11-2002

Protein Folding Studies of Adipocyte Lipid Binding Protein and Homologous Mesophilic and Thermophilic Type I DNA Polymerases

Allyn J. Schoeffler

Protein Folding Studies of Adipocyte Lipid Binding Protein and Homologous Mesophilic and Thermophilic Type I DNA Polymerases

by

Allyn J. Schoeffler

Undergraduate honors thesis under the direction of

Dr. Vince LiCata

Department of Biological Sciences

Submitted to the LSU Honors College in partial fulfillment of the Upper Division Honors Program

November, 2002

Louisiana State University

TABLE OF CONTENTS

ABSTRACT

This study explores the folding of two sets of proteins. Adipocyte lipid binding protein (ALBP), a small intracellular lipid binding protein found in adipocytes, is a member of a large family of structurally similar lipid binding proteins with divergent amino acid sequences and varied surface charge topologies. ALBP has a unique polar surface charge topology. We have constructed a charge mutant of ALBP (referred to as CM-ALBP) in which some of the positively charged residues near the ligand entry portal are neutralized to alanines. We have performed urea-induced chemical denaturations of WT- and CM-ALBP as a function of salt and found that increasing potassium chloride concentrations stabilize both proteins by ~2.5 kcal/mol, but the salt stabilization effect takes effect at different [KCl] for each protein. WT-ALBP is stabilized by 250 mM KCl, while CM-ALBP does not achieve the same stabilization until it is the presence of 1 M KCl. At low (50 mM) and high (1 M) concentrations of KCl, the two proteins have similar stabilities. This suggests that the patch of positive charges that have been neutralized in CM-ALBP are not intrinsically destabilizing, but the binding of anions to this patch may have a stabilizing effect.

We have also conducted studies on the homologous Type I DNA polymerases from *Thermus aquaticus*, a thermophile, and *Escherichia coli*, a mesophile. Taq polymerase and its large fragment (Klentaq), both from *T. aquaticus*, are structurally and functionally similar to the *E. coli* polymerases Pol 1 and its large fragment (Klenow), but the polymerases from *T. aquaticus* are significantly more resistant to thermal denaturation. Guanidine hydrochloride-induced chemical denaturations of the four polymerases reveal that the structural domains in Pol 1 interact in the folded state, while the structural domains in Taq polymerase do not. We have also shown that Klentaq has a free energy of unfolding that is nearly 6 times that of Klenow's. It is believed that denaturation m-values correspond to the amount of surface area exposed upon unfolding. The m-value associated with the denaturation of Klentaq is double that of Klenow, but hydrodynamic experiments in this lab reveal that there is no significant difference between the surface areas of the two proteins in their native or denatured states. Thus, it is not likely that the different m-values correspond to different exposures of buried surface area in the denaturations of these two proteins.

INTRODUCTION

Introduction to Protein Folding

Understanding protein folding, that is, how a string of covalently linked amino acids associates with itself and other small molecules (water, protons, divalent metals) to form an exquisitely functional molecular machine, is one of the Holy Grails of biochemistry. The pursuit has implications not only in our understanding of biophysical processes, but also in the fields of intelligent structure prediction and novel protein design.

Early experiments by Christian Anfinsen showed that for many proteins (those that assume their native states without the aid of chaperone proteins), all of the information necessary to form the unique functional tertiary structure of the folded state is contained in the amino acid sequence (1). In addition, the process of folding is reversible, that is, it is under equilibrium control (1). Therefore, thermodynamic parameters (enthalpy (ΔH), entropy (ΔS) and free energy (ΔG)) can be used to describe the process of protein folding. The change in free energy upon unfolding, or ΔG_U, is a global descriptor of protein stability. The noncovalent forces that contribute to ΔG_U change from protein to protein and (presumably) from mesophilic homologue to thermophilic homologue.

One mystery of protein folding was first described by Cyrus Levinthal, who performed a now-famous calculation known as Levinthal's paradox. Assuming a peptide chain of 100 residues, each of which can adopt one of three conformations in a folded state (α-helix, β-sheet, random coil), how long would it take the peptide to assume a native conformation? The total number of possible structures in this example is 100^3. If

a protein spends 10^{-13} seconds testing each possible structure, it would take 1.6×10^{27} years for it to test them all. This situation does not come close to approximating the actual complexity of folded protein states, in which each residue may adopt far more than three possible orientations in relation to other residues and solvent. Clearly, protein folding is not a random search. The more probable method by which proteins find their native states is by progressing through intermediates with vestiges of native structure (1).

Ken Dill has conceptualized the folding process as a "folding funnel," an energy landscape through which the protein progresses on its way to the most energetically stable native state (See Figure 1). The denatured state is not a unique conformation but an ensemble of many conformations. A denatured conformation may take any number of convergent paths to the unique folded conformation, a phenomenon that is represented by the narrowing of the funnel. It is possible for proteins to be trapped in semi-stable misfolded states on the way to the tip of this folding funnel, a phenomenon that can create disease states.

Protein folding is typically studied by observing protein *un*folding, or denaturation. Denaturation can be achieved *via* thermal denaturation, chemical denaturation, or acid denaturation, and it may be monitored by a number of spectroscopic or calorimetric techniques.

In this thesis, I present two experimental studies of protein folding. First, I have studied adipocyte lipid binding protein (ALBP), a small protein with a polar surface charge topology. My studies have focused on the role the large patches of like-charged

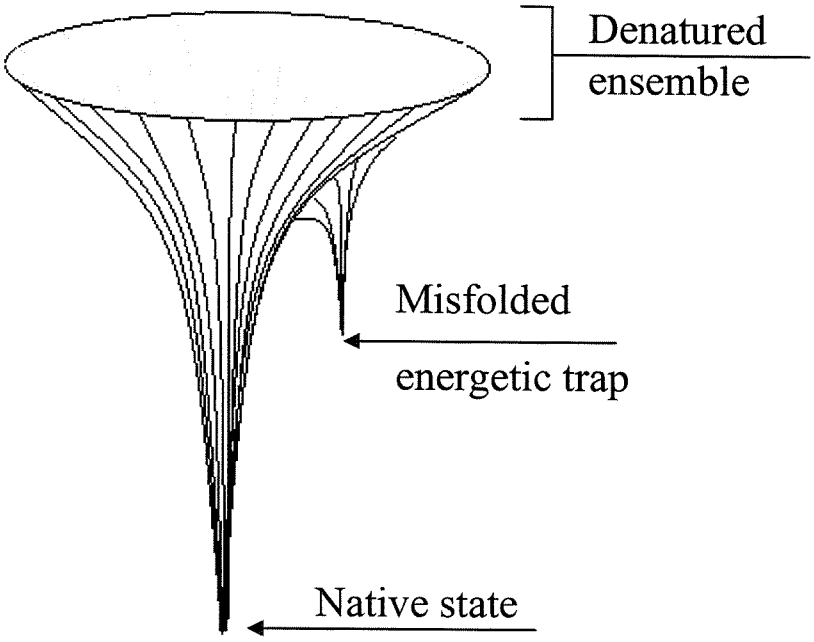


Figure 1: Idealized protein folding funnel, adapted from Ken Dill's models.

residues on the surface of ALBP play in the global stability of the protein. In addition, I have studied homologous pairs of DNA polymerases from *Escherichia coli* and *Thermus aquaticus*. The polymerases from *E. coli* are mesophilic, while those from *T. aquaticus* are thermophilic. My research has tried to determine the thermodynamic basis for the thermostability of the polymerases from *T. aquaticus*.

Introduction to Adipocyte Lipid Binding Protein

Adipocyte lipid binding protein (ALBP) is a member of the intracellular lipid binding protein (iLBP) family, also known as the intracellular fatty acid binding proteins (iFABPs). As its name suggests, ALBP is found predominantly in adipocytes and binds fatty acids and other hydrophobic ligands. The ability to bind lipids is common to the iLBP family, as is a highly conserved tertiary structure. Despite their similarities, the iLBPs have divergent amino acid sequences, unique surface charge potentials, and possibly divergent roles *in vivo*. ALBP has a polar surface charge character, which suggests that ionic conditions may play a role in its function. This study undertakes to determine the role of electrostatics in the stability of ALBP by determining the dependence of its stability on salt. In addition, I have constructed a mutant form of ALBP with part of its charged surface neutralized. I have conducted concomitant studies on the role of salt in the stability of the mutant

Members of the iLBP family have been discovered in a number of mammalian tissues, including liver, adipose, heart, brain, intestinal, and keratinocyte (2). The proteins are also found in amphibians, birds, fish and a variety of invertebrates (2). Most iLBPs are named for the tissue in which they were first discovered, but this should not mislead the reader to believe iLBPs are produced only in their namesake tissue. Liver fatty acid binding protein (LFABP), for example, is found not only in liver cells but also in intestine, kidney and colon cells (3). ALBP is the predominant iLBP of adipocytes, where it constitutes 1-5% of the total soluble protein, but it also exists in macrophages (4). In addition, adipocytes produce another member of the iLBP family, keratinocyte

FABP Tertiary Structure

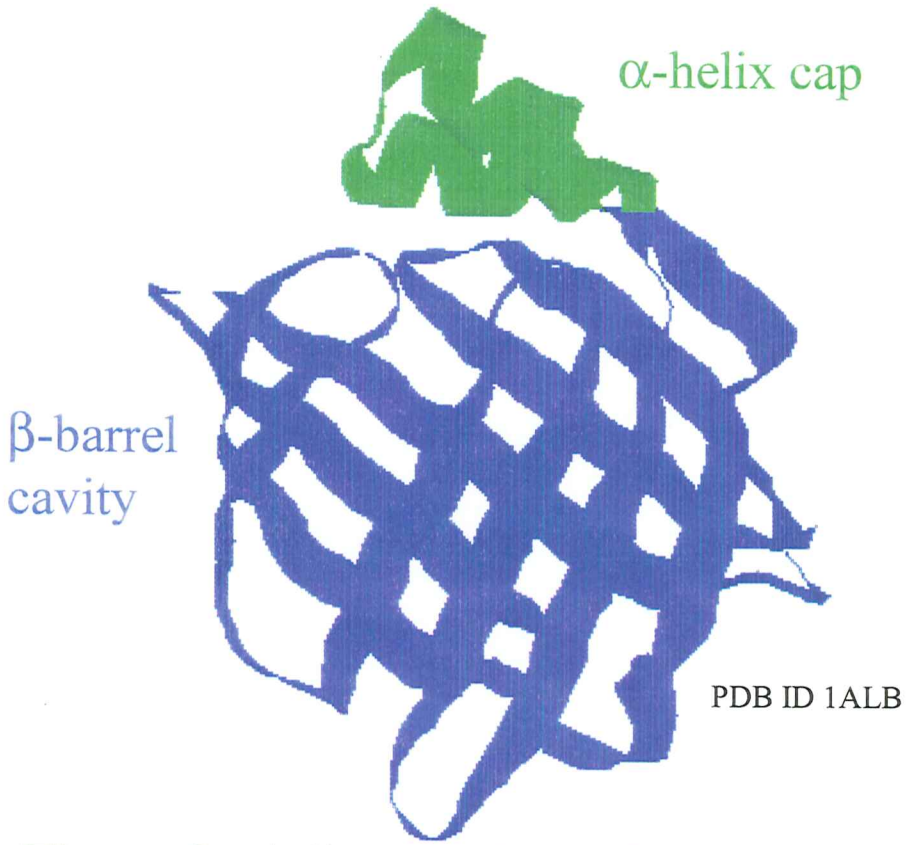


Figure 2: Adipocyte Lipid Binding Protein., shown here, exhibits the typical FABP fold, with a helix-turn-helix cap and a 10-stranded β-barrel cavity (6).

lipid binding protein (KLBP), albeit in very small amounts (1% of the levels of ALBP) (4).

Members of the iLBP family are characterized in part by their remarkably similar structures. The proteins are small, most weighing in near 15 kDa. The iLBP fold consists of a 10-strand antiparallel β-barrel and a helix-turn-helix cap (see Figure 2). Between family members, the RMS deviation between atoms ranges from 0.63 to 2.38 $Å^3$, two to seven times the deviation typically observed between different crystal structures of the same protein (5). They all possess a ligand binding cavity of approximately 1000 $Å^3$ located in the top half of the β-barrel (5). A ligand entry portal is formed by the helices and the loops connecting βC to βD and βE to βF (see Figure 3) (5,6). Interestingly, it is the βE-βF loop that exhibits the most structural divergence among iLBP family members (5). Also of interest are the βD and βE strands, which are the only strands not stabilized by direct hydrogen bonding (See Figure 3) (5). It is possible that this structural feature provides some flexibility by which the protein may expand to admit a lipid, since the portal is not large enough to admit a lipid without some sort of conformational change (5).

In ALBP, some of the features of this portal are worth noting. First, Phe^{57}, which is in the portal region, exhibits the greatest RMS deviation in position upon ligand binding (6). The hydrophobic side chain of Phe^{57} partially covers the portal of apo-ALBP (6). Putatively, the side chain moves to admit a ligand and does not move back. In addition, the portal region of ALBP, along with its entire "top" portion, is highly positively charged. Another unique feature of ALBP is its ability to be phosphorylated. Tyr^{19}, a residue situated near the center of the portal, can be phosphorylated by insulin

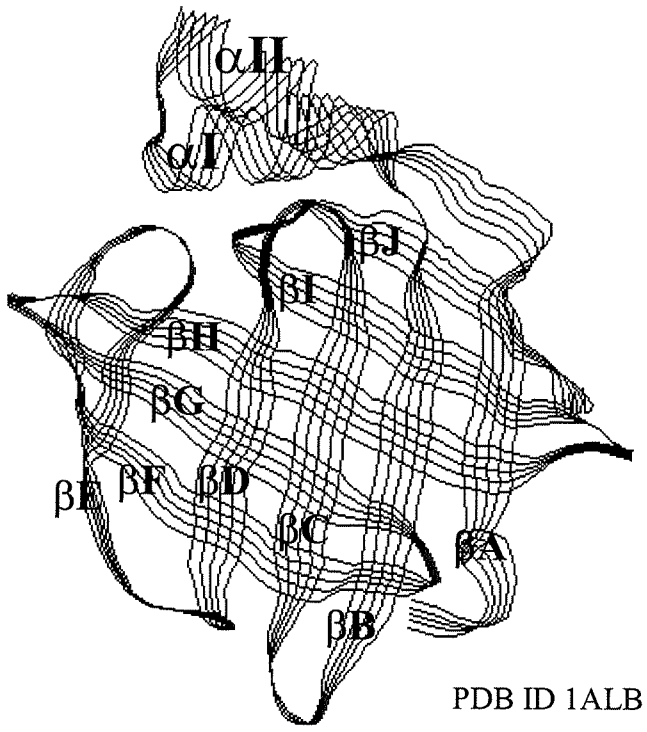


Figure 3: Ribbon diagram of ALBP with the major structural features labeled. The ligand entry portal is formed by the loop between βC and βD , the loop between βE and βF and helices I and II. Note also the larger gap between strands βD and βE (6).

receptor kinase, a tyrosyl kinase (7). The process is stimulated *in situ* by the addition of insulin (7). Phosphorylated apo-ALBP cannot bind ligand; phosphorylated holo-ALBP cannot release ligand (8). There is a steric explanation for this: The bulky phosphorylated residue would occlude a portion of the portal occupied by the hydrocarbon tail of a fatty acid in holo-ALBP (6). In addition, phosphorylation would place a large negative charge in the normally positively-charged portal region through with a negatively-charged ligand moiety would normally pass (6).

The similar tertiary structures of the iLBPs are in contrast to their highly variant amino acid sequences. Homology between family members ranges from 23 to 69%, with 39 highly conserved residues (5). Twenty-six of these conserved residues occupy a stabilizing structural backbone on the lower edge of the protein; the remainder are concentrated on αI (the posterior helix, see Figure 3) or speckled throughout the structure (5).

