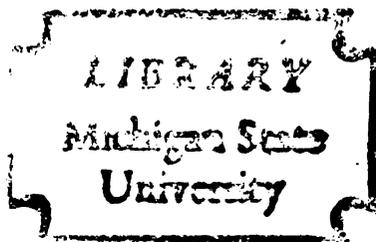




THESIS



This is to certify that the

dissertation entitled

Comparison of the Structures  
of the Two Independent Molecules  
of the  $\alpha$ -Chymotrypsin Dimer at 1.8 Å Resolution  
presented by

Mark A. Frentrup

has been accepted towards fulfillment  
of the requirements for

Ph.D. degree in Chemistry

*A. Tulinsky*

Major professor

Date Feb. 3, 1982.



RETURNING MATERIALS:  
Place in book drop to  
remove this checkout from  
your record. FINES will  
be charged if book is  
returned after the date  
stamped below.

--	--	--

COMPARISON OF THE STRUCTURES OF THE TWO  
INDEPENDENT MOLECULES OF THE  $\alpha$ -CHYMOTRYPSIN  
DIMER AT 1.8 Å RESOLUTION

By

Mark Andrew Frentrup

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Chemistry

1981

## ABSTRACT

### COMPARISON OF THE STRUCTURES OF THE TWO INDEPENDENT MOLECULES OF THE $\alpha$ -CHYMOTRYPSIN DIMER AT 1.8 Å RESOLUTION

By

Mark Andrew Frentrup

Alpha chymotrypsin (CHT) crystallizes at pH 4.0 in space group  $P2_1$  with two molecules per asymmetric unit. Significant differences between the two molecules of the asymmetric unit were observed at 2.8 Å resolution. The present work extends those observations to 1.8 Å resolution, yielding a detailed, quantitative comparison of the two independent molecules.

A Kendrew model of each molecule of the CHT dimer was built to the electron density. In order to retain small differences in model structure, the coordinates were measured with a very accurate polar coordinate measurement technique. A surveyor's transit and a cathetometer were used to record coordinates with a precision of 0.05 Å. These raw measured coordinates were then idealized, and the r.m.s. deviation between measured and idealized coordinates was 0.17 Å.

The idealized coordinates of the two independent molecules of CHT were compared, and the transformation between the two molecules was selected as that which produced the best least squares fit (0.43 Å r.m.s. deviation) of the

main chain of 68 residues which were built with substantially the same structure in the two molecules. The dimer molecules differ overall by 1.66 Å (r.m.s. deviation). The values for the exterior, interior, and dimer interface regions are 1.71, 1.25, and 1.96 Å, respectively. Even the interior of the molecule shows significant variability in structure.

The independent molecules of the CHT dimer were compared to  $\gamma$ -chymotrypsin, a monomeric form of chymotrypsin which crystallizes at pH 5.5. The structure of  $\gamma$ -CHT has recently been refined at 1.9 Å resolution. Large variations in structure were seen in the region Val 65-Ser 92 (3.3 Å and 7.0 Å r.m.s. deviation for main chain and side chain, respectively). The overall r.m.s. deviation between  $\gamma$ -CHT and  $\alpha$ -CHT is 2.4 Å. More detailed comparison can be made when the  $\alpha$ -CHT structure has been refined to an extent comparable to that of  $\gamma$ -CHT.

To Julie, and to our families.

## ACKNOWLEDGEMENTS

My sincere thanks to Professor Alexander Tulinsky for his guidance, patience, and support throughout this study and my years at Michigan State.

I am also grateful to Dr. N.V. Raghavan for his counsel and friendship through my undergraduate and graduate years, and for his collaboration in the initial stages of this work.

Dr. T.V. Atkinson has been indispensable for his invaluable role in the acquisition and maintenance of our computer facility. He has helped this laboratory run smoothly for years and his help is gratefully acknowledged.

My thanks also to Mr. Richard Blevins, for his help with some of the figures and tables.

## TABLE OF CONTENTS

	Page
LIST OF TABLES . . . . .	vi
LIST OF FIGURES . . . . .	vii
I. Introduction . . . . .	1
1. Non-crystallographic symmetry . . . . .	1
2. Alpha-chymotrypsin . . . . .	6
3. Gamma-chymotrypsin . . . . .	9
II. Coordinate Measurement of the Independent Molecules of $\alpha$ -Chymotrypsin . . . . .	10
1. Introduction . . . . .	10
2. Experimental . . . . .	12
a. The Transit . . . . .	14
b. The Measurements . . . . .	14
c. Coordinate Calculation and Transformation	16
d. Problem Areas . . . . .	22
3. Results . . . . .	25
III. The $\alpha$ -Chymotrypsin Model . . . . .	31
1. The 1.8 $\text{\AA}$ Resolution Electron Density Map . .	31
2. The Difference Density and the Model . . . .	35
3. The Two Molecules . . . . .	36

	Page
4. Solvent Structure . . . . .	40
5. Model Conformation . . . . .	43
IV. Comparison of the Structures of the Independent Molecules of $\alpha$ -Chymotrypsin Dimer . . . . .	49
1. Determination of the symmetry axis . . . . .	49
a. From the Electron Density Map . . . . .	49
b. From the Coordinates . . . . .	54
c. Comparison of the Two Transformations . . . . .	58
2. The two molecules of $\alpha$ -CHT . . . . .	64
3. Variability in $\alpha$ -CHT structure . . . . .	70
4. Significance of the differences . . . . .	74
V. Comparison to Other Structures . . . . .	81
1. $\alpha$ -CHT and $\gamma$ -CHT . . . . .	81
2. Concluding remarks . . . . .	91
LIST OF REFERENCES . . . . .	93

LIST OF TABLES

Table		Page
1	Measured Geometry of the Main Chain . . . . .	29
2	Partial List of Potential Solvent Molecules. Molecule One . . . . .	41
3	Partial List of Potential Solvent Molecules Present in Molecule One Prime But Not in Molecule One . . . . .	42
4	Transformation Parameters, Molecule One Prime to Molecule One . . . . .	63
5	Comparison of Molecule One and Molecule One Prime . . . . .	65
6	Comparison of $\alpha$ -CHT to $\gamma$ -CHT . . . . .	82
7	R.m.s. Deviations of Val 65-Ser 92 in $\alpha$ -CHT and $\gamma$ -CHT . . . . .	85

## LIST OF FIGURES

Figure		Page
1	Geometry of polar coordinate measurement . . .	13
2	Transformation from transit coordinates to unit cell coordinates. T - transit position; M - origin marker; O - unit cell origin. The parallelogram represents half-silvered mirror of the Richards box. a) Subtraction of origin marker coordinates; b) rotation to $\vec{a}^*$ , $\vec{b}$ , $\vec{c}$ , followed by subtraction of unit cell origin coordinates, c) rotation around $\vec{y}$ or $\vec{b}$ by $\beta-90^\circ$ . . . .	19
3	Graphics display of Trp 27 - Pro 28 - Trp 29, calculated directly from measured transit coordinates . . . . .	24
4	Histogram of Model Build results . . . . .	28
5	Ramachandran plot for molecule one . . . . .	44
6	Ramachandran plot for molecule one prime . . . . .	45
7	Main chain torsion angles of a protein . . . . .	47
8	Equation of the two fold axis . . . . .	52
9	The region Gly 216-Thr 219, viewed down the local two fold axis . . . . .	67
10	Comparison of alpha carbon positions in the two molecules of $\alpha$ -chymotrypsin . . . . .	69
11	Interaction of Leu 155 with Gly 140-Trp 141 in molecule one (open bonds) and one prime (dark bonds) . . . . .	73

Figure		Page
12	Superposition of the active sites of molecule one (open bonds) and one prime (dark bonds) of $\alpha$ -CHT . . . . .	76
13	Comparison of alpha carbon positions in $\gamma$ -CHT and molecule one of $\alpha$ -CHT . . . . .	86
14	Comparison of alpha carbon positions in $\gamma$ -CHT and molecule one prime of $\alpha$ -CHT . . . . .	87
15	Superposition of the active sites of $\gamma$ -CHT (dark bonds) and molecule one of $\alpha$ -CHT . . . . .	89
16	Superposition of the active sites of $\gamma$ -CHT (dark bonds) and molecule one prime of $\alpha$ -CHT . . . . .	90

## CHAPTER 1

### Introduction

Alpha chymotrypsin (CHT), a pancreatic serine protease, is a monomer in its active form near neutral pH. As pH is lowered, dimerization occurs concomitant with loss of enzyme activity. The low pH, dimeric form of CHT crystallizes readily, and yields crystals whose diffraction pattern extends to approximately 1.8 Å resolution and which are quite resistant to decay in an x-ray beam. Thus, they were an excellent candidate for study by x-ray crystallography and have been so investigated for many years, both by D.M. Blow and coworkers in Cambridge, England<sup>1,2</sup> and in this laboratory by A. Tulinsky and colleagues.<sup>3-5</sup>

#### 1. Non-crystallographic symmetry

CHT crystallizes at pH 3.5 in the non-centrosymmetric space group  $P2_1$  with four molecules per unit cell, which means that there are two molecules per asymmetric unit.

The two molecules of the asymmetric unit are related to each other by a non-crystallographic symmetry element. The groups at Cambridge and Michigan State have confirmed that the symmetry element is an approximate two fold rotation axis, i.e. a rotation of  $180^{\circ}$ . Since the symmetry element is non-crystallographic, there is no a priori reason for the two molecules to have exactly the same structure. Many other proteins crystallize as dimers, tetramers, hexamers, and even trimers, and in many cases the symmetry element which relates the individual subunits of the oligomeric system coincides closely or exactly with a crystallographic axis, so that the symmetry is constrained by the space group to be an exact symmetry axis.<sup>6-11</sup> It is not clear why some oligomeric systems crystallize with their symmetry axis along an exact crystallographic symmetry axis while CHT and others<sup>12-16</sup> do not. Since CHT crystallizes as a dimer in the asymmetric unit and because it has a diffraction pattern that extends to relatively high resolution, it affords an excellent opportunity to determine, by careful measurement and interpretation, the degree to which the individual subunits of CHT differ in structure. Through such study, one can hope to lay down the principles for structural variability in oligomeric protein systems.

It is possible to study non-crystallographic symmetry, even in systems which display exact crystallographic symmetry, by not assuming the symmetry axis and treating the crystal system as one of lower symmetry and larger

asymmetric unit. As an example, consider 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase which crystallizes in the cubic space group  $P2_13$ . The enzyme is a trimer, each subunit of which is related to the other two by a crystallographic three fold axis.<sup>9</sup> There are 12 molecules per unit cell, one molecule per asymmetric unit. If the three fold axes are not assumed, the space group becomes the orthorhombic group  $P2_12_12_1$ , still with 12 molecules per unit cell, but now with three molecules per asymmetric unit.

Exact crystallographic symmetry affects the diffraction pattern in one of two ways: 1) certain classes of reflections are rendered identically zero (these classes are called systematic absences), or 2) certain classes of reflections become equivalent in phase and intensity to certain other classes. The first effect is caused by symmetry elements such as screw axes and glide planes which have translational components. The second effect is the consequence of pure rotation, with no translational components. As an example of 2), consider the addition of a two fold axis of symmetry to a unit cell of the triclinic space group  $P1$ . This changes the space group to the monoclinic  $P2$  which results in a situation in which all reflections  $h\bar{k}l$  are now equivalent in phase and intensity to reflections  $hkl$  and so the former need not be collected since they are not unique. Thus, data collection time is considerably reduced (in this case by a factor of  $\frac{1}{2}$ ) and the structure solution is less work. In

such a case the electron density of the subunits related by the crystallographic symmetry would be calculated as exactly equal by symmetry. If the individual subunits of an oligomeric system are not exactly equal in structure, then the crystallographic symmetry breaks down, with the result that there will be small differences in what would have been equivalent reflections had the symmetry been exact. Thus, in the same example, reflections  $h\bar{k}l$  would be slightly different in intensity and phase from reflections  $hkl$ , the differences depending on the degree of difference in the structure of the subunits. As the difference in structure becomes smaller, the reflections  $hkl$  and  $h\bar{k}l$  become more nearly equivalent. Very accurate intensity measurements would be required to differentiate such small intensity differences. Since in addition, the amount of data to be collected would increase drastically, systems in which the oligomeric axis of symmetry lies along a crystallographic symmetry direction have always been treated as systems with exact symmetry,<sup>6-11</sup> and no doubt justifiably so in many if not most cases.

In other oligomeric systems, the intermolecular symmetry axis is not coincident with a crystallographic axis<sup>12-16</sup> and the asymmetric unit contains a larger part of or the entire oligomer. Such is the case with CHT. The two fold axis which relates the two subunits of CHT dimer can not be interpreted as a crystallographic two fold axis in any unit

cell setting. Therefore, it is referred to as a non-crystallographic, or local, two fold axis. Here the non-crystallographic symmetry has no effect on the equivalence of certain classes of reflections, and therefore an assumption of the exactness or inexactness of the local two fold rotation has no effect on the amount of data collected. It is in the interpretation of the data that that assumption is made. The electron density calculated from the data for the two regions related by the non-crystallographic symmetry would not be equal. Differences in density would be observed, and the significance of these differences could be assessed on the basis of the expected error in the electron density. Again, accurate intensity measurements are required to establish with confidence the significance of small differences in electron density.

Such non-crystallographic symmetry can and has been used as a method to solve structures.<sup>18-20</sup> The Fourier series, which is the electron density equation, is required to converge to the same values in each of the regions of the asymmetric unit corresponding to the boundaries of the individual subunits; this is a powerful constraint on the data.<sup>18,21</sup> The method has been used to great advantage, particularly in virus structures,<sup>22-25</sup> in which high order non-crystallographic symmetry is often present. A disadvantage of this method is the loss of whatever information there might have been in the data regarding the non-equivalence of subunits.

Another common technique for the treatment of non-crystallographic symmetry is to average the electron density of the individual subunits. Again, the method disregards the possibility of the existence of significant differences among subunits.<sup>2,26-28</sup> In fairness to these applications, it should be noted that in most of these systems, the assumption of exact symmetry is justified to the extent that the resolution is relatively low and/or the data are relatively imprecise. Thus in many of the investigations cited, the investigators worked with data collected on film, often at relatively low resolution ( $>2.5 \text{ \AA}$ ). Moreover, the phase angles for some of these structures have been determined from a single isomorphous derivative, seldom from more than two. In such cases the averaging process can actually improve what would otherwise be a poorer quality map. In contrast, the structure solution of CHT at Michigan State has revolved around innovative diffractometer measurements.<sup>5,29</sup> Phases were first determined at  $2.8 \text{ \AA}$  resolution from six isomorphous derivatives and then carefully extended to  $1.8 \text{ \AA}$  resolution by means of a very mild density modification procedure.<sup>5,30</sup>

## 2. Alpha-chymotrypsin

The structure of the dimeric form of CHT was determined in the late sixties and early seventies by Blow and co-workers,<sup>1,2</sup> when work by Tulinsky et al. was also well in progress. Blow et al. concluded that their data did not

suggest significantly different structures for the molecules of the dimer, and so they calculated an average map as described above, and interpreted CHT as an exact dimer.<sup>1,2</sup> However, they did allow for the variation of the conformation of certain of the exterior side chains. The experience at Michigan State was different from the outset. Here, evidence of non-equivalence of the two molecules was observed in the asymmetric binding of the heavy atoms used in the phase determination.<sup>3,4</sup> In this context symmetry and asymmetry refer to the equivalence or difference, respectively, of the response of the two independent molecules of CHT to various conditions. Further work with the binding of competitive and reversible inhibitors,<sup>31,32</sup> pH changes,<sup>33,34</sup> and other studies<sup>35,36</sup> confirmed the asymmetry of CHT, sometimes in striking ways. As an example, when the pH of the crystal soaking solution was raised to 8.3, the startling observation was made that in one molecule the ion pair between the carboxylate group of Asp 194 and the positively charged amino group of the B chain N terminus Ile 16 was disrupted by deprotonation of the latter, while in the other molecule the ion pair remained intact.<sup>34</sup> The interpretation of such a result was that the pK of the B chain terminal amino group is different in the two molecules, and this implies that the structures are different. Asymmetry of response of CHT to small molecule binding has thus been overwhelmingly documented.<sup>32</sup> The purpose of the current

study is to begin the documentation of the structural basis which underlies and is manifested in the expression of asymmetry by preparing a set of CHT coordinates. These coordinates would also be suitable for further refinement by least squares methods.<sup>37</sup>

The electron density map used in the current work is the 1.8 Å resolution map calculated by Raghavan and Tulinsky.<sup>5</sup> Intensities were collected with a wandering count 6-drop 2  $\omega$  step scan.<sup>29</sup> Crystal alignment was carefully maintained by monitoring the intensities of three reflections suitably distributed in reciprocal space. These intensities were measured approximately every hour during the data collection, and when any of the monitor intensities decreased by 10%, the crystal was realigned by accurately centering eight selected reflections and finding the crystal orientation by the method of least squares from the Miller indices and diffractometer angles of these reflections. Phases were determined at 2.8 Å resolution by the method of multiple isomorphous replacement,<sup>4</sup> and the phases were refined and extended to 1.8 Å resolution by density modification.<sup>5,30</sup> The details of the method are described elsewhere.<sup>5,38</sup>

A crucial point is that the method does not require the supposition of an atomic model for the phase extension. Thus the final map is not biased with respect to an assumed model, and should truthfully display the structural differences between the two molecules.

### 3. Gamma-chymotrypsin

It was mentioned above that chymotrypsin is a monomer near pH 7. This monomeric form is called  $\gamma$ -CHT and crystallizes at pH 5.5 in the tetragonal space group  $P4_22_12$ . The 1.9 Å refined structure of  $\gamma$ -CHT has recently been reported and compared with the averaged  $\alpha$ -CHT structure of Blow.<sup>39</sup> Earlier, a low resolution electron density map of  $\gamma$ -CHT was compared with Blow's averaged  $\alpha$ -CHT structure<sup>40</sup> and no important differences were found. The most significant observed differences in the higher resolution study were in residues 74-78 and in the side chain orientation of the active site Ser 195.<sup>41</sup> Furthermore, several residues are disordered in  $\gamma$ -CHT which are not disordered in  $\alpha$ -CHT, and these residues were not compared. It will be of interest to compare both molecules of our 1.8 Å model to  $\gamma$ -CHT and to analyze this comparison in light of the previous comparison of  $\gamma$ -CHT to an averaged structure of  $\alpha$ -CHT.

