This is to certify that the

thesis entitled

AN X-RAY CRYSTALLOGRAPHIC STUDY OF THE CHEMICAL DENATURATION OF a-CHYMOTRYPSIN

presented by

LYNDON STANLEY HIBBARD

has been accepted towards fulfillment
of the requirements for

PH.D. degree in CHEMISTRY

A. Tulinsky

Major professor

Date August 17, 1977.

O-7639

AN X-RAY CRYSTALLOGRAPHIC STUDY OF THE CHEMICAL DENATURATION OF α-CHYMOTRYPSIN

By

Lyndon Stanley Hibbard

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

1977

ABSTRACT

AN X-RAY CRYSTALLOGRAPHIC STUDY OF THE CHEMICAL DENATURATION OF α-CHYMOTRYPSIN

By

Lyndon Stanley Hibbard

Protein denaturation is a process in which the three-dimensional structure of a protein molecule is changed from that of the native conformation to a more disordered conformation. It is an aspect of the phenomenon of protein folding and has been studied for a number of proteins in solution to determine the mechanism by which the polypeptide chains of a protein are organized to form a particular tertiary structure. Urea and guanidine HCl (GdnHCl) are particularly powerful protein denaturants and it is the effects of these compounds on the three-dimensional structure of the enzyme α-chymotrypsin (CHT) that form the principal results reported in this dissertation.

Crystals of native CHT in solutions 75% saturated in ammonium sulfate at pH 3.6 were exposed to gradually-increasing concentrations of the denaturants and 2θ-scans of the axial reflections were recorded at various concentration increments and compared with those of the native protein crystals. When significant changes were observed, three-dimensional X-ray intensity data were measured to 2.8 A. resolution, and "best" difference

electron density maps[1] were calculated using the coefficients

$$m\Delta F\exp(i\alpha_N),$$

where ΔF is the difference between the derivative and the native structure amplitudes, α_N is the native phase angle, and m is a function of error in α_N (figure of merit) which is included to weight the terms of the difference Fourier.

The difference Fourier maps for two derivatives, 2.0 M. GdnHCl and 3.0 M. urea, were very complicated revealing large numbers of changes over many parts of the CHT molecule. A graphical representation of the difference peaks, the difference diagonal plot (DDP),[2] was used to characterize the locations of, and the extent of the most significant changes in the derivative maps with respect to the CHT amino acid sequence. Several GdnHCl difference map peaks near the aromatic substrate binding site created a distinctive array of DDP features similar to those observed in the DDP's of inhibitor-CHT complexes.[3] An examination of the urea DDP revealed several groups of changes which occurred in the interior of the CHT molecule.

The peaks which appeared in the GdnHCl difference map were located exclusively on the surface of the protein. A number of difference peaks appeared in the dimer interface region near the substrate binding site and represent changes in the local solvent structure with the possible binding of a guanidinium ion (Gdn^+) near sulfate ions

hydrogen bonded to the Ser 195 hydroxyl groups of both CHT molecules. A Gdn^+ was also observed in the uranyl binding site region[4,5] forming an intermolecular salt bridge between two CHT molecules.

The urea difference displayed a complex picture of changes occurring at sites in the nonpolar interior of the CHT molecule as well as on the surface. The nature and disposition of many of the peaks strongly suggests that urea molecules penetrated the interior of the protein to interact with a number of aliphatic side chains. Disruptions of the native structure were observed in the A-chain/B-chain contact region, in a segment of B-domain β-sheet near the tryptophan cluster Trp 27, Trp 29, and Trp 237, and in a segment of A-domain β-sheet near the aromatic cluster Trp 51, Phe 89, and Trp 237.

1. D. M. Blow and F. H. C. Crick, Acta Cryst., 12, 794 (1959).
2. M. N. Liebman, Ph.D. Dissertation, Michigan State University, 1977.
3. A. Tulinsky, I. Mavridis, and R. Mann, J. Biol. Chem., in the press.
4. A. Tulinsky, N. V. Mani, C. N. Morimoto, and R. L. Vandlen, Acta Cryst., B29, 1309 (1973).
5. R. L. Vandlen and A. Tulinsky, Biochemistry, 12, 4193 (1973).

To my wife,

Susan

Acknowledgements

With deep appreciation and gratitude the author wishes to thank Professor Alexander Tulinsky for his guidance, support, and encouragement.

The author is grateful to Dr. N. V. Raghavan and to Dr. Stephen R. Ernst for many stimulating discussions and helpful contributions.

To Dr. M. N. Liebman for his contribution of the difference diagonal plot and for many useful discussions, the author is grateful.

To Mr. Lawrence Weber for his contributions to the difference diagonal plot calculations, the author extends his sincere thanks.

For their financial support, the author thanks the Molecular Biology Section of the National Science Foundation and the Department of Chemistry, Michigan State University.

The author wishes to thank his wife, Susan, for her constant support and encouragement through four difficult but rewarding years.

TABLE OF CONTENTS

Table of Contents (Continued)

LIST OF TABLES

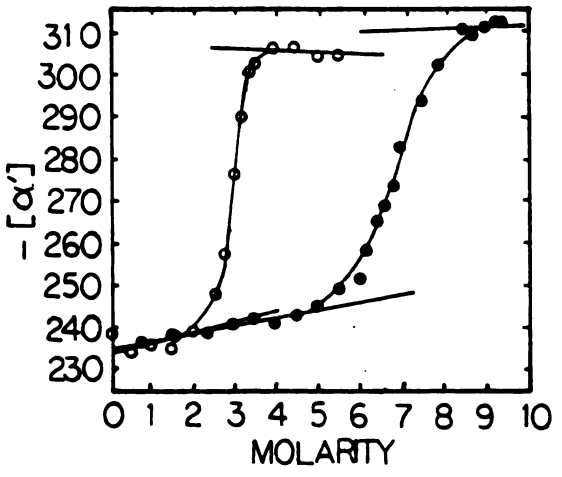
LIST OF FIGURES

List of Figures (Continued)

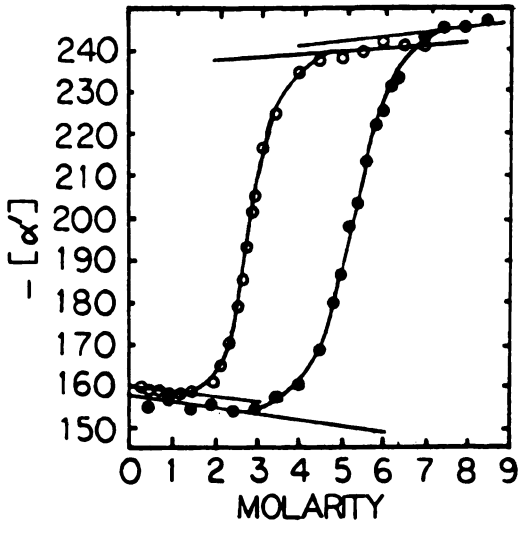
Chapter I

Denaturation and the Mechanism of Protein Folding: An Introduction

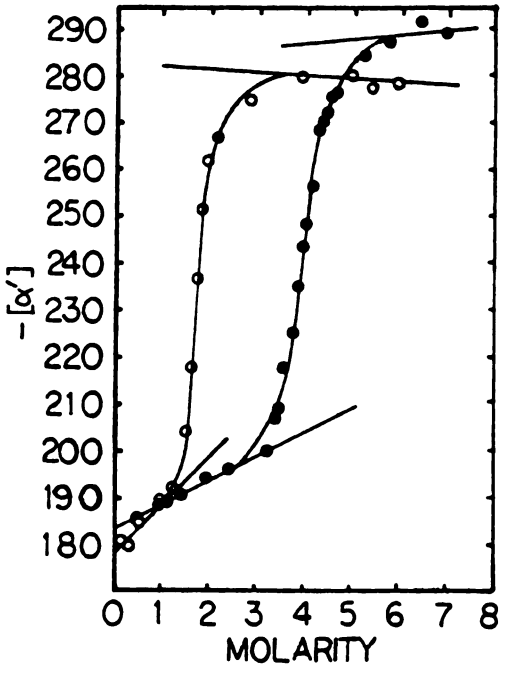
Protein denaturation is a process in which the three-dimensional structure of a protein molecule is changed from that of the native conformation to a more disordered structure.[1] This unfolding is accompanied by loss of the native function and by dramatic changes in many of the physical and chemical properties of the protein. Denaturation can be induced by exposing the protein to heat, extremes of pH, solvent-air boundaries, and a wide variety of chemicals which includes some inorganic salts (e.g., LiBr and KSCN), detergents, organic solvents (especially halogenated acetic acids and ethanols), and a class of organonitrogen compounds (urea, guanidine HCl (hereafter referred to as GdnHCl), and related compounds). Susceptibility to denaturation by each of these varies from protein to protein. The regeneration of the native conformation from the fully-unfolded (random coil) form also varies from one protein to another, and with the denaturant for a given protein. Acids, heat, GdnHCl, and, urea are the most frequently used protein denaturants with GdnHCl and urea being more effective than the former. A greater degree of unfolding is generally observed with GdnHCl than with urea solutions of the same concentration.[2] Two of these observations are illustrated by Figure 1 in which the reduced specific rotation at 365 nm. is plotted against denaturant concentration for the proteins lysozyme, ribonuclease,



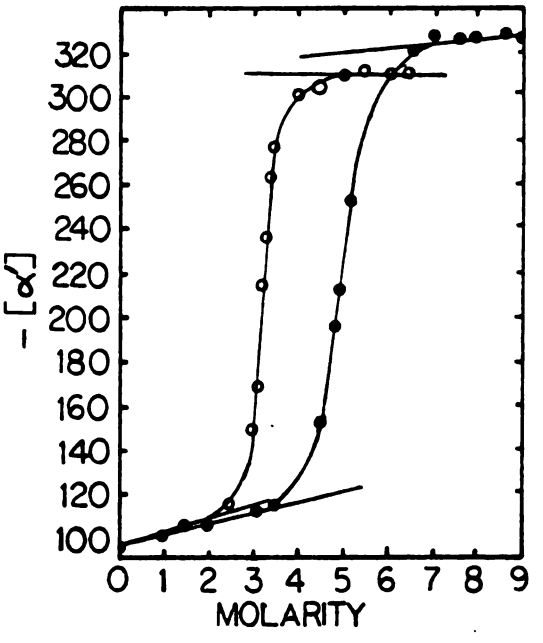
(a) Ribonuclease at pH 6.6.



(b) Lysozyme at pH 2.9.



(c) α -Chymotrypsin at pH 4.3.



(d) β -Lactoglobulin at pH 3.2.

Figure 1. The optical rotation of four proteins at 365 nm. is shown as a function of Urea (●) and Guanidine HCl (○) concentration. (From Reference 3)

α-chymotrypsin, and β-lactoglobulin.[3] All the proteins undergo significant structural changes at lower concentrations of GdnHCl than urea, and the concentrations at which the transitions occur vary somewhat among the four proteins. The literature in this field is extensive and a number of very helpful review articles have been published including those of Edelhoch and Osborne,[4] Franks and Eaglund,[5] Pace,[2] Tanford,[6,7] and Kauzmann.[1]

The mechanism of protein denaturation can be considered half of one of the fundamental problems of biochemistry--the mechanism of protein folding. Because many proteins can undergo denaturation reversibly <u>in vitro</u> to regain the fully functional native structure from the denatured conformation, it is presumed that the amino acid sequence contains all the information necessary to determine the three-dimensional structure. The manner in which this is accomplished from the randomly-coiled polypeptide chain is not well-understood; a random search by the molecule of all the possible conformations is not the mechanism as such a search would take prohibitively long.[8] The most popular theory currently is that of a folding mechanism in which the polypeptide chain is guided to the final conformation through the formation of local structures in different parts of the molecule (nucleation). It is generally assumed that the folding nuclei are elements of secondary structure (α-helix, β-sheet, β-bends) which are formed earliest and which then direct the rest of the protein molecule into place.[9]

Thermodynamic studies have generally supported a two-state folding/unfolding mechanism in which the native conformation (N) and the denatured conformation (D) are directly interconvertible as indicated by the equilibrium expression[7,10,11]

$$N \rightleftharpoons D.$$

For example, Privalov and Khechinashvili[12] studied the thermal denaturation of five proteins (metmyoglobin, ribonuclease, cytochrome c, lysozyme, and α-chymotrypsin) by scanning microcalorimetry techniques. They found that to a good approximation the thermal denaturation of these proteins could be described by the two-state mechanism, and that the small deviations from the model which were observed must be due to folding intermediates present in small concentrations. Further, they produced the interesting result that the differences in free energy between the native and the denatured states for the five proteins are all in the range 7-13 kcal./mole of protein at 37°C.

Kinetic studies of denaturation processes present a more complicated picture. The presumption that a protein molecule must proceed through partly-folded intermediates in transit between the native and denatured states is supported by the results of many kinetics experiments.[2] Commonly observed in many thermal-, urea-, and GdnHCl-induced denaturation processes are simultaneous fast and slow unfolding and refolding reactions whose rates may be different by two orders of magnitude.[13] As one example, Baldwin and coworkers[14]

observed two reactions with rates differing by a factor of 450 inthe refolding of GdnHCl-denatured ribonuclease at pH 5.8 and at 25°C.

Brandts and coworkers[13] have suggested the possibility that the slow folding (or unfolding) reaction may be due to cis-trans isomerism of the amino acid residue proline. They found from a review of the known protein structures determined by X-ray crystallography that the trans-conformation of proline is that most often observed (90-95% of all prolines observed) in native proteins. Therefore, the normally-trans proline residues in the native protein may be in equilibrium with the cis-conformation in the random-coil protein, and the cis-prolines may impede the normal refolding rate of the protein molecule. They suggest further that the slow unfolding reaction is really a transition among two or more partly unfolded forms. The partially unfolded intermediates formed as a result of a cis-trans equilibrium should be detectable in kinetics experiments but not in calorimetric experiments, as the enthalpies of the cis-trans isomers are about the same.

The enzyme α-chymotrypsin (hereafter referred to as CHT except when distinguished from other chymotrypsins) and its enzymatically-inactive precursor chymotrypsinogen (hereafter referred to as CHTgen) have been the subjects of many denaturation studies. The native-denatured protein equilibrium of CHTgen was studied by Brandts[15] over the pH range 0.5-3.7 and over the temperature range 0-65°C. He found that the

heat capacities of the native and denatured forms differ by several thousand cal./deg. mole and that a salt present in the CHTgen solutions could act either to stabilize or denature the native protein depending on the pH of the solution. Biltonen and Lumry[16] studied the denaturation of CHT in the acid pH range and found two distinct native-denatured equilibria. Porter and Preston[17] found that CHT retains its catalytic activity in dilute solutions of the detergent sodium dodecyl sulfate (SDS). Finally, the change in free energy for the acid-thermal unfolding of CHT at pH 3 and at 25°C. has been determined to be 7.4 kcal./mole CHT.[15]

The study of denatured proteins by X-ray methods was first undertaken by Astbury and coworkers in 1935 who obtained and analyzed X-ray photographs of fibers of denatured egg albumin and seed proteins.[19] They observed that the conformation state of the denatured proteins was that of the extended polypeptide chain as opposed to the globular conformation exhibited by native proteins in solution. The first account of protein denaturation in a crystal is that of Snape, et. al.,[20] who observed changes in the X-ray diffraction pattern of lysozyme at several concentrations of urea. They were able to produce crystals which diffracted well even when soaked in urea solutions as concentrated as 9M. However, they concluded that the changes in structure, even at lower urea concentrations such as 3M, were so great as to preclude a detailed analysis based on conventional difference electron density maps. A low resolution (3.0 A.) difference map

revealed widespread changes throughout the lysozyme molecule More recently, the results of some crystallographic denaturation studies by Yonath and coworkers[21,22] have been published. They have examined the effects of several denaturants (urea, KCNS, bromoethanol, and SDS) on triclinic crystals of lysozyme cross-linked by glutaraldehyde. They find, for example, that SDS penetrates the hydrophobic interior of the lysozyme molecule causing the two "wings" to spread apart.

At the outset of the work to be reported here, the effects of several chemically-distinct denaturants on the three-dimensional structure of CHT were undertaken for investigation. Urea and GdnHCl were obvious choices. Trichloroacetic acid (TCA), which has been used in the separation of protein inhibitors from the other proteins present in many varieties of beans, was another choice presumably because it has special denaturing abilities not possessed by the more powerful mineral acids.

Chapter II

The Crystal and Molecular Structure of Chymotrypsin.

1. The Activation and Primary Structure of Chymotrypsin.

The function of the enzyme CHT is the degradation of other proteins during digestion. Specifically, CHT catalyzes the hydrolytic cleavage of peptide bonds whose carbonyl groups belong to amino acid residues with aromatic side chains (Trp, Tyr, Phe). It is a serine protease as its catalytic properties are due to an especially reactive serine residue. It is functionally and structurally similar to a number of other serine protease enzymes including the digestive enzymes trypsin and elastase, the blood clotting enzymes factor X and thrombin, and the bacterial proteases alpha-lytic protease and subtilisin.

α-CHT A is produced in the mammalian pancreas as the catalytically-inactive precursor (zymogen) CHTgen A. Bovine CHTgen A is a single-stranded polypeptide chain containing 245 amino acid residues and cross-linked by 5 disulfide bonds. The amino acid sequence, or primary structure, of CHTgen A was determined by Hartley[23,24,25] and confirmed Meloun, et. al.[26] There exists a second bovine CHTgen, CHTgen B, also found in pancreatic extracts, whose amino acid sequence contains 245 residues and is very similar to the sequence of CHTgen A,[27] and can be activated by trypsin to form the active enzyme CHT B.

Bovine CHTgen A can be activated to yield several CHT A's by the enzymatic cleavage of certain peptide bonds by

trypsin and CHT. π-CHT results from the cleavage by trypsin of the CHTgen peptide bond Arg 15-Ile 16. Subsequent cleavage by CHT of the peptide bond Leu 13-Ser 14 and the removal of the dipeptide Ser 14-Arg 15 gives δ-CHT. Further cleavage at Tyr 146-Thr 147 and Asn 148-Ala 149 by CHT results in the removal of the dipeptide Thr 147-Asn 148 and yields α-CHT.[28,29] The key cleavage step is the first one as both π- and δ-CHT are fully active enzymes having, in fact, significantly higher activities than α-CHT.[30] A fourth CHT, γ-CHT, is known and has been shown to be a pH conformer of α-CHT.[31]

X-ray crystallographic methods have shown that the π-, δ-, and γ-forms of CHT have isomorphous crystal structures indicating that these forms of the enzyme have very similar conformations.[32] In fact, the overall folding of the polypeptide chains in CHTgen and all the forms of CHT is similar.[33,34] As a result of comparisons of these protein structures, the changes which occur upon activation are relatively small in number and result mainly in an improved substrate binding site which is only partially formed in CHTgen[34] and small changes in the orientation of the active Ser 195 side chain.[35]

The α-CHT molecule is composed of 241 amino acid residues in three polypeptide chains--the A-chain (residues 1-13), the B-chain (residues 16-146), and the C-chain (residues 149-245)--with a molecular weight of approximately 25,300. The amion acid sequence is shown in Figure 2, with

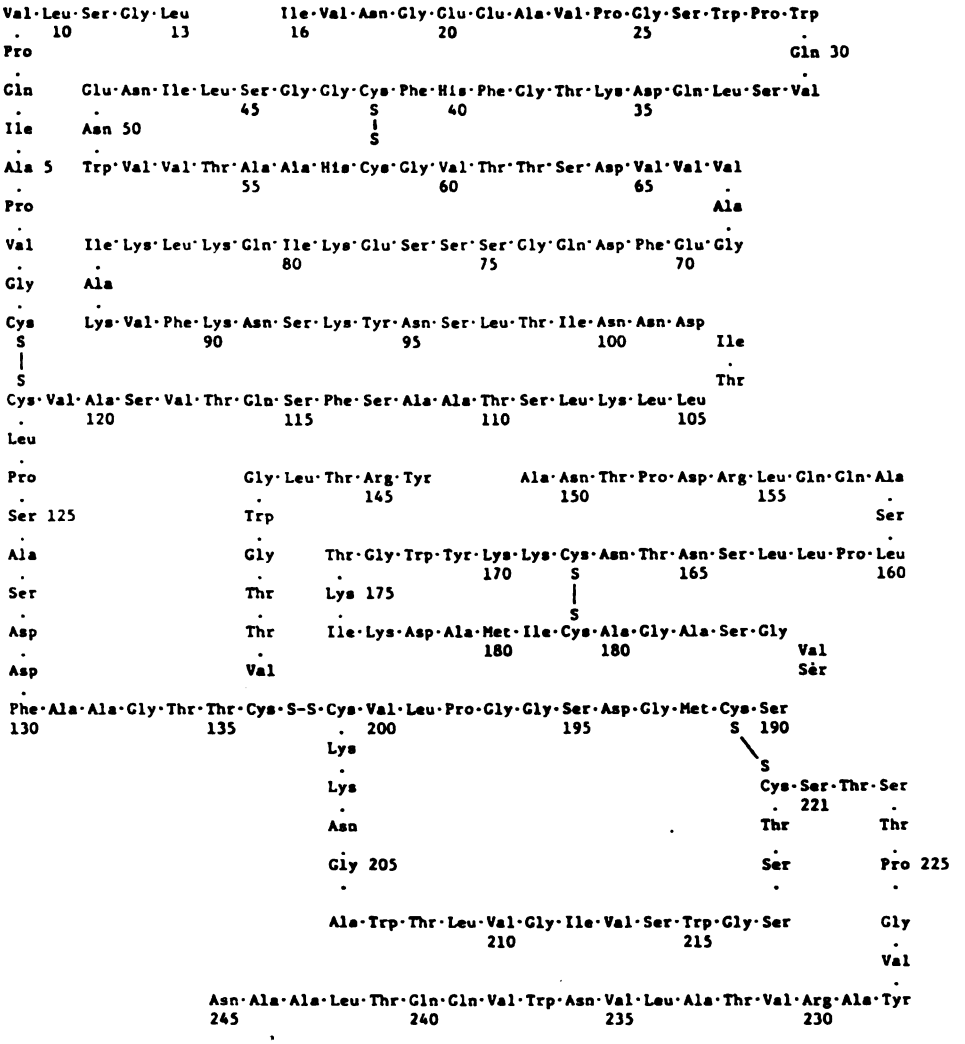


Figure 2. The amino acid sequence of CHT.

the numbering of the residues being the same as that of the amino acid sequence of CHTgen. Five disulfide bonds connect residues 1 and 122, 42 and 58,136 and 201, 168 and 182, and 191 and 220 forming both inter- and intra-chain linkages.

2. The Crystal Structure of Chymotrypsin.

CHT crystallizes as monoclinic plates from solutions 50% saturated in ammonium sulfate at pH 3.6. The crystal structure has the symmetry of space group $P2_1$ with four molecules per unit cell or two molecules per asymmetric unit. This arrangement is illustrated in Figure 3. The molecules form "chains" parallel to the c-axis such that the molecules of one chain are related to the molecules of neighboring chains by two-fold screw axes coincident with and parallel to the b-axis. Thus, Molecules I and I' are related to molecules II and II' respectively by the two-fold screw axes (Figure 3b). Molecules in the same chain are related by two different non-crystallographic or local two-fold rotation axes[36] labelled A and B (Figure 3a). The local two-fold axis A relates molecules I and II to I' and II' respectively. The local dyad B also relates I to I' but in a different way.