In addition to their divergent primary structures, iLBPs have very different surface charge potentials. Surface charge potential calculations reveal that iLBPs of different mammalian tissues have unique charge topologies despite their nearly identical structures (see Figure 4) (9). ALBP, for example, has a polar surface charge potential—positively charged on top, negatively charged on bottom (9). The iLBP found in p2 myelin has an all-over negative charge, and liver fatty acid binding protein (LFABP) has a more hydrophobic surface character (9). *A priori*, the varied charge topologies of the iLBPs suggest that surface charge may play an important role in their different functions *in vivo*.

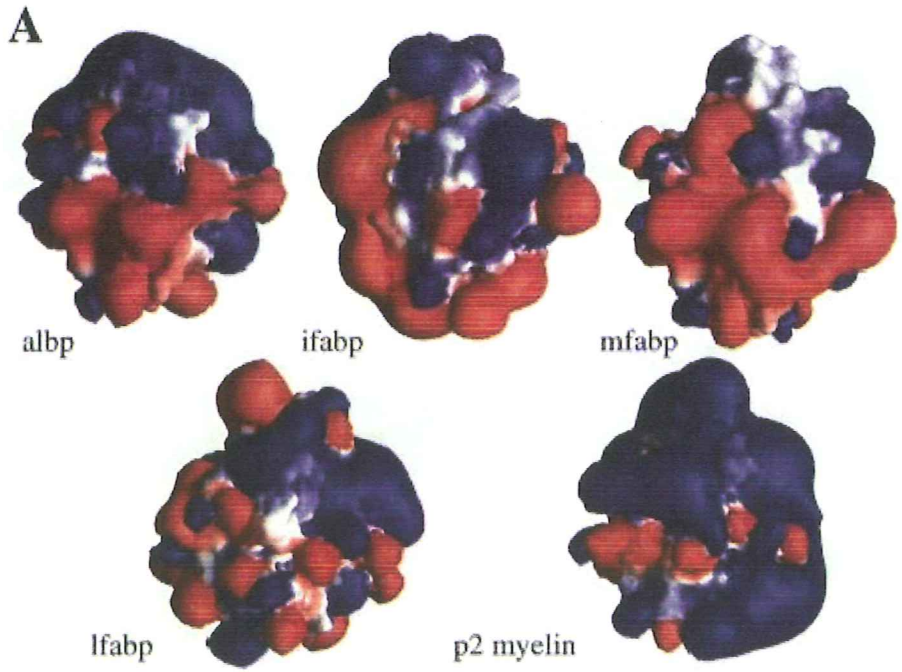


Figure 4: Electrostatic surface potentials of selected mammalian iLBPs. They are, from right to left, top to bottom: albp: adipocyte lipid binding protein, ifabp: intestinal fatty acid binding protein, mfabp: muscle fatty acid binding protein, lfabp: liver fatty acid binding protein, and p2 myelin lipid binding protein. Blue areas represent positive potential, red areas represent negative potential and white areas represent neutral or hydrophobic patches. Note the unique polar surface charge of ALBP (9).

Though their specific metabolic roles remain undetermined, the iLBPs share, as their family name suggests, the ability to bind lipids. Specifically, ALBP can bind a number of fatty acids of at least 14 carbons (5,10,6,11) as well as other hydrophobic ligands (12) in its large, water-filled cavitiy. In order to bind, a ligand must have both a hydrophobic moiety and a charged carboxylate or sulfonate head group. The head group coordinates with three amino acids in the center of the cavity: Arg^{126}, Tyr^{128}, and Arg^{106} (*via* a water molecule) (6,10). Among the ligands shown to bind to ALBP are palmitate (a 16-carbon carboxylic acid) and its sulfonic acid homologue, the 16-carbon hexadecanesulfonic acid (10). ALBP has been shown to effectively bind stearic acid (6), oleic acid (6,13) and arachidonic acid (see Figure 5) (11,13). The protein also binds retinoic acid, albeit weakly (13). Short-chain fatty acids, such as octanoic acid, and ligands without charge groups, such as retinol, cannot bind to ALBP (13). In addition to these physiologically relevant molecules, ALBP has been shown to bind the fluorescent probe 1-anilinonaphthalene-8-sulfonate (ANS) (see Figure 6) (12). ANS is a competitive inhibitor of fatty acid binding and produces a large increase in fluorescence upon binding (12,13). This allows binding constants of ANS to ALBP to be followed fluorometrically (12).

The distribution of different iLBPs in different tissues, along with their varied surface charge character, suggests that the proteins may play quite different metabolic roles. The metabolic role of ALBP is beginning to be understood by observing its interactions with other proteins involved in fat metabolism. ALBP has been demonstrated to interact directly with the N-terminus of hormone-sensitive lipase (HSL) (14). HSL is the enzyme responsible for freeing fatty acids from triacylglycerol in

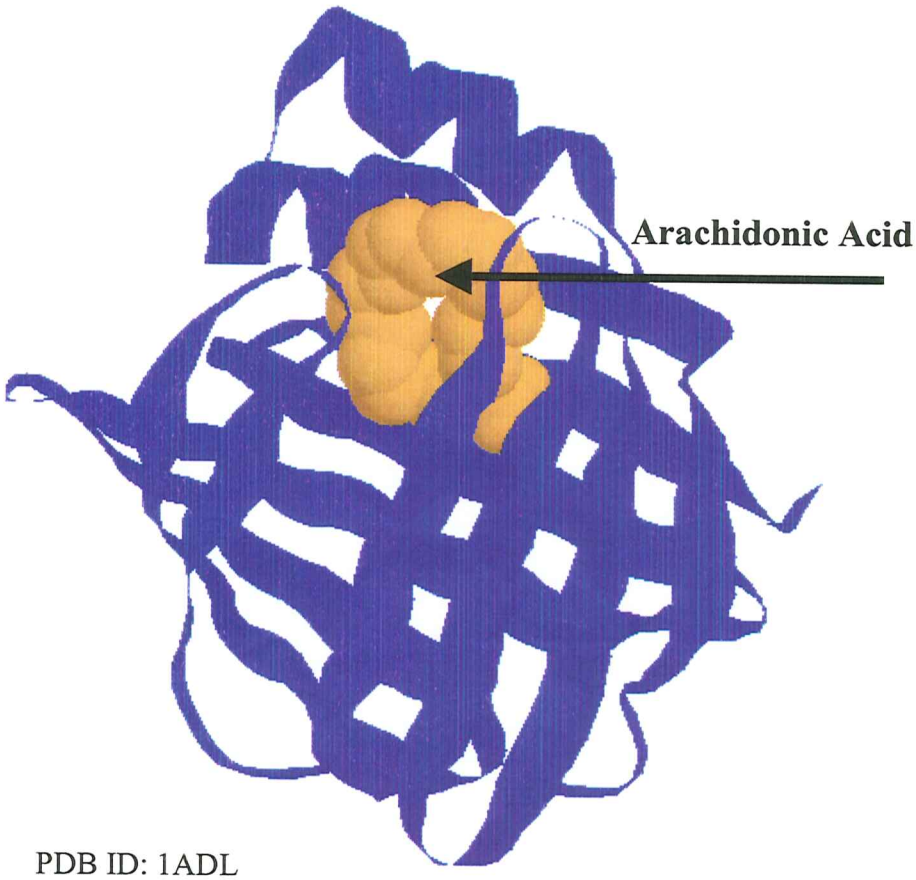


Figure 5: ALBP with arachidonic acid bound (11)

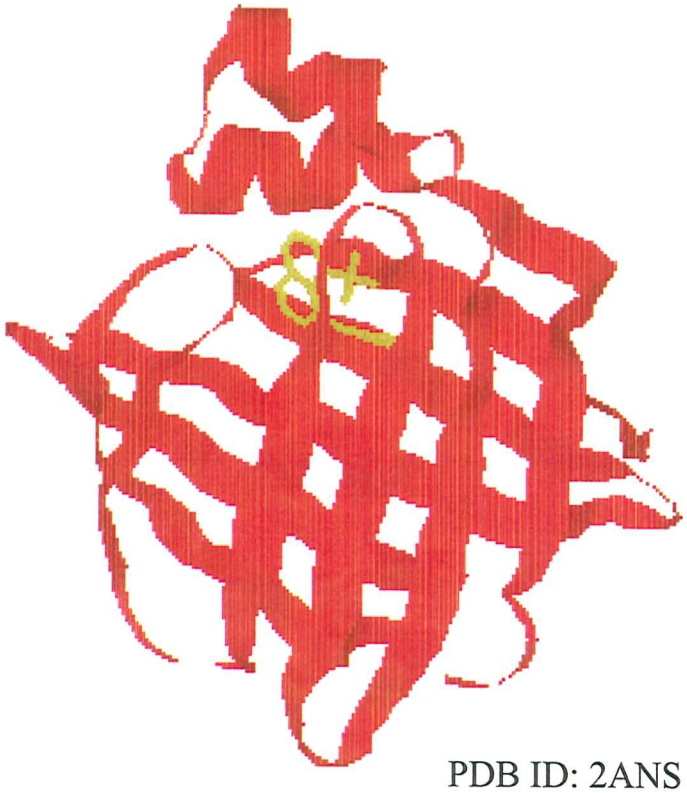


Figure 6: ALBP with ANS (shown in yellow) bound (12).

response to lypolytic hormonal signals (15). It has been postulated that ALBP binds to a "docking region" of HSL, rapidly binding freed fatty acids in order to 1) transport them to another region of the adipocyte and 2) remove them from the presence of HSL to prevent substrate inhibition of lipolysis (14).

ALBP has additional physiological importance in its role in the development of type II diabetes. In a well-known study of ALBP-knockout mice, Hotamisligil and associates found that obese ALBP-knockout mice did not develop insulin resistance or diabetes, while their wildtype companions did (16). Insulin levels in wildtype obese mice were highly elevated but produced no concomitant mediation of glucose levels (16). In contrast, insulin levels in the obese ALBP-knockout mice were indistinguishable from lean wildtype or lean knockout mice, and the glucose levels of these obese mice were depressed (16). The precise role of ALBP in the development of diabetes and the feasibility of targeting diabetes drugs to the protein remain undetermined (16).

The supposition that surface charge plays a role in iLBP function, specifically in ALBP function, has yet to be extensively tested, but Storch and associates have conducted a series of experiments involving charged lipid vesicles that are of interest. Surface lysines on ALBP have been shown to affect the interaction of the protein with charged lipid vesicles (17). Wildtype ALBP shows stronger binding to lipid vesicles composed of anionic phospholipids than to those that are dominated by positive charges or are neutral (17). ALBP also shows a faster fatty acid transfer rate to anionic vesicles over cationic vesicles (17). Neutralizing the surface lysines of ALBP by acetylation causes the protein to exhibit the same behavior toward anionic vesicles as unacetylated protein displays towards cationic vesicles (17).

The polar nature of ALBP's surface charge suggests that it may play an important role in the function of the protein. To aid in our determination of the role of surface charge in the stability of ALBP, we have created a mutant form of ALBP in which some of the positive residues have been neutralized. We chose four positive residues that line the portal region in WT-ALBP—Lys^{31}, Arg^{30}, Lys^{37}, and Lys^{58}—and mutated them to alanines, forming a quadruple charge-mutant henceforth referred to as CM-ALBP (see Figure 14 on page 40). Three of these residues (31, 37 and 58) are among those that are highly conserved in the iLBP family (5). Studies on the stability of CM- and wild-type (WT-) ALBP as a function of salt follow.

Introduction to the Type I DNA Polymerase Family

Type I DNA polymerases, known to remove RNA primers and conduct DNA repair *in vivo*, share a characteristic tertiary structure consisting of three structure/function domains (1, 18). The C-terminal region contains the polymerase and 3' -> 5' exonuclease (proofreading) domains, while the smaller N-terminal region has 5' exonuclease activity (see Figure 7) (19). The hand-shaped large domain grasps DNA during synthesis, as depicted in Figure 7 (20).

This study focuses on two members of the Type I DNA polymerase family, one from the eubacterium *Escherichia coli*, a mesophile that thrives at 37°C, and one from the eubacterium *Thermus aquaticus*, a thermophile. The optimal growth temperature of *T. aquaticus* is between 70 and 72°C (21, 22). The Type I polymerase of *E. coli* is commonly known as Pol 1, and the homologous polymerase from *T. aquaticus* is called Taq polymerase. The independent large fragments of these proteins are called Klenow and Klentaq, respectively (See Figure 8). The pairs of proteins are structurally and functionally homologous (21, 23, 18, 20, 24). The Klenow and Klentaq fragments share almost 50% amino acid sequence homology while the N-terminal (or "tail") domains share less amino acid sequence homology (24).

Despite their functional and structural homology, the polymerases derived from *T. aquaticus* are significantly more resistant to heat than their *E. coli* counterparts. It has been known for some time that Taq polymerase and Klentaq can withstand the high temperatures (95°C) employed in the polymerase chain reaction (PCR) and retain their functionality (23,25). Taq polymerase has been revolutionary to the field of molecular

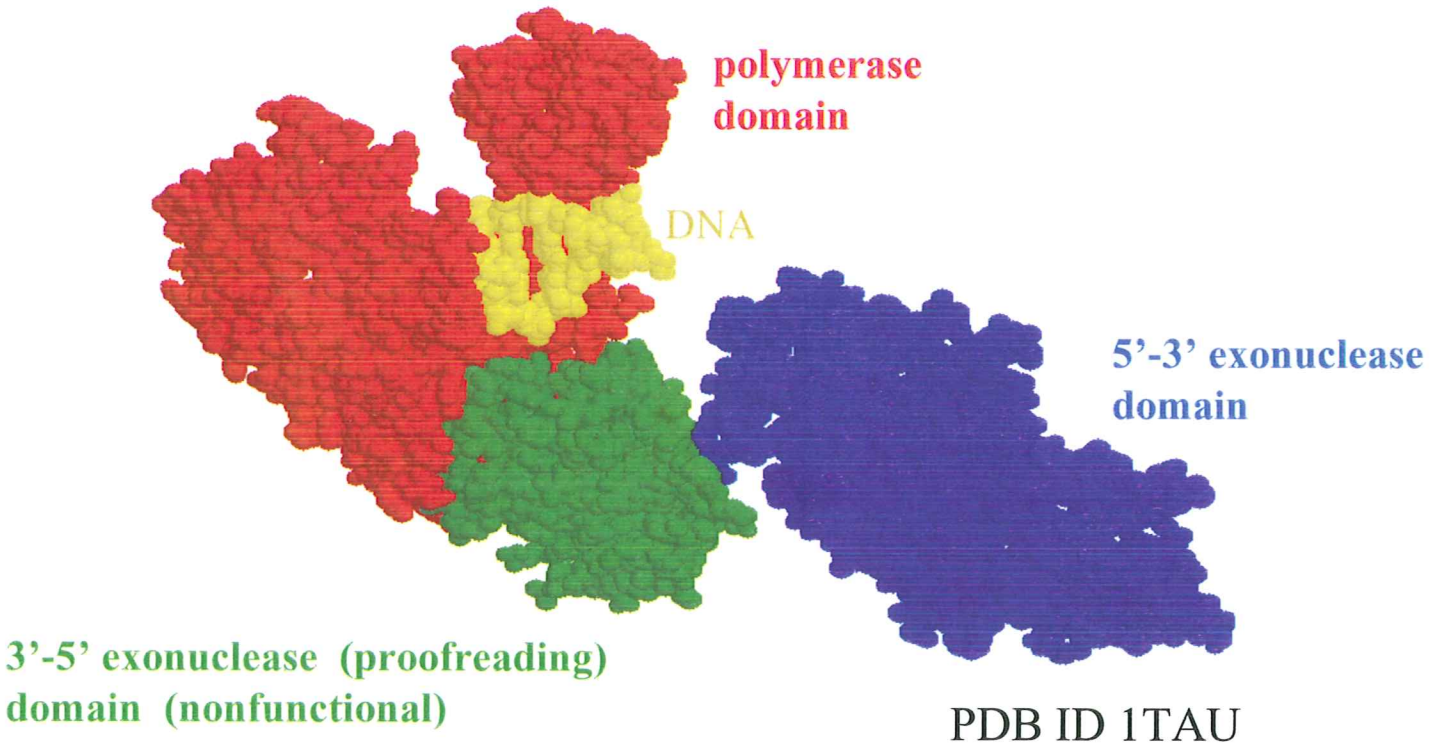


Figure 7: Taq DNA polymerase with DNA bound. The polymerase and proofreading domains (in red and green, respectively) make up the large C-terminal region, while the N-terminal region contains 5' to 3' exonuclease activity. DNA (in yellow) is shown here bound to the polymerase domain (20).