## CHAPTER 2

### Coordinate Measurement of the Independent Molecules of $\alpha$ -Chymotrypsin.

#### 1. Introduction

The phase extended 1.8 Å resolution map permits an excellent fit of a model to the density: side chain densities are usually clear and unambiguous, and many of the carbonyl oxygen atoms of the main chain are visible as bumps protruding at regular intervals from the main chain density. A Kendrew model of each of the subunits of CHT was built to fit this map with the aid of a Richards optical comparator.<sup>42</sup> Kendrew model parts consist of brass rods approximately 2 mm in diameter which are machined to a scale of 2 cm per Ångstrom into pieces which represent parts of protein in a standard geometry, e.g. planar hexagonal phenyl rings, planar peptides, and tetrahedral carbon atoms.

Coordinate measurement of a Kendrew model of a protein is usually a prerequisite to further studies by computer

methods. With the advent of interactive computer graphics facilities, Tsernoglou et al.<sup>43</sup> demonstrated the feasibility of obtaining a set of molecular coordinates without building a Kendrew model. However, such facilities are not commonly available and might even prove difficult to use with more complicated systems than a snake venom  $\alpha$ -neurotoxin, which has a molecular weight of only 13,000; the CHT dimer has a molecular weight of 50,000 daltons. Since it was also desirable for pedagogical reasons to have a physical model of CHT, it was necessary to devise a rapid, convenient, and accurate technique to record coordinates of such a model. The requirement for accuracy was especially important in our case because of the quality and high resolution of the electron density map; significant features in the map were surely incorporated into the model, and accurate coordinates would be required in order to avoid losing the significance of the features as a result of large uncertainties in measured coordinates. Since the measured coordinates were to be used subsequently as guide coordinates for Diamond's coordinate idealization procedure,<sup>44</sup> even greater accuracy for measured coordinates was required in order to avoid loss of significant features owing to idealization of coordinates.

We considered the use of a plumb line imprecise and unsatisfactory to achieve our purpose. In addition, the method is very cumbersome in practice, because it requires

physical contact with the model and introduces the likelihood of inadvertent readjustment. We had previously measured the coordinates of a model of CHT based on a 2.8 Å resolution map from a calibrated grid inserted in the Richards box at the same depth as the atom to be measured, but this method was rejected on similar grounds. On the other hand, more precise and accurate methods, including an elegant computerized "coordinate hunting engine" designed by Saleme and Fehr<sup>45</sup> appeared to involve considerable time in the design and construction of a measuring device.<sup>46</sup> We found a satisfactory compromise and solution to the problem by successfully employing a surveyor's transit as a coordinate measuring instrument.<sup>47</sup> The transit is capable of high accuracy and precision and is conveniently complemented by a laboratory cathetometer. Together, these two instruments measure the polar coordinates of a Kendrew model very accurately and reproducibly.

## 2. Experimental

The geometry of the measurement of the atomic positions of a Kendrew model with a surveyor's transit is shown in Figure 1. The polar angles  $\phi$  and  $\chi$  are measured with the transit located at the origin while the height coordinate,  $y$ , is recorded with a cathetometer. Thus, the method records polar coordinates directly whereas previous methods for measuring Kendrew model coordinates have generally been based on the direct measurement of Cartesian coordinates.<sup>2,45,46,48-50</sup>

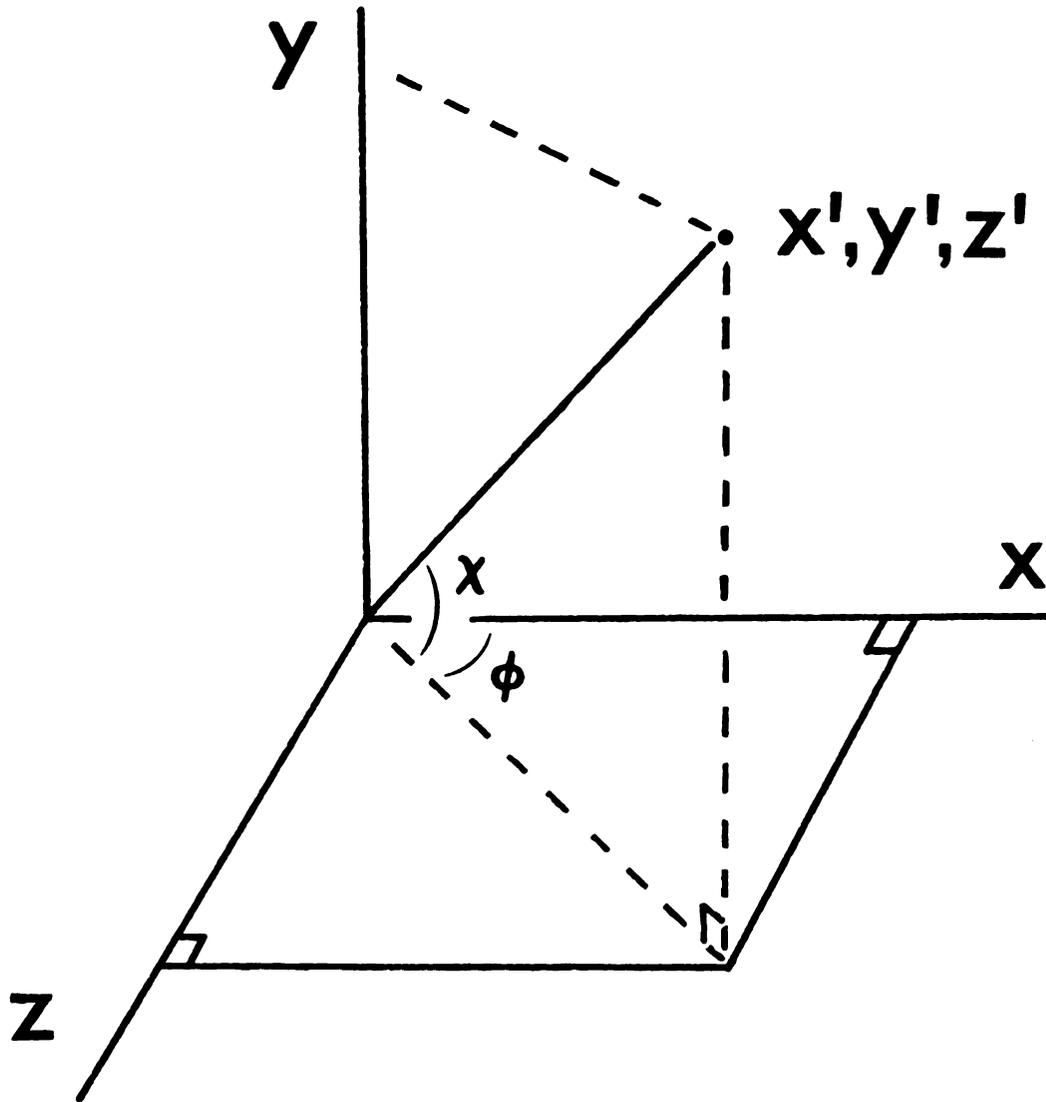


Figure 1. Geometry of polar coordinate measurement.

a. The Transit

The transit used in the present work was a David White/Path Model TR303 made by Fuji-Koh of Japan. This transit has a 27x power telescope with crosshairs (minimum focal distance 1.7 meters) mounted on rotatable horizontal and vertical circles. The two circles can be locked and moved with tangent screws, which provide a fine adjustment, or they can be unlocked to rotate freely. The divisions of the horizontal circle are in 20 minutes of arc ( $0.33^{\circ}$ ) with a vernier reading to 20 seconds ( $0.006^{\circ}$ ); those of the vertical circle are 30 minutes ( $0.5^{\circ}$ ) and 1 minute ( $0.017^{\circ}$ ), respectively. The magnification of the cathetometer telescope is similar to that of the transit and the height range is 100 cm with a vernier measurement capable of 0.05 mm. The price of the transit in 1979 was approximately \$1500.

b. The Measurements

The actual measurement by the transit or cathetometer is accomplished by sighting in the telescope cross hairs as closely as possible to the intersection of the Kendrew model rods representing the various bonds between atoms. An exception is the carbonyl oxygen positions, which are measured at the terminus of the rod representing the carbon-oxygen bond. All heights recorded by the cathetometer are corrected for the height of the transit above the origin of the cathetometer. This height is found as the cathetometer vernier reading when the cathetometer telescope is viewing

directly into the eyepiece of the transit with both telescopes leveled. With care, this measurement is reproducible to  $\pm 0.1$  cm.

Angular measurements were made to the nearest minute of arc so that the horizontal and vertical angle measurements might have comparable precision. These measurements are reproducible to  $\pm 1$  minute by different observers and reproducible virtually to the minute by the same observer. At a distance of 2 m, a one minute uncertainty in angular position corresponds to a deviation of about  $\pm 0.06$  cm ( $0.03 \text{ \AA}$ ) in atomic position. This correlates approximately with errors expected from the dimensions of the model itself: the diameter of the model parts is about 2.5 mm or about 4 x the uncertainty; it is possible to reproduce the position of the crosshair to about  $1/5$  the diameter of the rods. However, when the line of sight is not perpendicular to the bonds of an atom, it becomes progressively more difficult to estimate the position of the intersection of two bonds and the uncertainty in atomic position increases, perhaps by up to a factor of two.

Sometimes an atom is partially obscured because of intervening model parts. In most cases, it is still possible to obtain a reliable measurement by noting the two bonds which intersect at the desired atomic position and visually extrapolating to their (non-visible) intersection point. Only rarely does such an indirect determination yield a result which affects the ideal geometry as noted on a

graphics display of the measured coordinates (planar peptides and aromatic rings, tetrahedral angles). In fact, it is surprising that such problems are not encountered more often with our Kendrew model of 241 residue complexity. An atom is completely obscured approximately once every three residues and only about 20% of these atoms had to be disregarded because of an unacceptable coordinate. Thus, a total of about 15 atom coordinates were unmeasurable due to such problems. In principle, it is possible to reposition the transit and/or cathetometer so that such atoms are no longer obscured. On the other hand, careful positioning of the transit and cathetometer at the beginning of a series of measurements can avoid such difficulties. Finally, the transit and cathetometer telescopes can "focus around" intervening models parts that are far from the plane of focus so that more atoms are visible through the telescopes than to the eye.

c. Coordinate Calculation and Transformation

The measured polar coordinates are converted to Cartesian coordinates according to

$$\begin{aligned}
 y' &= y - h \\
 x' &= y' \cos\phi / \tan\chi \\
 z' &= y' \sin\phi / \tan\chi
 \end{aligned}
 \tag{2.1}$$

where  $y$  is the height measurement,  $h$  is the height of the transit telescope above the origin of the cathetometer scale

and  $\phi$  and  $\chi$  are the horizontal and vertical polar angles, respectively. The problem which remains is to convert these Cartesian coordinates, referred to an arbitrary origin and based on a general orientation with respect to the axes of the unit cell of the crystal, to Cartesian coordinates based rationally on unit cell axes. In the present monoclinic case, the latter Cartesian coordinates are taken as: x parallel to  $\vec{a}$ , y parallel to  $\vec{b}$ , and z parallel to  $\vec{c}^*$ .

The transformation can be accomplished in two steps: 1) measurement of the unit cell origin coordinates in the transit frame and subtraction of these from each of the atomic coordinates to refer the coordinates to the unit cell origin and 2) multiplication of the origin shifted coordinates by a rotation matrix that will rotate the transit x axis ( $\phi = 0^\circ$ ) into crystallographic  $\vec{a}$ , y into  $\vec{b}$ , and z into  $\vec{c}^*$ .

Two aspects of our Richards box prevented proceeding in such a straightforward fashion. First, the position of the unit cell origin (O of Figure 2) was outside the frame which contains the molecule and was thus inaccessible for direct measurement. Second, the direction  $\vec{a}$  was not readily measureable, whereas  $\vec{a}^*$  was, the latter being perpendicular to the half-silvered mirror ( $\vec{a}^*$  also coincides closely with a local two fold axis relating two molecules of  $\alpha$ -CHT in the asymmetric unit).

To circumvent the inaccessibility of the origin, we chose an arbitrary point on the local two fold axis for

Figure 2. Transformation from transit coordinates to unit cell coordinates. T - transit position; M - origin marker; O - unit cell origin. The parallelogram represents half-silvered mirror of the Richards box. a) Subtraction of origin marker coordinates; b) rotation to  $\vec{a}^*$ ,  $\vec{b}$ ,  $\vec{c}$ , followed by subtraction of unit cell origin coordinates, c) rotation around y or  $\vec{b}$  by  $\beta-90^\circ$ .

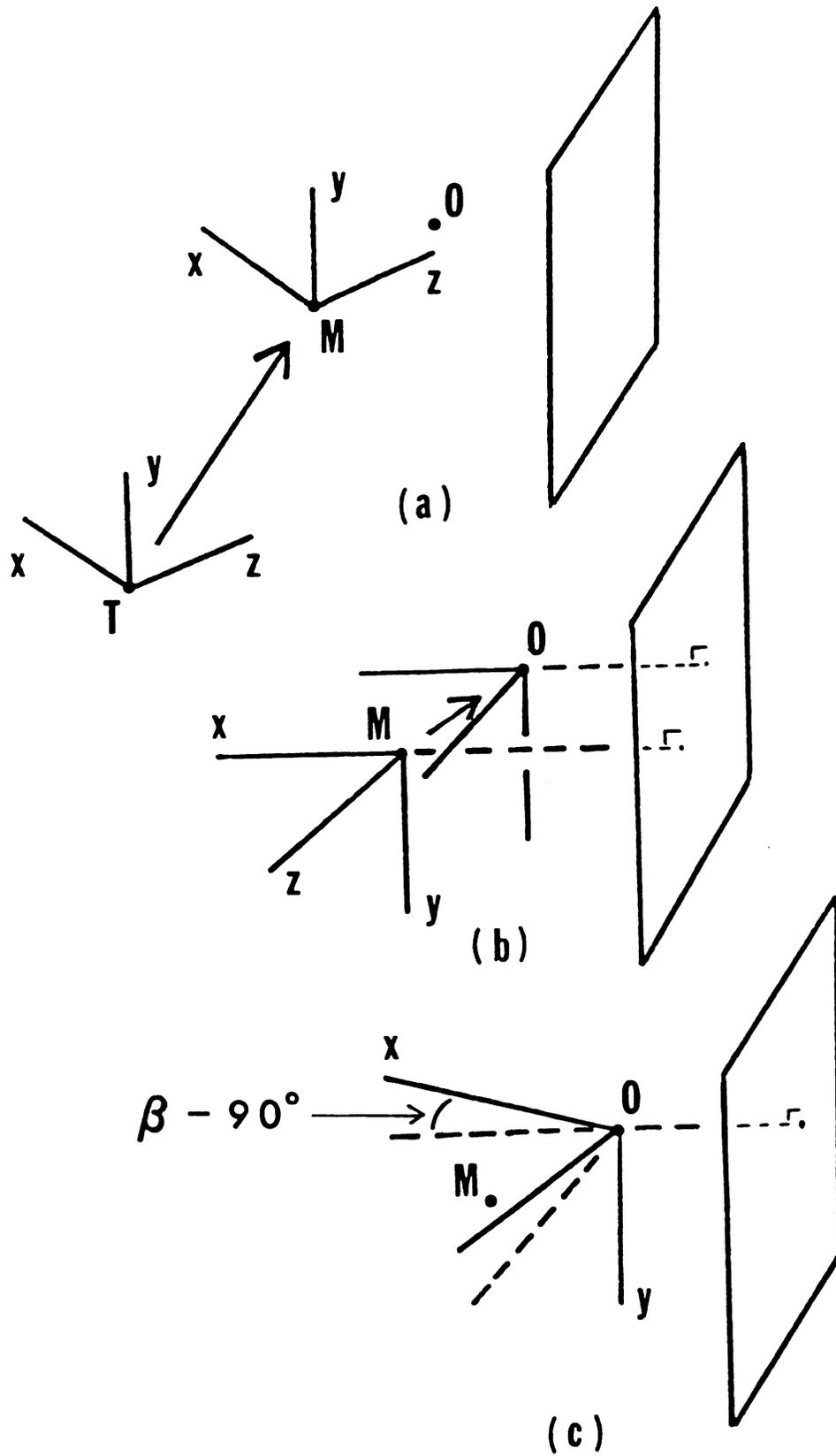


Figure 2

which precise unit cell Cartesian coordinates were obtainable and designated this point the origin marker (M of Figure 2).

In order to perform the translation-rotation transformation, the coordinates of the origin marker are measured in the transit coordinate frame and then subtracted from the transit frame coordinates of each atom. (Figure 2(a)). The orientation of this new coordinate system with respect to the unit cell axes is then found by determining the unit vectors in the transit reference frame parallel to  $\vec{a}^*$ , to  $\vec{b}$ , and to  $\vec{c}$ . In the present work,  $\vec{a}^*$  is perpendicular to the mirror while  $\vec{b}$  and  $\vec{c}$  are parallel to the mirror and mutually perpendicular. The  $\vec{a}^*$  direction was fixed by placing four tetrahedral Kendrew model parts in a line perpendicular to the mirror. In a similar way, five additional tetrahedra located a line parallel to  $\vec{c}$  and the mirror. The coordinates of these nine tetrahedra are measured in the transit frame and orthogonal unit vectors parallel to  $\vec{a}^*$  and  $\vec{c}$  are determined by the method of least squares. The actual angle between  $\vec{a}^*$  and  $\vec{c}$  before orthogonalization was measured consistently to be  $91^\circ$ . The unit vector parallel to  $\vec{b}$  is given by the cross product  $\vec{c} \times \vec{a}^*$ . The components of these three unit vectors can then be used to form a matrix which rotates the three transit frame axes to a set of axes parallel to  $\vec{a}^*$ ,  $\vec{b}$ , and  $\vec{c}$  (Figure 2(b)). The Cartesian components of the origin marker M along  $\vec{a}^*$ ,  $\vec{b}$ , and  $\vec{c}$  are then subtracted to refer the measured coordinates to the

origin of the unit cell. Finally the entire coordinate system is rotated by the amount  $\beta-90^\circ$  about  $\vec{b}$  to bring  $x$  parallel to  $\vec{a}$  rather than  $\vec{a}^*$ ,  $z$  parallel to  $\vec{c}^*$  rather than to  $\vec{c}$ , and which leaves  $y$  unchanged and parallel to  $\vec{b}$  (Figure 2(c)). Thus, the transformation can be written as

$$\vec{x} = R_\beta [R(\vec{x}' - \vec{t}) - \vec{x}_o], \quad (2.2)$$

where  $\vec{x}$  is the atomic position in unit cell Cartesian coordinates,  $R_\beta$  is the matrix for rotation of  $(\beta-90)^\circ$  about  $y$ ,  $R$  is the rotation matrix made up of the unit vectors parallel to  $\vec{a}^*$ ,  $\vec{b}$  and  $\vec{c}$  in the transit frame,  $\vec{x}'$  is the atomic position in the transit frame,  $\vec{t}$  is the position of the origin marker in the transit frame and  $\vec{x}_o$  is the position of the origin marker based on  $\vec{a}^*$ ,  $\vec{b}$ , and  $\vec{c}$ .

