Copyright

By

Travis Parke Schrank

2010

**The Dissertation Committee for Travis Schrank Certifies that this is the approved version of the following Dissertation:**

# INVESTIGATIONS OF LOCAL UNFOLDING LIKE CONFORMATIONAL EXCURSIONS IN THE NATIVE STATE OF ADENYLATE KINASE FROM E. COLI

**Committee:**

Vince Hilser, Supervisor

Wayne Bolen, Co-Supervisor

Wlodzimierz Bujalowski

Vicente Resto

Junji Iwahara

Werner Braun

John Olson

Dean, Graduate School

# Investigations of Local Unfolding Like Conformational Excursions in the Native State of Adenylate Kinase

by

**Travis Parke Schrank, B.A.**

**DISSERTATION**

Presented to the Faculty of

**THE GRADUATE SCHOOL OF BIOMEDICAL SCIENCES**
**OF THE UNIVERSITY OF TEXAS MEDICAL BRANCH**

in partial fulfillment
of the requirements
for the degree of

**Doctor of Philosophy**

**THE UNIVERSITY OF TEXAS MEDICAL BRANCH**
**AT GALVESON**
**June 1, 2010**

## Dedication

For my grandfather, Auline R. Schrank.

We are much the same.  I miss you.

**Acknowledgements**

Many thanks to my mentor, Drs. Bolen and Hilser, who have taught me what it means to be a scientist. Thanks to Dr. Jorg Rosgen and Mikhail Sinev for their expert instruction. Thanks to W. Austin Elam for being a great colleague and a better person.

**Investigations of Local Unfolding Like Conformational Excursions in the Native State of Adenylate Kinase**

Publication No. _____

Supervisors: Davis Wayne Bolen and Vincent J. Hilser

Abstract: Elucidating the interplay between protein structure and dynamics is a prerequisite to an understanding of both function and adaptation in proteins. It has been difficult to experimentally separate these effects because it is challenging to rationally design mutations that will selectively effect dynamics. Here we adopt a mutation approach that is based on a thermal adaptation strategy observed in nature, and we use it to study the binding interaction of *Escherichia coli* adenylate kinase (AK). We rationally design several single-site, surface-exposed glycine mutations to selectively perturb the excited state conformational repertoire, leaving the ground-state X-ray crystallographic structure unaffected. The results not only demonstrate that the conformational ensemble of AK is significantly populated by a locally unfolded state that is depopulated upon binding, but also that the excited-state conformational ensemble can be manipulated through mutation, independent of perturbations of the ground-state structures. The implications of these results are twofold. First, they indicate that it is possible to rationally design dynamic allosteric mutations, which do not propagate through a pathway of structural distortions connecting the mutated and the functional sites. Secondly, the results reveal a general strategy for thermal adaptation that allows enzymes to modulate binding affinity by controlling the amount of local unfolding in the native-state ensemble. These findings open new avenues for rational modulation of protein function and fundamentally illuminate the role of local unfolding in function and adaptation.

## TABLE OF CONTENTS

**LIST OF ILLLUSTRATIONS**

**LIST OF TABLES**

**LIST OF ABBREVIATIONS**

**CHAPTER 1: Enzymes, The State of Knowledge**

**Introduction**

**Enzymatic Adaption to Temperature**

**Enzyme response to Denaturants and Stabilizing Osmolytes**

**Enzyme Dynamics and Adenylate Kinase**

*Structural Changes and Catalysis*

*Time Scale Independent Measures of Conformational Disorder*

*Time Scale Dependent NMR and FRET Measurements Native State Dynamics*

*Ensemble View of the Protein Native State*

*Experimental Goals and Mutation Design*

**CHAPTER 2: Inhibitor Binding and Stability**

**Materials and Methods**

*Preparation of E. Coli Adenylate Kinase*

*Determination of Solubilizing Conditions*

*ITC Experimental Design*

*Circular Dichroism Experimental Design*

**Results and Discussion**

*ITC Data Quality and Comparison to the Literature*

*Linkage Explanation of Enthalpy/Entrophy Compensation AK/Ap5A Binding Reaction*

*Changes in Global Stability with Mutation*

*Solvent Conditions Effects Thermodynamics fo Ap5A Binding*
*Changes in Thermodynamics of Binding with Stabilizing Osmollytes*

**Summary and Conclusions**

**CHAPTER 3: Temperature Dependent and Structural Mapping of Locally Unfolding in AK from E. Coli**

**Introduction**

**Materials and Methods**

*Non-linear Least Squares Analysis of ITC Data*

*Simulating the Three Stae Model of Apo-Adenylate Kinase Unfolding*

*NMR Methods*

**Results and Discussion**

*Fitting of ITC Data*

*Agreement of External Experimental Observables Validates Fitted Model*

*Structural Distribution of Local Unfolding*

*Similarity of the BI State to Local Unfolding*

*Implications of Local Unfolding for the adaption of Binding Affinity*

**Summary and Conclusions**

**CHAPTER 4: Structural Investigation of AK Mutants**

**Introduction**

**Materials and Methods**

*X-Ray Crystallography*

*NMR Methods*

**Results and Discussion**

*Structural Effects of Studied Mutations and Crystal Packing*

*Structural Effects in the Unbound Conformation*
*Propagation of Mutational Effects in the Absence of Structural Change*

**Summary and Conclusions**

**CHAPTER 5: Direct Comparison of the Kinetics of Local Unfolding of the AK LID and the Catalytic Rate**

**Introduction**

**Materials and Methods**

*Measurement of $k_{cat}$*

*CPMG Relaxation-Dispersion NMR and Data Analysis*

**Results and Discussion**

*Measurement of $k_{cat}$*

*Relaxation Dispersion NMR Experiments*

**Summary and Conclusions**

**CHAPTER 6: Computational Investigations of Native State Ensemble of Kinase Alllostery and Ligand Binding**

**Introduction**

**Materials and Methods**

**Results and Discussion**

*Calculation Residue Specific Stability Constants*

*Structural Corollaries to Local Stability*

*Distribution of native Conformers in Term of Δ H and Δ Cp*

*Effects of Ligand Binding on the AK Ensemble*

*Calculation of Residue Specific Allosteric Connectivity*

*Calculation of Calorimetrically Observable Ensemble Effects on Ligand Binding*

**Summary and Conclusions**

**CHAPTER 7: Concluding Remarks**

**Conclusions and Future Directions**

## LIST OF ILLUSTRATIONS

**Figure**

Figure 1.5.1: Mutational Strategy and Binding Associated Conformational Changes in Adenylate Kinase

Figure 2.2.1.1: Representative ITC Titration Data and Model Fitting

Figure 2.2.2.1: Apparent Changes in Thermodynamics of Binding on Mutation

Figure 2.2.2.2: Simulation of Expected Trends of Enthalpy/Entropy Compensation for Folding Linked to Binding

Figure 2.2.2.3: Temperature Dependence of the AK/Ap5A Binding Reaction

Figure 2.2.3.1: Representative Circular Dichrosim Unfolding Experiments

Figure 2.2.3.2: Simulation of Destabilizing Loop Mutation Effects on Protein Global Stability

Figure 2.2.4.1.1: Heat of Binding as a Function of Buffer Heat of Ionization

Figure 3.3.1.1: Fitting of $\Delta H_{conf,app}$ for Comformational Thermodynamic Parameters

Figure 3.3.1.2: Comparison to the Binding Reaction and Temperature Denaturation of a Structurally invasive mutant, s129a

Figure 3.3.2.1.1: Agreement of External Experimental Data with the Fitted Conformational Model and Parameters

Figure 3.3.3.1: $^{1}H$-$^{15}N$ HSQC of WT and v142g AK (33°C)

Figure 4.3.1.1: Alignment of the Asymmetric Unit of WT, v148g, and s129a AK

Figure 4.3.1.2: Alignment Chains from Three Different Crystallographic Environments

Figure 4.3.1.3: Mutations Preserves Structure of Neighboring Residues

Figure 4.3.1.4: Structural conservation in v148g AK

Figure 4.3.2.1: $^{1}H$-$^{15}N$ HSQC of WT and v142g AK (19°C).

Figure 5.3.1.1: Eyring Plot of $k_{cat}$ as a Function of Temperature.

Figure 5.3.1.1: $^{15}N$ Relaxation Dispersion Profiles for Select Residues in the LID Region of v142g AK.

Figure 6.3.1.1: Regional Stability and Structural Corollaries in AK

Figure 6.3.3.1: Analysis of Enthalpy and Heat Capacity Differences of Native State Ensemble Constituents

Figure 6.3.4.1: Effects of Ligand Association on the AK Ensemble

Figure 6.3.5.1: Closed Ensemble Allosteric Response to Perturbation

Figure 6.3.5.2: Open Ensemble Allosteric Response to Perturbation

Figure 6.3.6.1: Predicted Ensemble Effects on the Calorimetric Observables of Ap5A Binding to the Closed Conformation.

Figure 6.3.6.2: Predicted Ensemble Effects on the Calorimetric Observables of Ap5A Binding to the Open Conformation.

# CHAPTER 1

# ENZYMES: THE STATE OF KNOWLEDGE

## 1.1 INTRODUCTION

Enzymes are the ubiquitous mediators of the chemical reactions of biology. Virtually all of biological reactions are enzymatically mediated. Furthermore, enzymes are often the targets of bioactive compounds and commercially available pharmaceuticals. As such, the fundamental importance of detailed and predictive understanding of enzymatic function is critical to the progress of biology and medicine in general. Much progress has been made in the understanding of the structure and organic mechanism of many enzymes. Such progress has even allowed some to even computationally design novel enzymes (1). However, early investigation into enzyme function focused on the kinetics of enzymatic reactions and the thermodynamic properties of enzyme substrate interaction (Dixon Webb, Segel). Somewhat paradoxically, a predictive physical understanding of such parameters is all but absent from the literature. In fact, most physical understanding of catalytic rates and association constants has been inferred from purely biological investigations of enzymatic adaptation (2, 3). The present inability of the field to rationally modulate classical enzymological parameters by mutation, highlights the true shallowness of our understanding of such basic phenomina. To

the best of our knowledge, no study to date has rationally designed mutations based on a mechanistic/biophysical hypothesis that were able to increase the catalytic rate of an enzyme.

The somewhat unspoken consensus in the field is that our lack of understanding of native state dynamics is at least one of the critical barriers to true physical understanding of enzymatic function. Of late, many investigators have taken up this challenge (4-8). However, even accepting this hypothesis much is still to be learned. Many key questions might be answered by studies that investigate the physical nature the individual conformations that by dynamically inter-converting, constitute biologically relevant native state dynamics. Also, as critical are experiments that will probe the nature of the transition state ensemble that governs the kinetics of these dynamic inter-conversions (9, 10).

## 1.2 ENZYMATIC ADAPTATION TO TEMPERATURE

Much insight into enzyme function has been gained based on investigations of homologous enzymes which are adapted to function under various physical conditions (2, 5, 11-14). From this literature many important generalizations can be made. The classical kinetic parameters Km and $k_{cat}$ are generally highly conserved at physiological conditions (15). Furthermore, the sequence composition and structural integrity of the active site are highly to perfectly conserved (16, 17). Therefore, evolution has revealed a robust and intriguing property of enzymes, namely that mutations remote from physical contact with the substrate can mediate

changes in catalytic rate and binding affinity. The exact origins of these mutational effects are as of yet unknown.

Several lines of evidence support a hypothetical relationship between native state flexibility or dynamics and enzyme functional adaptation. From an enzymological perspective, the activation energy barrier for a catalytic reaction is the key determinate of the catalytic constant, $k_{cat}$. According to absolute rate theory this energy barrier can be taken as a free energy barrier ($\Delta G^{\ddagger}$), and therefore split into enthalpic and entropic contributions, i.e. the activation enthalpy ($\Delta H^{\ddagger}$) and activation entropy ($\Delta S^{\ddagger}$) of the reaction (18). Studies of the temperature dependence of $k_{cat}$ have revealed that psychrophilic enzymes as compared to mesophiles invariably decrease the unfavorable $\Delta H^{\ddagger}$, but also suffer a counter acting increase in unfavorable $\Delta S^{\ddagger}$. Thus, $k_{cat}$ is increased but not to the degree that the change in $\Delta H^{\ddagger}$ would suggest. Somero and colleagues have used this thermodynamic data to set down a general theory of enzymatic adaptation based on an ensemble view of the protein native state flexibility (3).

This theoretical framework was presented in the context of investigations of lactate dehydrogenase from nothenoid fishes (3), but has been quite generally applied and heavily cited. In this work, Fields et al. present enzymological and sequence data for a set of orthologous lactate dehyddrogenases (LDHs). Cold adapted LDHs were found to have sequence changes, notably several mutations to glycine, which suggested increased structural flexibility (3). As this hypothesis is foundational to most modern investigation into enzyme function/dynamics relationships it merits restatement here. As eloquently stated by Fields et al. (3):

"The flexibility in any enzyme, including A4-LDH, necessary for function causes each molecule to occupy an ensemble of conformational states. A change in temperature is likely to alter the number of states available, the rate at which molecules shift among these states and the time spent in any given state. Starting from these basic principles, we postulate that differences among orthologs in $K_m$ and $k_{cat}$ derive from inherent differences in conformational entropy, that is, in the number of conformational states accessible to an ortholog at any given temperature.

According to this hypothesis, temperature-adaptive increases in $k_{cat}$ will occur concomitantly with increases in $K_m$. This can be seen by comparing cold- and warm-adapted enzymes at a common temperature. This higher conformational mobility required for rapid catalysis (high $k_{cat}$) in the cold-adapted form will lead to a higher number of conformational states available to the molecule and, as a result, to a larger proportions of enzyme conformational that bind substrate poorly or not at all. This will yield a higher $K_m$ in the cold-adapted enzyme. Conversely, at this common temperature, the enzyme adapted to function at high temperatures will have less inherent conformational flexibility, the proportion of binding-competent states will be higher, and $K_m$ (as well as $k_{cat}$) will be lower.

The inherent differences in conformational flexibility between cold-and warn-adapted enzymes also will be manifested in their responses to acute temperature changes. Q10 is a direct measure of the rise in $k_{cat}$ with temperature and reflects thermal effects on rate-limiting conformational changes. Q10 values of cold-adapted orthologs are lower than those of warm-adapted ones, indicating lower activation enthalpies ($\Delta H^{\ddagger}$) that we argue are caused by increased localized flexibility. Conversely, activation entropies ($\Delta S^{\ddagger}$) for cold-adapted enzymes are higher, revealing the greater degree of ordering these orthologs must undergo to form the ternary complex."

In summary, adaptive changes in sequence residing primarily outside of the active site have led many to a view here protein conformational dynamics are the dominant factor in determining enzyme kinetic parameters. As such, much present day research, both computational and experimental is focused on this approach to uncovering the physical determinants of enzyme kinetics.

## 1.3 ENZYME RESPONSE TO DENATURANTS AND STABILIZING OSMOLYTES

Other key evidence linking enzyme conformational dynamics to activity has come from the observation of strong effects of denaturants and stabilizing osmolytes on enzyme function. Namely, urea and trimethylamine N-oxide (TMAO) are small molecules called osmolytes that are known to have a large effect on protein stability. These molecules affect protein stability by non-specific interactions with the peptide backbone and side chains (19, 20). The interaction with the backbone has been demonstrated to be the dominant effect (19, 20). Urea, a denaturant, interacts favorably with the peptide backbone, thus making transitions that expose a great deal of backbone surface area to the solvent more probable. TMAO, a counteracting osmolyte, has an unfavorable interaction with the backbone, thus favoring states with the maximal amount of backbone surface area hidden from the solvent. Hydrogen/Deuterium exchange (HX) and Nitrogen-15 ($^{15}N$) relaxation experiments have also proven that urea shifts the native ensemble toward more solvent exposed and conformationally diverse states; TMAO has the opposite effect (21, 22). The current hypothesis is that osmolytes affect structural

fluctuations by the same mechanisms known to effect global protein stability. Osmolytes are well known to have dramatic effects on enzyme function (22, 23). TMAO generally decreases $K_m$ (increases binding affinity) and increases $k_{cat}$ (23). Urea generally has the opposite effects. These effects are hypothesized to arise from shifting the native ensemble to states with differing functional properties, such as binding affinity and catalytic rate (23). These trends, however, are not without exception. Interestingly, urea *increases* the $k_{cat}$ and $K_m$ of rabbit muscle adenylate kinase (24). These changes seemingly recapitulate the expected effects of increased conformational dynamics observed for cold-adapted enzymes that are affected by changes in sequence. Accordingly, the authors reporting these findings interpret the kinetic changes imparted by increasing urea concentration as arising from increased flexibility of the structures surrounding the active site (24).

## 1.4 ENZYME DYNAMICS AND ADENYLATE KINASE: THE STATE OF KNOWLEDGE

The model protein used in our investigations of the dynamics/function relationships in the present work is adenylate kinase from E. Coli. This protein has been the subject of investigation with virtually all major structural and biophysical approaches presently applied to proteins. Therefore, the current physical picture of native state dynamics will be summarized in terms of the literature reported for adenylate kinase.

### 1.4.1 Structural Changes and Catalysis

Clear structural differences between liganded and un-liganded forms of proteins have been observed since the genesis of bio-molecular crystallography. Such observations are exemplified by the two well conformational states observed crystallographically for AK. In the unliganded crystal structure of AK the active site residues are highly exposed to the solvent, providing a mechanism for substrate dissolved in the solvent to access the active site (25). The crystal structure of AK in complex with a bi-substrate analog inhibitor, $P^1$,$P^5$-Di(Adenosine-5')Pentaphosphate (Ap5A), demonstrates that binding is associated with a conformational rearrangement of residues surrounding the active site. Such rearrangements totally occlude any pathway for substrate to escape the binding pocket, without atoms passing through atoms (26). Although the effect on the accessibility of the active site is great, only two relatively small regions of structure are shifted in this conformational change. The mobile region associated with the adenosine-5'-triphosphate (ATP) binding site has been termed the "LID" domain, and that associated with the adenosine-5'-monophosphate (AMP) binding site, the "AMP binding domain (AMPbd)." In reference to the effect on the active site these conformations are often referred to as "open" and "closed." Because of the need for cycling of reaction substrate and in and out of the active site, many have reasoned that AK must cycle through these conformations to perform catalysis.

### 1.4.2 Time Scale Independent Measures of Conformational Disorder

Modern hydrogen/deuterium exchange (HX) experiments are performed by preparing protonated protein sample spin labeled with $^{15}N$, and then moving the protein into a high concentration $D_2O$ solution. In such an experiment $^{1}H$-$^{15}N$ spin pairs in the backbone amide groups provide a signal for each residue of the protein as observed in a two-dimensional hetero-nuclear single quantum coherence (HSQC) experiment. Interestingly, once moved into deuterium oxide ($D_2O$) these signals decay in intensity as protons in the protein are replaced (exchanged) with deuterons in the solution. In general, highly buried amide proton in the protein structure exchange more slowly than those on the protein surface. One interpretation of this effect is that only a small population of conformations in the native state expose such amides to the solvent.

Furthermore, since the intrinsic rates of hydrogen exchange for exposed amides is well known, the apparent modulation of these rates by structure can be (under certain conditions) analyzed to estimate the probability, or free energy, of such highly solvent-exposed states, at every position along the protein backbone (27). HX experiments on AK have revealed that the LID and AMPbd exchange rapidly with the solvent (28). This suggests that native AK experiences conformational excursions in these regions, of relatively high probability.

Another time independent measurement that has been interpreted in terms of native state dynamics is crystallographic temperature factors (B-factors). Such parameters define the volume (usually a sphere or ellipsoid) in which an atom's electron density is distributed are estimated as a part of the structure refinement process. Thus, atoms where the density is distributed over a larger volume are in a

sense more disordered. This apparent disorder can arise from crystal imperfections, model imperfections, static variations in the structure of individual molecules, or from more interesting thermal fluctuations of the protein (29).

In the case of AK the relative temperature factors of different regions of the protein seem to change between the open and closed conformations (25). More specifically, regions surrounding the N and C termini, which is remote from the binding site, become more disordered upon closing (25). The authors of this work hypothesize that these increasing fluctuations of this region of the protein somehow mitigate the energy difference between the bound and unbound states of the molecule. Therefore they have termed the region an "energetic counterweight"(25). A more conservative interpretation would be that these data agree with that from HX, in that the native state of AK appears to be composed of a set or ensemble of many conformations which surround the average structure of the molecule as typified by the crystal structure.

### 1.4.3 Time Scale Dependant NMR and FRET Measurements of Native State Dynamics

The obvious disadvantage of the above-discussed measures of native state disorder is the lack of kinetic information. This is a critical omission, as biologically important conformational phenomenon must take place on a timescale that is relevant to protein-mediated processes in the cell, such as binding and catalysis. Therefore, more recent investigations of protein dynamics have focused on nuclear

magnetic resonance (NMR) relaxation experiments that can provide such kinetic information.

$^{15}N$ relaxation experiments have been applied to many proteins as a method to examine internal pico- to nano-second motions of backbone amide groups (30). The results of these studies allow determination of the squared order parameter ($S^2$), which is a measure of the distribution of angular orientations of a particular bond vector, with respect to the molecular frame. Values range from one, representing a fixed vector, to zero representing completely isotropic distribution. Such studies of AK have revealed that on this time scale the ligand-associated sub-domains are highly dynamic, in both the free and complexed state (10, 31). Notably, the outermost portions of the LID domain (~ residue 150) and AMPbd have much lower values of $S^2$ (squared order parameter) as compared to the remainder of the protein (10). Furthermore the correlation time of the motions of these residues was found to be faster on average than the core of the protein (31).

More interesting still are relaxation-dispersion experiments, which are able (in certain cases) to directly measure the rates of dynamic conformational processes on the micro- to millisecond time scale (32, 33). The significance of such experiments is that many important bio-molecular processes take place on this timescale, for example catalysis is usually a millisecond phenomenon. Such experiments have been recently applied to AK in effort to quantify the rates of the open/close conformational reaction discussed above (5). Interestingly, these authors find that the opening rate is coincident with $k_{cat}$ for both E. Coli AK and a hyper-thermophilic ortholog as well (5). The authors infer from this finding that the

necessary conformational re-arrangements required for catalysis to proceed may indeed be highly rate limiting, much as Somero and colleagues hypothesized (3, 5).

Single molecule Forester Resonance Energy Transfer measurements (FRET) also have the advantage of providing information about the kinetics of protein structural dynamics. The notable disadvantage of this technique is the requirement of attaching (rather bulky) flourophors to the protein. FRET experiments on AK have monitored the conformational motions of the LID domain of AK (34). The LID opening rate was found again by this methodology to be on the approximate timescale of catalysis. Also, important information revealed by FRET is the measurement of relatively large intra-molecular distances. In this manner, Hanson et al. also demonstrate that the presence of substrate not only shifts AK further towards the closed conformation, but also decreases the width of the distance distribution surrounding both open and closed states (34). This finding is important because it suggests that structural conformations such as open and closed in AK are indeed themselves complex collections of inter-converting microstates.

### 1.4.4 Ensemble View of the Protein Native State

From the above-discussed methodologies, a large volume of evidence has accumulated suggesting that the native state of proteins is a heterogeneous collection of states. What is more elusive, however, is a clear physical picture of what the constituent states are, and how they are harnessed by proteins to influence/allow biological function. One approach to this problem has been to consider the functionally relevant native state of proteins to be approximated by the

set of all possible states of local unfolding of the protein (35, 36). Such an approach has been successful in agreeing with and even explaining many intuitively elusive properties of the native state, such as HX protection factors, allosteric communication within protein domains, and the interaction between pH and denaturants effects on protein global stability (37-40).

## 1.5 EXPERIMENTAL GOALS AND MUTATION DESIGN

Therefore, considering the relative success of this ensemble model, we have designed experiments to probe the native ensemble of AK for states that are similar to local unfolding. Our goals, in general terms, are to investigate how similar fluctuations of the native state are to unfolding reactions, and to uncover the functional significance of such local unfolding for catalysis in AK. The goal of our mutation strategy, therefore, is to selectively probe the conformational manifold of states that contribute to the dynamics of the LID domain, in search of states that are both similar to local unfolding and relevant to substrate binding. To accomplish this, we have carefully chosen mutations that should promote locally unfolded states, without disturbing the structural properties of fully folded ground state.

We note that, in principle, the promotion of unfolded states could be effected purely by changes in the disordered state itself, with no energetic perturbation to the ordered conformation. A good example of this principle is the thermodynamic mechanism by which denaturants destabilize protein folds (41). Urea, for example, stabilizes *both* the native and denatured states of proteins in water. The apparent

destabilization of the folded conformer is only a result of the greater stabilization of the denatured state, due to the much larger amount of surface area exposed therein (41). To effect such a change by mutation, we must consider which properties of amino acids are greatly expressed in the unfolded state and weakly expressed in the folded conformation. In the unfolded state, peptides randomly search through conformational degrees of freedom available because of the presence of many rotatable bonds. The restriction of such conformational freedom is manifest as in the very large unfavorable entropy of protein folding (42). The beta-branched amino acids, valine (Val) and isoleucine, are known to have a relatively small decrease in conformational entropy upon unfolding, which is due to steric collisions of their bulky side chains in the denatured state (42). The absence of a side chain in glysine (Gly), on the other hand, results in a comparatively large increase in the conformation entropy of unfolding (42). Therefore, Val to Gly mutations should, in principle, decrease the stability of a protein by increasing the conformational entropy of the denatured state.

Val to Gly mutations at various surface exposed positions in the LID domain were therefore selected. This type of mutation is expected to increase the conformational entropy associated with unfolding this region of the protein, if indeed locally unfolded states are relevant. We have defined several criteria for selecting mutation sites such that the effects should be exclusively in the conformational entropy, as described above, rather than in a perturbation of the structure of the fully folded state of the protein. The criteria were: 1) the side chains of the mutated residues should be highly surface exposed, involving few to no intra-

molecular contacts, 2) the side chain should contain no important moieties such as charged or highly polar groups, which could contribute to long range interactions, and 3) the mutated sites should be greater than 8 Å from the active site, thus making no contact with the ligand. As no stabilizing contacts within the native fold or between the protein and ligand will be disturbed by such a mutation, little or no perturbation should be made to the structure of the bound complex. Three positions were selected in the LID domain of AK that satisfy these three criteria: v142g, v148g, and v135g.