Molecules of CHT form dimers in solution at lower pH and from comparisons of the results of solution dimerization equilibrium experiments with the X-ray crystallographic structure, it has been concluded that the dimer in solution is the one formed by dyad A in the crystal.[37,38]

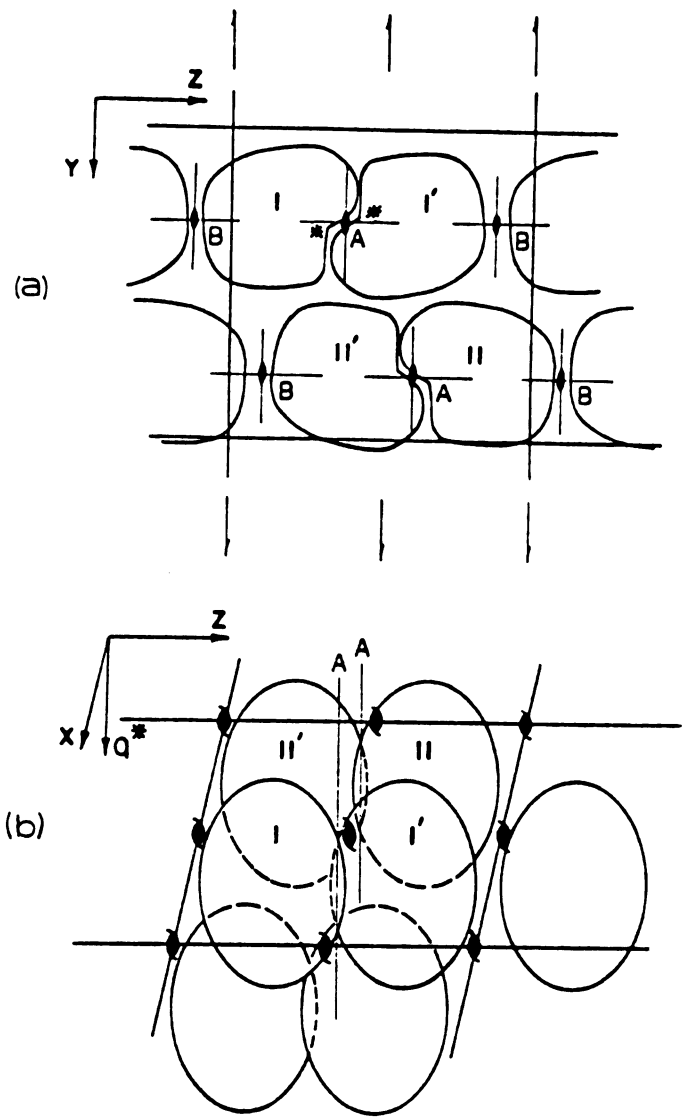


Figure 3. The packing of CHT molecules in the crystal. (a) Projection on the yz-plane. The asterisks denote the active site regions. (b) Projection on the xz-plane.

Therefore, all further references to the CHT dimer refer to this pair of molecules.

Another important aspect of the crystal structure is the solvent channels parallel to the c-axis which result from the manner in which the CHT molecules pack in the crystal. The channels are large enough to permit the diffusion of small molecules, i.e., heavy-atom compounds, enzyme inhibitors, denaturant molecules, fluorescent probes, etc., into the crystal lattice to interact with the protein.

3. The Molecular Structure of Chymotrypsin.

The structure of the CHT molecule has been determined independently using X-ray crystallographic methods by Blow and coworkers at the Medical Research Council Laboratories, Cambridge, England[39,40] and by Tulinsky and coworkers at Michigan State University.[41,42] The details of the structure are recorded in the papers cited. The intention here is to describe briefly those aspects of the CHT structure most pertinent to the results to be presented later.

(a) Secondary Structure.

The CHT molecule consists largely of extended polypeptide chains, often running parallel or antiparallel to one another, forming several sections of ill-formed β-sheet. The extended chains are often terminated by β-bends--sharp turns in the polypeptide backbone in which a hydrogen bond may occur involving a peptide carbonyl oxygen and an α-amino

hydrogen three residues away. There are two short helices in the molecule: three turns of α-helix made up of the polypeptide segment Leu 234-Asn 245 occurs at the carboxy-terminal end of the C-chain on the exterior of the molecule on the side opposite the active site: the other helix is one-and-a-half turns of distorted α-helix between Ser 164 and Gly 173, also occurring on the surface of the molecule.

(b) Folding Domains and the Distance Diagonal Plot.

The tertiary structure of a protein arises from the organization by the protein of its elements of secondary structure to form a globular molecule. From an examination of its β-sheet structure, it is evident that the CHT molecule is organized into two folding domains. These have been described in the literature by Birktoft and Blow[40] and called Cylinders 1 and 2. Cylinder 1 is made up of β-sheet from polypeptide segments of the B-chain. The axis of Cylinder 1 lies nearly in the plane defined by the local two-fold axes at an angle of about 40° with respect to the two-fold axes (Figure 4). The axis of Cylinder 2 makes an angle of about 45° with respect to the ac-plane of the crystal at right angles to the axis of the Cylinder 1. The two cylinders, or domains, thus formed divide the CHT molecule roughly into halves such that the molecule may be bisected by a plane parallel to the bc-plane of the crystal close to the active site. The arrangement of the cylinders with respect to the overall shape of the CHT molecule is illustrated schematically in Figure 4.

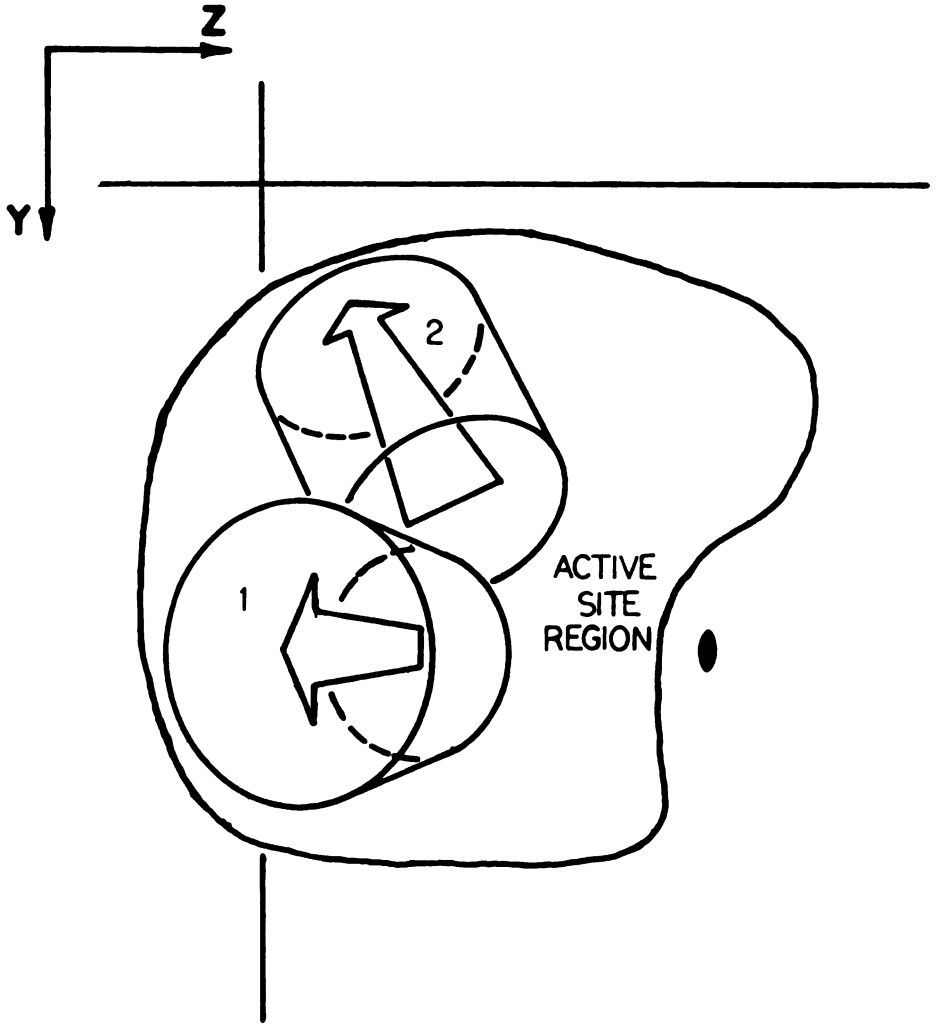


Figure 4. Schematic illustration of the folding domains in the CHT molecule. Cylinders 1 and 2. Denoted, respectively.

The interiors of both cylinders consist almost entirely of aliphatic side chains (Val, Leu, Ile) except for a small number of polar side chains occurring in both cylinders near the center of the molecule. Two serines--Ser 54 and Ser 104--occur in the innermost end of Cylinder 1. Three polar side chains--Thr 138, Ser 190, and Tyr 228--are found in the innermost end of Cylinder 2.

A concise and dramatic representation of protein tertiary structure is the distance diagonal plot.[43,44] The distance diagonal plot is an NxN matrix, the elements of which are r_{ij}, the distances between the i-th and the j-th α-carbon atoms, where N is the number of amino acid residues in the protein molecule. The array thus calculated may be contoured at convenient distance levels. The distance diagonal plot for CHT is shown in Figure 5. Since $r_{ij} = r_{ji}$, the matrix is symmetrical and only one-half is unique. Features of the distance diagonal plot correspond to the close approach (below an arbitrary limit) of one α-carbon atom to another and are a unique two-dimensional representation of the three-dimensional folding.

Elements of secondary structure are immediately evident from the distance diagonal plot. An α-helix appears as a thick feature on the diagonal and the C-terminal α-helix (residues 234 to 245) of CHT is distinctly visible. Segments of parallel β-sheet are features parallel to the diagonal at some distance from it. Anti-parallel β-sheet is represented by lengths of density perpendicular to the

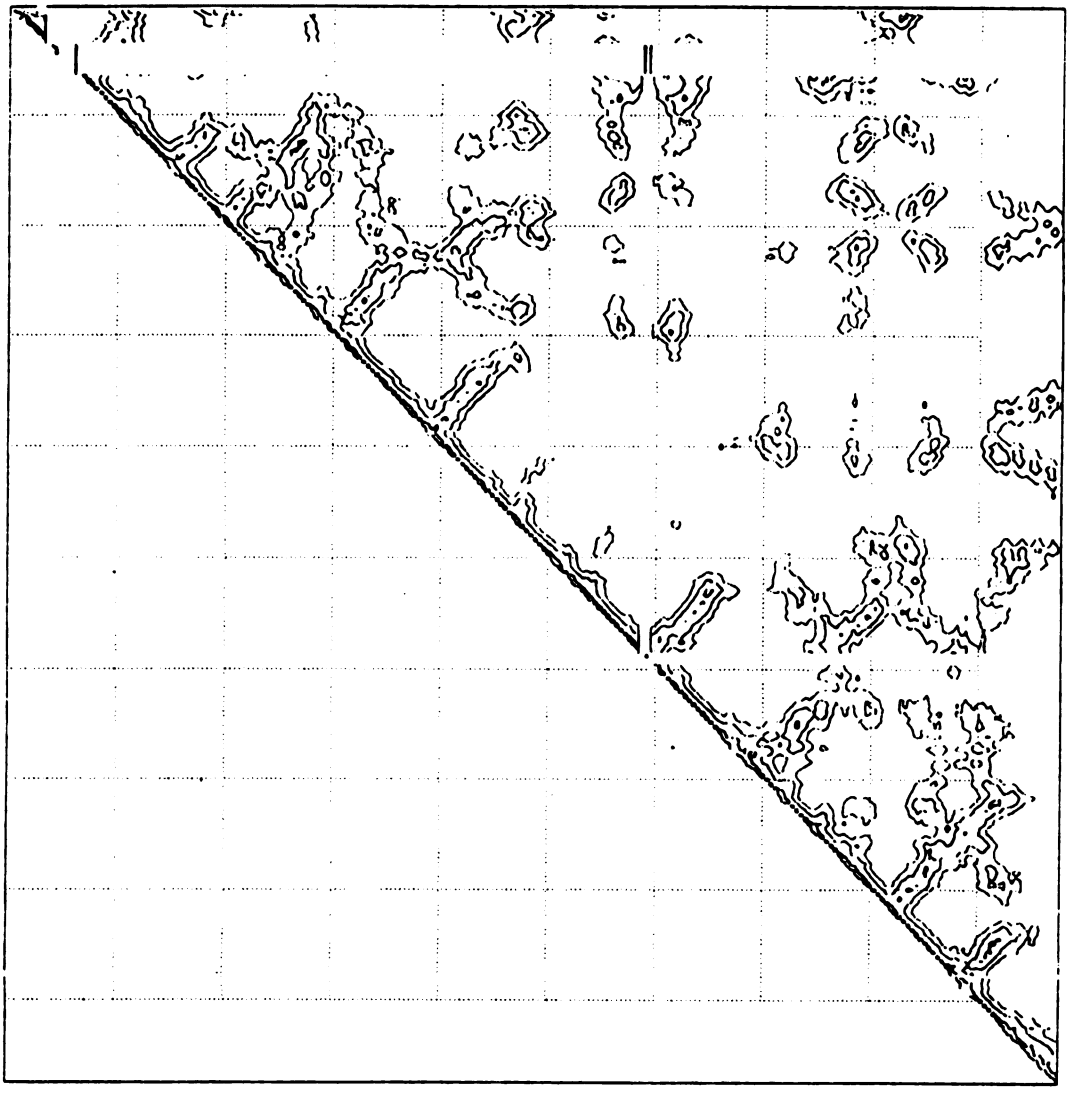


Figure 5. The distance diagonal plot of CHT. Plot contours are 5.0, 10.0, and 15.0 A. Interatomic distances, rij, were calculated and plotted using the CEMCOMGRAF facility, Department of Chemistry, Michigan State University. Plot courtesy of Dr. S. R. Ernst.

diagonal and connected to the diagonal. Other features occurring further from the diagonal represent the close approach of polypeptide strands and loops which are not hydrogen-bonded secondary structural features. Detailed interpretations of these have been given elsewhere.[45,46]

The CHT folding domains formed from segments of parallel and antiparallel β-sheet appear in the distance diagonal plot as two groups of blocks stacked together on the diagonal - the two largest features in the plot. A further interesting feature of the CHT distance diagonal plot is the approximate mirror of symmetry perpendicular to the diagonal at about residue 123. The two domain features are clearly mirror images suggesting that the amino acid sequence 1-122 might be homologous to the reversed sequence of 123-245 (palindromic homology). Conventional measures of sequence homology,[47] however, have not borne out this relationship.[48,49]

That folding domains are a widespread feature among proteins, and that the distance diagonal plot may be useful for studying them has been pointed out by Rossman[50] and by Kuntz.[45] Rossman and Liljas[50] showed that particular structural domains may be associated with a particular function, and that the corresponding parts of different enzymes performing the same function may have the same three-dimensionalfolding (as indicated by the distance diagonal plot) while exhibiting little amino acid sequence homology. Further, Liebman[46] has discovered an example of similar folding occurring in two proteins having

opposing functions, i.e., CHT (enzyme) and pancreatic trypsin inhibitor (substrate) where the overall folding of the latter approximates that of the first half of Cylinder 1 of CHT.

(c) Nonpolar Interdomain Region.

The two β-sheet cylinders with interior nonpolar amino acid side chains constitute two large hydrophobic volumes in the CHT molecule. There exists also a third hydrophobic region--a semicircular band of nonpolar side chains wrapping around the inner ends of both cylinders. In contrast to the cylinders which are filled with aliphatic side chains, this region has in addition to aliphatic residues a large number of aromatic groups, most of which are arranged in clusters. This region also contains a nonpolar cavity, about 8 A. in diameter, just inside the molecule near the C-terminal α-helix. The nonpolar amino acid residues which form this unusual feature are highly conserved among the other serine proteases whose tertiary structures are similar to that of CHT.[42] The arrangement of the nonpolar groups and the nonpolar cavity are illustrated in Figure 6.

The CHT molecule contains four clusters of aromatic side chains, three of which occur in the interdomain region. The tryptophan cluster--Trp 27, Trp 29, and Trp 207--along with the three prolines--Pro 4, Pro 8, and Pro 28--constitutes a binding site for some substrate-like molecules[51] as well as capping one end of the hydrophobic cavity. The residues Trp 51, Phe 89, and Trp 237 cap the other end

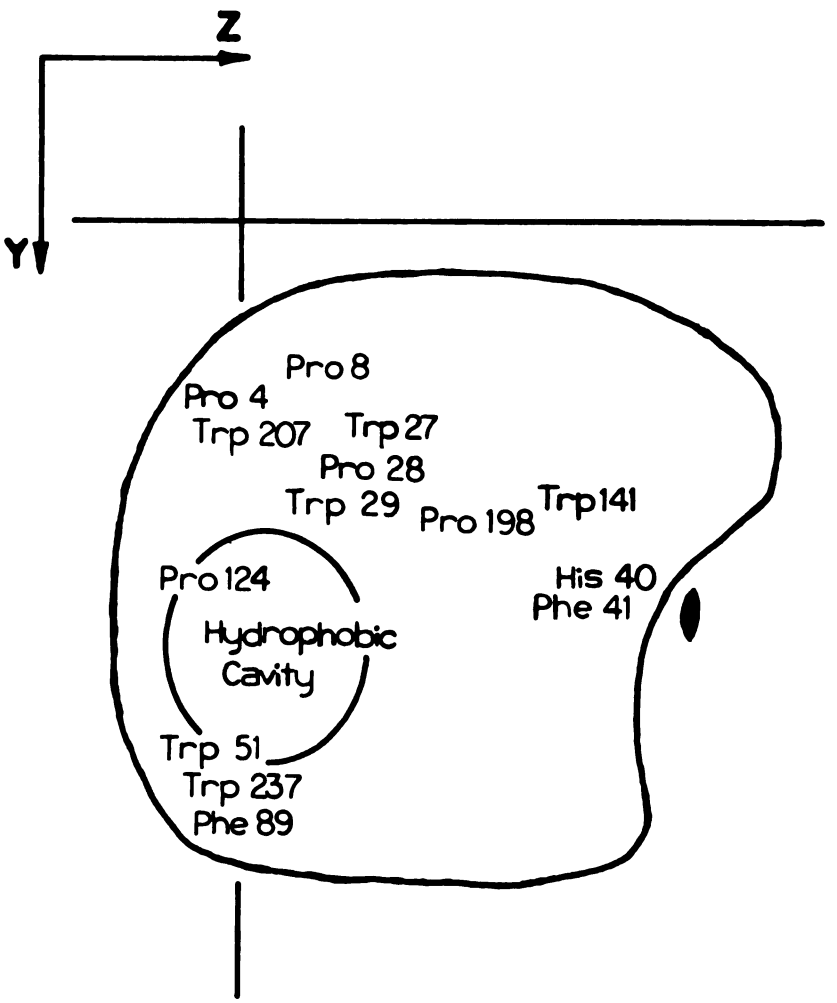


Figure 6. The nonpolar interdomain region of CHT shown is projection along the local two-fold rotation axis. The residues shown occur in the x-interval 45/76 to 65/76.

of the hydrophobic cavity. The group His 40, Phe 41, and Trp 141 extend toward the active site region completing the semicircular interdomain arrangement of aromatic residues. A fourth aromatic cluster--Tyr 171, Trp 172, and Trp 215--occurs just inside the surface of the molecule near the 1.5 turns of distorted α-helix formed by residues Ser 164-Gly 173.

(d) The Dimer Interface Region.

The dimer interface region, or the region in which the two CHT monomers make contact about dyad A as shown in Figure 3a, is formed as a result of the complementary fit of the monomers. It is characterized by a number of mutual close contacts between the amino acid side chains of both molecules involving van der Waals interactions, hydrogen bonding, and ion-pair bond formation. These interactions are summarized schematically in Figure 7. The polar interactions involve pairs of residues in a mutually reciprocal manner while the nonpolar interactions involve residues of one molecule with their exact counterparts in the other molecule.[52]

At pH 3.6, the ε-amino group of lysine is protonated making possible the two ion-pairs Lys 36'-Asp 153 and Lys 36-Asp 153', where the primes indicate that a particular amino acid residue belongs to CHT molecule I'. The terminal amino group at Ala 149 is also protonated at this pH resulting in the ion-pairs Asp 64'-Ala 149 and Asp 64-Ala 149'. Extensive hydrogen bonding also occurs. The B-chain

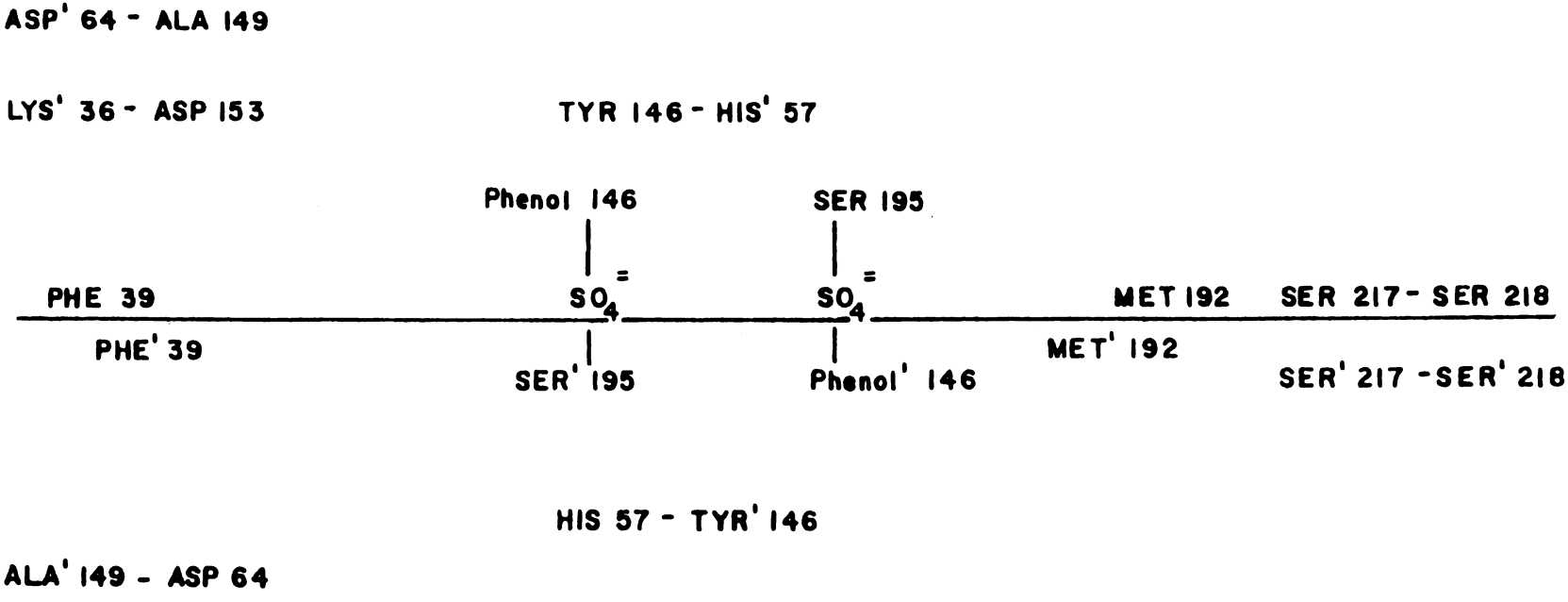


Figure 7. Schematic representation of the dimer interface interactions projected on a plane perpendicular to the yz-plane and containing the local two-fold rotation axis (horizontal line through center of figure). (From Reference 52).

carboxy-terminal residue Tyr 146 is located such that it may form hydrogen bonds with the carbonyl of His 57 of the other CHT molecule and with the hydroxyl-group of Ser 195 of the other CHT molecule through a sulfate ion bridge. Finally, van der Waals forces determine the equilibrium positions of the phenyl groups of Phe 39 and Phe' 39 so as to minimize close contacts between them.