A convenience of the method is that, since the transformation between transit and unit cell Cartesian coordinates is easily programmed, there is no need for keeping the transit fixed at the same position throughout all the coordinate measurements. All that is required with each change in transit position is the remeasurement of the position of the origin marker  $M$  and of the nine marker tetrahedra along the  $\vec{a}^*$  and  $\vec{c}$  directions from which a new rotation matrix and translation vector are calculated. Thus, the transit position can be changed from day to day, or more often if necessary, without seriously disrupting the course of the measurements. As noted above, it is in fact desirable at times to reposition the transit so as to obtain a less

obstructed view of the atoms to be measured. The cathetometer can be relocated similarly to provide an optimal view. In this case, it is necessary only to remeasure the height of the transit telescope relative to the cathetometer.

We calculated Cartesian coordinates from the two transit angles and the height measurement with an interactive program operating on a Cal Data 135 computer (PDP 11/40 equivalent) which can also drive a Vector General Interactive Graphics System. The connectivity and geometry of the model was then checked by displaying a model of the measured coordinates on the Vector General display. Interatomic distances and bond angles are calculated at the same time and are used to isolate any larger measurement errors (which are remeasured) or typographical errors (which are corrected). A typical graphics display is shown in Figure 3.

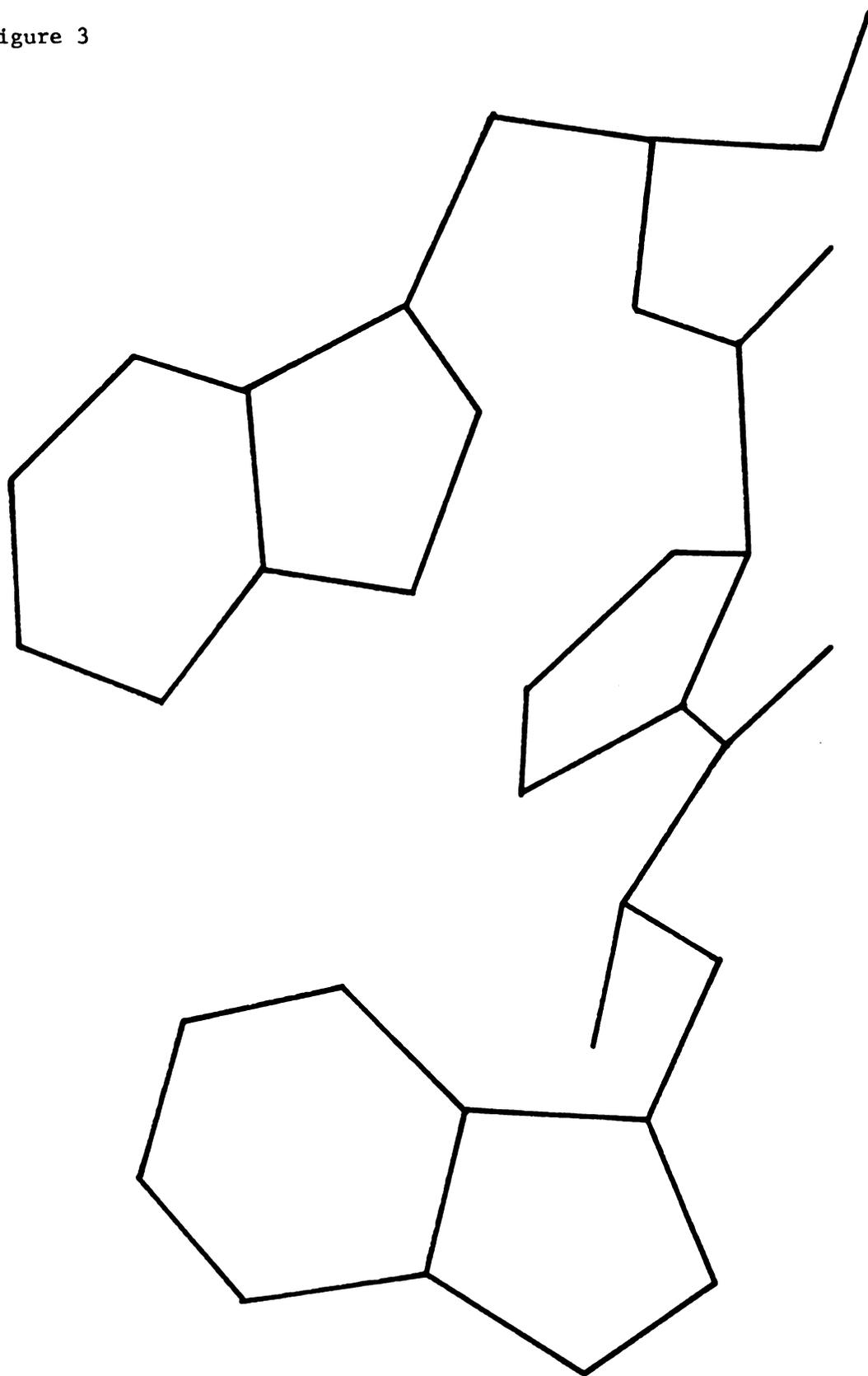
It is possible for one person to measure and calculate the coordinates of about 10 residues in four hours. Since the measurements are made with two instruments, use of a second person effectively halves this time.

#### d. Problem Areas

Vertical angle measurements of  $\chi$  less than about  $10^\circ$  have yielded Cartesian coordinates with large errors and measurements of less than  $5^\circ$  have led to coordinates so inconsistent as to be unacceptable. This is because of the functional dependence of the conversion from polar to Cartesian coordinates on  $1/\tan \chi$  (equation 2.1). The

Figure 3. Graphics display of Trp 27 - Pro 28 - Trp 29,  
calculated directly from measured transit  
coordinates

Figure 3



derivative of  $1/\tan \chi$  as  $\chi$  becomes small is very large so that a small error in  $\chi$  is translated into large errors in the calculated position. Bond distances of 3.5 Å and greater are not uncommon when calculated from measurements involving  $\chi$  angles of about 4-5°. Similarly, horizontal angles ( $\phi$ ) close to 0° or 90° are also to be avoided since  $\cos\phi$  and  $\sin\phi$  vary slowly with angle near these respective angles. Polar coordinate measurements in these regions also lead to calculated Cartesian coordinates with large errors. Therefore, all measurements described here were made at vertical angles of 10° or greater. Similarly, the horizontal angle measurements were made around 30-40°.

### 3. Results

An accurate set of Kendrew atomic model coordinates of both molecules of the dimeric structure of CHT fitted to the 1.8 Å resolution electron density map has been measured by using the polar to Cartesian coordinate conversion method just described. This entailed the measurement of the polar coordinates of about 3,600 atoms at least one time. The model structure was then idealized by using Diamond's model build program,<sup>44</sup> but with the standard amino acid structure dictionary of Hendrickson and Konnert's restrained least squares program.<sup>51</sup> The model build program uses the measured coordinates as guide atoms through which to thread an idealized polypeptide chain. The idealized chain is built from individual residues, the coordinates of which come from

x-ray or neutron diffraction structure determinations of the twenty amino acids. The ideal chain is initially layed down in an  $\alpha$ -helical conformation, and then the torsion angles are adjusted to give the best least squares fit between the guide coordinates and the altered idealized chain. Because the problem is drastically non-linear, convergence is a problem if an angle has to be changed by more than approximately  $90^\circ$  from its  $\alpha$ -helical value. Therefore, advantage was taken of a program option which allows unwinding of the ideal  $\alpha$ -helix before the least squares process begins. Values of all torsion angles were calculated from the "raw" measured coordinates, and these angles were applied to the ideal  $\alpha$ -helix in such a way as to give the ideal structure approximately the same conformation as the measured model. Thus, the idealized coordinates matched the measured coordinates fairly well even before the least squares refinement of the torsion angles. Because of this, convergence was possible and usually very rapid (2-4 cycles).

The r.m.s. difference between measured and idealized coordinates was  $0.17 \text{ \AA}$  for all atoms, while the average deviation was  $0.15 \text{ \AA}$ . For main chain atoms only, these values were  $0.16$  and  $0.14$ , respectively. The closeness in value of the average deviation and the r.m.s. value indicates a sharp distribution of error; of the 3054 atoms which were used as guide coordinates, 78% fit the final

idealized coordinates within  $0.2 \text{ \AA}$ , and only 2.5% had deviations greater than  $0.4 \text{ \AA}$ . The largest observed difference was  $0.82 \text{ \AA}$ , with only five other deviations greater than  $0.6 \text{ \AA}$ . These results are summarized in Figure 4. A comparison of main chain torsion angles calculated from the measured coordinates with the values of these angles in the idealized coordinate set shows no difference greater than 20 degrees between the two. Thus, the final idealized coordinates are a faithful representation of the measured coordinates, which gives us confidence that the idealized structure accurately reflects the conformation of the Kendrew model as it was built. A summary of the measured geometry of the 478 peptide bonds of both molecules of CHT is given in Table 1, from which it can be seen that the mean bond lengths agree exceptionally well with expected values. In the search for longest and shortest bonds, all bonds were included for which the positions of both atoms could be precisely determined or reasonably estimated. Similar criteria were used for inclusion of bond angles and torsion angles, so that each average in Table 1 is the mean value of approximately 450 determinations.

The observation that the average N-CA and CA-C bond lengths are slightly greater than expected while those of C-O and C-N are as expected is consistent with the fact that the N-CA and CA-C bonds are built by joining two Kendrew model pieces with a connector, while the C-O and C-N bonds



R.m.s. deviation of measured coordinates from idealized coordinates.

Figure 4. Histogram of Model Build results.

Table 1. Measured Geometry of the Main Chain\*

<u>Distances</u>	<u>Average (<math>\sigma</math>)</u>	<u>Largest</u>	<u>Smallest</u>	<u>Standard Values</u>
N-CA	1.50 (.09)	1.86	1.23	1.463
CA-C	1.56 (.09)	2.00	0.70	1.522
C-O	1.24 (.10)	1.52	0.82	1.234
C-N	1.32 (.09)	1.81	0.81	1.314
<u>Angles</u>				
N-CA-C ( $\tau$ )	110.5 (5.0)	126.4	94.0	109.54
Ca <sub>i</sub> -C-N-CA <sub>i+1</sub> ( $\omega$ )	179.5 (7.3)	197.7	153.3	180.1

\*Distances in Angstroms, angles in degrees.

are machined in a single piece. The natural error caused by the joint is to increase the bond length. While this is not evident from inspection of individual residues, the general trend is clear.

## CHAPTER 3

### The $\alpha$ -Chymotrypsin Model.

In chapter two, the results of the coordinate measurement of both molecules of the CHT dimer were described. Here, the structure of CHT will be discussed as well as observed differences between the two molecules. In the following chapter, the differences will be discussed in a more quantitative manner.

#### 1. The 1.8 Å Resolution Electron Density Map

In order to begin the fit of a model of CHT, the 1.8 Å resolution map was contoured and traced onto plexiglass sheets, and these sheets were placed in a Richards box for model construction purposes. The model fitted to the 1.8 Å resolution density map will henceforth be referred to as the 1.8 Å model, while the model built with the previous 2.8 Å resolution map will be called the 2.8 Å model. The two independent molecules of the CHT dimer will be designated as molecule one and molecule one prime. Molecule one occupies approximately the region  $y = 0 \rightarrow \frac{1}{2}$ ,  $z = 0 \rightarrow \frac{1}{2}$  in

the unit cell. The 2.8 Å model corresponded to molecule one; molecule one prime was not built at 2.8 Å resolution. Therefore, the 1.8 Å model will here refer to molecule one. Later, molecule one prime will be discussed in terms of differences from molecule one.

Since the 2.8 Å model was already in place in the Richards box, the 1.8 Å model was built by readjustment of the 2.8 Å model. The rebuilding was total-every atom of the molecule was adjusted in the process of fitting to the new density. To accomplish this, a typical procedure was to remove 3-5 residues from the 2.8 Å model and build these residues in the conformation suggested by the 1.8 Å resolution map. Thus the 1.8 Å model was built without reference to the details of the 2.8 Å model, and any structural differences between the 2.8 Å and 1.8 Å models should reflect either a different interpretation of the electron density map, or improved resolution which allowed a more certain fit of the 1.8 Å model. The presence of the 2.8 Å model helped, however, in the fit of the 1.8 Å model in that, while attention was not paid to the fine details of the 2.8 Å model, advantage was taken of the presence of the unadjusted parts of the 2.8 Å model by noting interatomic interactions between different parts of the CHT model. Thus, very close contacts in the constructed model could be avoided better by referring to the electron density and the unadjusted parts of the model than by referring to the electron density alone.

The unadjusted portions of the model also gave indications of possible hydrogen bonds and ionic interactions to the portion under adjustment. These indications influenced the choice of main chain and side chain conformation, and again the influence was greater with the model present than with only the electron density map for reference. In summary, the 1.8 Å model is close in structure to the 2.8 Å model, especially with regard to the gross features such as overall folding. However, there is no systematic equivalence in fine detail between the two, and the 1.8 Å model must be considered an independent interpretation of the structure of CHT. This interpretation should also be an improved one, because of the higher resolution of the electron density map, and because of the greater care that was taken in the building of the model in light of that higher resolution.

Just as the 1.8 Å model of molecule one of the dimer was built by readjustment of the 2.8 Å model, the 1.8 Å model of molecule one prime was built by refitting the model of molecule one. To accomplish this, the plexiglass sheets on which were drawn the electron density contours of molecule one prime were rotated by 180° and inserted into the Richards box. With careful alignment, it was thus possible to replace the molecule one density with that of molecule one prime, without rotating the model. It could then be determined by inspection where the two molecules differed in structure and

where the molecule one model would have to be rebuilt to molecule one prime.

The electron density map to which the 1.8 Å model was fit is in general of very high quality. Most of the side chain densities are well defined; only a few of the surface side chains are not visible to the end. It is presumed in these cases that the side chain in question is disordered and adopts more than one conformation; with the result that the electron density observed is an average of these conformations and is therefore too weak to be seen. Alternatively, the side chain could be completely disordered with no preferred orientation. This would have the same effect on the observed electron density. Thus in molecule one, the side chain densities of Lys 87 and Lys 107 were not visible at all, and the side chain of Glu 49 was not seen past the beta carbon. It should be noted, however, that the side chains absent in the molecule one density were present in molecule one prime. Thus, Lys 87 and Lys 107 are visible in their entirety in the molecule one prime density, and Glu 49 is present past the beta carbon. However, the side chain density for Glu 21 and Asp 72 is missing in one prime but present in molecule one. In addition, the side chain of Arg 145 is weaker in molecule one prime than in one. There was very weak density for residues Ser 11-Gly 12-Leu 13 in both molecules, so weak that no model could be built with confidence in this region. This was also observed in the

Cambridge structure and in  $\gamma$ -CHT.<sup>39</sup> Another poor region of electron density in both molecules is Gly 74-Ser 77. Here, there seems to be poor connectivity of the density, with perhaps many solvent molecules present and with close contacts to the neighboring molecule in the crystal. No attempt was made to compare structure of the two molecules in this region. In fact, the molecule one prime model was not refit in this region, so that the two molecules have been built with the same structure here, even though it should be realized that the structure of both molecules in this region is uncertain. In molecule one, there are slight breaks in electron density between Gly 184 and Ala 185, and between Ser 190 and Cys 191. Also in molecule one, the side chain density of Leu 143 is not continuous with the main chain. In both molecules, the side chain density of Met 192 is not continuous with the main chain, and the density is not as well defined as in Met 180, the other methionine residue in the molecule.

## 2. The Difference Density and the Model

When the molecule one prime density with the difference density traced onto it was inspected with the molecule one model in place in the Richards box, it was necessary to rebuild the main chain of approximately 110 of the 238 residues in order to fit the density. An additional 50 residues needed only side chain adjustment or minor main chain refitting. Thus, roughly two thirds of the molecule has differing conformation in the two molecules. The

following is a brief qualitative description of the differences between the two molecules of the CHT dimer. Since molecule one was built first and then one prime was constructed by making changes in the molecule one model, it is natural to discuss differences in the sense of differences in molecule one prime from molecule one. Therefore, unless otherwise noted, in the following discussion, the differences described will be differences of molecule one prime from molecule one. In the same way, all references to changes or readjustments will be to the rebuilding of one to one prime.

### 3. The Two Molecules

The A chain has a slightly different conformation throughout its length in the two molecules. Especially significant is the position of the Pro 8 side chain in molecule one prime where it forms a closer contact with the Trp 27-Pro 28-Trp 29 nest. This was also noted in the 2.8 Å work, where it was observed that N-pipsyI-L-phenylalanyl chloromethyl ketone binds near this aromatic cluster in molecule one but not in one prime, because the closer contact of Pro 8 with surrounding residues blocks binding at this site.<sup>3</sup>

The side chain of Ile 16 has been interpreted differently in the two molecules. In molecule one, the side chain extends parallel to the local two fold axis, whereas in one prime it has been built more perpendicular to that axis.

The density into which the one prime Ile 16 side chain has been built is also present in molecule one, where it was interpreted as possible solvent in the interior of the protein. However, the density into which the Ile 16 side chain of molecule one was built is not present in one prime, so the two side chains were built differently. This region clearly requires more study to determine the exact nature of the structural differences present. The surface residues Val 17-Pro 24 differ in conformation in the two molecules. The deviations between the two are not large, but they are obviously significant from inspection of the electron density. This is typical of all regions of CHT in which there are observed differences in structure. From Pro 24, the main chain goes into the interior of the molecule and emerges at Asp 35. Slight side chain adjustments were required for Asp 35 and Lys 36. A larger difference was observed in the side chain of Phe 39. The Phe 39 side chain in molecule one resides directly on the local two fold axis, so that in molecule one prime, the Phe 39 side chain must be rotated to avoid superposition of the two. In addition, Phe 41 has a significantly different side chain orientation.

