THE ROLE OF INTERNAL WATER MOLECULES IN THE STRUCTURE AND STABILITY OF BACTERIOPHAGE T4 LYSOZYME

by

KANE LIBRUM ANDERSON

A THESIS

Presented to the Department of Chemistry
and the Robert Donald Clark Honors College of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Bachelor of Arts

May 1998

An Abstract of the Thesis of

Kane Librum Anderson for the degree of Bachelor of Arts

in the Department of Chemistry to be taken June 1998

Title: THE ROLE OF INTERNAL WATER MOLECULES IN THE STRUCTURE AND STABILITY OF BACTERIOPHAGE T4 LYSOZYME

Approved: [redacted]

Michael L. Quillin, Ph. D.

Proteins are macromolecules that participate in nearly all biological functions. For proteins to function they must be stable in their active form. While the stable forms of many proteins have densely-packed hydrophobic interiors, others have internal cavities in which water molecules reside. The purpose of this study was to analyze through genetic manipulation the role of the internal water molecules in the stability of a specific protein, bacteriophage T4 lysozyme. In addition, the structures of the mutant lysozymes will be determined by x-ray crystallography to help elucidate the physical nature of protein stability by revealing its structural consequences. Two mutants were constructed. The Ile 58 → Thr mutant was designed to create a cavity and introduce a water molecule, but failed to do so. The change in stability of this mutant, $\Delta\Delta G$ was -3.24 kcal/mol. The Gln 105 → Met mutant was designed to eliminate hydrogen bonds to a water molecule. This mutant formed a cavity by displacing water and had a $\Delta\Delta G$ was -1.0 kcal/mol. The results of these experiments lead to the conclusion that because the nature of hydrated cavities varies drastically, they must be extremely dependent on their local environments.

Acknowledgments

I would like to extend my thanks to Dr. Michael L. Quillin for his infinite patience with me throughout the course of this project and his guidance and discussion regarding the concepts involved. I thank Nadine Gassner for her assistance with the mutagenesis and Leslie Gay and Hong Xiao for expression and purification of the lysozyme mutants. Finally, I appreciate the graciousness of Dr. Brian Matthews in allowing me the privilege of performing research in the Matthews' Laboratory.

Table of Contents

List of Figures

List of Tables

Introduction

Proteins are essential to nearly every biological function. To work properly however, they must be stable. Misfolded and unstable proteins are the biochemical basis for many diseases. Therefore, understanding the forces controlling protein stability is a fundamental step towards comprehending larger biological considerations. Globular proteins, so called because they have a compact shape when folded, are the focus of my study. Like other proteins, they are large, complex molecules which are composed of amino acids linked end-to-end, like beads on a string. They can fold into unique three-dimensional structures based on the order and the identity of their amino acids. The folding process is so complex that it is currently impossible to predict the structure of a protein on the basis of order and identity of amino acids alone. One thing that is known about the folding of globular proteins is that their apolar amino acids, which contain mainly carbon and hydrogen in their side chains, are usually found in the interior of the protein, while their polar amino acids, those that contain nitrogen and oxygen in their side chains, point outwards and interact with water. In 1977, Richards helped illustrate the point that globular proteins typically have well-packed apolar cores that are devoid of cavities and solvent. Some protein molecules, however, are imperfectly packed and have internal cavities, which may or may not contain water molecules. It is the goal of this study to address the role of these internal water molecules in the structure and stability of proteins by attempting to engineer water-binding sites into and out of proteins.

This project explores protein stability by altering a well-studied protein, bacteriophage T4 lysozyme, that degrades bacterial cell walls after the bacteriophage has finished replicating inside. Lysozyme has internal cavities in which four water molecules reside, coordinated by hydrogen bonds with the side chain atoms of polar amino acids and main chain atoms lining the cavities (see Figure 1). While many proteins use water to catalyze reactions, the internal water molecules of the protein in this study are too far from

the active site to participate in catalysis. Therefore, if they are not present for catalysis they might have some different role.

To examine the stability of lysozyme I attempted to add and remove the internal water molecules by using site-directed mutagenesis to alter the neighboring amino acid side chains. This technique relies on the fact that genes direct the synthesis of proteins, such as lysozyme. First, a gene is transcribed into RNA and then translated into a protein. Mutagenesis is the process by which I made changes to the DNA that resulted in alteration in the protein itself. Site-directed mutagenesis allows precise control over the specific amino acids that are mutated. Since these changes take place near the internal water molecules in the folded protein, disturbing these amino acids might perturb the water molecules as well. In one mutation, I attempted to introduce another water molecule inside the protein, in the other, to remove one. After the changes were made, I analyzed the effect of adding or removing the water molecules on the stability and structure. By studying these forces we can better understand proteins and the complicated systems in which they participate as a whole.

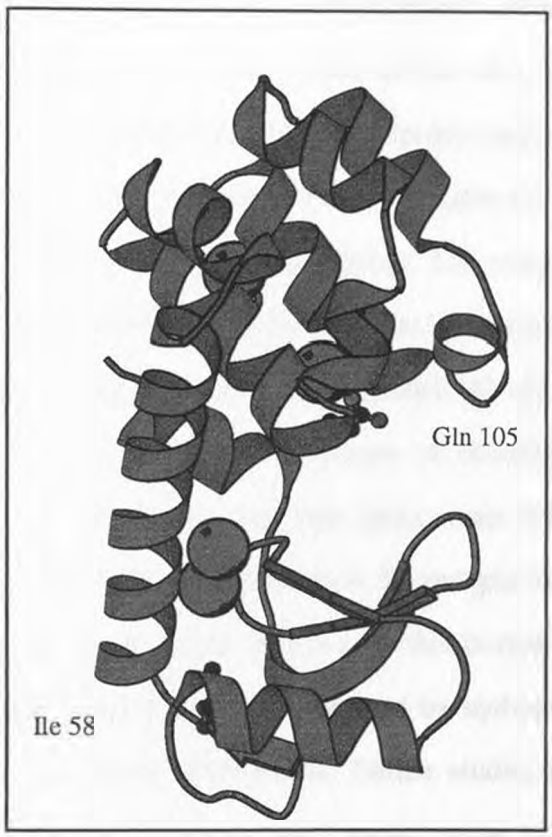

Figure 1: Wild-type bacteriophage T4 lysozyme. Internal water molecules are represented by large spheres. Locations of site-directed mutagenesis are indicated by smaller spheres.

Three separate studies on T4 lysozyme, the protein involved in this study, have demonstrated that the effect on protein stability from different amino acid substitutions can be highly dependent on the identity of the amino acid replacement. One such study illustrated that introduction of polar amino acids that lack hydrogen-bonding partners into the core of lysozyme destabilizes the folded state by 1-3 kcal/mol (Blaber *et al.*, 1993).

Another study analyzed the effect of truncating long amino acids into shorter ones, thereby creating cavities in the core of the protein. Erikkson *et al.*, showed that there is a linear decrease in the stability of lysozyme as the volume of the newly-created cavity increases (1993). They determined this loss of stability to be 24 - 33 cal/mol/Å^3.

Finally, others have examined the role of water molecules in the stability of proteins. They replaced threonine, a small, polar amino acid, with glycine, the smallest amino acid, on the surface of lysozyme. This substitution made enough space available to allow the introduction of a water molecule which presumably served the role of the missing threonine by fulfilling lost hydrogen bond interactions. This mutant was nearly as stable as the wild-type protein. Apolar replacements at this same site, lacking the ability to hydrogen bond, however, destabilize the protein by 1-3.1 kcal/mol (Alber, 1987).

I selected amino acids near internal waters of bacteriophage T4 lysozyme for mutagenesis to specifically address the role that these water molecules play in lysozyme stability. The two sites I identified were isoleucine 58 and glutamine 105. I chose the first site because isoleucine is a large apolar amino acid that borders a cavity containing two water molecules. The water molecules are coordinated by hydrogen bonds to a network of neighboring side chains and main chain atoms. Earlier studies examined the role of this isoleucine by replacing it with alanine, a small, apolar amino acid. This process created a 18 $Å^3$ cavity (Xu, 1998). In a study of different proteins containing internal waters, Rashin, *et al.*, found the minimum cavity volume in which a water can be seen using crystallography was between 11.5 $Å^3$ (the van der Waals' volume of water) to 30 $Å^3$ (1986). Therefore, this mutant might be expected to introduce water into the cavity. Williams points out that the criteria for inclusion of water vary dramatically. He found that internal waters within proteins form an average of three hydrogen bonds with the cavity walls (Williams, 1994). This experiment combines the approaches of past experiments. By replacing the large, apolar isoleucine with the small, polar threonine, this study examines the effect of simultaneously adding a hydrogen-bonding partner and increasing the cavity volume to determine if the combination is sufficient to incorporate water into the protein structure.

The second prospective site in this experiment was glutamine 105, a large polar amino acid. This amino acid forms a hydrogen bond with an adjacent, internal water

molecule via its amide functionality (a specific arrangement of carbon, nitrogen, and oxygen). Previous work found that mutating this glutamine to alanine (small apolar), glutamate (similarly-sized, negatively charged), and glycine (smallest apolar) all decreased the stability of the protein by 0.6-1.5 kcal/mol (Pjura, 1993). In this study however, I introduced a large, apolar amino acid, methionine, which is approximately the same size and shape as glutamine, to analyze the effect of disrupting the hydrogen bonding partner of one of the water molecules in the cavity by introducing a thioether (sulfur-containing) group in the place of the amide while maintaining the size of the cavity.

Materials and Methods

Site-directed mutagenesis is the process by which alteration in DNA change the protein product of that DNA. The DNA of the gene that directs the production of the lysozyme protein has previously been isolated and placed, or cloned, into the DNA of a virus. Viral DNA is a good host for foreign DNA because it is small and designed to be easily replicated. While DNA is often double-stranded, the virus utilized in this technique produces single-stranded DNA that can be isolated. This single-stranded DNA can be used as a template which can be mutated and turned into double-stranded DNA by using techniques of molecular biology. The DNA can then be introduced into a new strain of bacteria where the protein is made and isolated from the bacterial cells. A more detailed description of the materials and methods of experiment follows.