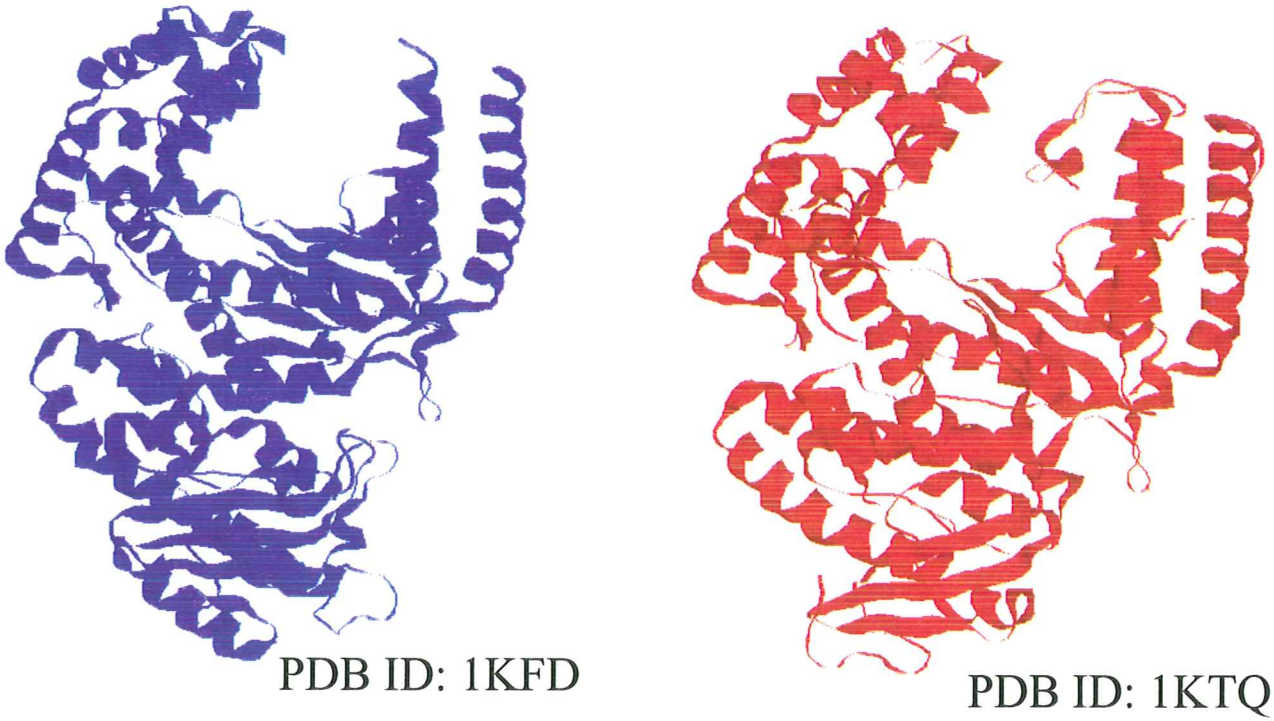


Figure 8: Large fragments of *E. coli* and *T. aquaticus* Type I DNA polymerases. Klenow (left, in blue) is the large fragment of *E. coli* Pol 1, and Klentaq (right, in red) is the large fragment if Taq polymerase. Note the high degree of structural homology between the two polymerases (26, 24).

biology, allowing PCR users to operate the reaction relatively "hands-free" for 30 cycles or more. Most of the "stability studies" done on Taq polymerase were simple activity checks, in which the protein was heated up to various temperatures and assayed for activity in DNA synthesis. Studies such as these were useful in revealing the functional nature of Taq polymerase's stability.

Attempts to determine the structural basis for the high thermostability of Taq polymerase and Klentaq have been made, though none of them have been tested. The Steitz and Waksman laboratories have noted structural features of the enzymes (in comparison to their mesophilic homologues) which they believe account for the high thermostability of the proteins. Steitz and colleagues note an exchange of two ionic interactions in Klentaq for two hydrophobic interactions (18). The Waksman laboratory points out that of the ~450 primary sequence differences between Klentaq and Klenow, there are 79 highly non-conservative substitutions, 19 of which are substitutions of charged amino acids with oppositely charged residues (24).

The proofreading domain of Taq polymerase is non-functional (27). According to the Steitz laboratory, this is due to structural deletions in the remnants of the proofreading domain of Taq polymerase (18). Compared to the proofreading domain of Klenow, that of Taq polymerase is missing four loops of between 8-27 residues, as well as four carboxylate-containing amino acids that normally bind divalent metal ions in Klenow (18). The researchers note that the truncation of the proofreading domain (which falls structurally between the polymerase and exonuclease domains) does not alter the structural relationship between the two larger domains (18).

Thermal Denaturations of Taq polymerase, Pol 1, Klentaq and Klenow

Despite the wealth of information on the functional and putative structural elements of Klentaq's stability, there are no published studies exploring the thermodynamic means by which Taq polymerase is stabilized. Previous studies in the LiCata laboratory conducted by Irene Karantzeni have, for the first time, characterized the thermal unfolding of all four polymerases. Differential scanning calorimetry (DSC) monitors the heat released upon the unfolding of a protein. Figure 9 shows the analyzed DSC curves representing the thermal denaturations of all four polymerases. After non-linear fitting, this information can be used to find the temperature midpoint of denaturation (Tm) and the change in enthalpy upon unfolding (ΔH_U) of a protein. These data are organized in Table 1. Karantzeni's work shows that there is a 63° increase in the Tm of Klentaq versus Klenow. Denaturations of the full-length polymerases reveal that while the interactions between the N- and C-terminal regions of Taq polymerase and Pol 1 are *structurally* identical, they are not *thermodynamically* identical. First, the two structural domains of Taq polymerase seem to denature independently of one another. In contrast, the two structural domains of Pol 1 appear to denature simultaneously, though the single excess heat capacity peak given during denaturation can be mathematically deconvoluted into two separate transitions. This allows the differences in the domain relationships to be analyzed. Klenow alone denatures at 37°C. The inclusion of the exonuclease domain in full length Pol 1 produces a 3.6°C Tm stabilization of the Klenow region in Pol 1. The opposite effect is observable for Klentaq. Alone, Klentaq denatures at 100°C. In full length Taq polymerase, the Tm of the large C-terminal region is destabilized by 1°C.

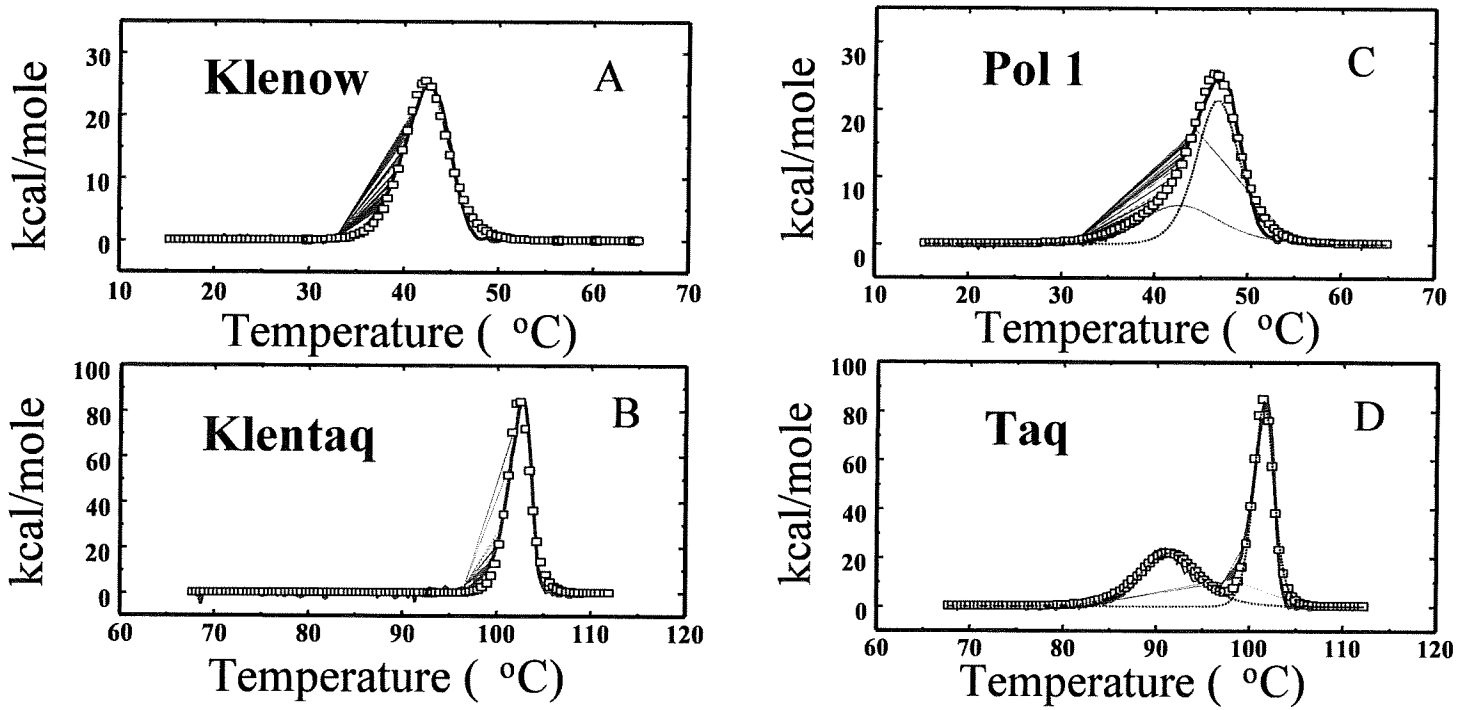


Figure 9: Excess heat capacity curves for the thermal denaturations of Klenow (A), Klentaq (B), Pol 1 (C) and Taq polymerase (D). Solid lines represent experimental data. Squares and dotted lines show the best fit of the given model to the data. A, B and D show fits of non 2-state models; Pol 1 (C) was fit to the sum of two 2-state transitions. In the case of Pol 1 and Taq, the squares show the sum of the fitted deconvolutions. Fitted ΔH_U and Tm values are given in Table 1.

	Tm (°C)		ΔH^{app}_{cal} (kcal/mol)		ΔH^{app}_{vH} (kcal/mol)	
	Peak 1	Peak 2	Peak 1	Peak 2	Peak 1	Peak 2
Taq	88.9 ± 0.9	99.1 ± 0.1	147 ± 5	250 ± 4	125 ± 8	334 ± 10
Klentaq	---	100.2 ± 0.1	---	244 ± 17	---	311 ± 8
Pol 1	36.7 ± 0.9	40.4 ± 0.9	56 ± 9	111 ± 2	NA	NA
Klenow	---	36.8 ± 0.5	---	131 ± 2	---	127 ± 2

Table 1: Fitted values for temperature midpoint of unfolding (Tm) and apparent calorimetric enthalpy of unfolding (ΔH^{app}_{cal}) for all four polymerases. (Karantzeni & LiCata, unpublished)

Given a pair of homologous proteins, one of which has a much higher Tm, there are three *a priori* thermodynamic possibilities for the stabilization of the more thermostable protein. They are (see Figure 10) (28, 29, 30):

1) The total thermodynamic energy (ΔG) stabilizing the thermophilic protein is *increased* relative to its mesophilic homologue (green line).
2) The stable temperature range for the thermophilic protein is *broadened* relative to its mesophilic homologue (red line).
3) The stable temperature range for the thermophilic protein is *shifted* relative to its mesophilic homologue (blue line).

These curves are described by the Gibbs-Helmholtz equation, given below (28, 29, 30):

$$\Delta G_U (T) = \Delta H_m [(Tm-T)/Tm] - \Delta Cp \{Tm-T[1-\ln(Tm/T)]\},$$

where ΔG_U is the free energy of unfolding at a given temperature (the dependent variable), T is any temperature (the independent variable), Tm is the temperature midpoint of unfolding, ΔH_m is the enthalpy of unfolding at the Tm, and ΔCp is the change in heat capacity upon unfolding.

In this study, I have conducted chemical denaturations of all four DNA polymerases and compared them to the thermal denaturations already performed. This makes it possible to make suggestions regarding the thermodynamic strategy by which Klentaq is stabilized.

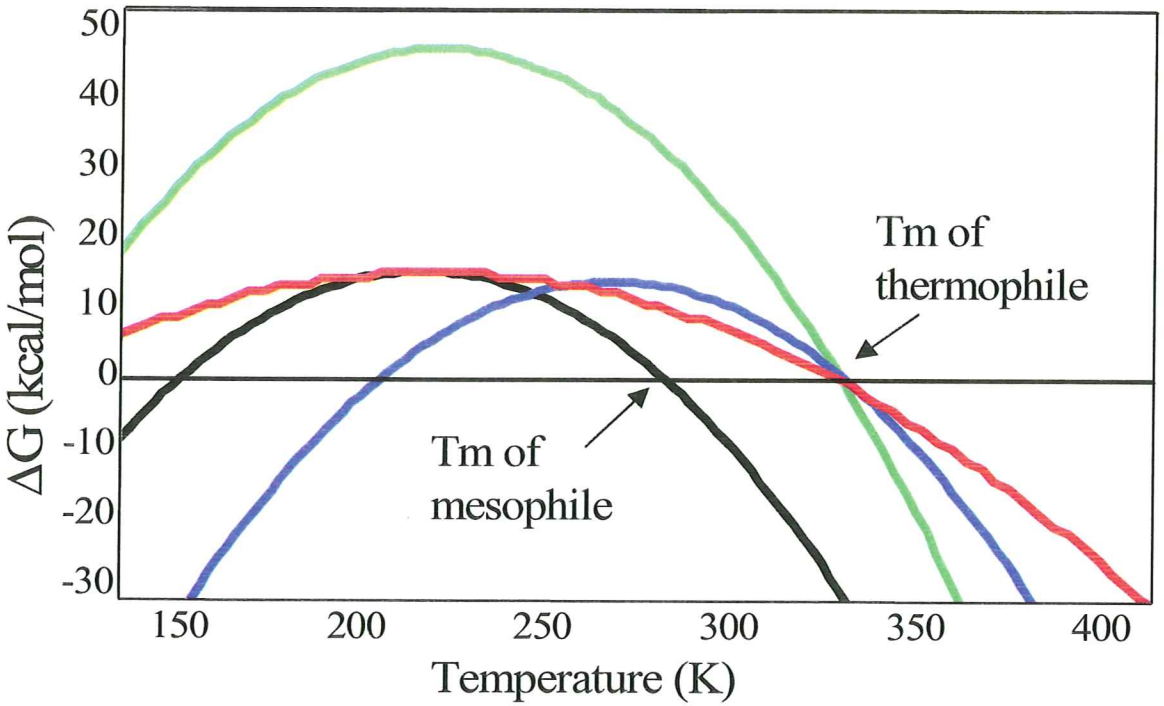


Figure 10: Theoretical Gibbs-Helmholtz plots describing thermodynamic models for thermostability. The black line describes a mesophilic protein. Relative to the mesophilic protein, the *red* line represents a thermophilic protein with a *broadened* stable temperature range. The *blue* line represents a thermophilic protein with a *shifted* stable temperature range, and the *green* line represents a thermophilic protein with *increased* stability across the same temperature range.