The next significant difference in the structure is in the Asn 48-Trp 51 loop. Here the peptide carbonyl group between Glu 49 and Asn 50 was turned through  $180^{\circ}$  so that the carbonyl oxygen points in opposite directions in the two molecules. This readjustment gives this loop the more

normal type I  $\beta$ -bend configuration,<sup>52</sup> whereas in molecule one the loop had been in the type II configuration. Since the type II  $\beta$ -bend is energetically unfavorable when the third residue in the loop is not glycine (in this case, it is Asn 50), it is possible that the main chain was built incorrectly in molecule one; in any event the structures are different in this region, as evidenced by the significant differences in density in the region of Glu 49-Asn 50.

At Trp 51, the main chain enters the interior of the molecule again, and there are no large differences until His 57. Here, the side chain is in a slightly different orientation. Another large difference in structure is in the residues Ala 68-Gly 69-Glu 70. These are partially exposed to solvent and required large adjustment of the main chain conformation. A major main chain difference also occurs in the loop Asn 95-Ser 96-Leu 97-Thr 98. These residues are part of the dimer interface. There were minor readjustments to be made for residues Asn 100-Asn 101, and the carboxylate group of the side chain of the active site Asp 102 was rotated through  $90^\circ$ . This latter adjustment is really a different interpretation of the map, as the electron density of the two molecules is substantially the same at Asp 102.

The entire region from Thr 110 to Val 118 required readjustment. These residues are on the exterior of the molecule. The Cys 1-Cys 122 disulfide bridge was moved slightly, as was Pro 124. The bend Ala 126-Ser 127-Asp

128-Asp 129 required minor adjustment. The main chain of Gly 140-Trp 141 adopts a different configuration in molecule one prime, and this change is concerted with an adjustment of the Leu 155 side chain, which points toward the Trp 141 side chain in molecule one prime. The main chain of the deformed  $\alpha$ -helix Thr 166-Tyr 172 has a slightly altered configuration in the two molecules, and many of the side chains have different orientations, the most dramatic of which is Tyr 172. The phenyl ring of this side chain differs in orientation by  $90^\circ$ . Minor revisions were made to Gly 173-Val 188. The Cys 191-Cys 220 disulfide bridge was moved slightly. The large active site loop of Met 192-Leu 199 has several structural differences, but none that are very large. Thus, even though Gly 196-Pro 198 and Met 192-Gly 193 required substantial readjustment to fit the molecule one prime density, Asp 194 and Ser 195 needed considerably less. The Ser 195 side chain seems to be in a slightly different orientation, but final confirmation of this will have to await further refinement. Another important region of structural difference is the loop Trp 215-Ser 223. This is also part of the dimer interface. Finally, the C-terminal helix shows very small differences in the main chain, but rather larger changes in some of the side chain configurations. Here, the difference electron density clearly indicated differences, but the changes required were very minor. A feature of interest is the ion pair of the Asn 245 C-terminal carboxylate

group with the  $\epsilon$ -amino group of Lys 90. This is the major non-hydrophobic interaction between the C-terminal helix and the rest of the molecule.

#### 4. Solvent Structure

During the building of the two independent molecules of CHT, some density was observed that could not be accounted for by the polypeptide chain. Some of this density has already been shown to be due to sulfate ions.<sup>36</sup> Since protein crystals are approximately 50% solvent by volume,<sup>53</sup> it is customary to identify such "extraneous density" as solvent molecules, especially when the presence of the solvent molecule makes chemical sense, as in hydrogen bonding interactions. Table 2 is a partial list of such possible solvent molecules in molecule one. It can be seen from the 31 entries of Table 2 that these solvent molecules represent a potentially important source of scattering material in the crystal, and any complete description of the CHT crystal structure will have to take this electron density into account. It is usual to include solvent molecules in further stages of protein refinement, and such refinements of other proteins have yielded positions of scores of solvent molecules surrounding the protein.<sup>54,55</sup> Likewise, Table 3 is a partial list of solvent molecules present in molecule one prime but not in molecule one. Even though this is only a partial list, it is apparent that there exists a large degree of asymmetry in the solvation of the two molecules.

Table 2. Partial List of Potential Solvent Molecules.  
Molecule One.

Peak Number	Near Residue(s)	Comments
W1†	Val 3	surface
W2	Ala 5	surface
W3	Ala 5 - Ser 119	possibly more than 1 H <sub>2</sub> O
W4	Gln 7	extended density
W5	Gln 7 - Pro 8	extended density
W6	Ile 16	possibly several H <sub>2</sub> O, interior
W7	Gly 19	in channel
W8†	Gly 25, Glu 70	H-bond connection
W9	Gly 19	close to main chain
W10	Ser 32, His 40	between side chains
W11	Gln 34	near side chain
W12†	Asn 48, Glu 49, Asn 50	in middle of β-bend
W13	Thr 62	surface, H-bond to side chain
W14	Glu 70, Phe 71	
W15	Val 88, Leu 106	between β-strands
W16	Thr 117, Glu 70	side chains
W17	Ser 113	side chain
W18	Val 118	main chain
W19	Gln 156, Arg 154	
W20†	Gly 133, Pro 161	
W21	Ala 131	carbonyl oxygen, surface
W22†	Trp 172, Ser 223	H-bond between side chains
W23†	Arg 230, Asp 178	
W24	Pro 161, Ser 186	
W25	Trp 141, Asp 194	possibly more than 1 H <sub>2</sub> O
W26†	Thr 138, Pro 198	
W27	Ser 221	side chain H-bond
W28	Ser 190, Gly 216, Gly 226	specificity pocket
W29	Ser 190, Tyr 228	0 atoms of side chains
W30	Lys 199, Ala 202	across a bend

†These solvent peaks are not present in molecule one prime.

Table 3. Partial List of Potential Solvent Present in Molecule One Prime But Not in Molecule One.

<u>Peak Number</u>	<u>Near Residue(s)</u>	<u>Comments</u>
WP1	Val 9, Trp 27	
WP2	Cys 42	main chain
WP3	Gly 140	main chain
WP4	Tyr 171	on OEH of side chain
WP5	Arg 230	side chain
WP6	Ser 96, Leu 97	side chain
WP7	Asn 91, Thr 103	side chain
WP8	Leu 242	side chain
WP9	Asn 48	near 2 fold axis
WP10	Thr 134	surface, side chain
WP11	Trp 215	main chain

Since enzyme molecules function naturally in aqueous solution, it is reasonable to assume that water molecules and other ions in solution play an important role in the structure and action of these proteins. Thus, variability in solvent structure is another manifestation of the lack of exact equivalence between the two molecules.

The solvent molecules are found mainly on the exterior of the enzyme. In Table 2, only W6, W10, W26, W27, and W30 appear in the interior. In Table 3, all the peaks listed are on the exterior of the molecule.

The solvent structure seems to be very sensitive to the structure of the individual subunits of the CHT dimer. A clear example of this is W22 (Table 2), which is present in molecule one but not in one prime. It forms a hydrogen bond with NE1 of Trp 172 and with OG1 of Thr 224. In molecule one prime, these two residues are in closer contact, and the water molecule does not bind because there is no space for it. Thus, a minor structural difference between subunits can have a dramatic influence on solvent structure. Further refinement of CHT should reveal even more solvent molecules and more examples of such asymmetry in solvent structure.

##### 5. Model Conformation

The Ramachandran plots<sup>56</sup> for molecule one and molecule one prime are shown in Figures 5 and 6, respectively. Each point in the plots represents the conformation of an amino

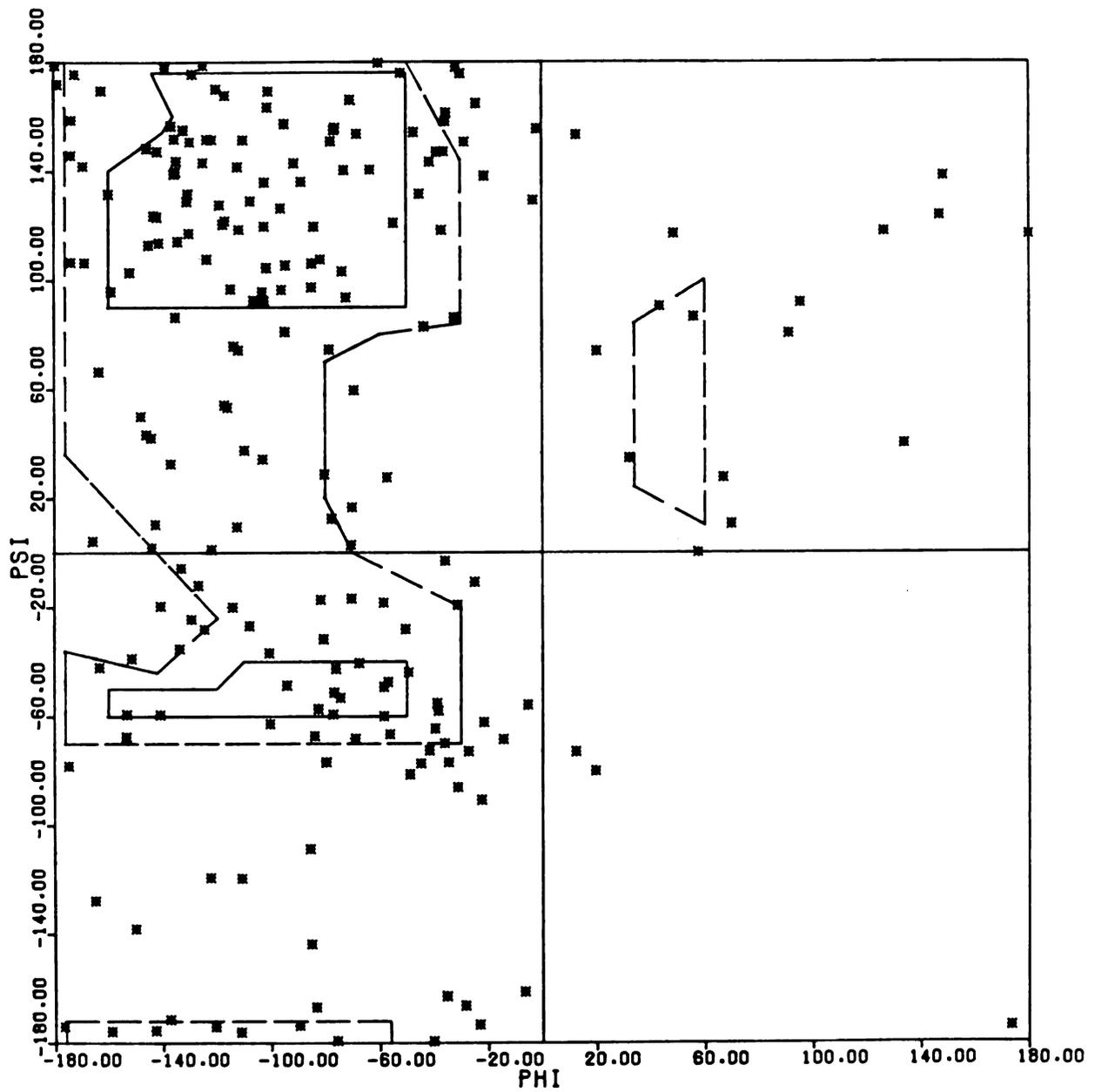


Figure 5. Ramachandran plot for molecule one.

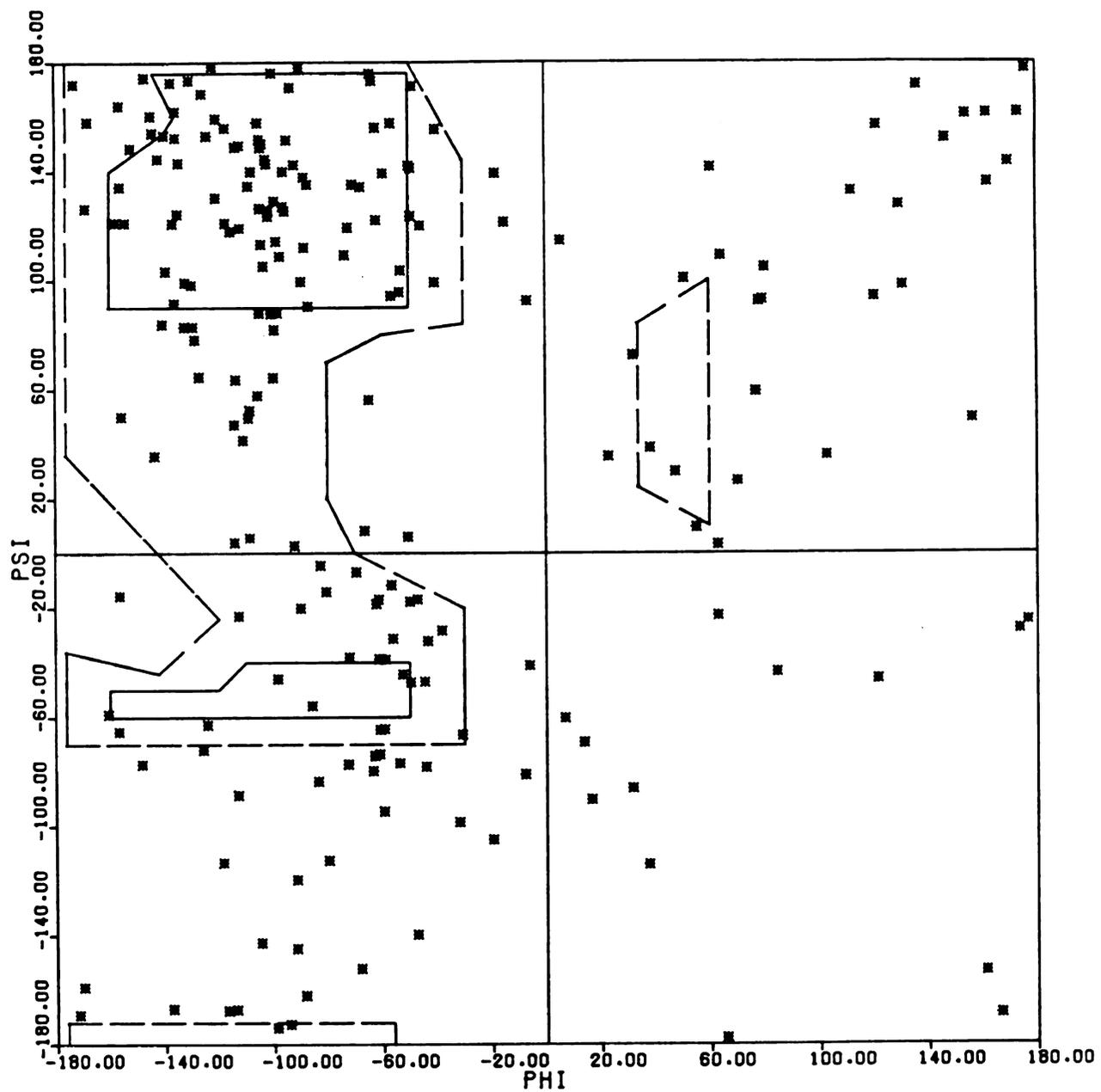


Figure 6. Ramachandran plot for molecule one prime.

acid residue. The torsion angles  $\phi$  and  $\psi$  are defined in Figure 7. The majority of possible conformations are not stereochemically reasonable, as they represent configurations in which atoms of the molecule make non-bonded approaches which are too close. The closest allowed non-bonded distance between two atoms is the sum of the van der Waals radii of the atoms. The dashed lines of Figures 5 and 6 enclose the allowed regions. Glycines, because they have no side chain, can adopt a wider range of allowed conformation. Many Gly residues, therefore, fall outside the allowed regions of Figures 5 and 6, and have been omitted. It can be seen that in both molecules, approximately 10% of the residues have been built in a formally disallowed conformation. This appears to be typical of unrefined structures.<sup>48,57</sup>

In molecule one, it was noted that the C chain had a smaller percentage of residues with disallowed conformation than did the B chain. This improvement was ascribed to experience of the model builder, the C chain having been built after the B chain. It is therefore surprising that molecule one prime does not show an "improvement" over molecule one. This is not to say, however, that portions of the molecule were built with grossly unreasonable stereochemistry. Most deviations of the model from a formally allowed conformation result in non-bonded contacts which are only a few tenths of an Ångstrom less than the van der Waals distance. Apparently, even for an experienced builder, a perfectly allowed model is difficult to build

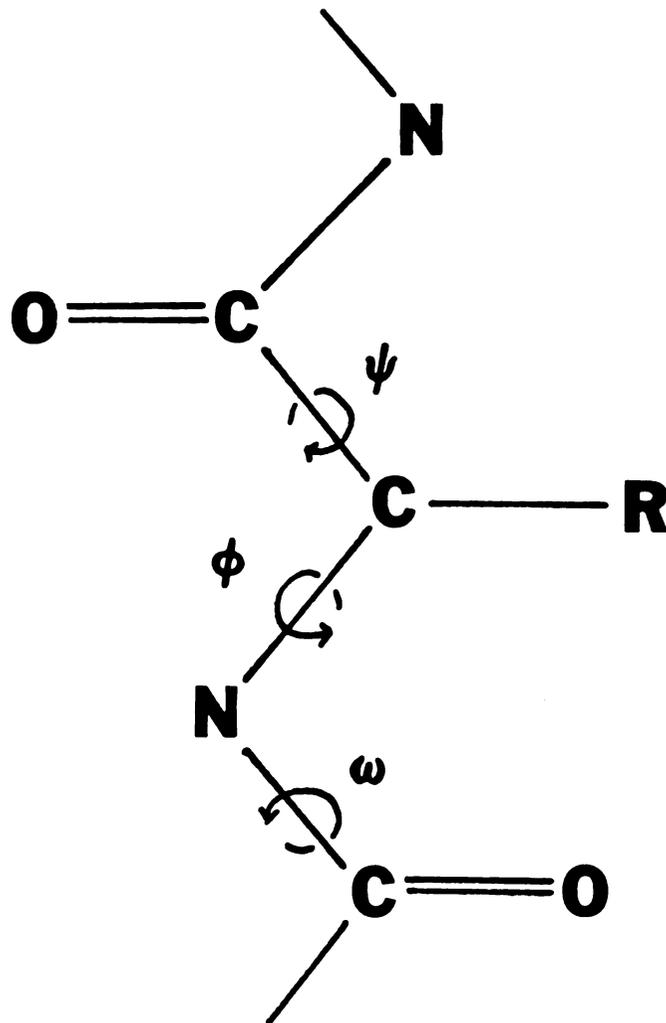


Figure 7. Main chain torsion angles of a protein.

with only indirect attention paid to the absolute values of the conformational angles of the model being constructed. The residues which lie in the formally disallowed region of the Ramachandran plot will be given special attention in further refinement, so that they might adopt a more energetically favored conformation in the final, refined structure.