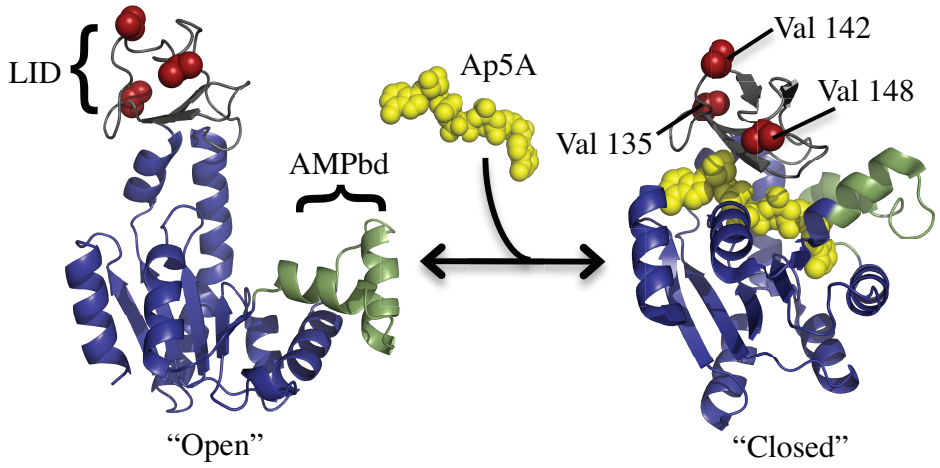



**Figure 1.5.1: Mutational Strategy and Binding Associated Conformational Changes in Adenylate Kinase.**

AK from *E. coli* is a much studied enzyme, that catalyzes the reaction 2ADP ↔ ATP + AMP. ***Open***. Graphic of Apo-AK (PDB ID 4AKE). ***Ap5A***. $P^1$,$P^5$-Di(Adenosine-5')Pentaphosphate, a non-hydrolysable bi-substrate analogue inhibitor of AK. ***Closed***. Graphic of the Ap5A/AK complex, PDB ID 1AKE. ***LID*** (grey). The "LID domain" is a highly dynamic structural appendage associated with the ATP binding site, as defined by Shapiro et al. ***AMPbd*** (green). The "AMP binding domain" is a highly dynamic structural appendage associated with the AMP binding site, as defined by Shapiro et al.(29) ***Red Spheres***. Indicate the location of selected mutations sites.

# CHAPTER 2

# INHIBITOR BINDING AND STABILITY

## 2.1 MATERIALS AND METHODS

### 2.1.1 Preparation of E. Coli Adenylate Kinase

#### 2.1.1.1 *Expression System*

The plasmid, PEAK91, was obtained from Dr. Mickahail Sinev of the D. Wayne Bolen laboratory of UTMB (43). PEAK91 contains the AK sequence downstream from its native promoter in E. Coli. Transformation of competent E. coli was performed as recommend by the heat shock method. Protein expression was found to be quite unreliable in BL21 cells. Yields varied from 250-30 mg/L of LB media. Outgrowths were performed for 14hours at 37C with no induction. To solve this problem several cell lines were investigated. HMS174 were found to be genetically most similar to the K12 bacteria that the plasmid was originally used with. HMS174 reliably yields 100 mg/L AK using the above protocol.

#### 2.1.1.2 *Mutagenesis and Purification Procedures*

Mutagenesis was performed with Quick Change™ kit, obtained from Stratagene. Mutations were confirmed by sequencing and mass spectroscopy. Bacterial out growths were performed as above, using approximately 6-8L LB per

preparation. The general purification scheme utilized has been previously described in the literature(43, 44). However, for the purpose of reproducibility the exact conditions applied will be recorded here. E. coli cell pellets were re-suspended in 50mM Tris (pH8), 100mM NaCl, 2mM EDTA, 2mM DTT, and disrupted by sonication for a total exposure time of 10 minutes. A disposable 50ml column of pre-swollen DEAE-cellulose from Whatmann™ was used as an initial fractionation step. Lysate cleared of cell debris was located onto the column equilibrated in the above disruption buffer. The protein was passed through the column at 2ml/min and initial 100ml of unbound species was retained. The retained volume was diluted 2:1 by volume with 10mM Tris (pH 8), 1mM DTT, to reduce the ionic strength. Protein was then loaded onto a Cibracron Blue F3GA Agarose column equilibrated in 50mM Tris (pH 7.5), 0.1 mM EDTA, 1mM DTT. When unbound species were cleared from the column with a flow rate of 2ml/min. Elution of AK was then performed with a linear gradient elution, increasing the final concentration of NaCl to 2M. Fractions containing AK by SDS/PAGE electrophoresis were then pooled and dialyzed against 4L 20mM Tris (pH 8), 0.1 mM EDTA, 1mM DTT for 12-24 hours. A Resource-Q ion exchange column was equilibrated in the above dialysis buffer. The protein was loaded to the column, and linear gradient elution, increasing the NaCl to a final concentration of 0.3 M, was used to bring the protein into solution. AK was stable for months at -20°C in 25mM Na-phosphate (pH 7), 0.2 M NaCl, 2mM EDTS, 2mM DTT, 50% weight per volume glycerol. Protein purity was assayed by checking for a single band of SDS-PAGE of the final product.

### 2.1.2 Determination of Solubilizing Conditions

Initial investigations using Calorimetry (ITC) demonstrated gross aggregation at 1mg/ml AK solution within the calorimeter. Such aggregation may well be the factor delaying calorimetric investigation of this classic biophysical model system. To address this issue, thermal unfolding experiments were carried out at 0.1 mg/ml AK concentrations, with various salt concentration, buffer, and pH. Grossly observable aggregation as well as the reversibility of thermal unfolding as assayed by return to baseline Circular Dichroism (CD) signal was used to determine the most permissive solution conditions. Of all parameters assayed, pH seemed to be the most critical. pH > 7.75 was determined to be compatible with reversible thermal denaturation with a maximal temperature of ~70°C.

### 2.1.3 ITC Experimental Design

Isothermal Titration Calorimetric experiments were carried out using the MicroCal VP-ITC system. Solution conditions were 60 mM PIPES acid, 2mM EDTA, and pH 7.85, unless otherwise noted. Protein was prepared by dialyzing into the above buffer 3 times over 18 hours, and diluted to a final concentration of ~0.04 mM. Concentration was measured by absorbance at 277 nm, 0.55 $AU_{277}^{-1}Molal^{-1}cm^{-1}$, or 12,925 $AU_{277}^{-1}M^{-1}cm^{-1}$. Ap5A was then dissolved in the final dialysis buffer to a final concentration of ~0.35 mM. Concentration was measured using absorbance (30.8AU/M/cm). pH measurement was performed to ensure pH of the ligand solution was not perturbed by Ap5A.

### 2.1.4 Circular Dicrhoism Experimental Design

Variable temperature Circular Dichroism (CD) experiments were carried out using 0.1 mg/ml AK in the above-described ITC conditions with a 1mm cuvette. Optical rotation was recorded at 228 nm from 20-70°C. All data were fit two a two state thermodynamic model, as according to the modified Gibbs-Helmholtz equation. CD signal for the folded or unfolded states was taken to be linear with temperature. The parameters fit to each thermal unfolding curve were: ΔH, ΔS, ΔCp, as well as two baseline slope/intercept pairs.

## 2.2 RESULTS AND DISCUSSION

### 2.2.1 ITC Data Quality and Comparison to the Literature

Using the above protocol, high quality ITC titration data was achieved. Figure 2.2.1.1 demonstrates the quality of the data and agreement of the data with the two state binding model provided with the MicroCal instrument. Table 2.2.1.1 summarizes the results of fitted parameters analyzed for each ITC titration experiment, at 37°C. From this data and analysis it is clear that not only can interpretable ITC binding curves be obtained for the AK/Ap5A interaction, but also all fitted parameters are well defined.

Fitted values of number of sites (N) ranged from 0.95-1.08, at temperatures below 30°C (where non-reversible aggregation was not observed), thus strongly confirming not only one to one binding but also the correct quantification of the protein and ligand Ap5A. The applied extinction coefficient (277nm) is not identical to that used in the literature i.e. (12,925 $M^{-1}cm^{-1}$ vs. 10,000 $M^{-1}cm^{-1}$). The rigorous

agreement of our extinction coefficient with the fitting analysis from ITC and the >10% disagreement between the two reported values undermines the key finding of Wolf-Watz et al. (5). More specifically, the measurement of $k_{cat}$ is directly dependant on the measurement of protein concentration. Therefore, one should expect that the catalytic rate of AK is actually less than reported in this paper (5). The key finding of this work was that the conformational "opening" rate measured by Carr-Purcell-Meiboom-Gill (CPMG) relaxation-dispersion methods is coincident with the catalytic rate of AK (5). Therefore, the opening rate is most likely slightly faster than the actual catalytic rate, indicating a much more reasonable conclusion that AK is not perfectly efficient. In other words, most but not all substrate dissociation events are associated with the release of forward reaction products.

### 2.2.2 Linkage Explanation of Enthalpy/Entropy Compensation AK/Ap5A Binding Reaction

At 37°C (and above) all mutations decreased binding affinity. However, the mutations induce relatively small changes in free energy of binding in comparison to a large and opposing changes in enthalpy and entropy (Figure 2.2.2.1). Such entropy/enthalpy compensation phenomena are common in biophysical chemistry. Actually, a quite extensive literature explores the meaning of such findings (45, 46). In particular, binding reactions are often found to have a ΔG, which is robust to perturbations to the protein or ligand in question, while radical changes in ΔH and ΔS are effected (47-50). However an rarely discussed, but applicable and simple explanation of such phenomena is a conformational equilibrium linked to the

binding reaction (51). Eftink et al. have elegantly demonstrated how perturbations in conformational equilibria, which are common in proteins, can account for such compensation induced by either direct perturbations to the macromolecule or by a series of cognate ligands (51).

In the case of the mutant AKs under study, the obvious choice of model would be a perturbation to a conformational equilibrium within the protein molecule itself. The simplest model which could explain our data, is a mandatory conformational coupling model, where the ligand has significant binding affinity for only one of two possible states of the macromolecule (51). The equations describing the expected thermodynamic observations of a binding reaction for such a system are available in the literature (51):

$$K_{app} = \frac{K_0}{1 + K_c} \qquad \text{Equation 2-1}$$

$$\Delta H_{app} = \Delta H_0 - \Delta H_c \left( \frac{K_c}{1 + K_c} \right) \qquad \text{Equation 2-2}$$

$$\Delta S_{app} = \Delta S_0 - R \ln(1 + K_c) - \Delta H_c \left( \frac{K_c}{1 + K_c} \right) \qquad \text{Equation 2-3}$$

The results of the above equations are plotted in Figure 2.2.2.2, perturbing only $\Delta S_c$, i.e. the entropy change of the linked conformational process. This demonstrates the observed effect of perturbing only the conformational equilibrium, revealing large changes in $\Delta S_{app}$ and $\Delta H_{app}$, which might be mistakenly interpreted as compensation intrinsic to the association event itself.

It is well known that protein folding in general is an entropically unfavorable process, in which many conformational degrees of freedom within the peptide must be constrained. Therefore in general, folding must be enthalpically favorable in conditions where the protein is stable. Such trends have been quantitatively parameterized for arbitrary proteins with known sequence and structure (52, 36). This along with the above theoretical considerations suggest that our data are consistent with a model where the mutations in question promote a locally unfolded state having little to no binding affinity, and the addition of ligand promotes the folding of this region through a linked process as discussed above. Furthermore, it seems that similar effects are mediated by all mutations, suggesting that the mutational effects are brought about by a shared mechanism. Figure 2.2.2.3 shows the fitted values of $\Delta H$ and $\Delta G$ plotted as a function of temperature. From this plot, it is clear that $\Delta H$ is non-linear with temperature, and that the trend observed is similar for all mutant AK. This is also consistent with a linked folding reaction (53). Taken together these qualitative results strongly suggest that a shared conformational reaction present in all mutant proteins where one state has little or no affinity for Ap5A. This hypothesis will be quantitatively investigated in following sections.

### 2.2.3 Changes in Global Stability with Mutation

Circular dichroism temperature scan experiments were used to measure the thermal stability of AK and the glycine mutants, see Figure 2.2.3.1. The results of two state fits to the data are recorded in Table 2.2.3.1. From these data it is clear

that the mutations in question make a quite minimal perturbation to the global stability of the molecule. This result is unexpected, as the entropic destabilization of a valine (Val) to glycine (Gly) mutation on a folding reaction are well known. This is consistent, however, with the idea that most of the entropy conferred by the mutations is expressed in the native state prior to the global thermal unfolding transition. This principle has been previously discussed in the literature (54, 55). However, to demonstrate this principle here we can simulate the properties of a thermodynamically three state unfolding system using the following linkage equations,

$$Q = 1 + K_{loc} + K_{loc} \cdot K_{unf} \qquad \text{Equation 2-4}$$

$$P_{loc} = \frac{K_{loc}}{1 + K_{loc} + K_{loc}K_{unf}} \qquad \text{Equation 2-5}$$

$$P_{glob} = \frac{K_{loc}K_{glob}}{1 + K_{loc} + K_{loc}K_{glob}} \qquad \text{Equation 2-6}$$

$$K_{loc} = e^{\frac{\Delta G_{loc}}{-RT}} = \frac{P_{loc}}{P_{native}} \qquad \text{Equation 2-7}$$

$$K_{glob} = e^{\frac{\Delta G_{glob}}{-RT}} = \frac{P_{glob}}{P_{loc}} \qquad \text{Equation 2-8}$$

where the temperature dependence of the ΔG terms defined by the modified Gibbs-Helmholtz equation,

$$\Delta G\ (T) = \Delta H \left( \frac{T_m - T}{T_m} \right) - (T_m - T)\Delta C_p + T\Delta C_p \ln\left( \frac{T_m}{T} \right) \qquad \text{Equation 2-9}$$

with ΔH, $T_m$, $\Delta C_p$ values corresponding to each unfolding transition in the model,

*local* and *global*. Figure 2.2.3.2 plots the populations of the globally unfolded states as a function of temperature, and $T_{m,\ local}$ for a hypothetical wild type and mutant protein where the mutation selectively destabilizes the region of the protein able to undergo local unfolding. Published values for the entropic destabilization expected for a Val to Gly mutation were applied (52). From this simulation we can see that where $T_{m,\ local}$ is equal to or greater than $T_{m,\ global}$ the mutation effects a large change in the apparent melting temperature of the protein ($T_{m,\ apparent}$). However, where $T_{m,\ local}$ is much less than $T_{m,\ global,}$ the change in ($T_{m,\ apparent}$) approaches zero. Therefore, it can be seen that mutational effects on protein stability in general are a function not only of the energetic cost of the mutation itself, but also the relative local stability at the mutation site.

Therefore, the small magnitude of the shift of $T_m$ suggests the mutated region of AK may well be unstable in the native state. Also in agreement with this assertion is the relatively large reduction in the apparent enthalpy of unfolding the mutant proteins, which suggests a change in the apparent cooperative unit of unfolding AK. This change is furthermore on the same scale of magnitude as the gross changes in heat of binding. Suggesting that these two results may well be reflecting different aspects of the same physical phenomenon.

Furthermore, it is interesting to note the small magnitude of change of the baseline for the folded state, i.e. 20-50°C, between the wild type (WT) and mutant proteins. From this data, it appears that the native fold of the protein is grossly unperturbed by the mutation. The intuitive disagreement the relatively large thermodynamic changes in the binding reaction with this data will be addressed

below. However, some small change in the slope of the baseline is consistent with the hypothesis that the mutations have changed the probability of locally unfolded states that are an important component of the native state ensemble.

### 2.2.4 Solvent Conditions Effects Thermodynamics of Ap5A Binding

#### 2.2.4.1 *Contributions of Ionization to $\Delta H_{app}$*

It is common for ionization to be linked to biological binding reactions. As such, it is important to investigate whether the linked conformational process in question can be given such trivial explanation. In this case proton transfer will effect the observed ITC data by the following equation:

$$\Delta H_{app} = \Delta H_0 - \Delta H_{ion}\left(n_{H+}\right) \qquad \text{Equation 2-10}$$

where $\Delta H_o$ is the heat associated with the intrinsic binding reaction, $\Delta H_{ion}$ is the heat of ionization of the buffer in use, and $n_{H+}$ is the number of protons transferred per binding reaction. Therefore, experiments in multiple buffers were performed to deduce the relative contribution of ionization to the measured ΔH of binding, and the excess heat associated with the conformational process in question, see Figure 2.2.4.1.1. Equation 2-10 was fit to the data by the least squares method. $n_{H+}$ fit for the WT and v142g proteins were -0.04±0.07 and -0.17±0.17, respectively. Therefore considering the small heat of ionization of the buffer used for most experiments (PIPES) and these values, it is clear that ionization is a minor to negligible contributor to the reported heats of binding.

### 2.2.5 Changes in Thermodynamics of Binding with Stabilizing Osmolytes

Several natural occurring osmolytes (e.g. Sarcosine, Betaine, Sorbitol, Tri-methyl Amine Oxide) are well known to stabilize the native fold of proteins (20). This effect is predominantly caused by a larger free energy penalty for transferring peptide backbone surface area to from the core of the protein to a solvent (water) containing the osmolyte vs. water alone (20, 19). As such, conformational transitions involving large changes in surface area exposure of the peptide backbone, such as folding, are the most sensitive to osmolytes. To qualitatively determine whether the conformational transition in question is associated with significant changes in solvent exposed surface area, we have performed ITC titrations in the presence of 2.5 molar sarcosine.

These data, summarized in Table 2.2.5.1, show several interesting trends. First, the overall affinity of AK for Ap5A is reduced in the presence of sarcosine. As sarcosine is known to stabilize more compact structures of proteins, this result is slightly counterintuitive. However, the importance of the solvent in defining the bound and un-bound states cannot be overstated. For example, salt is known to radically decrease the affinity of AK for Ap5A. This is probably due to screening of the charge/charge interactions in the binding interface. At high concentrations, such as 2.5 molar, even a relatively small proportion of ionized sarcosine is likely to have a strong impact on charge-charge interactions as well.

More interestingly, we can see that the difference in the binding affinity between the wild type and mutant, $\Delta\Delta G_{val\text{-}gly}$, is decreased. This is interesting because, solvent effects on such as charge screening should be manifest in the intrinsic term, which should remain identical in both AKs (54). The decreased

difference is consistent with increasing the stability of the binding competent state. This reduced $\Delta H_{app}$ of the mutant relative to the WT protein in the presence of sarcoine is also consistent with depopulating the BI state.

## 2.3 SUMMARY AND CONCLUSIONS

From the synthesis of these data, we have strong reason to infer that the thermodynamic changes observed upon mutation of AK are the result of promoting change in population of a conformational state of AK that is modulated by inhibitor binding. The large magnitude of the associated enthalpies, the changes in global unfolding, and the response of the system to stabilizing osmolytes are consistent with the promoted conformational state is a locally unfolded state of AK involving the mutated residues. This assertion will be explored in detail in the following chapters.

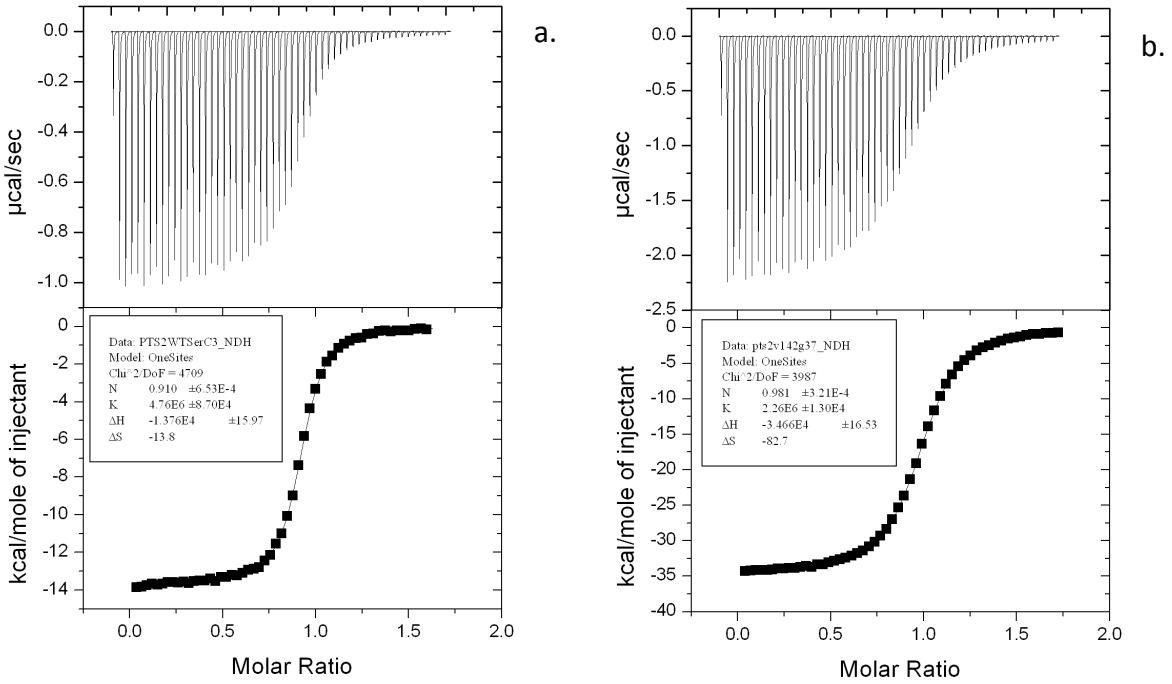



**Figure 2.2.1.1: Representative ITC Titration Data and Model Fitting.**

**a.** Wild-type AK, **b**. v142g AK. Data were gathered under identical conditions: 37°C, 60 mM PIPES acid, 2mM EDTA, and pH 7.85.

**Table 2.2.1.1: Representative Summary of Fitted Parameters for ITC Data.**

| (37°C) | $K_a \cdot 10^{-6}$ | $\Delta G_{app}$(cal/mol) | $\Delta H_{app}$(cal/mol) | $T\Delta S_{app}$(cal/mol) | N[a] |
|---|---|---|---|---|---|
| Wild Type | 4.9±0.14 | -9,490±18 | -13,600±140 | 4,110 | 1.07±0.02 |
| v142g | 2.21±0.035 | -9,001±9.7 | -34,300±440 | 25,299 | 1.03±0.044 |
| v148g | 3.87±0.01 | -9,346±1.6 | -28,920±80 | 19,574 | 1.05±0.02 |
| v135g | 1.69±0.05 | -8,840±18 | -28,490±20 | 19,650 | 1.02±0.032 |

[a]Number of sites, i.e. the stoichiometry of binding.
[±]Errors are standard errors calculated from direct experimental repeats

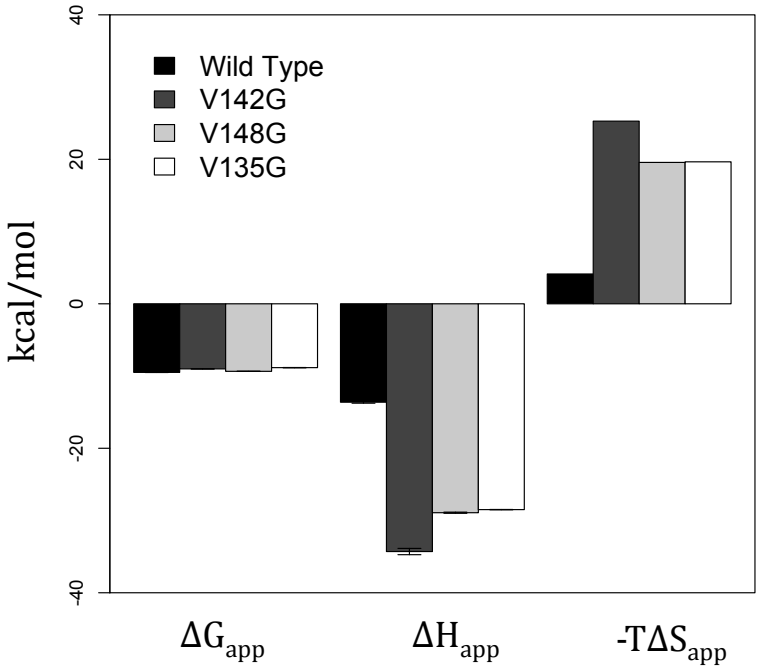



**Figure 2.2.2.1 Apparent Changes in Thermodynamics of Binding Upon Mutation.** *black*-wild type AK, *charcoal*- v142g AK, *grey*- v148g AK, *white*- v135g AK. All data were gathered under identical conditions: 37°C, 60 mM PIPES acid, 2mM EDTA, and pH 7.85.

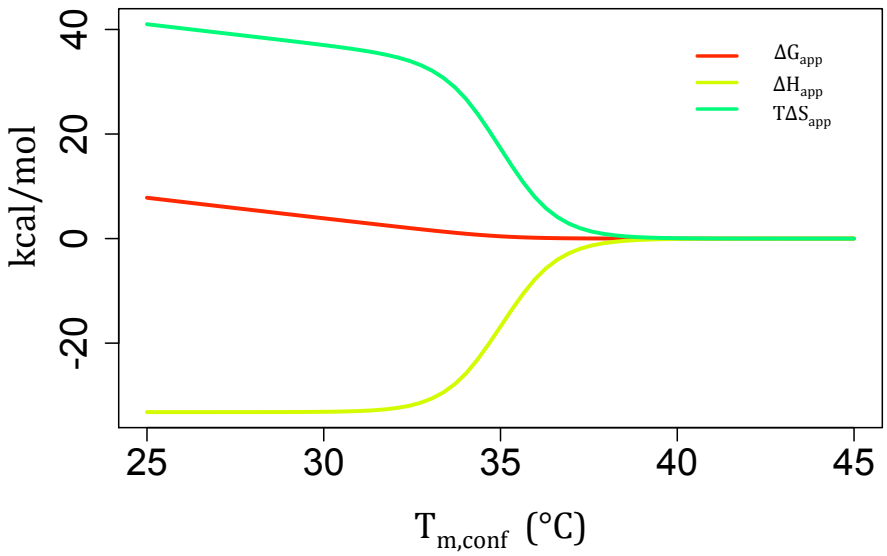



**Figure 2.2.2.2 Simulation of Expected Trends of Enthalpy/Entropy Compensation for Folding Linked to Binding.** *red*-The additive contribution of free energy of a linked conformational equilibrium to the measured free energy of binding ($\Delta G_{conf,app}$), defined as $\Delta G_{conf,app} = -RT\ln(1/(1+K_{conf}))$.*yellow*- The additive conformational contribution of enthalpy to a measured enthalpy of binding ($\Delta H_{conf,app}$), defined as $\Delta H_{conf,app} = \Delta H_{conf}(1/(1+K_{conf}))$, *green*- The additive conformational contribution of -TΔS to the measured -TΔS of binding, defined as $-T\Delta S_{conf,app} = \Delta G_{app,conf} - \Delta H_{app,conf}$. Simulation parameters were as follows: $\Delta H_{conf}(55°C)=33$ kcal/mol, $\Delta Cp_{conf}=660$ cal/mol·K, $T_{m,conf}$ as indicated.

