta C_{p,\text{int}} + \Delta C_{p,\text{conf,app}}, \quad (9.14)$$

where $C_{p,\text{int}}$ is the change in heat capacity for the BC/ligand association reaction, and $C_{p,\text{conf,eff}}$ is the equilibrium dependent contribution to the observed heat capacity. This conformational term can be written as (Eftink et al., 1983; Prabhu and Sharp, 2005),

$$\Delta C_{p,\text{conf,eff}} = \Delta C_{p,\text{conf}} P_{\text{BI}} + \frac{\Delta H_{\text{conf}}^2}{RT^2} P_{\text{BI}} P_{\text{BC}}, \quad (9.15)$$

as the sum of the probability weighted change in heat capacity associated with the conformational reaction itself ($C_{p,\text{conf}}$), as well as the mean squared fluctuation in the enthalpy (H_{conf}) of the available conformational states.

Using the above expressions (Eftink et al., 1983), it is possible to simulate the expected effect of a linked folding reaction on the experimental observables provided by ITC (Fig. 9.4A). ITC data measuring the apparent free energy and enthalpy of the binding between AK and its nonhydrolyzable bisubstrate analog P¹,P⁵-di(adenosine) pentaphosphate (Ap5A) are also presented for comparison (Fig. 9.4B and C). The qualitative agreement between the simulation and the experiment is clear and can be used to highlight key features that can serve as a potential diagnostic of linked conformational/binding reactions.

First, we note that the enthalpy of many protein–ligand binding reactions is linear with temperature (as in the case with the low temperature data, Fig. 9.4C). However, the onset of a curvature of H with increasing population of the BI state is expected based on the linkage equations. This phenomenon represents an apparent nonconstant $C_{p,\text{app}}$ (the derivative of H_{app} with respect to temperature). Although not directly observed by ITC, the $C_{p,\text{app}}$ at a given temperature range can be estimated from the slope of the line connecting two measurements of H_{app} , which differ (slightly) in temperature. Such nonconstant $C_{p,\text{app}}$ values may be indicative of a linked process in a binding reaction of interest. The pronounced curvature of the H_{app} of AK binding provides a clear example of this effect.

As stated above, increasing the population of a BI (nonbinding) state in the unbound ensemble will decrease the apparent binding affinity of the system, regardless of the free

energy of the intrinsic association reaction. The decreased affinity of all AK mutants (above 37 °C) is consistent with the expected effects of mutations that increase the probability of a BI state. The temperature dependence of the free energy for such reactions can be more completely understood by considering three population regimes; (1) the *nonpopulated* regime, where the BI state is not significantly populated; (2) the *partially populated* regime, where the BI state is partially populated, and (3) the *fully populated* regime, where BI is populated >99%. As expected, because the BI state never becomes significantly populated in the *nonpopulated* regime, binding reactions that fall within this regime are easily described by a two-state equilibrium, which provides direct access to the intrinsic association reaction. On the other extreme, the *fully populated* regime can also be described by a two-state equilibrium. Although in this regime, the apparent H , S , and C_p represent the sums of the contributions from the intrinsic binding reaction and the conformational transition (see the high temperature data in Fig. 9.4A). Within this regime, it can be difficult to decouple the intrinsic energies and those arising from conformational changes.

Only within the partially populated regime does the energetics provide unambiguous evidence of linked conformational and binding reactions. Owing to the evolving difference in the amount of BI at different temperatures, within this regime, the free energy function within this regime cannot be adequately described by a two-state model. However, in practice such a determination can be difficult to make, owing in large part to the wide temperature range over which small (i.e., low enthalpy) conformational transitions occur, as well as the corresponding level of precision that would be required to make such a determination.

As discussed previously (Eftink et al., 1983), another hallmark feature of coupled thermodynamic processes, which may also serve as an indicator of a linked equilibrium, is the observation of entropy/enthalpy compensation, whether such effects are manifested is highly dependent on the thermodynamic parameters describing the conformational reaction. In general, protein folding is enthalpically favorable and entropically unfavorable. As a result, induced folding via ligand association should produce a considerable degree of compensation. Indeed, we note that for AK the relatively small change in G_{app} is accompanied by relatively large changes in H_{app} (see Fig. 9.4B and C).

In summary, it appears that mutants and WT AK proteins demonstrate several key thermodynamic features of linked folding and binding. However, it should be clear that other models could fit the data equally well. For this reason, qualitative agreement between the model and the data should not be considered unequivocal evidence in support of the model. Instead, we use this result as the starting point and take steps to validate the model through structural and dynamic analyses.

4. Strategies for Quantitative Interpretation of Measured Enthalpies for a Linked Folding and Binding System

A significant benefit of investigating ligand binding reactions with ITC is that in addition to binding affinities, which can be obtained from numerous techniques, ITC provides direct

access to the enthalpy of the reaction. Access to the H of a binding reaction renders ITC unique in its ability to determine the cooperativity of the reaction.

To demonstrate this unique potential, we rewrite the conformational contribution to the observed enthalpy of binding (H_{app} , as in Eq. (9.11)) as,

$$\Delta H_{conf,app}(T) = \left(\frac{e^{\Delta G_{conf}(T)/-RT}}{1 + e^{\Delta G_{conf}(T)/-RT}} \right) (\Delta H_{m,conf} - \Delta C_{p,conf}(T - T_{m,conf})), \quad (9.16)$$

where the temperature dependent free energy ($G_{conf}(T)$) is defined by the Gibbs–Helmholtz equation (Eq. (9.9)). This form illustrates that this observable quantity is determined not only by the enthalpy but also all other parameters determining G_{conf} , namely the changes in heat capacity ($C_{p,conf}$) and conformational entropy (S_{conf}). Moreover, two distinct copies of H_{conf} (and $C_{p,conf}$) are retained in the expression, both within and outside of the exponential term. The exponential term represents the rate of change of the probability of the BI state with temperature (also known as the van't Hoff enthalpy, H_{vH}), while the preexponential copy represents the magnitude of the calorimetrically observable heat evolved (also known as the calorimetric enthalpy, H_{cal}). As a consequence of having both the calorimetric and the van't Hoff enthalpies in the expression, use of Eq. (9.16) as a fitting function carries the added advantage of rigorously imposing two-state thermodynamic behavior by requiring that $H_{vH} = H_{cal}$. Indeed, if Eq. (9.16) produces a satisfactory fit, a thermo-dynamically two-state process is highly suggested.