The two residues Phe 39 and Phe' 39 display a further general property of the dimer interactions. The local two-fold symmetry in this region is far from exact and differences in conformation occur between the two molecules. When CHT molecules dimerize, the phenyl rings of Phe 39 and Phe' 39 must rotate about their respective C_α-C_β bonds because of close contacts and do not display local two-fold symmetry. This lack of two-fold symmetry has been termed variability in structure and has been discussed in some detail.[41,42] The variability in structure is often reflected in chemical behavior as well, both in the asymmetric binding of small molecules to the protein and in changes in the protein molecules accompanying derivative formation. This phenomenon is displayed by the TCA, the GdnHCl, and the urea derivatives.

Chapter III.

Experimental.

1. The Crystal Soaking Experiments.

The denaturation experiments were performed by exposing crystals of CHT to gradually-increasing concentrations of several chemical denaturants and by observing the diffraction patterns of individual crystals at periodic intervals of time. If substantial changes were evident with respect to the diffraction pattern of crystals of native CHT, the three-dimensional X-ray intensities were measured at 2.8 A. resolution, reduced to structure amplitudes, and scaled to the native intensities, and difference electron density maps were calculated and examined.

The CHT used in all the denaturation experiments was bovine α-CHT, three times crystallized, from Worthington Biochemical Corporation. All the denaturants and other chemicals were commercially-available, reagent grade quality, and used without further purification. Crystals of CHT were grown from aqueous solutions approximately 50% saturated in ammonium sulfate at pH 3.6 with a protein concentration of about 5 mg./ml. of solution. The tubes containing the mother liquor were stored at room temperature and the crystals--diamond-shaped plates--usually appeared within four to seven days after the solutions were prepared. After two weeks time, crystals of suitable size (1.0-1.5 mm. along the longer face diagonal) and apparent quality (clear and without cracks) were harvested and placed in tubes

containing aqueous solutions 75% saturated in ammonium sulfate with the pH adjusted to 3.6. The crystals were then stored in a water bath at 14-16°C. Crystals of CHT may be stored in this manner indefinitely. Those having a regular shape and as thick as possible (0.2-0.3 mm.) were chosen for further soaking experiments from this stock of stored crystals.

The soaking experiments were carried out in two-chambered soaking cells (Figure 8). The cells consisted of a lower chamber (a 5 ml. snap-cap vial) containing the crystals bathed in the storage solution, covered by a cellophane dialysis membrane, and an upper chamber fashioned from a short length of glass tubing having the same outside diameter as the vial. To assemble the cell, the crystals selected for the experiments were placed in the vial and covered by the 75% saturated ammonium sulfate solution. The vial was filled completely and covered by the membrane so that no air bubbles were trapped under the membrane. The membrane was then secured in place by a sleeve made from Tygon tubing. The upper chamber was then filled with solution and capped with a rubber stopper or parafilm. The total volume of the cells were about 10 ml. with the volumes of the upper and lower chambers being about equal.

The soaking experiment is begun with the upper chamber filled with the 75% saturated ammonium sulfate solution. A 1-2 ml. aliquot of denaturant dissolved in 75% ammonium sulfate at pH 3.6 is exchanged for an equal volume of

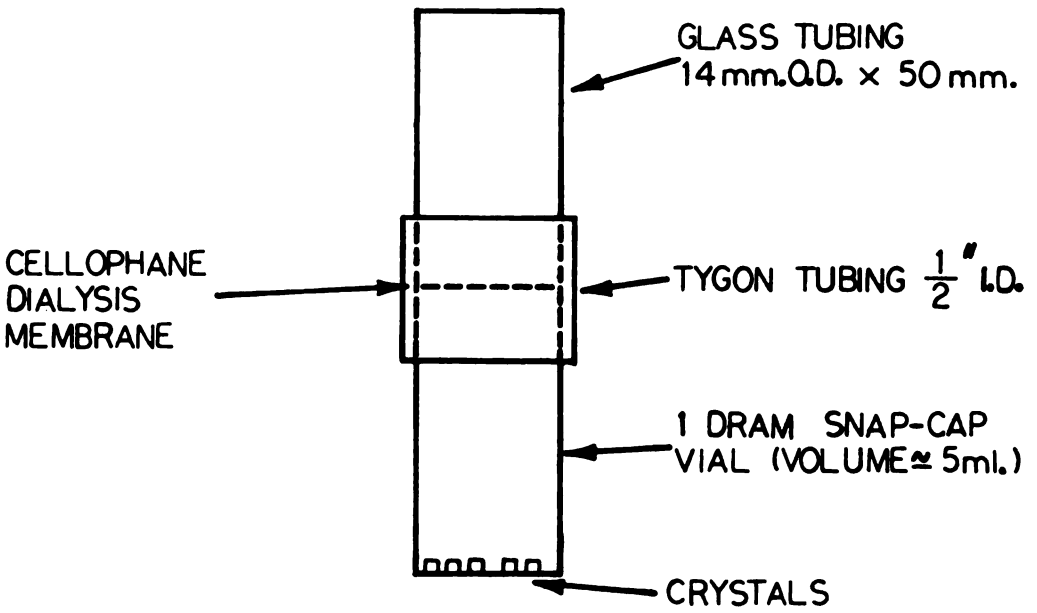


Figure 8. The soaking cell used for the preparation of the denatured derivative CHT crystals.

the solution already in the upper chamber. After this has been done daily for several days and the denaturant concentration in the cell approaches that of the stock denaturant solution, the solutions may be exchanged by simply pouring out the upper chamber and refilling it with the stock at the end of 1 1/2-2 weeks, the concentration of the denaturant in the cell is essentially that of the stock denaturant solution. At this time, a crystal may be withdrawn for an examination of its diffraction pattern. If no changes are evident, either with respect to the native pattern or that of some previous concentration derivative, the soaking cell is reassembled and the aliquot-exchanging process is resumed with the denaturant at some higher concentration. This procedure is continued until substantial changes have appeared in the diffraction pattern of the denatured protein crystals. The pH of the solution in the upper chamber was measured before each aliquot was exchanged to provide a check on the course of the experiment. A variation in pH measurements of 0.1-0.2 pH units from one day to the next was not unusual, but the pH was never allowed to drift from 3.6 by more than 0.3 pH units.

There are several advantages to the use of this soaking cell over the use of a test tube. The dialysis membrane impedes the physical mixing of solutions reducing the possibility of damage to the crystals by a sudden change in the solvent environment. The crystals are physically isolated from the pH meter electrode and transfer pipettes reducing

the danger of their being accidentally jarred or crushed. Finally, exchanging aliquots is convenient when one can simply pour out the upper chamber without jostling the crystals about.

The choice of starting concentration for each denaturant was determined from literature accounts of protein denaturation experiments in solution along with the expectation that changes in the structure of CHT may occur at lower denaturant concentrations than those necessary to obtain a fully-unfolded, random-coil conformation. The urea and GdnHCl experiments began at 1.0 M. concentrations and proceeded to higher concentrations, going in 1.0 M. increments for the urea and 0.5 M. increments for the GdnHCl. Both these compounds are very soluble in 75% saturated ammonium sulfate solutions such that 3.0 M. GdnHCl and 5.0 M. urea solutions are easily produced. TCA also dissolved readily in the ammonium sulfate solutions, but because less was known about its denaturing abilities, a starting concentration of 0.5 M. was chosen. The pH of the TCA solutions was 3.6 like all the other derivative soaking solutions.

The urea experiment proceeded in 1.0 M. steps beginning at 1.0 M. and terminating at 5.0 M. Substantial intensity changes first appear in the axial reflections of the 2.0 M. derivative and reappear in a generally more pronounced way in the 3.0 M. derivative. Further changes which appear in the 4.0 M. diffraction pattern are generally continued in

the 5.0 M. derivative. Since only small changes are evident between the 2.0 and the 3.0 M. 2θ-scans, the three-dimensional intensity data for the 3.0 M. derivative were measured at 2.8 A. resolution.

Scans were recorded and examined for GdnHCl derivatives at the following concentrations: 0.5 M., 1.0 M., 1.5 M., 2.0 M., and 2.5 M. Small changes first appear in the axial reflections of the 1.0 M. crystals so that three-dimensional intensity data were measured for this derivative at 2.8 A. resolution. At 2.0 M. GdnHCl, further changes with respect to the native diffraction pattern were observed so three-dimensional intensity data were also measured for this derivative. At 2.5 M., the crystals had become opaque with cracks and the diffraction pattern had sharply deteriorated making data collection impractical.

Crystals of CHT soaked in 0.5 M. TCA solutions 75% saturated in ammonium sulfate at pH 3.6 became badly cracked and disintegrated within 24 hours. Crystals soaked in 0.1 M. TCA solutions became thoroughly cracked but withstood mounting in a capillary and the 2θ-scans were recorded. Dramatic differences with respect to the native diffraction pattern, even at low 2θ-angles, were evident on all three axes. The rapid decline in intensity with increasing 2θ-angle as well as the high background, which itself had a peak at about 22° in 2θ, indicated the relatively poor quality of this crystal. Finally, crystals soaked in 0.01 M. TCA solutions produced good quality diffraction patterns

exhibiting significant changes with respect to the native diffraction pattern, and three-dimensional X-ray intensities were measured for this derivative at 2.8 A. resolution.

The axial 2θ-scans for the 2.0 M. GdnHCl, the 3.0 M. urea, and the 0.01 M. TCA derivatives along with those of the native CHT crystals are shown in Figure 9.

2. X-ray Intensity Measurements.

Accurate measurement of the X-ray intensities which are subsequently converted to structure factor amplitudes requires the careful consideration of a number of experimental problems, many of which arise because of the delicate nature of a protein crystal. Protein crystals must be enclosed in a sealed glass capillary in contact with the mother liquor because if allowed to dry, the crystal lattice breaks down and the diffraction pattern is lost. Protein crystals mounted in this manner[53] are held against the inside wall of the capillary by surface tension of the mother liquor and occasionally move becoming misaligned during the course of data collection. If not corrected, this misalignment will cause the intensity measurement to be in error, reduced from the true value. Protein crystals are sensitive to radiation damage so that in developing an optimal data collection procedure, the reduction in intensity due to radiation damage must be balanced against an adequate amount of exposure for reliable precision. These and other such problems have been overcome to the extent that it is now routine practice in our laboratory to measure about 6400

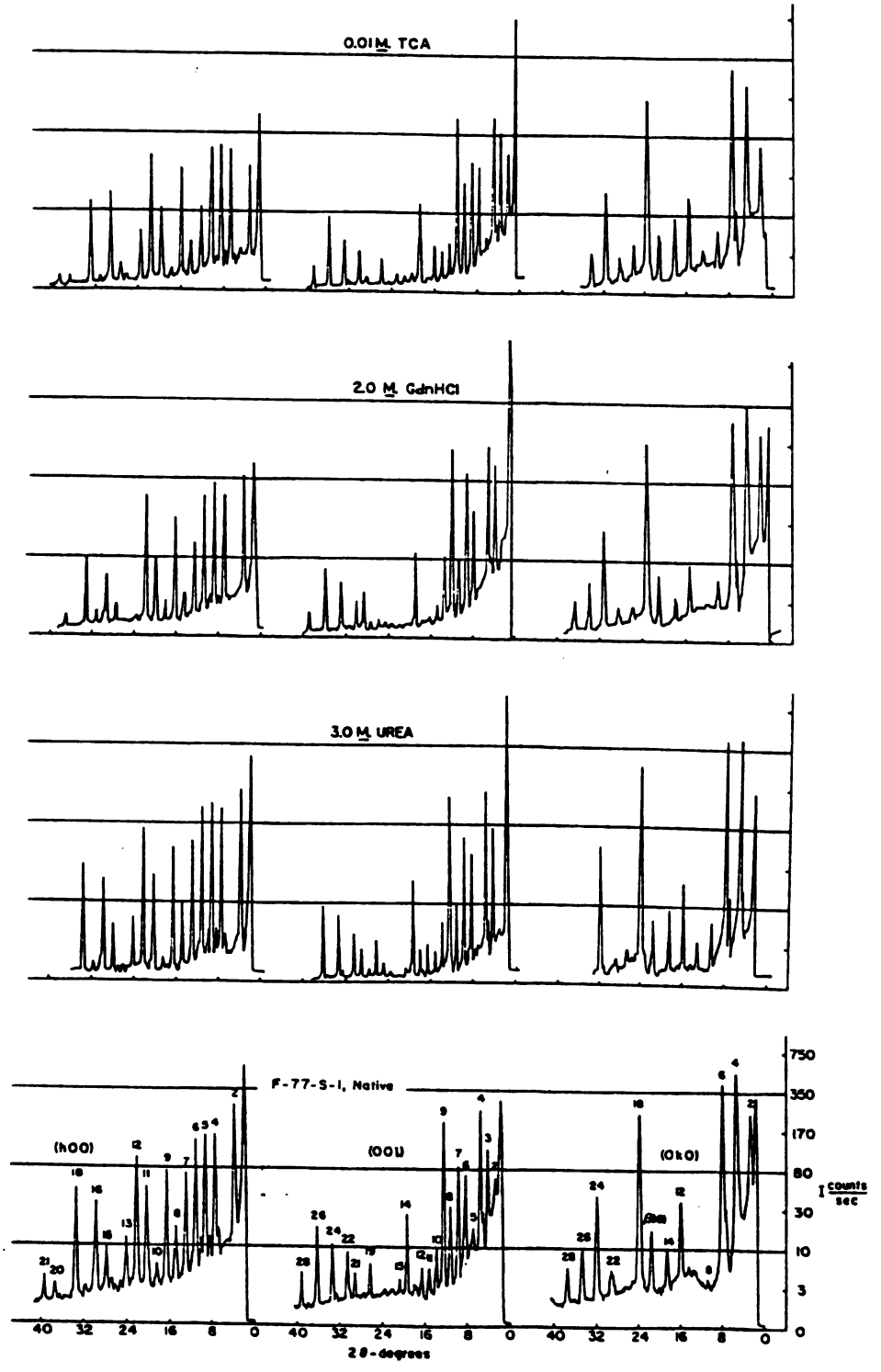


Figure 9. The axial 2θ-scans for the 0.01 M. TCA, 2.0 M. GdnHCl, and 3.0 M. urea derivatives, and native CHT, respectively.

X-ray intensities from a single crystal corresponding to about 75 hours of X-ray exposure and typically experiencing 10 to 20% decay in intensity due to X-ray damage. Many of the details of the experimental methods used in our laboratory have already been described elsewhere.[41,54]

The X-ray intensity measurements to be described here were made using a Picker FACS-1 4-Circle Automatic Diffractometer controlled by a Digital Equipment Corporation (DEC) PDP-8 computer interfaced also with a DEC 32K DF32 Disc File and an Ampex TMZ 7-track tape transport. The software system used in our laboratory is based on the program package originally supplied with the FACS-1 system by Picker but adapted to perform protein data collections more suitably.[54] The data collected by this system is written on magnetic tape and transferred to the Michigan State University Computer Center for calculations on the CDC 6500 computer.

Due to the fact that protein crystals frequently move during data collection, an automatic crystal aligning subroutine is an essential part of the data collection software.[54] Motion of the crystal is monitored by periodically measuring the intensities of several reflections during the course of data collection. The monitored reflections are chosen such that their intensities are very sensitive to small changes in the angular coordinates, and in such a way that they define the orientation of the crystal. The (0,18,0) reflection at the two ϕ-angles corresponding to the $\vec{a}^*$ and the $\vec{c}^*$ axes (which are 78° apart in ϕ) is used

to define the coincidence of the crystal $\vec{b}^*$ axis and the ϕ-axis of the diffractometer. The intensities of this reflection at these two ϕ-angles are extremely sensitive to small changes in ω (0.05° change in ω leads to about 10% loss in intensity). Another reflection in general reciprocal space (non-special angles) at high 2θ is chosen to monitor the ϕ-angles to protect against a twisting motion about the $\vec{b}^*$ axis. These three standard reflections are measured about once every hour during data collection, and if the intensity of any one of them drops below a preset limit (90% of their initial intensity) the data collection program transfer control to the crystal aligning subroutine.

The data collecting program is based on an ω step-scan procedure using balanced Ni/Co filters to obtain a correction for background radiation. The step-scan is performed with respect to the most sensitive angle, ω. The diffractometer is driven to the calculated 2θ, χ, and ϕ coordinates of a reflection and scans through a 0.15° range in beginning at -0.07° and going to $+0.08^\circ$ in ω in 0.03° steps. The intensity is measured for a preset time at each step and the four largest intensities are summed to give the intensity of the reflection (count-six-drop-two step scan).[55] The crystal is then returned to the ω-angle corresponding to the highest intensity and the background is measured with a balanced Co filter. This background measurement is multiplied by 4 to obtain the total background intensity.

Occasionally, the crystal will undergo a slight misalignment such that the reflection peak is not at $\omega = 0.00^{\circ}$. To compensate for this, the step-scan program checks the ω-angle for the step of highest intensity. If the maximum is at, or next to, either end of the ω-step range, the intensities of one or two additional 0.03° steps are measured beyond the end of the step range. If the maximum is one step from either end of the range, only one additional step is measured. If the maximum is at the end of the range, two additional steps are taken. The final reflection intensity is still taken as the sum of the four highest intensities.

3. The Reduction of Measured Intensities to Structure Amplitudes.

In principle, the relative structure amplitude $|\vec{F}(hkl)|$ for a single crystal having a completely isotropic absorption behavior may be expressed as a function of the X-ray intensity by the relation

$$F(hkl) = \left[\frac{I(hkl)}{Lp} \right]^{\frac{1}{2}}$$

where I(hkl) is the measured X-ray intensity and L and p are the Lorentz and polarization corrections, respectively. The polarization correction is necessary because the efficiency of reflection of a non-polarized incident X-ray beam varies with the diffraction angle 2θ since it becomes plane-polarized

at 90^{o} and has the functional form

$$p = \frac{1 + \cos^2(2\theta)}{2} .$$

which is independent of the method of intensity measurement. The functional form of the Lorentz factor depends upon the geometry of the diffraction experiment and for diffractometer measurements is given by

$$L = \frac{1}{\sin(2\theta)} .$$

The need for the Lorentz correction arises because the incident beam is divergent.

Several other corrections are necessary for the proper reduction of CHT data. The most important source of systematic error in the X-ray intensities is the absorption of X-rays by the mother liquor, glass capillary, and by the crystal itself. The absorption correction used in our work is semi-empirical,[56] based on the variation of the relative transmission of X-rays with respect to the azimuthal angle (ϕ). The intensities of four (0,k,0) reflections (k = 2, 4, 6, 18) are measured versus ϕ and are used to calculate an absorption correction for all general (hkl) reflections depending upon ϕ, 2θ, and k-index (reciprocal lattice level). Since crystals of CHT are twinned along the $\vec{c}^{*}$ axis, the (0kl) intensities are the sum of the intensities of the larger crystal and its twin. This systematic effect is

Table 1. Summary of X-ray Intensity Data Collection, Reduction, and Scaling Parameters.

Derivative	Unit Cell Dimensions* a	b	c	β	Twin Ratio (kt)	Decay (%)	Scale (k)	B(A.^{2})
1.0 M. GdnHCl	49.37(5) A.	67.44(6) A.	66.12(8) A.	101.8(1)°	1.49	0.0	1.47	-6.0
2.0 M. GdnHCl	49.27(5)	67.49(6)	66.05(8)	101.9(1)	1.00	2.0	2.13	0.0
3.0 M. Urea	49.49(7)	67.77(8)	66.20(11)	101.9(2)	1.68	10.0	1.81	0.0
0.01 M. Trichloro-acetate	49.18(8)	67.75(10)	65.82(12)	101.6(1)	2.48	3.0	1.83	+1.85
Native CHT	49.24(7)	67.20(10)	65.94(9)	101.79(8)	-	-	-	-

*Numbers in parentheses are the standard errors calculated by the least squares routine.

corrected in data reduction by use of the relationship

$$I(0kl) = \frac{k_t}{1 + k_t} \, 'I(0kl)',$$

where k_t is the ratio of the crystal intensity to the twin intensity, I(0kl)' is the raw intensity, and I(0kl) is the corrected intensity. Background radiation is measured so that a correction to the observed intensity can be made. In our experiments, background radiation is the result of coherent and incoherent scattering and is measured through the use of balanced Ni/Co filters. In actual practice, however, the filters are not quite balanced, and a lack-of-balance correction is applied to the observed background.

A summary of pertinent data collection and reduction parameters is given in Table 1.

4. Scaling the Reduced Structure Amplitudes.

The data reduction calculations just described place the individual intensities in a data set on a common basis. The entire three-dimensional data set, however, must be placed on a scale comparable to that of the native structure amplitudes. There is also the further problem that the X-ray scattering power of a crystal decreases due to disorder with increasing 2θ-angle and that the rate of decrease varies from native to derivative sets of data. Scaling is accomplished by fitting the distribution of average reduced derivative intensities $\langle |\vec{F}_D|^2 \rangle$ to the distribution of average native reduced intensities $\langle |\vec{F}_N|^2 \rangle$ which have already been placed on an absolute scale

(with a B = 27 A.^{2}, see reference 41). This is accomplished by fitting the distribution of relative average derivative intensities $\langle|\vec{F}_D|^2\rangle$, in intervals of 2θ, to the native intensity distribution varying the parameters k and B in the equation

$$\langle|\vec{F}_D|^2\rangle(\text{abs}) = \langle|\vec{F}_D|^2\rangle\ (\text{rel})\ \cdot k^2 \exp(2B\sin^2\theta/\lambda^2)$$

to obtain the best fit (Figure 10). Table 1 contains the k and B values for the denatured derivatives where $\langle|\vec{F}_D|^2\rangle$(abs) and $\langle|\vec{F}_D|^2\rangle$(rel) are the scaled and relative average derivative intensities, respectively.