Site-directed mutagenesis- The Kunkel template method of oligonucleotide-mediated mutagenesis requires two single-stranded pieces of DNA: a uracil-containing template strand and a mutation introducing oligonucleotide primer. To create the template, a culture of *E.coli* strain CJ236 (a *dut- ung-* strain that lacks the DNA repair enzyme uracil N-glycosylase and therefore cannot degrade uracil-containing DNA) was infected with M13 bacteriophage (see Figure 2, step 1). The M13 phage subsequently replicate their DNA in the uracil-permissive environment. The single-stranded (ss) DNA from the culture of phage-infected CJ236 was isolated and purified by aliquoting 1.5 mL of culture into eppendorf tubes and centrifuging at 14,000 RPM for 2 minutes at 4° C. 300 µL of 5X polyethylene glycol/NaCl solution (15 grams PEG 8000, 14.6 grams NaCl/100 mL) was added to the remaining mixture to precipitate the ss DNA. After vortexing, the reaction mixture was placed on ice for 15 minutes, centrifuged at 14K for 5 minutes at 4° C, and the pellet was resuspended in 400 µL of low TE. 3 phenol/chloroform and 1 chloroform extractions are performed to remove the residual protein and RNA. 50 µL of 5 M NH_4OAc pH 4.8 and 1mL of 95% ethanol were added to precipitate the DNA. The mixture was

allowed to sit at -80° C for 15 minutes, then was centrifuged at 14K for 10 minutes at 4° C. The pellet was washed with 500 µL 70% ethanol and again centrifuged at 14K for 10 minutes at 4° C (step 2). The final concentration of the uracil-containing ss DNA template was approximately 0.1 pmoles/µL.

Oligonucleotide primers were designed to introduce the mutations (step 3). The primers were analyzed with the program Oligo (v5.0 Rychlik, 1996) for possible problems involving hair-pin and duplex formation, false priming sites, and introduction of additional restriction sites. The melting temperatures of the primers were analyzed to insure their stability during the annealing reactions. The primer for the first mutant, Ile 58 → Thr, was 21 nucleotides long and involved a one base pair change from wild-type (noted in bold). The second mutant, Gln 105 → Met, used a primer 27 nucleotides long, incorporating 3 base changes to cause the mutation. Both primers were synthesized by Macromolecular Resources (Colorado State University, Fort Collins, CO).

Primer Ile 58 → Thr: ATC TTT TGT A**G**T TAC ACC ATT

Wild-type strand: ATC TTT TGT AAT TAC ACC ATT

Primer Gln 105 → Met: GGT TTC TCC CAT **CAT** GAA AAC CAT ATT

Wild-type strand: GGT TTC TCC CAT TTG GAA AAC CAT ATT

The oligonucleotide primers were phosphorylated on the 3' hydroxyl terminus to permit the elongation reaction (step 4). The mixture consisted of 10 µL of 20 pmol/µL oligo, 3 µL of 10X polynucleotide kinase buffer (50 mM Tris-HCl, 1mM DTT, 0.1 mM EDTA, 1 µM ATP, 50% glycerol, pH 7.5), 1 µL 10mM ATP, 15 µL ddH_2O, and 1 µL polynucleotide kinase. The reaction was mixed gently and incubated for 2 hours at 37° C, and stored at -20° C.

The phosphorylated oligonucleotide was annealed to the uracil-containing, ss M13 template strand by combining 1 μL template DNA, 1 μL oligo, 1 μL annealing sequencing buffer (0.2 M Tris HCl pH 7.5, 0.1 M $MgCl_2$, 0.5 M NaCl) and 7 μL ddH_2O (step 5). The reaction mixture was heated to 98° C for 3 minutes then allowed to cool overnight to room temperature.

The oligo was elongated along the template strand by adding 15 μL of 2X elongation buffer (40 mM Tris-HCl, 4 mM DTT, 20 mM $MgCl_2$, 1 mM of each dNTP, 2 mM ATP) to above annealing reaction tubes with 1 μL T4 DNA polymerase and 1 μL T4 DNA ligase, cooled on ice for 5 minutes and incubated at 37° C for 2 hours (step 6).

Glycerol stocks of calcium competent cells were produced by inoculating a 100 mL culture of 2 X YT media with 0.5 mL of JM101 overnight and incubating at 37° C for 2 hours to an $O.D._{595}$ of 0.3. Sequential 5K centrifugations for 5 minutes at 4° C of the cultures were separated by resuspension in 25 mL of 0.1 M $MgCl_2$ and cooling on ice for 15 minutes. After the second centrifugation, 2 mL of 0.1 M $CaCl_2$ were added and the cells were cooled for 1 hour on ice. 200 μL of cells were combined with 100 μL 50% glycerol, vortexed briefly, then snap frozen in liquid nitrogen and stored at -80° C.

Elongation reaction mixtures were transformed into calcium competent cells of *E.coli* strain JM101 by adding 1 μL of elongation mixture to 50 μL of bacterial cell glycerol stock (step 7). The transformation reactions were placed on ice for 30 minutes before being heat shocked at 42° C for 1 minute, to allow the foreign DNA to enter the cells. 51 μL transformed JM101 cells were plated on 2 X YT agar plates. 200 μL JM101 overnight culture were added on top of the agar and 3.5 mL top YT agar were added and swirled to mix. The cultures were incubated at 37° C overnight.

Plaques indicating the location of phage containing the M13 template DNA were visually identified, then picked with a pipette tip, added directly to 1 mL of low TE and vortexed vigorously (step 8). 30 μL of this solution was used to inoculate a 3 mL culture of

2X YT media with 15 μL of overnight culture of JM101. Cultures were incubated at 37° C for 6-8 hours with agitation.

The isolation of ss M13 DNA was performed as above except that the pellet of the first centrifugation was saved as double-stranded (ds) DNA and 300 μL of the supernatant was decanted and saved as infective phage stock.

Single-stranded DNA sequencing was performed using the USB Sequenase protocol which relies upon incorporating radioactive S^{35}-labeled chain-terminating ddNTPs into the growing DNA polymer, followed by analysis of the products on a polyacrylamide gel.

The previously stored ds DNA pellet of the confirmed mutant was prepared for enzymatic digest by using the Qiagen miniprep DNA purification kit. Pre-prepared P1, P2, and N3 buffers were used to resuspend the cell pellet, lyse the bacterial cell walls and degrade the RNA and protein, and neutralize the reaction, respectively. After each buffer addition, the tube was inverted 4-6 times to mix. Buffers were added in the following order: 250 μL of P1 (Tris, RNase, and EDTA); 250 μL of P2 (SDS and NaOH); 350 μL of N3. The mixture was centrifuged at 14K for 10 minutes at 4° C to separate cell particulate from DNA in the supernatant solution. The supernatant was removed and added to a Qiagen column suspended above a collection tube and centrifuged for 1 minute at 14K at 4° C to specifically bind ds DNA to separate it from the rest of the lysate. The collection tube was drained after each centrifugation and 0.5 mL buffer PB was added to the column to remove trace nuclease activity which would degrade the DNA. 0.75 mL buffer PE was added as a final wash and centrifuged. The columns were centrifuged again to remove final traces of buffers and ethanol that would inhibit enzymatic activity, before 50 μL of 70° C water was added to elute the DNA from the column. The water was allowed to sit on the column for 2 minutes before centrifugation. Concentration of ds DNA after purification was determined by absorption at 260 nm using an extinction coefficient of 50 μg/mL/OD. Final concentration ranged between 0.3 mg/mL and 1 mg/mL.

The pure ds DNA was digested by restriction endonucleases BamH I and HinD III (step 9). These enzymes recognize and cleave DNA at either side of the gene of interest within the M13 plasmid, thereby allowing the isolation of the DNA fragment containing the mutagenized gene from the host plasmid. 1 μL of each ice-cold enzyme was added to 16 μL of ds DNA and 2 μL of 10X NEB Buffer for BamH I (150 mM NaCl, 10 mM $MgCl_2$, 10 mM Tris-HCl, 1mM DTT, pH 7.9) and 0.1 μL of 100X (100 μg/mL) bovine serum albumin. Reactions were incubated at 37° C for 2 hours.

Upon completion of the restriction digest the DNA fragment was purified via low-melt agarose gel electrophoresis (step 10). 0.4 mg FMC BioProducts SeaPlaque GTG agarose was added to 40 mL TAE electrophoresis buffer (40 mM Tris acetate, 1 mM EDTA) and heated to melt the agarose. Approximately 8-10 μL of restriction digest mixture ds DNA were added to 2-3 μL of agarose gel loading buffer (30% glycerol and 0.25% tracking dyes-bromophenol blue and xylene cyanol). This mixture was typically analyzed against a 1 kBase DNA standard and uncut ds DNA for size comparison. Gels were typically run at 70 V for 30-45 minutes. The gel was stained with ethidium bromide to exposed to UV radiation to visualize the DNA fragments. The fragment containing the mutation was cut from the gel and stored at -20° C.

The pHS 1403 expression vector was prepared by inoculating 3 mL of LB containing 0.5 mg/mL ampicillin with 20 μL of JM101 glycerol stock containing the pHS 1403 plasmid. The culture was grown for 6-8 hours and then purified with Qiagen miniprep and enzymatically digested with BamH I and HinD III to prepare the vector for the addition of the mutagenized insert (step 11). The digested pHS 1403 was gel purified as described previously.

Both the gel-purified expression vector and the mutagenized insert were heated at 70° C for 15 minutes to liquefy the gel. At the same time 7 μL of water was added to a new tube and equilibrated to 37° C. 5 μL each of the vector and insert were added to the water in combination with 2 μL of 10X T4 DNA ligase buffer (20 mM Tris-HCl, 1 mM EDTA, 5

mM dithioerythitol, 60 mM KCl, 50% glycerol, pH 7.5) and 1 μL ice-cold T4 DNA Ligase (Boehrigner Mannheim). Ligation mixtures were incubated at 16° C overnight (step 12).

The ligation mixture was transformed into calcium competent cells of *E.coli* strain RR1 by adding 5 μL of ligation mix to 100 μL of cells (step 13). The mixture was placed on ice for 30 minutes then heat shocked at 42° C for 2 minutes and rescued with 900 μL of SOC media and agitation at 37° C for 1 hour. Mixture was plated on an LB + 4X ampicillin culture plate and grown overnight. Because the pHS 1403 expression vector confers ampicillin resistance, only bacteria which contain the plasmid of interest can grow in the presence of ampicillin.