MATERIALS AND METHODS

Site-Directed Mutagenesis of WT-ALBP

An expression vector containing the cDNA of murine wildtype ALBP was the gift of David Bernlohr of the University of Minnesota, Twin Cities. The pRSET type vector (obtained from Invitrogen) contains the cDNA for ALBP, a gene for ampicillin resistance, an M13 bacteriophage origin of replication, and the *lac* repressor.

Site-directed mutagensis was performed on the pRSET-WTALBP vector to create CM-ALBP. The single-stranded DNA technique (described below) was used to create four mutations in WT-ALBP: R30A, K31A, K37A and K58A (see Figure 14 on page 39). First, pRSET-WTALBP was transformed into *E. coli* strain CJ236. This strain contains mutations that disable uracil-DNA glycosidase and deoxyuridine transferase. These mutations, referred to as the *dut/ung* mutations, have the effect of forcing the replication machinery of the cell to incorporate uracil in place of thymidine during DNA replication. Colonies containing the plasmid were selected by ampicillin resistance and then grown in a larger culture. This culture was infected with M13KO7 bacteriophage. The bacteriophage recognizes the M13 origin of replication on the plasmid and replicates and packages only the coding strand. The phage particles were isolated by centrifugation (bacteria and lysed cell debris form a pellet while the phage particles remain in solution). The single-stranded plasmid was then purified from the phage particles by precipitating the phage particles with polyethylene glycol and salt, then purifying the DNA by phenol-chloroform extraction. Mutagenic oligonucleotides were annealed to the ssDNA plasmid and the remainder of the strand was replicated using *E. coli* DNA polymerase I. The result of this process was a hybrid plasmid, one with a wildtype sequence and

deoxyuridine, the other with a mutant sequence and deoxythymidine. The hybrid plasmid was transformed into *E. coli* strain DH5α, where the normal cellular machinery attempted to remove the deoxyuridines and in the process destroyed the wildtype strand. This left only the mutant coding strand, whose complement was created by the DH5α replication enzymes. The resulting plasmid was purified from *E. coli* cells and subjected to restriction enzyme digestion to confirm the presence of mutations. Plasmids that appeared to have incorporated all mutations were then sequenced for confirmation.

Mutagenic Oligonucleotides

Mutagenic oligonucleotides (see Figure 11) were designed to produce the desired mutations and to contain new restriction enzyme digest sites. For the R30A/K31A mutation, the oligonucleotide CAT GCC TGC CAC GGC CGC TGT GGC AAA GCC was used. This oligonucleotide contains a recognition site for the restriction enzyme *Eag* I. For the K37A mutation, the oligonucelotide GAT CAT GTT GGG CGC GGC CAT GCC TGC, which contains a restriction enzyme recognition site for the restriction enzyme *Bst* UI, was used. For the K58A mutation, the oligonucleotide AAT CTC GGT GTT CGC GAA AGT ACT CTC TGA C was used. It contains a recognition site for the restriction enzyme *Nru* I. (Mutated codons responsible for introducing mutations are underlined.) All oligonucleotides were obtained from Integrated DNA Technologies, Inc

Figure 11: Mutagenic Oligonucleotides

* Restriction enzymes are shown in blue above the mutant strand
* Restriction enzyme recognition sites are boxed
* Mutated nucleotides are shown in red

Purification of WT and CM ALBP

WT- and CM-ALBP were both purified according to the same protocol, adapted from the protocol published by the Bernlohr laboratory (31). The appropriate vectors were heat-transformed into *E. coli* strain BL21 cells. Colonies containing the plasmid were selected by ampicillin resistance. Viable colonies were grown in 5 mL of LB media for a period of 10 to 14 hours. One of these cultures was then used to innoculate five liters of M9 minimal media (1 mL/ liter). Cultures were grown in a shaking incubator at 37°C and 250 RPM until the absorbance of media at 550 nm reached 0.4-0.6 OD (about 6 hours). Cells were then induced with 0.5 mM IPTG, a lactose analog that induced the expression of protein by deactivating the *lac* repressor. Four hours after induction, cells were harvested by centrifugation. Cell pellets (typically 10-20 g total) were resuspended in lysis buffer (25 mM imidazole HCl, 50 mM NaCl, 1 mM β-mercaptoethanol, 5 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride) and sonicated to lyse cells. Sonication was performed using a Branson Sonifier 250 sonicater at a setting of 5 to maximum. Sonication was performed in 60-second bursts with cooling periods (variable length) on ice between bursts.

The lysed cell slurry was then centrifuged to remove cell debris and the supernatant was collected. The supernatant was subjected to a protamine sulfate precipitation (1% final concentration) to remove DNA. The precipitated DNA was removed by centrifugation and the supernatant was reserved. The supernatant was then subjected to a pH-5 precipitation step, in which sodium acetate at pH 5 was added slowly to the protein solution until it reached pH 5. The mixture was allowed to stir overnight at 4°C, after which precipitated proteins were removed by centrifugation.

The supernatant from this step was concentrated to ~15 mL and loaded onto a Sephadex G-75 column with a bed volume of ~2 liters. The column was swollen and equilibrated with column buffer (12.5 mM HEPES, 250 mM NaCl, at pH 7.5 for WT-ALBP and at pH 6.7 for CM-ALBP). Protein was eluted from the column based on size, and elution peaks were detected by their absorbance at 280 nm. The column was run at a flow rate of 1.7 mL/minute. Individual fractions were read at 280 nm in a Cary-UV/Vis spectrometer, and an elution profile was plotted. Fractions from throughout the profile were chosen and run on a 15% polyacrylamide denaturing gel (SDS-PAGE). Based on the gels and the elution profile, the absorbance peak corresponding to ALBP was chosen and those fractions were pooled and concentrated. Concentrated protein was then aliquoted into microfuge tubes and stored at $-70^{\circ}C$.

Purification of Polymerases

All polymerases were purified in-house according to published procedures with modifications as described below.

Taq polymerase Purification

The cDNA encoding Taq polymerase was amplified using the set of primers described by Engleke, *et al.*, and the amplified product was cloned into an expression vector (pTrc99A, obtained from Amersham) using *Eco* R1 and *Bam* H1 restriction enzymes (32). The cDNA was then recut from this construct using *Nco* I and *Sal* I restriction enzymes and ligated into the expression vector pET-15b (Novagen). This plasmid was transformed into *E. coli* strain BL21(DE3) for protein expression. Cells were grown and harvested according to Engleke, *et al.*, then lysed by incubating with lysozyme (32). Many contaminating proteins were removed by heating the crude lysate to 75°C for one hour. Nucleic acids were removed by PEI precipitation in the presence of 250 mM KCl. Three chromatography columns followed. First, the protein was applied to Bio-Rex 70 ion exchange resin column equilibrated with KTA buffer (20 mM Tris, 22 mM $(NH_4)_2SO_4$, 1 mM DTT, 0.1 mM EDTA, 10% glycerol) at pH 7.9. The flow-through from this column was run on a heparin-sepharose column, and protein was eluted with an $(NH_4)_2SO_4$ gradient. Taq polymerase eluted from this column was applied to a second Bio-Rex 70 ion exchange resin column equilibrated with KTA buffer at pH 8.8. All surfactants were omited from the preparation and storage of Taq polymerase.

Klentaq Purification

The Klentaq clone was obtained from ATCC (Manassas, VA, ATCC # 69244). The protein was purified according to the published protocol (33, 34). Alterations to the

published protocol include the omission of surfactants from all preparatory and storage buffers and the addition of a final column, a Bio-Rex ion exchange resin column equilibrated at pH 9.1.

Klenow and Pol I Purification

Clones for both *E. coli* proteins were obtained from Catherine Joyce of Yale University. The proteins were expressed according to the published protocol (35). Modifications include the addition of a final heparin column, from which the protein was eluted using a KCl gradient. The column was run in 10 mM potassium phosphate buffer with 1 mM DTT at pH 7.

Protein Concentration Determination

The concentrations of the polymerases were determined by the Bradford assay, with reagents obtained from Bio-Rad. The concentrations of CM- and WT-ALBP were determined using the BCA assay, obtained from Pierce. An extinction coefficient for CM-ALBP (10923 $M^{-1}cm^{-1}$) was determined in-house and used thereafter. A published extinction coefficient for WT-ALBP (13600 $M^{-1}cm^{-1}$) was verified in-house and used thereafter (31).

Chemical Denaturations

Proteins were dialyzed extensively (3 buffer changes of 1 liter each, dialyzed for at least 4 hours each) against appropriate buffers. For WT- and CM-ALBP, the buffers were 10 mM potassium phosphate at pH 7.5, with varying potassium chloride concentrations from 0 to 1 M. For the polymerases, the buffer was 10 mM Tris, 125 mM KCl, and 5 mM $MgCl_2$ at pH 7.5. Concentrated urea and guanidine hydrochloride (GdnHCl) stocks were prepared by first dissolving denaturants (obtained from Amresco) in deionized water. Urea stocks were then deionized by stirring with AG 501-X8 deionizing resin (Bio-Rad) for one hour. The concentrations of deionized urea and GdnHCl were determined by refractive index as described by Nozaki (36). Urea and GdnHCl stocks were then incorporated into the appropriate buffers. A Bausch and Lomb ABBE-3L refractometer was used for all measurements

Stepwise chemical denaturations were performed by incubating individual aliquots of protein at 0.1 to 0.2 mg/mL with buffered urea or GdnHCl for one hour, well past the time required to reach equilibrium for all proteins. The amount of secondary structure remaining after incubation was determined by circular dichroism spectroscopy. Each sample was scanned in an Aviv 62DS circular dichroism spectrometer. For WT- and CM-ALBP, spectra were recorded from 225 to 213 nm. For the polymerases, spectra were recorded from 225 to 215 nm. A quartz cuvette with a pathlength of 0.2 cm was used for all measurements.

Analysis of the chemical denaturations was performed by first determining the raw CD signal (in millidegrees) of each sample at specific wavelengths characteristic to

each protein. For WT- and CM-ALBP, CD signals at 216, 217 and 218 nm were used. For the polymerases, CD signals at 218, 219, 220 and 221 nm were used.

Raw CD signal was then transformed into molar ellipticity ($\Delta\varepsilon$) by the following equation:

$$\Delta\varepsilon = \frac{(S_{sample} - S_{blank}) * K}{c * l},$$

where S_{sample} is the raw CD signal (in milldegrees) of an individual sample at a given wavelength, S_{blank} is the raw CD signal of buffer at the same wavelength in millidegrees, K is a constant conversion factor (30.3214), c is the concentration of protein in moles/ liter, and l is the pathlength of the cell in cm (37).

Molar ellipiticities of samples at each wavelength were plotted as a function of denaturant conentration, producing sigmoidal denaturation curves. These curves were analyzed using the nonlinear form of the linear extrapolation method (LEM) (38):

$$\Delta\varepsilon = \frac{(\Delta\varepsilon_N + m_N[D]) + (\Delta\varepsilon_U + m_U[D]) \times e^{-(\Delta G^{\circ}_{N->U}/RT + m_G[D]/RT)}}{1 + e^{-(\Delta G^{\circ}_{N->U}/RT + m_G[D]/RT)}}$$

Values for this equation are as follows:

$\Delta\varepsilon$ = molar ellipticity at a given wavelength (the dependent variable)

[D] = denaturant concentration in molar (the independent variable)

$\Delta\varepsilon_N$ = y-intercept of the native state baseline

m_N = slope of the native state baseline

$\Delta\varepsilon_U$ = y-intercept of the unfolded state baseline

m_U = slope of the unfolded state baseline

$\Delta G^{\circ}_{N->U}$ = free energy of unfolding in the absence of denaturant

R = Gas constant (8.31451×10^{-3} kJ/(mol*K)

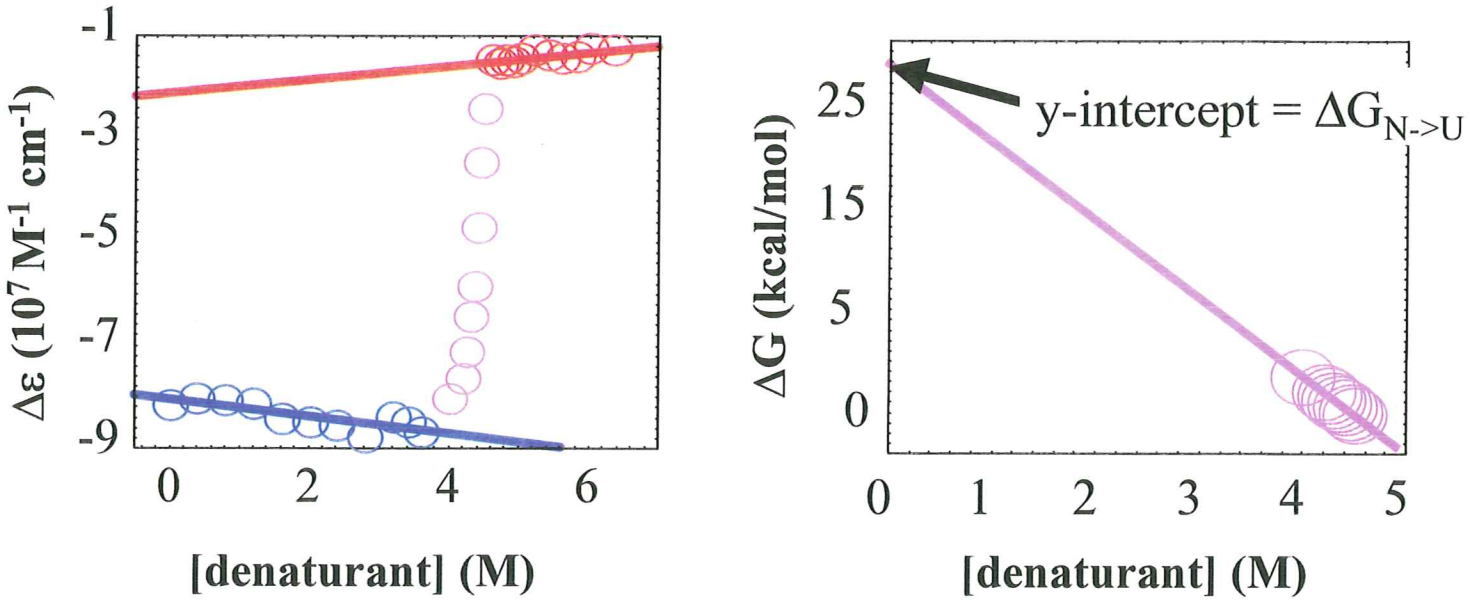
T = temperature in Kelvin

m_G = slope of the calculated dependence of ΔG on [D] (aka m-value)

This equation fits the native-state baseline, the denatured-state baseline and the transition region simultaneously. The LEM considers each point in the transition as an equilibrium constant, defined by its relative position between the native and denatured state baselines. Assuming that the relationship between equilibrium constant (and thus free energy of unfolding) and denaturant concentration is linear, the equation extrapolates to the free energy of unfolding of the protein in the absence of denaturant (see Figure 12). The assumptions of this method are: 1) the native and denatured states of the protein yield a CD signal that changes linearly with denaturant concentration, 2) the free energy of unfolding of the protein changes linearly with denaturant concentration when it is not in a fully native or fully denatured state, and 3) unfolding takes place *via* a two-state process. It should be noted that this makes the analysis model-dependent, that is, the analysis will only work if the protein unfolds *via* a two-state process. The equation was fit to the data using the program KaleidaGraph.