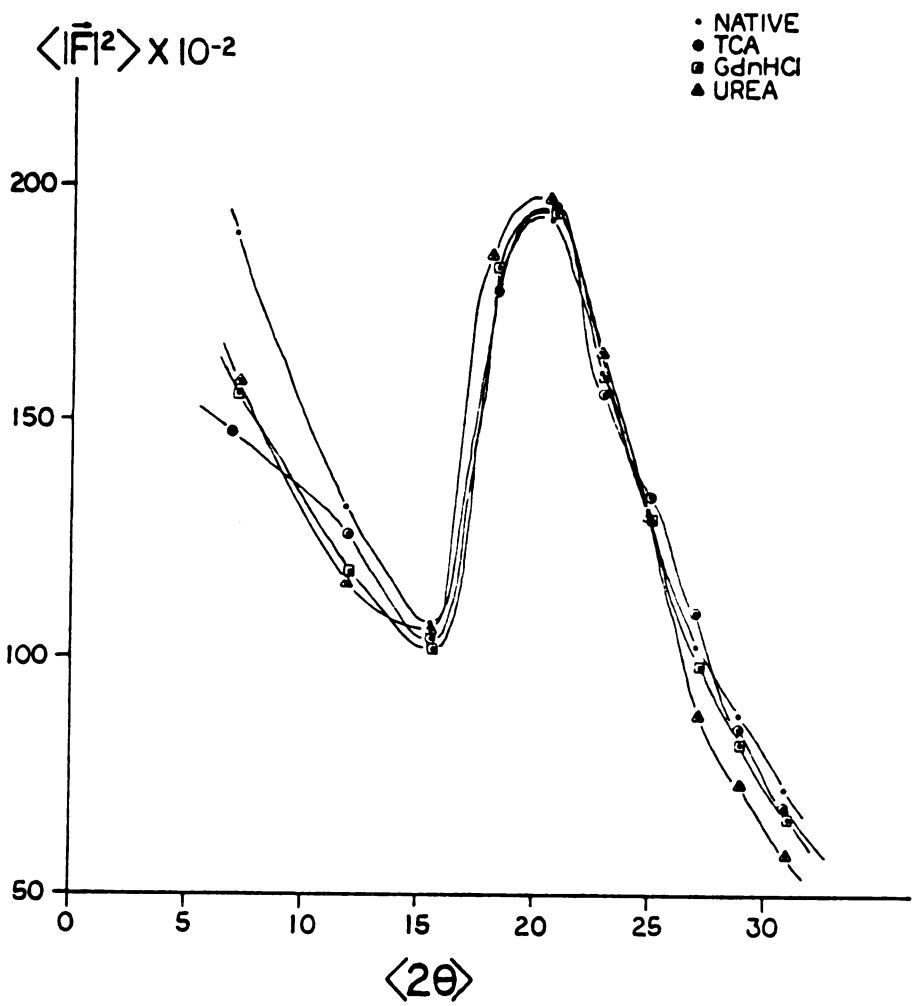


Figure 10. The distributions of average reduced derivative and native intensities in intervals of 2θ.

Chapter IV

The Difference Fourier Synthesis.

1. The Difference Fourier Synthesis.

The large volumes of solvent which exist in the protein crystal lattice enable small molecules or ions to diffuse throughout the lattice and interact with the protein molecules. The chemist may exploit this property to obtain information about the protein beyond that immediately deriveable from the structure of the native protein alone. The term native structure denotes the parent protein structure determined by the method of multiple isomorphous replacement at a particular set of solvent conditions. The term derivative structure will denote structures differing from the native due to changes in solvent composition or the binding of small molecules or ions regardless of whether or not a covalent bond is formed between the ligand and the protein. The changes in structure accompanying the formation of the derivative may be examined by the difference Fourier synthesis.

The electron density at a point (x,y,z), $\rho(x,y,z)$, in a unit cell of volume V in a monoclinic crystal is given by the Fourier summation

$$\rho(x,y,z) = \frac{4}{V} \sum_{h=-\infty}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \vec{F}(hkl) \exp\left[-2\pi i(hx+ky+lz)\right], \quad (1)$$

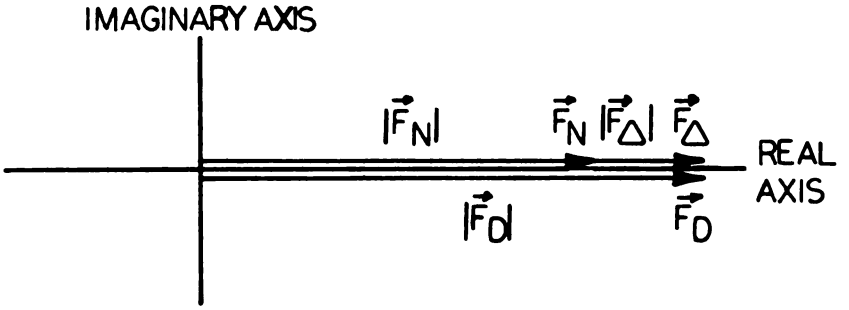
where each term contains a complex structure factor $\vec{F}(hkl)$ and the limits of the summation are such that the sum is confined to the unique reflections. A difference Fourier synthesis is used to compare a derivative electron density map to a native map,

$$\Delta\rho(x,y,z) = \frac{4}{V} \sum_h \sum_k \sum_l \left[\vec{F}_D(hkl) - \vec{F}_N(hkl) \right] \cdot \exp\left[-2\pi i(hx+ky+lz)\right], \quad (2)$$

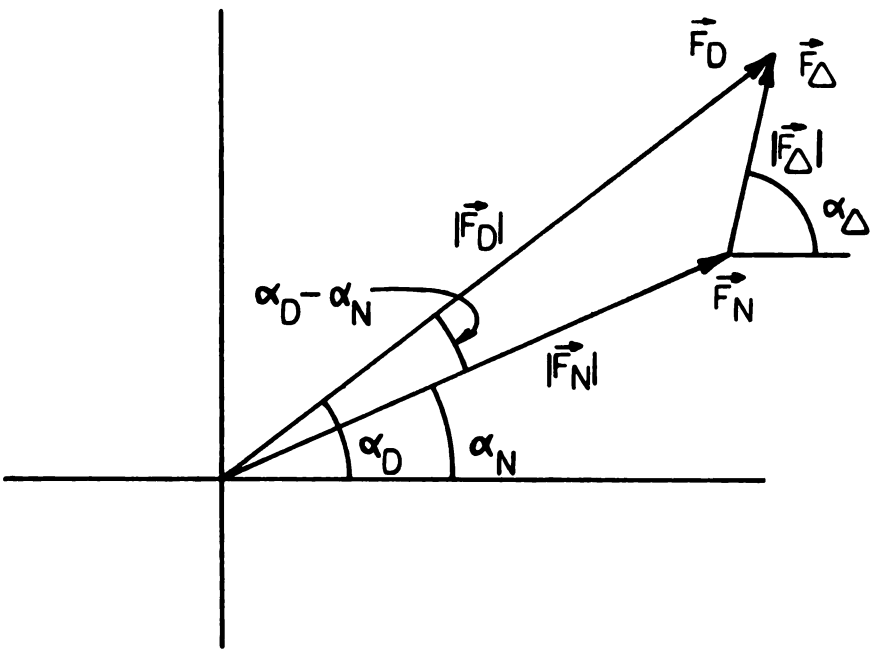
where the coefficients are the differences between the derivative structure factor, $\vec{F}_D(hkl)$, and the native structure factor, $\vec{F}_N(hkl)$. Figure 11 shows the vectors $\vec{F}_D$ and $\vec{F}_N$ for a given (hkl) in the complex plane with the phase angles α_D and α_N, and moduli $|\vec{F}_D|$ and $|\vec{F}_N|$, respectively. Equation (2) may be rewritten with the structure factors expressed as the products of structure amplitudes and phase factors

$$\Delta\rho(x,y,z) = \frac{4}{V} \sum_h \sum_k \sum_l \left[|\vec{F}_D| \exp(i\alpha_D) - |\vec{F}_N| \exp(i\alpha_N) \right] \cdot \exp\left[-2\pi i(hx+ky+lz)\right], \quad (3)$$

where the (hkl) notation has been dropped for convenience. The Fourier synthesis of equation (3) is the exact difference Fourier because both the derivative and native phases are included. In practice, however, α_D is not known and the assumption is made that since the differences between the derivative and the native structure amplitudes are small,



(a) The centrosymmetric case.



(b) The noncentrosymmetric case.

Figure 11. Argand diagram of the native and derivative structure factor vectors.

the corresponding differences between α_D and α_N will also be small. Therefore, α_N is used for α_D in equation (3). Further, each phase has associated with it a figure of merit, m, which is a measure of the error in α_N, and is used to weight the terms of the Fourier series to give the "best" electron density map. The "best" electron density has the lowest r.m.s. error with respect to a map calculated with exact phases.[57] The difference Fourier synthesis used in the derivative studies reported here correspond to the "best" difference electron density and are

$$\Delta\rho(x,y,z) = \frac{4}{V} \sum_h \sum_k \sum_l m\left[|\vec{F}_D| - |\vec{F}_N|\right] \exp(i\alpha_N) \cdot \exp\left[-2\pi i(hx+ky+lz)\right]. \quad (4)$$

2. Errors and Peak Heights in the Difference Fourier Synthesis.

There are three main sources of error in the difference Fourier--the error in the native phases, the errors in the intensity measurements, and the errors due to the use of the coefficients $\Delta F\exp(i\alpha_N)$ (where $\Delta F = |\vec{F}_D| - |\vec{F}_N|$) instead of the exact coefficients $\left[|\vec{F}_D|\exp(i\alpha_D) - |F_N|\exp(i\alpha_N)\right]$.

(a) Errors in the native phases. Henderson and Moffat[58] have shown that the average contribution of the difference coefficient $m\Delta F\exp(i\alpha_N)$ to the height of features in difference electron density maps is $m^2\Delta F$. The figure of merit is a function of 2θ-angle and varies among reflections according to the heavy-atom phasing so that only an approximate

average estimate of the error is possible. The average figure of merit for 2.8 A. resolution CHT phases with $m \geq 0.7$ is 0.89.

(b) Errors in the intensity measurements. The mean square error, δ^2, in an electron density map resulting from errors in the amplitudes is estimated by the formula[59,60]

$$\delta^2 = \frac{4}{V^2} \sum_h \sum_k \sum_l \sigma^2 (|\vec{F}|), \tag{5}$$

where $\sigma(|\vec{F}|)$ is the standard error in $|\vec{F}|$ and the summation is over the independent reflections for a noncentric monoclinic space group. For a difference Fourier with amplitudes $(|\vec{F}_D| - |\vec{F}_N|)$, the mean square error is

$$\delta^2_\Delta = \frac{4}{V^2} \sum_h \sum_k \sum_l \left[\sigma^2(|\vec{F}_D|) + \sigma^2(|\vec{F}_N|)\right]. \tag{6}$$

For difference coefficients to be significant, they must at least satisfy

$$(|\vec{F}_D| - |\vec{F}_N|) > \left[\sigma^2(|\vec{F}_D|) + \sigma^2(|\vec{F}_N|)\right]^{\frac{1}{2}}. \tag{7}$$

(c) Errors due to the use of inexact coefficients. Figure 11a illustrates the relationships among $\vec{F}_N$, $\vec{F}_D$, and $\vec{F}_\Delta$ for the centrosymmetric case where the phase angles are restricted to the values 0 and π. Changes of phase upon derivative formation are rare for centric reflections so that α_D and α_N are probably the same. The coefficient, $\Delta F \exp(i\alpha_N)$ is then the exact difference structure factor.

In addition, the figures of merit for most centric reflections are close to 1.0 so the centric terms in equation (4) contribute at full weight and exactly to the difference density.

For noncentrosymmetric reflections, the situation is more complicated. Referring to Figure 11b, the structure amplitudes $|\vec{F}_D|$, $|\vec{F}_N|$, and $|\vec{F}_\Delta|$ are related by

$$|\vec{F}_D|^2 = |\vec{F}_N|^2 + |\vec{F}|_\Delta^2 - 2|\vec{F}_N||F_\Delta| \cos\left[\pi - (\alpha_\Delta - \alpha_N)\right]. \quad (8)$$

Rearranging (8) and using the relation

$$\cos(\alpha_\Delta - \alpha_N) = \tfrac{1}{2}\left[\exp\{i(\alpha_\Delta - \alpha_N)\} + \exp\{-i(\alpha_\Delta - \alpha_N)\}\right],$$

an expression for the difference Fourier coefficient is obtained[61] from which several observations can be made.

$$\Delta F \exp(i\alpha_N) = \frac{|\vec{F}_\Delta|^2}{|\vec{F}_D| + |\vec{F}_N|} \exp(i\alpha_N) \qquad \text{(i)}$$

$$+ \frac{|\vec{F}_N||\vec{F}_\Delta|}{|\vec{F}_D| + |\vec{F}_N|} \exp(i\alpha_\Delta) \qquad \text{(ii)} \quad (9)$$

$$+ \frac{|\vec{F}_N||\vec{F}_\Delta|}{|\vec{F}_D| + |\vec{F}_N|} \exp\left[i(2\alpha_N - \alpha_\Delta)\right] \qquad \text{(iii)}$$

On the average, $|\vec{F}_N|$ and $|\vec{F}_D| >> |\vec{F}_\Delta|$ and since $|\vec{F}_D| \simeq |\vec{F}_N|$, term (i) will not contribute significantly to the electron density. Term (iii) is approximately

$$\frac{|\vec{F}_\Delta|}{2} \exp\left[i(2\alpha_N - \alpha_\Delta)\right]$$

and should simply contribute to the background since α_N and α_Δ are uncorrelated and their difference is, in general, a random number. Term (ii) is approximately

$$\frac{|\vec{F}_\Delta|}{2} \exp(i\alpha_\Delta)$$

and makes the greatest contribution to the difference electron density. The form of term (ii) indicates that the peak heights in the difference Fourier using the coefficients $\Delta F \exp(i\alpha_N)$ will be about one-half the heights of peaks in an exact difference Fourier. This is consistent with the observation by Luzzati[62] that the peak heights of atoms in electron density maps not included in the phasing will appear with peak heights approaching about one-half their true height.

Blundell and Johnson conclude that the error due to the use of approximate coefficients is proportional to the r.m.s. ΔF.[61] Henderson and Moffat[58] considered the problem in detail, but the formula they derived to estimate the error in difference maps has since been shown to calculate values that are low by a factor of 2. Ford, et.al.,[63] report a corrected formula for the mean square error in difference maps calculated with the coefficients $m\Delta F \exp(i\alpha_N)$ compared to maps calculated with the true coefficients $\frac{1}{2}\left[|\vec{F}_D| \exp(i\alpha_D) - |\vec{F}_N| \exp(i\alpha_N)\right]$ which is

$$\langle \Delta\rho^2 \rangle = \frac{4}{V^2} \sum_h \sum_k \sum_l \left[\frac{(\Delta F)^2}{2} + \frac{\delta_\Delta{}^2}{4} \right], \qquad (10)$$

where $\delta_\Delta{}^2$ is the mean-square error in ΔF due to errors in intensity measurements given by equation (6), and where the summation is over the unique monoclinic reflections.

An upper limit to the error in the difference map is the observed r.m.s. difference density

$$\langle \Delta\rho^2 \rangle_{obs}^{\frac{1}{2}} = \frac{1}{V} \left[\sum_x \sum_y \sum_z \Delta\rho(x,y,z)^2 \right]^{\frac{1}{2}} \qquad (11)$$

calculated over the unit cell.

Finally, a discussion of the errors in the difference Fourier should include the observation that the difference Fourier is a sensitive function whose error level is normally a small fraction of that of the parent protein electron density. This is partly due to the cancellation of series termination errors. Secondly, the error in a Fourier series is in general proportional to the magnitudes of the coefficients and because ΔF is on the average much smaller than either $|\vec{F}_N|$ or $|\vec{F}_D|$, the error in the difference Fourier is correspondingly less than in either the $\vec{F}_N$ or the $\vec{F}_D$ electron density maps. Difference Fouriers commonly contain significant features with peak heights that are less than one-tenth the features in the parent electron density.[58]

Chapter V.

The 0.01 M. Trichloroacetate Derivative.

1. The 0.01 M. Trichloroacetate Difference Electron Density.

The difference electron density map of the 0.01 M. TCA derivative revealed several unexpected features. First, there is no evidence in this map that the structures of the CHT molecules are significantly affected by the presence of the TCA anion even though the b-axis shrinks by 0.45 A. There are only three significant difference peaks in the asymmetric unit, all positive, and all suggesting binding of TCA to the protein. There are no significant negative peaks to indicate movement by any part of either CHT molecule in response to the TCA. Secondly, two of the binding sites are in the active site regions of Molecule I and I' near the substrate specificity sites in which the aromatic side chains of a substrate protein are located. These two difference peaks are approximately related by the local two-fold rotation axis. The other binding site is on the surface of Molecule I. This site is intermolecular and is formed by amino acid side chains of surface residues of both Molecule I and a neighboring CHT molecule. The unit cell coordinates and heights of these peaks are given in Table 2. The observed r.m.s. $\Delta\rho$ calculated over a featureless region of the asymmetric unit is 0.03 $eA.^{-3}$.

Peak 1, the TCA binding site on the surface of Molecule I, is 3.0-4.0 A. from the carbonyl group of Thr 134 and 4.0-5.0 A. from the pyrrolidine ring of Pro 161. The

Table 2. The Most Significant Features in the 0.01 M. TCA Difference Map

Peak	X	Y	Z	Height (e.A.$^{-3}$)	Comments
1	0.368	0.057	0.192	0.30	Surface of Molecule I
2	0.632	0.283	0.385	0.27	Active Site of Molecule I
3	0.645	0.330	0.947	0.29	Active Site of Molecule I'

shape of the peak is roughly spherical and does not suggest an orientation for the TCA anion. This peak does not have a local two-fold equivalent in Molecule I', as the intermolecular binding site itself does not have a local two-fold related counterpart. This behavior is similar to that of the uranyl heavy atom derivative used in the determination of the 2.8 A. phases (binding site U, Table 3, Reference 41).

The two active site substitution peaks are also approximately spherical, but show a slight tapering along the positive direction of the a* axis (local two-fold direction). Kendrew models of the TCA anion were constructed and placed into the difference peaks using a Richards' box[64] with the carboxyl groups in the tapered side of the peaks. The TCA anions are approximately related by the local dyad both with respect to location and with respect to orientation of the anions. The TCA corresponding to Peak 2 in Molecule I is oriented such that the CCl_3 group is about 3.0 A. from the peptide bond joining Trp 215 and Gly 216 and the carboxyl group is 2.0-3.0 A. from the main chain atoms of Cys 191 and Met 192 and 2.0-3.0 A. from the SO_4^{2-}-195 group. In Molecule I', the TCA of Peak 3 is oriented such that the CCl_3 group is about 5.0 A. from the α-carbon atom of Gly' 216 and the TCA carboxyl group is 2.0-3.0 A. from both the SO_4^{2-}-195' and the γ-OH of Ser' 195. The carboxyl group is also about 3.0 A. from the main chain atoms at Ser' 195. In both binding sites, the TCA carboxyl groups

closely approach carbonyl groups of the protein. However, both CHT molecules accomodate the TCA anions without any apparent difficulty as the difference map shows no evidence of movement by any part of either CHT molecule.

The pK_a of TCA is about 1 (25° C.) so that the carboxyl group is charged at pH 3.6. The forces responsible for the binding may involve either hydrogen bonding or ion-dipole interactions between the TCA carboxyl group and the neighboring protein carbonyl groups, and hydrogen bonding interactions with the sulfate groups. Van der Waals interactions are probably responsible for orienting the CCl_3 group near the main chain atoms at Trp 215 and Ser 216.

2. The 0.01 M. Trichloroacetate Difference Diagonal Plot.

The representation of the difference density on a diagonal plot, or the difference diagonal plot (hereafter abbreviated as DDP), is a two-dimensional representation of the changes occurring in the three-dimensional difference map and is an extension of the concept of the distance diagonal plot described in Chapter II. The DDP wasconceived and originally used by Dr. M. N. Liebman[46] of our laboratory. The DDP is an NxN matrix whose elements r_{ij} are the moduli of the vectors $\vec{r}_{ij}$ which are

$$\vec{r}_{ij} = (\vec{p} - \vec{C\alpha}_i) + (\vec{p} - \vec{C\alpha}_j)$$

where $\vec{C\alpha}_i$, $\vec{C\alpha}_j$, and $\vec{p}$ are the coordinate vectors of the i-th and the j-th alpha-carbon atoms and a difference map peak, respectively, in an orthogonal coordinate system.

The number N designates the amino acid residue. For a difference map containing m peaks, there will be m vectors $\vec{p}$ and m vectors $\vec{r}_{ij}$ for each ij-pair of α-carbon atoms. In our DDP program only the smallest of the moduli of the m vectors $\vec{r}_{ij}$ is recorded. The plot thus calculated may then be contoured at convenient distance intervals. The features which appear on the DDP represent the close approach (below an arbitrary limit) of a difference peak to a given pair of α-carbon atoms, and are a unique two-dimensional representation of the three-dimensional difference density associated with the molecule. Because the vectors $\vec{r}_{ij}$ and $\vec{r}_{ji}$ are of equal length, only one-half of the NxN matrix is unique. Therefore, the full matrix may be used to represent changes occurring in both CHT molecules of the dimer. In all the DDP's calculated in our work, the area below the diagonal is devoted to Molecule I and the area above the diagonal to Molecule I'. The α-carbon coordinates of Molecule I' were obtained by performing a two-fold rotation about the local dyad A of Figure 3a on the α-carbon atom coordinates of Molecule I.