## CHAPTER 4

### Comparison of the Structures of the Independent Molecules of $\alpha$ -Chymotrypsin Dimer.

#### 1. Determination of the symmetry axis

##### a. From the Electron Density Map

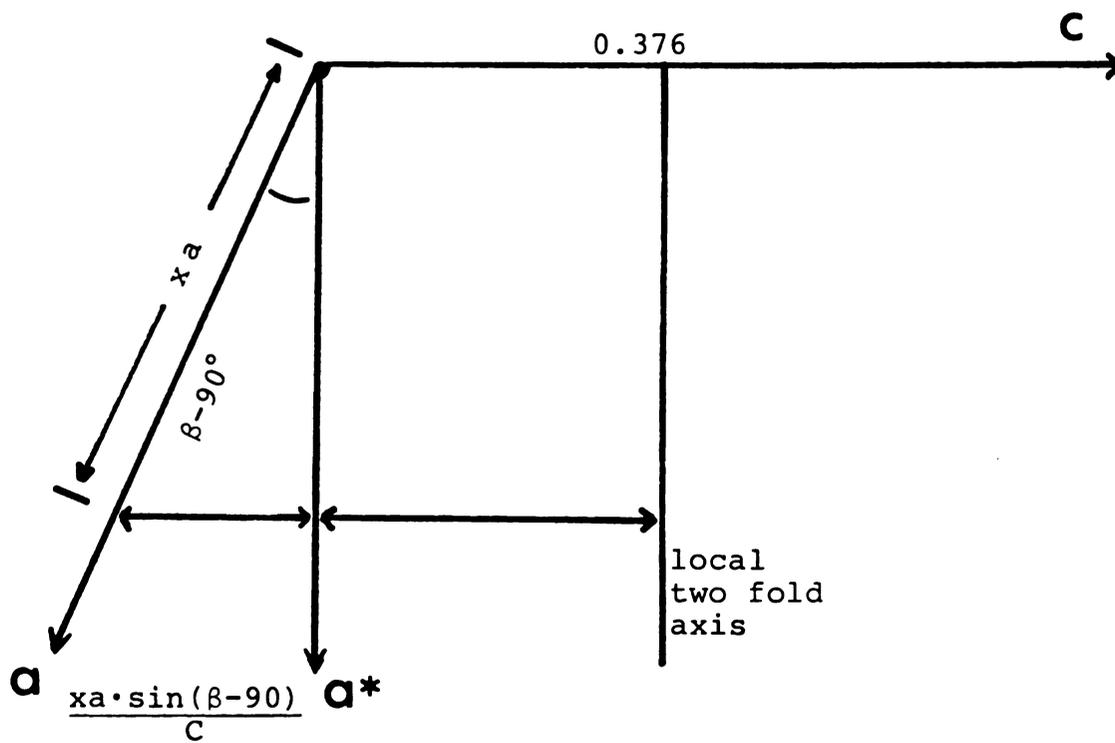
In Chapter 3, differences in structure of the two molecules of the CHT dimer were discussed in a qualitative way. The observed differences were deduced from inspection of the electron density map and the difference electron density map between the two molecules. The latter was calculated as  $\Delta\rho(\vec{r}_1) = \rho(\vec{r}_1) - \rho(\vec{r}_1')$ , where  $\vec{r}_1$  ranges over the entire volume of one molecule,  $\vec{r}_1'$  is the position in the electron density map which is related to  $\vec{r}_1$  by the non-crystallographic symmetry, and  $\rho$  is the electron density. Thus,  $\vec{r}_1'$  refers to the second molecule of the dimer. The difference density calculation is straightforward once the proper transformation is known. The transformation is that which when applied will, in effect, lead to the smallest root mean square value of  $\Delta\rho$  over the entire molecule. It is found by a least squares minimization of  $(\Delta\rho)^2$  by adjustment of the

transformation parameters. The transformation parameters consist of three linearly independent rotation angles, from the values of which a rotation matrix can be calculated, and three translations, which are the components of a translation vector in an orthogonal coordinate system. From the transformation parameters, the equation of the rotation axis of symmetry and the amount of rotation about that axis can be calculated.

The equation of the line is calculated in the following way. The values of the transformation parameters used in the calculation of the difference map between the molecules were:  $\psi = 0^\circ$ ,  $\theta = 180^\circ$ ,  $\phi = 0^\circ$ ,  $d_1 = 0.0$ ,  $d_2 = 39.86 \text{ \AA}$ ,  $d_3 = 49.56 \text{ \AA}$ . The angles  $\psi$ ,  $\theta$ , and  $\phi$  are the three Eulerian angles defined as:  $\psi$  is a rotation about  $z$ ,  $\theta$  is a rotation about the new  $x$ , and  $\phi$  is a rotation about the new  $z$ . The translations  $d_1$ ,  $d_2$ , and  $d_3$  are in the  $x$ ,  $y$  and  $z$  directions respectively. In this case, the  $x$  direction is along  $\vec{a}^*$  of the crystal. Thus the Eulerian angles specify a  $180^\circ$  rotation about the  $\vec{a}^*$  ( $= x$ ) direction, while  $d_2$  and  $d_3$  are related to the offsets in the  $y$  and  $z$  directions of the position of the two fold axis as follows. After rotation about  $\vec{a}^*$  at  $z = 0$ ,  $y = 0$ , translations of  $d_2$  and  $d_3$  are required to complete the symmetry operation. This is equivalent to a pure rotation about the point  $(d_2/2, d_3/2)$  with no translation. Based on cell dimensions of  $a = 49.24 \text{ \AA}$ ,  $b = 66.91 \text{ \AA}$ ,  $c = 65.83 \text{ \AA}$  and  $\beta = 101.8^\circ$ , the value of  $d_2/2$

in fractional coordinates is  $39.86/(2*b) = 0.298$ , and that of  $d_3/2$  is  $49.56/(2*c) = 0.376$ . Moreover, since the axis of rotation is not parallel to the x direction in the crystal, but is parallel to  $\vec{a}^*$ , the z position of the two fold axis varies with respect to x, shifting  $0.153x$  in each section of x (see Figure 8). The parametric equation for the two fold axis along  $\vec{a}^*$  is therefore  $z = 0.153*x + 0.376$ ,  $y = 0.298$ . Note that  $0.153 = \frac{a(\sin(\beta-90^\circ))}{c}$ , independent of  $d_2$  and  $d_3$ . The slightly different value of this equation reported by Raghavan and Tulinsky<sup>5</sup> is due to their use of slightly different values of a, b, and c from those reported here.

From the values of the transformation parameters thus obtained, a difference map between the two molecules was calculated and traced onto the plexiglass sheet which also had the electron density of molecule one prime. These sheets were rotated by  $180^\circ$  and inserted into the Richards box, so that rotation of the model was not necessary. However, because the two fold axis did not pass through the center of the plexiglass sheet as drawn, it was necessary to raise the map four inches after rotation to bring the two fold axis of the one prime electron density into coincidence with the two fold axis on the model of molecule one. This was accomplished by unbolting the model frame from the map box, jacking up the box and placing it on four inch stacked plywood blocks, taking great care to raise the box vertically and leave the model precisely in place. In this



$$z = 0.153x + 0.376$$

Figure 8. Equation of the two fold axis.

way it was possible to replace the molecule one density with the molecule one prime electron density onto which was superimposed the difference density between the two molecules. The difference density was drawn in red to distinguish it from the black electron density contours, and drawn beginning at  $\pm 0.8 \text{ e}/\text{\AA}^3$ , with contours at each additional  $0.4 \text{ e}/\text{\AA}^3$  increment. The standard error of the multiple isomorphous replacement electron density map was reported to be  $0.18 \text{ e}/\text{\AA}^3$  (reference 4). From this, the expected error in the difference density is  $\sqrt{2} * (0.18)$ , or  $0.25 \text{ e}/\text{\AA}^3$ . Thus,  $0.8 \text{ e}/\text{\AA}^3$  is approximately  $3\sigma(\Delta\rho)$  and is an appropriate level to begin contouring.

The electron density and the difference density were consistent throughout. Portions of the molecule one prime density that were not present in molecule one were indicated by negative contours, while positive contours in the difference map represented density that had been present in molecule one but not in one prime. Thus, negative contours were always found superimposed on electron density contours, while positive difference contours were observed on weak or zero density regions of molecule one prime. That these were so consistent with the model is an indication that the two fold rotation axis which best relates the two electron density regions also applies to the atomic coordinates. It is also an indication that the model was built into the density well, with features significant in the density being retained in the model.

b. From the Coordinates

The coordinates of the two models built by reference to these maps were measured as described in Chapter 2. From these coordinates a transformation between the molecules can be calculated. As noted above, the consistency of the difference map with the electron density map and the model suggested a very good match of the electron density and the model. In such an event, the transformation which relates the coordinates of the two molecules should be the same transformation that relates the electron density of the two molecules. However, here one must consider the possibility that, while the model matches the map very well, the coordinates are not an error-free representation of the model built to the map. To the extent that errors do exist in the coordinates, especially if the errors are systematic, the coordinates will not exactly match the map, and the transformation between coordinates will be different from the transformation which best relates the electron density.

In the measurement of coordinates described in Chapter 2, a systematic error is introduced in the measurement of the unit cell origin in the transit frame. The tetrahedral model part which represents the origin marker (Figure 2) is manually placed on the model and then the unit cell fractional coordinates of this origin marker are measured from the apparent position of the reflection of the marker superimposed on the electron density. The z and y unit cell

coordinates of the marker are determined from the measurement of the distance of the origin marker's reflected superposition on the electron density from the edge and from the top, respectively, of the plexiglass sheet which is at the same depth in the map as the origin marker is on the model. The x unit cell coordinate is obtained from the x coordinate of the plexiglass sheet (the sheets are stacked in sections perpendicular to x). Thus the origin marker must be placed precisely at the depth of a given plexiglass sheet, and this introduces an error of as much as  $\pm 0.4$  cm because of lack of perfect depth perception. The y and z coordinate measurement of the origin marker is affected by parallax, but this is not as great an uncertainty as the depth error. The uncertainty in depth perception is added to by the bowing of the plexiglass sheets in the box, so that the x coordinate of the middle of the sheet is slightly different from that of the top and bottom. With care, the error in the measurement of fractional coordinates of the origin marker can be minimized, but the imprecision with which the origin marker position is known is a significant source of error in measured coordinates of the molecules, and thus in the match or consistency of the measured model with the map. The origin marker in molecule one prime was placed independently of that of molecule one. That is, even though the model was not moved but rather the density maps were rotated to fit the model, it was considered desirable

to reposition the origin marker. New fractional coordinates were determined for the new position, and these coordinates have the same kind of systematic error as described above. Because the two origin markers were placed and measured independently, the systematic error is not necessarily cancelled when coordinates of molecule one are compared with those of molecule one prime. Thus, the transformation between the two sets of coordinates will be affected by this systematic error in coordinate measurement.

During the rebuilding of the molecule one model to molecule one prime, approximately 68 amino acids were not substantially altered in main chain conformation, even though the side chains of many of these 68 did require adjustment. Because these 68 residues fit the electron density of molecule one prime even before the adjustment of the molecule one model, they are taken as exactly equivalent. Therefore, any transformation between the two molecules should result in a fit such that the difference between these residues in molecule one and the transformed molecule one prime would be minimal. To calculate this transformation, the alpha carbon atoms of these 68 residues were used as input to a program by Rossmann and Argos<sup>58</sup> which determines by iterative least squares refinement the parameters which best transform the coordinates of one molecule to another and hence of molecule one prime to molecule one. The function minimized is  $D = \sum (\vec{x}_1 - \vec{x}'_{1p})^2$  where  $\vec{x}_1$

are the coordinates of molecule one,  $\vec{x}'_{1p}$  are the coordinates of molecule one prime which have been transformed by the three Eulerian angles and three translation vector components, and the sum is over all atoms in the structure to be compared, in this case over all the alpha carbons of the 68 equivalent residues. The residues used were: Trp 27, Pro 28, Val 31-Asp 35, Cys 42-Ile 47, Trp 51-Ala 56, Asn 100-Leu 108, Cys 136-Thr 139, Asn 150-Leu 162, Leu 199-Lys 202, Thr 208-Ile 212, Pro 225-Val 227, and the C terminal  $\alpha$ -helix Leu 234-Ala 244. The r.m.s. deviation of the coordinates of the 68 transformed molecule one prime alpha carbons from the molecule one coordinates is 0.41 Å. The transformation parameters that produce such a fit are  $\psi = 87.6^\circ$ ,  $\theta = 155.9^\circ$ ,  $\phi = 87.4^\circ$ ,  $d_1 = -10.79$  Å,  $d_2 = 38.83$  Å,  $d_3 = 48.37$  Å where the terms are as defined earlier in this chapter. It can be seen that these transformation parameters do not match the parameters previously determined from the electron density map. This is because the two are based on different axial coordinate systems.

Thus, the electron density map derived transformation parameters operate on a coordinate system based on the  $\vec{a}^*$ ,  $\vec{b}$ , and  $\vec{c}$  directions in the crystal, while the parameters obtained from the alpha carbon coordinate comparison are based on the directions  $\vec{a}$ ,  $\vec{b}$ , and  $\vec{c}^*$ . The choice for the electron density map was natural because the map was calculated and drawn in sections perpendicular to  $\vec{a}^*$ .

## c. Comparison of the Two Transformations

The two transformations can be compared in the following way. Any two sets of coordinates can be related to each other by a rotation matrix and a translation vector. In the present case, let  $m_2$  be the coordinates of molecule one prime and  $m_1$  be the coordinates of molecule one. They are approximately related to each other by

$$m_1 = Rm_2 + t \quad (4.1)$$

For the case when  $m_1$  and  $m_2$  are not exactly equal in structure,  $R$  and  $t$  are the matrix and vector, respectively, that relate the two in a least squares sense;  $R$  and  $t$  are found such that  $D = \sum [m_1 - (Rm_2 + t)]^2$  is a minimum, where the sum is over all atoms in  $m_1$ . We wish to find values of  $T$  and  $d$  from  $R$  and  $t$ , where  $T$  and  $d$  give the transformation about the axial system based on  $\vec{a}^*$ ,  $\vec{b}$ , and  $\vec{c}$ , so that we might compare them with values of these parameters obtained from the electron density map. In order to do this, we separate the general transformation (equation 4.1) into three steps: 1) convert the coordinates of molecule one prime to coordinates based on  $\vec{a}^*$ ,  $\vec{b}$ , and  $\vec{c}$ ; 2) apply the transformation of  $T$  and  $d$ ; and 3) convert the coordinates back to the system based on  $\vec{a}$ ,  $\vec{b}$ , and  $\vec{c}^*$ . This can be expressed as

$$m_1 = B(T(B^{-1}m_2) + d) \quad (4.2)$$

where  $B^{-1}$  is the matrix which rotates the coordinate system from  $\vec{a}$ ,  $\vec{b}$ ,  $\vec{c}^*$  to  $\vec{a}^*$ ,  $\vec{b}$ ,  $\vec{c}$ , and  $B$  is the inverse of  $B^{-1}$  and

converts the coordinate system from  $\vec{a}^*$ ,  $\vec{b}$ ,  $\vec{c}$  to  $\vec{a}$ ,  $\vec{b}$ ,  $\vec{c}^*$ .

Equating (4.1) with (4.2), we have

$$Rm_2 + t = BTB^{-1}m_2 + Bd \quad (4.3)$$

Equating terms in (4.3) yields

$$R = BTB^{-1}; \quad t = Bd \quad (4.4)$$

and, solving for T and d, we obtain

$$T = B^{-1}RB; \quad d = B^{-1}t \quad (4.5)$$

The matrix B is known, so that T and d can be calculated from (4.5). B is a rotation about y of  $\beta - 90^\circ$ , where  $\beta = 101.77^\circ$  is a unit cell parameter. It is therefore given by

$$B = \begin{pmatrix} \cos 11.77 & 0 & -\sin 11.77 & .979 & 0 & -.204 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ \sin 11.77 & 0 & \cos 11.77 & .204 & 0 & .979 \end{pmatrix} \quad (4.6)$$

The inverse matrix  $B^{-1}$  is found by substituting  $-11.77^\circ$  for  $11.77^\circ$  in equation (4.6). The inverse in this case is simply the transpose. R and t are known from the transformation parameters found by comparison of the coordinates. R is given from the Eulerian angles  $\psi$ ,  $\theta$ , and  $\phi$  by

$$\begin{array}{ccc}
 \cos\psi\cos\phi & \sin\psi\cos\phi & \sin\psi\sin\theta \\
 -\sin\psi\sin\phi\cos\theta & +\cos\psi\sin\phi\cos\theta & \\
 \\
 -\cos\psi\sin\phi & -\sin\psi\sin\phi & \cos\phi\sin\theta \\
 -\sin\psi\cos\phi\cos\theta & +\cos\psi\cos\phi\cos\theta & \\
 \\
 \sin\psi\sin\theta & -\cos\psi\sin\theta & \cos\theta
 \end{array}
 \tag{4.7}$$

$$\begin{array}{ccc}
 .9134 & .0061 & .4071 \\
 = & -.0014 & -.9998 & .0182 \\
 .4071 & .0172 & -.9132
 \end{array}$$

Now T and d can be calculated:

$$T = B^{-1}RB = \begin{array}{ccc} 1.000 & .002 & .008 \\ .002 & -1.000 & .020 \\ .008 & - .020 & -1.000 \end{array}
 \tag{4.8}$$

$$d = B^{-1}t = \begin{array}{ccc} -0.7 & & d_1 \\ 38.83 & = & d_2 \\ 49.56 & & d_3 \end{array}$$

Thus T represents, within the precision obtainable, a  $180^\circ$  rotation around x, which is just as inferred from the electron density map. The equation of the two fold axis calculated from these parameters is  $z = 0.153x + 0.376$ ,  $y = 0.290$ . The parameters  $d_1$  and  $d_2$  differ by  $-0.7 \text{ \AA}$  and  $-1.03 \text{ \AA}$ , respectively, from those obtained from the electron density map. The non-zero value of  $d_1$  means that the symmetry axis has a slight screw component:  $x' = x + d_1/a = x - .014$ . This represents a translation in the  $a^*$  direction of  $0.7 \text{ \AA}$ . This is approximately the distance between the plexiglass sheets in Richards box. Such an error in x could be caused

by uncertainty of half a section in the depth placement of the origin marker in molecule one, and by a half section misplacement in the opposite sense in molecule one prime. As discussed above, this amount of uncertainty is expected in the x direction. The discrepancy between the parameter  $d_2$  determined from the electron density map and from the coordinates is somewhat larger and is perhaps more than the expected uncertainty in measuring the position of the origin marker. However, because the positioning of the origin marker in both molecules was subject to some error and because the measurement was made from only one observation in each case, it can not be ruled out that the observed difference between the two values of  $d_2$  might be due to random error. It should also be noted that the y coordinate for the two fold axis reported in the 2.8 Å resolution work<sup>4</sup> was  $y = 0.291$  as compared to  $y = 0.290$  found here. The standard error reported for y in the 2.8 Å resolution work was  $\pm 0.002$ , so the difference here is not significant. That  $d_3$  determined by the two methods is equal to four significant figures is fortuitous, but it is not surprising that it is in the z-coordinate that the least discrepancy between model coordinates and map is found. The z position of the origin marker was the easiest to determine precisely, as it entailed the measurement of the origin marker from the edge of the plexiglass sheet and involved no problem with depth perception. The y coordinate was measured from a point less

precisely determined than an edge, and, as was discussed above, the x-coordinate measurement was uncertain by approximately the distance between sheets. In view of all of the above, it can be concluded that the transformations calculated from the electron density and from the coordinates of the two models built to the density do not differ significantly. In any event, the discrepancy is not larger than that found between transformations determined from electron density maps at two different resolutions.