One can determine three useful values from chemical denaturation curves. The first, and most important, is the free energy of unfolding of the protein in the absence of denaturant, or ΔG_U. Also important is the m-value, or the slope of the line describing the relationship between ΔG_U and denaturant concentration (See Figure 12 for a graphical

representation of this quantity). M-values have been shown to correlate with the amount of surface area exposed upon the chemical unfolding of a protein (39). The third parameter is the chemical midpoint of denaturation, or Cm. This value has no thermodynamic significance, but can be useful in comparing proteins.



$$\Delta\varepsilon = \frac{(\Delta\varepsilon_N + m_N[D]) + (\Delta\varepsilon_U + m_U[D]) * e^{-(\Delta G^\circ_{N->U}/RT + m_G[D]/RT)}}{1 + e^{-(\Delta G^\circ_{N->U}/RT + m_G[D]/RT)}}$$

Figure 12: Pictorial representation of the linear extrapolation method. The linear extrapolation method (shown above) fits the the native and denatured state baselines (blue and red, respectively)to straight lines, as shown in the left panel. The equation simultaneously fits to a calculated linear relationship between ΔG and [denaturant] in the transition region, as shown in purple in the right panel. The y-intercept of the latter line corresponds to the ΔG of unfolding in the absence of denaturant.

RESULTS

Construction and Purification of CM-ALBP

The construction of the expression vector for CM-ALBP was screened by restriction enzyme digests, and the results of that screening are shown in Figure 13. Each picture is of a 5% agarose gel stained with ethidium bromide, showing digestion fragments of the CM-ALBP expression vector after treatment with restriction enzymes. From top to bottom, the gels depict digests with *Eag*I, which screens for the R30A and K31A mutations, *Bst*UI which screens for the K37A mutation, and *Nru*I which screens for the K58A mutation. The introduction of additional restriction enzyme sites often produces fewer bands of DNA on an agarose gel due to the creation of two smaller DNA fragments were there was once one. When screening for mutations using restriction digest sites, it is most useful to observe the difference between the wild-type digestion pattern and the digestion patterns of the samples under scrutiny. In this case, only sample E in Figure 13 displays digestion patterns that differ from wild-type for all three digestions, suggesting that the corresponding plasmid accepted all three mutagenic oligonucleotides. This plasmid was sequenced to verify the integrity of the sequence. Sequence integrity (with mutations) was confirmed, and the plasmid was used to express CM-ALBP according to the protocol described in "Materials and Methods". CM-ALBP was successfully purified using the WT-ALBP purification protocol described in "Materials and Methods." Figure 14 depicts the structure of CM-ALBP as determined (theoretically) by SWISS-PROT.

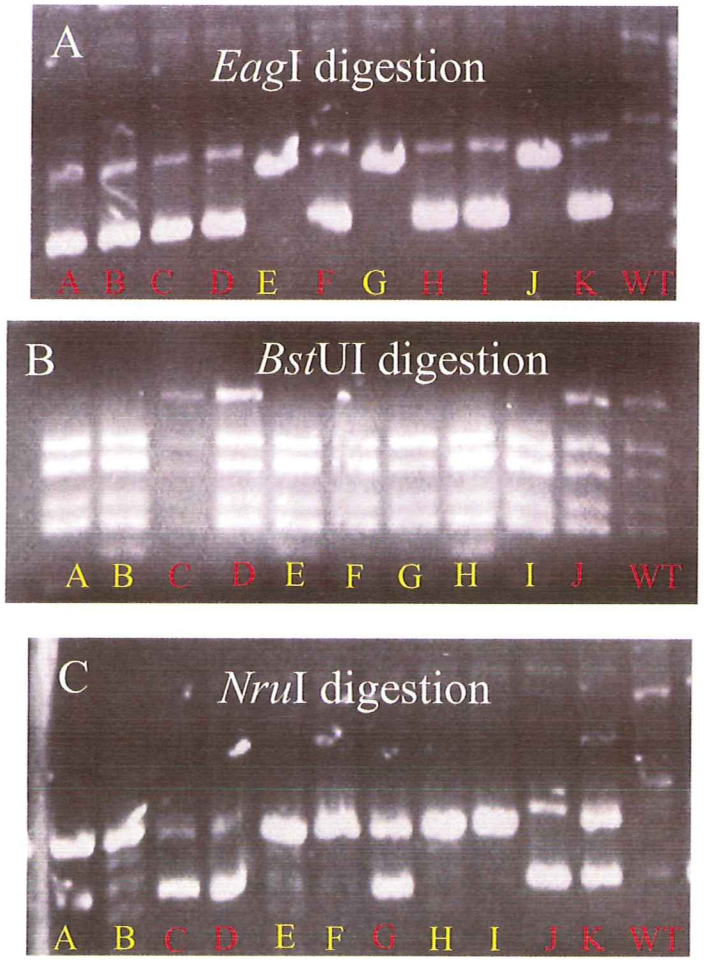


Figure 13: Restriction digests of mutated plasmids. Individual plasmid samples from single colonies are represented by letters; WT refers to wildtype pRSET-ALBP. Samples that accepted a mutation are labeled yellow; those that did not accept a mutation are labeled red. Here, only sample E unambiguously contains all three altered restriction digest sites and therefore all four mutations. In **panel A**, the *Eag*I digestion screens for the R30A and K31A mutations. In **panel B**, the *Bst*UI digestion screens for the K37A mutation, and in **panel C**, the *Nru*I digestion screens for the K58A mutation.

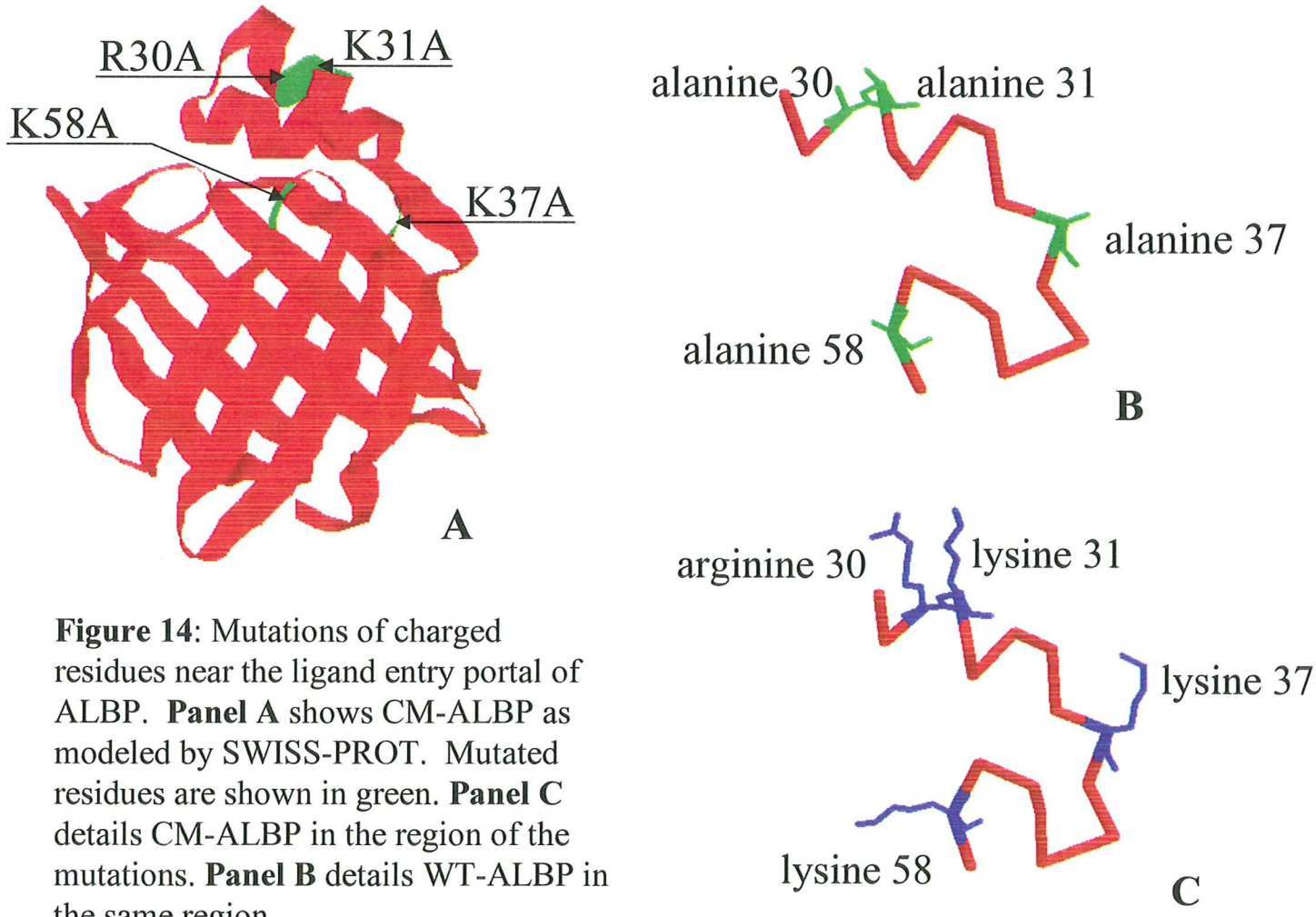


Figure 14: Mutations of charged residues near the ligand entry portal of ALBP. **Panel A** shows CM-ALBP as modeled by SWISS-PROT. Mutated residues are shown in green. **Panel C** details CM-ALBP in the region of the mutations. **Panel B** details WT-ALBP in the same region.

Chemical Denaturations of WT-ALBP

Figure 15 depicts the loss of secondary structure in WT-ALBP with the addition of the chemical denaturant urea. Note that the trough of the spectra are broad; therefore, we chose three wavelengths (216, 217, and 218 nm) at which to monitor the loss of secondary structure in WT- and CM-ALBP. Chemical denaturations of WT-ALBP monitored at 216 nm (nearest the base of the through) are shown in Figure 16. The denaturation curves, at potassium chloride concentrations from 0 to 1 M, are shown fit to the LEM. There is a significant trend in the ΔG_U values obtained from this nonlinear fitting, and the average ΔG_U values and m-values for each titration at 216, 217 and 218 nm are shown in Table 3. From these data it is clear that there is an almost 2.5 kcal/mol stabilization of WT-ALBP upon increasing the KCl concentration from 50 to 250 mM. This is larger than any salt-induced stabilization currently reported in the scientific literature. The addition of more KCl beyond 250 mM does not lead to additional significant stabilization of the protein. Interestingly, if the salt effect on stability is treated as an isotherm, the midpoint of the transition is ~150 mM, near the physiological concentration of potassium.

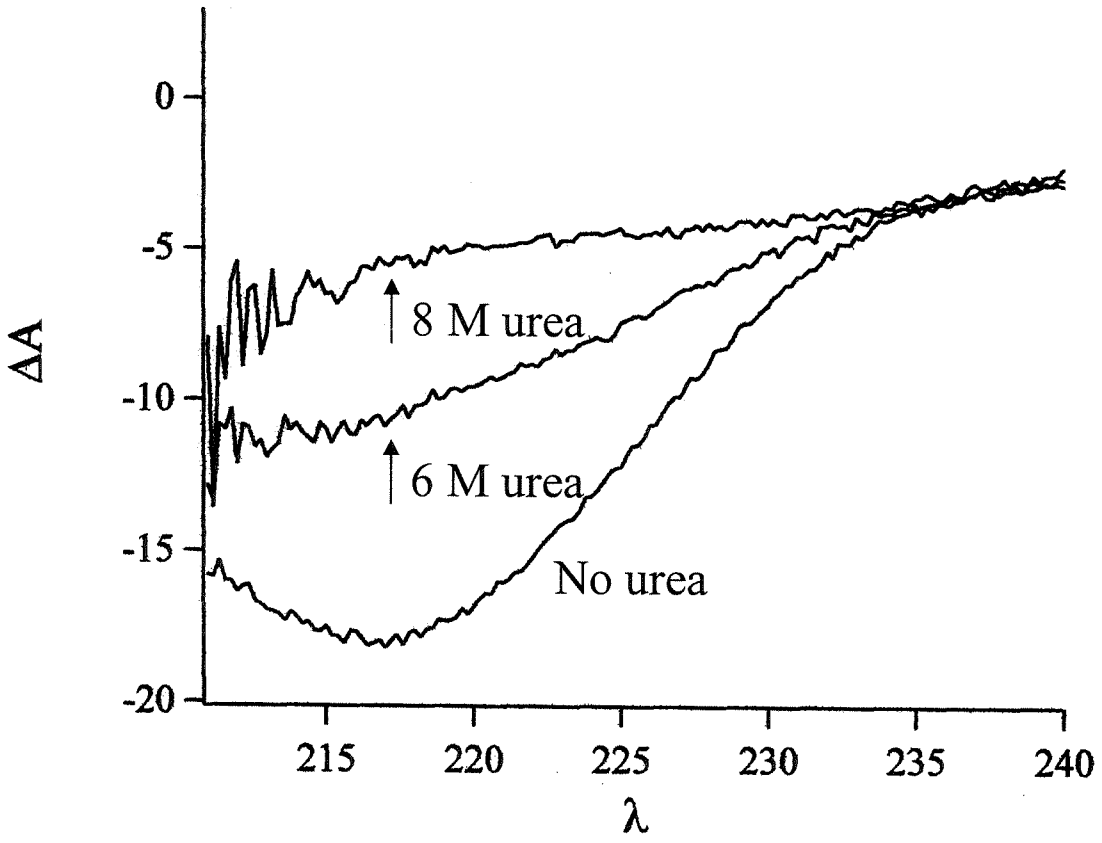


Figure 15: Progressive loss of secondary structure in WT-ALBP upon the addition of urea. The curves above are circular dichroism spectra of WT-ALBP in the presence of increasing amounts of urea. ΔA on the y-axis is the CD signal in millidegrees, or the difference between the amounts of right- and left-circularly polarized light absorbed by the sample at a given wavelength. Wavelengths are given in nanometers on the x-axis.

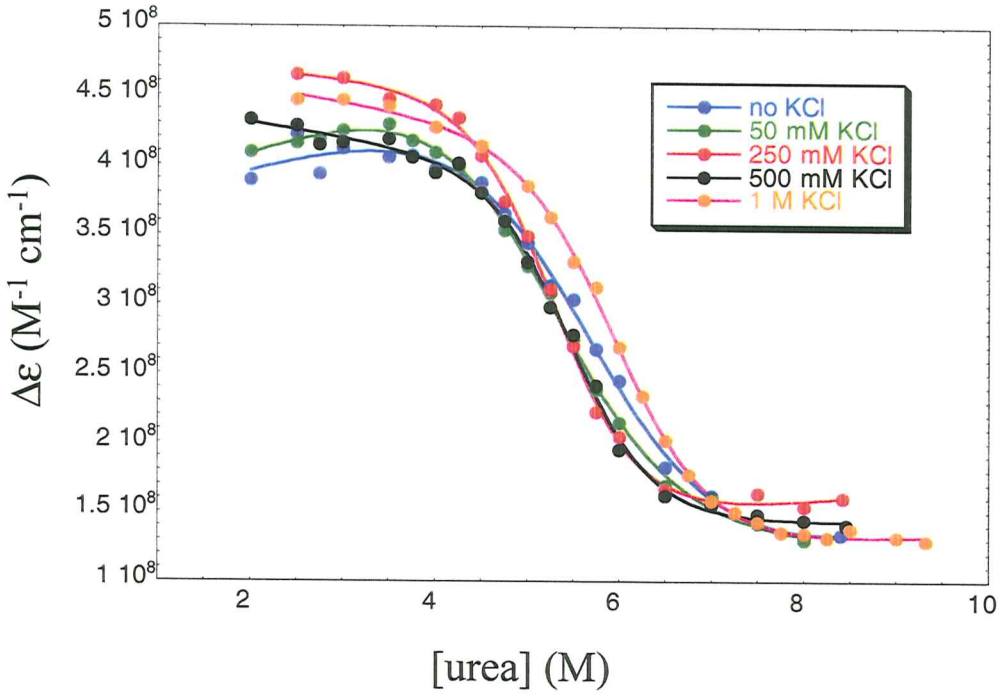


Figure 16: Urea-induced denaturations of WT-ALBP with various concentrations of potassium chloride. All data shown is at 216 nm.