Colonies were isolated and glycerol stocks were made from cultures grown by picking individual colonies. In addition, ds DNA was prepared for a restriction endonuclease digest to confirm successful ligation. Agarose gel electrophoresis was performed as described previously. Ligated inserts had the same lengths as the mutagenized restriction fragment from M13 previously. DNA was submitted to the DNA Sequencing Facility to confirm the identity of the mutants.

Once bacterial cells carrying the mutation-containing ds DNA plasmids were identified, they were allowed to grow and induced to produce protein.

Protein expression and purification- A 100 mL LB + 0.10 mg/mL ampicillin overnight culture was grown from glycerol stocks of the mutants to inoculate 3 L LB with 100 μg/mL ampicillin in a fermenter. The growth proceeded at 37° C, 750 rpm rotor speed, 15 psi until an OD_{600} of 0.9-1.0 was reached. 250 mg/L IPTG was used to induce protein expression for 2 hours while rotor speed and pressure were reduced to 225 rpm and 7.5 psi, respectively. Lysed cells were thinned using 20 mL of 0.5 M EDTA for 15 minutes with agitation. In addition, 20 mL of 1 M $MgCl_2$ and 5-10 mg of DNase I were used to complete the lysing of the cells and to degrade the DNA.

Cultures were centrifuged at 10K for 35 minutes and each liter was dialyzed against 8 liters to a conductivity of less than 4 siemens. The dialyzed product was loaded on to a CM Sepharose CL weak cation exchange column by gravity. 400 mL of 50 mM NaCl and 400 mL of 300 mM NaCl in 50 mM Tris-HCl, 1 mM EDTA, pH 7.3 buffers were used to make a gradient to elute the protein from the column. Collected fractions were dialyzed against water for 1 hour and then concentrated using an SP Sephadex column and eluted with $NaPO_4$, NaCl, and NaN_3.

At the conclusion of this process I received approximately 100 mg of each of the mutant proteins at high purity. From the DNA sequencing it was clear that I had performed the effective truncation of the long, apolar isoleucine 58 to the short, polar threonine and the replacement of the hydrogen-bonding glutamine 105 to apolar methionine. I submitted these proteins for thermodynamic analysis. I also attempted to obtain protein crystals to analyze the structural effects of these mutations.

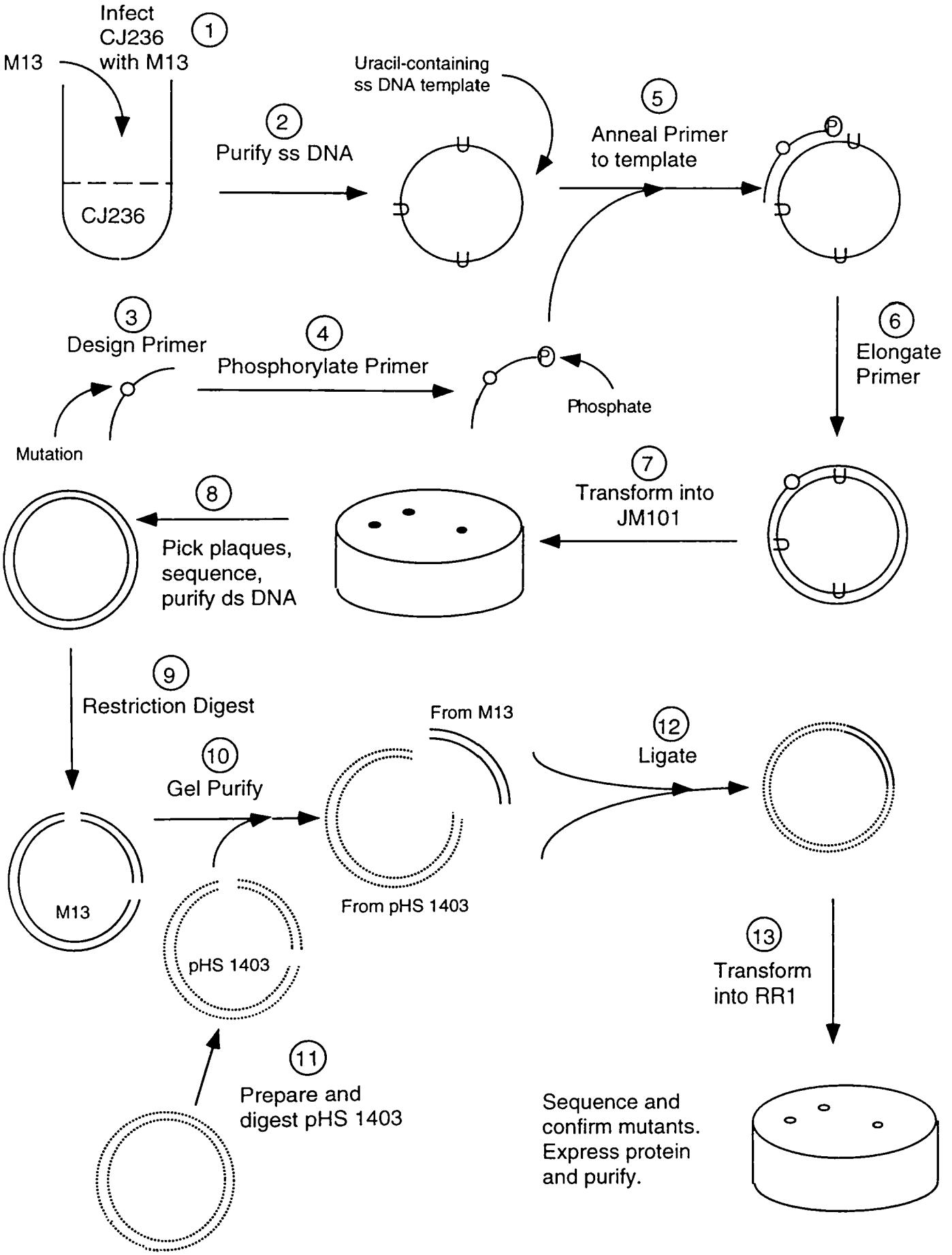

Figure 2: Site-directed mutagenesis.

X-ray crystallography

X-ray crystallography allows the construction of a atomic model based on the diffraction of x-rays from a crystal of the molecule. The model can then be examined to see the structural manifestations of the mutation (see Figure 3). X-ray crystallography requires that molecules of interest come together in an ordered fashion to form crystals. Protein crystals are formed by combining a concentrated solution of protein with a precipitant (a chemical that makes the protein aggregate, sometimes to form a crystal) on a glass cover slip, then suspending the drop above a well containing only precipitant. Water leaves the drop on the cover slip through diffusion and gradually concentrates the protein and the precipitant in the drop. Previous studies have determined the optimum conditions of pH and precipitant concentration for crystallization of lysozyme. Crystals grow to full size within one or two weeks and can reach a millimeter on a side. When the crystals reach an appropriate size they can be mounted in a quartz capillary tube, then exposed to a beam of x-rays. The electrons in a protein crystal interact with the x-rays to form a diffraction pattern that can be collected on an x-ray sensitive detector. Each diffracted ray has a certain amplitude. By combining this amplitude information from the diffraction pattern with the model of wild-type lysozyme, structural differences can be analyzed between mutant and wild-type proteins. These differences can be modeled as changes in amino acid conformation or as rearrangements of water molecules using a molecular graphics program. The revised model is then iteratively fit to the measured amplitudes with the constraint that it satisfies the requirements of appropriate chemical geometry.

Crystallization conditions were initially screened for in 1.8 - 2.2 M Phosphate at pH 6.3 - 7.24 with 50 mM β-mercaptoethanol and 50 mM hydroxyethyl disulfide (HEDS). 5 µL of 10 - 20 mg/mL protein were mixed with 5 µL of well solution and placed on a silanized glass cover slip using the hanging drop method of vapor diffusion crystallization. Optimization of crystallization conditions followed initial trials. Final conditions for Ile

58→Thr were pH 7.1, 2.1 M phosphate, 50 mM HEDS, 10 μL of 13 mg/mL protein, combined with 10 μL of well solution. After 10 days crystals measured 0.4 mm x 0.4 mm x 0.1 mm. Final conditions for Gln 105 → Met were pH 6.7, 1.85 M phosphate, 50 mM HEDS, 5 μL of 9.7 mg/mL protein, combined with 5 μL of well solution. After 13 days these crystals measured 0.2 mm x 0.2 mm x 0.03 mm. A quartz capillary tube was filled with well solution, except for the top 0.75 inches. Crystals were harvested from the drop on the cover slip and inserted into the end of the capillary. Because lysozyme crystals are less dense than the well solution, inverting the capillary allows the crystal to float to the meniscus. Subsequently, the excess well solution was withdrawn and the capillary cut at 1.5 inches for positioning on the x-ray goniometer head. In addition, the excess liquid was wicked away, leaving only the crystal. Capillaries were sealed with a small plug of well solution and then reinforced with epoxy. Capillaries were mounted on the goniometer and visually aligned in the x-ray path. A 60-second x-ray exposure was taken to check quality of diffraction. High and low-resolution detectors were moved to optimally collect data. Orientation data was collected in three 2-degree sweeps to allow alignment of the crystal axis and determination of unit cell parameters. After determining optimal diffraction strategies, 15 sweeps were performed with an oscillation range of 0.15° and 0.2°, for Ile 58 → Thr, and for Gln 105 → Met, respectively.

Diffraction data from each of the 15 sweeps were merged and scaled, and outlying data rejected in an iterative process. The intensities determined through processing were converted to structure factor amplitudes, $|F_o|$, and combined with phase information and amplitude, $|F_c|$, from the wild-type* (Cys 54 → Thr, Cys 97 → Ala, PDB code 1L63, referred to as wild-type) model to make initial $|F_o-F_c|$ and $|2F_o-F_c|$ electron density maps.