Although the benefits of using Eq. (9.16) for nonlinear least squares (NLS) fitting are clear, the calorimetrically observed apparent enthalpy, H_{app} , contains at least one more term, representing the intrinsic heat of ligand association to the BC state (H_0). In principle, this term could be retained in the fitting equation, although at the cost of two more free parameters (i.e., a line). However, to simplify the analysis of AK, we subtracted the extrapolated low temperature baseline enthalpy prior to fitting.

Finally, all other possible linked contributions to the observed enthalpy should be considered. For example, protonation or deprotonation events that are coupled to the BI to BC transition will impact the results and possibly alter the analysis. Likewise, the globally unfolded state (i.e., the state where all residues are unfolded) can contribute enthalpy to the measured heat of binding at high temperatures. For such a case, the heat corresponding to the unfolding of additional residues can be estimated from CD unfolding experiments, and subtracted from the data prior to fitting, as was done for AK (Schrack et al., 2009).

NLS fitting of Eq. (9.16) to the conformational contribution to the enthalpy of binding ($H_{conf,eff}$) for the WT and three Val to Gly mutants of AK is shown in Fig. 9.5, and the parameter estimates are summarized in Table 9.2. Interestingly, all four AK variants could be described by a single value of H_{conf} and $C_{p,conf}$ (determined for the most complete data set, v142g), indicating that the conformational process that is being monitored is the same for all of the mutants and the wild type. This agreement suggests that the mutation

strategy was successful in perturbing only the entropy of the local folding reaction in question. Furthermore, the conservation of the enthalpy of local folding suggests that the same cooperative unit (subset of residues) of local folding is shared between all mutants and the wild-type protein. As noted above, the high quality of the fits provides strong evidence supporting a two-state thermodynamic model of the local folding reaction.

In summary, NLS fitting of Eq. (9.16) is a powerful method for the extraction of thermodynamic information for linked folding and binding reactions. In addition, the entropy-enhancing mutation strategy has both allowed us to amplify the native state local unfolding to an extent that it can be rigorously analyzed, while the same time preserving both the temperature dependence of the reaction (H_{conf} and $C_{\text{p,conf}}$), as well as the cooperative unit of the local unfolding event found in the wild-type protein.

5. Interplay of Local Mutational Effects, Global Stability, and Binding Affinity

In general, destabilizing mutations are expected to affect the thermodynamic stability of the native state with respect to the globally unfolded state. However, when mutations are placed in a locally unstable region, the effect on the global stability will depend not only on the destabilizing impact of the mutation itself, but also the relative stability of the region where the mutation is placed (Ferreon et al., 2004). This slightly counterintuitive result can be understood by considering a three-state unfolding reaction. The relevant partition function can be written as

$$Q = K_{\text{conf}} + K_{\text{conf}}K_{\text{unfold}}, \quad (9.17)$$

where

$$K_{\text{conf}} = e^{\Delta G_{\text{conf}}' - RT} = \frac{P_{\text{BI}}}{P_{\text{BC}}} = \frac{P_{\text{LU}}}{P_{\text{F}}}; \quad K_{\text{unfold}} = e^{\Delta G_{\text{unfold}}' - RT} = \frac{P_{\text{U}}}{P_{\text{LU}}}, \quad (9.18)$$

and F represents the fully folded state, LU the BI and locally unfolded state discussed above, and U the state where all residues are unfolded. The probability of all states in the system is therefore given by

$$P_{\text{F}} = \frac{1}{Q}; \quad P_{\text{LU}} = \frac{K_{\text{conf}}}{Q}; \quad P_{\text{U}} = \frac{K_{\text{conf}}K_{\text{unfold}}}{Q}. \quad (9.19)$$

As written, G_{unfold} represents the free energy difference between a locally unfolded state and the state where all residues are unfolded. For a mutation perturbing K_{conf} only, the impact on the stability of the globally unfolded state will depend on which terms dominate the partition function. For instance, in the extreme case that $K_{\text{conf}} \gg 1$, increasing K_{conf} will not appreciably change the probability of the unfolded state (P_{U}), as all dominant terms in

both numerator and denominator are multiplied by this value. In the opposite case, where the locally unfolded state is highly improbable, mutational effects are fully expressed in the globally unfolded state. These trends are visualized in Fig. 9.6A, where the effect of an identical thermodynamic mutation is simulated in the context of several different local stabilities (according to Eqs. (9.17)–(9.19)). To summarize, the apparent effect of the mutation on global stability is inversely correlated with the local stability of the region surrounding the mutation site (Ferreon et al., 2004).

To assess the effects of the Val to Gly mutations in AK, global unfolding experiments were performed by monitoring CD as a function of temperature (see Fig. 9.6B). The observed destabilization was relatively minimal (a decrease in the unfolding transition temperature (T_m) < 1.5 °C for all mutants), suggesting as noted above that the region in question is relatively unstable to local unfolding (see Fig. 9.6B). In terms of the three-state unfolding system presented in Eqs. (9.17)–(9.19), the observed CD signal for AK can be described by,

$$\langle \theta_{228} \rangle(T) = \theta_{228,F}(T)[P_F(T) + P_{LU}(T)] + \theta_{228,U}[P_U(T)], \quad (9.20)$$

where there is no (or little) difference in the observed value for the F and LU states. This inference is reasonable as the LU state of the mutant AKs is known to be highly populated before the global unfolding transition ($T_{m,local} = 35\text{--}41$ °C); the CD signal is minimally affected in this range. In the case of v142g, the LU state is populated to a level of ~85% before the initiation of the observed global unfolding transition. Therefore, the two-state fit to the CD data should provide a suitable estimate of $K_{unfold}(T)$, that is, the thermodynamics of the LU to U transition. Having estimates of $K_{unfold}(T)$, $K_{conf,wt}(T)$, and $K_{conf,v142g}(T)$, the probabilities of all three states can be calculated from Eq. (9.19; see Fig. 9.7A). Interestingly, this calculation reproduces the measured (CD) temperature dependence of the globally unfolded state of WT AK (Fig. 9.7A, *dot-dashed line*), although no WT unfolding data is included. Such a prediction, using an independent observable, strongly validates the accuracy of the parameter estimates describing the local unfolding reaction, as well as the thermodynamic analysis applied to both the ITC and CD data. We note that a significant population of the locally unfolded state is revealed for the WT AK at biological temperatures (~5% at 37 °C).