[KCl] (mM)	ΔG_U (kcal/mol)	m-value (kcal/(mol*M)
0	4.8±0.8	-0.8±0.2
50	4.6±0.6	-0.9±0.1
250	6.8±0.7	-1.3±0.1
500	6.7±0.7	-1.2±0.1
1000	6.9±0.4	-1.2±0.1

Table 3: Fitted ΔG_U values and m-values for WT-ALBP.

Chemical Denaturations of CM-ALBP

Chemical denaturations of CM-ALBP monitored at 216 nm (nearest the base of the trough) are shown in Figure 17. The denaturation curves, at potassium chloride concentrations of 50, 250, 500 and 1000 mM, are shown fit to the LEM. There is a significant trend in the ΔG_U values obtained from this nonlinear fitting, and the average ΔG_U values and m-values for each titration at 216, 217 and 218 nm are shown in Table 4. As for WT-ALBP, there is a significant stabilization of CM-ALBP upon the addition of potassium chloride; however, the effect occurs at much higher KCl concentrations for CM-ALBP. As shown in Figure 18, the stability of WT-ALBP increases by 2.5 kcal/mol upon the addition of 250 mM KCl. The addition of the same amount of KCl to CM-ALBP leads to only a 0.5 kcal/mol stabilization, which is within the error of the measurements. Only in the presence of 1 M KCl is CM-ALBP as stable as WT-ALBP in the presence of 250 mM KCl. Clearly, KCl has a stabilizing effect on both CM- and WT-ALBP, but the effect occurs at much higher salt concentrations for CM-ALBP.

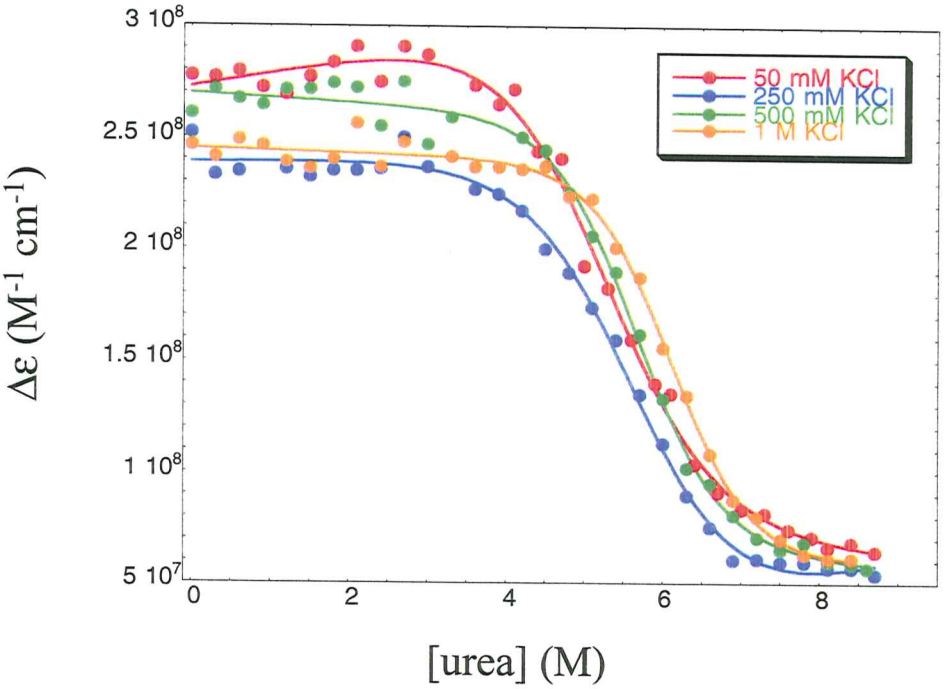


Figure 17: Urea-induced denaturations of CM-ALBP with various concentrations of potassium chloride. All data shown is at 216 nm.

[KCl] (mM)	ΔG_U (kcal/mol)	m-value (kcal/(mol*M)
50	4.5±0.5	-0.9±0.1
250	5.0±0.5	-0.9±0.1
500	6.2±0.7	-1.1±0.1
1000	7.1±0.8	-1.2±0.1

Table 4: Fitted ΔG_U values and m-values for CM-ALBP.

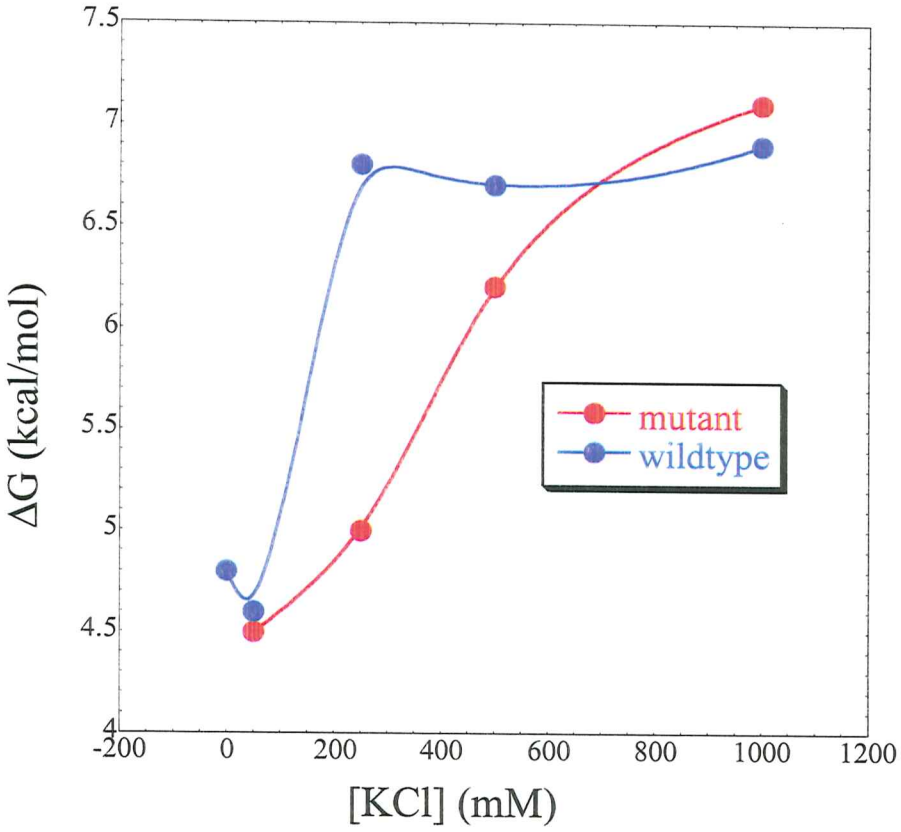


Figure 18: Free energies of unfolding of CM- and WT-ALBP as a function of potassium chloride concentration. Note that the salt-stabilization of WT-ALBP takes effect at a lower KCl concentration than it does for CM-ALBP.

Equilibrium Times of Denaturation for CM- and WT-ALBP

To ensure that the incubation times used in the chemical denaturations were sufficient to achieve equilibrium, we monitored unfolding of both proteins as a function of time. Unfolding was found to be almost immediate and well within the times used in the experiment. Figure 19 shows the change in CD signal at 216 nm for WT-ALBP in the presence of denaturing conditions as a function of time; Figure 20 reports similar data for CM-ALBP.

Refolding of CM- and WT-ALBP

Refolding was accomplished by denaturing CM- and WT-ALBP (at ~10-20uM) in 9 M urea, then dialyzing extensively to remove denaturant. CD spectra of fresh native protein were compared to those of renatured protein. Figure 21 shows native, denatured, and refolded scans of WT-ALBP; Figure 22 shows similar data for CM-ALBP. At 216 nm, the recovery of native state signal was 84% for WT-ALBP and 90% for CM-ALBP.

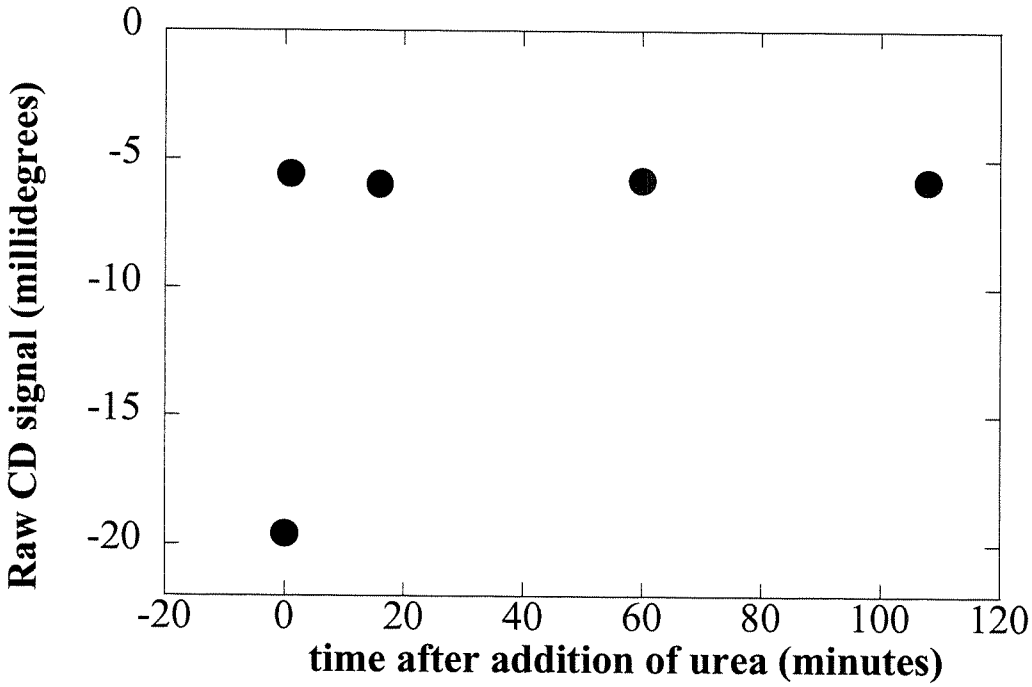


Figure 19: The time required to reach WT-ALBP unfolding equilibrium is illustrated in the graph above. The raw CD signal of WT-ALBP (~ 6μM, pathlength 0.2 cm) at 216 nm is shown as a function of the time after addition of urea (chemical denaturant). Note that denaturation is complete almost immediately.

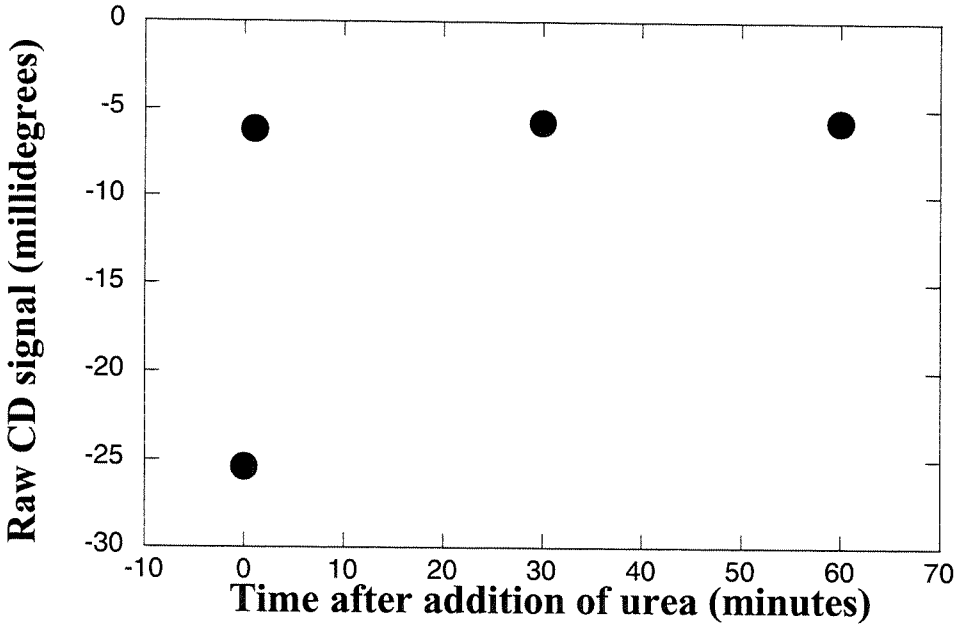


Figure 20: The time required to reach CM-ALBP unfolding equilibrium is illustrated in the graph above. The raw CD signal of CM-ALBP (~14μM, pathlength 0.2 cm) at 216 nm is shown as a function of the time after addition of urea (chemical denaturant). Note that denaturation is complete almost immediately.

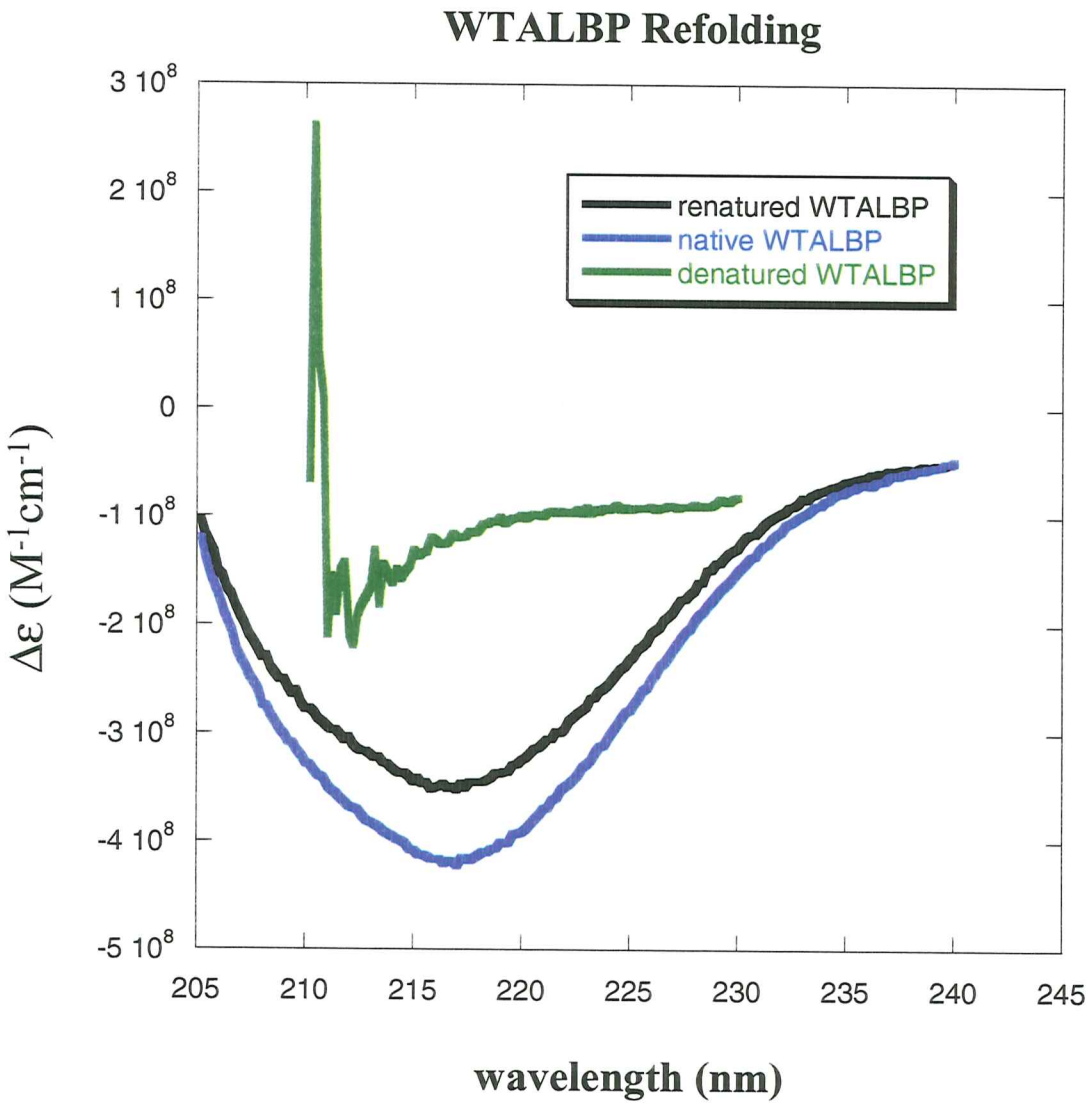


Figure 21: Refolding of WT-ALBP. WT-ALBP was denatured in 9 M urea, then refolded by extensive dialysis in non-denaturing buffer (10 mM $KHPO_4$, 50 mM KCl, pH 7.5). Note the almost complete (84%) recovery of native-state signal.