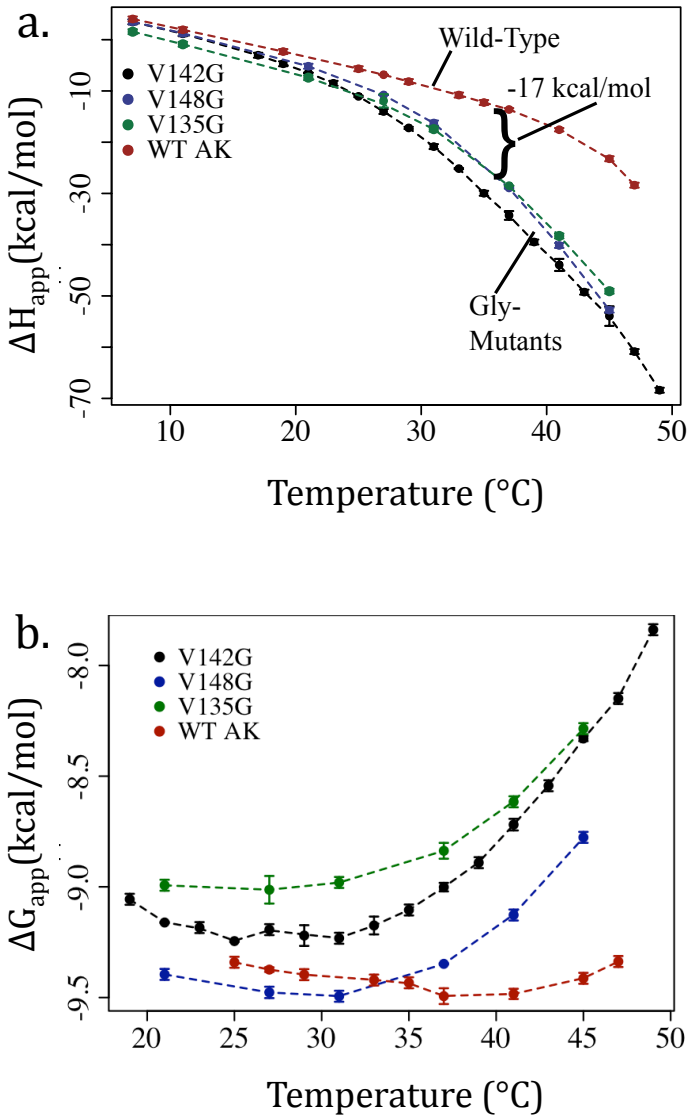



**Figure 2.2.2.3 Temperature Dependence of the AK/Ap5A Binding Reaction.** Data from WT (*red*), v142g (*black*), v148g (*blue*), and v135g (*green*), binding to the inhibitor Ap5A. Temperature was 37°C. See methods for buffer conditions.

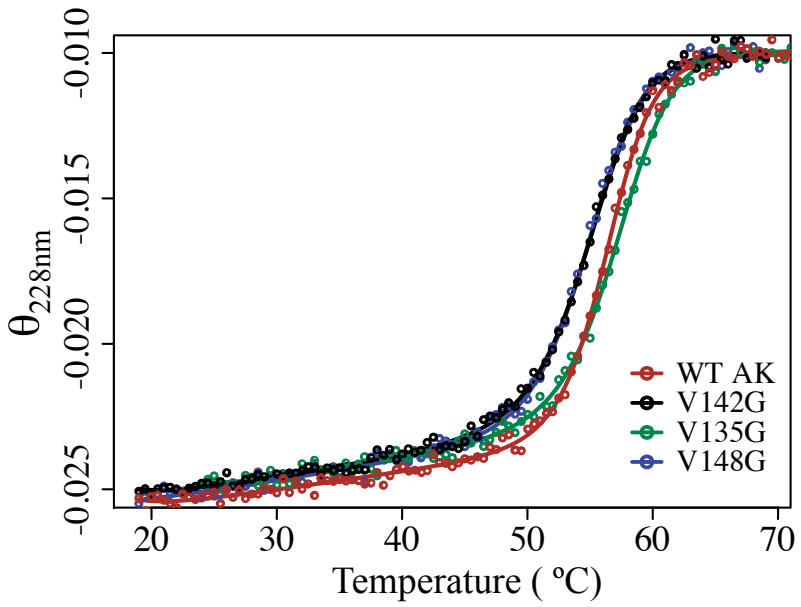



**Figure 2.2.3.1. Representative Circular Dichrosim Unfolding Experiments.** Molar eliptisity at 288nm, $\theta_{288}$, (deg·cm²/dmol·res), recorded by circular dichroism spectroscopy. WT (*red*), v142g (*black*), v148g (*blue*), and v135g (*green*), binding to the inhibitor Ap5A. Temperature was 37°C. See methods for buffer conditions. Lines are fitting functions, representing a two state thermodynamic model. Fitted parameters are recorded in Table 2.2.3.1. Experimental conditions are identical to that used for ITC, see methods.

**Table 2.2.3.1. Two-State Fitting of Temperature Denaturation Data**

| | **ΔH[a]** | **ΔCp[b]** | **Tm[c]** |
|---|---|---|---|
| Wild Type | 115.1±4.6 | 4.1±1.9 | 55.9±0.5 |
| V142G | 95.4±3..5 | 4.1±0.2 | 55.1±0.5 |
| V148G | 94.6±1.2 | 3.8±0.4 | 55.0±0.2 |
| V135G | 97.6±2.8 | 4.6±0.3 | 56.6±0.6 |

[a]Apparent heat of unfolding (kcal/mol)
[b]Apparent heat capacity change of unfolding (kcal/mol/K)
[c]Transition midpoint temperature (C)
± 95% confidence interval

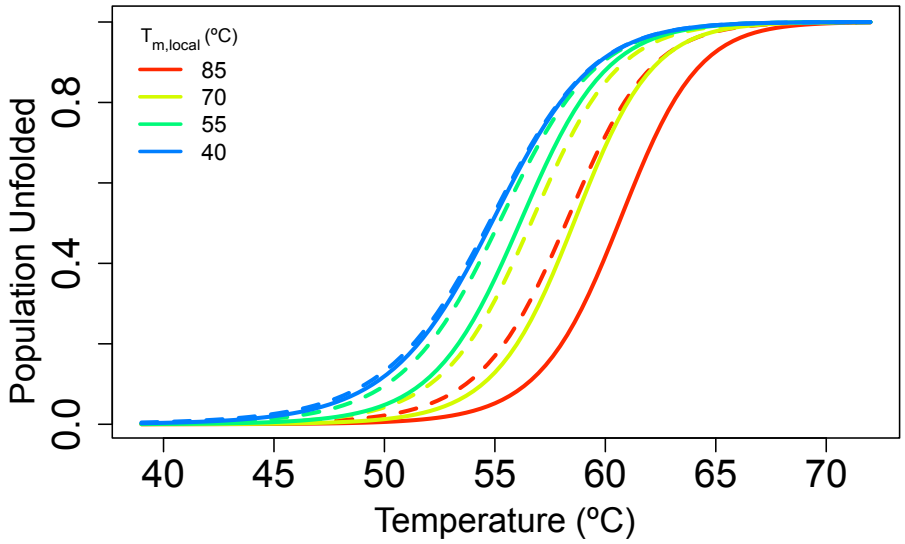



**Figure 2.2.3.2. Simulation of Destabilizing Loop Mutation Effects on Protein Global Stability.** Plot of the population unfolded, $P_{glob}$ , from Equations 2-4 through 2-8. *Solid Lines* represent the un-mutated protein. *Dashed lines* represent the protein with a destabilizing mutation applied to the locally unstable region (i.e. changing $K_{loc}$).

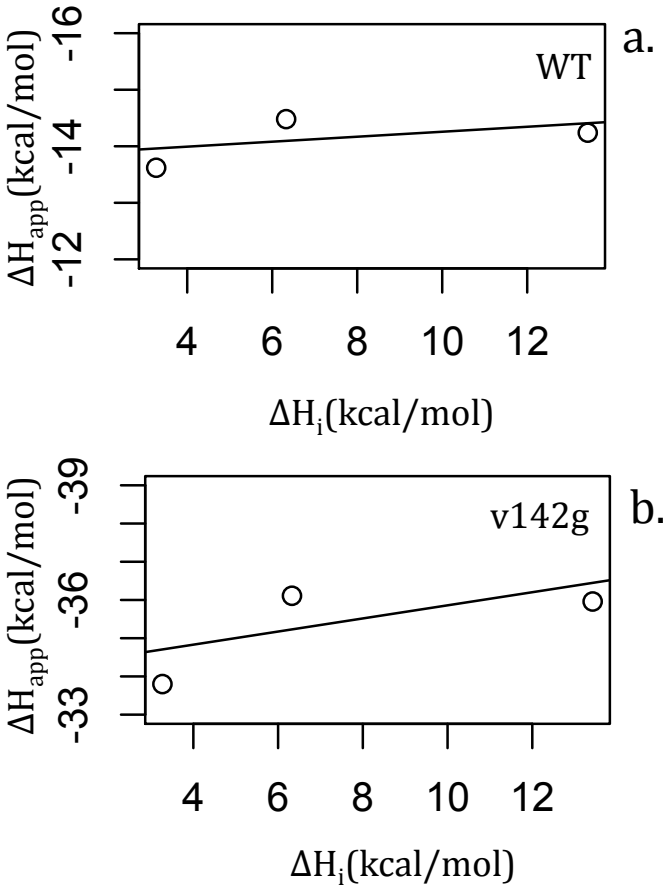



**Figure 2.2.4.1.1 Heat of Binding as a Function of Buffer Heat of Ionization.** Measured Heat of binding of WT and v142g AK in several buffers are plotted with the heat of ionization ($\Delta H_i$) of the corresponding buffer. Linear fits to the data by the least squares method were performed according to Equation 2-10.

**Table 2.2.5.1: Changes in Thermodynamics of Binding with Mutation and Increasing Sarcosine Concentration.**

| | $\Delta G_{app}$ (cal/mol) | $\Delta H_{app}$ (cal/mol) | $-T\Delta S$ (cal/mol) |
|---|---|---|---|
| WT(0M) | -9,950±25 | -8,770±18.0 | -11,80±31 |
| WT(2.5M) | -9,760±24 | -9,295±20.0 | -465±31 |
| *difference* | 190±34 | -530±27 | 720±43 |
| v142g(0M) | -9,770±31 | -19,350±51 | 9,580±59 |
| v142g(2.5M) | -9,630±15 | -14,440±23 | 4,812±28 |
| *difference* | 140.0±34 | 4,910±55 | -4770±65 |

Experiments performed in Tris pH 7.5, 30°C, in the presence of 0 or 2.5 M sarcosine
Differences are 2.5M -0M.

# CHAPTER 3

# TEMPERATURE DEPENDENCE AND STRUCTURAL MAPPING OF LOCAL UNFOLDING IN AK FROM E. COLI

## 3.1 INTRODUCTION

In Chapter 2 data from multiple techniques were presented which qualitatively suggest that the mutations v142g, v148g, and v135g promote a population change in a conformational process within the native state of AK. Furthermore, this conformational reaction appears to be thermodynamically linked to the Ap5A inhibitor binding reaction. Because of the availability of high quality calorimetric data, it seemed reasonable to attempt to quantitatively treat this data in order to better characterize the state itself and attempt to construct a model that will capture the relationship of this state to the binding of Ap5A.

## 3.2 MATERIALS AND METHODS

### 3.2.1 Non-linear Least Squares Analysis of ITC Data

The ITC results suggest that the binding reaction is coupled to a conformational equilibrium that is apparently affected by the mutations. The simplest conformational model that can explain these observations is one in which a binding incompetent (BI) state with no affinity for the ligand is in equilibrium with a

binding competent (BC) state, and that the addition of ligand promotes therefore the population of the BC state through mass action. We approach the application of this model to our data by fitting the temperature dependence of the observed enthalpy of binding. The necessary theory is developed below.

According to this model, the apparent binding constant will be

$$K_{app}(T) = \frac{[BCX]}{[BC + BI][X]}. \qquad \text{Equation 3-1}$$

Since $K_{conf}$ (T)= [BI]/[BC], Equation 3.1 reduces to

$$K_{app}(T) = \frac{K_o(T)}{(1 + K_{conf}(T))} \qquad \text{Equation 3-2}$$

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

$$K_o(T) = \frac{[BCX]}{[BC][X]}. \qquad \text{Equation 3-3}$$

The corresponding free energy of binding will thus consist of two terms:

$$\Delta G_{app}(T) = -RT \ln K_o(T) + RT \ln(1 + K_{conf}(T)) \qquad \text{Equation 3-4a}$$

$$\Delta G_{app}(T) = \Delta G_0(T) - \Delta G_{conf,app}(T), \qquad \text{Equation 3-4b}$$

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 difference between the BI and BC states. We note that in expression 4a

$$K_{conf}(T) = \exp\left( {-\Delta G_{conf}}/{RT} \right) \qquad \text{Equation 3-5}$$

where

$$\Delta G_{conf}(T) = \Delta H_{conf}\left(\frac{T_{m,conf} - T}{T_{m,conf}}\right) - (T_{m,conf} - T)\Delta C_{p,conf} + T\Delta C_{p,conf} \ln\left(\frac{T_{m,conf}}{T}\right) \qquad \text{Eq. 3-6}$$

and $\Delta H_{conf}$, $T_{m,conf}$, and $\Delta Cp_{conf}$ are the enthalpy, transition temperature, and heat capacity difference between the BI and BC states.

The apparent enthalpy for the binding process can be obtained through a temperature derivative of Equation 3.2 with respect to 1/T,

$$\Delta H_{app}(T) = -R \cdot \frac{d \ln K_{app}(T)}{d\left(1/T\right)}, \qquad \text{Equation 3-7}$$

to yield

$$\Delta H_{app}(T) = \Delta H_0(T) + \frac{K_{conf}}{\left(1 + K_{conf}\right)} \Delta H_{conf}(T) \qquad \text{Equation 3-8}$$

where, once again, 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:

$$\Delta H_{app}(T) = \Delta H_0(T) + P_{BI} \cdot \Delta H_{conf}(T) \qquad \text{Equation 3-9a}$$

$$\Delta H_{app}(T) = \Delta H_0(T) + \Delta H_{conf,app}(T) \qquad \text{Equation 3-9b}$$

Removal of $\Delta H_o$ (T) can be facilitated by subtracting out the baseline enthalpy change for each mutant (i.e. the slope and intercept of Figure 2.2.2.3-b below 20°C) resulting in an expression for the apparent conformational enthalpy effect from the BC to the BI transition:

$$\Delta H_{conf,app}(T) = \left( \frac{e^{\frac{\Delta G_{conf}(T)}{-RT}}}{1 + e^{\frac{\Delta G_{conf}(T)}{-RT}}} \right) \bullet \left( \Delta H_{m,conf} - \Delta C_{p,conf}(T - T_{m,conf}) \right) \qquad \text{Equation 3-10}$$

The form of equation 3.10 is particularly useful for fitting because $\Delta H_{conf}$ (T) is contained both in the exponential and a pre-exponential term, thus determining both the magnitude and the rate of change of the apparent enthalpy, $\Delta H_{conf,app}$ (T), with temperature. Therefore, if identical values of ΔH can be used in fitting, this validates the classic test of two-state behavior, namely $\Delta H_{van't\ Hoff} = \Delta H_{calorimetric}$.

However, because the upper temperature range of the ITC data approaches that of the global unfolding temperature (see Figure 2.2.2.3) we must consider a slightly more complicated formulation of the observed enthalpy of binding that includes the contribution of global unfolding at high temperatures in order to be able to satisfactorily fit the experimental data. A three state model accounting for this effect will include a fully folded state with affinity for the binding competent ligand, and two binding incompetent states that are locally (BIL) and globally (BIU) unfolded. The partition function for such a system will be:

$$Q = 1 + K_{conf} + K_{conf} \cdot K_{unfold} \qquad \text{Equation 3-11}$$

where,

$$K_{conf} = e^{\frac{\Delta G_{conf}}{-RT}} = \frac{P_{BIL}}{P_{BC}};\ K_{unfold} = e^{\frac{\Delta G_{unfold}}{-RT}} = \frac{P_{BIU}}{P_{BIL}}. \qquad \text{Equation 3-12}$$

The probability of all states in the system are therefore,

$$P_{BC} = \frac{1}{Q};\ P_{BIL} = \frac{K_{conf}}{Q};\ P_{BIU} = \frac{K_{conf} \cdot K_{unfold}}{Q}. \qquad \text{Equation 3-13}$$

For this system, as above, the observed heat of binding in a calorimetric binding experiment will have contributions from all relevant linked equilibrium:

$$\Delta H_{app}(T) = \Delta H_0(T) - \mathrm{P}_{BIL} \cdot \Delta H_{conf}(T) - \mathrm{P}_{BIU} \cdot (\Delta H_{conf}(T) + \Delta H_{unfold}(T)) \qquad \text{Equation 3-14}$$

or equivalently,

$$\Delta H_{app}(T) = \Delta H_0(T) - \frac{K_{conf}}{\left(1 + K_{conf} + K_{conf} \cdot K_{unfold}\right)} \Delta H_{conf}(T) - \frac{K_{conf} \cdot K_{unfold}}{\left(1 + K_{conf} + K_{conf} \cdot K_{unfold}\right)} (\Delta H_{conf}(T) + \Delta H_{unfold}(T))$$ Eq. 3-15

By grouping terms,

$$\Delta H_{app}(T) = \Delta H_0(T) - \frac{K_{conf}(1 + K_{unfold})}{\left(1 + K_{conf} + K_{conf} \cdot K_{unfold}\right)} \Delta H_{conf}(T) - \frac{K_{conf} \cdot K_{unfold}}{\left(1 + K_{conf} + K_{conf} \cdot K_{unfold}\right)} \Delta H_{unfold}(T)\ ,$$ Equation 3-16

where $\Delta H_{conf}$ represents the heat of the BIL/BC transition, and $\Delta H_{unfold}$ is the heat of the BIL/BIU transition. The three terms in Equation 3.16 represent the enthalpy contribution of the association reaction of the ligand with the BC state, the BIL to BC transition, and the BIU to BIL transition, respectively. Therefore to investigate the thermodynamic properties of the local unfolding within the native state, it is necessary to isolate an experimental measure of the second term of Equation 3.16. The methods used to estimate and subtract terms one and three are described below.

The first term of Equation 3.16 ($\Delta H_0$) is simply linear where the temperature of the corresponding $\Delta Cp_0$ is constant. Therefore to subtract this contribution from our data, we use two data points at low temperatures (7 and 11°C) to define a line. This baseline heat is extrapolated to higher temperatures and subtracted from each data point.

The third term of Equation 3.16 represents the probability of the fully unfolded state multiplied by the enthalpy difference between the partially and globally unfolded states. This value can be estimated for AK by utilizing the

presented CD thermal unfolding data (Table 2.2.3.1). From the raw CD data plotted it is evident that the BC to BIL transition has little observable change in CD signal. Specifically, in the temperature range 30-40°C all mutants have high probabilities of the BIL state, and the CD signal is very similar to the WT AK. Therefore, the unfolding transitions seen by CD principally monitor the unfolding of the remainder of the protein, i.e. the BIU state. Although, the presented fits of the CD data use the incorrect model (two-state), they do provide suitable estimates of the probability of the BIU state for each protein as a function of temperature. However, the heat estimated by these fits is quite different for each protein, although we expect the heat of the true BIL to BIU transition to be highly similar for all proteins. The reason for this is clear. When fitting a two-state model to a three-state system, the probability of the intermediate state (even if not observable) influences the probability of globally unfolded state, as can be seen in Equations 3.11-13.

However, if the BC/BIL and BIL/BIU transitions are sufficiently separated, a two state fit will yield thermodynamic parameters that are a reasonable approximation of the parameters describing $K_{unfold}$ (T) in the three-state model. It is clear from the raw data that v142g should have the most separated BC/BIL and BIL/BIU transitions. Therefore, the $\Delta H$ and $\Delta C_p$ obtained for this fitting was used to approximate the BIL/BIU heat for all proteins. Therefore, the correction subtracted from the raw enthalpy data representing the third term of Equation 3.16, is the probability of the unfolded state (as estimated by two state fits of each protein) multiplied by the enthalpy of unfolding estimated for v142g. This correction was

applied to data from temperatures above 25°C, and was very small (< 1 kcal/mol) for all data below 42°C.

After these corrections are applied the remaining enthalpy has the form

$$\Delta H_{conf,app}(T) = \frac{K_{conf}(1+K_{unfold})}{\left(1+K_{conf}+K_{conf}\cdot K_{unfold}\right)}\Delta H_{conf}(T). \qquad \text{Equation 3-17}$$

This equation is readily reduced to Eq. 3.10a. For our high temperature data (45-49°C) where global unfolding becomes most significant, $K_{conf}$ is large (> 7/3) and $K_{unfold}$ is small (< 1/9). Accordingly, the maximal expected disagreement between the two values computed from the two state model (Equation 3.10) and three-state model (Equation 3.17) is less than ~5%, even for these most affected data points (T > 45°C), the simplified two-state model is therefore acceptable. The presented fits are fits of the Equation 3.10 to the enthalpy data corrected as described in this section.

### 3.2.2 Simulating the Three State Model of Apo-Adenylate Kinase Unfolding

The simulation of the three-state model discussed below was calculated with the identical equations and assumptions as discussed in the ΔH fitting section above. To review, $K_{conf}$ was estimated using the thermodynamic parameters defining this equilibrium. $K_{unfold}$ was estimated using the thermodynamic parameter estimates from the CD unfolding data of v142g, for the above stated reasons.

### 3.2.3 Estimating the Surrogate Adaptive Temperatures of AK Mutants

The simulation of $K_{app}(T)$ as calculated according to the equations developed for the fitting and ΔΔG analysis presented below for the WT and v142g pair. Where,

$$K_{app} = e^{\frac{\Delta G}{-RT}}. \qquad \text{Equation 3-18}$$

$K_{app}$ functions were also defined for v135g and v148g by adjusting the $T_{m,\ conf}$ to the values estimated for each mutant. Therefore, these functions model the effect of the local unfolding event and exclude any changes to the intrinsic binding reaction. These functions were then set equal to the $K_{app}$ for WT AK at 37°C and solved numerically for temperature. This yields the temperature at which these proteins' $K_{app}$ is equal to that of the WT.

Using $K_{app}$ of Ap5A as an adaptive endpoint is clearly artificial, but nevertheless instructive. In fact, the energetic contribution of a linked equilibrium to an apparent binding affinity is independent of the free energy of the intrinsic association reaction. Therefore, the adaptive potential of the mechanism and mutations investigated depends only on the balance between the magnitude of the mutations effect on $K_{conf}$ and the temperature dependence of the intrinsic binding reaction. Although this illustration is constructed using a non-biological example, it clearly demonstrates that the mechanism in question can be a very efficient adaptation strategy, if the thermodynamic properties of the system are appropriate.

**3.2.4 NMR Methods** $^{15}N$ labeled AK was prepared using Cambridge Isotopes Bio-Express media, and purifying as above. Experimental conditions were identical to ITC conditions, with 2 mM EDTA and ~1.0 mM AK concentration. All experiments

were performed on a Varian 800 MHz spectrometer. Partial assignments were made from previously published values(28). All data were processed in NMR pipe and visualized with nmrView.

## 3.3 RESULTS AND DISCUSSION

### 3.3.1 Fitting of ITC Data

Fitting the equation for $\Delta H_{conf,\ obs}$ to the v142g data yields good estimates of $T_{m,\ conf}$, $\Delta C_{p,\ conf}$, and $\Delta H_{m,conf}$ (Figure and Table 3.3.1.1). Two noteworthy aspects of the calorimetric data emerge from the fitted values. First, the enthalpy difference between the BI and BC states is large (~33kcal/mol) amounting to what would be expected for the enthalpy of unfolding of the entire LID domain (as described below). Second, the data obtained for all of the proteins can be fit with a common $\Delta Cp_{,conf}$ and $\Delta H_{conf}$, each protein differing only in the fitted $T_{m,conf}$, the midpoint temperature for the BC to BI state transition. Furthermore, the effect of the mutation (to Gly) on $T_{m,conf}$ is quite large, shifting the local unfolding transition midpoint from ~53°C (WT) to as low as ~35°C (v142g).

In terms of the presented model and analysis, the significance of these calorimetric results is three-fold. First, the fact that only the $T_{m,conf}$ differs between each protein indicates that the Gly mutations have indeed selectively increased the entropy of the BI state, consistent with the assertion that the mutational effects are primarily manifested as increases in the conformational degeneracy of a locally unfolded state. Second, by strong inference the results suggest that the

thermodynamic and conformational character of the BI state is common to the wild type and all three Gly mutants. Third, the data indicate that the transition between the BI and the BC state is a highly cooperative "two-state" process. This point is supported by the demonstration that the van't Hoff and calorimetric enthalpy changes for the local unfolding transitions are equal, as demonstrated by the high quality fit of Equation 3-10, see methods section.

To underscore the excellent agreement of the thermodynamic properties observed for all three Val to Gly mutations, it is instructive to examine the gross disagreement of these results with those from s129a, a mutation designed to destroy a key hydrogen bond in the LID domain, see Figure 3.3.1.2. In this case, the cooperative unit of the local unfolding is grossly different. And accordingly, all attempts to fit this data with using the shared enthalpy and heat capacity of the glycine mutants failed. Interestingly, the excess enthalpy and change in apparent binding constant can, in this case be accounted for by a three state unfolding model, that includes a bound and partially unfolded state (analysis not presented). Furthermore, the s129a mutant appears to change the magnitude of the native state CD signal at 228 nm to a greater degree than the Val to Gly mutants, suggesting that the surface exposed glycine strategy is superior in terms of gross structural preservation. This result is also an important demonstration of the basic result that the WT LID is inherently unstable with respect to local unfolding, and any type destabilizing mutation can result in populating the locally unfolded state. In other words, the Val to Gly mutations under discussion are not unique in their ability to destabilize the LID, but do appear to be unique in their ability to enhance local

unfolding while preserving the cooperative substructure of the WT native state ensemble.

### 3.3.2 Agreement of External Experimental Observables Validates Fitted Model

#### 3.3.2.1 *Changes in the Apparent Free Energy of Binding*

As described above (Equations 3.4a-b), the observed free energy of the inhibitor binding reaction will depend on both the intrinsic association of the protein with ligand and the position of the relevant conformational equilibrium, $K_{conf}(T)$. If a mutation affects only the conformational equilibrium and leaves the intrinsic binding event unchanged, the difference in the apparent binding affinity between the two proteins will be

$$\Delta\Delta G_{app,wt-mut}(T) = RT\ln\left(\frac{1+K_{conf,wt}(T)}{1+K_{conf,mut}(T)}\right) \qquad \text{Equation 3-19}$$

where the two identical terms representing $\Delta G_0(T)$ cancel. The raw binding data shows that at low temperature, where the $P_{BI}$ is small, the difference in $\Delta G_{obs}$ approaches zero for the v142g mutant/wild-type pair. Accordingly, the approximation can be made that $\Delta G_0$ is not changed for this mutation. Therefore, we can use the above estimated thermodynamic parameters describing the BI state for the wild-type and v142g proteins together with Equation 3.18, to directly predict the difference in the observed binding affinity of these two proteins. This calculation was performed by first fitting the Equation 3.4a to the v142g binding affinity data, where;

$$K_0 = e^{\frac{\Delta G_0}{-RT}} = \frac{P_{BCX}}{P_{BC}}$$ Equation 3-20

The temperature dependence of $\Delta G_0$ was defined with the modified Gibbs-Helmholtz Equation (Equation 3.6). The $\Delta C_p$ used for the intrinsic reaction was defined by the slope of the ΔH measured for temperature points 7 and 11°C, as above. The fit was performed with two free parameters, the enthalpy and reference entropy of the intrinsic binding reaction. This was necessary as the experimental parameters measured at low temperatures failed to sufficiently model the ΔG data. This is probably due to linkage of the intrinsic reaction to protonation events, and relative high error of determining $\Delta G_{app}$ at these temperatures. We note, that because affinity data were very poor, titration experiments with ΔH of < 3.5 kcal/mol (low temperatures) were not included in the analysis. The discussed fit is shown in Figure 3.3.2.1.1 (*black line*). ΔΔG was then calculated as a distance from this fitted curve using Equation 3.18 and the fitted thermodynamic parameters that define $K_{conf}$ as a function of temperature (Table 3.3.1.1). The upper and lower error bounds for the prediction were calculated using (two times) the average standard error of measurement of $\Delta G_{app}$, calculated from all repeated binding experiments. 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. It is reasonable to note that the assumption of low to no binding affinity for the BI state is strongly validated by this simulation. Also, although this prediction can not be quantitatively applied to v148g and v135g because of some small perturbation to the intrinsic free energy of binding (Figure 2.2.2.3, low

temperature), the fact that each of the Val to Gly mutations shows similar temperature dependence of the binding affinity (see Figure 3.3.2.1.1-a, *inset*) further supports the validity of the model for all mutants investigated.