After the construction of the initial electron density map, TNT (Tronrud, 1987) was used to perform structural refinement. The structure was refined in several stages. Steepest

descent rigid body refinement of the whole structure was followed by refining the C-terminal, central helix, and N-terminal domains. Next, secondary structural elements were refined as separate rigid bodies and B-factors were combined. Subsequently, 240 rounds of conjugate direction refining occurred with manual rebuilding every 50 rounds using Chain (Pflugrath, 1984).

Analysis of cavities in wild-type and mutant lysozymes was performed using MSROLL (Connolly, 1989) with a probe radius of 1.4 Å. Surface points were visualized as cavity surfaces in Chain and graphic representations of these structures were made in Molscript (Kraulis, 1997). In addition, putative water positions at optimal hydrogen bonding distances were analyzed for contacts in Chain.

This process allowed me to construct a model of the mutant proteins. In addition, I could examine the volume of the cavities within them, and look for the addition or removal of water molecules. By placing water molecules in the model at locations were they would normally be expected, I could examine the reasons for their exclusion.

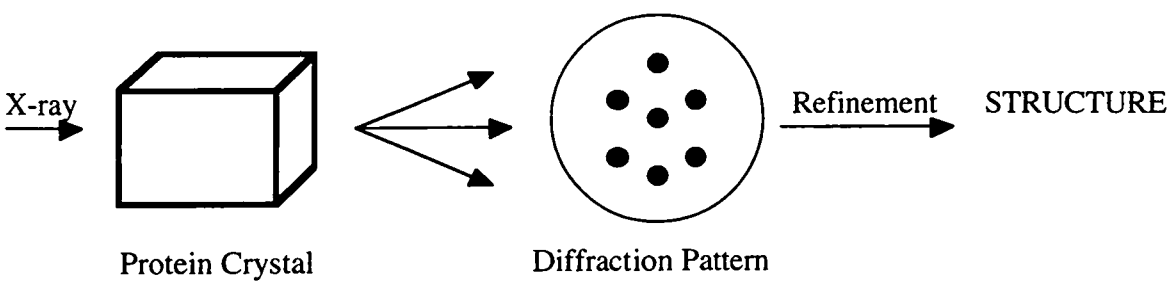

Figure 3: Representation of basic steps in protein x-ray crystallography.

Results

Thermodynamic Data

The mutant proteins in this experiment were less stable than the wild-type protein. In any reaction, A → B, which has reached equilibrium, there is a ratio which describes how many molecules of A and B are present. The equilibrium ratio is a reflection of the intrinsic stability of each of these compounds. The more stable a compound, the more prevalent it is. Proteins behave similarly. One way to think about their stability is to consider the balance between unfolded (compound A) and folded proteins (compound B) in a solution. The higher its stability, the more the balance shifts towards the folded form. The thermodynamic formula, $\Delta G = -RT \ln K$, relates the stability change to a change in the ratio of these forms. The formula explains that a 1.4 kcal destabilization ($\Delta\Delta G$ = -1.4 kcal/mol) results in a 10-fold shift of the balance to the unfolded form. The first mutant, Ile 58 → Thr, lost 3.24 kcal/mol of stabilization. Therefore, at a given temperature (300 K), unfolded form of the mutant protein is present over 200 times more than the unfolded form of the wild-type protein. For Gln 105 → Met, only a 1.0 kcal/mol change was measured, meaning only a 5-fold change in this ratio.

To assess the stability of the mutant lysozymes, our collaborator Dr. Walter Baase performed reversible thermal denaturation as determined by CD spectroscopy in 20 mM KPO_4, 25 mM KCl, pH 3. As desribed above, the Ile 58 → Thr mutation resulted in a $\Delta\Delta G$ of -3.24 kcal/mol compared to wild-type, while the Gln 105 → Met exhibited a $\Delta\Delta G$ of -1.0 kcal/mol compared to wild-type.

Structures of Mutant Lysozymes

Data collection and structure refinement statistics for both mutants are presented in Table 1. The structures of the proteins were both globally very similar to the wild-type protein, but locally, quite different. The examination of the structures on a finer scale, particularly near the sites of the mutations revealed these changes. The first mutant, Ile 58 → Thr, was designed to increase the volume of the nearby water-filled cavity and introduce a new water molecule in the additional space. The cavity, however, was not enlarged, and the water molecule did not appear. Instead, the rest of the protein moved toward residue 58, filling the additional volume introduced by the mutation.

In the second mutant structure, Gln 105 → Met, the water molecule near the site of mutation was no longer present. Moreover, the absence of water created an unfilled cavity in the interior of the protein which was enlarged by the shift of a nearby amino acid. The mutant protein did not collapse and refill the cavity as the previous mutant did. A detailed explanation of the results follows.

Global conformation of Ile 58 → Thr - The global protein conformation of Ile 58 → Thr remained very similar to the wild-type structure. The root-mean-square shift of the main chain atoms between mutant and wild-type structures was 0.13 Å. Figure 4 shows the differences in main chain atom positions. An examination of this plot reveals the local disturbances surrounding the site of mutation. Among the largest main chain movements are Ala 63 Cα (0.39 Å) and Val 57 C (0.34 Å). Overall, however, these changes are rather small.

	Ile 58 → Thr	Gln 105 → Met
Cell Parameters		
Space group	P3221	P3221
a = b	60.93	61.12
c	96.80	96.76
Data collection		
Resolution	1.9 Å	2.2 Å
No. of measured reflections	51962	39357
No. of unique reflections	16299	10738
Data completeness (%)	96	97
Rmerge	0.058	0.1053
Refinement statistics		
Rcryst	0.163	0.162
RMS deviations		
Bonds (Å)	0.021	0.020
Angles (degrees)	2.865	2.961
Trigonal Planes (Å)	0.016	0.016
General Planes (Å)	0.016	0.015

Table 1: X-ray data collection and structure refinement statistics for the mutants, Ile 58 → Thr and Gln 105 → Met. This table shows the agreement between the observed data and the refined model.

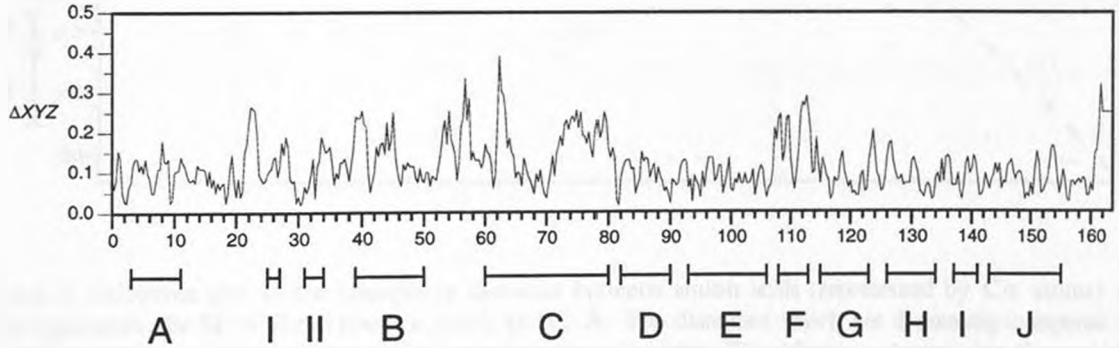

Figure 4: Main chain atom shifts (Å) between wild-type and the mutant, Ile 58 → Thr. This figure indicates that there were not many large movements of the protein backbone atoms.

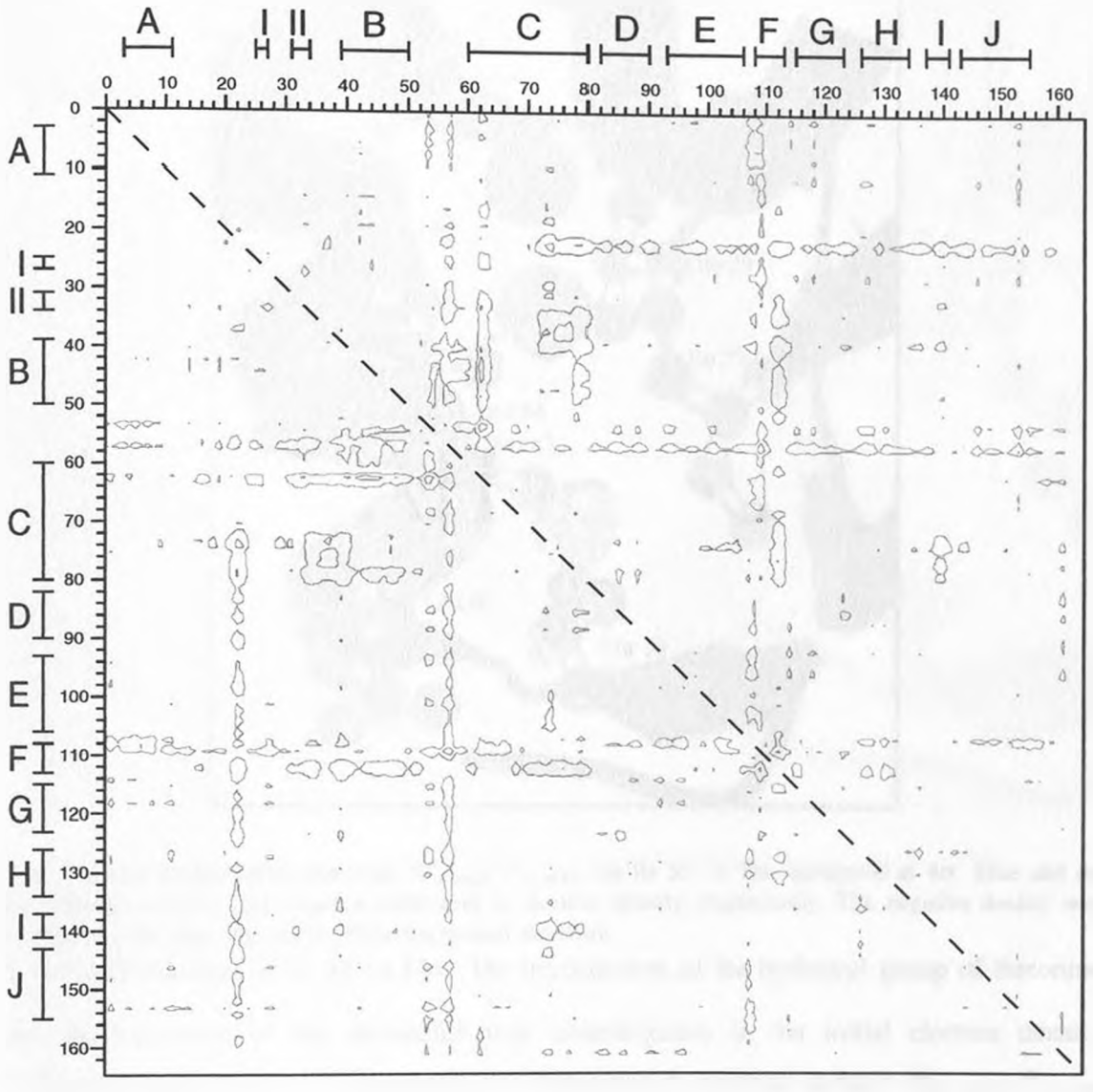

Figure 5: Difference plot in the changes in distances between amino acids (represented by Cα atoms) in wild-type *versus* Ile 58 → Thr at contour levels of 0.2 Å. The distances which are decreasing compared to wild-type are shown in red. Those which are increasing are blue. This figure indicates that the protein contracts toward residue 58 in the mutant structure.