One other thermodynamic relationship can be developed, which can provide additional means of validating the analysis. As described above (Eq. (9.6)), the observed free energy of binding will also depend on the position of the relevant linked conformational equilibrium, $K_{conf}(T)$. If mutational effects are manifested exclusively in the conformational equilibrium, leaving the intrinsic binding interaction unchanged, the difference in the apparent binding affinity between the two proteins will be

$$\Delta\Delta G_{app, wt - mut} = RT \ln \left(\frac{1 + K_{conf,mt}}{1 + K_{conf,mut}} \right), \quad (9.21)$$

where the two identical terms representing $G_0(T)$ cancel. For clarity, this can be rewritten as

$$\Delta\Delta G_{\text{app, wt} - \text{mut}} = RT \ln \left(\frac{1 + K_{\text{conf, wt}}}{1 + K_{\text{conf, wt}} \Omega_{\text{mut}}} \right), \quad (9.22)$$

where Ω_{mut} represents the effect of the mutation (i.e., the change in ratio of accessible conformational space to Val and Gly in the disordered state, Eq. (9.1)). Considering these relationships, it becomes clear that when the mutated region is highly unstable ($K_{\text{conf, wt}} \gg 1$), then the effect of the mutation on G will be maximized. Conversely, low probabilities of the BI state diminish the mutational effects. In the case of AK, low temperatures suppress local unfolding ($1 > K_{\text{conf, wt}}$). Accordingly, the difference in G_{obs} between mutant and wild-type AKs decreases with decreasing temperature. For v142g AK, the difference in affinity as compared to WT approaches zero, indicating that the G_0 of these proteins are equal. Therefore, we can use the above estimated thermodynamic parameters describing the BI state for the wild type and v142g proteins together with Eq. (9.21), to directly predict the difference in the observed binding affinity of these two proteins.

This calculation was performed by first fitting Eq. (9.6) to the v142g data (the most extensive data set), fixing all parameters given ($C_{\text{p, intrinsic}}$) or estimated (H_{conf} , $C_{\text{p, conf}}$, and $T_{\text{m, conf}}$) in the fitting of the observed enthalpy. G was then calculated as a distance from this fitted curve. The agreement of the experimental data and this prediction validates both the proposed binding model and the quality of the fitted parameters obtained from the enthalpy data (Fig. 9.7B). We note that the assumption of low to no binding affinity for the BI state is strongly validated by this simulation. Also, although this prediction cannot be quantitatively applied to v148g and v135g because of apparent small perturbations to the intrinsic free energy of binding (Fig. 9.4B, low temperature), the fact that each of the Val to Gly mutations shows similar temperature dependence of the binding affinity (see Fig. 9.7B, *inset*) further supports the validity of the model for all mutants investigated.

In previous work from our group, Ferreon et al. proposed an alternate experimental approach for determining local stability, using the identical mutation strategy under discussion. In this work, they showed that by combining HX and ITC data, the principle of inversely related effects on stability and affinity can be leveraged to determine local stability (Ferreon et al., 2004). Such an approach only requires measurements at a single temperature or solution condition. By this method, they demonstrated that the RT-loop of SEM-5 SH3 domain (a flexible binding loop) behaves as it was 25% in an unfolded conformation (Ferreon et al., 2004). We note that many local unfolding reactions may have small or poorly determined enthalpy. In this case, this affinity/stability method may be more applicable. The enthalpy-based approach demonstrated here is preferred when possible, because of the relatively complete thermodynamic characterization available from this method.

6. Success of the Strategy in Preserving Structure

One important goal of our mutation strategy was to selectively perturb the properties of local unfolded-like conformational excursions, while leaving the ground-state structure unperturbed. For this reason, we have chosen highly surface-exposed residues, involved in few molecular contacts. If our mutation design is successful at achieving the goal of structural preservation, then we can be much more certain that any observed changes in binding are a consequence of modulating the stability of highly disordered or unfolded states. This inference becomes very important when experimental isolation of the intrinsic and conformational contributions to measured binding affinity or enthalpy is not possible.

As a reference, it is well known that the fold shared by homologous proteins is robust to sequence changes. For example, AKs from *Bacillus* sp. share the fold of *E. coli* AK. Furthermore, an interesting series of *Bacillus* AKs, sharing ~70% sequence identity, have been examined by crystallography (Bae and Phillips, 2004). Structural alignment of the LID region (residues 128–159) of the molecules revealed alpha carbon RMSD values of 0.43–0.46 Å (Bae and Phillips, 2004). By extension, the structural effects of a single point mutation, as in the strategy under discussion, are expected to be subtle and not propagate by largely perturbing the conformation of the protein backbone. Of course, local structural consequences of disrupted interactions are expected. To evaluate the success of the proposed (surface-exposed Val to Gly) mutation strategy on preserving the structure of AK, we performed crystallography on two AK mutants (v142g, v148g) and the wild-type molecule.

Only the v148g mutant crystallizes in the identical space group ($P2_12_12$) and asymmetric unit (ASU) as the wild type, and thus merits detailed comparison. When comparison of small structural changes is the goal, one must carefully consider the effects of crystal packing on the resultant structure. In this case, the ASU happens to contain two distinct copies of AK. This provides an excellent opportunity to examine the magnitude and distribution of crystal packing effects. Interestingly, a distinct conformation of the outer region of the LID domain is associated with each crystallographic environment investigated (also observed in the distinct ASU of v142g, data not presented). We note that this region (residues ~130–155) has been demonstrated to be the most dynamic region of the protein on the nano- to picosecond time scale (Shapiro et al., 2000). Therefore, structural plasticity observed in this region may be indicative of a relatively heterogeneous conformational manifold that can be dynamically explored in solution (Shapiro et al., 2000). Alignment demonstrates that the influences of crystal packing are largely identical for both WT and v148g AK (Fig. 9.8A).