#### 3.3.2.2 *Changes in the Apparent Global Protein Stability*

We can also predict the population of the globally unfolded state as a function of temperature for the WT from the above-described three state model. Such a prediction relies on the identical assumptions necessary to allow the fitting performed above, namely that the change in CD measured for AK upon unfolding monitors only the non-LID residues, and also that the unfolding of v142g is an approximately two-state reaction transition between the BIL and BIU states, as defined above. This prediction according to Equation 3.13 is plotted in Figure 3.3.2.1.1-b. Interestingly, this simulation nicely reproduces the measured temperature dependence of the globally unfolded state of the wild type; thus validating the above assumptions, the model used and the parameters estimated.

Additionally, the two state estimate of $\Delta H_m$ and $\Delta C_p$ of global unfolding from fitting CD unfolding data were lower and higher, respectively, than the experimental sums calculated as estimates of unfolding all residues of AK (see following sections). This indicates that the unfolding curve of WT AK is wider and more asymmetric than should be expected for a true two-state unfolding reaction of all residues. This simulation demonstrates that the increased population of the BI state before the global unfolding transition determines the decrease in apparent $\Delta H$ of unfolding for wild-type AK, as well as the further decreased apparent enthalpy of unfolding for the mutant proteins.

### 3.3.3 Structural Distribution of Local Unfolding

To obtain a site-resolved view of the mutational effects, $^{15}N$-$^{1}H$ HSQC NMR spectra were gathered at 33°C for the wild type and v142g proteins (Figure 3.3.3.1). As is evident, a large number of resonances clearly seen in the wild-type spectra have no measurable intensity in the v142g spectra. No new peaks appear concomitant to this loss, indicating that the resonances for residues affected by the BC to BI transition are severely broadened due to a chemical exchange process.

Of note, all but two of the broadened and assignable residues are included in a contiguous chain, from residue 109 to 165, consistent with the mutation at position 142 affecting a large region including, but not limited to, the LID domain. Qualitatively similar results are observed for mutations at positions 135 and 148 as well. We also note that not all residues experiencing chemical exchange need be directly involved in the BI to BC transition. Clearly, conformational processes can affect neighboring amides. In summary, the NMR data demonstrate that the mutations to the LID domain appear to bring about a cooperative conformational process that is consistent with the large enthalpy ($\Delta H_{conf,app}$) observed by ITC.

### 3.3.4 Similarity of the BI State to Local Unfolding

As noted, the mutation strategy implemented here is designed to increase the probability of excited states, while leaving the structures of the ground states unaffected. The results and analysis thus far indicate that the mutational effects are manifested as expected. But the question remains, how similar is the BI state to a

locally unfolded state, wherein those residues that are affected by the mutation are unfolded, with the remaining part of the structure being folded?

To address this question, we utilized the COREX algorithm, which uses the high-resolution structure as a template, and a long-standing surface area based parameterization of unfolding energetics, to predict the $\Delta H$ and $\Delta C_p$ of unfolding of different regions of the protein structure. If the BI state were indeed a locally unfolded state, a reasonable model based on the NMR data would be to assume that residues 110-164 are unfolded in the BI state. Using COREX, the predicted thermodynamics of unfolding this region (PDB ID 4AKE) were calculated and compared to the experimentally determined values (Table 3.3.4.1). The agreement is excellent. Surprisingly, the thermodynamic parameters estimated for global unfolding of v142g also match the COREX predicted values for unfolding the remaining residues, 1-109 and 165-214. Apparently the transition midpoints ($T_m$) of the local and global unfolding event are sufficiently separated for the v142g mutant, such that the thermal unfolding data (v142g CD experiment) provide an unambiguous experimental measure of all residues not involved in the local unfolding process.

Two important results emerge from the agreement between experimental and predicted values. First, the $\Delta H$ and $\Delta C_p$ predicted by COREX are directly calculated from expected changes in solvent exposed surface area. Therefore, in terms of changes in solvent exposure, the BC/BI transition appears to be indistinguishable from a local unfolding of residues 110-164. Second, the fact that the sum of the heat of the BC/BI and the native/denatured state transitions equal

the total enthalpy expected for unfolding all residues, suggests that fluctuation to the BI state is the dominant local unfolding process in the native state.

Although, maybe less clear to interpretation, we can also compare the expected change in conformational entropy of unfolding for a surface exposed valine to glycine those manifest in the present ITC results. Figure 3.3.1.1 shows simulations of the data expected for average surface exposed valine to glycine mutations base the expected changes in the conformational entropy gain of unfolding. Interestingly, the mutations under study here are larger than both of these estimates. The origin of the larger than expected effect is uncertain. However, we can say that these data are consistent will the mutated residues experiencing the full range of conformational freedom available to residues in fully denatured states of proteins (42). Therefore in terms of surface exposure and conformational entropy, it seems that the conformational state in question is highly similar to an unfolded state.

### 3.3.5 Implications of Local Unfolding for the Adaptation of Binding Affinity

To investigate the possible adaptive significance of this mechanism we have compared the temperature dependence of the apparent association constant ($K_{app}$) of inhibitor binding for the wild-type and v142g proteins as a function of temperature. This data is plotted along with the populations of the locally unfolded state in Figure 3.3.5.1 to illustrate the origin of the observed effect. From this we can see that the $K_{app}$ of v142g becomes equal to that of the WT protein at physiological temperature (37°C) at approximately 27°C. Thus, a significant

adaptive change can be mediated by this mechanism via a single point mutation. We note that, this result and adaptive mechanism is generalizable to any protein-ligand binding reaction with similar thermodynamic properties. This adaptative mechanism is also consistent with the adaptive changes seen in nature. The change in binding occurs in the absence of structural changes in the active site or structural core of the protein, consistent with the high conservation of active site residues. By extension, our data suggest that the physical nature of the low affinity conformational states predicted by Somero, may be locally unfolded states in some proteins.

## 3.4 SUMMARY AND CONCLUSIONS

From these analyses we can see that the presented data and model yield an excellent quantitative explanation of the three experimental measurements provided, namely the apparent ΔH and ΔG of binding as well as the CD unfolding data. In addition, NMR experiments have allowed us to map the region involved in the cooperative local unfolding process. The quantitative agreement of the ITC derived thermodynamic parameters with those predicated from COREX, suggests that the state in question is indeed quite similar to an unfolded state. Moreover, since the COREX energy function is parameterized in terms of solvent exposed surface area, we can further infer that the change in solvent exposure for the folding of the LID domain, within the context of an otherwise folded protein, is identical to

what would be expected for folding the same structural unit (i.e. the LID) free in solution.

This is a clear demonstration that partial unfolding can be a member of the set of native state conformations available to proteins. However, it would be interesting to make such measurements on other systems where the structure of the locally unstable region is not domain like, as in LID of AK. The work of Manson et al. seems to indicate that conformational excursions of smaller regions (as compared to the LID) might be highly affected by contacts with stable structural elements. It is possible that in cases such as these, the conformational and thermodynamic quality of the high-energy state may be less similar to true unfolding than the locally unfolded state described here.

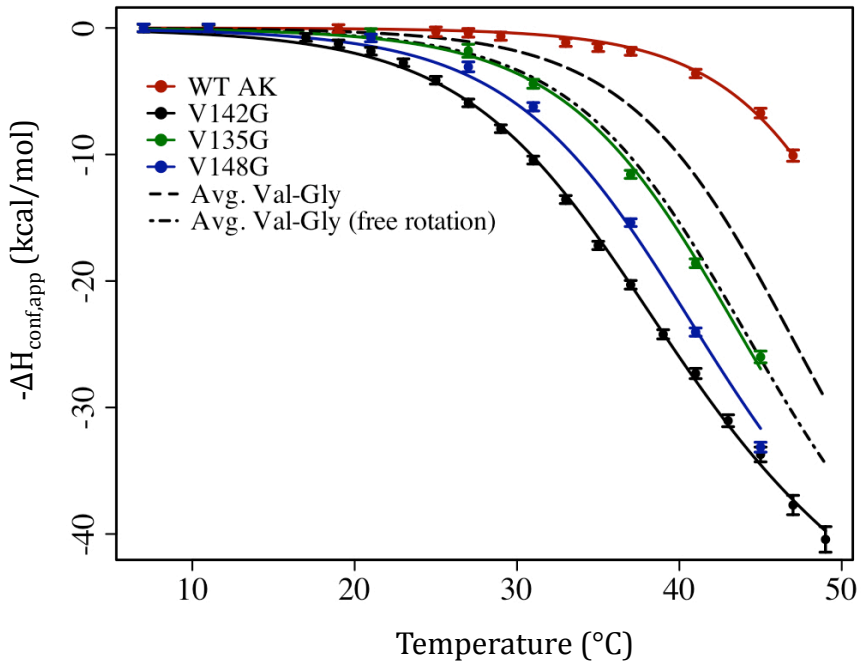



**Figure 3.3.1.1: Fitting of $\Delta H_{conf,app}$ for Comformational Thermodynamic Parameters.**

Corrected data and fitting functions representing $\Delta H_{conf,app}$ as described in materials and methods. 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 expected data for an average valine to glycine mutation based on the fitted conformational ΔH and ΔCp (40). *Dot-dashed line* represents is a simulation of the expected data for a Val to Gly mutation where folding places no conformational restriction on the side chain, i.e. only the backbone contribution is taken into account (40).

**Table 3.3.1.1: Thermodynamic Parameter Estimates Describing the Binding Associated Conformational Transition.**

| | $\Delta H_{m,\ conf}$ [a] | $\Delta C_{p,\ conf}$ [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 |

[a] Transition enthalpy at $T_{m,\ conf}$ of v142g (kcal/mole)
[b] Change in heat capacity (kcal/mole/kelvin)
[c] Transition midpoint temperature (degrees Celsius)
[±] 95% confidence interval (estimated from the fitted model, by profile likelihoods)
[nv] Value not varied in the fitting routine (v142g value applied)

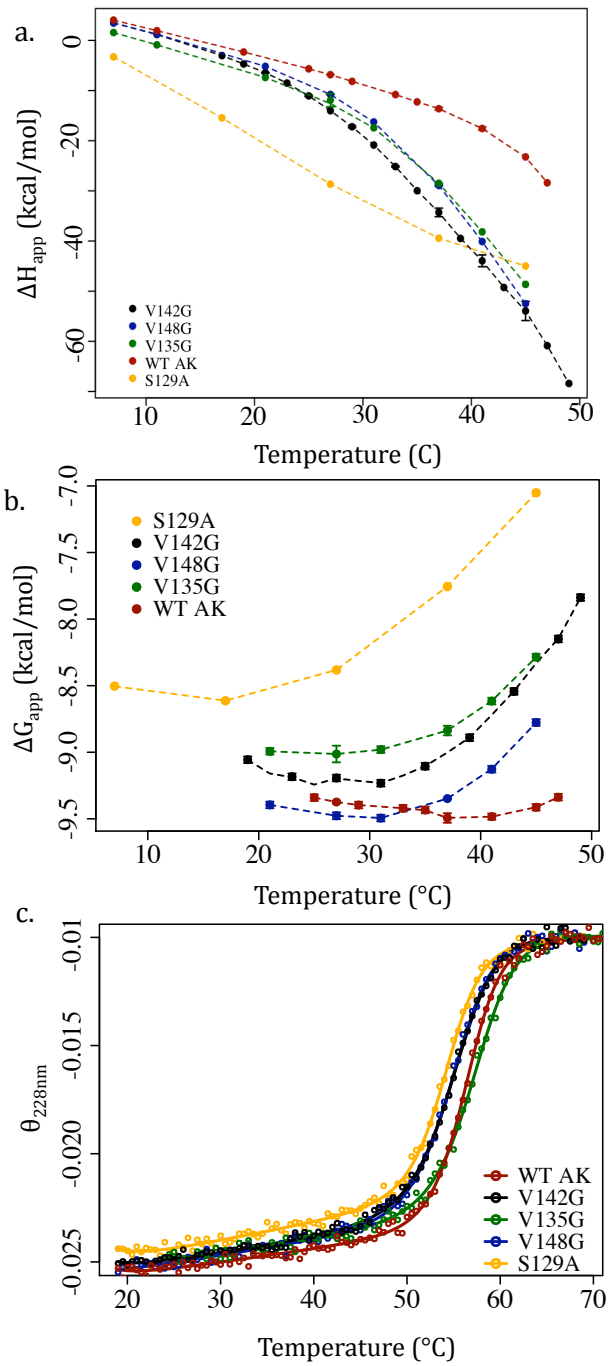



**Figure 3.3.1.2: Comparison to the Binding Reaction and Temperature Denaturation of a Structurally invasive mutant, s129a.**

WT (*red*), v142g (*black*), v148g (*blue*), v135g (*green*), and s129a (*orange*). **a.** The apparent enthalpy of binding to Ap5A ($\Delta H_{app}$). **b.** The apparent free energy of binding Ap5A($\Delta G_{app}$). **c.** Molar eliptisity at 288nm, $\theta_{288}$, (deg·cm$^2$/dmol·res), recorded by circular dichroism spectroscopy. Buffer conditions were as above.

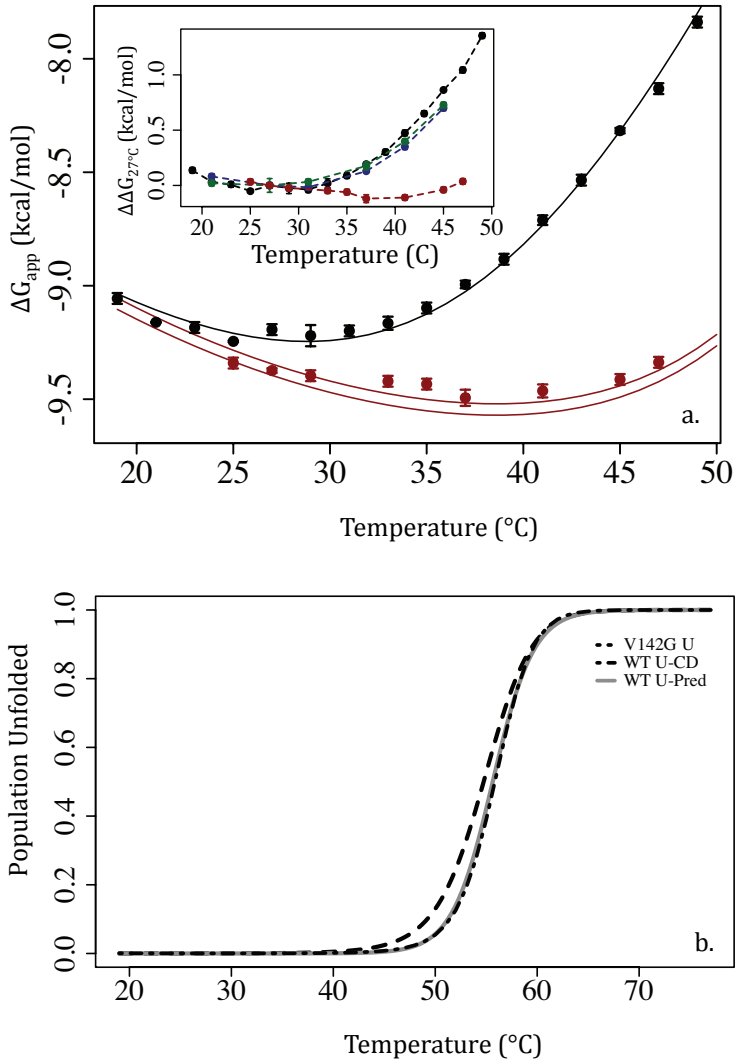



**Figure 3.3.2.1.1: Agreement of External Experimental Data with The Fitted Conformational Model and Parameters**.

**a.** Prediction of WT data based on the fitting of $\Delta H_{conf,app}$, see materials and methods. Lines represent the prediction, plus and minus the average standard error of determining $\Delta G_{app}$. *Inset* shows change in $\Delta G_{app}$ for each protein, with temperature, referenced from 27°C. $\Delta\Delta G_{27\cdot C}$(Temp) = ΔG(Temp) - ΔG(27°C). **b.** Prediction of population of the globally unfolded state for WT AK based on the fitted model, and v142g unfolding data. *Dashed line* shows measured (CD) probability of the globally unfolded (U) state for v142g. *Dot-dashed line* shows prediction of WT U state. *Grey line* shows measured probability of the globally unfolded state for WT AK.

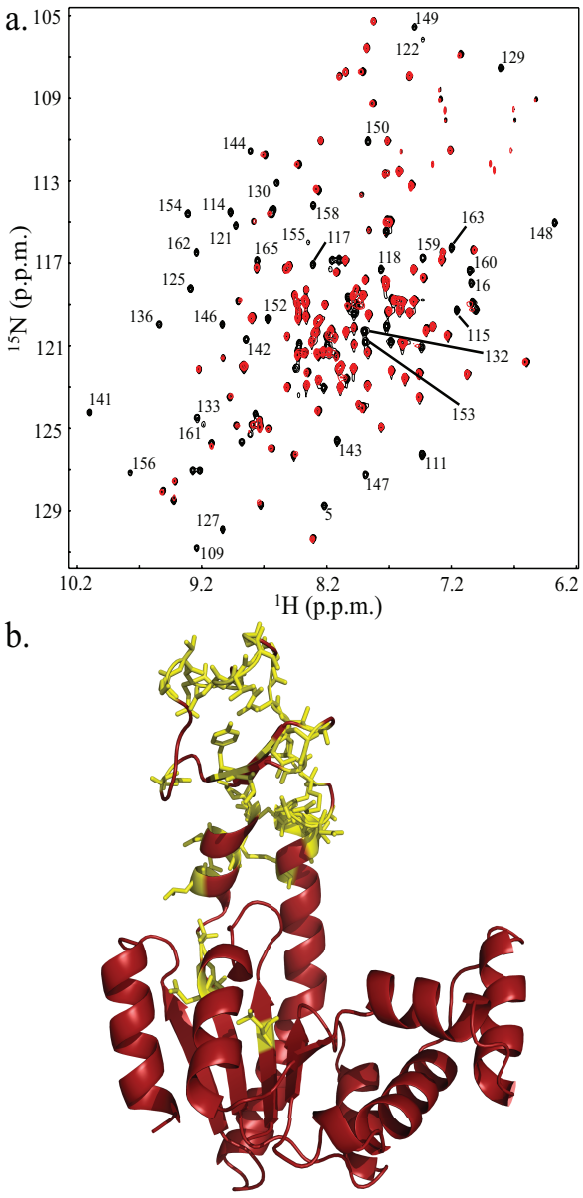


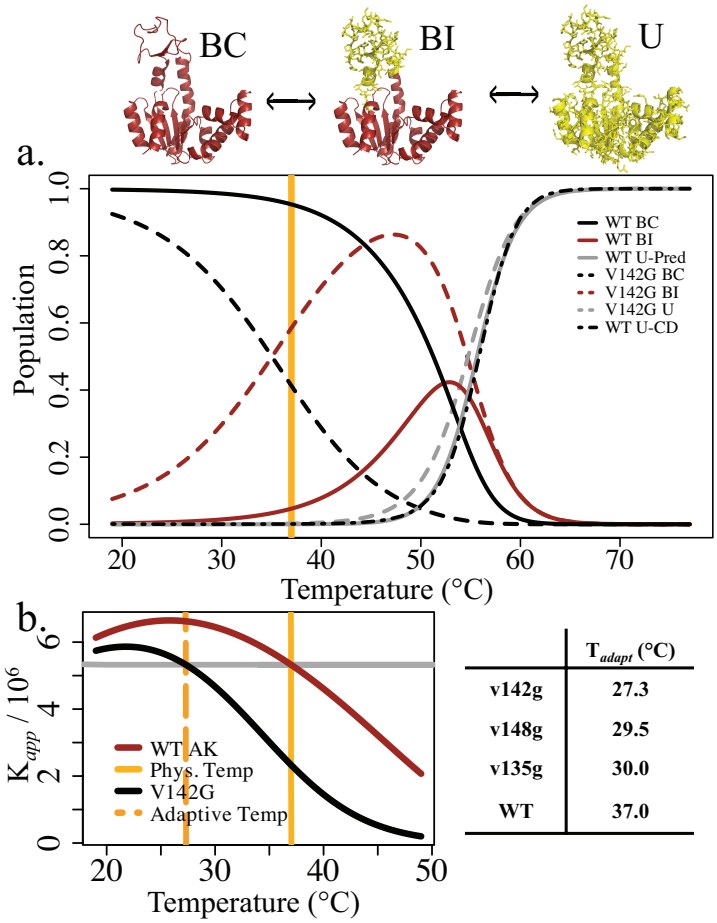

**Figure 3.3.3.1: $^{1}$H-$^{15}$N HSQC of WT and v142g AK (33°C).**

**a.** $^{1}$H-$^{15}$N HSQC spectra of wild type (black) and v142g (red) AK at 33°C. Labels are provided for all residues that were assignable from published values (26), and have measurable intensity in the wild-type spectrum, but no measurable intensity in v142g. **b.** The same residues (labeled in *a*) are shown in yellow with "sticks", projected on the open conformation of AK, PDB ID 4AKE.

**Table 3.3.4.1: Experimental and Predicted Thermodynamics of Local and Global Unfolding.**

| | | Local[a] | | Global[b] | | Sum[c] | |
|---|---|---|---|---|---|---|---|
| | T(°C) | ITC[d] | COREX[‡] | CD(v142g)[e] | COREX[‡] | Experimental[f] | COREX[‡] |
| **ΔH** | 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 |
| **$\Delta C_p$** | | 0.66±0.03 | 0.7 | 2.8±0.2 | 2.4 | 3.5±0.2[†] | 3.1 |

[a] Experimental and computational measures of a possible local unfolding reaction (residues 110-164).
[b] Experimental and computational measures of unfolding the remainder of the protein(residues 1-109, 165-214).
[c] Sum of local and global reactions. Represents unfolding all residues in the protein (residues 1-214).
[d] Parameters estimated from the ITC data fitting, or calculated from these estimates.
[e] Parameters estimated from the CD data fitting, or calculated from these estimates.
[f] Sum of heat or heat capacity estimated for the BI/BC transition (ITC) and the native/denatured transition (CD).
[‡] Values calculated from the expected change in exposed polar and apolar surface area upon unfolding, using the COREX energy function and PDB ID 4AKE, confidence is generally ± 10% for estimation of $\Delta C_p$ and ΔH for a differential scanning calorimetric unfolding experiment
[±] Experimental 95% confidence interval (CD), or Standard Error (ITC)
[†] Propagated Error




| | $T_{adapt}$ (°C) |
|---|---|
| v142g | 27.3 |
| v148g | 29.5 |
| v135g | 30.0 |
| WT | 37.0 |

**Figure 3.3.5.1: Adaptive Significance of Locally Unfolded States in Adenylate Kinase.**

**a.** Populations are calculated from the BI/BC ($K_{conf}$) equilibrium estimated from the ITC data, and the Unfolded/BI ($K_{unf}$) equilibrium estimated from the CD unfolding experiments on v142g (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). Therefore, the partition function for the system is: $Q = 1 + K_{conf} + K_{conf} \cdot K_{unf}$. The probability of BC, BI and Unfolded (U) states are calculated by: $1/Q$, $K_{conf}/Q$, and $K_{conf} \cdot K_{unf}/Q$, respectively. ***a- Dot-dashed line.*** Population of the U state for WT as calculated directly from the CD unfolding parameter estimates (Fig. 2c-Table). ***Orange line***. Marks optimal growing temperature for *E. Coli.* **b.** Plot of the apparent binding affinity ($K_{app}$) for Ap5A, calculated from the ITC data fitting (Fig. 2a-b). ***b-Dashed orange line***. Marks temperature where $K_{app}$ is equal for v142g and WT (37C), ie. the new surrogate homeostatic temperature of v142g. ***b-Table.*** Surrogate homeostatic temperatures, $T_{adaptive}$, calculated for all mutants. $K_{app,mut}(T_{adaptive}) = K_{app,wt}(37°C)$.

# CHAPTER 4

# STRUCTURAL INVESTIGATIONS OF AK MUTANTS

## 4.1 INTRODUCTION

In general, it is well known that the fold shared by homologous proteins is robust to sequence changes. For example, adenylate kinases from *Bacillus sp.* share the fold of *E. coli* AK. Furthermore, an interesting series of *Bacillus* AKs having approximately 70% sequence identity have been examined by crystallography (56). Structural alignment of the LID region (residues 128-159) of the molecules revealed carbon-alpha RMSD values of 0.43-0.46 angstroms (56). By extension, the structural effects of a single point mutation are expected to be subtle and not prorogate by largely perturbing the conformation of the protein backbone. However, local structural consequences of disrupted interactions are, of course expected.

One primary goal of our mutation strategy was to selectively perturb the properties of conformational excursions while leaving the structure unperturbed. For this reason, we have chosen highly surface exposed residues, involved in few molecular contacts (see Chapter 1). If our mutation design is successful at achieving the goal of structural preservation, then we can be much more certain that any functional effects observed are a consequence of modulating the ensemble of native

state fluctuations. To evaluate the success of the above described mutation strategy on preserving the structure of AK we have performed crystallography on three AK mutants (v142g, v148g, and s129a) and the wild-type molecule, as well as performing supporting studies using NMR.

## 4.2 MATERIALS AND METHODS

### 4.2.1 X-Ray Crystallography

Crystals were grown in by the sitting drop method, in conditions similar to those previously reported (26). Protein was concentrated to 20-45 mg/ml with ~2mM Ap5A in 50 mM MES, pH 6.7, 1 mM EDTA. Precipitant solutions prepared in the above buffer with 3% w/v PEG 2K with 1.8-2.3 M Ammonium Sulfate, and pH 7.0-7.3. In 50% precipitant solution, crystals grew in 24-48hrs, equilibrating against a 0.5 ml reservoir. Crystals were cryo-protected in Paratone, and frozen in place. Data was processed using HKL2000. Phases were determined by molecular replacement using PHASER, using PDB ID 1AKE as a probe. Initial refinement was performed with COOT and Refmac. Further refinement was performed using PHENIX. Optimal R-free was achieved by using TLS refinement with 10 TLS units per protein chain. v142g crystals were grown under similar conditions, however, PMB/CNS refinement software was utilized in this case. No TSL refinement was necessary in this case.