The transformation was also calculated for groups of atoms other than the 68 residues mentioned above. The groups used were: 1) all residues; 2) all residues of domain 1 (residues 1-122); 3) all residues of domain 2 (residues 123-245); and 4) all residues not in the original 68. The results of these comparisons are given in Table 4. Here, level 1 means that only alpha carbons were used in the comparison. Level 3 indicates that the alpha carbon, carbonyl carbon, and nitrogen atoms of the main chain of each residue were used. "Interior" designates the 68 equivalent residues, while "exterior" refers to all the rest. It can be seen from Table 4 that the inclusion of the carbonyl carbon and nitrogen atoms had minimal effect on the values of the transformation parameters calculated. The parameters calculated in the various groups of Table 4 are also seen to be roughly equivalent, especially when the uncertainty in the measurements is considered. The transformation of the 68 interior

Table 4. Transformation Parameters, Molecule One Prime to Molecule One.

Group of Atoms	Level <sup>†</sup>	$\psi$	$\phi$	$d_1$	$d_2$	$d_3$	r.m.s. <sup>†</sup>	N <sup>*</sup>
All	1	89.9	89.5	-11.1	39.3	47.4	1.09	238
All	3	89.8	89.5	-11.1	39.3	47.4	1.05	714
Domain 1	1	89.8	89.3	-11.0	38.9	48.0	1.13	117
Domain 1	3	89.7	89.0	-11.0	38.9	47.9	1.07	351
Domain 2	1	90.3	90.3	-10.9	39.6	47.3	1.04	121
Domain 2	3	90.2	90.2	-10.9	39.7	47.3	1.03	363
Exterior	1	90.9	90.5	-11.1	39.5	47.1	1.26	170
Exterior	3	90.9	90.4	-11.1	39.5	47.1	1.21	510
Interior	1	87.6	87.4	-10.8	38.8	48.4	.41	68
Interior	3	87.5	87.4	-10.8	38.8	48.3	.43	204

<sup>†</sup> Level 1 - only alpha carbons used. Level 3 - alpha carbons, nitrogen, and carbonyl carbons used in comparison.

<sup>†</sup> Root mean square deviation of all alpha carbons in the group.

\* Number of atoms in the group.

atoms, level 1, was chosen as the "best" because it was known that these 68 residues have nearly the same conformation in both molecules. This fact is evident from the r.m.s. values in Table 4, which are the root mean square deviations of all alpha carbon atoms in the group. The value for interior atoms, level 1, is less than half as large as any others. Here, level 1 was chosen over level 3 on the basis of its (insignificantly) lower overall r.m.s. value.

## 2. The two molecules of $\alpha$ -CHT

For comparison of the two molecules, molecule one prime was transformed by the "best" parameters of Table 4: interior, level 1. Then the two molecules were examined atom by atom and the distance between equivalent atoms was calculated for each of the 1764 non-hydrogen atoms of CHT. A summary of the detailed comparison is given in Table 5. It is of interest to compare certain classes of residues. To this end the residues of CHT have been divided into three groups: 1) the residues in the dimer interface; 2) those residues in the interior of the molecule, not in contact with the solvent; and 3) the residues that are on the exterior and exposed to solvent in the dimer. These groups have no members in common, and every residue belongs to one of the three groups. The dimer interface region exhibits some very large deviations from exact two fold symmetry. This is expected because in the dimerization process, as the molecules approach, the

Table 5. Comparison of Molecule One and Molecule One Prime.

<u>Group of Atoms</u>	<u>r.m.s. Deviation, A</u>	<u>Number of Atoms</u>
All Atoms	1.66	1764
Main Chain*	1.08	723
Side Chain	2.06	797
Alpha Carbons	1.13	241
Carbonyl Oxygens	1.66	244
His 57, Asp 102, Ser 195	0.99	24
Dimer Interface		
All Atoms	1.96	288
Main Chain*	1.37	117
Side Chain	2.42	132
Interior Residues		
All Atoms	1.25	530
Main Chain*	0.80	213
Side Chain	1.48	246
Exterior Residues		
All Atoms	1.71	808
Main Chain*	1.05	333
Side Chain	2.13	364

\* Calculation for main chain atoms does not include carbonyl oxygens.

dimerization is accompanied by generally asymmetric rearrangement of the entire interface, especially where the contact about the two fold axis is very close. The three regions of closest contact are Phe 39, Met 192, and Gly 216-Thr 219. The r.m.s. deviation of the Phe 39 side chain is  $6.5 \text{ \AA}$ . The phenyl ring of the Phe 39 in molecule one lies directly on the two fold axis, while in molecule one prime, the ring is rotated to avoid contact. The region Gly 216-Thr 219 is shown in Figure 9, where the two fold axis is appropriately marked, and the view is parallel to that axis. It can be seen that the close approach of molecule one to the two fold axis in this region precludes the other molecule's occupying the exact two fold related position. The r.m.s. deviation of the main chain from exact two fold symmetry in this region is  $1.65 \text{ \AA}$ .

As expected and as observed in the  $2.8 \text{ \AA}$  work, the exterior, solvent exposed residues have a larger variability in structure than do the interior residues. This is especially true in the side chains. An extreme example of this is the Arg 145 side chain. The r.m.s. deviation of this side chain in the two molecules is  $5.8 \text{ \AA}$ , while that of the main chain of this residue is only  $0.72 \text{ \AA}$ . Clearly the surface residues have fewer restraints on their orientation than do the interior side chains. However, the interior side chains are observed to show variability, although less than the exterior. The main chain in all three groups has considerably

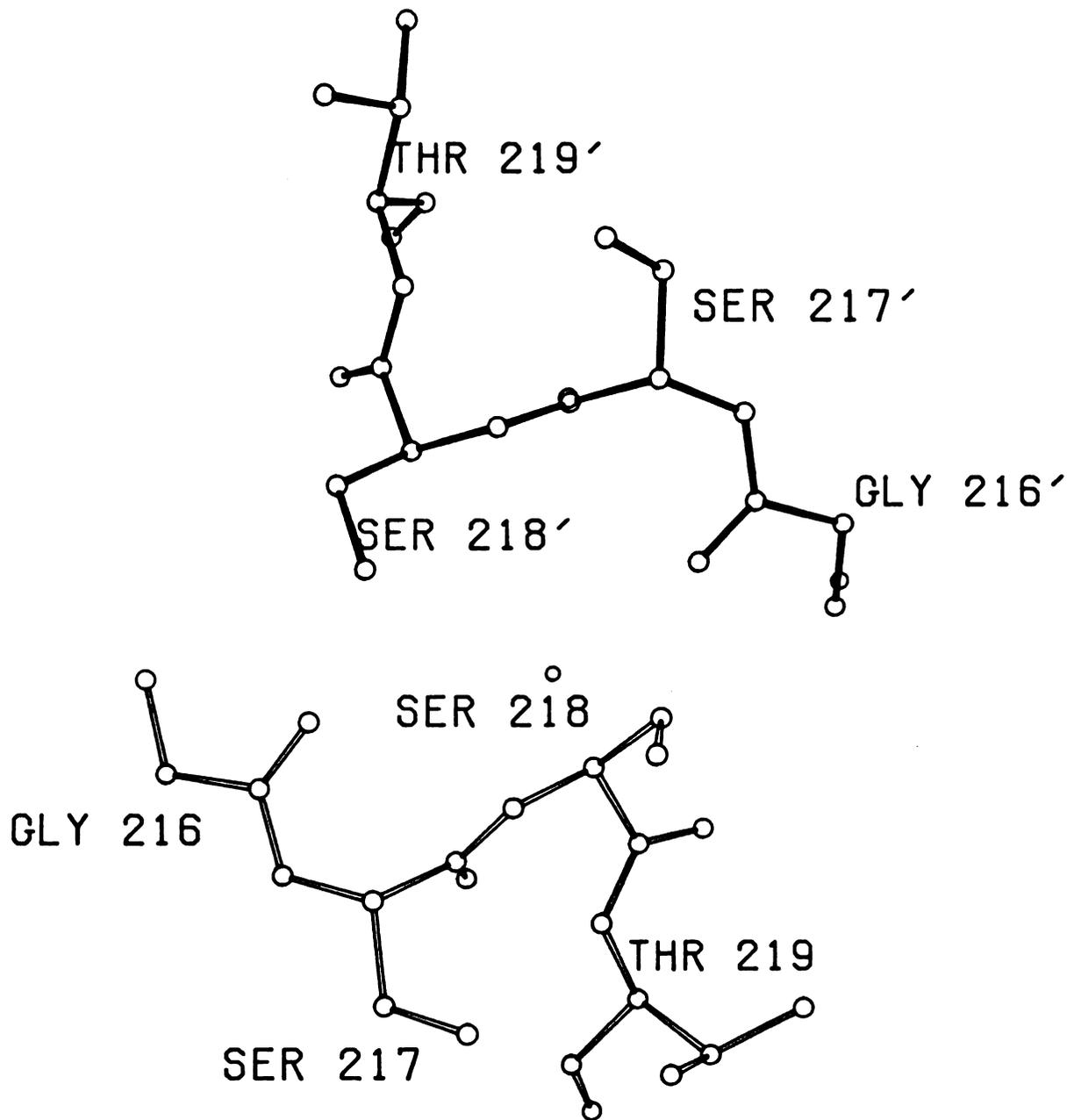


Figure 9. The region Gly 216-Thr 219, viewed down the local two fold axis.

less variation in structure than do the side chains.

But while there are regions of large differences in structure, there are also certain portions of the dimer molecule which do not deviate greatly from two fold symmetry. For instance, the entire main chain, with a few exceptions, has a very similar conformation in the two molecules (r.m.s. = 1.1 Å). The largest deviation in main chain positions is at Gly 69 (3.4 Å). This part of the chain is on the surface exposed to solvent. Other regions with large deviations from two fold equivalence of the main chain are: 1) the loop Asn 95-Ile 99 (at the dimer interface); 2) Ala 112-Asp 129; 3) Gly 140-Trp 141; 4) Tyr 171-Trp 172; 5) Met 192-Gly 193 and Gly 196-Gly 197; and 6) Gly 216-Ser 223. Groups 1, 6, and part of group 5 are in the dimer interface. Group 3 and a portion of 2 and 5 are located in the interior, while group 4 and part of group 2 are on the exterior of the dimer molecule. The r.m.s. deviation of each of these regions is greater than 1.7 Å, the value of r.m.s. difference in structure of the complete molecule including side chains. The differences in main chain conformation are depicted graphically in Figure 10.

A feature of some interest in Table 5 is the large r.m.s. deviation of the carbonyl oxygen atoms relative to the rest of the main chain. This is because the oxygen position is sensitive to minor adjustment of main chain conformation, while the overall position of the main chain is less affected.

ALPHA CARBON DEVIATION  
MOLECULE ONE - ONE PRIME

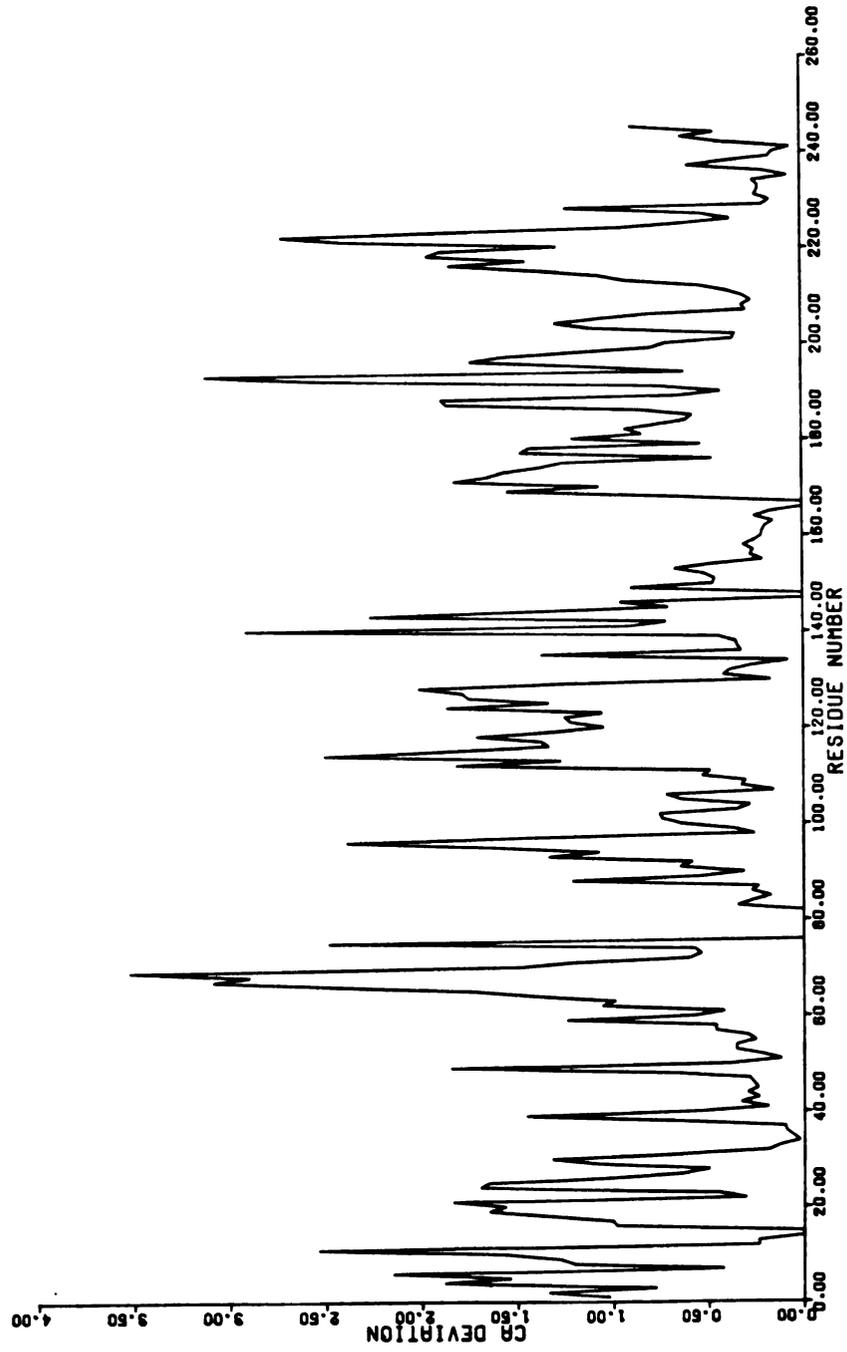


Figure 10. Comparison of alpha carbon positions in the two molecules of  $\alpha$ -chymotrypsin.

The side chains are like the carbonyl oxygens in that they protrude from the main chain, and thus their position is also very sensitive to main chain conformation.

### 3. Variability in $\alpha$ -CHT structure

From the reverse point of view, this argument suggests a model for protein variability that is consistent with the observations in CHT. The structure of a protein is uniquely determined by its amino acid sequence. The side chains of the various amino acid residues provide functional groups for stability and to bind and carry out chemical reactions on substrate molecules. Almost all of the space in and around the enzyme molecule is occupied by side chains. Whatever stabilization of structure the main chain might contribute through formation of intrachain hydrogen bonding in  $\alpha$ -helix,  $\beta$ -sheet, and  $\beta$ -bends, it is clear that a major role of the main chain is to place precisely the side chains so that the molecule might remain stable and react with an appropriate substrate. It is therefore natural that protein structural variability should be expressed in the side chain structures. Small changes in side chain configuration can be accomplished with very little perturbation of the main chain configuration. If greater side chain adjustment is required, then the main chain can also change conformation.

In CHT, the perturbations which cause structural variability are: 1) the dimer interface; 2) natural surface variability; and 3) neighboring molecules in the unit cell

(crystal packing forces). There are no clear indications of 3) in the present work, although it can not be ruled out that crystal packing plays some role in the variability of the CHT structure. The role of packing is probably small because the molecules in the CHT crystal, as in all protein crystals, do not make extensive close contacts, although it was noted in Chapter 3 that the region Gly 74-Ser 77 was involved in close intermolecular contact. Protein crystals are approximately 50% solvent by volume,<sup>53</sup> so most of the surface of the molecule in the crystal is exposed to solvent and not to a neighboring molecule. By natural surface variability, we mean the freedom of side chains which extend into solvent to adopt different configurations by rotations about single bonds. This probably occurs on the interior also, but to a lesser extent since in the interior the side chains are involved in more intimate interaction with other parts of the molecule. Variability in side chain structure could not occur without disruption of those interactions, which are obviously important for the stability of the enzyme structure. It is when the dimer forms that the structure of CHT is perturbed most. Some examples of large deviations in the dimer interface were given above, and it was seen in Table 5 that the dimer interface region had the largest r.m.s. deviation from exact symmetry in the molecule.