However, we can also examine the structural changes effected by our mutations (where crystallographic influences should be minimal) by comparing identically oriented copies of WT and v148g from the shared ASU. The similar backbone structures of analogous chains of v148g and WT can be seen in Fig. 9.8A. Two copies within the ASU also provide another benefit when using medium resolution data (as here), in that true observable structural perturbations effected by mutation should be reproducible in the data from chain A to chain B. Using a very small threshold of perturbation ($>0.3\text{Å}$), we identified all atoms that are potentially perturbed in both copies of the ASU, comparing WT versus v148g structures.

Such atoms are highlighted in Fig. 9.8B. Notably, the structures surrounding the mutation sites do not show a significant increase in the number of perturbed atoms. Furthermore, most of the identified atoms belong to surface-exposed side chains where motility of the side chain results in poor electron density. The positions of such atoms are simply poorly determined. Therefore, within the available resolution of measurement, it appears that our mutation strategy is highly successful at preserving the fully folded structure of the protein. These data have been corroborated by the preservation of ^1H - ^{15}N chemical shifts (HSQC experiment) of WT and mutant AKs, as well as similar NMR studies of the SH3 SOS-Y binding reaction (Ferreon and Hilser, 2003b; Schrank et al., 2009).

7. Comparison of Interaction Versus Entropy Based Mutation Strategy

The s129a mutation represents an interesting control study, as this mutation should destabilize the LID by a completely distinct mechanism. According to the crystal structure (Schrack et al., 2009), the hydroxyl group of serine 129 is a hydrogen bond donor to a nearby histidine side chain nitrogen. Mutation to Ala disrupts this interaction, while preserving the β -carbon at position 129, and therefore should make very minimal perturbations to the sterically allowed conformational freedom of this position.

Identical ITC experiments were carried out for this mutant. As expected, the LID region appears to be highly destabilized, as a very large increase in the favorable enthalpy of binding is observed. However, it is interesting to note that the gross shape of the change in $H_{\text{conf,app}}$ with temperature is very different from the shared behavior of the Val to Gly series, see Fig. 9.9A. Furthermore, this data cannot be fitted by the shared conformational scheme (with or without shared H_{conf} and $C_{p,\text{conf}}$) that so nicely describes Val to Gly mutants (see Fig. 9.9A). Also, the magnitude in the observed differences is inconsistent with what would be expected for local mutational effects, ~ 10 kcal/mol. Therefore, the thermodynamic model describing the local unfolding reaction must be fundamentally different. We again note the utility of Eq. (9.16) as a test of a two-state system, which in this case is clearly excluded.

Therefore, the case of s129a mutation demonstrates the desirable properties of the investigated Val to Gly mutation strategy. It appears that preservation of energetic relationships (cooperativity) between the many possible states of unfolding is a unique property of the noninvasive, entropy-enhancing strategy that we utilize. Furthermore, we note that crystallographic investigation of the s129a mutation reveals minimal changes, similar to those reported for v148g (unpublished results). Therefore, the presented data support the greater utility of entropy-based, surface-exposed mutations as compared to even structurally conservative interaction-directed mutations.

8. How Similar Are Local and Global Unfolding?

One might expect that a locally unfolded region of an otherwise folded protein may have important differences from an isolated unfolded peptide free in solution. Both the spatially constrained peptide ends and the relatively high likelihood of interactions with the surrounding structures should in principle alter the conformational manifold. The analysis

presented here gives us a rare opportunity to compare several different thermodynamic properties of local unfolding within the native state of AK with what would be expected for global unfolding of that same region.

Published values are available for the expected entropic consequences of a Val/Ala to Gly mutation for an unfolding reaction (D'Aquino et al., 1996). In addition to the backbone contributions emphasized in Section 2, unfolding of Val increases the rotational freedom of the side chain (Lee et al., 1994). We have used published values to simulate the expected destabilizing effect of Val to Gly mutation in AK, considering the background of the measured thermodynamics of the local unfolding in WT AK. Because of the expected variability in side chain contributions, we have simulated both the case of an average surface-exposed Val to Gly as well as the backbone contribution of Val to Gly mutation alone (see Fig. 9.5). In both cases, the effects of all mutations were found to be greater than or equal to the expected effect. We reiterate that the destabilization induced by the mutations appears to be strictly entropic. As demonstrated above, mutation effects should be decreased relative to the prediction if residues in the locally unfolded regions are partially constrained. Although the origin of the additional entropy gain is unknown, our data are in agreement with a model where the mutated positions, upon local unfolding, gain the full amount of conformational freedom available to the same residue in the fully unfolded state.

This assertion is also buttressed by success of the v142g local unfolding data to describe the difference between the v142g and WT AK in terms of global stability. If additional conformational freedom were gained at the mutated site upon global unfolding, as might be expected, the three-state unfolding model presented would underestimate the effect of the mutation on global unfolding. In this case, the wild-type protein should be more stable than predicted. Thus, the ability of the local unfolding parameters to describe the global unfolding data also supports the equivalence of the entropy (and conformational diversity) of local and global unfolding for the mutated positions in question.

Similarly, it is interesting to investigate the solvent associated enthalpy and heat capacity change of local unfolding to determine if these values are also compatible with an unfolding model. Fortunately, excellent parametric equations exist that relate changes in solvent exposed (polar and apolar) surface area (ASA) upon unfolding to the heat and heat capacity of the unfolding reaction (Murphy et al., 1992). These relationships are based on calorimetric unfolding experiments. Estimation of the H and C_p expected for a given local unfolding reaction can be made, based on the crystallographic structure, by calculating the expected ASA of transferring a segment of the protein from the context of the fold into the solvent (i.e., the surface area of the molecular interface) as well as the ASA of unfolding the isolated segment. Recent reviews have described the full computational procedure for making such calculations (COREX algorithm), so the details of these methods will not be discussed further here (Hilser and Freire, 1997; Hilser et al., 2006).