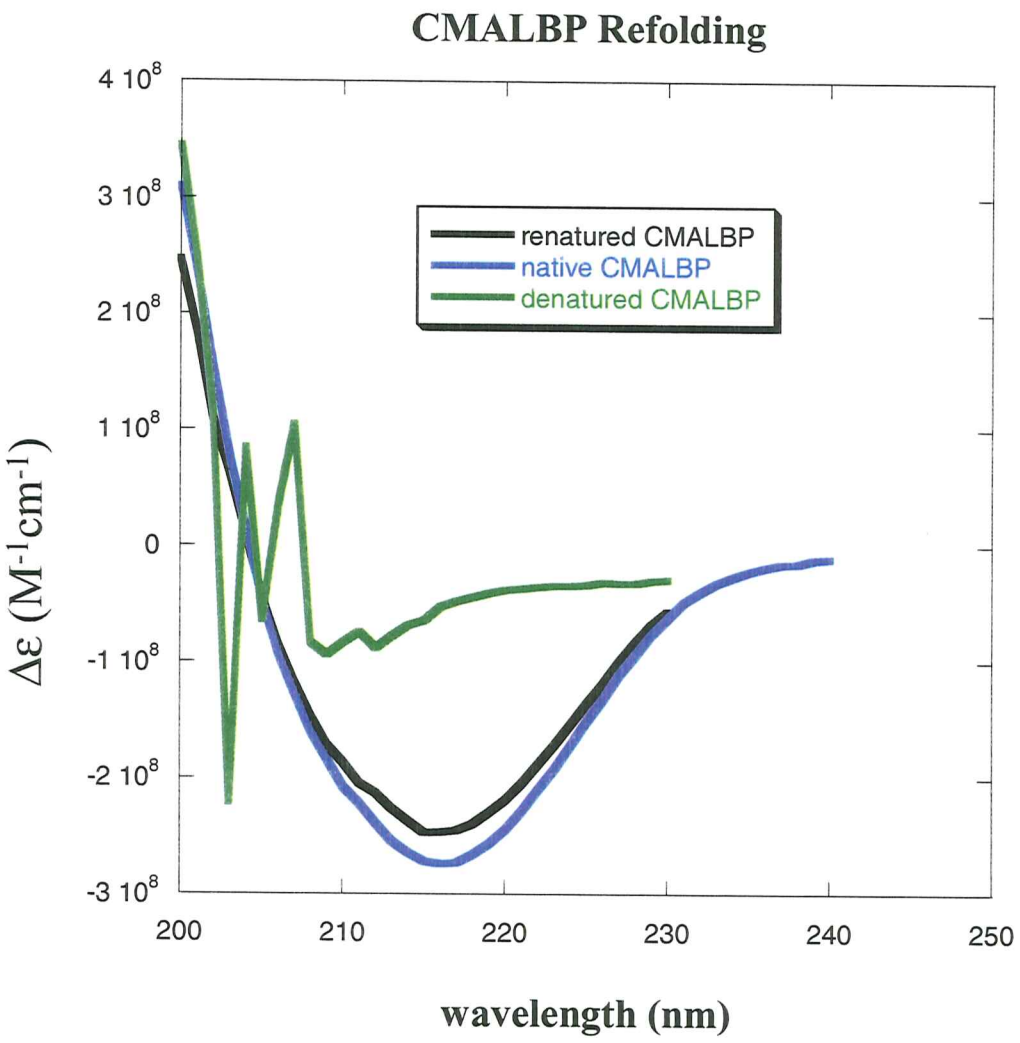


Figure 22: Refolding of CM-ALBP. CM-ALBP was denatured in 9 M urea, then refolded by extensive dialysis in non-denaturing buffer (10 mM $KHPO_4$, 100 mM KCl, pH 7.5). Note the almost complete (90%) recovery of native-state signal.

Chemical Denaturations of Polymerases

Figure 23 shows the native-state CD spectra of Klentaq and Klenow. The broad two-trough shape of the spectra is typical of predominantly α-helical proteins such as these. Chemical denaturations of Klenow and Klentaq DNA polymerases by GdnHCl are shown in Figure 24. Denaturation curves of both proteins at 218, 219, 220 and 221 nm are shown fit to the LEM. Nonlinear analysis by the LEM at all four wavelengths yields the average free energies of unfolding and m-values listed in Table 5. Given these values, Klentaq is more stable than Klenow by nearly 6-fold, and its Cm (midpoint of chemical unfolding) is higher by 3 M GdnHCl. In addition, the dependence of ΔG_U on the denaturant concentration (represented by the m-value) for Klentaq is nearly double that for Klenow.

Chemical denaturations of the full-length polymerases are shown in Figure 25. The denaturation of Taq polymerase is clearly not two-state, given the distinctly biphasic nature of the denaturation profile. In Figure 25, the denaturation curve for Taq polymerase is fit to a double-transition form of the linear extrapolation method only to aid in the visual inspection of the graph. The double-transition model contains too many parameters to yield uniquely determined thermodynamic values for this system; thus, no values for the fit are reported here. The denaturation of Pol 1 appears to be possibly two-state, and fits reasonably well to the two-state form of the linear extrapolation method. However, given that the thermal denaturation of Pol 1 gives a monophasic (single-peak) transition that can be successfully deconvoluted into two transitions, it is likely that the chemical denaturation of Pol 1 is also not a two-state process. Thus, no values for the fit are reported here.

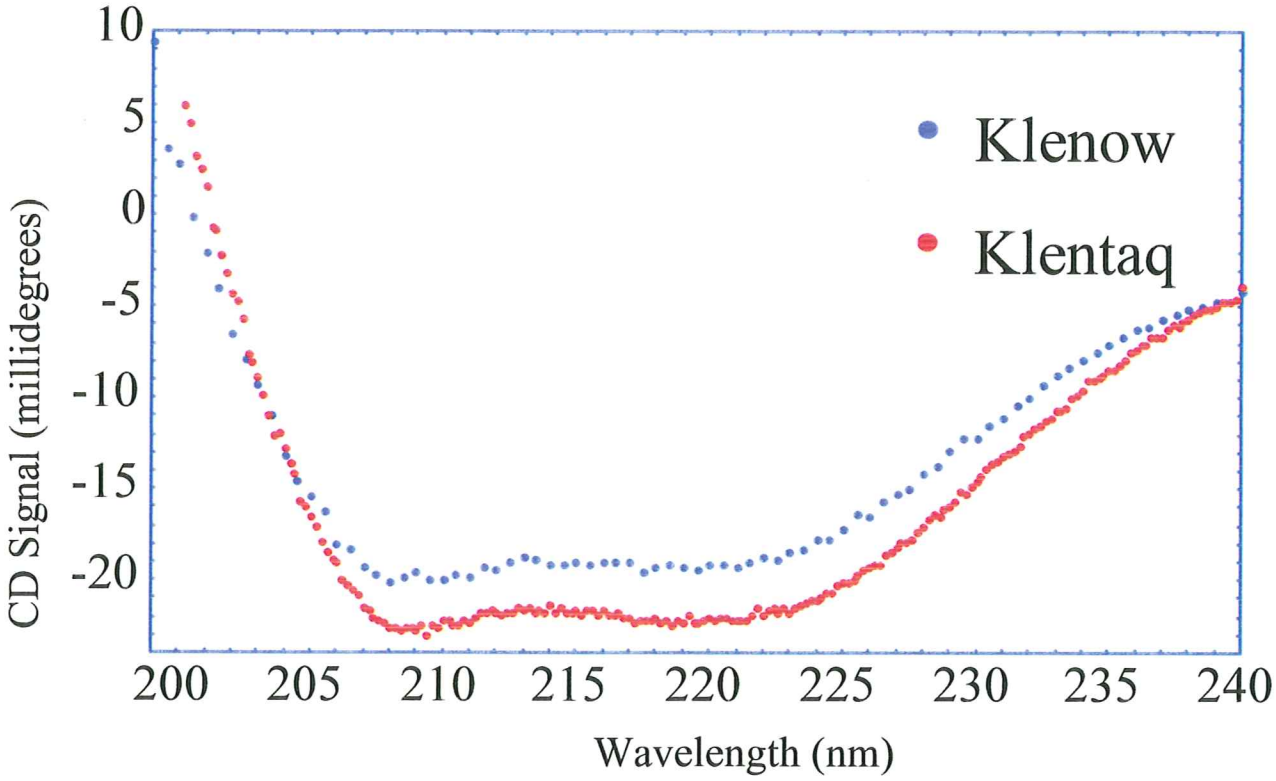


Figure 23: Native circular dichroism spectra of Klenow (blue) and Klentaq (red). Note the two-trough shape of the curves, typical of predominantly α-helical proteins such as these.

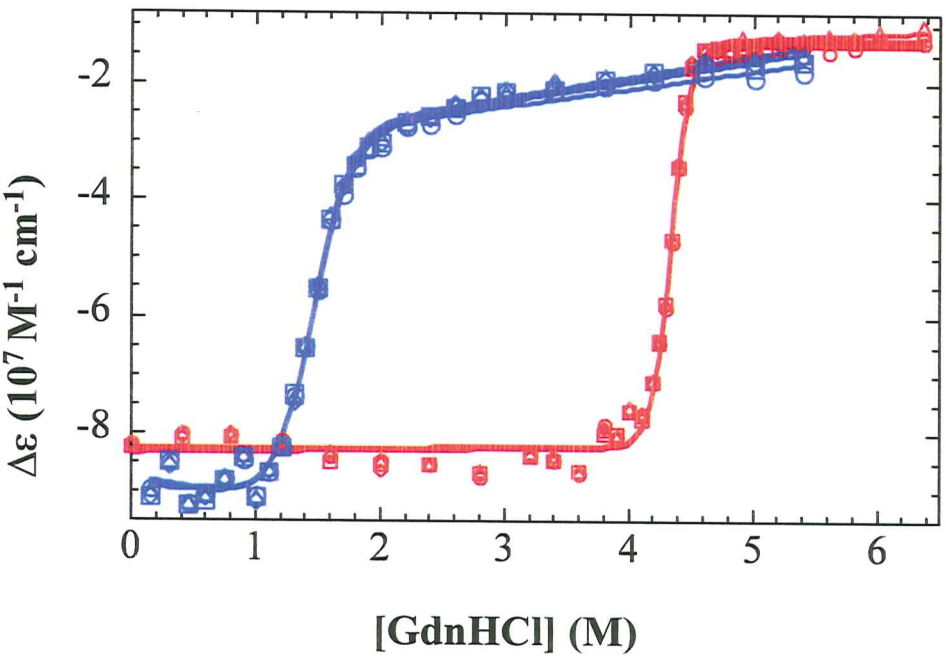


Figure 24: Guanidine hydrochloride denaturations of Klenow (blue) and Klentaq (red). The graph shows molar ellipticity as function of denaturant concentration. Circles and diamonds represent individual samples; lines are the best fits of the linear extrapolation method to the data. Data collected at 218, 219, 220 and 221 nm are shown for both polymerases. Fitted ΔG values, m-values, and Cms are shown in Table 5.

	ΔG (kcal/mol)	Cm (M) (denaturation midpoint)	m-value (kcal/(mol*M)
Klenow	4.7 ± 1.2	1.5 ± 0.01	3.3 ± 0.8
Klentaq	27.8 ± 4.3	4.3 ± 0.05	6.5 ± 1.0

Table 5: Fitted parameters for chemical denaturations of Klenow and Klentaq. Note the ~6 fold increase in free energy of unfolding (ΔG_U) for Klentaq relative to Klenow, as well as the tripling of the Cm and the doubling of the m-value.

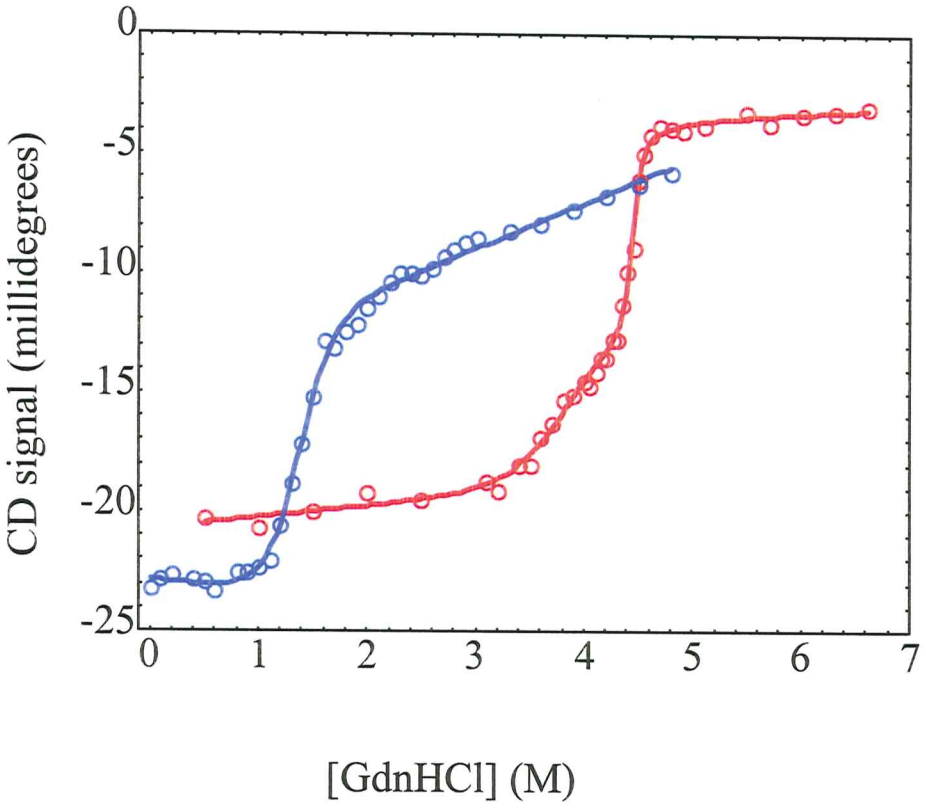


Figure 25: Guanidine hydrochloride denaturations of full length polymerases. The graph shows CD signal as a function of denaturant concentration ([GdnHCl]), with circles representing individual samples. The lines through the data are the best fits of 2-state (for Pol 1, in blue) and double-transition (for Taq, in red) forms of the linear extrapolation method. Data collected at 218 nm is shown.