The variability in structure is confined mostly to the surface, but the interior is also affected by surface

variability. An example is the interaction between Thr 224 and Trp 172, which was mentioned in Chapter 3. The conformation of Thr 224 is perturbed by the dimer interface. The resulting change in structure affects the position of the Trp 172 side chain and the water molecule which lies between the two side chains in molecule one. As was noted earlier, this water molecule is absent in molecule one prime. Thus, variability at the dimer interface is transmitted to the interior (Trp 172). Another example is the interaction of Leu 155 and Trp 141 in molecule one prime shown in Figure 11. Some surface effect causes the Leu 155 side chain to adopt a different configuration in molecule one prime such that it points into the interior toward the Trp 141 side chain. The close contact is relieved by a reorganization of the structure of Gly 140 and Trp 141 in the interior of the molecule. Thus the interior structure seems not to have been "insulated" from surface perturbations, but rather to have responded by a network of minor inter-related side chain adjustments.

It is also of interest and significance that certain structural features are retained to a high degree of fidelity in the two molecules. For example, the five disulfide bridges match in structure as a group to within an r.m.s. deviation of  $0.8 \text{ \AA}$ . Similarly, the tryptophan nest of Trp 27, Trp 29, and Trp 207 differs in structure by only  $0.65 \text{ \AA}$ . The catalytic residues of the active site itself (His 57, Asp 102, Ser 195) deviate from exact symmetry by  $1.0 \text{ \AA}$ , even though

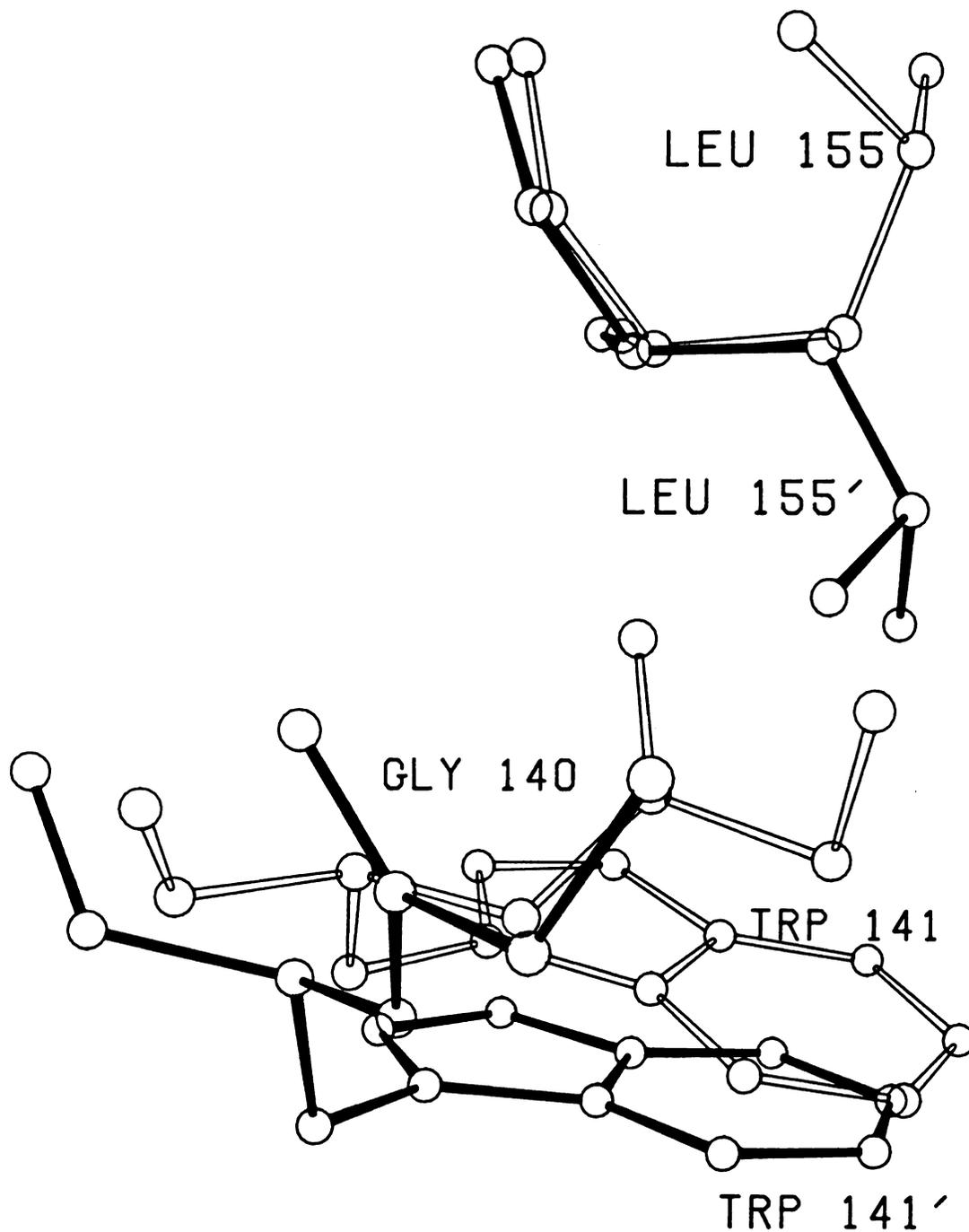


Figure 11. Interaction of Leu 155 with Gly 140-Trp 141 in molecule one (open bonds) and one prime (dark bonds).

this region is on the dimer interface, a region of high structural variability.

The active sites of the two molecules are superimposed in Figure 12. It can be seen that much of the deviation in side chain structure consists of a translation of molecule one prime to the left in Figure 12 relative to molecule one, and that the relative configuration of the active sites is very similar in the two molecules. Here the fidelity of the coordinate measurement technique is clearly demonstrated. It was mentioned earlier that the His 57 side chain in molecule one prime was rotated relative to molecule one. It can be seen in Figure 12 that this small difference in structure has been faithfully translated into the coordinates. Thus, the requirement of accuracy in coordinate measurement has been met with the transit and cathetometer, and it can be appreciated that the combination of this measurement procedure with the excellent quality high resolution electron density map has made the current work possible and capable of providing detailed structural insight.

#### 4. Significance of the differences

The comparisons carried out here have been of two models that were built independently from a high quality electron density map, and the differences discussed are differences between these models. The model coordinates have not been refined. There are two levels of significance to consider.

Figure 12. Superposition of the active sites of molecule one (open bonds) and one prime (dark bonds) of  $\alpha$ -CHT.

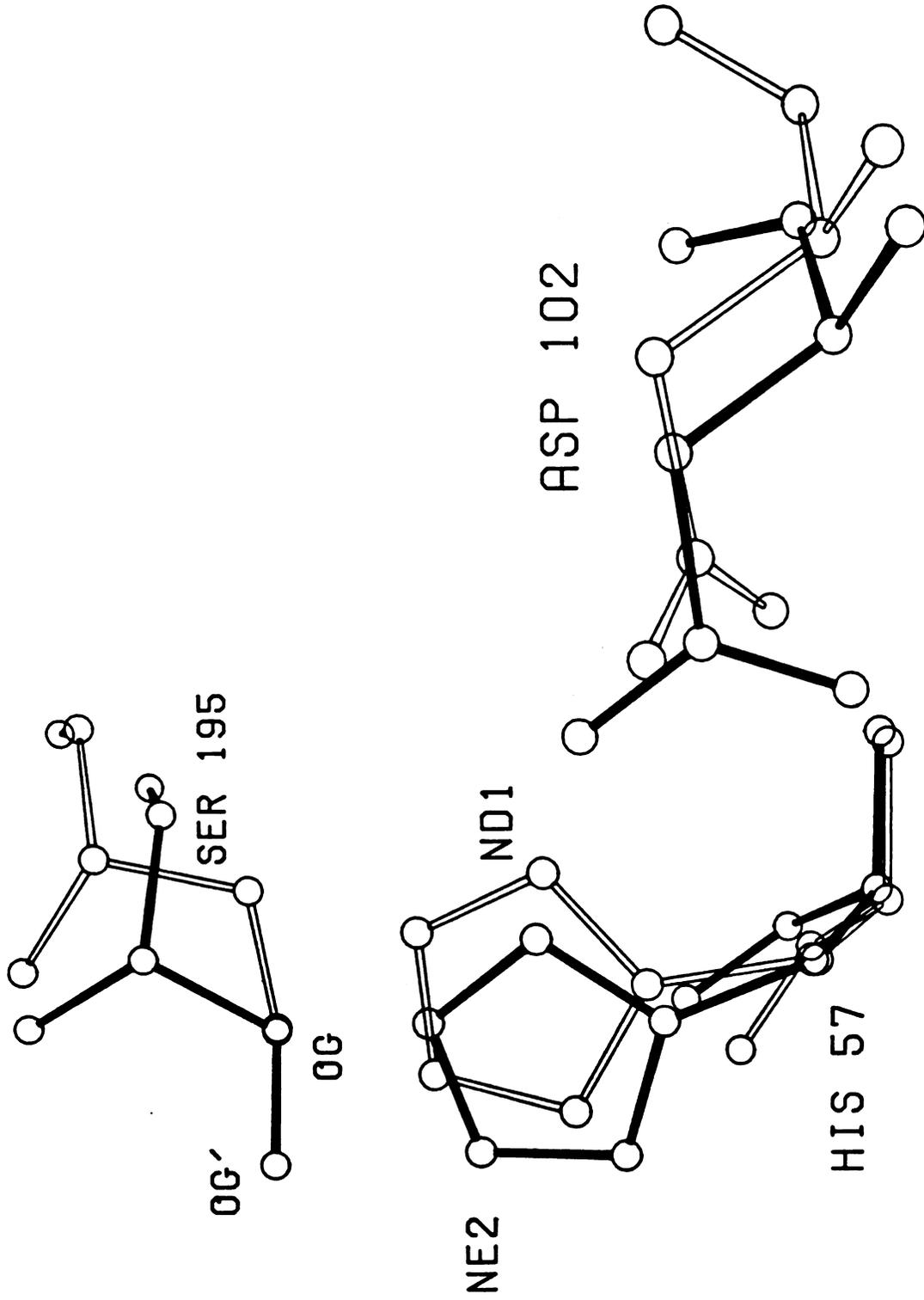


Figure 12

The first is the magnitude of r.m.s deviation in coordinates that corresponds to a real difference in the models as they were built independently to the electron density. The second is the correspondence of these model adjustments, which were the subjective judgement of the experimenter, to significant differences in the actual structure of CHT in the crystal.

The first level has to do with how faithfully the idealized measured coordinates of the molecule reflect the conformation of the Kendrew model. A unique feature of this work is the high accuracy of the coordinate measurement technique. As was seen in Chapter 2, the standard deviation of bond distances calculated from the measured coordinates is approximately  $0.1 \text{ \AA}$ . This implies a standard error in measured coordinates on the order of  $0.05 \text{ \AA}$ . The raw measured coordinates were subjected to the idealization procedure of Diamond<sup>44</sup> and, as was noted earlier, the r.m.s. deviation between measured model coordinates and final idealized coordinates was  $0.17 \text{ \AA}$ . In this way, it can be roughly estimated that the final idealized coordinates match the model within  $0.2 \text{ \AA}$ , or somewhat less than the sum of the above numbers. The error expected in the difference between the coordinate sets is then given by  $\sqrt{2} * 0.2 \sim 0.3 \text{ \AA}$ . This estimate of the minimum significant difference is consistent with the observed r.m.s. deviation ( $0.43 \text{ \AA}$ ) of the coordinates of the main chain of the 68 residues which were used to determine the transformation between molecules (Table 4). In addition,

of the 68 residues which have a total r.m.s. deviation of 0.43 Å, a subset exists which has an even lower r.m.s. value. This subset contains all those residues whose backbone was not readjusted at all from molecule one to molecule one prime and includes residues Leu 33-Asp 35, Trp 51, Val 52, and Asn 150-Leu 162. This subset has an r.m.s. deviation of 0.33 Å, in very good agreement with the expected error. From such an analysis of error, it can be concluded conservatively that a r.m.s. deviation of 0.5 Å or greater represents a different conformation of the molecule in the two models. This is very consistent with the observation of r.m.s. deviation against residue number (see Figure 10) and with the notes made during the refitting of molecule one to molecule one prime. That is, those regions which were readjusted have r.m.s. deviations above 0.5 Å, while regions not readjusted have an r.m.s. deviation of about 0.4 Å. Therefore the values shown in Table 5 represent significant variations in structure between the two models.

The second question concerns which of the deviations in model structure reflect a real difference in the structure of CHT in the dimer. Obviously the largest differences are significant, and there is no doubt that the two molecules differ in structure in the dimer interface, and at certain regions of the exterior such as Gly 69 and Arg 145. On the other hand, the r.m.s. deviation of the C-terminal helix Leu 234-Asn 245 is only 0.43 Å, yet some significant

difference in density is observed in this region, which indicates that there is significant structure difference in this region. It is possible that real difference in structure exists here, but that it is so small that it is beyond the power of the current coordinate measurement technique to resolve. A somewhat similar situation obtains at His 57, where the side chain deviation is only 0.74 Å and where there is also good and consistent indication of structural difference in the difference density map. Therefore there exists significant difference density in regions of the molecule where the observed r.m.s. deviation between the two molecules is about equal to or only slightly greater than the standard error of the r.m.s. deviation. From this we conclude that even the value of 0.8 Å for the r.m.s. deviation of the main chain of the interior can represent a significant difference in structure between the two molecules. Also, because the values for the r.m.s. deviations of the various groups are consistent with the amount of rebuilding done to the exterior as opposed to the interior, we conclude that the exterior of the molecule, including the dimer interface, differs in structure significantly more than the interior residues.

It is therefore likely that on further refinement, the differences observed here between the two molecules will be in large part retained. Certainly not every detail of the structure and of the differences between the two molecules is

correct and significant, but it can be stated with confidence that the two independent molecules of CHT dimer differ in structure, and that qualitatively if not quantitatively, the differences are those described in this work.

## CHAPTER 5

### Comparison to Other Structures

#### 1. $\alpha$ -CHT and $\gamma$ -CHT

As noted earlier,  $\gamma$ -CHT is a monomeric form of chymotrypsin, and crystallizes at a higher pH (5.5). The structure has recently been refined by least squares methods at 1.9 Å resolution.<sup>39</sup> Refined  $\gamma$ -CHT coordinates were obtained from the Protein Data Bank, Brookhaven National Laboratory. The  $\gamma$ -CHT molecule was compared to molecule one and molecule one prime of  $\alpha$ -CHT by the method described in Chapter 4 for the comparison of molecule one to one prime. Table 6 presents a summary of that comparison. The transformation selected for comparisons in Table 6 used all residues, level 3, in the comparison, since there was no indication of greater or lesser similarity in structure for any particular region as there was in  $\alpha$ -CHT.

The same general trends are seen in Table 6 as in Table 5; the main chain matches much better than the side chains, and the carbonyl oxygen atoms have a higher r.m.s. deviation

Table 6. Comparison of  $\alpha$ -CHT to  $\gamma$ -CHT.

Group of Atoms	r.m.s. Deviation ( $\text{\AA}$ ) of $\gamma$ -CHT from		Number of Atoms
	$\alpha$ -CHT One	$\alpha$ -CHT One Prime	
All	2.4 (1.7)	2.4 (1.8)	1764 (1551)
Main Chain*	1.4 (1.0)	1.5 (1.2)	723 (639)
Side Chain	3.1 (2.2)	3.1 (2.2)	797 (696)
Alpha Carbons	1.5 (1.0)	1.6 (1.3)	241 (213)
Carbonyl Oxygens	2.1 (1.7)	2.2 (1.9)	244 (216)
His 57, Asp 102, Ser 195	1.0	1.0	24
<b>Dimer Interface</b>			
All	1.8	1.9	288
Main Chain*	1.2	1.4	117
Side Chain	2.2	2.3	132
<b>Interior Residues</b>			
All	1.2	1.4	530
Main Chain*	0.8	0.9	213
Side Chain	1.5	1.7	246
<b>Exterior Residues</b>			
All	2.6 (1.9)	2.6 (2.1)	808 (595)
Main Chain*	1.4 (1.1)	1.6 (1.4)	333 (249)
Side Chain	3.3 (2.5)	3.3 (2.5)	364 (263)
<b>Domain 1</b>			
All	3.0 (1.8)	3.0 (1.9)	887 (674)
Main Chain*	1.8 (1.0)	1.8 (1.3)	360 (276)
Side Chain	3.9 (2.3)	3.8 (2.4)	406 (306)

Table 6 (cont'd)

Domain 2			
All	1.6	1.8	877
Main Chain*	1.0	1.2	363
Side Chain	2.0	2.1	393

\* Calculation for main chain atoms does not include carbonyl oxygens.

Note: Values in parentheses do not include Val 65-Ser 82, which have r.m.s. deviations of 3.3 Å and 7.0 Å for main and side chains, respectively.

than does the rest of the main chain. Again, the structurally significant regions match in structure better than other parts of the molecule. Thus the r.m.s. deviations of the active site region, the disulfide bridge sulfur atoms, and the tryptophan nest of Trp 27, Trp 29, and Trp 207 are 1.0 Å, 1.0 Å, and 0.8 Å, respectively. These are the values of the molecule one,  $\gamma$ -CHT comparison, but the values for molecule one prime,  $\gamma$ -CHT are comparable. These values are all well within the overall r.m.s. deviation of 1.7 Å. The values of r.m.s. deviation of  $\gamma$ -CHT from molecule one and from molecule one prime do not differ greatly, but the general trend in Table 6 is a slightly better fit to molecule one.

The region Val 65-Ser 92 gives a very poor fit between  $\alpha$ -CHT and  $\gamma$ -CHT, with an overall r.m.s. deviation of 3.3 Å in the main chain, and 7.0 Å in side chains (Table 7). Since almost all of these residues are on the exterior, these residues make a large contribution to the r.m.s. deviation of the exterior residues, and the effect is seen in Table 6. It is also seen in Figures 13 and 14, which are plots of alpha carbon deviation against residue number for  $\gamma$ -CHT with  $\alpha$ -CHT molecule one and one prime respectively. This region of large structural difference also contributes greatly to the higher r.m.s. deviation of domain 1 over domain 2, since residues 65-92 are all in domain 1. However, determination of transformation parameters from domain 1 of  $\gamma$ -CHT and  $\alpha$ -CHT and then from domain 2 yielded parameters which were

Table 7. R.m.s. Deviations of Val 65-Ser 92 in  $\alpha$ -CHT and  $\gamma$ -CHT.