### 4.2.2 NMR Methods

$^{15}N$ labeled AK was prepared using Cambridge Isotopes Bio-Express media, and purifying as above. Experimental conditions were identical to ITC conditions, with 2 mM EDTA and ~1.0 mM AK concentration. All experiments were performed on a Varian 800 MHz spectrometer. Partial assignments were made from previously published values (28). All data were processed in NMR pipe and visualized with nmrView.

## 4.3 RESULTS AND DISCUSSION

### 4.3.1 Structural Effects of Studied Mutations and Crystal Packing

Table 4.3.1.1 and 4.3.1.2 summarize of crystallographic refinement statistics for the 4 AK molecules investigated. WT, v148g, and s129a proteins crystallize in the identical space group ($P2_12_12$) and asymmetric unit (ASU). In this case, the asymmetric unit happens to contain two distinct copies of AK, see Figure 4.3.1.1. The v142g mutation disrupts a key crystal contact in the WT lattice, not surprisingly this molecule crystallizes in a different (I222) space group and ASU, which contains only one copy of AK. We note that at least ridged body TLS refinement of molecules in the $P2_12_12$ lattice is critical, as the molecules appear be aligned somewhat heterogeneously along one primary axis of rotation. This observation probably accounts for the relatively poor refinement statistics reported in most crystallographic studies of AK in the literature. The novel crystal lattice of v142g apparently does not permit such rotational freedom as observed in the WT lattice.

Correspondingly, the resolution of the v142g was slightly improved and no TLS refinement was necessary.

Whenever comparing small structural changes is the goal, one must carefully consider the effects of crystal packing on the resultant structure. In the case of our series of AK mutants we have several opportunities to examine the magnitude and distribution of such crystal packing effects. We have observed the fold of AK in three distinct crystallographic environments. Interestingly, a distinct conformation of the outer region of the LID domain is associated with each crystallographic environment investigated, see Figure 4.3.1.2. This region (~residues 130-155) has been demonstrated to be the most dynamic region of the protein on the nano- to picosecond time scale (31). Therefore, structural plasticity observed in this region may be indicative of a relatively heterogeneous conformational manifold that can be dynamically explored in solution (31). It is important realize that the meaning of the changes between v142g and the WT, in terms of possible mutational effects, would be far less clear if only one copy of the WT were available.

As a counterpoint, we can examine the structural changes effected by our mutations where crystallographic influences should be minimal, by comparing WT, v148g and s129a. Aligning, one copy of within the asymmetric unit, it is clear that the influences of crystal packing are largely identical in these cases, see figure 4.3.1.1. Furthermore, it is clear that differences between chains are reproducible between datasets/mutants (fig. 4.3.1.4-a and Table 4.3.1.3). From the data it also clear that the mutations in question have very minimal structural effects. In all cases of the $P2_12_12$ space group (even in the case of s129a, where a buried hydrogen

bond is disrupted) the structural orientation of residues surrounding the mutation site seem to be quite preserved, see Figure 4.3.1.3. To more rigorously examine the data for correlated changes in structure we have identified all atoms that move greater than a prescribed threshold (>0.3 Å) in both copies of the ASU, comparing WT versus v148g structures. Such atoms are highlighted in Figure 4.3.1.4 (*grey spheres*). Notably, the region of mutation site itself does not seem to have any increase in the number or 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 by the experiment.

### 4.3.2 Structural Effects in the Unbound Conformation

Unfortunately crystallographic investigation of unbound AK is fraught with experimental difficulties (25). Crystals grown and used in other labs were unstable and no storage condition could be found (25). Furthermore, the apo-state crystallizes in a highly interlocking conformation, were the opened active site of AK is filled by the protein atoms from another copy of AK in the crystal lattice, thus calling into question the relevance of this conformation (25). Therefore, we have used NMR observed chemical shift values as a qualitative measure of structural conservation. Although chemical shifts alone cannot define the structure, they are in general highly sensitive to changes in molecular environment. Figure 4.3.2.1 shows the overlaid HSQC spectra of v142g and WT AK at low temperatures, which suppresses local unfolding. The agreement between the two spectra is excellent.

Furthermore, the small changes noticeable at this temperature seem to be confined to the immediate region of mutation site. Differences in chemical shift based on the sequence change alone are expected in the region +/- two residues of the mutation site (57, 58). Therefore, the observations are consistent with the hypothesis that the unbound conformation of the folded state is minimally perturbed by the v142g mutation. Qualitatively similar results were found for v148g and v135g (data not shown).

### 4.3.3 Propagation of Mutational Effects in the Absence of Structural Change

One very important result of the structural studies is that through local unfolding equilibria, allosteric information can be propagated in the absence of a structural pathway. The great structural conservation of our mutants has been validated for both the bound and unbound state of AK, by crystallography and NMR, respectively. In particular, these data reinforce the notion that the observed allosteric mutational effects cannot be reconciled in the context of the structural changes, or in terms of a pathway of structural distortion between the mutated site and the active site. This is demonstrated in Figure 4.3.1.4, where it is clear that no network of structural distortions (> 0.3 Å) can be found which link the mutated site and the binding cleft. Although mutations or binding to some allosteric proteins do manifest themselves as structural distortions, the results presented here argue directly against an exclusively mechanical interpretation of energy propagation.

## 4.4 SUMMARY AND CONCLUSIONS

The results presented in this chapter demonstrate that the outermost portion of the LID domain (residues 130-155) is structurally plastic. However, the effects of crystal packing are the key determinant of the conformational deviations observed. In comparison, the effects of the mutations themselves are quite minimal. Furthermore, at the resolution of the experiment no pathway of structural distortions is observed that could impact functionally critical residues in the binding pocket. Therefore, these data strongly support our hypothesis that the effects of the described mutations are interpreted as the result of changes in the probability, kinetics or chemical properties of conformational excursions that surround the fully folded/ordered ground state.

**Table 4.3.1.1: Data Collection and Refinement Statistics for WT and v148g Adenylate Kinase.**

| Crystallographic Data and Refinement: wild type and v148g | wild type | v148g |
|---|---|---|
| *Data collection* | | |
| Wavelength (Å) | 1.542 | 1.542 |
| Temperature (K) | 100 | 100 |
| Space group | $P2_12_12$ | $P2_12_12$ |
| Unit cell axes (Å) | 83.65/72.64/78.69 | 83.47/72.62/78.48 |
| Molecules/asymmetric unit | 2 | 2 |
| Resolution (Å) | 50.0-1.9 | 50.0-2.0 |
| No. of measured reflections | 180,220 | 120,739 |
| No. of unique reflections | 36,651 (3,048) | 32,187 (3,144) |
| Completeness (%) | 98.2 (82.8) | 97.2 (96.1) |
| Redundancy | 4.9 | 3.8 |
| $R_{merge}$ | 0.046 (0.159) | 0.052 (0.195) |
| $I/\sigma$ | 47.55 (9.06) | 41.59 (10.32) |
| *Refinement* | | |
| Resolution (Å) | 15.0-2.0 | 15.0-2.0 |
| No. of atoms | 4,098 | 3,829 |
| No. of water molecules | 661 | 381 |
| R-factor (%) | 19.9 | 20.6 |
| $R_{free}$ (10% data) (%) | 24.5 | 24.3 |
| RMSD bond lengths (Å) | 0.011 | 0.008 |
| RMSD bond angles (°) | 1.51 | 1.46 |
| Mean *B*-factor protein ($Å^2$) | 33.4 | 36.4 |
| Mean *B*-factor water ($Å^2$) | 39.2 | 39.2 |
| Ramachandran regions (%) | | |
| Preferred | 94.3 | 92.9 |
| Additionally allowed | 5.7 | 6.9 |
| Generously allowed | 0 | 0.3 |
| Disallowed | 0 | 0 |

Values for the highest-resolution shell are in parentheses. wild-type: PDB ID 3HPQ. v148g: PDB ID 3HPR.

**Table 4.3.1.2: Data Collection and Refinement Statistics for v142g and s129a Adenylate Kinase.**

| Crystallographic Data and Refinement: s129a and v142g | | |
|---|---|---|
| | s129a | v142g |
| *Data collection* | | |
| Wavelength (Å) | 1.542 | 1.542 |
| Temperature (K) | 100 | 100 |
| Space group | $P2_12_12$ | I222 |
| Unit cell axes (Å) | 83.05/72.6/78.48 | 73.32/77.92/83.71 |
| Molecules/asymmetric unit | 2 | 2 |
| Resolution (Å) | 50.0-2.2 | 50.0-1.95 |
| No. of measured reflections | 113,974 | 69,941 |
| No. of unique reflections | 24,484 | 17,284 (1096) |
| Completeness (%) | 99.5 (98.8) | 97.2(92.3) |
| Redundancy | 4.7 | 4.0 |
| $R_{merge}$ | 0.119 (0.404) | 0.048 (0.183) |
| $I/\sigma$ | 5.11 (1.52) | 30.44(5.91) |
| *Refinement* | | |
| Resolution (Å) | 15.0-2.35 | 15.0-1.95 |
| No. of atoms | 3,741 | 3,482 |
| No. of water molecules | 306 | 93 |
| R-factor (%) | 22.6 | 20.2 |
| $R_{free}$ (10% data) (%) | 27.13 | 24.2 |
| RMSD bond lengths (Å) | 0.003 | 0.022 |
| RMSD bond angles (°) | 0.731 | 2.18 |
| Mean *B*-factor protein ($Å^2$) | 31.81 | 20.97 |
| Mean *B*-factor water ($Å^2$) | 25.87 | 18.24 |
| Ramachandran regions (%) | | |
| Preferred | 94.5 | 92.3 |
| Additionally allowed | 5.5 | 7.7 |
| Generously allowed | 0 | 0 |
| Disallowed | 0 | 0 |

Values for the highest-resolution shell are in parentheses.

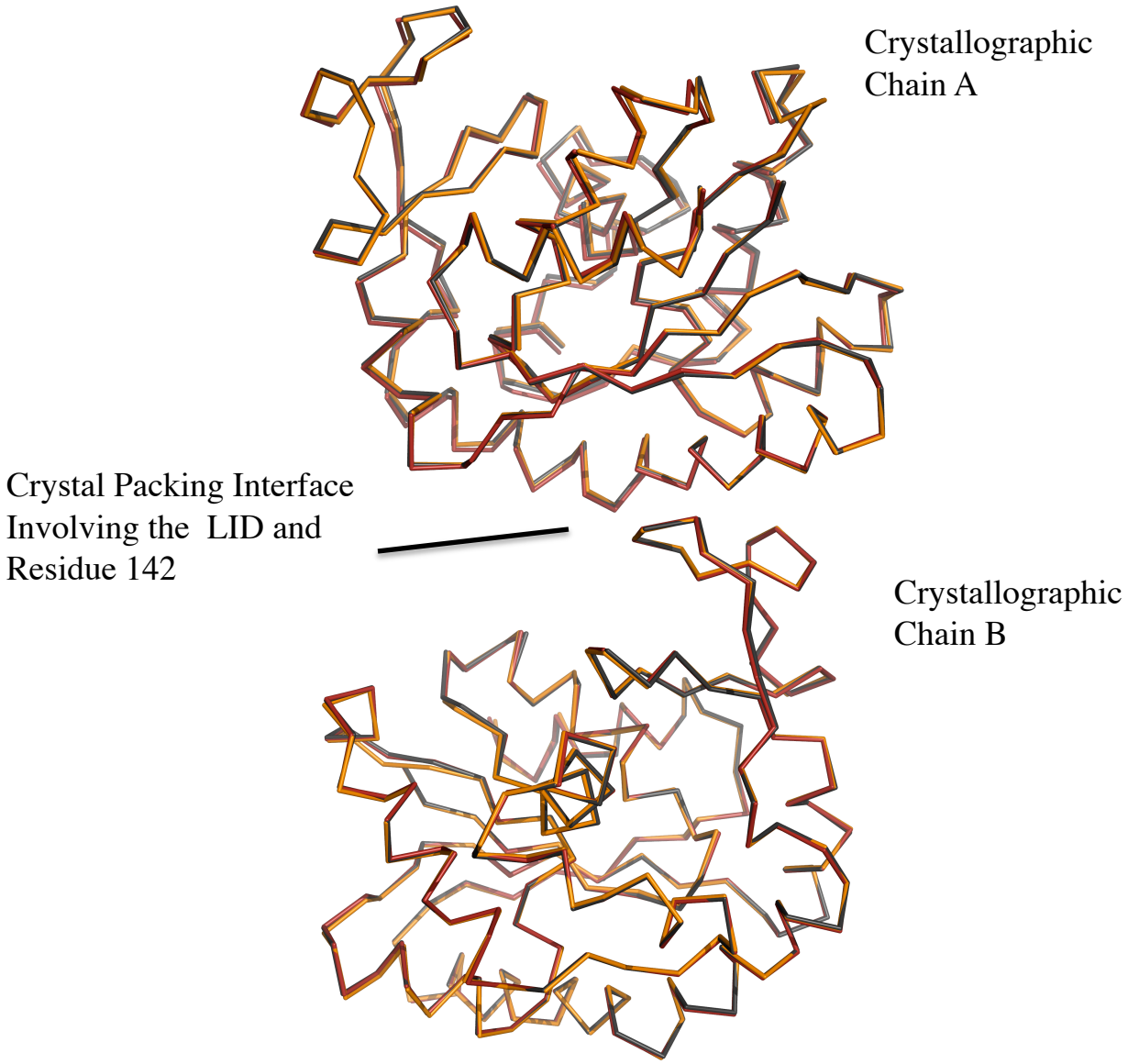



**Figure 4.3.1.1: Alignment of the Asymmetric Unit of WT, v148g and s129a AK.**

***grey***- the wild type AK. ***red***- v148g AK. ***orange***- s129a AK. Alignment is performed on chain B.

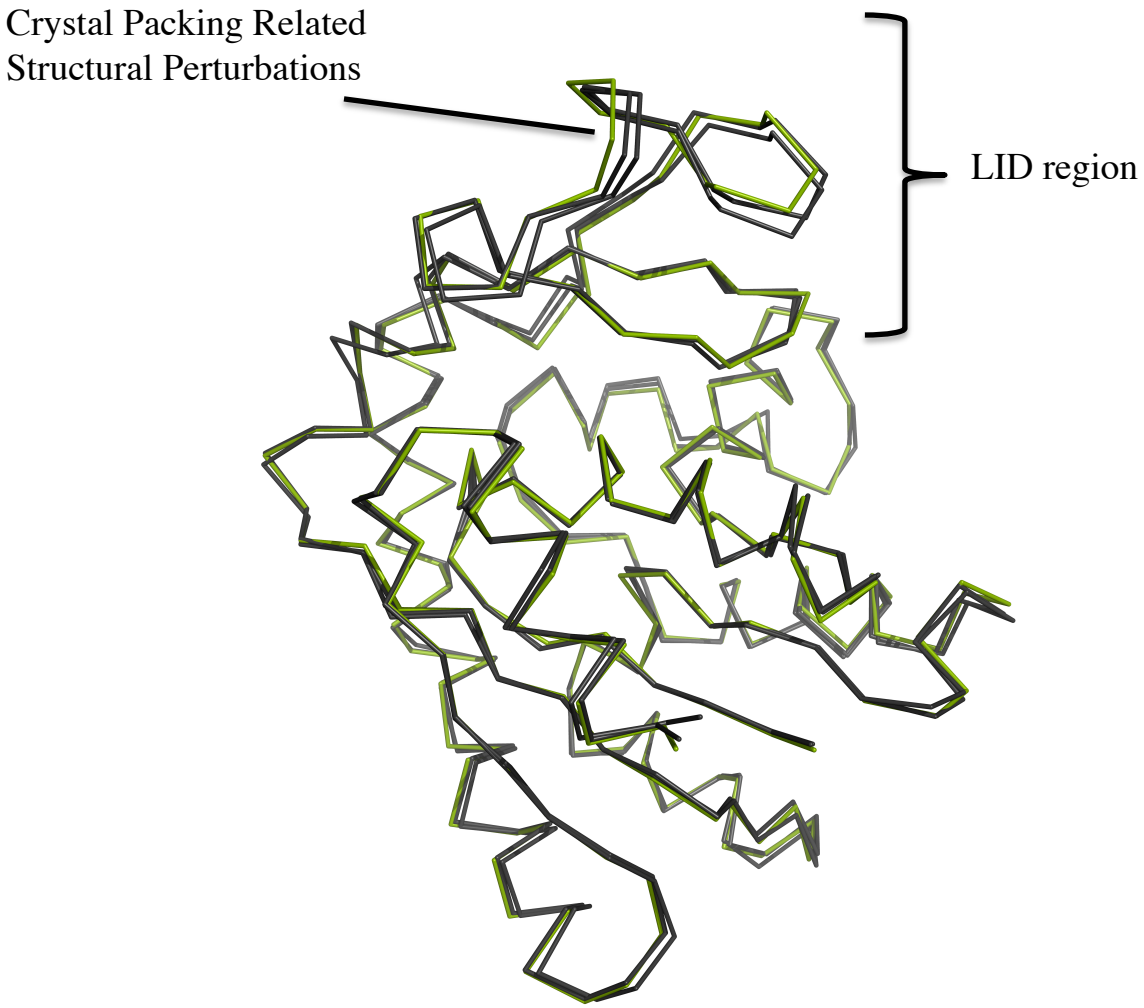



**Figure 4.3.1.2: Alignment Chains from Three Different Crystallographic Environments.**

***grey***- the wild type AK, both chains A and B. ***green***- v142g AK.

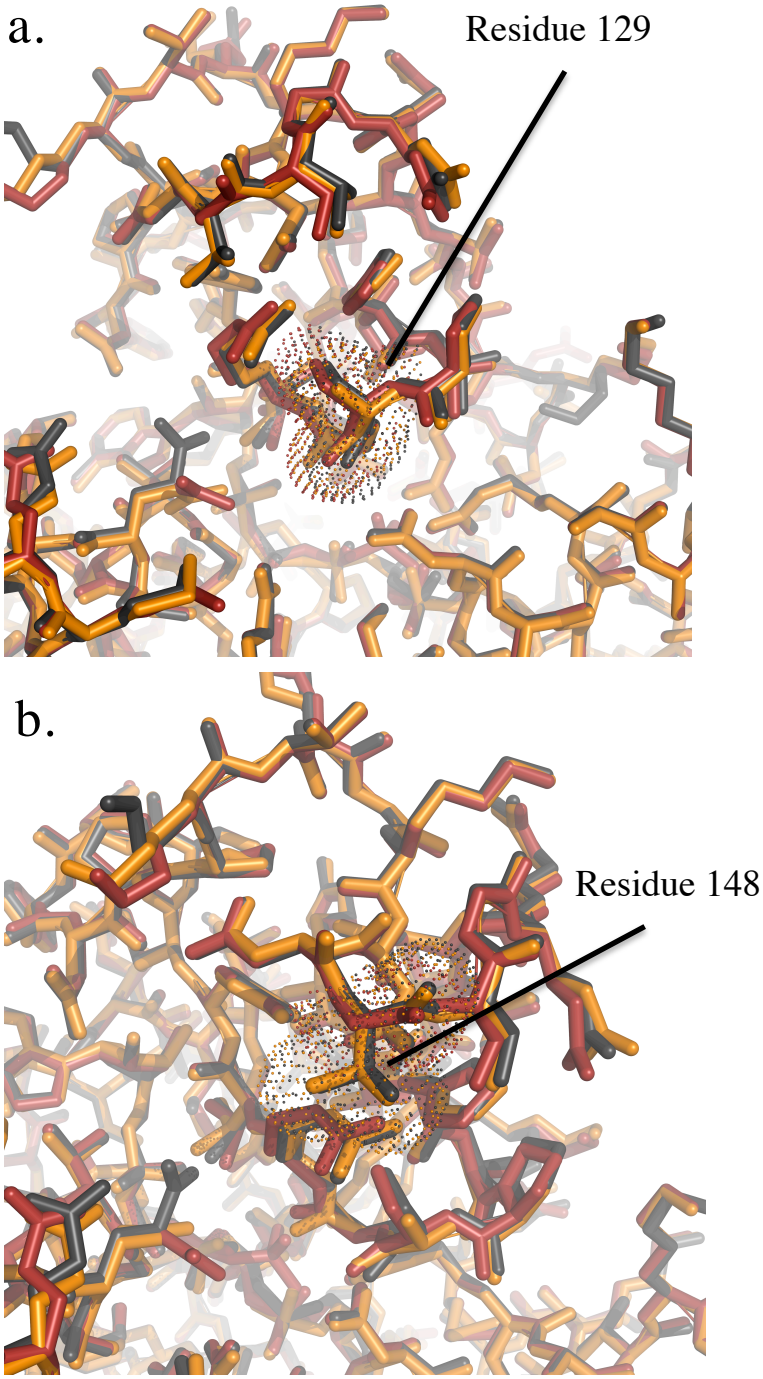



**Figure 4.3.1.3: Mutations Preserves Structure of Neighboring Residues.**

Structural alignments of the crystallographic B-chain was examined to assess the local structural impact of the mutations. B-chain was in general had more well-defined electron density in the LID region. ***grey***- the wild type AK. ***red***- v148g mutant. ***orange***- The s129a mutant AK. ***a-*** The s129a mutations site. ***b-*** The v148g mutations site.

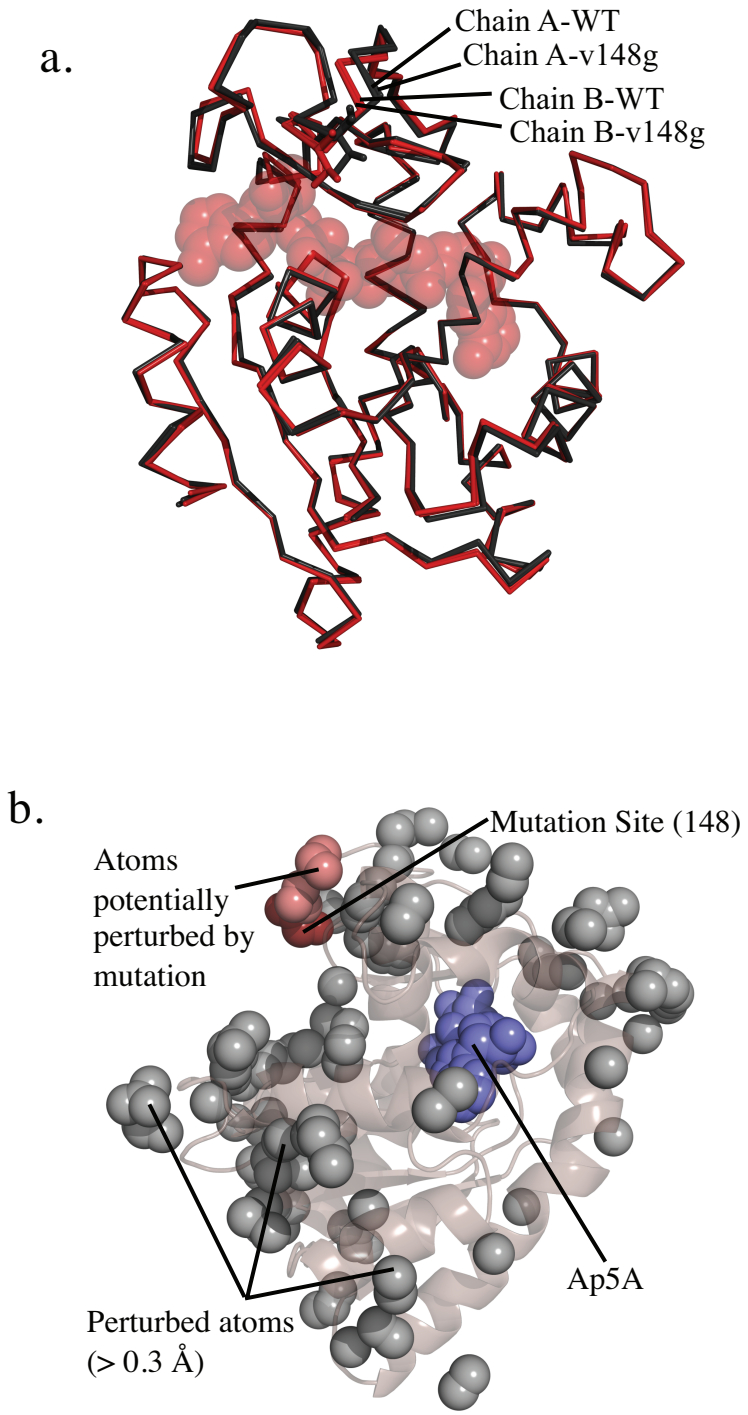



**Figure 4.3.1.4: Structural conservation in v148g AK.**

**a.** Alignment of the reported crystal structures of WT and v148g AK. ***A-red.*** Chains (WT and v148g) from position A within the asymmetric unit. ***a-black.*** Chains from position B within the asymmetric unit. **b.** Analysis of structural possible structural perturbations effected by mutation. ***b-Grey spheres.*** Represent all atoms that move > 0.3 Å from the WT to mutant structure in both copies within the ASU. ***b-Red spheres.*** The mutation site (148). ***b-Salmon spheres.*** All perturbed atoms (Grey) that can be connected to the mutation site by a continuous chain (< 6 Å per step) of other perturbed atoms. ***b-Blue spheres.*** Ap5A.

**Table 4.3.1.3: Summary of All Atom RMSD Comparing WT and v148g AK.**

**R.M.S.D. Comparison**

| | WT – Chain A | v148g – Chain B |
|---|---|---|
| WT – Chain B | 0.49 | 0.17 |
| v148g – Chain A | 0.20 | 0.42 |

All atom RMSD calculated using PyMol

# CHAPTER 5

## Direct Comparison the Kinetics of Local Unfolding of the AK LID and the Catalytic Rate.

### 5.1 INTRODUCTION:

To summarize, in the above-presented work, we have rigorously tested our hypothesis that changes in local unfolding alone, i.e. in the absence of changes to the fully folded ground state structure can mediate important functional changes in binding affinity. Furthermore, it is clear that local unfolding of the LID domain is an not just a result of mutation, but rather an important albeit relatively low probability state within the WT AK native ensemble.

Enzymes adapted to different temperature generally are found to have opposing changes in binding affinity and catalytic rate. Furthermore Somero and colleagues have discussed these changes as two sides of the same mechanistic process (3). Thus we interested to see if the expected changes in catalytic rate were present, as the AK glycine mutations quite nicely recapitulated the type of affinity changes seen in enzymatic adaptation. Therefore we planned experiments to investigate any perturbation to $k_{cat}$ induced by our mutants as well as attempt to related any such changes to the local unfolding of the LID region.

## 5.2 MATERIALS AND METHODS

### 5.2.1 CPMG Relaxation Dispersion NMR and Data Analysis

CPMG based $^{1}H$-$^{15}N$ relaxation dispersion experiments were performed (32). Conditions were 50mM MOPS buffer, pH 7.85, 50mM NaCl, 5% w/v glycerol, 10mM DTT, 10mM TCEP. Air in the sample tube was replaced with argon gas to further regard oxidation of the sample, as oxidation was found to be associated with gross aggregation of the sample. Under these conditions, 1.5-2 mM concentrations of AK were achievable.