The trend of these changes is represented in Figure 5, a difference plot of inter Cα distances. The red band between nearly all residues and residue 58 clearly indicates that the global conformation of the mutant adjusts to fill the cavity created by the truncation of Ile 58 → Thr.

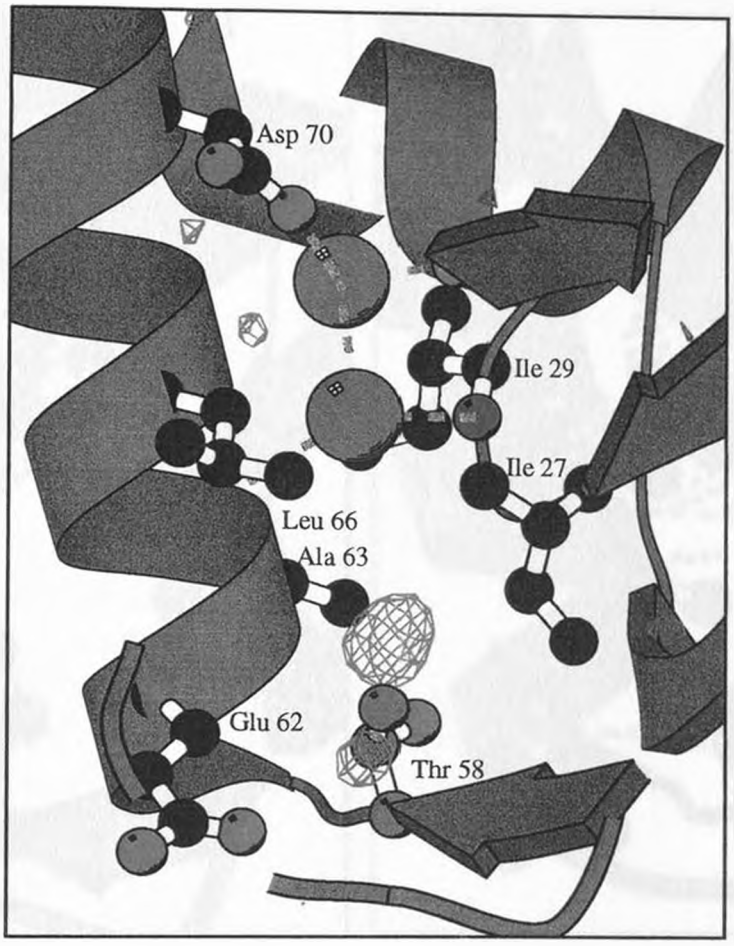

Figure 6: Electron density difference map, $|F_{O,\ MUT} - F_{O,\ WT}|$, for Ile 58 → Thr, contoured at 4σ. Blue and red contours indicate positive and negative differences in electron density, respectively. The negative density near Thr 58 indicates the truncation of Ile 58 in the mutant structure.

Local conformation of Ile 58 → Thr - The introduction of the hydroxyl group of threonine and the truncation of the isoleucine was unambiguous in the initial electron density difference maps. Figure 6 represents the difference in electron density $|F_{O,\ MUT} - F_{O,\ WT}|$ contoured at 4σ. In addition, the *gauche* (g+) conformation of the isoleucine is retained in the threonine mutant. No additional water molecules entered the volume near the mutation. In addition, there were relatively insignificant differences between the volumes of the hydrated cavity near the site of mutation (see Figures 7 and 8). Although the structure of the protein was globally conserved, the conformation near the mutation was significantly altered. Table 2 shows that 6 atoms within 4 Å of Ile 58 moved 0.4 Å or greater. The largest of these local changes was Ile 27 Cδ1 which shifted 0.88 Å into the volume near 58.

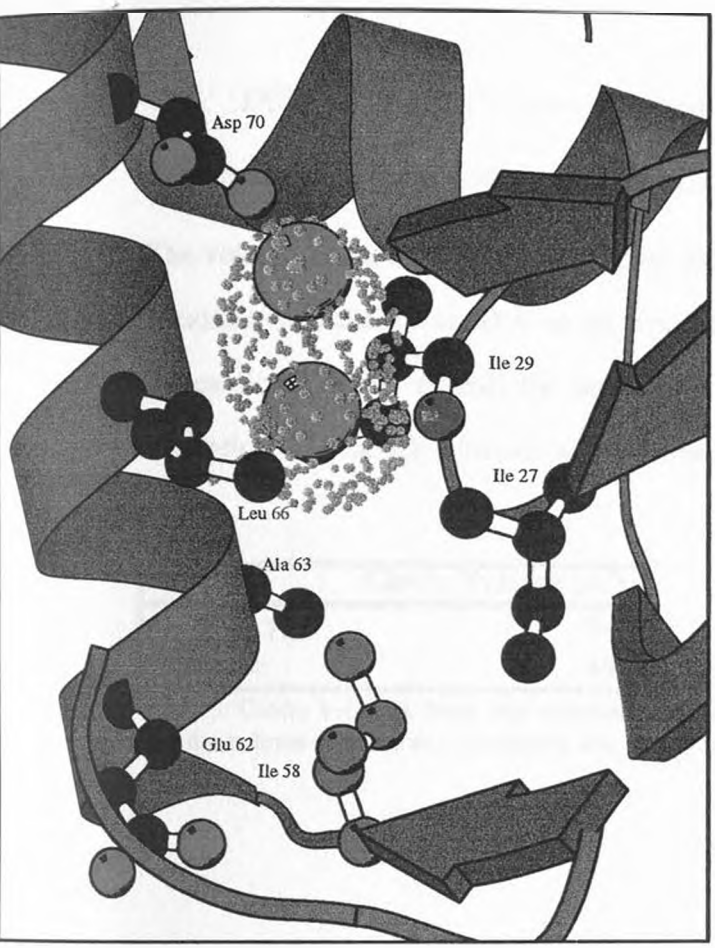

Figure 7: Wild-type lysozyme near residue Ile 58.

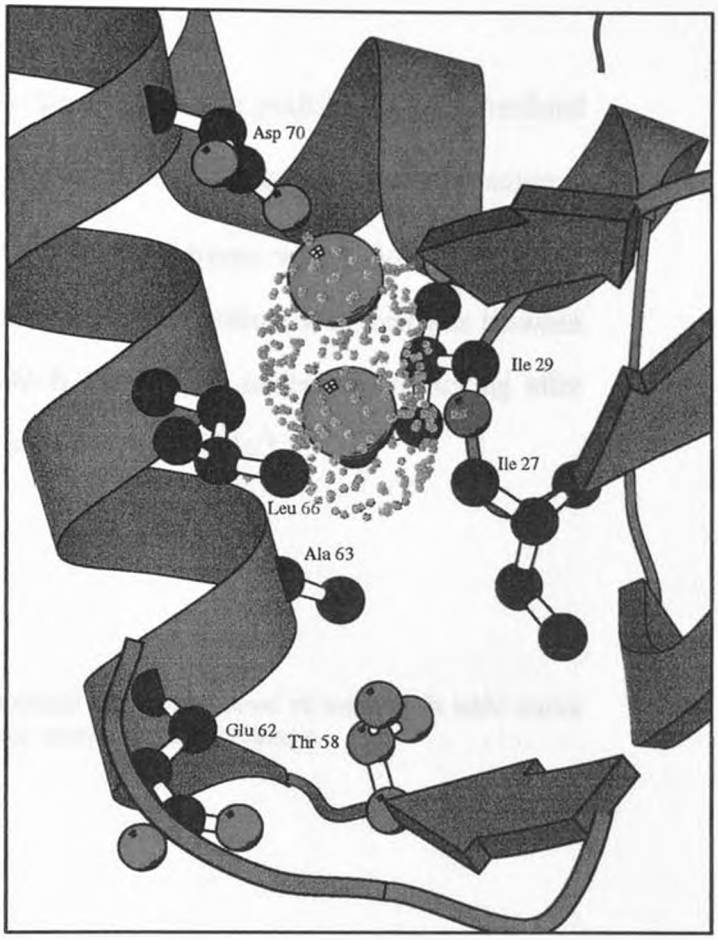

Figure 8: Mutant lysozyme, Ile 58 → Thr.

Atoms within 4 Å of Ile 58	Distance (Å)
Leu 15 CD2	0.45
Leu 15 O	0.16
Ile 27 CG1	0.79
Ile 27 CD1	0.88
Leu 46 CD2	0.47
Ile 50 CG1	0.28
Val 57 C	0.34
Val 57 O	0.61
Val 57 CA	0.24
Thr 59 0	0.16
Thr 59 CA	0.13
Thr 59 N	0.15
Glu 62 CB	0.03
Glu 62 OE1	0.17
Ala 63 N	0.22
Ala 63 CB	0.57
Ala 63 CA	0.39

Table 2: Shifts of atoms within 4 Å of Ile 58 between wild-type and the Ile 58 → Thr mutant.