In the case of the local unfolding in AK, the amplification of the locally unfolded state has allowed us to map the residues involved in the local unfolding reaction using NMR, simply assigning resonances with (greatly) enhanced chemical exchange broadening in the HSQC spectrum of the v142g mutant (Schrack et al., 2009). This well-described effect is due to the

local unfolding reaction, which happens to be on the chemical shift time scale. Therefore, using the surface area based approximation described above, predictions of the local and global unfolding of AK were made, and are summarized in Table 9.3. As is evident, the calculation reproduces the measured heats of local unfolding as well as unfolding the remainder of the protein, as approximated by the global unfolding experiment for v142g. Therefore, in terms of solvent accessibility and the related thermodynamic parameters H and C_p , the locally unfolded state appears to be indistinguishable from a true protein unfolding event.

These data demonstrate that local unfolding within the context of a folded protein at times can be highly similar in thermodynamic and conformational character to much studied global unfolding reactions. This result, of course, does not imply that the protein is completely extended. Instead, it simply means that the thermodynamic character of local and global unfolding is similar.

9. Summary

The exploration of the structural, energetic, and kinetic details of the native state ensemble is one of the key goals of modern biochemistry. However, the experimental tool box available to the investigator is still rather limited, especially when searching for techniques that will yield physical insight into the conformational nature of fluctuations and/or allow the investigation of minor populations typically hidden in ensemble (average property) based experimental approaches.

In this work, we have demonstrated an experimental strategy that allows one to selectively probe the native state ensemble for highly disordered states. The advantages of the surface-exposed, binding site associated Gly mutation strategy used are threefold. protein (i.e., the intrinsic binding affinity) need not be altered, thus allowing unambiguous interpretation of the functional effects of high-energy states. Second, studying binding site associated regions of the protein allows depopulation of high-energy states by the addition of ligand. This external tool for population modulation greatly facilitates the calorimetric/thermodynamic investigation of partially populated states. Third, Gly is a somewhat selective probe for highly disordered states, as large regions of ϕ/ψ space must be explored for a strong effect. Thus, the effectiveness of the mutation strategy itself provides qualitative conformational information.

The example mutations discussed here have been successful for the study of conformational fluctuations within the native state of AK (Schrack et al., 2009). We have found that in this case, our strategy preserves the cooperative substructure of local unfolding present within the WT protein. Furthermore, the H and C_p of native state unfolding of the LID are preserved between WT and mutant proteins. This greatly controlled set of mutations has allowed the otherwise impossible estimation of the population of LID unfolding in WT AK (~5% at 37 °C; Schrack et al., 2009). In ongoing research in our group, these mutant proteins have allowed us to enhance the population of local unfolding above the detection limit of relaxation dispersion NMR experiments (unpublished results), thus revealing detailed kinetic and conformational information about these states.

This experimental framework has also been successfully applied to conformational fluctuations of the SEM-5 SH3 domain, as well as experiments designed to probe the polyproline-II content of peptides free in solution (Ferreon and Hilser, 2003b; Ferreon et al., 2004; Manson et al., 2009). Therefore, this approach is likely to be widely applicable to many other systems and compatible with many other experimental techniques such as NMR, FRET, and single molecule techniques.

REFERENCES

- Babu CR, Hilser VJ, and Wand AJ (2004). Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation. *Nat. Struct. Mol. Biol* 11, 352–357. [PubMed: 14990997]
- Bae E, and Phillips GN (2004). Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. *J. Biol. Chem* 279, 28202–28208. [PubMed: 15100224]
- Bai Y (2006). Protein folding pathways studied by pulsed- and native-state hydrogen exchange. *Chem. Rev* 106, 1757–1768. [PubMed: 16683753]
- Baldwin AJ, and Kay LE (2009). NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol* 5, 808–814. [PubMed: 19841630]
- Boehr DD, McElheny D, Dyson HJ, and Wright PE (2010). Millisecond timescale fluctuations in dihydrofolate reductase are exquisitely sensitive to the bound ligands. *Proc. Natl. Acad. Sci. USA* 107, 1373–1378. [PubMed: 20080605]
- Cliff MJ, Williams MA, Brooke-Smith J, Barford D, and Ladbury JE (2005). Molecular recognition via coupled folding and binding in a TPR domain. *J. Mol. Biol* 346, 717–732. [PubMed: 15713458]
- D'Aquino JA, Gómez J, Hilser VJ, Lee KH, Amzel LM, and Freire E (1996). The magnitude of the backbone conformational entropy change in protein folding. *Proteins* 25, 143–156. [PubMed: 8811731]
- Eftink MR, Anusiem AC, and Biltonen RL (1983). Enthalpy–entropy compensation and heat capacity changes for protein–ligand interactions: General thermodynamic models and data for the binding of nucleotides to ribonuclease A. *Biochemistry* 22, 3884–3896. [PubMed: 6615806]
- Englander SW (1975). Measurement of structural and free energy changes in hemoglobin by hydrogen exchange methods. *Ann. NY Acad. Sci* 244, 10–27. [PubMed: 1056161]
- Ferreon JC, and Hilser VJ (2003a). Ligand-induced changes in dynamics in the RT loop of the C-terminal SH3 domain of Sem-5 indicate cooperative conformational coupling. *Protein Sci* 12, 982–996. [PubMed: 12717021]
- Ferreon JC, and Hilser VJ (2003b). The effect of the polyproline II (PPII) conformation on the denatured state entropy. *Protein Sci* 12, 447–457. [PubMed: 12592015]
- Ferreon JC, Hamburger JB, and Hilser VJ (2004). An experimental strategy to evaluate the thermodynamic stability of highly dynamic binding sites in proteins using hydrogen exchange. *J. Am. Chem. Soc* 126, 12774–12775. [PubMed: 15469262]
- Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, and Kern D (2007). A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* 450, 913–916. [PubMed: 18026087]
- Hilser VJ, and Freire E (1997). Predicting the equilibrium protein folding pathway: Structure-based analysis of staphylococcal nuclease. *Proteins* 27, 171–183. [PubMed: 9061781]
- Hilser VJ, García-Moreno EB, Oas TG, Kapp G, and Whitten ST (2006). A statistical thermodynamic model of the protein ensemble. *Chem. Rev* 106, 1545–1558. [PubMed: 16683744]
- Huyghues-Despointes BM, Langhorst U, Steyaert J, Pace CN, and Scholtz JM (1999). Hydrogen-exchange stabilities of RNase T1 and variants with buried and solvent-exposed Ala → Gly mutations in the helix. *Biochemistry* 38, 16481–16490. [PubMed: 10600109]