$R_{2,eff}$ was estimated for each resonance by the difference in intensity (*I*) between a reference spectra taken in the absence of a CPMG element and another with including the pulse train at a particular frequency, according to the following equation.

$$R_{2,eff}^{\exp} = \frac{1}{T_{relax}} \ln\left(\frac{I_{cmpg}}{I_{ref}}\right);$$

Equation 5-1

Error (standard error) of the measured intensities was taken as being equal to the standard deviation of the intensity of the nmr spectra baseline. This value was estimated using the NMRView software package. Error of $R_2$ ($\sigma_{R2}^{exp}$) was therefore estimated by propagating the error of *I* thought the above equation. Data were analyzed by directly fitting a numerical solution of the Bloch-McConnel Equations for the dependence of transverse relaxation (R2) on the pulsing speed of the CPMG pulse train (*νCPMG*). Other authors have successfully used this approach (59).

MATLAB™ direct-search/genetic algorithm package was used numerical for minimization of the fitting-equation as standard derivative based methods of non-linear least squares require analytical expressions. The minimized equations are as follows:

$$x^2 = \sum_{\nu CPMG} [R_{2.eff}^{\exp}(\nu CPMG) - R_{2.eff}^{\exp}(\nu CPMG)]^2 / [\sigma_{R_{2,eff}}^{\exp}(\nu CPMG)]^2, \qquad \text{Equation 5-2}$$

for individual resonances (*res*), or

$$x^2 = \sum_{res.} \sum_{\nu CPMG} [R_{2.eff}^{\exp}(\nu CPMG) - R_{2.eff}^{\exp}(\nu CPMG)]^2 / [\sigma_{R_{2,eff}}^{\exp}(\nu CPMG)]^2, \qquad \text{Equation 5-3}$$

where multiple resonances were fit globally as a group. Errors of fitted parameters were taken as the standard error of direct experimental repeats where possible. Repeats were performed as separate, with respect to both experimental and data processing procedures. Fitting analysis and error propagation were performed by in house scripts written by T. Schrank, using MATLAB™ and the Python™ language Uncertainties Package.

### 5.2.4. Measurement of $k_{cat}$

Enzyme activity was measured in the direction of ATP production, utilizing a Hexokinase/Glucose-6-phosphate Dehydrogenase coupling system. Coupling enzymes were procured from Sigma (H8629-500UN) and diluted in 500ul of 50%

w/v glycerol, 10 UN per 1ml reaction mixture were added. 1.2 mM NADP was added to reaction mixture and increasing absorbance was measured at 340nm. The extinction coefficient 6220 AU $cm^{-1}$ $M^{-1}$ was used for calculation of $k_{cat}$. Final concentrations of the reaction mixture was 100 mM KCl, 20mM $MgCl_2$, 50 mM HEPES, pH 8.0, 0.5 mg/ml BSA, 10 mM D-Glucose. All enzyme concentrations were calculated as the fitted number of sites from an Ap5A ITC titration experiment on the stock solution. This procedure greatly reduces artifacts in the estimation of protein concentration as a result of inactivated protein. Final AK concentrations were in the range of $4x10^{-6}$ mM. Dilutions from stock concentration (0.04 mM) were made in Tris pH 7.5, 0.5 mg/ml BSA. Final ADP concentrations were in the range of 15 mM, and were increased until no appreciable change in activity was noticed. Initial velocities were non-linear in the first 50 seconds of observation. Although unexplained, this observation was highly reproducible. Therefore, the first 50 seconds (the non-linear region) of observation were excluded in all analysis.

## 5.3 RESULTS AND DISCUSSION

### 5.3.2 Measurement of $k_{cat}$

$k_{cat}$ was quantified in the direction of ATP production, to avoid complications of high-substrate (AMP) inhibition that are typically seem in with this enzyme (60). High quality data was achieved by this method. Reference values and experimental errors are summarized in Table 5.3.1.1. Interestingly, two of the four surface exposed Val or Ala to Gly mutations investigated (a55g mutation to the AMP binding

domain was included in this analysis) increased the catalytic rate of AK. These designed mutations therefore partially recapitulate the adaptive trends commonly observed for psychrophilic enzymes (which are adapted for increased native state dynamics), i.e. they express the characteristic concomitant decrease in binding affinity and increase in catalytic rate. However, this data seems to open new question. What are the critical differences between the activating mutations (v135g and a55g) and those that preserve $k_{cat}$ (v142g and v148g). A clue to the answer of this dilemma is gained from ITC experiments with the a55g mutant. This mutant, in the AMPbd, induces no gross change in the observed enthalpy of Ap5A binding (data not shown), as opposed to the LID mutations. Thus, induced unfolding of the AMPbd is unlikely a strong feature of the native ensemble. Therefore, these effects might be mediated by a different set of conformational processes, most likely within the sub-ensemble of states where these residues remain folded.

Eyring analysis of the energy barrier of $k_{cat}$ was also performed. Figure 5.3.1.1 shows the expected linear trend of a (log) rate constant with respect to (inverse) temperature. Such data can be fit for the enthalpy of activation according to the linear Eyring equation.

$$\ln(k_{cat}) = \frac{\Delta S^{\#}}{R} - \frac{\Delta H^{\#}}{R \cdot T} \qquad \text{Equation 5-4}$$

Furthermore, the measured $k_{cat}$ was converted into a free energy barrier using the classic relationship of transition state theory;

$$k = \left(\frac{k_B T}{h}\right)\exp\left(\frac{-\Delta G^{\#}}{RT}\right), \qquad \text{Equation 5-5}$$

where $k$ is the relevant rate constant, and $k_B$, $h$ and R are the Boltzmann's constant, Plank's constant and the ideal gas constant, respectively. A summary of these conversions of the measured catalytic rates are recorded in Table 5.3.1.1. From this data we learn that the barrier to catalysis in the case of AK is primarily enthalpic. Furthermore, the perturbations that result in activation of a55g and v135g are primarily entropic in nature. This result is opposed to the trend seen for psychrophiles, were primary gains in catalytic rate come from reducing the enthalpy of activation (61). Nevertheless, in the context of our mutation strategy (Chapter 4) these changes are intriguing, as a dynamic explanation for these effects is most likely.

### 5.3.2 Relaxation Dispersion NMR Experiments

Relaxation-dispersion experiments were conducted on the v142g mutant. These studies were intended to quantify the kinetics of the local unfolding process under investigation, further investigate the similarity of this state to unfolding, as well as investigate a possible role for local unfolding in the catalytic cycle of AK (62, 63). Experiments conducted at 19°C on the substrate free state revealed that as expected the transverse relaxation rate of several of the resonances associated with the locally unfolding region are decreased by the refocusing pulses of the CPMG experiment, see Figure 5.3.2.1. This is the hallmark behavior of resonances that are involved in a chemical (conformational) exchange process on the milli- to microsecond timescale.

Considering the strong thermodynamic data suggesting a two state (thermodynamic) conformational process it seemed reasonable to attempt fitting the relaxation data to a two state (kinetic) chemical exchange model. Such a model depends on four parameters, namely the rate of unfolding ($k_{unf}$), the rate of folding ($k_{fold}$), the chemical shift difference between the folded and unfolded states ($\Delta\omega$), and the intrinsic transverse relaxation rate of the resonance in question ($R_{2,o}$). To isolate the effects of local unfolding, resonances known to be affected by the opening/closing conformational transition (determined by CPMG experiments on WT AK under substrate saturating conditions) were excluded from this analysis.

Interestingly, all remaining CPMG sensitive resonances assignable to the LID region can be fit using a $\Delta\omega$ calculated as a difference from the measured spectra and experimentally determined (parameterized) random coil chemical shifts, see Figure 5.3.2.1 (57). Furthermore the fitted rate constants for the individual resonances were similar, suggesting that a global treatment of the data as a concerted kinetic process may be warranted. Figure 5.3.2.1-b demonstrates the reasonable quality of such a global fit. Furthermore, the reduced Chi-Squared value for the global fit was not greatly increased from those achieved when fitting individual resonances. Furthermore, the fitted estimates of the rates of folding and unfolding were highly consistent between three data sets using the global fitting approach, see Table 5.3.2.1. However, the reduced Chi Squared is slightly higher (2.0 vs. < 1.5) than would be expected for Gaussian error (of the size estimated) surrounding the model function. This could be explained by some small but real asynchrony of residues in the folding process, or relaxation due to other un-

accounted for exchange processes. Several experimental factors, not considered, such as pulse imperfections or off resonance effects could also account for this discrepancy.

The fitting results suggest that the population of the locally unfolded state is 4.3±0.27% (standard error) at 19°C. Based on the calorimetric and fitting analysis presented in Chapter 3, this probability is estimated to be 7.5±0.74% (95% CI). This is reasonable qualitative agreement, considering the substantial differences in solution conditions for the two experiments. It is important to note that not all resonances identified as unfolding by NMR in Chapter 3 have measurable relaxation dispersion profiles in this experiment. It is possible, if not likely, that simply not all residues experience a fully unfolded like state in this excited conformation of the LID (possibly just at low temperatures). Nevertheless we interpret the finding of 8 resonances experiencing random coil chemical shifts as being a strong indication of state highly similar to local unfolding.

### 5.3.3 SUMMARY AND CONCLUSIONS

The experiments presented here demonstrate that the rate of local unfolding is slower that the catalytic rate, we have also confirmed this result in the presence of substrate at 25°C (data not shown). Therefore, it is clear that local unfolding is not an obligatory intermediate for catalysis, as some authors have suggested (62, 63). However, the time scales of the two processes are not mismatched with a great severity. Unfortunately, we were not able to gather temperature dependence data

for the unfolding rate by NMR (although in principle it might be possible). Therefore, we are not able to predict to what degree the rate of unfolding might catch up with or $k_{cat}$ with increasing temperature, possibly becoming a significant (yet auxiliary) pathway of substrate release. However, we can at least infer that this hypothesis is not totally unreasonable. The $\Delta H^{\ddagger}$ of $k_{cat}$ is ~12 kcal/mol, and the $\Delta H_{conf}$ of the local unfolding process is ~33 kcals/mol. A peculiar feature of folding reactions is that the enthalpy of activation of an unfolding process is generally intermediate to the end states (64). Nevertheless, even within this constraint, it is clear that the temperature dependence of $k_u$ could very well be greater than that of $k_{cat}$. Clearly, the presence of substrate is also expected to have some significant effect on this rate, and this effect would needs to be carefully considered.

Although substrate dissociation via local unfolding might be plausible at higher temperatures, the results of our kinetic experiments seem to indicate that $k_{cat}$ can be modulated by some mutations which do not induce partial unfolding as seen in the LID domain (i.e. a55g), furthermore the Val to Gly mutations which seem to have very similar effects on local unfolding do not mediate a consistent effect on $k_{cat}$. Therefore the context of the presented data it seems likely that other types of conformational processes, possibly faster time scale dynamics are the key to the adaptive modulation of $k_{cat}$ in AK.

**Table 5.3.1.1: Representative Measured Values of $k_{cat}$ in the Direction of ATP Formation, for WT, v142g and v135g.**

| Temperature (°C) | $k_{cat}$-WT | $k_{cat}$-v142g | $k_{cat}$-v135g |
|---|---|---|---|
| 15 | 67±8 | 66±4[a] | 100±3 |
| 20 | 99.3±0.9 | 96±4[a] | 138.5±0.2 |
| 25 | 138±1 | 133.9±0.1 | 204±6.5 |
| 30 | 194.5±0.5 | 191.5±0.4 | 292±7 |

[a]The average experimental error of all $k_{cat}$ measurements is reported as no direct repeat of this experiment is available

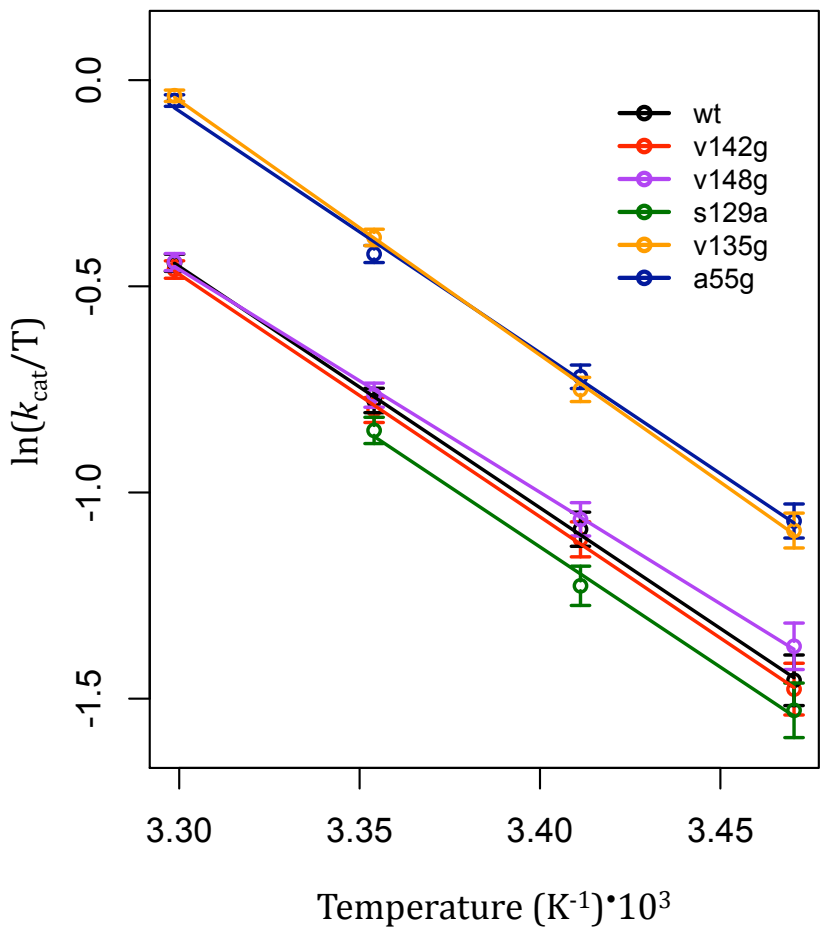



**Figure 5.3.1.1: Eyring Plot of $k_{cat}$ as a Function of Temperature.**

*Lines* – linear fits. *Error bars* – average experimental (standard) error of the data set.

**Table 5.3.1.1: Eyring Analysis of the Free Energy of Activation of Catalysis in AK ($k_{cat}$).**

| | $\Delta H^{\ddagger}$ (kcal/mol)[a] | $\Delta G^{\ddagger}$ (kcal/mol)[b] | $T\Delta S^{\ddagger}$ (kcal/mol)[c] |
|---|---|---|---|
| WT | 11.7±0.19 | 14.52±0.017 | -2.8±0.20 |
| v142g | 11.7±0.18 | 14.54±0.018 | -2.9±0.18 |
| v148g | 10.7±0.20 | 14.52±0.017 | -3.8±0.20 |
| v135g | 12.3±0.20 | 14.29±0.011 | -2.0±0.20 |
| s129a | 11.5±0.79 | 14.57±0.018 | -2.8±0.79 |
| a55g | 11.7±0.39 | 14.31±0.012 | -2.6±0.39 |

[a]Activation free energy barrier at 25°C. Error is estimated by profile likelihood method for least-squares fitting of the linear Eyring equation.
[b]Error is estimated by error propagation of the experimental error of determining $k_{cat}$.
[c]Error is estimated by error propagating the errors of $\Delta H^{\ddagger}$ and $\Delta G^{\ddagger}$; $\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger}$.

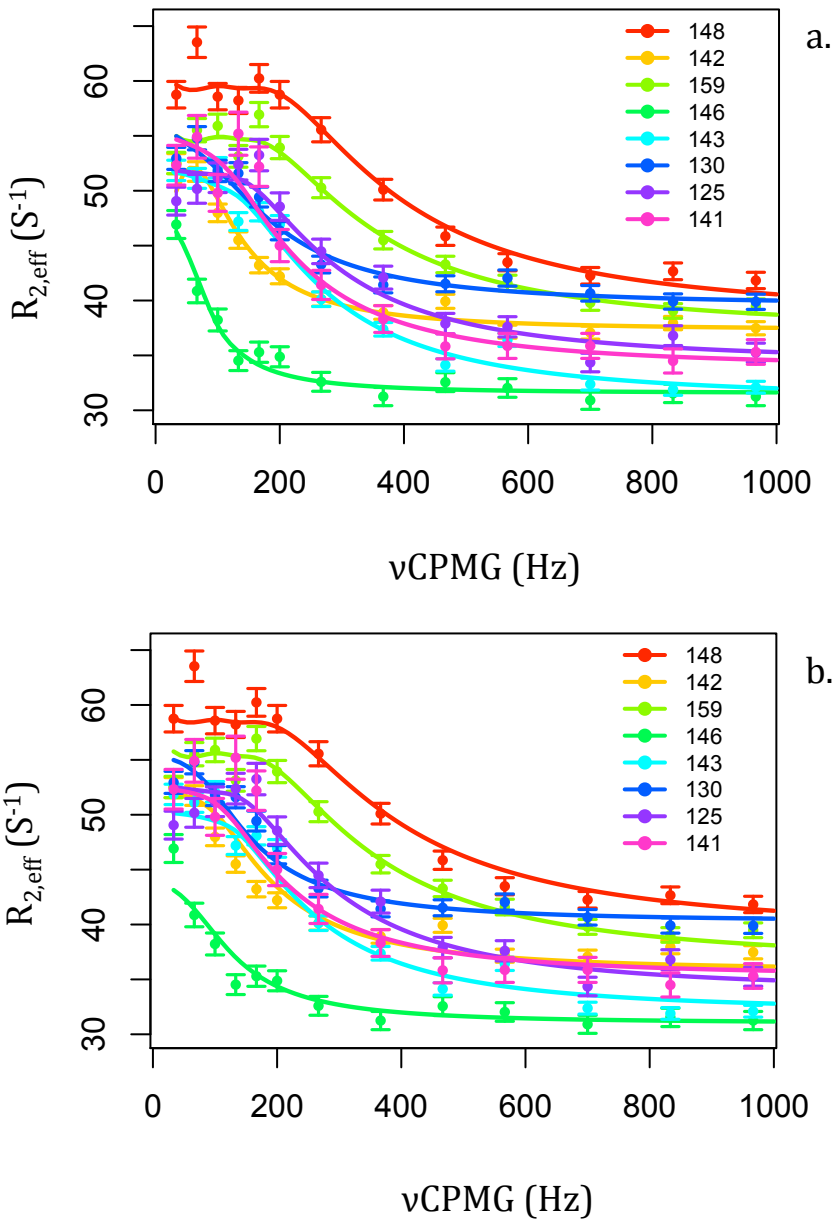



**Figure 5.3.2.1: $^{15}$N Relaxation Dispersion Profiles for Select Residues in the LID Region of v142g AK.**

Data represents one of three repeat experiments under this condition; 50mM MOPS, pH 7.85, 50mM NaCl, 20mM $MgCl_2$, 5% glycerol w/v, 8mM DTT, 8mM TCEP, stored under argon gas. ***points*** - $R_{2,eff}$ calculated from Equation 5.1. ***error bars*** - standard error estimated from the spectral baseline noise. ***lines*** - fitting functions. **a.** Relaxation profiles of amide resonances are fit individually. **b.** Relaxation profiles of amide resonances are fit globally as a group. See Table 5.3.2.1 for fitted parameter values.

**Table 5.3.2.1: Fitting Analysis of v142g Apo-State CPMG Relaxation Dispersion Data.**

| Residue | $k_{unfold}$[a] | $k_{fold}$[a] | $P_{unfolded}$[b] (%) | $\Delta\omega_{unf}$[c] | $\chi^2/N$[d] |
|---|---|---|---|---|---|
| 125 | 18±1.3 | 570±100 | 3.0±0.56 | 3.29 | 1.43±0.40 |
| 130 | 31±2.6 | 670±46 | 4.4±0.45 | 1.70 | 1.86±0.036 |
| 141 | 28±4.8 | 600±120 | 4.5±0.11 | 2.35 | 1.2±0.34 |
| 142 | 13±1.4 | 140±36 | 8.5±0.22 | 2.08 | 2.8±0.35 |
| 143 | 24.3±0.55 | 660±22 | 3.5±0.14 | 2.86 | 3±1.7 |
| 146 | 12±4.4 | 200±150 | 5.7±4.5 | 1.30 | 1.1±0.75 |
| 148 | 20±1.3 | 290±72 | 6.5±0.15 | 4.60 | 1.4±0.36 |
| 159 | 19±5.8 | 530±430 | 3.5±0.29 | 4.10 | 0.7±0.10 |
| Global | 19±0.30 | 420±27 | 4.3±0.27 | NA | 2.0±0.36 |

[a]Average rate constants and standard error from three direct repeats of complete experimental and fitting procedure. Experimental Conditions were 19°C, in the absence of substrate.
[b]Caculated from the presented rate constants, with propogated standard error.
[c]The expected change in chemical shift upon unfolding this residue based on random coil values.
[d]Average and standard deviation of the reduced Chi-Squared value for the 3 individual fits.

# CHAPTER 6

# COMPUTATIONAL INVESTIGATIONS OF NATIVE STATE ENSEMBLE OF KINASE: ALLOSTERY AND LIGAND BINDING

## 6.1 INTRODUCTION

The analysis presented to this point considers local unfolding within the native state ensemble as a single excited state of the adenylate kinase (AK) native state. However, Hilser and colleagues developed an algorithm, COREX, which allows the computational modeling of the protein native state as the set of all possible states of local unfolding (36, 66, 65). We have employed this technique to examine several different aspects of the AK native state, to allow comparison with the presented data, and to suggest avenues for further experimental investigation. In particular we wish to look for features of the native state ensemble that might yeald a testable hypothesis as to why only v135g and a55g increase the catalytic rate of AK.

## 6.2 MATERIALS AND METHODS

COREX Simulations: Monte-Carlo sampled COREX ensembles were generated with a window size of 8 residues and a minimum window size of 4 residues. 5000 states were generated per partition, yielding 40,000 states per ensemble. Ensembles were

generated for the "open" conformation using PDBID 4AKE and for the "closed " conformation using PDBID 1AKE (26, 25). All analysis of COREX ensembles presented was calculated using source code programmed by T. Schrank. A clear summary of the COREX algorithm has been recently published; therefore, only a highly simplified explanation will be presented here (65). Briefly, using a surface area based parameterization of the energetics of unfolding; COREX can be used to approximate the $\Delta G$, $\Delta H$ and $\Delta C_p$ of unfolding of different regions of the protein structure. Applying this calculation to all possible states of folding yields an energetically related ensemble states which can be analyzed together to model the behavior of the native state as a whole. A scaling factor, entropy-weighting factor (EWF) is applied to the conformational entropy terms utilized in COREX. This value is generally calibrated to reproduce the experimental melting temperature ($T_m$) of the protein in question. This can be achieved by one of two methods, either using only two states of the protein, folded and unfolded, and adjusting the EWF to achieve equal probabilities at Tm, by an ensemble approach, where a target of an ensemble average of 50% of residues are unfolded is optimized.

## 6.3 RESULTS AND DISCUSSION

### 6.3.1. Calculation Residue Specific Stability Constants

Residue specific stability constants were calculated for both the open and closed states of AK. The stability constant is a measure of the probability of the protein being in a state where the residue in question is unfolded. Formally,

$$\kappa_{f,j} \equiv \frac{\sum P_{f,j}}{\sum P_{u,j}} \qquad \text{Equation 6-1}$$

where $P_{f,j}$ represents the probability of microstates where residue *j* is folded, and conversely $P_{u,j}$ represents the probability of microstates were residue *j* is unfolded (36). Figure 6.3.1 compares the stability constant for open and closed conformational of AK. It is clear that in the open conformation, the region surrounding residues 50 and 150, roughly corresponding respectively to the Adenosine-5'-monophosphate binding domain (AMPbd) and LID regions of AK are also regions of relatively low stability in the native state ensemble. This is an experimentally reasonable result as these regions have been demonstrated to be highly dynamic on a wide range of timescales, however this result agrees particularly with strongly decreased order parameters calculated from $^{15}N$ relaxation experiments (10, 31). Furthermore, domain-swapping experiments in AKs from *Bacillus sp.* have demonstrated that swapping of these domains can transfer the $k_{cat}$ from one ortholog to another (67). Therefore, the thermodynamic stability of these domains might be important for AK function. Furthermore, the AMPbd is preferentially stabilized in the closed conformation. The outermost portion of the LID domain, residues 138-155, on the other hand are apparently destabilized by the open to closed transition. This counter intuitive effect will be discussed below.

### 6.3.2 Structural Corollaries to Local Stability

To investigate the local structural determinants of the COREX ensemble, we have extracted the local contributions (+/- 7 residues) of the three parameters which enter into the COREX calculation of energy for each state, namely the average per residue change in conformational entropy up unfolding and change in solvent-exposed surface area, both polar and apolar, for any given residue, see Figure 6.3.2.1. From these data, we can see the general structural reasonability of the COREX stability profile, for example residues 140-150 appear to be a local minimum in change in accessible surface area ΔASA upon unfolding, thus confirming the expected correlation between solvent exposure and structural instability. Furthermore, it appears that the local instability of the LID and AMPbd regions is primarily driven by solvent exposure, rather than conformational entropy. More specifically, the regions surrounding residues 50 and 145 appear to be associated with local minima in the per residue contribution to the $T\Delta S_{conf}$ (conformational entropy) of local unfolding, opposing the surface area effects which are imposed by the shape of the protein fold.

### 6.3.3 Distribution of Native Conformers in Terms of ΔH and ΔCp

To investigate the similarity between the measured ΔH and ΔCp values reported for the local unfolding reaction of the LID domain region and states within the native state ensemble as predicted by COREX, we have organized the native state ensemble according to change in heat and heat capacity relative to the fully folded state. In Figure 6.3.3.1 the free energy difference between the sub-ensembles in and out-side of the ΔH/ΔCp subspace queried are plotted for the COREX ensemble

for the open state of AK, see legend for details. In this data we can see two main energy wells in the calculated ensemble, one corresponding to the highly folded, native like states and one to highly unfolded, denatured like states. Interestingly, the most probable ΔH in the native energy well (relative to the crystal structure) is quite large, approximately 16 kcal/mol. This is in qualitative agreement with the presented experimental data that demonstrates native state fluctuations to high enthalpy states.

However, the measured value (~33 kcal/mol) seems to be in significant disagreement with the predicted energy minimum for the native state. The reason for this disagreement is an unexplained high degree of cooperativity between residues comprising the approximate LID region. As discussed in Chapter 3, the COREX energy function does an excellent job of reproducing the experimental ΔH and ΔCp of the experimentally determined cooperative unit of local unfolding in the LID region. Therefore, together this notable quantitative agreement and gross disagreement begs explanation. Several authors using molecular dynamics simulations have suggested that mechanical strain as a consequence of gross structural motions of the LID and AMPbd may increase the probability of unfolding in these regions (62, 63, 68). Possibly, such mechanical contributions, which are not accounted for in COREX analysis, could explain these discrepancies. However, our data suggest that some degree of local instability of the binding associated sub-domains of AK should be expected based on the sequence and static structures of AK alone.