Since the TCA difference map was relatively simple, the DDP of TCA is correspondingly simple and is shown in Figure 12. The TCA DDP is normally contoured at 4.5, 6.0, and 9.0 A. levels, but only the 9.0 A. features are shown in Figure 12. The vertical and horizontal lines dividing the plot into quarters are drawn in at residue 123 and denote the folding domains A (residues 1-122) and B

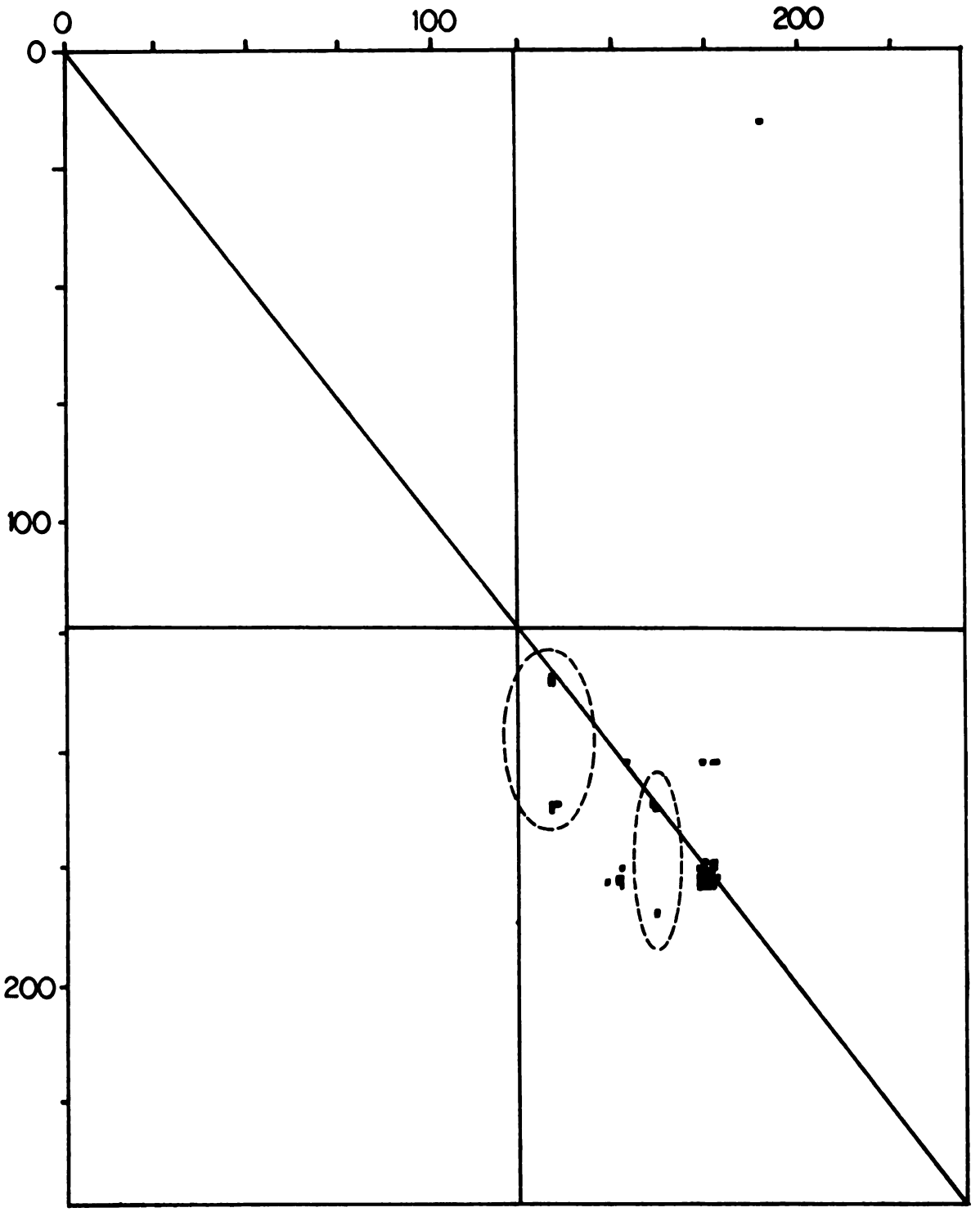


Figure 12. The difference diagonal plot of the 0.01 M. TCA derivative.

(residues 123-245) of CHT. Features in the DDP which are reflected across the diagonal represent changes occurring in both CHT molecules at the same amino acid residues. DDP features not having a diagonal-mirror plane equivalent represent changes occurring in one CHT molecule but not the other, and are usually related to the variability in the native structure. The TCA DDP contains examples of both kinds of features.

The two difference peaks in the active site regions of Molecules I and I' are the dominant features in the DDP and appear in the lower-right quarters of the plot. The two features on both sides of the diagonal at residues 191 to 196 represent the close approach of both TCA difference peaks to the α-carbon atoms of the polypeptide segment Cys 191-Met 192-Gly 193-Asp 194-Ser 195-Gly 196. The two features are not exact mirror images since neither the CHT molecules nor the two difference peaks are exactly alike. However, the similar binding behavior of the two TCA molecules gives rise to two very similar features in the DDP. The two off-diagonal features nearby also show an approximate symmetry with respect to the diagonal. Thus, similar structural changes are occurring in the two CHT molecules.

The TCA substitution on the surface of Molecule I, Peak 1, creates the group of features in the diagonal plot enclosed by broken lines. These features are contoured at 12.0 A. level. Peak 1 is located in the secondary binding site observed for the phenylalanine-CHT derivative,

and the pattern of peaks observed for this TCA substitution is very similar to those created by the Phe secondary substitution in the Phe-CHT DDP. Clearly, the DDP provides a convenient method for comparing or contrasting the binding behavior of small molecule substituents.

The DDP provides a quick and concise means for displaying changes observed in a difference Fourier. As will be seen with CHT, the DDP dramatically demonstrates structural variability in CHT derivatives. In addition, it may be used to estimate the location and extent of changes occurring in the derivatives. One difference peak near several α-carbon atoms, or a group of peaks near several α-carbon atoms, will result in more extended DDP features than those arising from a single peak near a small number of α-carbon atoms. Examples of both extended and localized interactions are found in the DDP of TCA.

Chapter VI.

The 2.0 M. Guanidine Derivative.

1. The Difference Fourier Map.

Approximately 60 peaks with heights $\geq$ 0.20 e.A.$^{-3}$ (4σ) were observed in the 2.0 M. GdnHCl difference map calculated over the CHT dimer and the surrounding solvent regions. Of these, about 30 peaks (both positive and negative) were located in or near the two CHT molecules 4.5 A. or less from an α-carbon atom as determined by a DDP calculation. The remaining peaks were located entirely in the intermolecular contact regions shared with neighboring CHT molecules; no significant density was observed in the solvent regions. The latter further enhanced the credibility of the map. The observed changes occurred in two regions--the dimer interface region and on the surface of the dimer. Most of the peaks were small in volume and occurred singly at scattered locations. The interiors of the molecules were essentially unchanged from the native structure.

Due to the nature of the changes displayed in the difference map, a detailed interpretation of many of the peaks was impossible. There were, however, several examples of extended density with relatively large peak heights (0.2-0.3 e.A.$^{-3}$) in both the dimer interface region and on the surface of the dimer. These were examined in some detail.

The observed r.m.s. of the difference density of the 2.0 M. GdnHCl map calculated over the entire asymmetric unit

is 0.05 e.A.$^{-3}$. Only those peaks with heights $\geq$ 0.20 e.A.$^{-3}$ (4σ) were examined.

2. The Difference Diagonal Plot.

The DDP calculated for all the difference peaks (63) in the asymmetric unit with heights $\geq$ 0.20 e.A.$^{-3}$ is shown in Figure 13. The DDP has been contoured only at the 9.0 A. level as before. Since the GdnHCl derivative is more complex than the TCA derivative, its DDP is correspondingly more complicated than that of TCA. However, several qualitative observations may be made immediately. The observed changes are wide-spread with respect to the amino acid sequence over both molecules in the dimer. Further, the changes occur almost entirely within one domain or the other since there are very few features in the lower-left and the upper-right quarters of the plot which would indicate changes involving amino acid residues of both domains. There are also several DDP features without a diagonal-mirror counterpart, reflecting asymmetric response and the structural variability between the independent CHT molecules.

The DDP can give an indication of whether key amino acid residues were affected by derivative formation. Some key residues include those involved in catalytic action, dimer formation, and other residues which proved to be prominent in earlier derivative studies. The 2.0 M. GdnHCl DDP contains numerous examples of structural changes near key residues. Difference peaks near His 40 and His' 40 gave rise to a set of DDP features in Figure 13 labelled

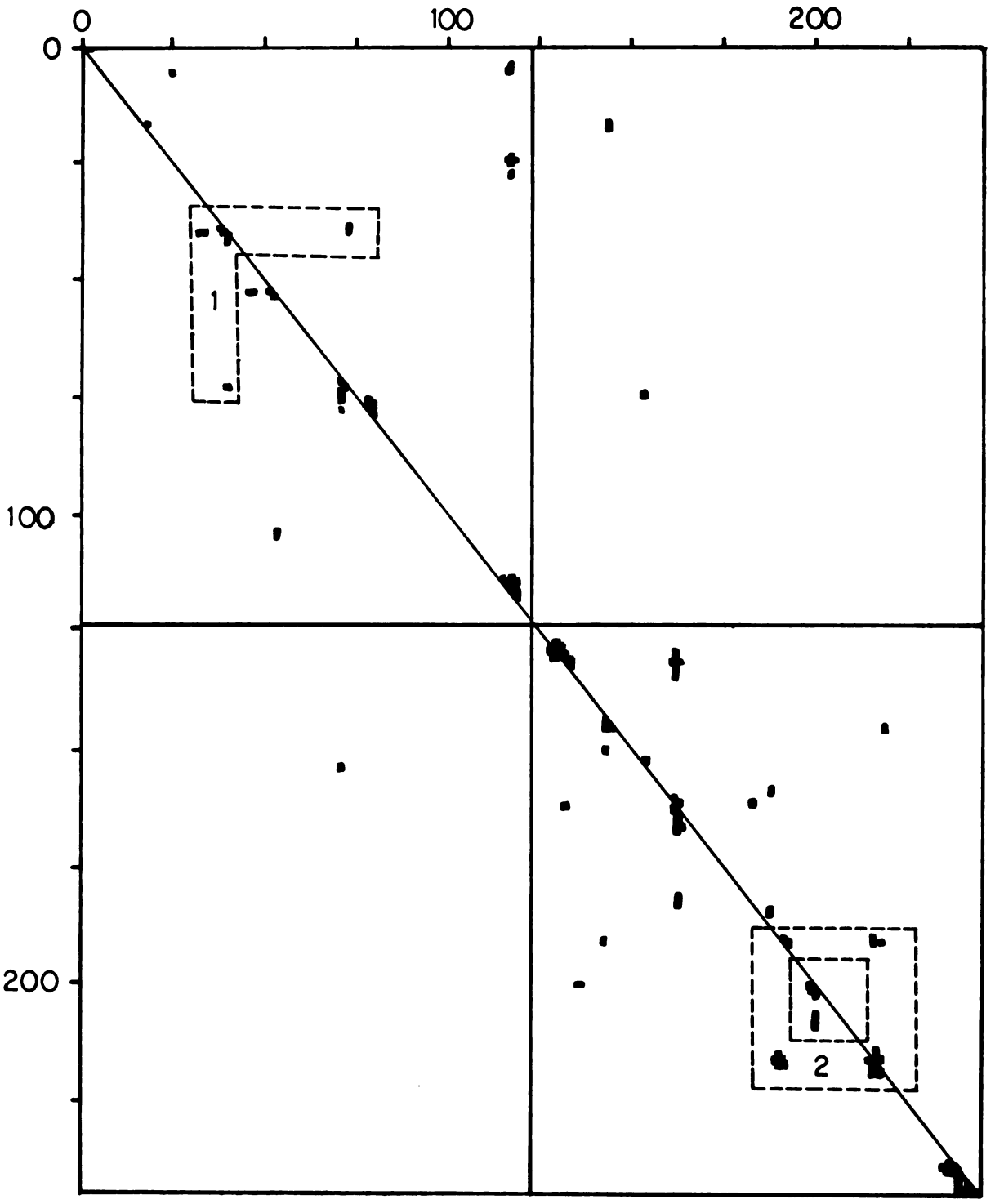


Figure 13. The difference diagonal plot of the 2.0 M. GdnHCl derivative.

"1". The aromatic substrate binding site is partly formed by two polypeptide segments located near one another--Ser 190-Cys 191-Met 192-Gly 193-Asp 194 and Trp 215-Gly 216-Ser 217. A characteristic set of features appear in the DDP for difference peaks occurring either in the substrate binding site (as in the TCA derivative) or in the dimer interface region near the substrate binding site (as in the GdnHCl derivative). These are the features in the rectangular array in Figure 13 labelled "2". There are other characteristic arrays which occur for difference peaks located in other parts of the CHT molecule.

The qualitative results of the DDP are readily-obtained and may be useful in characterizing, for instance, the members of a set of related CHT derivatives.[46,65] However, upon closer examination of the DDP and comparison with the difference Fourier, it is apparent that extensive DDP features do not always represent the most significant features in the difference map, and conversely, the most significant features in the difference map do not always create spectacular and extended features in the DDP. One of the most extensive DDP features in GdnHCl involves the C-terminal α-helix of Molecule I, residues 234 to 245 (see Figure 13). The features result from two small difference peaks--approximately 1.0 A. in diameter at the 0.2 e.A.$^{-3}$ contour--situated inside the α-helix. These two peaks do not comprise either significant or extended difference density features. Also, several very significant difference

peaks occurring on the surface of Molecule I in the uranyl binding site region[41] give rise to only a single point in the DDP. Thus, the DDP does not always represent the significance of difference peaks accurately.

This problem is partly alleviated by the use of the difference diagonal multiplicity plot (DDMP). This is an NxN matrix whose elements m_{ij} are the number of difference peaks for whom the inequality, $r_{ij} \leq k$, where k is an arbitrary limit, is satisfied for each ij-pair of α-carbon atoms. The quantities r_{ij} are calculated for the DDP as described in Chapter V. For each element r_{ij} which appears in the DDP, there is a corresponding element m_{ij} in the DDMP. For the 2.0 M. GdnHCl DDP, the limit of r_{ij} is 9.0 A. and each corresponding m_{ij} of the DDMP is the number of difference peaks whose distances to the i-th and the j-th α-carbon atoms add up to 9.0 A. or less. The number of difference peaks close to the ij-pair of α-carbons (below an arbitrary limit)--a multiplicity--gives an additional measure of the extent of change. The DDMP for the 2.0 M. GdnHCl derivative is shown in Figure 14, where only those elements $m_{ij} \geq 2$ have been plotted.

Comparing the GdnHCl DDMP to the DDP reveals two sets of DDP features containing points with peak multiplicities of 2 or more. Parts of four of the features in group "2" in the DDP (Figure 13) are due to groups of two or three peaks. The identity and map coordinates of each peak involved is listed along with the DDMP so that features in

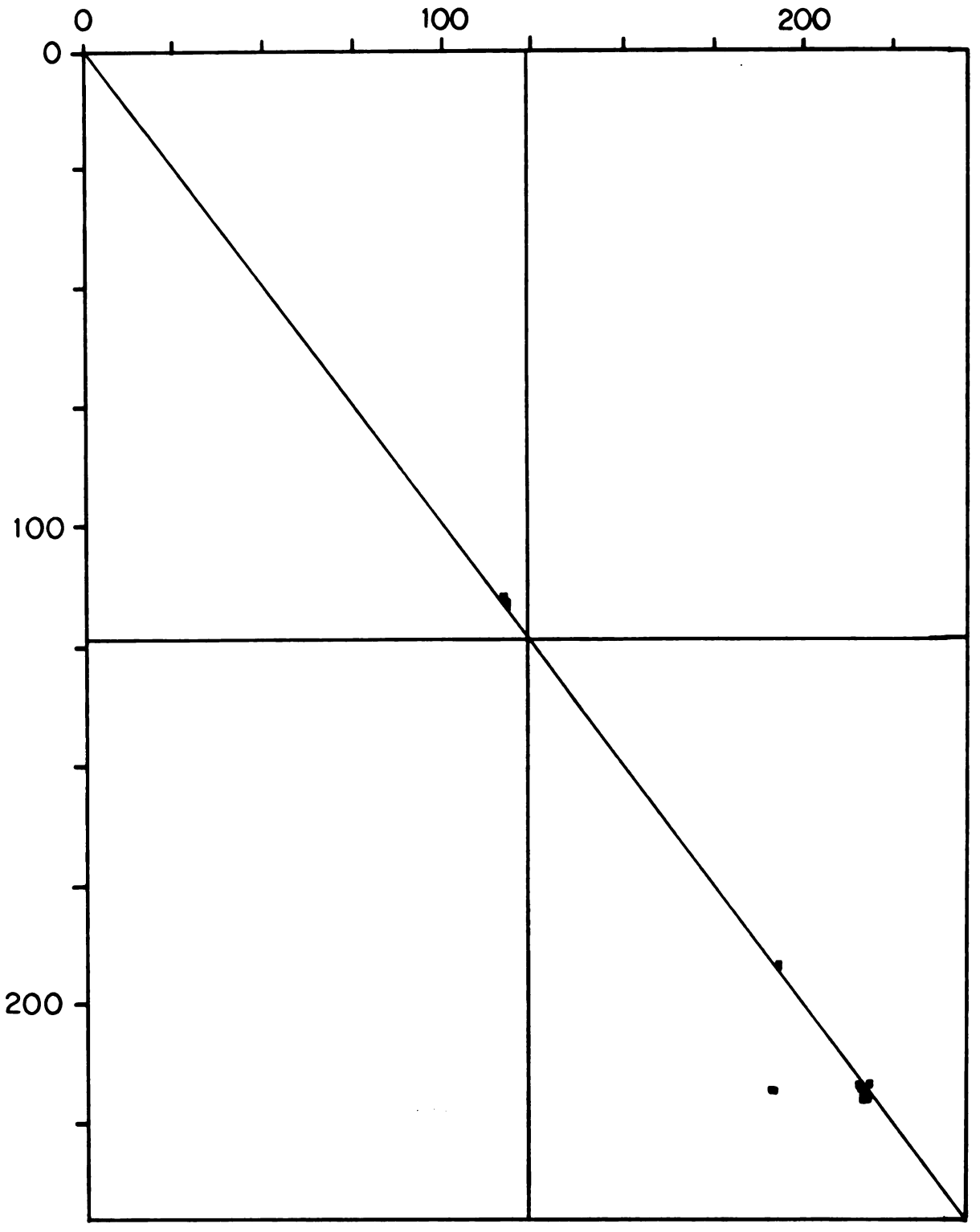


Figure 14. The difference diagonal multiplicity plot of the 2.0 M. GdnHCl derivative.

the DDP may be cross-referenced with peaks in the difference map. The peaks giving rise to multiplicities of 2 or more do in fact constitute the most significant set of changes in the dimer interface region. The only other feature in the DDMP occurs on the diagonal at residues Ser' 115 to Thr' 117 indicating that significant changes have occurred around this polypeptide segment which is located on the surface of Moleule I'.

The DDMP does not solve the problem of significant difference density occurring on the surface of the protein molecule which fails to contribute significantly to the DDP features. By raising the limit of r_{ij}, these surface peaks would contribute more heavily, but the DDP features already present would become much more enhanced with the concomitant loss of some of the discriminating ability of the DDP. The 9.0 A. limit is an optimal value determined by calculating DDP's at various limits. Thus, significant density occurring on the surface of the CHT molecule may be unavoidably underweighted by the DDP.

3. Changes in the Dimer Interface Region.

The dimer interface region contains a number of strained close contacts and is very sensitive to changes in the solvent environment.[38,67] This region is also notable for numerous features in the native electron density which do not correspond to protein groups and have been attributed to ordered solvent molecules and ions. The locations of four sulfate ions in the dimer interface region

have been unambiguously determined by a SO_4^{2-}:SeO_4^{2-} exchange experiment.[66] The unit cell coordinates and peak heights of the most significant difference peaks in the dimer interface region are listed in Table 3.

Peak 1 is about 3.5 A. in length (at the 0.20 e.A.$^{-3}$ contour) extending from a point 1.0 A. from the sulfur atom of Met 192, through the local two-fold axis, to a point 1.0 A. from the sulfur atom of Met' 192. Peak 2 is 3.0 A. from Peak 1 and 1.0 A. from SO_4^{2-}-195'. There is no negative density occurring on the native density at the methionine side chains, so there is no direct evidence that the side chains of either Met 192 or Met' 192 have moved. The two methionine sulfurs are 4.0 A. apart in the native dimer and are connected in the native electron density at the 0.5 e.A.$^{-3}$ contour, and it is in this density that Peak 1 is located. There is a cresent-shaped positive difference peak 2.0 A. from Peak 1 in the negative x-direction ranging in height from 0.16 to 0.20 e.A.$^{-3}$. This peak extends in an arc 6.0 A. in length (at the 0.16 e.A.$^{-3}$ contour) from a point 2.0 A. from the sulfur of Met' 192, passing within 1.0 to 2.0 A. of the sulfur of Met 192, and ending at a point 2.0 A. from SO_4^{2-}-195'. The total picture presented by Peaks 1 and 2 and the cresent-shaped positive peak is that of a reorganization of the solvent structure around the methionine side chains and SO_4^{2-}-195'. The native density joining the methionine side chains is probably due to close contact of the sulfur atoms, and the location of Peak 1

Table 3. The 2.0 M. GdnHCl Derivative - the Most Significant Difference Peaks in the Dimer Interface Region.

Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
1	0.632	0.292	0.469	-0.27	1.0 A. from Met 192 and Met' 192 side chains
2	0.618	0.302	0.523	-0.26	1.0 A. from SO_4^{2-}-195'
3	0.737	0.292	0.492	-0.27	3.0 A. from side chain of Leu' 143
4	0.776	0.292	0.515	+0.32	Double peak in solvent cavity in dimer
5	0.789	0.302	0.469	+0.29	interface
6	0.908	0.264	0.438	-0.25	1.0 A. from His 40 imidazole
7	0.921	0.330	0.585	-0.25	1.0 A. from His' 40 imidazole

suggests that these side chains may have been slightly disturbed.

Peak 3 is an oval peak 3.0 A. in length partly occupying a segment of native density which has been tentatively assigned to a counterion for SO_4^{2-}-195'. It is about 3.0 A. from the side chain of Leu' 143, 2.0 A. from the side chain of Met 192, and 2.0 A. from both Peaks 4 and 5. Peaks 4 and 5 are positive peaks whose centers are 2.0 A. apart but are joined at the 0.20 e.A.$^{-3}$ contour to form an extended segment of positive density whose projection in the yz-plane has the shape of an inverted "V", spanning a distance of about 5.0 A. at its widest part, and forming one of the most interesting features in the 2.0 M. GdnHCl difference map. Peaks 4 and 5 are shown in Figure 15. Peaks 4 and 5 are located in a solvent cavity in the dimer interface which is bounded by the side chains of Phe' 39, Thr 151, and Leu 143, and by the main chain segments His 40-Phe 41, Met 192-Gly 193, and His' 40-Phe' 41-Cys' 42. Peak 4 is an oval-shaped peak 3.0 A. in length, 3.0 A. from the peptide group joining Met 192 and Gly 193, and 3.0 A. from the main chain at His 40. Peak 5 is also 3.0 A. in length and is located 2.0 A. from the native peak attributed to the SO_4^{2-}-195' counterion and about 3.0 A. from the main chain atoms at Cys' 42. A detailed interpretation of the changes represented by Peaks 3, 4, and 5 is difficult. Since none of the surrounding protein groups appear to have been perturbed, these peaks must represent changes in the solvent. Either

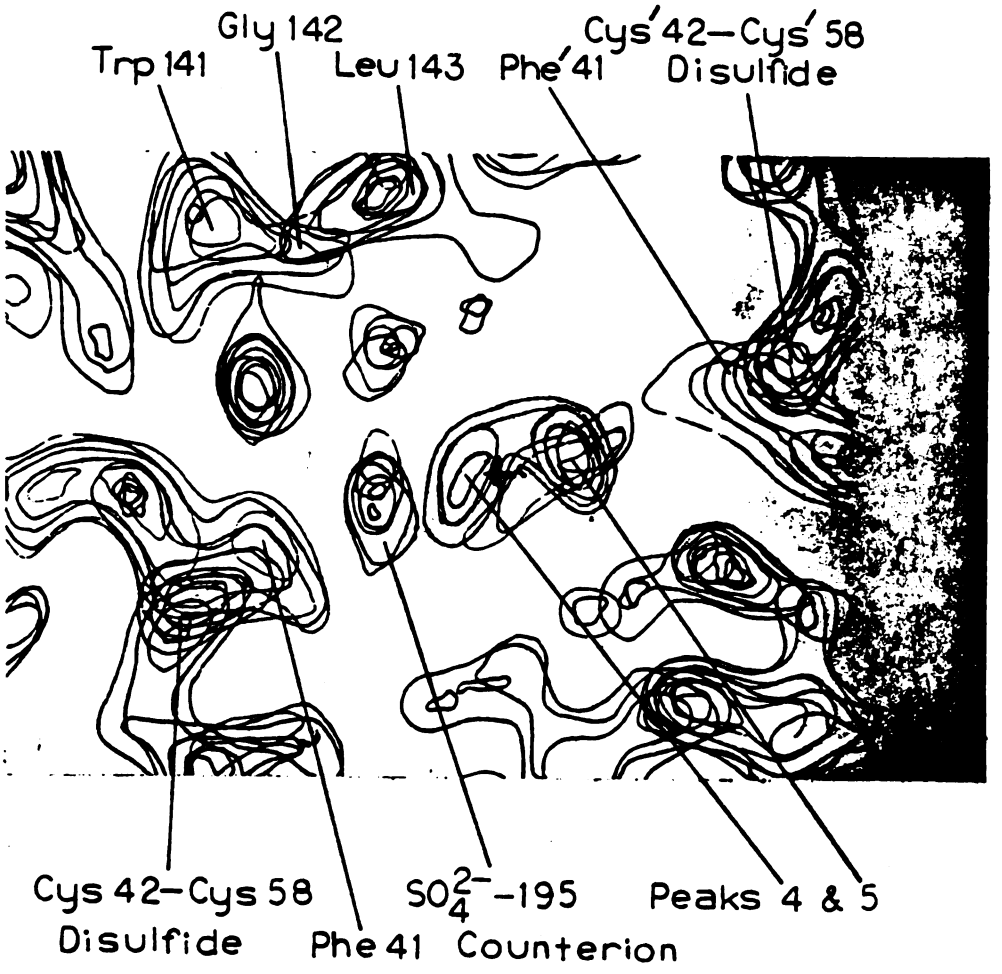


Figure 15. A feature of the 2.0 M. GdnHCl difference map located in the dimer interface region.