Differences in the volume occupied by the side chain at position 58 were modeled by truncating the residue 58 side chains at the Cα atom in wild-type and mutant structures. The volume occupied by the isoleucine side chain in wild-type was 116 $Å^3$ and by the threonine side chain was 80 $Å^3$ in the mutant structure. The difference in volumes between the cavities, 36 $Å^3$, reveals the degree to which the mutant underwent repacking after truncation. The cavity volumes without water are shown in Table 3.

Structure	Cavity Volume ($Å^3$)
Ile 58 → Thr	42.3
Wild-type	44.3

Table 3: Cavity volumes. Note that volumes were calculated with the removal of water. This table shows that the volume of the cavity containing the nearby water molecules did not change

Steric Exclusion of hydrogen-bonding partners to Thr 58- The Ile → Thr 58 mutation brought several local residues closer to Thr 58. Figure 9 shows the results of modeling a water at allowed positions (2.8 Å from the hydroxyl group of threonine at a bond angle of 109°).

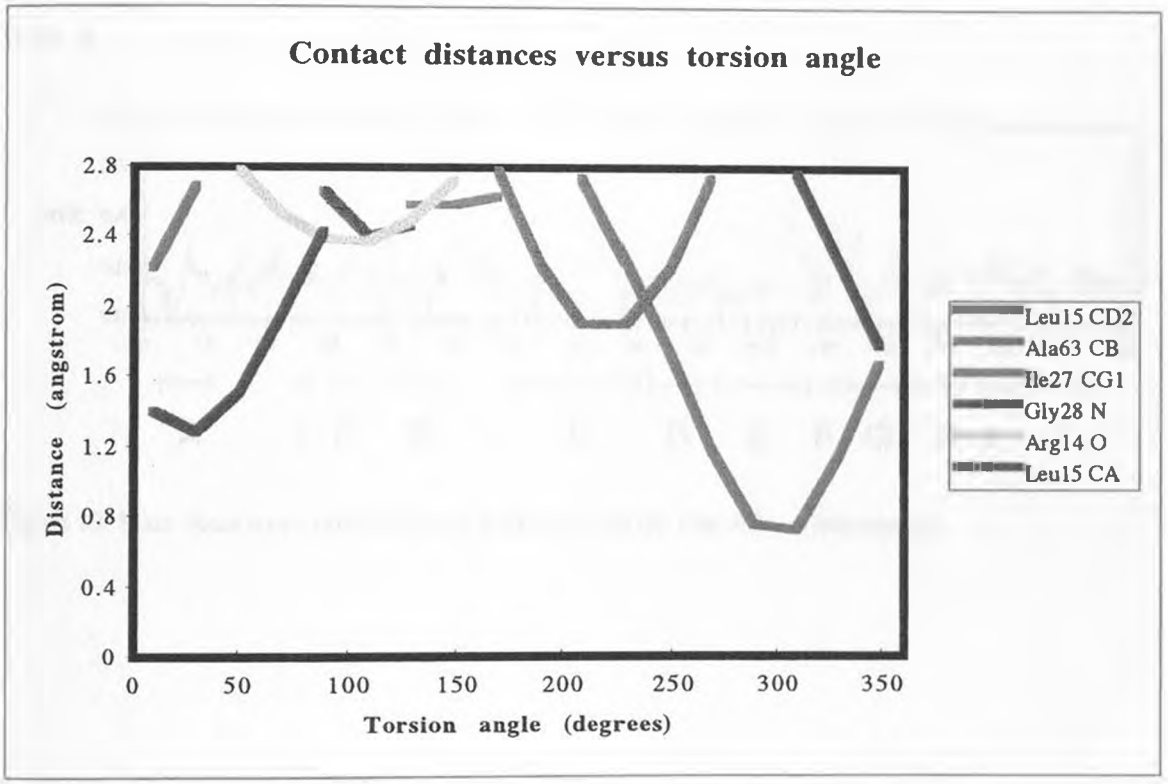

Figure 9: Graph of contact distances between the protein and a modeled water at different torsion angles. This figure shows that close contacts prevent water from occupying this region.

This analysis revealed close contacts at each of the preferred torsion angles of 60°- Leu 15 Cα, 1.5 Å; 180°- Leu 15 Cδ2, 0.8 Å; 300°- Ala 63 Cβ, 2.4 Å. In addition, at nonideal torsion angles, apolar residues were the closest contacts, all under 2.6 Å, except Arg 14 O between ~90° and 115° at 2.4 Å.

Global conformation of Gln 105 → Met - Similar to the first mutant, Gln 105 → Met was a globally conservative mutation. The root-mean-square shift between main chain atoms of wild-type and mutant structures was 0.19 Å. Figure 10 shows the main chain atom shifts between the mutant and wild-type. Most significant is the movement of Met 106 Cα by 0.48 Å.

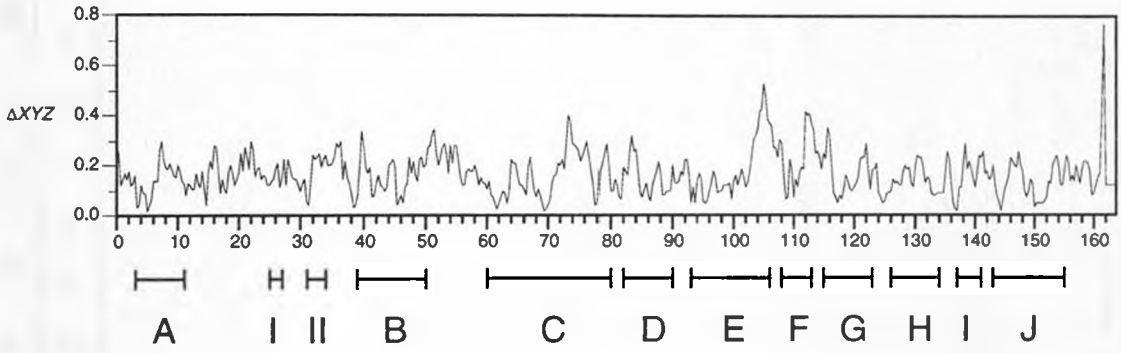

Figure 10: Main chain atom shifts between wild-type and the Gln 105 → Met mutant.

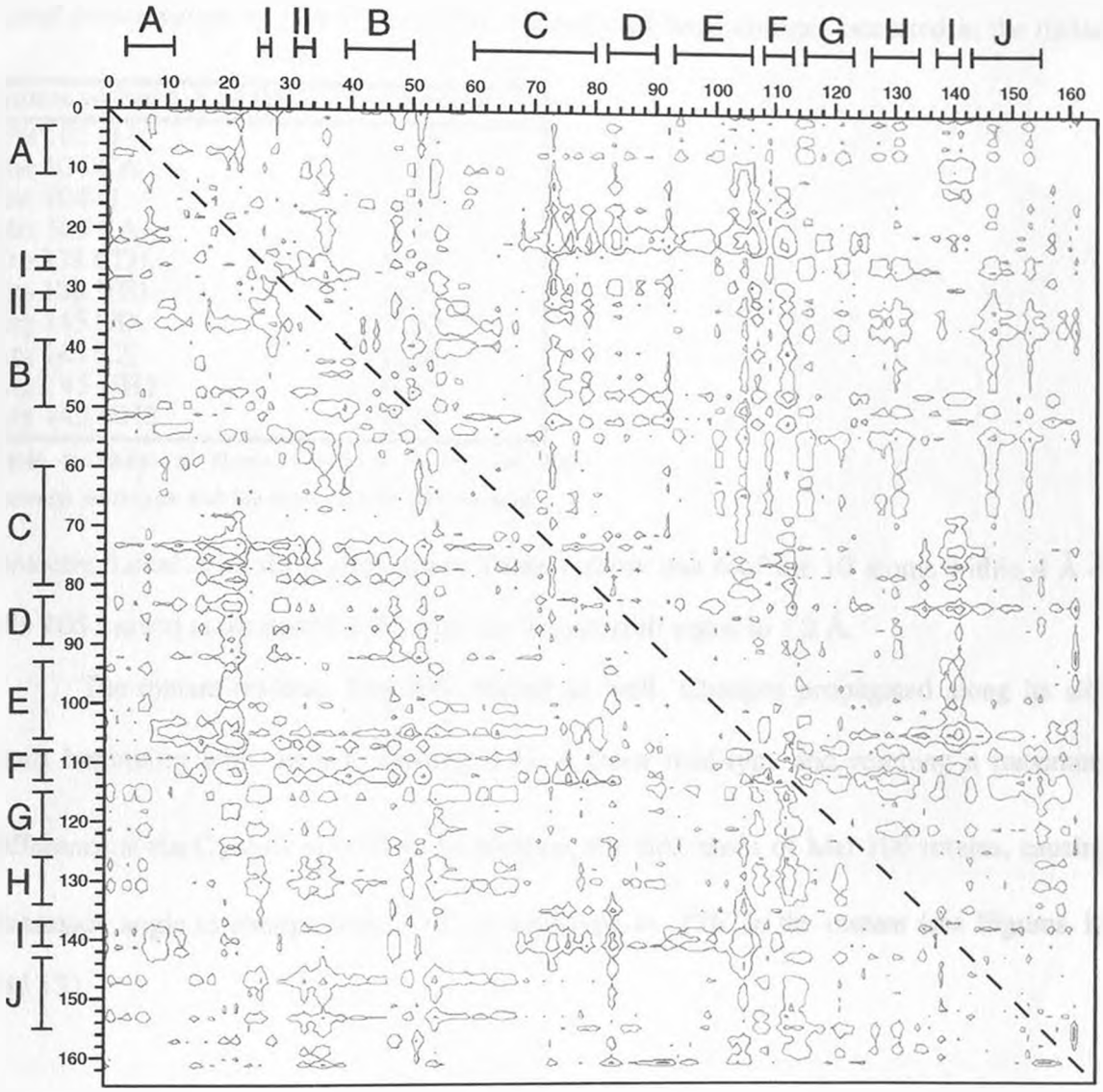

Figure 11: Difference plot of inter Cα distances between wild-type and the mutant, Gln 105→ Met at contour levels of 0.2 Å. The distances which are decreasing compared to wild-type are shown in red. Those which are increasing are blue. The absence of solitary bands indicates wide-spread movement in response to the mutation.

A plot of the inter Cα distances (Figure 11) shows larger and less concerted changes compared to Ile 58 → Thr.