### 6.3.4 Effects of Ligand Binding on the AK Ensemble

Ligand association has been known to change the dynamic properties of proteins for many years. Lieu et al. have demonstrated that residues in binding sites are indeed poised in thermodynamic space in such a way to have greater than average ability to modulate the stability of other positions in the protein (39). The effects of ligand binding can be modeled with COREX. We will consider that ligand association is only probable with states of the ensemble where all residues in the binding site are structured. Therefore, dividing the ensemble into states which can bind, binding competent (BC), and states which do not bind, binding incompetent (BI), we can calculate the effects of ligand association using the following equations:

$$K_{bc} = e^{\frac{\Delta G_{pert}}{-RT}} \qquad \text{Equation 6-2}$$

$$K_{corex} = e^{\frac{\Delta G_{corex}}{-RT}} \qquad \text{Equation 6-3}$$

$$Q = \sum_{bi} K_{corex} + \sum_{bc} K_{corex} \cdot K_{bc} \qquad \text{Equation 6-4}$$

$$P_{i,bc} = \frac{K_{i,corex} K_{i,pert}}{Q}; P_{i,bi} = \frac{K_{i,corex}}{Q} \qquad \text{Equation 6-5}$$

where $\Delta G_{corex}$ free energy of a given state according to the COREX energy function, and $\Delta G_{pert}$ represents the stabilizing perturbation a ligand effects on binding competent states.

To define the residues contributing to the binding pocket, I have selected all residues in the structure 1AKE that contain atoms that are within 4 angstroms of the associated inhibitor, Ap5A. The inhibitor Ap5A is a useful as it is a very close mimic of the biological substrate of AK. Specifically, it is chemically identical to two adenosine-5'-diphosphate (ADP) molecules, were the terminal phosphate groups are joined together by an additional phosphate group. Therefore, the binding sites for individual ligands such as Adenosine-5'-triphosphate (ATP), can be defined similarly, by simply removing the phosphate groups not corresponding to the ligand in question. The stability constants of all residues in the closed structure are plotted in figure 6.3.4.1, for both the bound and unbound states.

Inhibitor binding seems to, in general, promote the same trends in residue specific stability as the open to closed transition. This is a biologically reasonable result as ligand binding induces the open to closed transition. In particular, the region 138-155 becomes progressively more destabilized upon ligand association. This allosteric modulation of the stability of this region is very interesting from a thermodynamic perspective. It demonstrates that the native state ensemble of AK can use small changes in probability to effect allosteric work. Also we can see that, to a small degree, adenosine-5'-monophosphate (AMP) and ATP binding effect allosteric stabilization of residues involved in the other site.

### 6.3.5 Calculation of Residue Specific Allosteric Connectivity

In terms of an ensemble, allosteric communication is a property determined by energetic relationship of states within the ensemble. In other words, changing

the free energy of a subset of states within the ensemble redistributes the energetic differences between these and other states. In terms of the COREX representation of the native state, it is instructive to analyze energetic connectivities between residues by changing the free energy of all states where a residue of interest is unfolded, and observing the change in stability at other sites. Such a calculation was performed for each residue of AK, and was defined as follows (69):

$$\Phi_{f,k} = e^{\frac{\Delta G_{f,k}}{-RT}} \qquad \text{Equation 6-6}$$

$$\kappa_{f,j}^{k_{mut}} = \frac{\Phi_{f,j} \cdot \sum P_{f.j|f,k} + \sum P_{f,j|nf,k}}{\Phi_{f,j} \cdot \sum P_{nf.j|f,k} + \sum P_{nf,j|nf,k}} \qquad \text{Equation 6-7}$$

where all state with residue *k* folded are perturbed by an assigned free energy of $\Delta G_{f,k}$, and stability constant of other residues *j*, $\kappa_{f,j}$, are then evaluated.

The results are displayed in Figure 6.3.5.1. Interestingly, stabilizing residues 138-155 destabilizes the remainder of the protein. However, even applying a strong perturbation ($\Delta G_{f,k}$ = 5kcal/mol), no residues in the core of the protein seem destabilize the 138-155 region, as in the case of binding. Therefore, although anti-correlation of stability of 138-155 and other residues is seen in both cases, the ability of substrate to destabilize 138-155 is found to be a consequence of the set of particular residues in the binding site and stabilizing them as a group. In other words, arbitrary stabilization of other sites will not reproduce this effect. Therefore, this result highlights significance of ligand binding to destabilize this region.

Interestingly, similar calculations (unpublished results) have been done on several orthologous AKs from *bacillus sp.* and in all cases investigated, this region was found to also be negatively coupled to the rest of the protein, suggesting that some biological function might drive this evolutionary conservation.

### 6.3.6 Calculation of Calorimetrically Observable Ensemble Effects on Ligand Binding

From the above simulations we can see that ligand binding redistributes of states within the protein native state ensemble. As these states are states of local folding, each is associated with relatively large differences in solvent accessible surface area (ASA) relative to the fully folded crystal structure. The thermodynamic consequences of such changes have been well characterized. In fact, the parameterization of ΔH, ΔCp, and ΔS in terms of changes in polar and apolar ΔASA is the basis of the COREX energy function. Therefore, it is possible to calculate the expected contributions of ensemble re-distribution to calorimetric observables, namely the apparent free energy, enthalpy and heat capacity of binding. The free energy of binding can be defined as

$$\Delta G_{bind} = -RT \ln\left(\frac{P_{bound}}{P_{unbound}}\right). \qquad \text{Equation 6-8}$$

We can expand this definition to consider states of the ensemble that can and cannot bind the ligand. Binding competent (BC) or high-affinity states will be assumed to have all binding site residues folded, and alternately states of the ensemble that

disrupt the binding site by unfolding residues involved in stabilizing interactions will be considered binding incompetent (BI). Therefore it follows

$$\Delta G_{bind,app} = -RT\ln\left(\frac{K_{a,int}\sum K_{i,BC}}{\sum K_{i,BC} + \sum K_{i,BI}}\right) \qquad \text{Equation 6-9}$$

$$\Delta G_{bind,app} = -RT\ln(K_{a,int}) + RT\ln\left(\frac{\sum K_{i,BC} + \sum K_{i,BI}}{\sum K_{i,BC}}\right) \qquad \text{Equation 6-10}$$

where $K_{a,int}$ is the association constant for the intact binding site to the ligand, and $K_i$ is the statistical weight of state *i* in the ensemble, either BC or BI. The first term is recognizable as the free energy of association and second term, is the free energy cost of depopulating all BI states, or the ensemble contribution, $\Delta G_{bind,ens}$.

$$\Delta G_{bind,ens} = RT\ln\left(\frac{\sum K_{i,BC} + \sum K_{i,BI}}{\sum K_{i,BC}}\right) \qquad \text{Equation 6-11}$$

Similarly, the measured enthalpy of binding can be separated into terms corresponding to association and ensemble effects as follows;

$$\Delta H_{bind,app} = \Delta H_{int,assoc} - \left\langle \Delta H_{unf} \right\rangle_{unbound} + \left\langle \Delta H_{unf} \right\rangle_{bound} \qquad \text{Equation 6-12}$$

where $\langle\rangle$ denotes an ensemble average property. In terms of individual COREX states this can be written as;

$$\Delta H_{bind,app} = \Delta H_{int,assoc} - \sum\left(\Delta H_{i,unf}\frac{K_i}{Q}\right) + \sum\left(\Delta H_{i,unf}\frac{K_{i,BC}}{\sum K_{i,BC}}\right) \qquad \text{Equation 6-13}$$

where $\Delta H_{i,unf}$ represents the microstate enthalpy, as referenced from the fully folded state, BC denotes membership in the subset of all possible states which are binding competent. The ensemble associated change in entropy of binding can be easily calculated with the well know expression, $\Delta G = \Delta H - T\Delta S$.

Although heat capacity (Cp) can be defined in many ways, the most useful definition in our case is to consider Cp as the mean squared fluctuation in the enthalpy (70),

$$Cp = \frac{\langle \delta H^2 \rangle}{RT^2} = \frac{\langle (H - \langle H \rangle)^2 \rangle}{RT^2} \quad \text{Equation 6-14}$$

For the case of a protein binding reaction we can write:

$$\Delta Cp_{bind,app} = \Delta Cp_{int} + \Delta Cp_{ens,app} \quad \text{Equation 6-15}$$

$$\Delta Cp_{ens,app} = \langle Cp_{unbound} \rangle - \langle Cp_{bound} \rangle + \frac{\langle (\Delta H_i - \langle \Delta H_i \rangle)^2 \rangle}{RT^2} - \frac{\langle (\Delta H_{i,BC} - \langle \Delta H_{i,BC} \rangle)^2 \rangle}{RT^2} \quad \text{Eq. 6-16}$$

Therefore, the heat capacity contribution of the ensemble can also be calculated from the COREX ensemble.

The results of these calculations applied to the closed conformation of AK over a range of temperatures are summarized in Figure 6.3.6.1. As is plainly seen large changes in the apparent energetics are realized at the global unfolding transition, however, non-trivial contributions to all thermodynamic parameters of binding are contributed from ensemble effects.

Interestingly, in the case of 1AKE it appears that the contribution of enthalpy to the apparent binding reaction is positive; this is probably a consequence of the linkage discussed earlier in the chapter, where ligand binding promotes unfolding of the outermost portion of the LID domain. This is contrary to the general expectation that ligand association will stabilize binding pocket-associated regions of low stability. Furthermore, this result is in contradiction to our direct experimental

measurement of these contributions. Therefore again, we see that although COREX captures the pattern of regional local stability of AK, the exact cooperative unit is not well characterized by this method. This may be due simple to incorrect characterization of the solvent exposure of residues deep in the active site of AK, as the binding pocket is highly occluded from the solvent in the closed crystal structure.

Although, clearly a bit artificial this calculation can be repeated on the open structure, to probe the conformational manifold that would be expected to exist starting from this reference conformation. Examining Figure 6.3.6.2 we can see that a relatively large contribution of enthalpy (~1.1 kcal) can be contributed to the binding reaction, at stabilizing but biologically relevant temperatures (37°C). This value is in reasonable agreement with the excess conformational enthalpy measured for the Ap5A binding to the WT protein at 37°C ~ 1.7 kcal/mol. It is also interesting to note (Figure 6.3.5.2-a) that the ensemble average change in residues folded at 37°C is less than 1% of a 214 amino acid protein, a change that would easily be missed by spectroscopic techniques, such as circular dichroism, see Chapter 2. Therefore, in the interpretation of thermodynamic of binding reactions, one must be cognoscente that modulation of the conformational ensemble may well contribute to the outcome measurements. Thus the interpretations of enthalpies and entropies measured by ITC in terms of the binding interface must be done with extreme caution (71). Others experimentally investigating native state pico- to nanosecond phenomena by NMR have demonstrated that such motions distributed widely

through the protein can have important effects on the thermodynamics of ligand association (72).

## 6.4 SUMMARY AND CONCLUSIONS

From these investigations of AK based on the COREX model, it is clear that the native state of enzymes are expected to be a collection of many states. Furthermore, association of proteins with small molecules such as inhibitors or nucleotides will have significant impact on the distribution of states accessible to the macromolecule. This expanded native state ensemble can explain many counterintuitive aspects of the protein native state.

Important findings from these investigations include the notable differences in the cooperative unit of LID unfolding as seen in experiments and predicted by COREX merit further investigation. Specifically, residues 110-120 are found by experiment to be part of the LID unfolding cooperative unit, but appear to be quite stable by COREX analysis. Furthermore, residues 125-138 are predicted by COREX to be slightly more stable than the outermost portion of the LID, this difference alone is probably the major reason for the difference between the expected and calculated distribution of $\Delta H$ of unfolding within the native state as seen in Figure 2. In contrast, the COREX energy function does an excellent job of capturing the measured thermodynamics of a given local folding reaction (see previous chapters), therefore this disagreement in the cooperativity of LID unfolding is indicative of something missing in the COREX model, for this particular protein. It is possible that

the large amplitude conformational motions of the LID domain, moving relative to the remainder of the protein are the key to this discrepancy. Therefore AK may be a good system for the exploration of possible relationships between conformational dynamics and residue specific local unfolding stabilities in proteins.

Also important is the uncovering of strong allosteric behavior in the fold of AK. The predominant allosteric feature the Adenylate Kinase ensemble is the strong negative coupling between residues 138-155 and the rest of the protein. This region is destabilized upon inhibitor binding as well as the closing reaction. Conversely stabilization of this region such as might be effected by a protein-protein interaction is predicted to strongly destabilize the core and active site of the protein. The conservation of this ensemble feature in bacillus AKs suggests that this may well be a biologically important aspect of the AK native state. This phenomenon has been the subject of no experimental investigation, and is therefore ripe with possibilities for future studies.

**CHAPTER 7**

## CONCLUDING REMARKS AND FUTURE DIRECIONS

Our work has demonstrated that local unfolding, at least of the LID region, is an important fluctuation of the Adenylate Kinase native state. Furthermore, this fluctuation is similar to global unfolding in terms of solvent exposure, conformational entropy as well and $^{15}$N chemical shift. This highly temperature dependant state (mildly) modulates binding affinity of WT AK at the ATP site. More importantly the disordered state itself state can be strongly and robustly modulated by point mutation to shift the ensemble to a lower affinity, as demonstrated by the almost identical behavior of three Val to Gly mutations as distinct structural locations. The magnitude of the affinity modulation is also demonstrated to be significant in relation to the temperature dependence of Ap5A binding. The mutation strategy applied in this work is a generalizeable tool that can be utilized to probe the native state ensemble of any folded protein for highly disordered states such as the ones observed here. Furthermore, we have demonstrated that this strategy provides a method for functional modulation in the presence of great structural conservation, and separating (at least some) structural and conformational contributions to binding reactions.

The herein presented data, however, also leaves many questions to be answered. In particular, the origin of the strong differences in $k_{\mathrm{cat}}$ modulation

between four glycine mutations, designed with the same strategy, needs to be addressed. Also, the biological role of the strongly conserved negative coupling in the outer portion of the LID domain predicted by COREX needs to be investigated. One hypothesis that could explain these intriguing results is that the COREX algorithm is correctly capturing the thermodynamic relationships of conformational fluctuations within the ligand associated sub-ensemble, which this study has not addressed. Accepting this, the v135g and a55g mutations, which are relatively stable and close to the active site, should shift the bound ensemble to lower affinity states (i.e. those which lose contacts with the ligand). This lower intrinsic affinity might be associated with more rapid release of reaction products. Conversely, destabilizing the positions 142 and 148, which are predicted to be highly unstable and negatively coupled to the rest of the protein (including the active site) should have a weakly stabilizing effect on the active site (73). Interestingly, the modest but robustly determined mutational effects on the *intrinsic* binding for Ap5A correlate well with these predictions. v135g and a55g decrease intrinsic affinity, while v148g slightly increases it, and v142g has no effect. Based on these observations, one might also propose that molecular interactions that *stabilize* the 142-148 region should increase the catalytic rate of AK by allostericaly destabilizing the active site, thus providing an intuitive biological function for the predicted negative coupling in the LID.

In conclusion, the biological function of proteins should be expected to depend on many conformations thermodynamically or kinetically accessible to the macromolecule. As an example of this point, recent work comparing a thermophilic

and mesophilic pair of Adenylate Kinases was investigated in terms of milli- to microsecond dynamics, pico- to nanosecond dynamics, thermodynamic stability and $k_{cat}$. Somewhat surprisingly all, and even the most seemingly functionally irrelevant parameter, the free energy of global unfolding, were precisely conserved, at physiological temperature of each molecule (10). I take this result in combination with those presented here to indicate that biological enzymes should be viewed not as a static scaffold for the active site, but rather a set of molecular interactions which all contribute to tune the thermodynamic and kinetic relationships between the set of all available conformational states, which work in combination to optimize fitness.

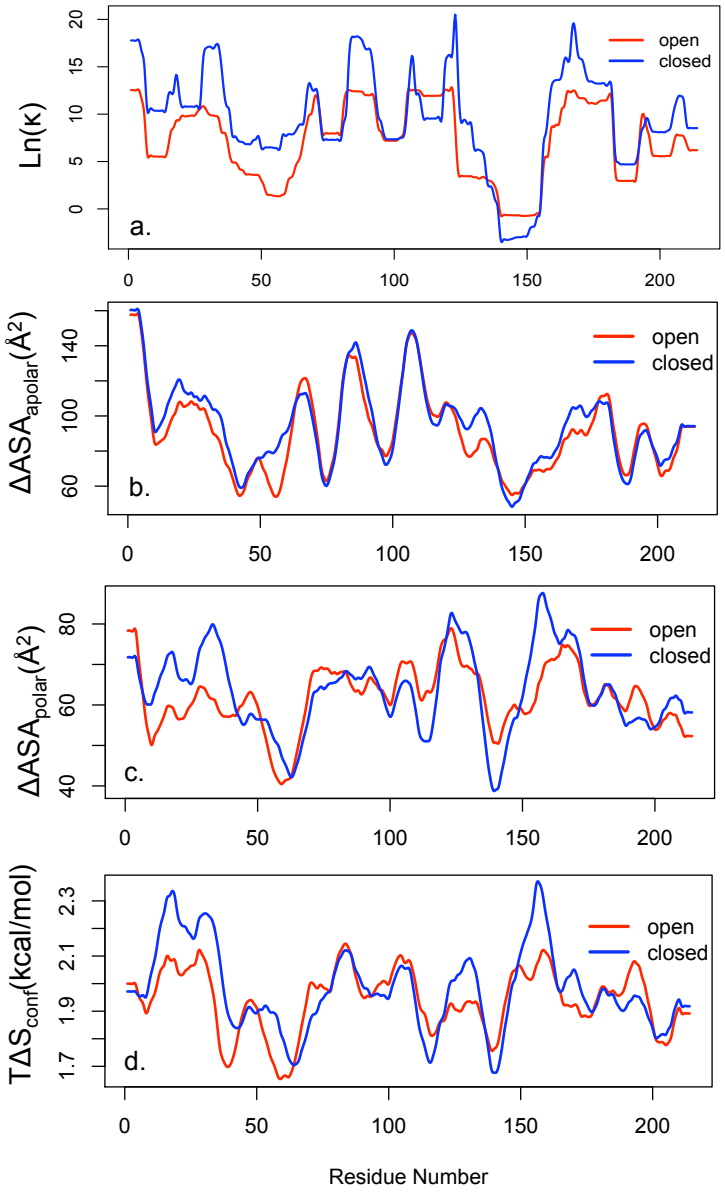



**Figure 6.3.1.1: Regional Stability and Structural Corollaries in AK.**

**a.** Residue specific stability constant, κ (Equation 6-1), plotted for the open (PDB 4AKE) and closed (PDB 1AKE) conformations. EWF = 0.9735 calibrated by the two state method for 4AKE, $T_m$ = 55°C. **b-c.** Structural properties calculated by COREX considering only states where only 8 contiguous residues, containing the probe residue in question (y-axis), may be unfolded. Effectively, this calculation produces a single sliding window (8 residues) unfolding, moving along the protein chain. Values are reported as per amino-acid averages of all such states in the ensemble.

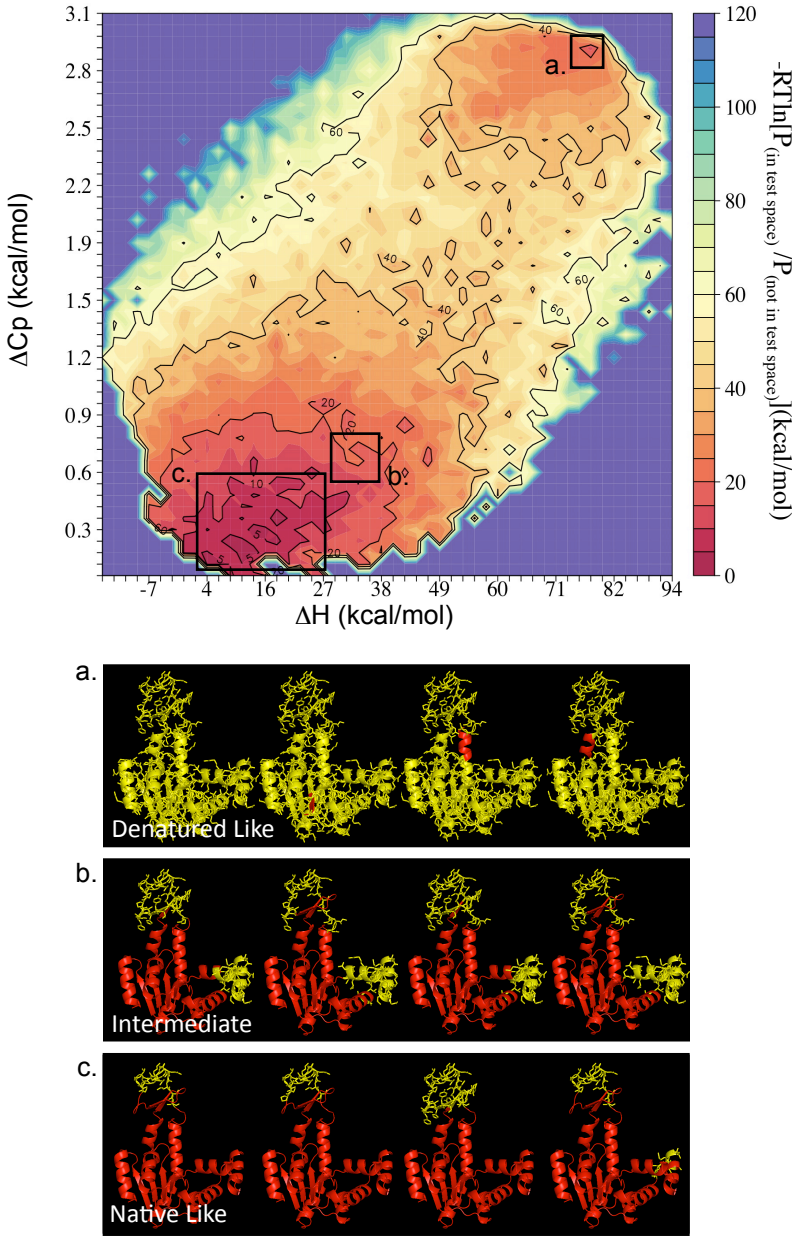



**Figure 6.3.3.1: Analysis of Enthalpy and Heat Capacity Differences of Native State Ensemble Constituents.**

The free energy of populating a given region of ΔH/ΔCp space was calculated as, $-RT \cdot \log (P_{in}/P_{out})$; were *in* represents all states with ΔH and ΔCp within the plotted subspace, and *out* represents the remainder of states within the ensemble. Un-sampled/impossible regions of thermodynamic space were assigned a value of 120 kcal/mol. **a-c.** Boxes highlight areas of interest in thermodynamic space. *c* and *a* correspond to the most probable states in the native and denatured states, respectively. *b* corresponds to the region of thermodynamic space compatible (within experimental error) of the measured thermodynamics of the dominant native state local unfolding event in AK. EWF = 0.9735 calibrated by the two state method for 4AKE, Tm = 55°C.

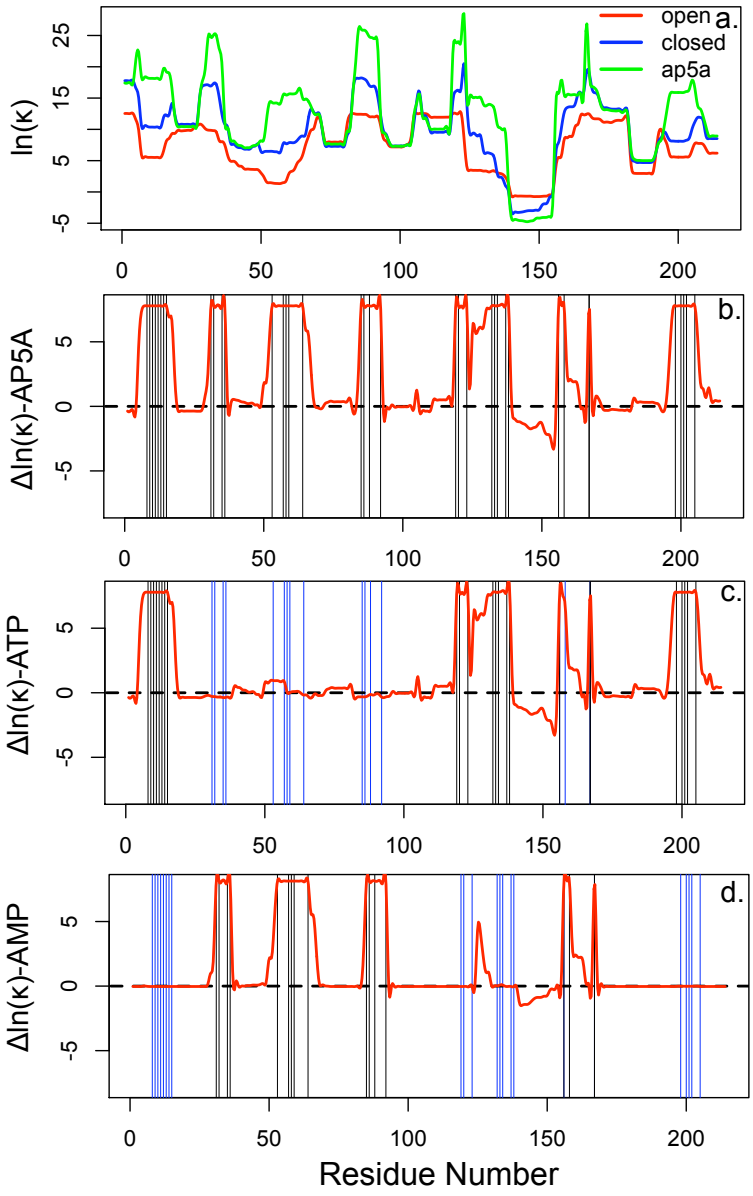



**Figure 6.3.4.1: Effects of Ligand Association on the AK Ensemble.**

Stability constants were calculated according to Equations 4.1-5, using $\Delta G_{pert}$ = 5 kcal/mol. **a.** *red*- the open (PDB 4AKE) conformation; *blue*- the closed (PDB 1AKE) conformation; *green*- closed plus Ap5A bound state. **b-c.** Change in ln(κ) for the stabilization of binding sites for three AK ligands, referenced from the closed conformation (as in a.). *Black vertical hashes*- represents the residues included in the binding site, i.e. those resides which must be folded to bind. *Blue vertical hashes*- residues of the other (non-perturbed) binding sites, marked to highlight site-to-site communication. EWF = 0.9735 calibrated by the two state method for 4AKE, Tm = 55°C.