Peak 4 or 5 could represent a guanidinium ion (hereafter abbreviated Gdn^+) as there is adequate room for the ion and numerous hydrogen bonding possibilities are available near either peak; however, little information is available from the peak's shapes to suggest an orientation for a Gdn^+ ion.

Peaks 6 and 7 represent changes in the imidazole side chains of His 40 and His' 40, respectively. Peak 6 is partly superimposed in the His 40 imidazole density, 1.0 A. from the Nδ1 atom. Peak 7 is 1.0 A. from the β-carbon atom of His' 40. There is no significant accompanying positive difference density for either negative peak. Histidine imidazolium ions have a pK_a of about 6.5 so that at pH 3.6, the imidazole groups both bear a positive charge. The difference peaks may represent some disruption of the solvent structure around the charged side chains. Alternatively, and more probably, the difference peaks represent small movements by the imidazole side chains away from the interface region and toward the interiors of their respective CHT molecules. If either of Peaks 4 or 5 are bound Gdn^+ ions the movement of the histidine side chains would be in response to the close approach of a like charge--Peak 4 is 5.0 to 6.0 A. from the His 40 side chain and Peak 5 is about 5.0 A. from the imidazole of His' 40.

Another group of changes were observed in the dimer interface region. There are several small difference peaks (<2.0 A.in diameter at the 0.20 $e.A.^{-3}$ contour) occurring

in and near the native electron density of the polypeptide segment Gly 216-Ser 217-Ser 218 in Molecule I and the corresponding segment of Molecule I'. These polypeptide segments display marked deviations from local two-fold symmetry in the native structure.[42] This region also contains 6 hydroxyl groups from various side chains in a volume about 6.0 A. in diameter. It may be significant that a region of the dimer which is crowded and stained, and which may have a complicated solvent structure because of the hydroxyl groups, should be affected by a chemical denaturant.

A DDP calculated for the difference peaks in the dimer interface region only is shown in Figure 16, contoured at the 9.0 A. level. The features created by the same difference peak, or by the same group of peaks, have been grouped together by broken lines. Groups 1 and 2 represent the changes occurring near His 40 (Peak 6) and His' 40 (Peak 7), respectively. Peaks 1, 2, 3, 4, and 5 cannot be separated because they all occur close together and near a number of the same α-carbon atoms. The features labelled "3" represent the close approach of Peaks 1 and 3 to the α-carbon atoms of Met 192, Gly 193, Ser 217, and Ser 218. The group labelled "4" arises from the close approach of Peaks 1, 2, 3, 4, and 5 to the polypeptide segments Cys 191-Met 192-Gly 193 and Trp 215-Gly 216-Ser 217-Ser 218. From this DDP, it is possible to isolate the contribution of difference peaks in the dimer interface region. The characteristic arrays of DDP features described in Section 2 are also

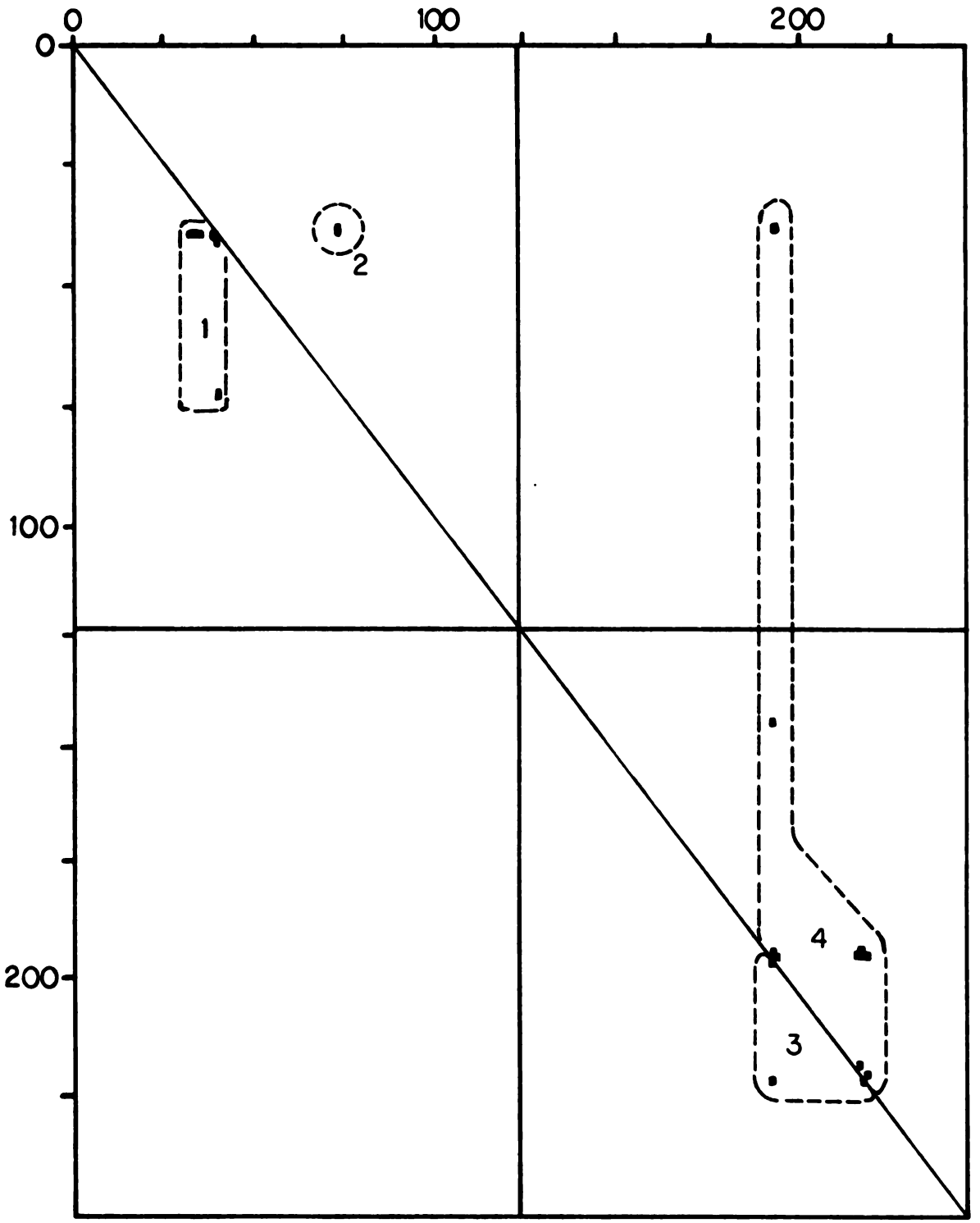


Figure 16. The difference diagonal plot of the 2.0 M. GdnHCl derivative, calculated only for the difference peaks in the dimer interface region. Distances r_{ij} calculated for each map point $\geq$ 0.2 e.A.$^{-3}$ within each peak.

evident here. Finally, the changes show an approximate local two-fold symmetry, but it is clear that the difference peaks in the dimer interface region are closer to the main chain atoms of Molecule I'. The full implication of this result is not clear at this time.

4. Changes on the Dimer Surface.

Most of the difference peaks observed on the surface of the CHT dimer with peak heights of 0.20 e.A.$^{-3}$ or more are small in volume and have no significant accompanying difference density of the opposite sign. These probably represent small reorientations of side chains, changes in the solvent structure surrounding a group of atoms, or small movements by main chain atoms, and are not interpretable in more detail. There are, however, several examples of more extended difference density peaks. The most significant difference features appearing on the dimer surface are listed in Table 4.

Peaks 4, 5, and 6 are two-fold screw equivalent peaks of 1, 2, and 3, respectively, occurring in two intermolecular contact regions which are two-fold screw related. Peak 1, about 4.0 A. in length, is located 2.0 A. from the carboxyl group of Glu 21 and about 3.5 A. from the Gdn^{+} group of Arg 154. Peak 4, identical to 1 in size and shape, occurs 2.0 A. from the carboxyl group of Asp' 153 and 5.0 A. from the carboxyamide group between Gly' 74 and Ser' 75. Peak 2 is less than 2.0 A. wide at the 0.20 e.A.$^{-3}$ contour, but is a "footprint"-shaped peak 5.0 A. from heel to toe

Table 4. The 2.0 M. GdnHCl Derivative - the Most Significant Difference Peaks on the Surface of the CHT Dimer.

Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
1	0.947	0.019	0.392	+0.33	2.0 A. from Glu 21 CO_2^--group
2	0.947	0.038	0.423	-0.24	1.0 A. from Arg 154 Gdn^+ side chain
3	0.066	0.000	0.392	-0.26	5.0 A. from Arg 154 side chain; 4.0 A. from Pro 24 pyrrolidine ring
4	0.053	0.519	0.608	+0.33	2.0 A. from Asp' 153 CO_2^--group
5	0.053	0.533	0.573	-0.24	In solvent, 4.0 A. from Asp' 153 CO_2^--group
6	0.934	0.500	0.608	-0.26	1.0 A. from Gdn^+ side chain of Arg' 154
7	0.408	0.481	0.231	-0.27	In the native density at SO_4^{2-}-177

at the 0.16 e.A.$^{-3}$ contour. It is located on the edge of, and along the entire length of, the native electron density of the Arg 154 side chain. Peak 5, identical to Peak 2, is located such that its peak is about 4.0 A. from the Asp' 153 carboxyl group with the rest of the peak extending away from the surface of Molecule I' toward a neighboring CHT molecule. Peak 3, which is approximately 2.0 A. in length, is located in the electron density of a neighboring CHT molecule, 9.0 A. from the Arg 154 Gdn^+ group and 7.0 A. from the pyrrolidine range of Pro 24. Peak 6, identical to Peak 3, is situated in the native electron density of the Arg' 154 side chain.

This intermolecular binding site has been prominent in previous work. The point of highest electron density in Peak 1 is 1.0 A. from the location of the dominant intermolecular binding site of the uranyl heavy atom derivative (site "U" in Table 3, Reference 41) which was notable for its high occupancy and its lack of a local two-fold equivalent. The structure of this binding site has been described in some detail previously.[38] In the uranyl derivative, the UO_2^{2+} ion complexes with the carboxylate groups of Glu 21 of Molecule I and Asp 153 of the neighboring CHT molecule (itself the two-fold screw equivalent of Molecule I'). Only one such binding site exists per dimer because the local two-fold related region is incapable of binding an ion because the CHT molecules are farther apart. Interestingly, Peaks 4, 5, and 6 are located only about 5.0 A. from the

positions they would have occupied had they been related to Peaks 1, 2, and 3 by the local two-fold rotation.

The uranyl binding site is also noted as one of the major areas in which changes in structure have been observed with change in pH.[38,52] Apparently, as the pH is raised from 3.6 to 5.4, the ion-pair consisting of the Arg 154 Gdn^+ ion and a sulfate ion is disrupted and a new ion-pair is formed involving the Arg 154 Gdn^+ ion and the carboxyl group of Glu 21. Additional difference map features for the pH 5.4 derivative create a complicated picture of changes in both the protein and solvent structures in this region.

The locations of Peaks 1 through 6 afford the convenient identification of segments of the native density belonging to the nearest neighbor CHT molecule which contribute to the formation of this intermolecular binding site. Because Peak 6 is near the Arg' 154 side chain, Peak 3 must also be near an Arg 154 side chain on the surface of the CHT molecule which is the nearest neighbor to Molecule I in the negative y-direction. A Kendrew model of the polypeptide segment Asp 153-Arg 154-Leu 155-Gln 156 was constructed and fitted into the native density using a Richards' box.[64] The result gives the structure of this binding site - a cavity 6.0 to 8.0 A. in diameter defined by the side chains of two Arg 154 residues and by the carboxyl side chains of Glu 21 of Molecule I and Asp 153 of the neighboring CHT molecule. The Arg 154 Gdn^+ groups are about 6.0 A. apart and the Glu 21 carboxyl group is 6.0 A. from the carboxyl group of Asp 153. The cavity thus formed is open to the solvent on two sides.

The picture which emerges from the disposition of the Gdn HCl difference peaks, the native structure of the binding site, its past history as an effective ligand of large positive ions, and its sensitivity to solvent changes, is that a Gdn^+ ion is binding in the uranyl binding site. Peaks 1 and 4 probably represent the Gdn^+ ions; a model of the Gdn^+ ion is readily accommodated by Peak 1 and is located at reasonable distances from the nearby Asp' 153, Arg' 154, Glu 21, and Arg 154 side chains. Peaks 2 and 5 represent the movement of the positively-charged Arg 154 Gdn^+ group of Molecule I away from the bound Gdn^+ ion in the solvent, and toward the sulfate ion with which it is presumed to form an ion-pair in the native structure.[38] Peaks 3 and 6 represent small movements of the Arg' 154 side chain on Molecule I'. The Gdn^+ ion so situated is between the carboxyl groups of Glu 21 (Molecule I) and Asp' 153 (the neighboring CHT molecule which is the two-fold screw equivalent of Molecule I'), and approximately half-way between them forming a kind of intermolecular salt bridge. Peak 7 is located squarely on the native density which has been shown to be a sulfate ion located 2.5 A. from the terminal amino group of the Lys 177 side chain (SO_4^{2-}-177, Reference 66). This peak shows simply that the sulfate ion has been removed from its position in the native structure. Since there is no positive density nearby, a shift to a new position is not indicated.

Chapter VII.

The 3.0 M. Urea Derivative.

1. The Difference Electron Density Map.

The difference Fourier map calculated over the CHT dimer and the surrounding solvent regions contained a larger number of peaks than was observed in the GdnHCl difference map. Approximately 160 peaks with heights $\geq$ 0.2 e.A.$^{-3}$ (3.0σ) were observed in the map, of which approximately 120 peaks occurred at a distance of 4.5 A. or less from an α-carbon atom. Numerous changes occurred on the dimer surface, in the dimer interface region, but unlike the GdnHCl derivative, changes also occurred in the hydrophobic interior of both CHT molecules. No significant difference density was observed in the solvent, but numerous peaks were observed in the intermolecular contact regions shared by Molecules I and I' with the neighboring CHT molecules. Though widely-distributed, the difference density was not randomly-distributed, occurring in specific regions of the dimer. The urea difference map contains fewer examples of extended, individual peaks like those observed in the GdnHCl map. The urea derivative difference peaks tended to be smaller in volume and clustered together at various locations in the molecule. Like the GdnHCl difference map, a detailed interpretation of many of the peaks is not possible, but the location and distribution of the peaks is significant and they will be discussed by groups.

Groups of peaks were observed in the following regions: the dimer surface, the dimer interface region, the A-Chain/B-Chain contact region, and in two areas of the nonpolar interdomain region. The changes occurring in the interiors of the CHT molecules were studied in detail for Molecule I only.

The observed r.m.s. $\Delta\rho$ for the difference map is 0.07 e.A.$^{-3}$ and only peaks with heights $\geq$ 0.20 e.A.$^{-3}$ were examined and input to the DDP calculation.

2. The Difference Diagonal Plot.

The DDP forthe 3.0 M. urea derivative is shown in Figure 17; the corresponding DDMP is shown in Figure 18. The features appearing in the DDP correspond to difference peaks for which $r_{ij} \leq 9.0$ A. The features appearing in the DDMP correspond only to those elements r_{ij} of the DDP for which 3 or more difference peaks satisfy the 9.0 A. limit.

The 3.0 M. urea DDP is considerably more complicated than the GdnHCl DDP. Changes have occurred throughout both folding domains of both CHT molecules. Also, there are considerably more changes occurring in the interdomain regions of both molecules (lower-left and upper-right quarters of the DDP) for the urea derivative. Generally, most plot features which occur in one CHT molecule are reflected across the diagonal, indicating that the deviations from local two-fold symmetry in the difference map are more a matter of degree than kind. The difference peaks are

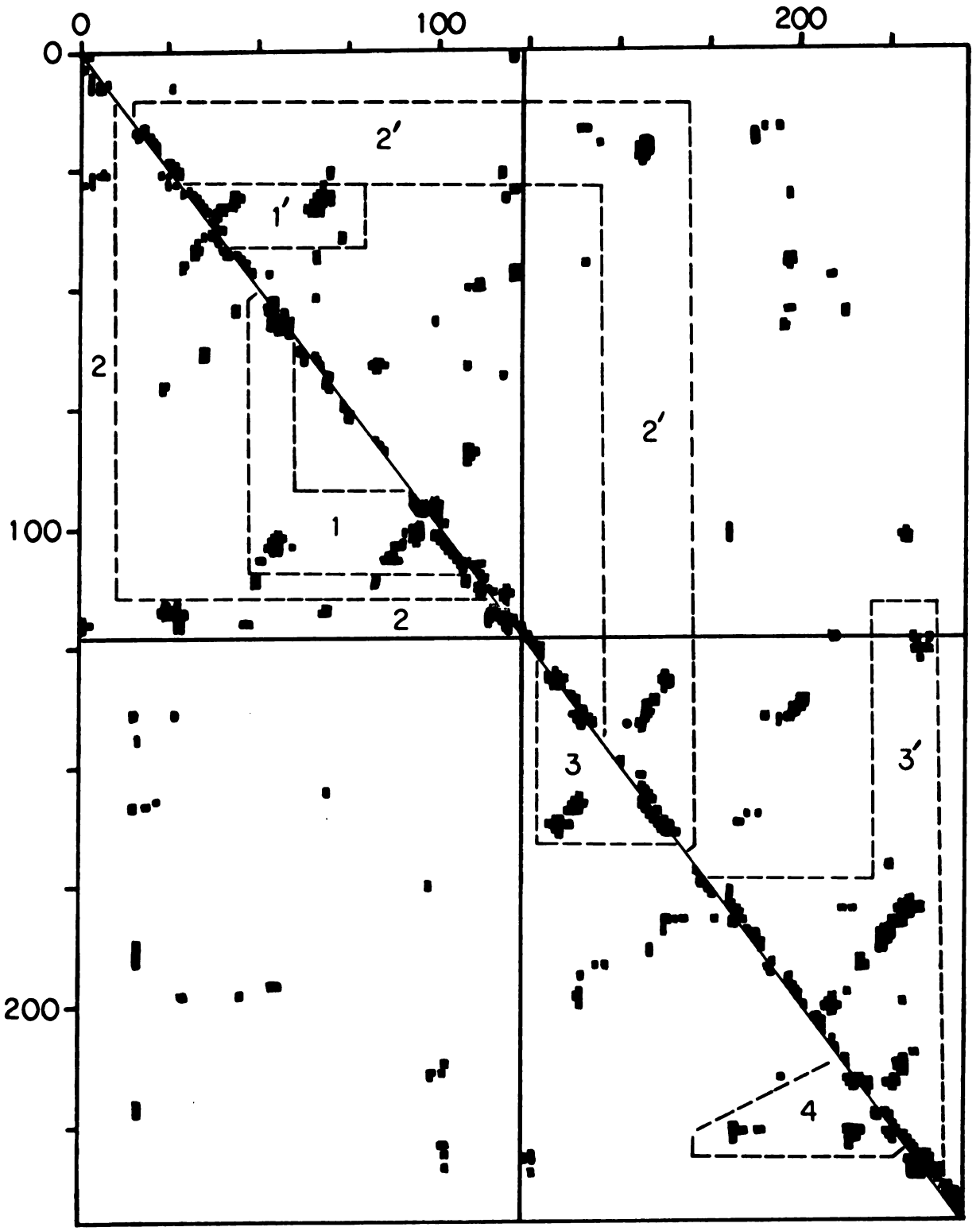


Figure 17. The difference diagonal plot of the 3.0 M. Urea derivative.