Local conformation of Gln 105 → Met - Significant local changes occurred in the mutant

Atoms within 4 A of Gln 105	Distance (A)
Met 102 O	0.21
Phe 104 CA	0.31
Phe 104 O	0.25
Met 106 CA	0.48
Trp 138 CD1	0.29
Trp 138 NE1	0.23
Arg 145 CD	1.20
Arg 145 CZ	0.36
Arg 145 NH1	0.41
Arg 145 NH2	0.29

Table 4: Shifts of atoms within 4 Å of Gln 105 between wild-type and the mutant, Gln 105 → Met.

structure. Local atom shifts depicted in Table 4 show that 6 of the 10 atoms within 4 Å of Gln 105 moved more than 0.2 Å, with the largest shift equal to 1.2 Å.

The mutant residue, Met 105, shifed as well. Changes propagated along its side chain beginning with the Cα shifting 0.42 Å from wild-type and reaching a maximum difference at the Cγ shift of 0.68 Å. In addition, the side chain of Met 106 rotates, causing its torsion angle to change from 112° in wild-type to -176° in the mutant (see Figures 12 and 13).

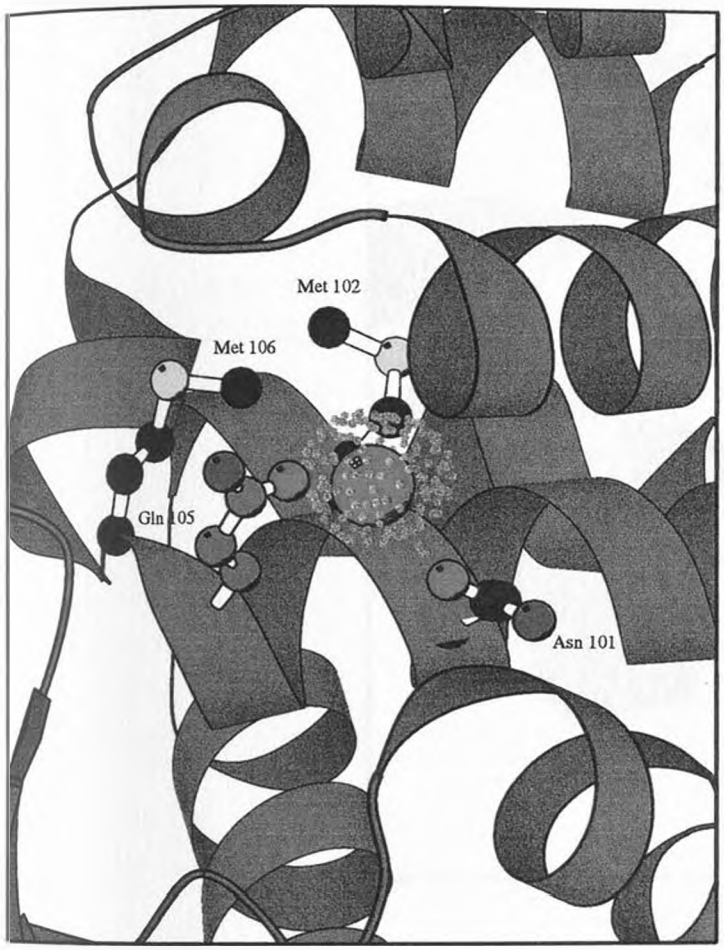

Figure 12: The structure of wild type lysozyme near residue Gln 105.

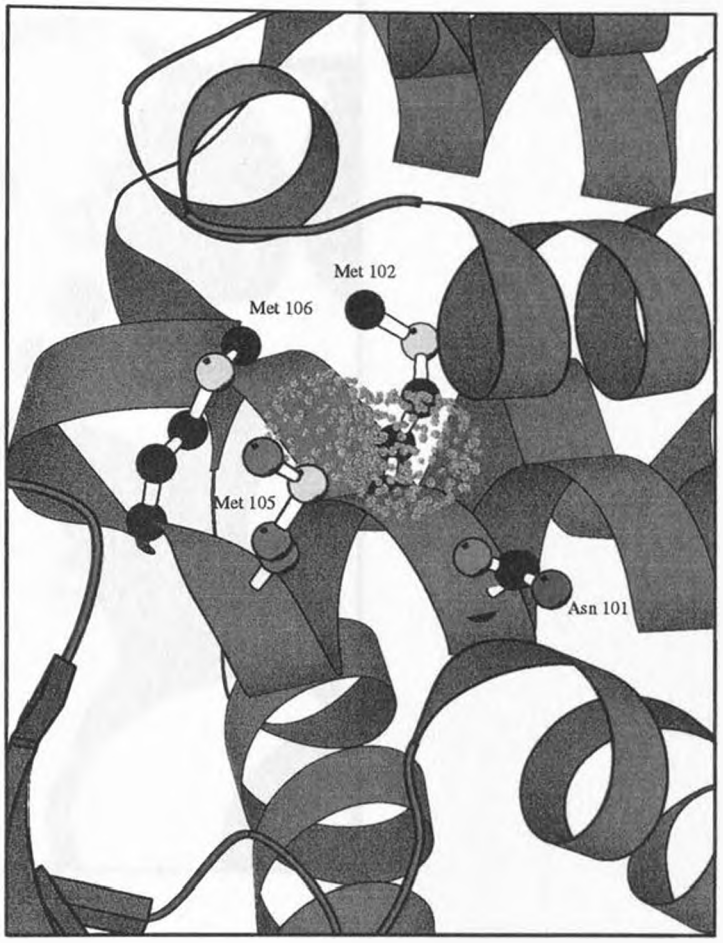

Figure 13: The structure of the Gln 105 → Met mutant lysozyme.

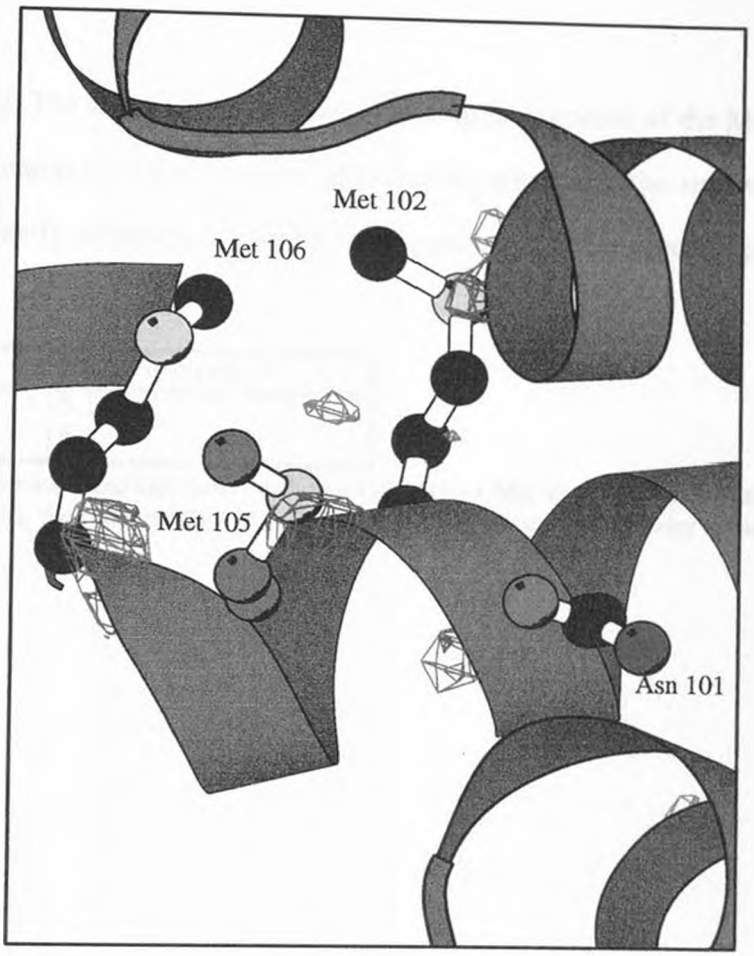

Figure 14: Electron density difference map, $|F_{O, MUT} - F_{O, WT}|$, for Gln 105 → Met, contoured at 4σ. Blue and red contours indicate positive and negative differences in electron density, respectively. The positive density near the sulfur atom of Met 105 reflects its greater electron density relative to carbon. The negative density near Met 105 indicates the removal of the oxygen atom of Gln 105.

Formation of an Unsolvated Cavity in Gln 105 → Met - Water 208, which was present in wild-type and coordinated by Gln 105 Oε1 and Asn 101 Oδ1, was absent in the initial electron density maps of the mutant. Figure 14 represents the difference in electron density $|F_{O, MUT} - F_{O, WT}|$ contoured at 4σ. The absence of water, in combination with other side chain rearrangements creates a 35 $Å^3$ cavity. Local residues do not shift substantially, leaving a cavity. Moreover, the original hydrated cavity increases in radius slightly and extends

towards Met 106. The effect of the previously mentioned rotation of the Met 106 side chain is the principle reason for the elongation of the cavity created by the removal of water 208. Table 5 shows cavity volumes of interest in the mutant and wild-type structures.

Structure	Cavity Volume (A^3)
Gln 105 → Met	35.2
Wild-type	15.1

Table 5: Cavity volumes. Note that cavity volumes Gln 105 → Met were calculated with removal of water. This figure shows that the mutant Gln 105 → Met caused an increase in the cavity volume

Discussion

Isoleucine 58 → Threonine

The loss of stability of Ile 58 → Thr can be attributed to many factors. The first consideration is the comparison of entropy, or disorder, that the different amino acids impart to the protein. While unfolded, isoleucine, the amino acid present in wild-type, has considerable ability to rotate around each of its bonds. In addition, isoleucine has the ability to order water molecules into cage-like structures called clathrates. The dissolution of these structures upon folding stabilizes the protein. Threonine, on the other hand, contains an oxygen atom which prefers to hydrogen bond with water rather than form clathrates. The unsatisfied bonds that result from the failure to introduce a water molecule destabilize the protein. These findings indicate that some areas of proteins are not as amenable to changes as others and react in different ways which are as yet, poorly understood.