DISCUSSION

Polymerases

The chemical denaturations of the polymerases confirm the observations of this lab based on the thermal denaturations of the polymerases. Namely, the two structural domains of Pol 1 interact significantly when folded, while the two structural domains of Taq polymerase appear to denature independently. Also, just as Taq polymerase and Klentaq are resistant to heat, so are they resistant to denaturing chemicals, as evidenced by their high Cms.

Given the Tms for Klenow and Klentaq obtained from thermal denaturations and the ΔG_Us of the proteins obtained by the chemical denaturations in this study, it is possible to calculate theoretical Gibbs-Helmholtz relationships for the two proteins. Given the Tm and ΔH_m from thermal denaturations and the ΔG_U at 25°C from the chemical denaturations, we used the Gibbs-Helmholtz relationship to solve for the ΔCp of denaturation for Klentaq and Klenow. Then, using the calculated ΔCp, we produced theoretical sets of ΔG_U *versus* temperature data which includes the single measured ΔG_Us. The resulting curves are shown in Figure 26. Referring back to Figure 10 on page 25, one can suggest which of the three thermodynamic models explains the thermostability of Klentaq. In fact, Klentaq seems to follow a combination of the increased stability model and the shifted stability model. That is, Klentaq displays a higher overall ΔG_U at almost all temperatures as a well as a right-shifted temperature range of maximal stability. The four experimental points for this analysis are shown as open circles on the theoretical curves.

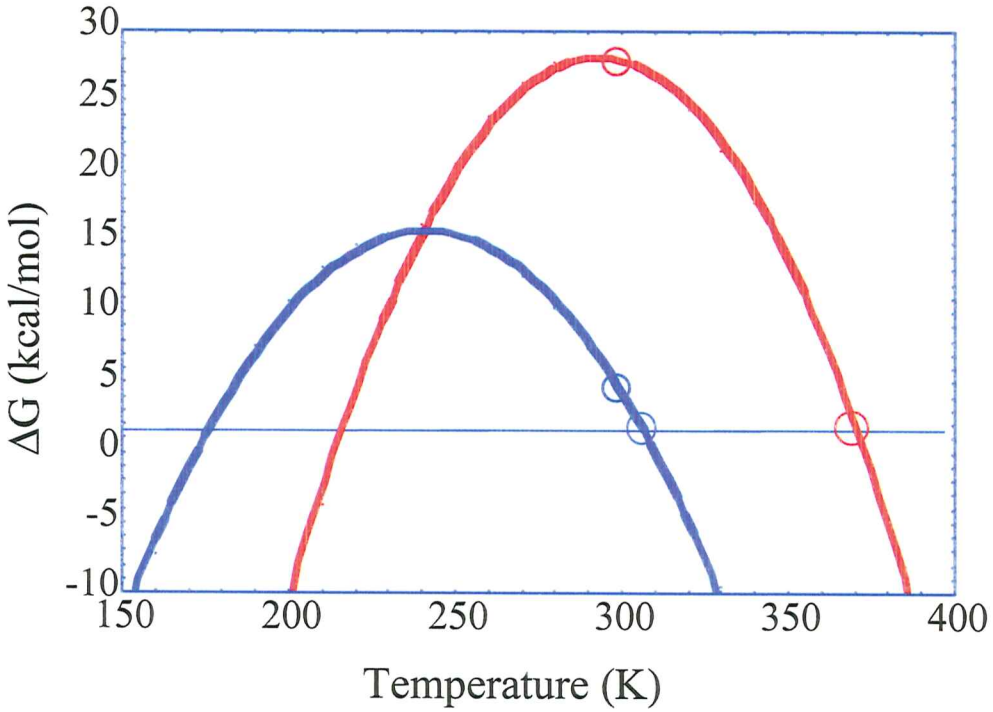


Figure 26: Projected Gibbs-Helmholtz plots describing the temperature dependence of free energy for Klentaq (red) and Klenow (blue). Theoretical data sets for both proteins were generated using the values of ΔH_{cal}, ΔG_U, and Tm reported in this study. Open circles represent experimental data.

To confirm this projection of the thermodynamic basis for the thermostability of Klentaq, additional experiments, namely chemical denaturations of Klentaq and Klenow at different temperatures, are currently in progress.

It has been suggested that the m-values obtained from chemical denaturations correlate with the change in exposed surface area of the protein upon unfolding (39). Given the doubling of Klentaq's m-value relative to Klenow's, this suggests that either the native or the denatured states (or both) of the two proteins differ significantly. Analytical ultracentrifugation experiments performed in this laboratory have found that this is not the case: Klenow and Klentaq have identical exposed surface areas in both native and denatured states (Khan and LiCata, unpublished). This suggests that, at least for this set of proteins, m-values do not correlate with the amount of surface area exposed upon unfolding.

CM- and WT-ALBP

Anthony Fink of the University of California, Santa Cruz, has suggested that patches of same-sign surface charges on marginally stable globular proteins are destabilizing due to unfavorable intramolecular charge-charge interactions, and that this destabilization is counteracted by the addition of salts (40). Our results for the chemical denaturation of WT-ALBP initially support this hypothesis. WT-ALBP is marginally stable by Fink's standards and it is clearly stabilized by the addition of potassium chloride. However, from this hypothesis one would expect that the removal of surface charge would have a stabilizing effect. We find that CM-ALBP is no more stable than WT-ALBP at low (50 mM) and high (1 M) KCl concentrations, and it is in fact less stable at intermediate (250-500 mM) KCl concentrations. A model that explains our results is that the positive patch of residues we neutralized is not inherently destabilizing, but counterion binding (anions, in this case) to the patch has a super-stabilizing effect.

The dependence of CM-ALBP's stability on KCl concentration demonstrates the need to characterize mutant proteins across a range of solution conditions. At 50 mM KCl, the difference between the stabilities WT-ALBP and CM-ALBP is negligible, but at physiological potassium concentration (~145 mM), the difference in their stabilities is more significant (> 1 kcal/mol). At 250 mM KCl, the stability difference jumps to almost 2 kcal/mol.

The fact that the stability of WT-ALBP is most sensitive to potassium chloride concentration at the physiological concentration of potassium suggests that ALBP is controlled in part by salt *in vivo*. It is believed that ALBP must maintain flexibility in order to bind a ligand. If salt-stabilization of the protein is achieved by a stiffening of the

protein and a concomitant loss of flexibility, the protein's ability to bind ligands may also be affected by salt concentration.

REFERENCES

1. Stryer, L. *Biochemistry, 4th ed.* (1995) W. H. Freeman and Co., New York
2. Stewart, J. M. (2000) *Cell. Mol. Life Sci.* **57**, 1345-1359.
3. Thompson, J.; Reese-Wagoner, A. and Banaszak, L. (1999) *Biochimica et Biophysica Acta* **1441**, 117-130.
4. Simpson, M. A.; LiCata, V. J.; Coe, N. R. and Bernlohr, D. A. (1999) *Mol. And Cell. Biochemistry* **192**, 33-40.
5. Banaszak, L.; Winter, N.; Xu, Z.; Bernlohr, D. A.; Cowan, S. and Jones, T. A. (1994) *Adv. Protein Chem.* **45**, 89-151.
6. Xu, Z.; Bernlohr, D. A. and Banaszak, L. J. (1993) *J. Biol. Chem.* **268**, 7874-7884.
7. Chinander, L. L. and Bernlohr, D. A. (1989) *J. Biol. Chem.* **264**, 19564-19572.
8. Buelt, M. K.; Xu, Z.; Banaszak, L. J. and Bernlohr, D. A. (1992) *Biochemistry* **31**, 3493-3499.
9. LiCata, V. J. and Bernlohr, D. A. (1998) *PROTEINS: Structure, Function, and Genetics* **33**, 577-589.
10. LaLonde, J. M.; Bernlohr, D. A. and Banaszak, L. J. (1994) *Biochemistry* **33**, 4885-4895.
11. LaLonde, J. M.; Levenson, M. A.; Roe, J. J.; Bernlohr, D. A. and Banaszak, J. L. (1994) *J. Biol. Chem.* **269**, 25339-25347.
12. Ory, J. J. and Banaszak, L. J. (1999) *Biophysical Journal* **77**, 1107-1116.
13. Kane, C. D. and Bernlohr, D. A. (1996) *Anal. Biochem.* **233**, 197-204.
14. Shen, W.-J.; Sridhar, K.; Bernlohr, D. A. and Kraemer, F. B. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 5528-5532.
15. Holm, C.; Langin, D.; Manganiello, V.; Belfrage, P. and Degerman, E. (1997) *Methods Enzymol.* **286**, 45-67.
16. Hotamisligil, G. S.; Johnson, R. S.; Distel, R. J.; Ellis, R.; Papaioannou, V. E. and Spiegelman, B. M. (1996) *Science* **274**, 1377-1379.
17. Smith, E. R. and Storch, J. (1999) *J. Biol. Chem.* **274**, 35325-35330.
18. Kim, Y.; Eom, S. H.; Wang, J.; Lee, D.; Suh, S. W. and Steitz, T. A. (1995) *Nature* **376**, 612-616.
19. Lyamichev, V.; Brow, M. A. D.; Varvel, V. E. and Dahlberg, J. E. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 6143-6148.
20. Eom, S. H.; Wang, J. and Steitz, T. A. (1996) *Nature* **382**, 278-281.
21. Perler, F. B.; Kumar, S. and Kory, H. (1996) *Advances in Protein Chemistry* **48**, 377-435.
22. Brock, T. D. (1974) in *Bergey's Manual of Determinative Bacteriology, 8th ed.* Eds Buchanan and Gibbons (Williams and Wilkins, Baltimore, MD) p. 285.
23. Lawyer, F. C.; Stoffel, S.; Saiki, R. K.; Myambo, K.; Drummond, R. and Gelfand, D. H. (1989) *J. Biol. Chem.* **264**, 6427-6437.
24. Korolev, S.; Nayal, M.; Barnes, W. M.; DiCera, E. and Waksman, G. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 9264-9268.
25. Gelfland, D. H.; Stoffel, S.; Lawyer, F. C. and Saiki, R. L. (1992) U. S. Patent 5,079,352.
26. Beese, L. S.; Derbyshire, V. and Steitz, T. A. (1993) *Science* **260**, 352-355.

27. Tindall, K. R. and Kunkel, T. A. (1988) *Biochemistry* **27**, 6008-6013.
28. Sterner, L. and Liebl, W. (2001) *Critical Reviews in Biochemistry and Molecular Biology* **36**, 39-106.
29. Jaenicke, R. (1993) in *Biocatalyst Design for Stability and Specificity*, Eds. Himmel & Georgiou (American Chemical Society, Washington, D. C.) pp. 53-67.
30. Rees, D. C. and Adams, M. W. (1995) *Structure* **3**, 251-254.
31. Xu, Z.; Buelt, M. K.; Banaszak, L. J. and Bernlohr, D. A. (1991) *J. Biol. Chem.* **266**, 14367-14370.
32. Engleke, D. R.; Krikos, A.; Bruck, M. E. and Ginsburg, D. (1990) *Anal. Biochem.* **191**, 396-400.
33. Barnes, W. M. (1992) *Gene* **112**, 29-35.
34. Barnes, W. M. (1995) U.S. Patent 5,436,149.
35. Joyce, C. M. and Derbyshire, V. (1995) *Methods Enzymol.* **262**, 3-13.
36. Nozaki, Y. (1972) *Methods Enzymol.* **26**, 43-50.
37. Pace, C. N. (1986) *Methods Enzymol.* **14**, 266-280.
38. Santoro, M. M. and Bolen, D. W. (1988) *Biochemistry* **27**, 8063-8068.
39. Myers, J. K.; Pace, C. N. and Schlotz, J. M. (1995) *Protien Science* **4**, 2138-2148.
40. Nishimura, C. Uversky, V. N. and Fink, A. L. (2001) *Biochemistry* **40**, 2113-2128.

ABBREVIATIONS

RMS: root mean squared
CM-ALBP: charge-mutant adipocyte lipid binding protein
DNA: deoxyribonucleic acid
E. coli: *Escherichia coli*
nm: nanometers
PDB ID: Protein Data Bank Identification Number
RPM: revolutions per minute
ssDNA: single-stranded DNA
T. aquaticus: *Thermus aquaticus*
WT-ALBP: wild-type adipocyte lipid binding protein
OD: optical density
IPTG: isopropyl-B-D-thiogalactopyranoside
g: gram
M: molar (moles/liter)
mM: millimolar (millimoles/liter)
HCl: hydrochloric acid
NaCl: sodium chloride
KCl: potassium chloride
$KHPO_4$: potassium phosphate
EDTA: ethylenediaminetetraacetic acid
°C: degrees Celsius
K: Kelvins
mL: milliliters
HEPES: N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
kcal/mol: kilocalories/mole
Cm: chemical midpoint of denaturation
cm: centimeters
$MgCl_2$: magnesium chloride
CD: circular dichroism
LEM: linear extrapolation method
ΔG_U: free energy of unfolding
$\Delta\varepsilon$: molar ellipticity
ΔCp: change in heat capacity
ΔH_{cal}^{app}: apparent enthalpy of calorimetric unfolding
ΔH_U: enthalpy of unfolding
ΔH_m: enthalpy of unfolding at the Tm
Tm: temperature midpoint of unfolding
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
cDNA: coding DNA
PEI: poly(ethyleneimine)
DTT: dithiothreitol
ΔS: change in entropy
BCA: bicinchoninic acid
$(NH_4)_2SO_4$: ammonium sulfate
µM: micromolar (micromoles/liter)

ACKNOWLEDGEMENTS

First, I would like to thank my thesis advisor Dr. Vince LiCata for his patience, guidance, and limitless support. Without him this project would have been impossible. His mentorship over the years has been a source of inspiration for me and I will forever be grateful for the effect he has had on my life.

I would also like to thank the members of the LiCata lab who have provided support, both moral and technical, throughout this project. I thank Carmen Ruiz, Kausiki Datta and Xuemei Yang for protein purification, Allison Joubert, Lauren Sohljoo and Fengang Peng for additional chemical denaturations that have confirmed the data reported here, and Greg Thompson for assistance in protein purification. In addition, Allison Joubert performed the denaturation of WT-ALBP in 1 M KCl that is reported in this study, and I offer my acknowledgement and thanks to her. Farheen Khan, Irene Karantzeni, and Xuemei Yang have performed experiments that have been the basis for this work, and to them I am grateful. I also thank Dr. David Bernlohr of the University of Minnesota for the use of his laboratory during the preparation of mutant ALBP. Members of his laboratory, namely Ann Hertzel and Anne Jenkins, were of immense help during the completion of that project. I also thank Dr. Bernlohr and Catherine Joyce of Yale University for their gifts of expression vectors for protein purification.

Special thanks goes to my committee members, Dr. Marcia Newcomer and Dr. David Spivak, to whom I am grateful for their willingness to serve on my committee.

My work would have been impossible without the technical support of the people at Aviv Instruments, especially Guy McCafferey and Glen Ramsey. Guy's patient assistance, which sometimes lasted 6 hours, kept the CD in working order, and I am eternally grateful. I also thank Martha Juban of the LSU Protein facility for the use of the CD spectrometer and for her patience.

Finally, I thank my family, whose moral support has been invaluable, and my friends, whose patience with me has been saint-like.

Note: Parts of the work reported in this thesis were performed outside of the 2001-2002 school year. They are described here because they are necessary to give an understanding of the project.