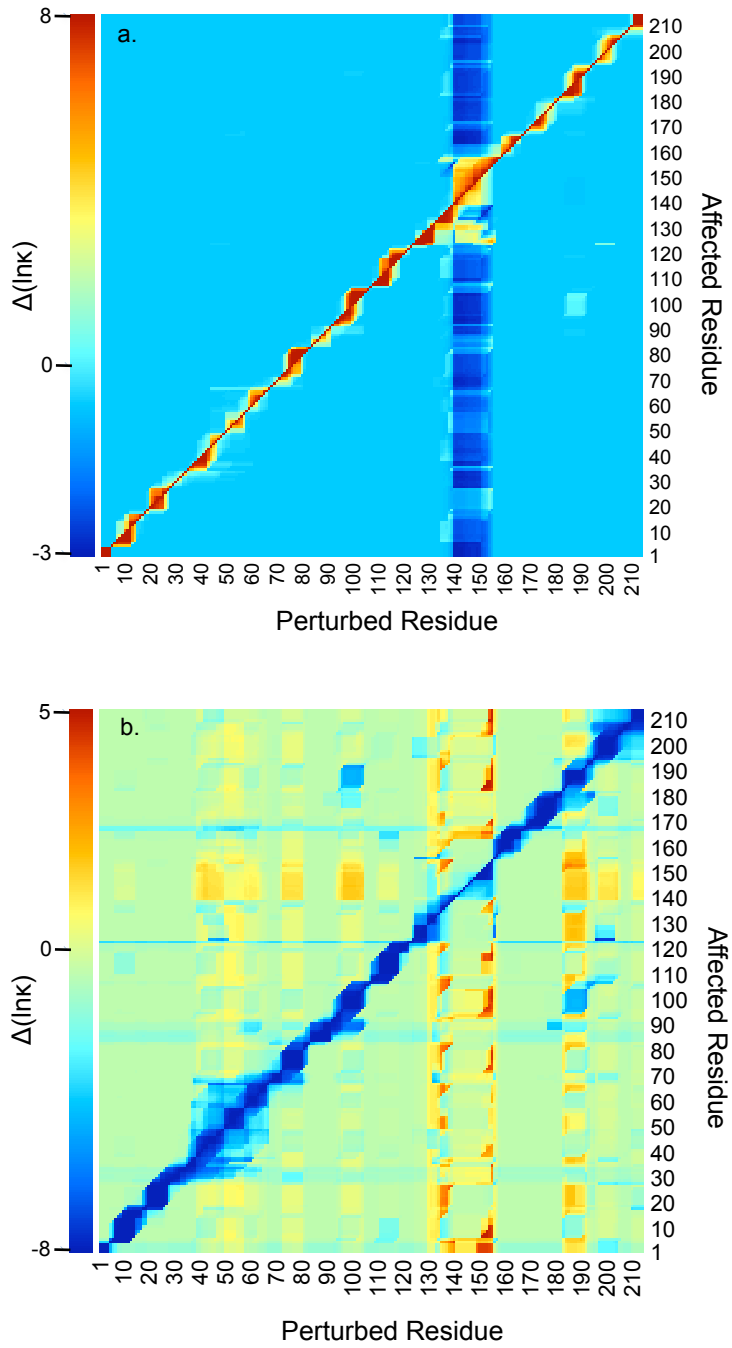



**Figure 6.3.5.1: Closed Ensemble Allosteric Response to Perturbation.**

Stability constants were calculated for the closed structure (1AKE) in the presence of a perturbing thermodynamic mutation to a given residue *j* according to Equations 6.6-7, using $\Delta G_{f,k}$ = +/-5 kcal/mol. Change in ln(κ) was calculated as a difference from the unperturbed closed (1AKE) ensemble. **a.** Stabilizing mutation to the perturbed residue. **b.** Destabilizing mutation to the perturbed residue. EWF = 0.9735 calibrated by the two state method for 4AKE, Tm = 55°C.

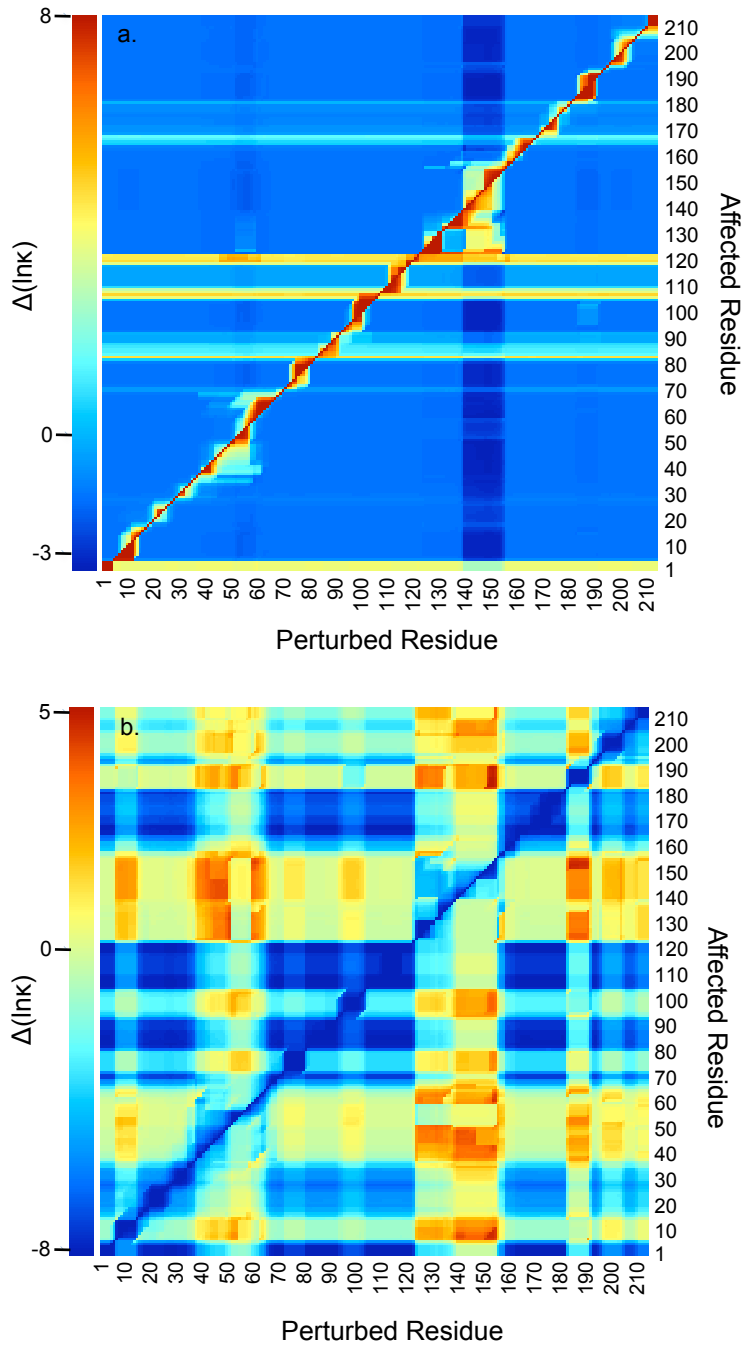



**Figure 6.3.5.2: Open Ensemble Allosteric Response to Perturbation.**

Stability constants were calculated for the closed structure (4AKE) in the presence of a perturbing thermodynamic mutation to a given residue *j* according to Equations 6.6-7, using $\Delta G_{f,k}$ = +/-5 kcal/mol. Change in ln(κ) was calculated as a difference from the unperturbed open (4AKE) ensemble. **a.** Stabilizing mutation to the perturbed residue. **b.** Destabilizing mutation to the perturbed residue. EWF = 0.9735 calibrated by the two state method for 4AKE, Tm = 55°C.

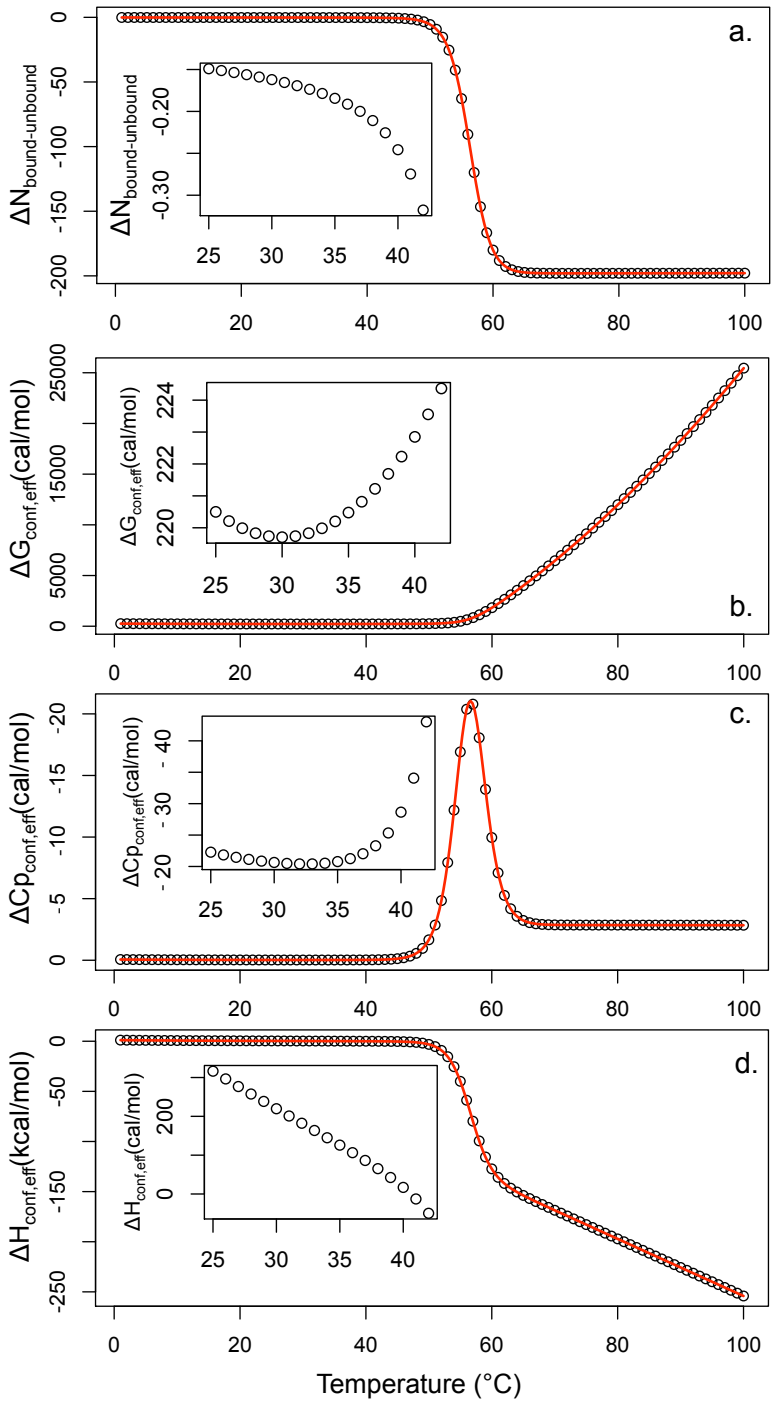



**Figure 6.3.6.1: Predicted Ensemble Effects on the Calorimetric Observables of Ap5A Binding.**

Simulations based on the closed conformation of AK (PDB ID 1AKE). **a.** The change in the ensemble average number of residues unfolded upon ligand binding, $\Delta N_{bound\text{-}unbound}$. **b.** The additive contribution of ensemble effects to the observed free energy of binding, $\Delta G_{conf,eff}$. **c.** The ensemble contribution to the observed change in heat capacity of binding, $\Delta Cp_{conf,eff}$. **d.** The ensemble contribution to the observed enthalpy change of binding, $\Delta H_{conf,eff}$.

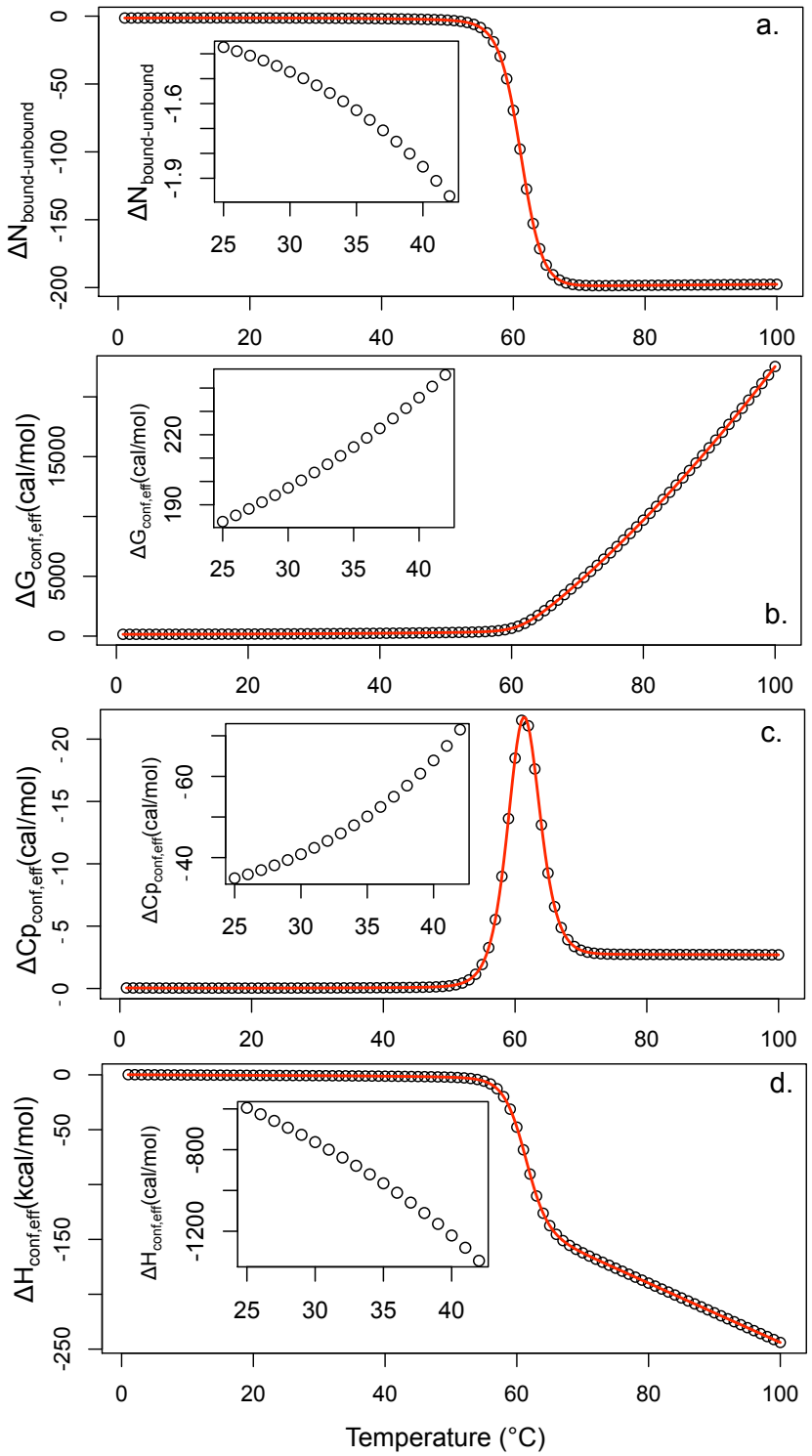



**Figure 6.3.6.2: Predicted Ensemble Effects on the Calorimetric Observables of Ap5A Binding.**

Simulations based on the open conformation of AK (PDB ID 4AKE). **a.** The change in the ensemble average number of residues unfolded upon ligand binding, $\Delta N_{bound\text{-}unbound}$. **b.** The additive contribution of ensemble effects to the observed free energy of binding, $\Delta G_{conf,eff}$. **c.** The ensemble contribution to the observed change in heat capacity of binding, $\Delta Cp_{conf,eff}$. **d.** The ensemble contribution to the observed enthalpy change of binding, $\Delta H_{conf,eff}$.

**REFERNECES**


1. Röthlisberger D et al. (2008) Kemp elimination catalysts by computational enzyme design. *Nature* 453:190-195.

2. Georlette D et al. (2004) Some like it cold: biocatalysis at low temperatures. *FEMS Microbiol. Rev* 28:25-42.

3. Fields PA, Somero GN (1998) Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. *Proc Natl Acad Sci U S A* 95:11476-81.

4. Závodszky P, Kardos J, Svingor, Petsko GA (1998) Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. U.S.A* 95:7406-7411.

5. Wolf-Watz M et al. (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat Struct Mol Biol* 11:945-9.

6. Henzler-Wildman K, Kern D (2007) Dynamic personalities of proteins. *Nature* 450:964-972.

7. Bosco DA, Eisenmesser EZ, Pochapsky S, Sundquist WI, Kern D (2002) Catalysis of cis/trans isomerization in native HIV-1 capsid by human cyclophilin A. *Proc Natl Acad Sci U S A* 99:5247-52.

8. Henzler-Wildman KA et al. (2007) Intrinsic motions along an enzymatic reaction trajectory. *Nature* 450:838-44.

9. Miyashita O, Onuchic JN, Wolynes PG (2003) Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci. U.S.A* 100:12570-5.

10. Henzler-Wildman KA et al. (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* 450:913-916.

11. Holland LZ, McFall-Ngai M, Somero GN (1997) Evolution of lactate dehydrogenase-A homologs of barracuda fishes (genus Sphyraena) from different thermal environments: differences in kinetic properties and thermal stability are due to amino acid substitutions outside the active site. *Biochemistry* 36:3207-15.

12. Hot spots in cold adaptation- localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes.pdf.

13. Závodszky P, Kardos J, Svingor, Petsko GA (1998) Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc Natl Acad Sci U S A* 95:7406-11.

14. Feller G, Gerday C (2003) Psychrophilic enzymes: hot topics in cold adaptation. *Nat Rev Micro* 1:200-208.

15. Somero GN, Hochachka, PW (1984) *Biochemical Adaptation* (Princeton, Princeton, New Jersy).

16. Deng H et al. (1994) Source of catalysis in the lactate dehydrogenase system. Ground-state interactions in the enzyme-substrate complex. *Biochemistry* 33:2297-305.

17. Aghajari N, Feller G, Gerday C, Haser R (1998) Structures of the psychrophilic Alteromonas haloplanctis [alpha]-amylase give insights into cold adaptation at a molecular level. *Structure* 6:1503-1516.

18. Glasstone S, Laidler K, Eyring H (1941) *The Theory of Rate Processes* (McGraw-Hill Book Company, Inc., New York, New York).

19. Auton M, Holthauzen LMF, Bolen DW (2007) Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proc. Natl. Acad. Sci. U.S.A* 104:15317-22.

20. Auton M, Bolen DW (2005) Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc. Natl. Acad. Sci. U.S.A* 102:15065-8.

21. Doan-Nguyen V, Loria JP (2007) The effects of cosolutes on protein dynamics: the reversal of denaturant-induced protein fluctuations by trimethylamine N-oxide. *Protein Sci* 16:20-29.

22. Zhang HJ, Sheng XR, Pan XM, Zhou JM (1997) Activation of adenylate kinase by denaturants is due to the increasing conformational flexibility at its active sites. *Biochem. Biophys. Res. Commun* 238:382-386.

23. Mashino T, Fridovich I (1987) Effects of urea and trimethylamine-N-oxide on enzyme activity and stability. *Arch. Biochem. Biophys* 258:356-360.

24. Zhang HJ, Sheng XR, Pan XM, Zhou JM (1997) Activation of adenylate kinase by denaturants is due to the increasing conformational flexibility at its active sites. *Biochem. Biophys. Res. Commun* 238:382-6.

25. Müller CW, Schlauderer GJ, Reinstein J, Schulz GE (1996) Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate

binding. *Structure* 4:147-56.

26. Müller CW, Schulz GE (1992) Structure of the complex between adenylate kinase from Escherichia coli and the inhibitor Ap5A refined at 1.9 A resolution. A model for a catalytic transition state. *J Mol Biol* 224:159-77.

27. Krishna MMG, Hoang L, Lin Y, Englander SW (2004) Hydrogen exchange methods to study protein folding. *Methods* 34:51-64.

28. Burlacu-Miron S, Gilles AM, Popescu A, Bârzu O, Craescu CT (1999) Multinuclear magnetic resonance studies of Escherichia coli adenylate kinase in free and bound forms. Resonance assignment, secondary structure and ligand binding. *Eur. J. Biochem* 264:765-74.

29. Sternberg MJE, Grace DEP, Phillips DC (1979) Dynamic information from protein crystallography , : An analysis of temperature factors from refinement of the hen egg-white lysozyme structure. *Journal of Molecular Biology* 130:231-253.

30. Palmer III AG (2001) NMR PROBES OF MOLECULAR DYNAMICS: Overview and Comparison with Other Techniques. *Annu. Rev. Biophys. Biomol. Struct.* 30:129-155.

31. Shapiro YE, Sinev MA, Sineva EV, Tugarinov V, 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.

32. Hansen DF, Vallurupalli P, Kay LE (2008) An improved 15N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J Phys Chem B* 112:5898-5904.

33. Loria JP, Rance M, Palmer AG (1999) A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J. Biomol. NMR* 15:151-155.

34. Hanson JA et al. (2007) Illuminating the mechanistic roles of enzyme conformational dynamics. *Proc. Natl. Acad. Sci. U.S.A* 104:18055-18060.

35. Hilser VJ, Freire E (1997) Predicting the equilibrium protein folding pathway: structure-based analysis of staphylococcal nuclease. *Proteins* 27:171-83.

36. Hilser VJ, Freire E (1996) Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. *J. Mol. Biol* 262:756-772.

37. Pan H, Lee JC, Hilser VJ (2000) Binding sites in Escherichia coli dihydrofolate reductase communicate by modulating the conformational ensemble. *Proc. Natl. Acad. Sci. U.S.A* 97:12020-12025.

38. Wooll JO, Wrabl JO, Hilser VJ (2000) Ensemble modulation as an origin of denaturant-independent hydrogen exchange in proteins. *J. Mol. Biol* 301:247-256.

39. Liu T, Whitten ST, Hilser VJ (2006) Ensemble-based signatures of energy propagation in proteins: a new view of an old phenomenon. *Proteins* 62:728-738.

40. Whitten ST, Wooll JO, Razeghifard R, García-Moreno E B, 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.

41. Auton M, Holthauzen LMF, Bolen DW (2007) Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proc. Natl. Acad. Sci. U.S.A* 104:15317-22.

42. D'Aquino JA et al. (1996) The magnitude of the backbone conformational entropy change in protein folding. *Proteins* 25:143-156.

43. Reinstein J, Brune M, Wittinghofer A (1988) Mutations in the nucleotide binding loop of adenylate kinase of Escherichia coli. *Biochemistry* 27:4712-4720.

44. Sinev MA, Sineva EV, Ittah V, Haas E (1996) Domain closure in adenylate kinase. *Biochemistry* 35:6425-6437.

45. Starikov EB, Nordén B (2007) Enthalpy-entropy compensation: a phantom or something useful? *J Phys Chem B* 111:14431-14435.

46. Sharp K (2001) Entropy-enthalpy compensation: fact or artifact? *Protein Sci* 10:661-667.

47. Lafont V et al. (2007) Compensating enthalpic and entropic changes hinder binding affinity optimization. *Chem Biol Drug Des* 69:413-422.

48. Yamniuk AP, Ishida H, Lippert D, Vogel HJ (2009) Thermodynamic effects of noncoded and coded methionine substitutions in calmodulin. *Biophys. J* 96:1495-1507.

49. Edwards AA et al. (2009) Altered Enthalpy−Entropy Compensation in Picomolar Transition State Analogues of Human Purine Nucleoside Phosphorylase. *Biochemistry* 48:5226-5238.

50. Frederick KK, Marlow MS, Valentine KG, Wand AJ (2007) Conformational entropy in molecular recognition by proteins. *Nature* 448:325-9.

51. Eftink MR, Anusiem AC, 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* 22:3884-96.

52. D'Aquino JA et al. (1996) The magnitude of the backbone conformational entropy change in protein folding. *Proteins* 25:143-56.

53. Cliff MJ, Williams MA, Brooke-Smith J, Barford D, Ladbury JE (2005) Molecular recognition via coupled folding and binding in a TPR domain. *J. Mol. Biol* 346:717-732.

54. Ferreon JC, Hamburger JB, 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* 126:12774-5.

55. Bradley CM, Barrick D (2002) Limits of cooperativity in a structurally modular protein: response of the Notch ankyrin domain to analogous alanine substitutions in each repeat. *J. Mol. Biol* 324:373-386.

56. Bae E, Phillips GN (2004) Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. *J. Biol. Chem* 279:28202-28208.

57. Braun D, Wider G, Wuethrich K (1994) Sequence-Corrected 15N "Random Coil" Chemical Shifts. *Journal of the American Chemical Society* 116:8466-8469.

58. Schwarzinger S et al. (2001) Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc* 123:2970-2978.

59. Vallurupalli P, Kay LE (2006) Complementarity of ensemble and single-molecule measures of protein motion: A relaxation dispersion NMR study of an enzyme complex. *Proceedings of the National Academy of Sciences* 103:11910-11915.

60. Tan Y, Hanson JA, Yang H (2009) Direct Mg(2+) binding activates adenylate kinase from Escherichia coli. *J. Biol. Chem* 284:3306-3313.

61. Fields PA, Somero GN (1998) Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. *Proc Natl Acad Sci U S A* 95:11476-81.

62. Miyashita O, Onuchic JN, Wolynes PG (2003) Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci. U.S.A* 100:12570-5.

63. Whitford PC, Miyashita O, Levy Y, Onuchic JN (2007) Conformational transitions

of adenylate kinase: switching by cracking. *J. Mol. Biol* 366:1661-71.

64. Dixon M, Webb E (1964) *Enzymes* (Academinc Press Inc. Publishers, New York, New York). 2nd Ed.

65. Hilser VJ, García-Moreno E B, Oas TG, Kapp G, Whitten ST (2006) A statistical thermodynamic model of the protein ensemble. *Chem. Rev* 106:1545-1558.

66. Hilser VJ, Freire E (1997) Predicting the equilibrium protein folding pathway: structure-based analysis of staphylococcal nuclease. *Proteins* 27:171-83.

67. Bae E, Phillips GN (2006) Roles of static and dynamic domains in stability and catalysis of adenylate kinase. *Proc Natl Acad Sci U S A* 103:2132-7.

68. Arora K, Brooks CL (2007) Large-scale allosteric conformational transitions of adenylate kinase appear to involve a population-shift mechanism. *Proc. Natl. Acad. Sci. U.S.A* 104:18496-18501.

69. Hilser VJ, Dowdy D, Oas TG, Freire E (1998) The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble. *Proc. Natl. Acad. Sci. U.S.A* 95:9903-9908.

70. Prabhu NV, Sharp KA (2005) Heat capacity in proteins. *Annu Rev Phys Chem* 56:521-548.

71. Ladbury JE, Klebe G, Freire E (2010) Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat Rev Drug Discov* 9:23-27.

72. Frederick KK, Marlow MS, Valentine KG, Wand AJ (2007) Conformational entropy in molecular recognition by proteins. *Nature* 448:325-329.

73. Hilser VJ, Thompson EB (2007) Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc. Natl. Acad. Sci. U.S.A* 104:8311-8315.

## Vita

Travis Parke Schrank was born to Elaine Marie Schrnak and Kenton Parke Schrank on the 29th of December, 1980 in San Antonio, TX. Travis graduated from Texas State University with a B.A. in music. In 2004 he matriculated to the M.D. combined degree program at the University of Texas Medical Branch in Galveston, TX. His graduate work was supported by the NIH through the Houston Area Molecular Biophysics Training Program. His work was acknowledged at national meetings of both the Protein Society and the Biophysical Society. Critical portions of the herein described work were published in the Proceedings of the National Academy of Sciences. Upon completing the combined degree program, Travis will be joining the Medical University of South Carolina as a resident in the department of Otolaryngology – Head and Neck Surgery.

Permanent address: 1112 Sandpiper Denton, TX 76205

This dissertation was typed by: Travis Parke Schrank