- Korzhnev DM, Neudecker P, Zarrine-Afsar A, Davidson AR, and Kay LE (2006). Abp1p and Fyn SH3 domains fold through similar low-populated intermediate states. *Biochemistry* 45, 10175–10183. [PubMed: 16922492]
- Krishna MMG, Hoang L, Lin Y, and Englander SW (2004). Hydrogen exchange methods to study protein folding. *Methods* 34, 51–64. [PubMed: 15283915]
- Leach SJ, Némethy G, and Scheraga HA (1966). Computation of the sterically allowed conformations of peptides. *Biopolymers* 4, 369–407. [PubMed: 5909467]
- Lee KH, Xie D, Freire E, and Amzel LM (1994). Estimation of changes in side chain configurational entropy in binding and folding: General methods and application to helix formation. *Proteins* 20, 68–84. [PubMed: 7824524]
- LeMaster DM, Anderson JS, and Hernández G (2009). Peptide conformer acidity analysis of protein flexibility monitored by hydrogen exchange. *Biochemistry* 48, 9256–9265. [PubMed: 19722680]
- Liu T, Whitten ST, and Hilser VJ (2007). Functional residues serve a dominant role in mediating the cooperativity of the protein ensemble. *Proc. Natl. Acad. Sci. USA* 104, 4347–4352. [PubMed: 17360527]
- Maity H, Lim WK, Rumbley JN, and Englander SW (2003). Protein hydrogen exchange mechanism: Local fluctuations. *Protein Sci* 12, 153–160. [PubMed: 12493838]
- Manson A, Whitten ST, Ferreon JC, Fox RO, and Hilser VJ (2009). Characterizing the role of ensemble modulation in mutation-induced changes in binding affinity. *J. Am. Chem. Soc* 131, 6785–6793. [PubMed: 19397330]
- Miller DW, and Dill KA (1995). A statistical mechanical model for hydrogen exchange in globular proteins. *Protein Sci* 4, 1860–1873. [PubMed: 8528084]
- Müller CW, and Schulz GE (1992). 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* 224, 159–177. [PubMed: 1548697]
- Müller CW, Schlauderer GJ, Reinstein J, and Schulz GE (1996). Adenylate kinase motions during catalysis: An energetic counterweight balancing substrate binding. *Structure* 4, 147–156. [PubMed: 8805521]
- Murphy KP, Bhakuni V, Xie D, and Freire E (1992). Molecular basis of cooperativity in protein folding III. Structural identification of cooperative folding units and folding intermediates. *J. Mol. Biol* 227, 293–306. [PubMed: 1522594]
- Pan H, Lee JC, and Hilser VJ (2000). Binding sites in *Escherichia coli* dihydrofolate reductase communicate by modulating the conformational ensemble. *Proc. Natl. Acad. Sci. USA* 97, 12020–12025. [PubMed: 11035796]
- Ponder JW, and Richards FM (1987). Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol* 193, 775–791. [PubMed: 2441069]
- Prabhu NV, and Sharp KA (2005). Heat capacity in proteins. *Annu. Rev. Phys. Chem* 56, 521–548. [PubMed: 15796710]
- Ramachandran GN, and Sasisekharan V (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem* 23, 283–438. [PubMed: 4882249]
- Ramakrishnan C, and Ramachandran GN (1965). Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. *Biophys. J* 5, 909–933. [PubMed: 5884016]
- Rogero JR, Englander JJ, and Englander SW (1986). Individual breathing reactions measured by functional labeling and hydrogen exchange methods. *Methods Enzymol* 131, 508–517. [PubMed: 3773770]
- Rundqvist L, Adén J, Sparrman T, Wallgren M, Olsson U, and Wolf-Watz M (2009). Noncooperative folding of subdomains in adenylate kinase. *Biochemistry* 48, 1911–1927. [PubMed: 19219996]
- Schrack TP, Bolen DW, and Hilser VJ (2009). Rational modulation of conformational fluctuations in adenylate kinase reveals a local unfolding mechanism for allostery and functional adaptation in proteins. *Proc. Natl. Acad. Sci. USA* 106, 16984–16989. [PubMed: 19805185]

- Shapiro YE, Sinev MA, Sineva EV, Tugarinov V, and Meirovitch E (2000). Backbone dynamics of Escherichia coli adenylate kinase at the extreme stages of the catalytic cycle studied by (15)N NMR relaxation. *Biochemistry* 39, 6634–6644. [PubMed: 10828981]
- Vallurupalli P, Hansen DF, Stollar E, Meirovitch E, and Kay LE (2007). Measurement of bond vector orientations in invisible excited states of proteins. *Proc. Natl. Acad. Sci. USA* 104, 18473–18477. [PubMed: 18006656]
- Whitten ST, Wooll JO, Razeghifard R, García-Moreno EB, and Hilser VJ (2001). The origin of pH-dependent changes in m-values for the denaturant-induced unfolding of proteins. *J. Mol. Biol* 309, 1165–1175. [PubMed: 11399086]
- Whitten ST, Yang H, Fox RO, and Hilser VJ (2008). Exploring the impact of polyproline II (PII) conformational bias on the binding of peptides to the SEM-5 SH3 domain. *Protein Sci* 17, 1200–1211. [PubMed: 18577755]
- Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, and Kern D (2004). Linkage between dynamics and catalysis in a thermophilic–mesophilic enzyme pair. *Nat. Struct. Mol. Biol* 11, 945–949. [PubMed: 15334070]
- Wright PE, and Dyson HJ (2009). Linking folding and binding. *Curr. Opin. Struct. Biol* 19, 31–38. [PubMed: 19157855]

Figure 1.