Protein stability of Ile 58 →Thr- The ~3 kcal/mol loss of stability of the Ile 58 → Thr mutant is partly attributable to the difference in the $\Delta\Delta G$ of transfer, or the difference in energy required to transfer an amino acid side chain between water and a nonpolar solvent. The value for the mutation of Ile to Thr is estimated at 2.53 kcal/mol (Tanford, 1962). This calculation, however, ignores other forces involved which can be further broken down into entropic and enthalpic contributions.

Absence of local cavity in Ile 58 → Thr- The truncation of Ile 58 to Thr was designed to extend the cavity near residue 58. The structure, however, shows that the cavity was not changed. This confirms the conclusion of Erikkson, et al., by showing that some areas of proteins are relatively rigid and remain fixed in place despite the formation of a cavity, while others, like this mutant, are more flexible and respond to cavity formation by

repacking. In addition, the conformation of the local side chains undergo rearrangement sufficient to eliminate the potential of forming satisfactory hydrogen bonds with a water molecule hydrogen bonded to Thr 58. Examination of these residues identifies prospective candidates for further site-directed mutagenesis (Figure 8). While mutagenesis cannot affect the main chain atoms of Leu 15, it can be used to truncate Leu 15 to eliminate the close contacts near -60°. The next closest contact is Leu 15 Cγ which could simultaneously be removed by mutating this residue to alanine. Furthermore, Cβ of Ala 63 is the closest contact at 180° and could be mutated to glycine in further studies. It is clear why graphic modelling before the experiment was unable to predict the collapse of this cavity.

This result also verifies the diversity of environments found in proteins. Although Williams has determined that an internal water molecule in a cavity makes, on average, three hydrogen bonds, a mutant of barnase was discovered to contain an internal water molecule in a cavity of 31 $Å^3$, that made only a single hydrogen bond (Buckle, 1996). Therefore, it is not entirely unexpected that this mutation in lysozyme fails to introduce a water.

Entropic Effects of Thr 58- The apolar residue isoleucine contributes to the formation of clathrates in the unfolded form, decreasing the entropy of the unfolded form. The disruption of these clathrates by removing isoleucine from its interactions with solvent and packing it in the interior stabilizes the folded form. Isoleucine orders water molecules to a much greater extent than threonine, so the mutation of Ile → Thr decreases this entropic benefit. In addition, the truncation of Ile → Thr results in a loss of favorable van der Waals interactions between the larger isoleucine and surrounding atoms. One reason to expect a water molecule to enter the cavity is that internal waters may stabilize proteins by reestablishing the missing van der Waals interactions (Williams *et al*, 1994).

Side chain entropy must also be considered in a discussion of stability. Various values exist for the side chain entropy of threonine. Sternberg, et al., illustrate that different ways of calculating $T\Delta S$ for threonine, give vastly different results from -0.71 to -1.63 kcal/mol (1994). The values for isoleucine, on the other hand, are in reasonable agreement, ~-0.85 kcal/mol. Therefore, with a calculated entropy change of -0.74 to 0.10 kcal/mol, it is uncertain as to whether or not this mutation entropically stabilizes or destabilizes the protein relative to wild-type.

Enthalpic effects of Thr 58- There is also an enthalpic component to the free energy of transfer. Previous research suggests that the loss of solvation, i.e. the lack of a hydrogen bonding partner, of the Thr residue in the interior destabilizes the protein by nearly 3 kcal/mol (Blaber, et al. 1993). The closest atom to the hydroxyl of Thr 58 is greater than 4 Å away. Therefore, without the introduction of a hydrogen-bonding water molecule, threonine has no other partners near enough to satisfy this interaction which results in destabilization of the folded form.

Glutamine 105 → Methionine

Gln 105 → Met displaces a water molecule from the core of the protein. The cavity that forms due to this displacement destabilizes the protein by removing the attractive forces (van der Waals interactions) between that non-bonded atoms in close proximity exhibit.

Two entropic factors also stabilize the protein. First, the removal of the internal water by the Gln 105 → Met mutation adds to the stability because the water is less constrained in solution than in the cavity. Second, while methionine and glutamine have similarly free rotation, methionine in the unfolded state forms clathrates to a much greater extent than glutamine, thereby stabilizing the protein.

Unsatisfied hydrogen bonds caused by the absence of the internal water molecule in this mutant destabilize the protein enthalpically. The Gln 105 → Met mutant results in one less unsatisfied hydrogen bond than the wild-type protein. This partially accounts for why this mutant is only slightly destabilized.

Protein stability of Gln 105 → Met- The free energy of transfer between glutamine and methionine is a 1.4 kcal/mol stabilization (Tanford, 1962). As explained previously this term is the sum of enthalpic and entropic terms. As in the case of the previous mutant, the formation of an unsolvated cavity in the interior of the protein due to this mutation must be considered. While Ile 58 → Thr failed to create a cavity, the opposite consequence occurred in Gln 105 → Met, contributing to its destabilization.

Cavity formation in Gln 105 → Met- The free energy of transfer only considers the identity of the original amino acid and its replacement. In this mutant, the rotation of the side chain of Met 106 and the removal of water 208 which previously occupied the cavity are the primary contributors to the formation of an internal cavity. Side chains lining the cavity, with the exception of Met 106, do little to change its volume. Because repacking can help stabilize a protein that has internal cavities, the absence of this process indicates that this is a relatively rigid part of the protein.

Previous studies on the energetics of cavity formation have predicted a 24-33 cal/mol/Å^3 loss of stability for formation of a cavity in the apolar core of a protein (Erikkson, 1992). Because the volume of the cavity created by this mutation is 35 Å^3, the predicted instability is between 0.85 and 1.1 kcal/mol from cavity formation alone. The energetic basis for this destabilization is the loss of many favorable van der Waals interactions present in wild-type.

Entropic effects of water 208- The water molecule previously coordinated by Gln 105 is no longer present, therefore, in the mutant, the water molecule is free in solution and is not as constrained. This increase in entropy from the release of water helps to stabilize the mutant lysozyme.

Entropic effects of Met 105- The entropic cost of constraining methionine and glutamine side chains to the well-ordered interior of the protein is nearly the same. Glutamine has slightly greater rotational entropy in solution, therefore the mutation stabilizes the protein by ~0.4 kcal/mol (Sternberg, 1994). In addition, methionine orders water into clathrates to a higher degree than glutamine. The difference in the increase in entropy upon folding of the two proteins is reflected in the additional stabilization of the mutant.

Enthalpic effects of hydrogen bonding interactions- Water has the ability to participate in four hydrogen-bonding interactions. It can both accept and donate two hydrogen bonds. In the wild-type structure, water 208 only fulfills two of these interactions, therefore it lacks 2 hydrogen bonding partners. In the mutant structure, only Asn 101 remains without interactions because glutamine has been eliminated and water has been removed there is one less unsatisfied hydrogen bond. It is estimated that this change results in a stabilization of the mutant protein.

Conclusion

This research has revealed that there is considerable variability in the environments near internal water molecules. In addition, a complicated set of interactions govern protein conformation in these areas. The results of the Ile 58 → Thr mutation illustrate that the introduction of one hydrogen bonding partner and the possibility of making additional hydrogen bonds with existing side chains and other internal water molecules does not lead to the introduction of a water molecule. Indeed, it points out that the influence of internal water molecules on protein stability is a complex phenomena. While some areas of proteins are prone to collapse upon truncation, the conformation of the Gln 105 → Met mutant did not undergo repacking to compensate for cavity formation. Furthermore, the elimination of a single hydrogen bond in this mutant is sufficient to displace water despite preservation of the cavity.

While the results from this study do not permit wide-spread conclusions about protein stability to be made, they do add to the growing body of knowledge about the subject. My hope is that others may build upon this research to further elucidate the subtleties of these interactions. It is evident that the connections to more overarching concerns, like understanding the cause of disease and possible cures, is quite distant.

References

1. Alber, T., Dao-Pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., Matthews, B. W. (1987) *Nature* **330**, 41-46.

2. Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D. W., Matthews, B. W. (1993) *Biochemistry* **32**, 11363-11373.

3. Buckle, A. M. Cramer, P., Fersht, A. R. (1996) *Biochemistry* **35**, 4298-4305.

4. Connolly, M. Msroll v3.6 (1989)

5. Erikkson, A. E., Baase, W. A., Zhang, X. J., Blaber, M., Baldwin, E., Matthews, B. W. (1992) *Science* **255**, 178-183.

6. Gassner, N. C., Baase, W. A., Matthews, B. W. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 12155-12158.

7. Kraulis, P.J. Molscript v2.0.1 (1997)

8. Matthews, B. W. (1993) *Annu. Rev. Biochem.* **62**, 139-160.

9. Matthews, B. W. (1995) *Advances in Protein Chemistry* **46**, 249-278.

10. Pjura, P., McIntosh, L. P., Wozniak, J. A., Matthews, B. W. (1993). *Proteins: SFG* **15**, 401-412.

11. Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng.* **6**, 151-176.

12. Sternberg, M.J.E., Chickos, J.S. (1994) *Protein Engineering* **7:2**, 149-155.

13. Tanford, C. (1962) *J. Amer. Chem. Soc.* **84**, 4240-4247.

14. Rashin, A. A., Iofin, M. Honig, B. (1986). *Biochemistry* **25**, 3619-3625.

15. Williams, M. A. Goodfellow, J. M., Thornton, J. M. (1994) *Protein Science* **3**, 1224-1235.

16. Xu, J. Baase, W. A., Baldwin, E., Matthews, B. W. (1998) *Protein Science* **7**, 158-177.

Glossary

Amide- a chemical group consisting of a carbon with a double bond to oxygen and a single bond to nitrogen, R-$CONH_2$

Ångstrom (Å)- 10^{-10} meters

Apolar amino acid- an amino acid that consists mainly of carbon and hydrogen

Apolar core- the interior of a protein which consists mainly of apolar amino acids

Bacteriophage- a virus which infects bacteria

Hydrogen bond- a non-covalent bond from oxygen or nitrogen to hydrogen

Lysozyme- a protein which degrades bacterial cell walls

Mutagenesis- a method by which changes are introduced into a gene

Non-covalent bond- a bond in which electrons are not shared between two atoms

Polar amino acid- an amino acid that contains oxygen and/or nitrogen

Thioether- a chemical group which contains sulfur between two carbon atoms