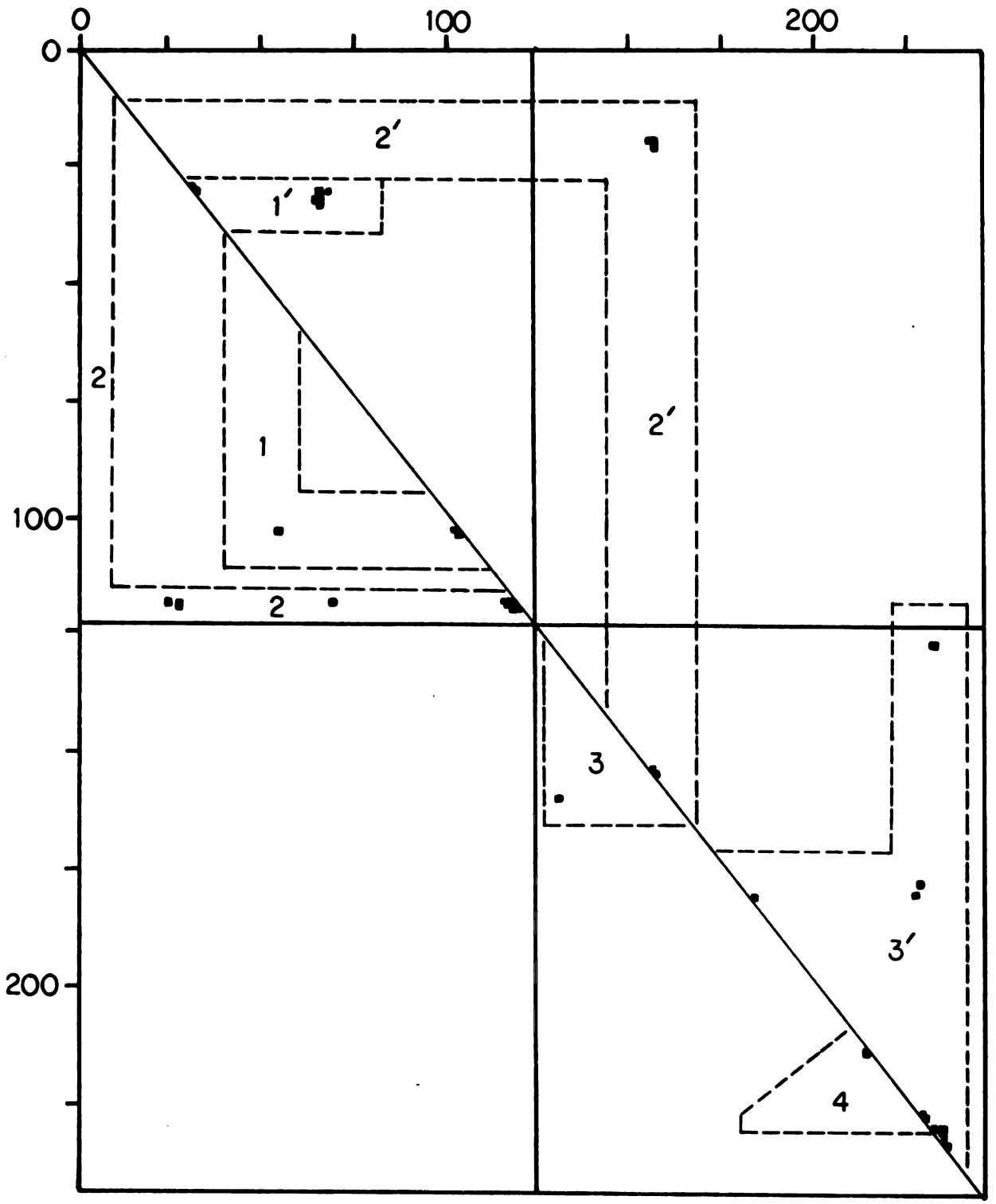


Figure 18. The difference diagonal multiplicity plot of the 3.0 M. Urea derivative.

closer to more pairs of α-carbon atoms in Molecule I' than in Molecule I; the number of $r_{ij} \leq 9.0$ A. is 538 for Molecule I' versus 369 for Molecule I. There is also a larger number of points above the diagonal (29) than below the diagonal (17) in the DDMP. These two results taken together indicate that the structure of Molecule I' is substantially more perturbed by the urea than Molecule I.

The DDP contains some indications of disruption of secondary structure in both molecules. There are several plot features, perpendicular to, and extending away from the diagonal, on both sides of the diagonal which represent difference density occurring near strands of anti-parallel β-sheet. There is some difference density near the C-terminal α-helix of Molecule I', but it may not represent significant changes since there is no feature in this region of the DDMP.

A most useful result obtainable from the DDP is an indication of those regions of the CHT molecule where difference peaks are concentrated. From the discussions of the TCA and GdnHCl derivative DDP's, difference peaks occurring in particular locations of the molecule may create distinct patterns of features. Examples include the rectangular array of features observed in both the TCA and the GdnHCl DDP's created by difference peaks in or near the substrate binding site, and other combinations of vertical and horizontal bands of DDP features occurring for difference peaks located in the interiors of the CHT molecules. However,

large features (or groups of features) may be created by small difference peaks so that the multiplicities of DDP features should also be considered. Therefore, to represent significant and extended difference density peaks, DDP features should satisfy two criteria: (1) the DDP feature should have a corresponding feature in the DDMP (subject to arbitrary limits appropriate for the derivative) and, (2) the DDP feature should extend over a number of amino acid residues, or even better, be one of several DDP features on approximately the same residue number ordinate or abscissa. In this way, one might "survey" a complex difference map to determine those regions of the CHT molecule in which the most significant changes might be expected to occur.

The 3.0 M. urea DDP and DDMP contain indications that there are several areas of extended difference density in both CHT molecules. In Molecule I of the DDMP, two features at (103, 55) and at (103-104, 102-103), where the coordinates are expressed as (row, column), correspond to two features in the DDP at the same positions. The first is an off-diagonal feature and the second is an extended feature on the diagonal. These two features, with two others nearby, join the L-shaped group labelled "1" in Figure 17. This group of DDP features was created by a group of peaks concentrated in and around a segment of anti-parallel β-sheet composed of four polypeptide segments - Gly 44-Ser 45-Leu 46, Trp 51-Val 52-Val 53-Thr 54, Thr 104-Leu 105-Leu 106-Lys 107-Leu 108-Ser 109 and Ile 85-Ala 86-Lys 87-Val 88-Phe 89-Lys 90. These

difference peaks make up one of the two principal changes observed in the hydrophobic core of the CHT molecule.

Four DDMP features at (118, 25), (118-119, 28), (118, 69-70), and (118-120, 115-119), all clustered in the abscissa interval 118-120 in Figure 18, lead to another L-shaped group of features in the DDP labelled "2" in Figure 17. This group of features indicates that changes have occurred involving the amino acid residue segments Cys 1-Pro 8, Val 23-Gln 30, and Gln 116-Cys 122. Examination of the difference map revealed that these DDP features are due to a group of difference peaks (including an example of one of the most extended difference peaks in the map) occurring in the A-Chain/B-Chain contact region.

Two smaller clusters of DDP features labelled "3" and "4" in Figure 17 have been suggested by DDMP features occurring at (162, 131) and (215, 214) in Figure 18, respectively. Cluster 3 arises from a number of difference peaks occurring in and around a segment of anti-parallel ß-sheet partly made up of the polypeptide sgements Leu 155-Leu 160 and Val 137-Gly 140. These difference peaks constitute the other major group of changes observed in the nonpolar interdomain region. Cluster 4 arises from a small number of difference peaks in the neighborhood of the polypeptide segments Try 215-Ser 217 and Gly 226-Tyr 228, and also near the aromatic cluster Tyr 171, Trp 172, and Trp 215. The group of difference peaks involved in Cluster 4 are less extensive than those groups of peaks responsible for Clusters 1, 2, and 3.

The features in the Molecule I' portion of the DDMP have been grouped together in three clusters, 1', 2', and 3'. Cluster 1' consists of four closely-spaced DDMP features at (30, 31), (31, 33), (31-34, 65-66), and (31,68). The corresponding DDP features indicate changes have taken place in a part of the A folding domain involving the polypeptide segments Glu' 30-Asp' 35, Gly' 38-Ser' 45, and Val' 65-Glu' 70. This portion of the A-domain borders the nonpolar interdomain region and is near (or includes) a number of nonpolar and/or aromatic residues including Phe' 39, His' 40, Phe' 41, and Trp' 141. Cluster 2' on the DDMP consists of two features at (20-22, 156-157) and (155-156, 156-157). The corresponding Cluster on the DDP indicates that there are changes involving the polypeptide segments Gly' 19-Val' 23, Cys' 136-Trp' 141, and Leu' 155-Ala' 158. These polypeptide segments are nearest neighbors and together form a small segment of twisted anti-parallel β-sheet bordering on the nonpolar interdomain region.

The final group of Molecule I' DDMP features corresponding to Cluster 3' in the DDP consists of seven separate features including the largest single feature in the 3.0 M. urea DDMP - a feature located at (231-235, 234-236). The coordinates of the other six features are: (126, 233), (182, 183), (180, 228), (182, 227), (228-229, 229-230), and (231, 232). The corresponding DDP features in Figure 17 are quite extensive and indicate that extensive changes in structure have probably occurred over a large volume of the

B-domain. The polypeptide segments especially implicated as the sites of these changes are Leu' 123-Ala' 126, Asp' 178-Val' 188, Ile' 212-Ser' 218, and Cys' 220-Trp' 237. The segment Leu' 123-Ala' 126 occurs on the surface of Molecule I' adjacent to a portion of the Cys' 220-Trp' 237 segment. The three segments Asp' 178-Val' 188, Ile' 212-Ser' 218, and Cys'220-Trp' 237 are adjacent to one another and together form a large segment of anti-parallel β-sheet, which is in turn a substantial portion of the B-domain.

With a detailed examination of the Molecule I' portion of the DDP, another result emerges - a greater degree of variability in local two-fold symmetry is now evident. The polypeptide segments implicated by each cluster are listed below and correspondences are indicated.

Molecule I		Molecule I'	
Cluster	Polypeptide Segment	Cluster	Polypeptide Segment
1	Gly 44-Leu 46	1'	Glu' 30-Asp' 35
	Trp 51-Thr 54		Gly' 38-Ser' 45
	Thr 104-Ser 109		Val' 65-Glu' 70
	Ile 85-Lys 90	2'	Gly' 19-Val'23
2	Cys 1-Pro 8		Cys' 136-Trp' 141
	Val 23-Glu 30		Leu' 155-Ala' 158
	Glu 116-Glu 30	3'	Leu' 123-Ala' 126
3	Leu 155-Leu 160		Asp' 178-Val' 188
	Val 137-Gly 140		Ile' 212-Ser' 218
4	Trp 215-Ser 217		Cys' 220-Trp' 237
	Gly 226-Tyr 228		

Clearly, substantial changes have taken place in both molecules which are not reflected in the other.

In conclusion, the DDP used in conjunction with the DDMP gives a reliable estimate of the location and extent of changes in structure represented by the peaks in a complicated electron density map. The 3.0 M. urea DDP indicated correctly the polypeptide segments where the difference peaks were most numerous - 3 of the 4 clusters of DDP features correspond to extensive sets of difference peaks discussed in separate subsequent sections of this chapter. The many small features of the DDP which have not been discussed can be categorized into two groups - features on or close to the ordinate or abscissa of a feature in a cluster already discussed, and features created by peaks isolated from other peaks or occurring near only a small number of α-carbon atoms. Peaks creating features in the first category may be members of significant and extensive groups of peaks. Significant difference peaks occurring in the intermolecular contact regions may create DDP features in the second category. In any case, the DDP does not contain all the information of the difference map, but may serve as an aid to the organization of the information contained in the difference map.

3. Changes on the Dimer Surface.

There are a substantial number of difference peaks occurring on the dimer surface and they appear to represent

mostly changes in the solvent structure, and small, localized changes in the positions of main chain and side chain atoms. Most of the peaks are small in extent ($\leq$ 2.0 A. in diameter at the 0.20 e.A.$^{-3}$ contour) and were observed at scattered locations over the surface. Peaks which appeared to belong to a group of peaks extending into the interior of the molecule are grouped with those peaks and are not discussed here. The unit cell coordinates and heights of the most significant surface peaks are listed in Table 5.

Peak 1 indicates that urea, like GdnHCl, displaces the SO_4^{2-} ion near Lys 177 from its position in the native structure. There is no significant positive density nearby to indicate that the sulfate ion has undergone a small change in position - it has simply been removed. Peaks 2 and 3 represent a small shift in the main chain atoms at the peptide link between Thr 174 and Lys 175 - no significant difference density was observed in the native density of the side chains of either residue. Peak 4, located just outside the native density at the Tyr 171 phenol may represent a change in the solvent structure rather than a motion of the side chain.

Peaks 5 through 8 all occur in a contact interface region shared by Molecule I with a neighboring CHT molecule (nearest neighbor in the negative y-direction). These four peaks, and an equal number of smaller peaks (<2.0 A. across), were observed in a small solvent cavity formed by the polypeptide segments Thr 135-Cys 136-Val 137 and

Table 5. The 3.0 M. Urea Derivative - The Most Significant Difference Peaks on the Surface of the CHT Dimer.

Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
1	0.394	0.481	0.215	-0.24	At* SO_4^{2-}-177
2	0.327	0.378	0.215	-0.26	At Lys 175 carboxyamide NH group
3	0.327	0.350	0.331	0.20	2.0 A. from Thr 174-Lys 175 peptide group
4	0.356	0.170	0.307	-0.28	2.0 A. from Tyr 171 phenol group
5	0.394	0.378	0.207	0.27	5.0 A. from Gly 133 main chain atoms; 5.0 A. from Ser 186 side chain
6	0.499	0.047	0.215	0.26	5.0 A from Thr 135 side chain; 2.0 A. from Ser 159 side chain
7	0.540	0.000	0.177	0.29	2.0 A. from Gly 12-Leu 13 peptide group; 3.0 A. fro Thr 135 side chain
8	0.671	0.038	0.215	-0.28	2.0 A. from the Ser 159 side chain
9	0.408	0.217	0.115	0.27	In solvent, 3.0 A. from main chain at both Phe 130 and Ser 164
10	0.065	0.330	0.117	0.28	2.0 A. from Ala 112 carbonyl group
11	0.040	0.330	0.222	-0.27	At Ala 112 C_α atom

Table 5. Continued

Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
12	0.868	0.019	0.415	0.26	2.0 A. from Asp 73 carboxyl group; 2.0 A. from Arg 154 Gdn^{+} group
13	0.132	0.519	0.585	0.26	5.0 A. from Asp' 153 carboxyl group

*"At...." a particular location means "In the native electron density at....."

Ala 158-Ser 159-Leu 160-Pro 161, and residues of the neighboring CHT molecule. Peaks 5, 6, and 7 are all about the same size, shape, and height, and may represent urea molecules engaged in hydrogen bonding interactions with groups on both proteins. Peak 8 is 2.0 A. from the hydroxyl side chain of Ser 159 and 2.0 A. from the peptide group joining Leu 10 and Ser 11 and may represent some reorganization of solvent structure; there was no significant difference density observed in the nearby native electron density. Peak 8 is also only several Angstroms from some of the peaks occurring in the A-Chain/B-Chain contact region and may be connected with the changes occurring in that region.

Peak 9 is shaped like a footprint, is about 3.0 A. in length, and is one of a number of similarly-shaped peaks observed throughout the map. It is located in a crevice on the surface defined by the polypeptide segments Ser 127-Asp 128-Asp 129-Phe 130-Ala 131 and Leu 163-Ser 164-Asn 165 and is about 2.0 A. from the carbonyl of Phe 130. No significant difference density was observed in any of the surrounding native density. Thus, peak 9 may represent a urea molecule hydrogen bonded to the Phe 130 carbonyl.

Peaks 10 and 11 form a gradient at the main chain atoms of Ala 112, clearly indicating that the polypeptide backbone at this residue has moved in the direction of the solvent.

Finally, one peak was observed in each of the uranyl binding sites discussed in Chapter VI. They are significant

features and are listed as Peaks 12 and 13 in Table 5. Peak 13 is the two-fold screw equivalent of Peak 12, and both peaks are about 3.0 A. from the GdnHCl Peaks 1 and 4 (Chapter VI, Section 4), respectively. There is no accompanying negative difference density at either site so that Peaks 12 and 13 may represent urea molecules hydrogen bonded to the Glu 21 and/or the Arg 154 side chains.

4. Changes in the Dimer Interface Region.

This region of the difference map contains a substantial number of difference peaks ($\sim$30) with absolute heights $\geq$ 0.2 e.A.$^{-3}$ There are no examples of single, extended difference features in this region, but rather groups of small peaks ($\sim$2.0 A. across at the 0.2 e.A.$^{-3}$ contour) were observed at several locations in the interface. Several of these groups of peaks appeared to represent significant changes in structure in both Molecules I and I'. These peaks are listed in Table 6.

One group of peaks occurred in and around the native density of the polypeptide segment Gly' 216-Ser' 217-Ser' 218-Thr' 219-Cys' 220-Ser' 221-Thr' 222-Ser' 223-Thr' 224. This segment contains seven hydroxyl groups within a relatively small volume and portions of it display large deviations from local two-fold symmetry.[42] The unit cell coordinates and heights of the most significant peaks in this region are Peaks 1-11 in Table 6. Interestingly, none of these peaks have a local dyad equivalent in Molecule I - there are no significant features in the difference map at

Table 6. The 3.0 M. Urea Derivative - The Most Significant Difference Peaks in the Dimer Interface Region.

Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
1	0.408	0.443	0.538	0.25	2.0 A. from Thr' 224-Pro' 225 peptide group
2	0.421	0.434	0.508	-0.22	At thr' 224 C_α atom
3	0.434	0.509	0.531	0.27	1.0 A. from Thr' 224 C_α atom
4	0.447	0.453	0.485	-0.27	At Ser' 221 carbonyl group
5	0.461	0.500	0.469	0.24	1.0 A. from Thr' 222 side chain
6	0.474	0.44β	0.523	0.25	At Ser' 159 side chain
7	0.487	0.514	0.546	-0.25	1.0 A. from Thr' 222 side chain
8	0.500	0.340	0.538	0.26	1.0 A. from main chain atoms of Gly' 216
9	0.579	0.358	0.431	-0.24	At Ser' 218 carbonyl group
10	0.566	0.425	0.462	-0.23	At main chain atoms of Cys' 220
11	0.592	0.302	0.508	0.27	2.0 A. from Gly' 216 carbonyl group
12	0.618	0.302	0.532	-0.27	2.0 A. from Tyr 146 phenol group; 2.0 A. from SO_4^{2-}-195'
13	0.671	0.264	0.477	0.26	1.0 A. from Met 192 side chain
14	0.711	0.292	0.515	-0.27	At SO_4^{2-}-195' counterion native density peak

Table 6. Continued

Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
15	0.724	0.292	0.462	-0.25	2.0 A. from SO_4^{2-}-195 counterion; 2.0 A. from Met 192 side chain
16	0.789	0.292	0.477	0.27	In solvent cavity 3.0 A. from both SO_4^{2-}-195 and SO_4^{2-}-195' counterions
17	0.829	0.453	0.592	0.28	2.0 A. from Pro' 152 pyrrolidine ring
18	0.947	0.292	0.488	-0.25	2.0 A. from main chain at Phe 39-His 40
19	0.105	0.311	0.508	0.29	2.0 A. from Gly 38 carbonyl group
20	0.000	0.387	0.554	0.28	1.0-2.0 A. from the Thr 37 side chain

residues Gly 216 to Thr 224. Due to the complicated solvent structure which exists around these residues, and because this part of the CHT molecule is strained by the close approach of its local dyad equivalent during dimer formation, this segment appears to be very sensitive to changes in the solvent environment, e.g., changes in pH (38, 67). Significant changes in this region were also observed in the GdnHCl derivative (Chapter VI).

Though the entire segment Gly' 216-Thr' 224 is perturbed by the urea, most of the difference peaks were clustered near residues Ser' 221-Thr' 222-Ser' 223-Thr' 224. They represent complicated changes involving both main chain and side chain atoms in a number of small adjustments in position. Changes in the solvent structure are evident as well and most of the peaks were small as well ($\sim$2.0 A. at the 0.2 e.A.$^{-3}$ contour). Residues Ser' 221-Thr' 222-Ser' 223-Thr' 224 form a β-bend with a hydrogen bond between the Ser' 221 carbonyl group and the Thr' 224 peptide amino group. Peaks 2 and 4 represent movements by the main chain atoms at Ser' 221 and Thr' 224 away from each other (they cannot move closer) probably breaking the β-bend hydrogen bond. Similarly, the negative peak at the Ser' 218 carbonyl group may represent a movement by the main chain at Ser'218-Thr' 219 resulting in the breaking of the hydrogen bond between the peptide amino group of Thr' 219 and SO_4^{2-}-217'. The SO_4^{2-}-217' itself has not been affected.

The three largest difference map features in this group of peaks are peaks 8, 10, and 11. Peak 10 is 3.5 A. in length at the 0.2 e.A.$^{-3}$ contour and is in the native density at the main chain at Cys' 220, indicating some disordering of the main chain atoms as there is no accompanying positive density. There is no indication that the disulfide Cys' 191-Cys' 220 has been affected. Peaks 8 and 11 are oval-shaped peaks, each about 2.5 A. in length, and neither have any significant accompanying negative density. They were observed on opposide sides of the main chain at Gly' 216 and may represent a pair of urea molecules binding to the CHT molecule since there are many polar groups on the protein here with which urea molecules could form hydrogen bonds.

Another set of difference peaks were observed near the active sites of both molecules in a pattern similar to that of some of the 2.0 M. GdnHCl difference peaks. Two of the peaks may represent bound urea molecules and the rest of the peaks represent changes in the solvent structure as there was no significant difference density observed in the native density of the protein in this region. The changes are those represented by Peaks 12-16 in Table 6.

Peaks 13 and 16 may represent urea molecules binding in the dimer interface. Peak 13 is an oval-shaped peak 2.0 A. in length; peak 16 is an oval about 2.5 A. in length. Peak 13 is 1.0 A. from the native density at the Met 192 side chain and may represent a urea molecule forming a hydrogen bond between an amino group and a non-bonding

pair of electrons on the Met 192 sulfur. Peak 16 may have come about because of a urea molecule hydrogen bonded to one of the counterions of either SO_4^{2-}-195 or SO_4^{2-}-195'. Peaks 12, 14, and 15 probably represent changes in the local solvent structure around the two sulfate ions. The locations of these peaks are somewhat similar to those of the GdnHCl peaks observed in this region. For example, peak 16 is located very near the position occupied by one lobe of the positive double peak observed in the GdnHCl difference map (peak 5, Chapter VI, Section 3). There correspondences observed between peaks of the two derivatives are listed below, where the GdnHCl peak numbers are those of the discussion in Chapter VII, Section 3.

Urea Peak	GdnHCl Peak
12	2
14, 15	3
16	5

Further, a small negative urea peak is located very close (<1.0 A.) to the negative GdnHCl peak 1, and the positive urea peak 13 is coincident with a portion of the positive, cresent-shaped (0.16 e.A.$^{-3}$ contour) GdnHCl peak described in Chapter VI near the Met 192 and Met' 192 side chains. These similarities between the urea and the GdnHCl derivatives are unusual - there are very few examples of peaks in one difference map coinciding with those of the other.

Finally, there are a number of peaks located in the dimer interface solvent cavity near residues Gly 38-Phe 39-His 40. These are Peaks 17-20 in Table 6. Peaks 19 and 20 are in the solvent and may represent urea molecules hydrogen bonded with the carbonyl oxygen of Gly 38 and the side chain hydroxyl group of Thr 37, respectively. The rest of the peaks observed here represent small changes in the protein and the solvent and are difficult to interpret.

5. Changes in the A-Chain/B-Chain Contact Region.

The A-chain (residues 1-13) is attached to the B-chain covalently at one point - the disulfide bond Cys 1-Cys 122- and by non-covalent interactions at many other points. The A-chain is composed chiefly of amino acid residues with nonpolar side chains and folds over the surface of the CHT molecule such that it forms a number of van der Waals contacts with nonpolar side chains of B-chain residues. The pyrrolidine side chains of Pro 4 and Pro 8 are directed toward the interior of the CHT molecule, and specifically toward a cluster of tryptophan side chains (Trp 27, Trp 29, Trp 237) which belong to the hydrophobic interdomain region described in Chapter II. This region was very much affected by the urea; the difference map contained a number of significant features in a relatively small portion of the CHT molecule. These features' unit cell coordinates and peak heights are listed in Table 7.