Expected effects of contact-based (classical) versus entropy based mutations on the native state conformational energy landscape. *Black*: The conformational space available to the wild-type protein. *Gray*: The conformational space available to the mutant protein. (A) The expected changes in a mutation designed to perturb contacts within the native fold of the molecule. (B) The idealized effect of the surface-exposed Gly mutation, which only perturbs the conformational entropy of excited conformational states and leaves the ground-state fold unaffected.

Figure 2.

Hard sphere collision simulation of the relative conformational space available to Ala and Gly in a disordered peptide. All simulations were performed using the MPMOD program (Manson et al., 2009; Whitten et al., 2008). Conformers were generated until 15,000 sterically allowed (Ponder and Richards, 1987; Ramachandran and Sasisekharan, 1968; Ramakrishnan and Ramachandran, 1965) conformations were found for each peptide (AAXAA) in question. The number of attempts required to achieve this goal was recorded for each peptide, as well as the ϕ and ψ angles of each acceptable conformer. Slightly reduced atomic radii were applied to the peptide backbone to increase sampling efficiency (Ponder and Richards, 1987). (A, B) Contour histogram of acceptable conformations of Gly and Ala. *Case-1*: A space accessible to a folded residue that will have similar conformational

freedom if mutated from Val or Ala to Gly. *Case-2*: A space accessible to a folded residue that may have different conformational freedom if mutated from Val or Ala to Gly (an example of a nonideal mutation site). *Points*: Favorable folded conformations of mutation sites selected for AK, PDB ID 1AKE (Müller and Schulz, 1992).

Figure 3.

Mutational strategy and binding associated conformational changes in AK. AK from *E. coli* is a much studied enzyme that catalyzes the reaction $2\text{ADP} \leftrightarrow \text{ATP} + \text{AMP}$. *Open*: Graphic of Apo-AK, PDB ID 4AKE (Müller et al., 1996). Ap5A: P^1, P^5 -di (adenosine-5') pentaphosphate, a nonhydrolysable bisubstrate analog inhibitor of AK. *Closed*: Graphic of the Ap5A/AK complex, PDB ID 1AKE (Müller and Schulz, 1992). *LID*: The “LID domain” is a highly dynamic structural appendage associated with the ATP binding site. *AMPbd*: The “AMP binding domain” is a highly dynamic structural appendage associated with the AMP binding site. Definitions of these regions have been previously published (Shapiro et al., 2000). *Spheres*: Indicate the location of selected Val (or serine) to Gly mutations sites. Adapted from previously published results (Schrack et al., 2009).

Figure 4.

Agreement of expected thermodynamic observables of linked folding and binding with those of from AK Val to Gly Mutants. (A) Simulation of expected calorimetric observables according to Eqs. (9.3)–(9.15) and populations of the relevant BI state. Simulation parameters defining the linked folding reaction were: $H_{conf}(35\text{ °C}) = 33\text{ kcal/mol}$, $C_{p,conf} = 660\text{ cal/mol K}$, $T_{m,conf} 35\text{ °C}$. Parameters defining the intrinsic binding (and conformational) reaction were chosen to approximate the AK data. (B, C) ITC measured (apparent) thermodynamic parameters of the AK/ P^1, P^5 -Di(adenosine-5') pentaphosphate (Ap5A) binding reaction. Ap5A is a nonhydrolysable bisubstrate inhibitor of AK. Solution conditions were 60 mM PIPES acid, 2 mM EDTA, and pH 7.85. Adapted from previously published results (Schrack et al., 2009).

Figure 5.

Fitting of $H_{\text{conf,app}}$ for thermodynamic parameters of the binding associated conformational reaction. Corrected data (see text) and fitting functions representing

$H_{\text{conf,app}}$. Error bars are two times standard experimental error or the average thereof. *Solid lines*: represent the fitting functions. Dashed line is a simulation of the data for the expected entropy gain Val to Gly mutation based published entropy estimates (D'Aquino et al., 1996; Lee et al., 1994), and the fitted conformational H and C_p . *Dot-dashed line*: represents a simulation of the expected data for a Val to Gly mutation where folding places no conformational restriction on the side chain, that is, only the backbone contribution is taken into account (D'Aquino et al., 1996). Adapted from previously published results (Schrack et al., 2009).

Figure 6.

Simulation of regional destabilization on global protein stability and representative circular dichroism unfolding experiments of AK mutants. (A) Simulation based on Eqs. (9.17)–(9.19). Parameters defining $K_{\text{unfold}}(T)$ were $T_{m,\text{unfold}} = 55$ °C, $H_{\text{unfold}}(T_m) = 89$ kcal/mol, $C_{p,\text{unfold}} = 2.8$ kcal/mol K. Parameters defining $K_{\text{conf}}(T)$ were: $T_{m,\text{conf}}$ as indicated, $H_{\text{conf}}(55$ °C) 33 kcal/mol, $C_{p,\text{conf}} = 0.66$ kcal/mol K. *Lines:* Simulated WT/mutant protein pairs. The applied thermodynamic mutation entropically destabilizes the locally unfolding region, $S_{\text{mutation}} = 5$ cal/mol K. (B) Experimental observation of global of AK mutants. Molar ellipticity at 288 nm, Θ_{288} (deg cm²/dmol res), recorded by circular dichroism spectroscopy. WT (*black*), v142g (*dark gray*), v148g (*gray*), v135g (*light gray*), and s129a (*very light gray*). Conditions were identical to ITC experiments. Lines are fitting functions, representing a two-state thermodynamic model. Adapted from previously published results (Schrack et al., 2009).

Figure 7.

Agreement of local unfolding parameter estimates with external experimental observables.

(A) Three-state model of unbound adenylate kinase. Populations are calculated from the BC/BI equilibrium, K_{conf} , estimated from the ITC data. The *locally unfolded* (LU)/*unfolded* (U) equilibrium, K_{unfold} , is estimated from the CD unfolding experiments on v142g ($T_{\text{m,unfold}} = 54.7^\circ\text{C}$, $H_{\text{unfold}}(T_{\text{m}}) = 89.4\text{ kcal/mol}$, $C_{\text{p,unfold}} = 2.8\text{ kcal/mol K}$). The reasonability of this approximation is seen in this plot, from the high probability of the BI state at 47°C , prior to the global unfolding transition. Calculations are based on Eqs. (9.17)–(9.19). *Dot-dashed line*: Population of the U state for WT as calculated directly from the CD unfolding parameter estimates. (B) Prediction of WT affinity data based on the fitting of $H_{\text{conf,app}}$, Eq. (9.22). Lines represent the prediction, plus and minus the average standard

error of determining G_{app} . *Inset*: change in G_{app} for each protein, with temperature, referenced from 27 °C. $G_{27\text{ °C}}(Temp) = G(Temp) - G(27\text{ °C})$. Adapted from previously published results (Schrack et al., 2009).

Figure 8.

Structural conservation of v148g and s129a AK. (A) Alignment of the crystal structures of WT and v148g AK (Schrack et al., 2009). *Black*: Chains (WT and v148g) from position A within the ASU. *Gray*: Chains from position B within the ASU. (B) Analysis of structural possible structural perturbations effected by mutation. *Gray spheres*: Represent all atoms that move $>0.3 \text{ \AA}$ from the WT to mutant structure in both copies within the ASU. *Asterisk*: The mutation site (148). *Cross*: All perturbed atoms (spheres) that can be connected to the mutation site by a continuous chain ($<6 \text{ \AA}$ per step) of other perturbed atoms. Adapted from previously published results (Schrack et al., 2009).

Figure 9.

ITC measurement of the s129a/Ap5A Binding Reaction. (A) Corrected data (see text) and fitting functions representing $H_{\text{conf,app}}$. Error bars are two times average standard error for all similar experiments. *Solid lines*: Two-state fitting functions. *Long dashed lines*: Failed fitting (best) of Eq. (9.16) to the s129a data using the shared H and C_p used to fit WT and Gly variants of AK. *Short dashed lines*: Failed fitting (best) of Eq. (9.16) to the s129a data varying H and C_p . (B) The apparent free energy of binding for s129a, v142g and WT AK. *Solid lines*: Prediction of G_{bind} for WT and v142g, as in Fig. 9.7. *Dashed lines*: Failure of either of the corresponding fits ($H_{\text{conf,app}}$) to reproduce the temperature dependence of the s129a binding affinity.

Table 1

Hard sphere collision approximation of the relative conformational space available to Ala and Gly in a disordered peptide

Test space (Φ, Ψ)	$\Omega_{\text{Ala-Gly}}$	$S_{\text{a,g}}$ (cal/mol K)	$\Omega_{\text{Val-Gly}}$	$S_{\text{v,g}}$ (cal/mol K)
All available ($0 \leftrightarrow 360, 0 \leftrightarrow 360$)	3.03, 2.20		4.28, 2.89	
Case-1 ($-150 \leftrightarrow -100, 100 \leftrightarrow 150$)	1.08, 0.14		1.16, 0.29	
Case-2 ($-120 \leftrightarrow -70, -75 \leftrightarrow -25$)	3.09, 2.24		7.27, 3.94	

All simulations were performed using the MPMOD algorithm (Manson et al., 2009; Whitten et al., 2008). Equal probabilities were assumed for all valine rotamers. $\Omega_{\text{Ala-Gly}}$ is as defined in Eq. (9.1). Van der Waals radii were taken from published values (Ponder and Richards, 1987; Ramachandran and Sasisekharan, 1968; Ramakrishnan and Ramachandran, 1965).

Table 2

Thermodynamic parameter estimates for local unfolding of the LID region of AK

	$H_{m,conf}^a$	$C_{p,conf}^b$	$T_{m,conf}^c$
V142G	33 ± 1.0	660 ± 60	35.1±0.3
V148G	nv	nv	37.9±0.3
V135G	nv	nv	41.1±0.2
Wild type	nv	nv	52.3±0.1

±95% Confidence interval (estimated from the fitted model, by profile likelihoods); nv, value not varied in the fitting routine (v142g value applied). Adapted from previously published results (Schrack et al., 2009).

^aTransition enthalpy at $T_{m,conf}$ of v142g (kcal/mol).

^bChange in heat capacity (cal/mol).

^cTransition midpoint temperature (°C).

Table 3

Experimental and predicted thermodynamics of local and global unfolding

	Local ^a		Global ^b		Sum ^c		
	<i>T</i> (°C)	ITC ^d	COREX [‡]	CD (v142g) ^e	COREX [‡]	Experimental ^f	COREX [‡]
<i>H</i>	35.2	33.3 ± 0.5	32.8	34 ± 9 [‡]	40.9	67 ± 10 [‡]	73.7
	54.7	46.2 ± 1.1 [‡]	46.5	89 ± 4	88.2	135.2 ± 5 [‡]	134.6
<i>C_p</i>		0.66 ± 0.03	0.7	2.8 ± 0.2	2.4	3.5 ± 0.2 [‡]	3.1

[‡]Values calculated from the expected change in exposed polar and apolar surface area upon unfolding, using the COREX energy function and PDB ID 4AKE (Müller et al., 1996), confidence is generally ±10% for estimation of *C_p* and *H* for a differential scanning calorimetric unfolding experiment.

[±]Experimental 95% confidence interval (CD), or standard error (ITC).

[‡]Propagated error. Adapted from previously published results (Schränk et al., 2009).

^aExperimental and computational measures of a possible local unfolding reaction (residues 110–164).

^bExperimental and computational measures of unfolding the remainder of the protein (residues 1–109, 165–214).

^cSum of local and global reactions. Represents unfolding all residues in the protein (residues 1–214).

^dParameters estimated from the ITC data fitting, or calculated from these estimates.

^eParameters estimated from the CD data fitting, or calculated from these estimates.

^fSum of heat or heat capacity estimated for the BI/BC transition (ITC) and the native/denatured transition (CD).