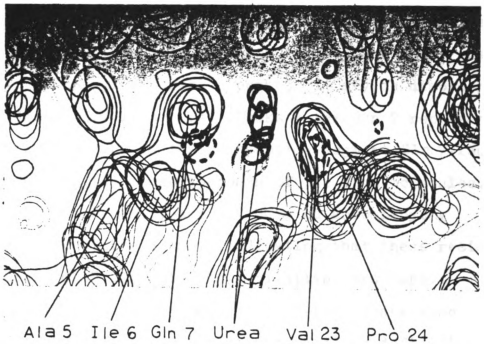
Peaks 1 and 2 join at the 0.2 e.A.$^{-3}$ contour to form a continuous feature about 5.0 A. in length, roughly perpendicular

Table 7. The 3.0 M. Urea Derivative - the Most Significant Difference Peaks in the A-Chain/B-Chain Contact Region.

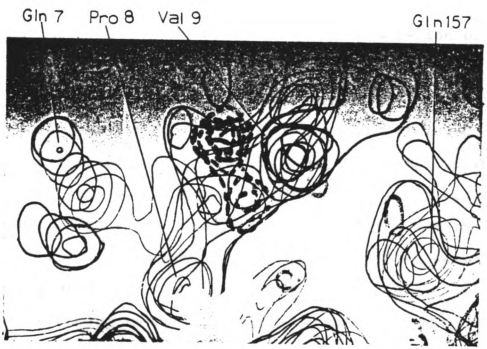
Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
1	0.895	0.009	0.223	0.25	2.0 A. from Val 9 and Val 23 side chains
2	0.947	0.038	0.223	0.27	2.0 A. from the Ile 6 side chain
3	0.895	0.038	0.262	-0.25	At Val 23 side chain
4	0.868	0.028	0.185	-0.25	At Glu 7 peptide carbonyl group
5	0.803	0.000	0.185	-0.31	At Val 9 C_α atom
6	0.776	0.038	0.192	-0.25	At Pro 8-Val 9 peptide group
7	0.671	0.038	0.215	-0.28	1.0 A. from Leu 10 carbonyl group

to the yz-plane. This feature is located in a crevice on the surface of the CHT molecule formed by the two polypeptide segments Ile 6-Glu 7-Pro 8-Val 9 and Ala 22-Val 23-Pro 24-Gly 25-Ser 26, and is about 2.0 A. from the side chains of Ile 6, Val 9, and Val 23. Peaks 3-7 represent changes in the orientation of the Val 23 side chain and a general disruption of the A-chain residues Glu 7-Pro 8-Val 9-Leu 10. The terminal methyl group of the Ile 6 side chain can be placed into the positive density of peak 2, but none of the neighboring side chains can be repositioned to fit peak 1. It is unlikely that peaks 1 and 2 were the result of the A-chain moving closer to the B-chain - the negative density around the A-chain indicates small, localized movements by individual groups. Peaks 1 and 2 were not formed by a movement of the B-chain toward the A-chain as there is no significant difference density in this region of the B-chain, except at the Val 23 side chain. The most satisfying explanation for these changes is that peaks 1 and 2 represent a pair of urea molecules - possibly hydrogen bonded to one another since their peak centers are bout 2.0 A. apart - binding by van der Waals forces in the aliphatic "pocket" formed by the side chains of Ile 6, Val 9, and Val 23. Stable urea-hydrocarbon complexes are known and the structures of some have been determined by X-ray crystallography.[68] These features appear in Figure 19(a).

Peak 3 arises because of a motion by the Val 23 side chain about the C_α-C_β bond to move the isopropyl group



(a) Sections of electron density in the x-interval 66/76 to 75/76.



(b) Sections of electron density in the x-interval 55/76 to 65/76.

Figure 19. Difference peaks observed in the A-chain/B-chain contact region. The difference map contours are the thick gray lines.

further from the bound urea. Its position in the derivative structure is unclear because there is no significant positive density accompanying negative peak 3; it may be disordered in the derivative structure. Peaks 4 through 7 represent movements by groups in the polypeptide segment Gln 7-Leu 10 in response to the close approach of the urea molecules. Like peak 3, however, peaks 4-7 have no accompanying significant positive density indicating that these residues have become disordered in the derivative structure. Interestingly, residues Ser 11-Gly 12-Leu 13 are also disordered but in the native structure as no definitive electron density exists for them.[52]

6. Changes in the Nonpolar Interdomain Region.

The striking structural features of this region of the CHT molecule have been described in Chapter II. It contains three clusters of aromatic residues, several proline side chains, and a number of aliphatic side chains arranged roughly in the shape of a "horseshoe" surrounding the inner ends of the two folding domains (cylinders). A concentration of difference peaks was observed in this region and the peaks occurred in two groups. One group of small peaks represents scattered changes occurring in a segment of β-sheet belonging to the B-domain near the tryptophen cluster Trp 27, Trp 29, and Trp 207. The second group of peaks were observed in and around a segment of β-sheet belonging to the A-domain and near the atomatic cluster Trp 51, Phe 89, and Trp 237.

The members of the first group of peaks were observed at scattered locations in and around a small segment of anti-parallel β-sheet composed of the polypeptide segments Val 137-Thr 138-Thr 139-Gly 140, Gln 156-Gln 157-Ala 158-Ser 159-Leu 160, and Gly 197-Pro 198-Leu 199-Val 200. The peaks of this group are peaks 1-5 in Table 8. Though the exact nature of the changes here are unclear, most of the difference peaks listed and an equal number of smaller peaks represent small local shifts in position of main chain and side chain atoms. Peaks 1 and 5 are similar in appearance - both are cylindrical and 2.0-3.0 A. in length and both are 1.0-2.0 A. from the native density of an aliphatic and an aromatic side chain, respectively. There are several other peaks observed in the difference map having sizes, shapes, and heights very similar to peaks 1 and 5, and are located close to, but not in, the native density at various residues. Because of the similarities of appearance, peaks 1 and 5 may represent urea molecules. Urea is capable of interacting with a variety of protein groups, including nonpolar groups so that the penetration of the hydrophobic core of the CHT molecule by urea is not at all unreasonable. If peak 1 is a urea molecule participating in van der Waals interactions with the Leu 155 side chain, peaks 2, 3, and a small negative peak at the Leu 155 α-carbon atom (all within about 5.0 A. of peak 1) represent small movements by the nearest protein groups to accomodate the urea molecule. Likewise peak 5 may represent a urea molecule sandwiched between the native

Table 8. The 3.0 M. Urea Derivative - the Most Significant Difference Peaks in the Nonpolar Interdomain Region.

Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
1	0.842	0.113	0.308	0.23	2.0 A. from Leu 155 side chain
2	0.816	0.142	0.223	-0.22	At Trp 27 C_α atom
3	0.711	0.170	0.269	-0.27	At Thr 139 C_α atom
4	0.763	0.236	0.185	0.23	2.0 A. from Trp 29 indole group; 2.0 A. from Pro 198 pyrrolidine ring
5	0.842	0.165	0.115	0.22	2.0 A. from Trp 29 indole group; 2.0 A. from Ser 119 side chain
6	0.816	0.443	0.215	-0.23	At Leu 106 peptide amino group
7	0.763	0.415	0.192	0.23	1.0 A. from Leu 105 C_α atom
8	0.750	0.415	0.223	-0.24	At Thr 104 C_α atom
9	0.724	0.368	0.246	-0.21	At Thr 54-Ala 55 peptide group
10	0.724	0.509	0.300	0.23	2.0 A. from Lys 90 side chain
11	0.711	0.462	0.338	0.24	1.0 A. from Ala 56 carbonyl group
12	0.632	0.443	0.192	0.25	1.0 A. from Ile 103 side chain
13	0.566	0.443	0.246	-0.30	At Asn 101-Asp 102 peptide group

Table 8. Continued

Peak	X	Y	Z	Height(e.A.$^{-3}$)	Comments
14	0.513	0.472	0.200	-0.24	At Asn 101 side chain carboxyamide group
15	0.632	0.387	0.192	0.23	1.0 A. from Ile 103 side chain
16	0.539	0.509	0.292	-0.25	At Asn 95 C_α atom
17	0.697	0.462	0.338	0.25	2.0 A. from Ala 56-His 57 peptide group

density of the Trp 29 indole side chain and that of the A-chain backbone at Gly 2. No significant negative difference density was observed in the native density of any of the nearby amino acid residues - apparently peak 5 represents a substitution requiring little accomodation by the protein.

The second group of peaks occurred in and around a segment of anti-parallel β-sheet in the A-domain made up of the polypeptide segments Trp 51-Val 52-Val 53-Thr 54, Thr 104-Leu 105-Leu 106-Lys 107-Leu 108-Ser 109, and Ile 85-Ala 86-Lys 87-Val 88-Phe 89-Lys 90. There are larger peaks in this group than in the previous group, and they combine to form more extended changes. They are listed in Table 8 as Peaks 6-17.

This group of peaks displays a feature not observed anywhere else in the urea difference map - an extended density feature at the 0.16 e.A.$^{-3}$ level under several small peaks at the 0.2 e.A.$^{-3}$ level. Three negative peaks, including peaks 8 and 9, and a small peak at the Phe 89 carbonyl group combine to form a negative feature approximately 11.0 A. in length involving peptide groups in three chains - peptide Thr 54-Ala 55, Thr 104-Leu 105, and Phe 89-Lys 90. There is one significant positive peak - peak 7 - approximately 2.0 A. from peak 8, which together with peak 8 may represent a movement by the main chain atoms at Thr 104 into the hydrophobic cavity described in Chapter II. The motions of the other two polypeptide segments are uncertain, but it is clear that this segment

of β-sheet has been disrupted. Three other nearby negative peaks indicate local shifts by main chain and side chain atoms at Asn 101 (peaks 13 and 14) and at Asn 95 (peak 16). These residues are on the surface of the CHT molecule and may be the result of local changes in solvent structure and not directly related to the extended negative feature just described.

This region also contains several positive, oval-shaped peaks having no significant accompanying negative density. The peaks are in the solvent near the Lys 90 side chain (peak 10), between the side chains of Ile 103 and Asn 91 (peak 12), and in the solvent 2.0 A. from the peptide group joining Ala 56 and His 57 (peak 17). These peaks may represent bound urea molecules (especially peak 12) or changes in the solvent structure at points close to the surface of the protein molecule.

Chapter VIII.

Summary of Results and Discussion

1. The 2.0 M. Guanidine HCl Derivative.

The changes observed in the 2.0 M. GdnHCl derivative were confined entirely to the surface of the molecule (including the dimer interface region) and no significant difference density was observed in the hydrophobic inter-domain region or in either of the folding domains. Most of the difference peaks were small in extent and indicative of localized movements of side chain and main chain atoms and changes in the local solvent structure. There is also evidence that the Gdn^+ ion binds to the proteins in the dimer interface near the sulfates SO_4^{2-}-195 and SO_4^{2-}-195', and in the uranyl binding site. However, the latter binding site is an artifact of the crystal packing and may not be significant to the chemistry of GdnHCl denaturation. The sulfate ion near Lys 177 is also removed from the surface of the molecule to the bulk solvent in the GdnHCl derivative.

The chief effect of the GdnHCl at the 2.0 molar concentration on the structure of the CHT molecule, seems to be to cause changes in the solvent shell around the molecule. The protein responded to the presence of the GdnHCl by undergoing structural adjustments in those parts of the molecule already known to be solvent-sensitive from the pH derivative studies. Concentrations of GdnHCl typically required to effect the complete unfolding of a protein

in solution are about 6.0 M.[69,70,71] Clearly, changes in structure occur at lower concentrations, and while the changes observed in the 2.0 M. GdnHCl derivative are significant and may accompany the onset of denaturation, it is not possible to extend these results to a model for GdnHCl-induced unfolding. Soaking CHT crystals in solutions with higher concentrations of GdnHCl will not provide answers either as an abrupt deterioration of the crystals occurs between 2.0 and 2.3 M. GdnHCl concentrations. With the extension of the CHT native phases from 2.8 A. to 1.8 A. resolution[72] it may be possible to learn more about the solvent-protein interactions in the native structure. In a 1.8 A. resolution 2.0 M. GdnHCl derivative, some of the details of the solvent-protein interactions in the dimer interface region may be elucidated.

2. The 3.0 M. Urea Derivative.

The 3.0 M. urea derivative is more interesting than the GdnHCl. Significant changes were observed not only on the surface of the CHT molecule, but in hydrophobic interior regions as well. The changes observed on the surface were local in nature affecting both polar and nonpolar groups, but no trends were evident in the behavior of the surface residues. A similar observation can be made with regard to the difference Fourier features in the dimer interface. The latter difference peaks represented local changes involving the amino acid residues in the polypeptide segment Trp' 215-Thr'224, the side chains of Met 192 and Met' 192, the

sulfate ions SO_4^{2-}-195 and SO_4^{2-}-195' and the surrounding solvent.

A special subset of the set of surface changes are those observed in the A-chain/B-chain contact region. The A-chain and the B-chain interact via van der Waals forces between aliphatic side chains, and these interactions are disturbed by what may be the intrusion of urea molecules into a space between the A-chain and segments of the B-chain. The A-chain folds over a portion of the hydrophobic interdomain region at the tryptophan cluster Trp 27, Trp 29, and Trp 207, and the changes observed in the A-chain/B-chain region are in close proximity to some of the difference peaks observed in the B-domain β-sheet. Here again, the difference map indicates small changes involving individual groups on the protein. There is also some evidence that urea molecules are binding near nonpolar side chains.

The final group of difference peaks observed in the interior of the CHT molecule were those near a segment of β-sheet in the A-domain near the aromatic cluster Trp 51-Phe 89-Trp 237. Several negative peaks, taken as a group, comprise the most extended difference feature in the entire map and indicate a general disruption of this segment of β-sheet. Here as previously, the appearance of substantial positive peaks without significant accompanying negative density suggests that urea molecules may be binding noncovalently to protein groups. Even if some of these peaks (especially the large peak near the Ile 103 side

chain) do represent urea molecules, it is not always clear how specific urea substitutions affect the protein groups nearest to them. Once again, this kind of detail may be obtainable from a 1.8 A. resolution 3.0 M. urea difference Fourier. Also, since 5.0 M. urea-CHT crystals diffract well, the changes observable in a 5.0 M. derivative may elaborate upon those observed in the 3.0 M. derivative.

LIST OF REFERENCES

LIST OF REFERENCES

1. W. Kauzmann, Adv. Prot. Chem., 14, i (1959).
2. C. N. Pace, CRC Crit. Rev. Biochem., 3, 1 (1975).
3. R. F. Greene, Jr., and C. N. Pace, J. Biol. Chem., 249, 5388 (1974).
4. H. Edelhoch and J. C. Osborne, Jr., Adv. Prot. Chem., 30, 183 (1976).
5. F. Franks and D. Eaglund, CRC Crit. Rev. Biochem., 3, 165 (1975).
6. C. Tanford, Adv. Prot. Chem., 23, 121 (1968).
7. C. Tanford, Adv. Prot. Chem., 24, 1 (1970).
8. D. B. Wetlaufer, Proc. Nat. Acad. Sci. USA, 70, 697 (1970).
9. A. V. Finkelstein and O. B. Ptitsyn, J. Mol. Biol., 103, 15 (1976).
10. R. Lumry and R. Biltonen, Structure and Stability of Biological Macromolecules, vol. 2, S.N. Timasheff and G. D. Fasman, eds., Marcel Dekker Inc., New York, 1969, p. 65.
11. J. F. Brandts, Structure and Stability of Biological Macromolecules, vol. 2, S. N. Timasheff and G. D. Fasman, eds., Marcel Dekker Inc., New York, 1969, p. 213.
12. P. L. Privalov and N. N. Khechinashvili, J. Mol. Biol., 86, 665 (1974).
13. J. F. Brandts, H. R. Halvorson, and M. Brennan, Biochemistry, 14, 4953 (1975).
14. J. R. Garel, B. T. Nall, and R. L. Baldwin, Proc. Nat. Acad. Sci. USA, 73, 1853 (1976).
15. J. F. Brandts, J. Am. Chem. Soc., 86, 4291 (1964).

16. R. Biltonen and R. Lumry, J. Am. Chem. Soc., 91, 4251 (1969).

17. W. H. Porter and J. L. Preston, Anal. Biochem., 66, 69 (1975).

18. R. Biltonen and R. Lumry, J. Am. Chem. Soc., 91, 4256 (1969).

19. W. T. Astbury, S. Dickinson, and K. Bailey, Biochem. J., 29, 2351 (1935).

20. K. W. Snape, R. Tjian, C. C. F. Blake, and D. E. Koshland, Jr., Nature (London), 250, 295 (1974).

21. A. Yonath, A. Sielecki, J. Moult, A. Podjarny, and W. Traub, Biochemistry, 16, 1413 (1977).

22. A. Yonath, A. Podjarny, B. Houig, A. Sielecki, and W. Traub, Biochemistry, 16, 1418 (1977).

23. B. S. Hartley, Nature (London), 201, 1284 (1964).

24. B. S. Hartley and D. L. Kaufmann, Biochem. J., 101, 229 (1966).

25. D. M. Blow, J. J. Birktoft, and B. S. Hartley, Nature (London), 221, 337 (1969).

26. B. Meloun, I. Kluh, V. Kostka, L. Moravek, Z. Prusik, J. Vanacek, B. Keil, and F. Sorm, Biochim. Biophys. Acta, 130, 543 (1966).

27. L. B. Smilie, A. Furka, N. Nagabhushan, K. J. Stevenson, and C. O. Parkes, Nature (London), 218, 343 (1968).

28. W. J. Dreyer and H. Neurath, J. Biol. Chem., 217, 527 (1955).

29. M. Rovery, A. Curnier, and P. Desnuelle, Biochim. Biophys. Acta, 17, 565 (1955).

30. P. Desnuelle, The Enzymes, vol. IV, Second Ed., P. D. Boyer, H. Lardy, and K. Myrback, eds., Academic Press, New York, 1960, p.93.

31. R. B. Corey, O. Battgay, D. A. Breuckner, and F. G. Mark, Biochim. Biophys. Acta, 94, 535 (1965).

32. J. Kraut, H. T. Wright, M. Kellerman, and S. T. Freer, Proc. Nat. Acad. Sci. USA, 58, 304 (1967).

33. H. T. Wright, J. Kraut, and P. E. Wilcox, J. Mol. Biol., 37, 363 (1968).

34. S. T. Freer, J. Kraut, J. D. Robertus, H. T. Wright, and Ng. H. Xuong, Biochemistry, 9, 1997 (1970).

35. J. J. Birktoft, J. Kraut, and S. T. Freer, Biochemistry, 15, 4481 (1976).

36. D. M. Blow, M. G. Rossman, and B. A. Jeffrey, J. Mol. Biol., 8, 65 (1964).

37. K. C. Aune and S. N. Timasheff, Biochemistry, 10, 1609 (1972).

38. R. L. Vandlen and A. Tulinsky, Biochemistry, 12, 4193 (1973).

39. B. W. Matthews, P. B. Sigler, R. Henderson, and D. M. Blow, Nature (London), 214, 652 (1967).

40. J. J. Birktoft and D. M. Blow, J. Mol. Biol., 68, 187 (1972).

41. A. Tulinsky, N. V. Mani, C. N. Morimoto, and R. L. Vandlen Acta Cryst., B29, 1309 (1973).

42. A. Tulinsky, R. L. Vandlen, C. N. Morimoto, N. V. Mani, and L. H. Wright, Biochemistry, 12, 4185 (1973).

43. D. C. Phillips, British Biochemistry Past and Present, T. W. Goodwin, ed., Academic Press, London, 1970, p. 11.

44. K. Nishikawa, T. Ooi, Y. Isogai, and N. Saito, J. Phys. Soc. Japan, 32, 1331 (1972).

45. I. D. Kuntz, J. Am. Chem. Soc., 97, 4362 (1975).

46. M. N. Liebman, Ph.D. Dissertation, Michigan State University, 1977.

47. M. O. Daynoff, Atlas of Protein Sequence and Structure, vol. 5, National Biomedical Research Foundation, Silver Spring, Maryland, 1972.

48. K. Nishikawa and T. Ooi, J. Theor. Biol., 43, 351 (1974).

49. N. V. Raghavan and A. Tulinsky, unpublished results of this laboratory.

50. M. G. Rossman and A. Liljas, J. Mol. Biol., 85, 177 (1974).

51. A. Tulinsky, P. W. Codding, and R. L. Vandlen, Abstracts, American Crystallographic Association, Columbia, S.C., Feb. 1972, No. K3.

52. R. L. Vandlen, Ph.D. Dissertation, Michigan State University, 1972.

53. M. V. King, Acta Cryst., 7, 601 (1954).

54. R. L. Vandlen and A. Tulinsky, Acta Cryst., B27, 437 (1971).

55. H. W. Wyckoff, M. Doscher, D. Tsernoglou, T. Inagami, L. N. Johnson, K. D. Hardman, N. M. Allewell, D. M. Kelly, and F. M. Richards, J. Mol. Biol., 27, 563 (1967).

56. A. C. T. North, D. C. Phillips, and F. S. Mathews, Acta Cryst., A24, 351 (1968).

57. D. M. Blow and F. H. C. Crick, Acta Cryst., 12, 794 (1959).

58. R. Henderson and K. Moffat, Acta Cryst., B27, 1414 (1971).

59. D. W. J. Cruickshank, Acta Cryst., 2, 65 (1949).

60. D. W. J. Cruickshank, Acta Cryst., 3, 72 (1950).

61. T. L. Blundell and L. N. Johnson, Protein Crystallography, Academic Press, New York, 1976.

62. V. Luzzati, Acta Cryst., 6, 142 (1953).

63. L. O. Ford, L. N. Johnson, P. A. Machin, D. C. Phillips, and R. Tjian, J. Mol. Biol., 88, 349 (1974).

64. F. M. Richards, J. Mol. Biol., 37, 225 (1968).

65. A. Tulinsky, I. Mavridis, and R. Mann, J. Biol. Chem., in the press.

66. A. Tulinsky and L. H. Wright, J. Mol. Biol., 81, 47 (1973).

67. A. Mavridis, A. Tulinsky, and M. N. Liebman, Biochemistry, 13, 3661 (1974).

68. A. E. Smith, Acta Cryst., 5, 224 (1952).

69. C. Tanford, K. Kawahara, and S. Lapanje, J. Am. Chem. Soc., 89, 729 (1967).

70. Y. Nozaki and C. Tanford, J. Am. Chem. Soc., 89, 742 (1967).

71. C. Tanford, K. Kawahara, S. Lapanje, T. M. Hooker, Jr., M. H. Zarlengo, A. Salahuddin, K. C. Aune, and T. Takagi, J. Am. Chem. Soc., 89, 5023 (1967).

72. N. V. Raghavan and A. Tulinsky, Abstracts, American Crystallographic Association, East Lansing, Michigan, August, 1977, No. PB22.