<u>Residue</u>	r.m.s. Deviation ( $\text{\AA}$ ) of $\gamma$ -CHT from			
	<u>Molecule one</u>		<u>Molecule one prime</u>	
	<u>main chain</u>	<u>side chain</u>	<u>main chain</u>	<u>side chain</u>
Val 65	2.1	4.2	1.2	1.2
Val 66	2.1	5.5	1.4	1.5
Val 67	3.1	7.3	1.4	2.6
Ala 68	3.2	5.0	1.0	1.0
Gly 69	3.2	-	1.2	-
Glu 70	2.4	9.0	2.2	8.6
Phe 71	3.8	4.8	2.9	3.1
Asp 72	4.9	8.9	4.6	8.7
Gln 73	3.2	2.8	3.0	3.6
Gly 74	3.1	-	3.0	-
Ser 75	2.9	3.5	3.5	2.9
Ser 76	2.8	2.2	2.8	2.2
Ser 77	4.7	7.3	4.7	7.3
Glu 78	1.4	1.0	1.4	1.0
Lys 79	2.2	5.5	2.2	5.5
Ile 80	3.0	6.8	3.0	6.7
Gln 81	3.6	9.6	3.7	9.6
Lys 82	3.3	8.8	3.3	8.9
Leu 83	3.3	7.0	3.4	6.9
Lys 84	3.2	8.2	3.4	8.1
Ile 85	2.7	6.1	2.8	6.0
Ala 86	3.3	6.2	3.3	6.1
Lys 87	3.2	-*	3.2	-*
Val 88	3.8	7.5	3.5	6.6
Phe 89	3.2	7.3	3.1	7.0
Lys 90	3.3	10.4	3.4	10.4
Asn 91	2.4	6.3	2.8	7.2
Ser 92	3.1	5.8	3.3	5.6

\* Lys 87 side chain atoms were left out of the comparison due to poor electron density at this residue in  $\alpha$ -CHT.

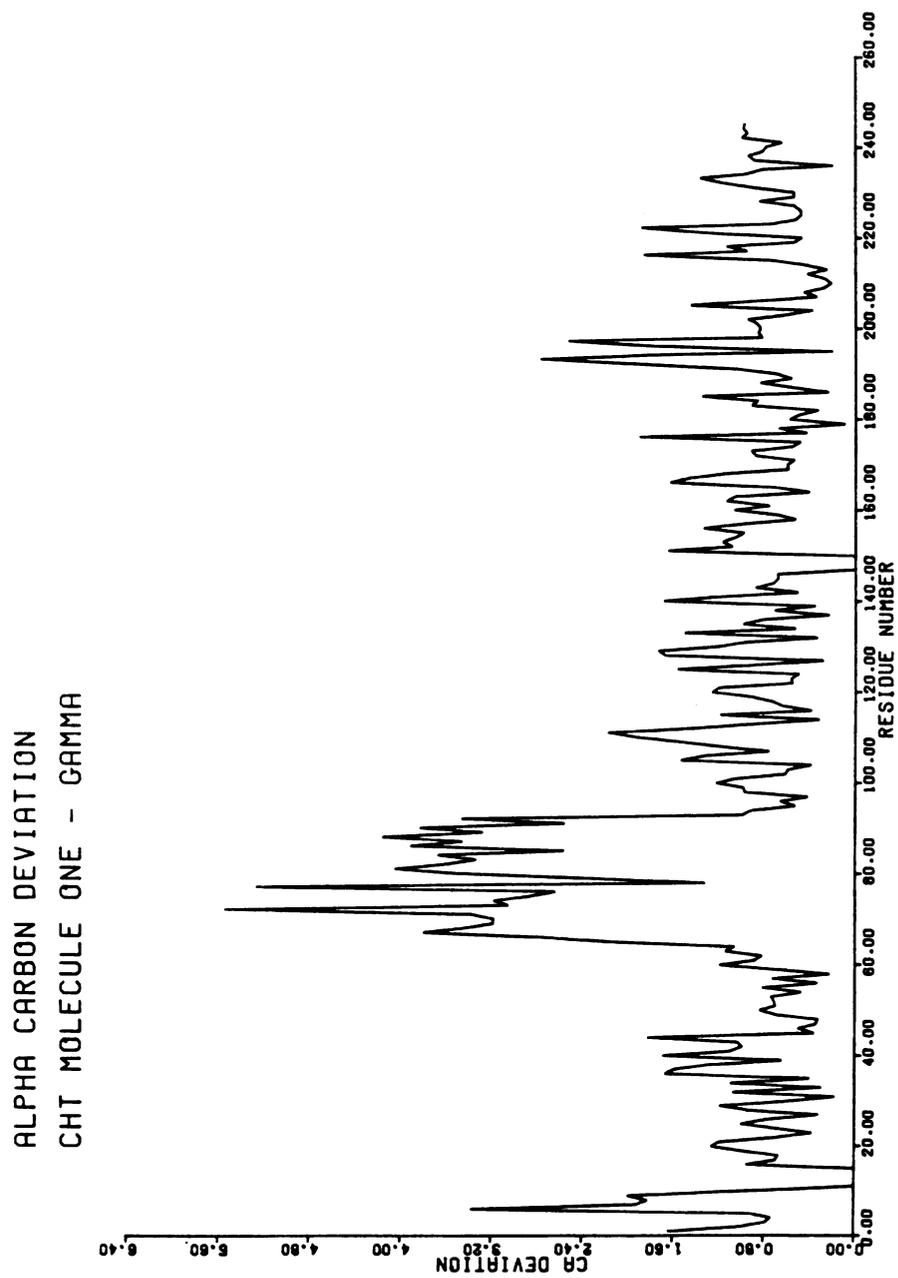


Figure 13. Comparison of alpha carbon positions in  $\gamma$ -CHT and molecule one of  $\alpha$ -CHT.

ALPHA CARBON DEVIATION  
CHT ONE PRIME - GAMMA

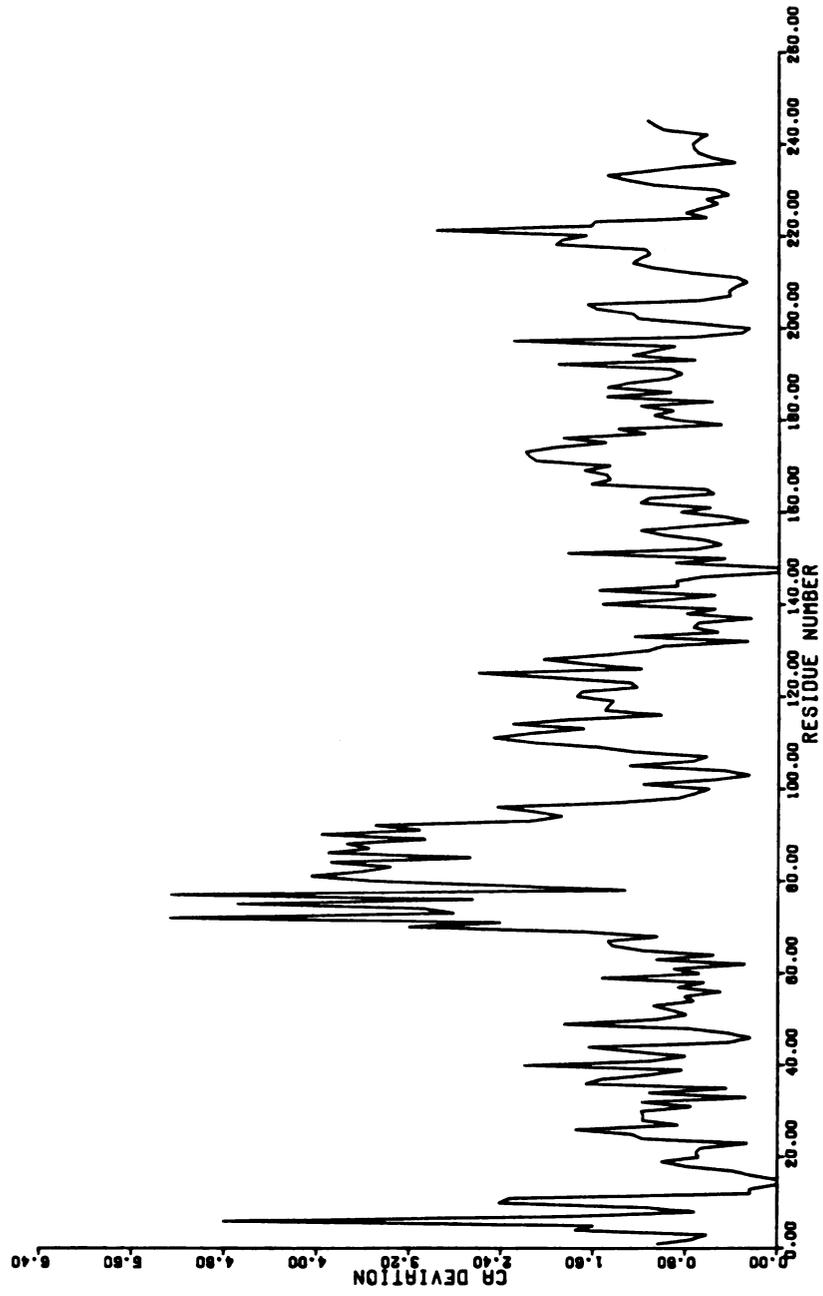


Figure 14. Comparison of alpha carbon positions in  $\gamma$ -CHT and molecule one prime of  $\alpha$ -CHT.

essentially identical. When the contribution of residues 65-92 is removed from the calculation, the values shown in parentheses in Table 6 are obtained. It can be seen that removal of the 28 residues from consideration results in values of r.m.s. deviations that are smaller by approximately 30%. Residues 65-70 correspond to a region of large difference in configuration in  $\alpha$ -CHT, and the electron density for residues 76-82 is very poor. However, residues 83-92 are quite well defined in the  $\alpha$ -CHT electron density, and so the large difference in this region between  $\alpha$ -CHT and  $\gamma$ -CHT is highly significant.

Superpositions of the active sites of  $\alpha$ -CHT molecules one and one prime with that of  $\gamma$ -CHT are shown in Figures 15 and 16, respectively. It can be seen that there are slightly larger differences between  $\alpha$ -CHT and  $\gamma$ -CHT than between the two molecules of  $\alpha$ -CHT. Of particular interest is the orientation of the Ser 195 side chain. In  $\gamma$ -CHT, it is seen to point "up", away from the His 57 side chain relative to Ser 195 in  $\alpha$ -CHT. The significance of this result can be assessed better and a more detailed comparison of  $\alpha$ -CHT to  $\gamma$ -CHT can be made after a refinement has been performed on the  $\alpha$ -CHT structure comparable to that carried out on  $\gamma$ -CHT.

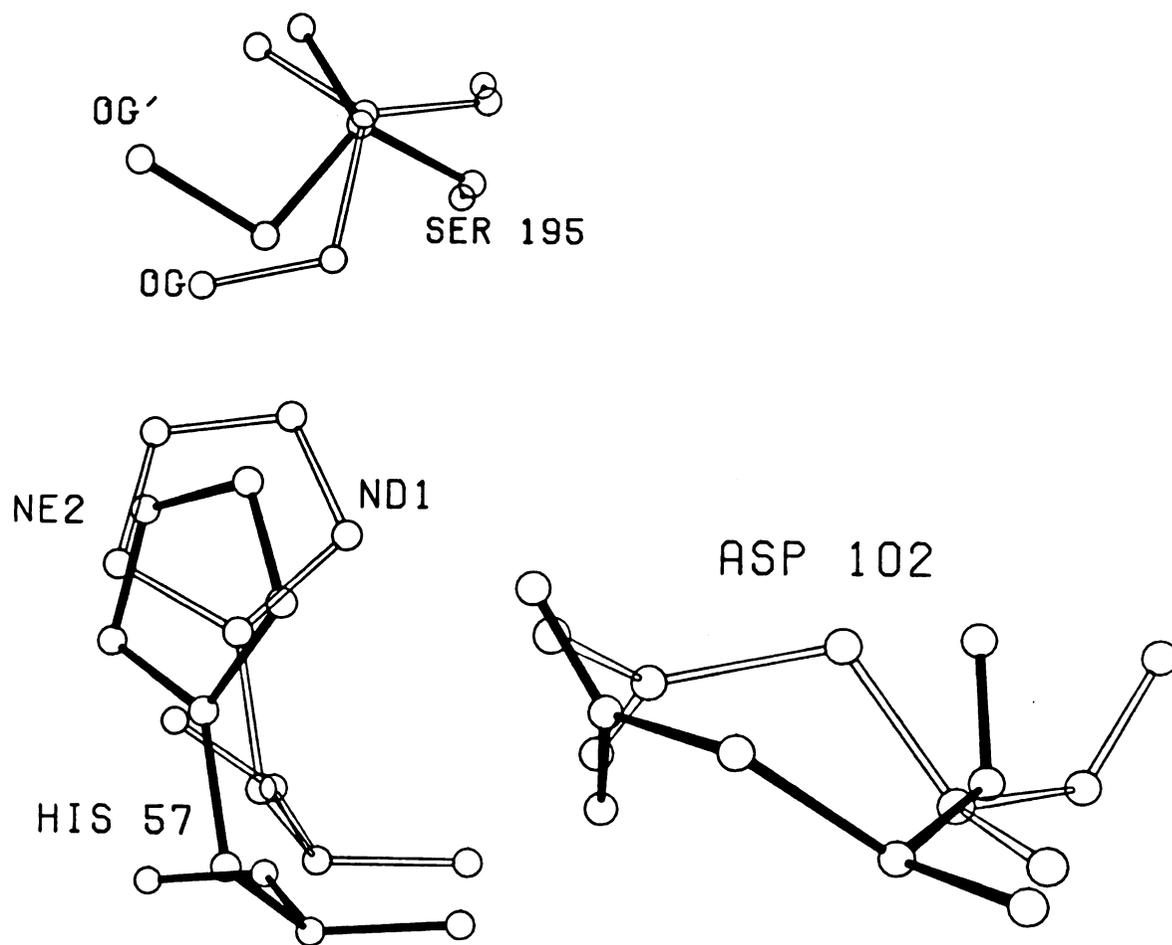


Figure 15. Superposition of the active sites of  $\gamma$ -CHT (dark bonds) and molecule one of  $\alpha$ -CHT.

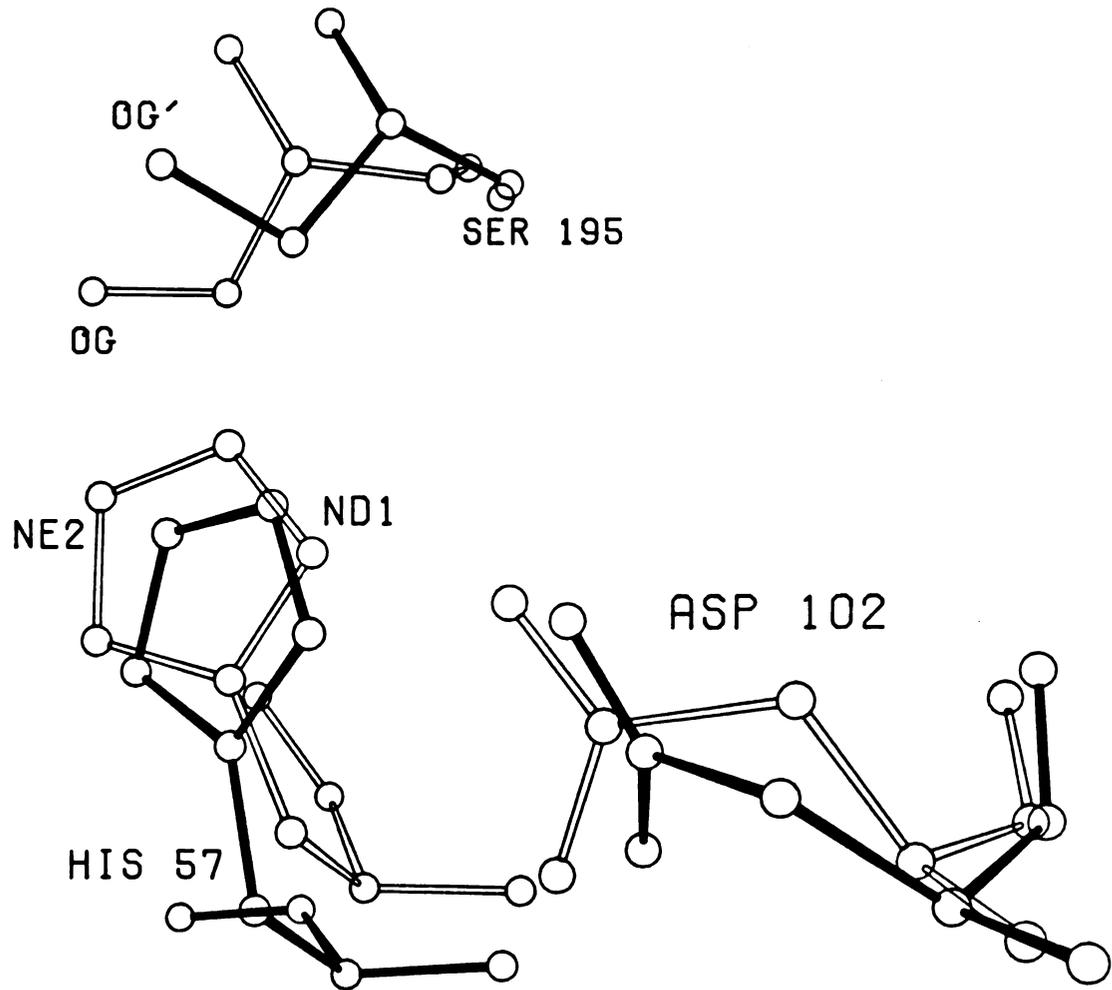


Figure 16. Superposition of the active sites of  $\gamma$ -CHT (dark bonds) and molecule one prime of  $\alpha$ -CHT.

## 2. Concluding remarks

At least one other set of coordinates for a dimer related by non-crystallographic symmetry has been reported. The structure is that of triose phosphate isomerase (TIM).<sup>15</sup> The coordinates of both molecules were compared by the method of Rossmann and Argos,<sup>58</sup> as was also done here with the two molecules of CHT. The authors report an overall r.m.s. deviation of 1.2 Å in the alpha carbons between the two molecules of TIM, while that of alpha carbons involved in regular secondary structure ( $\alpha$ -helix,  $\beta$ -sheet) was 0.78 Å. These two values are close to the corresponding values listed in Table 5 for the overall main chain and the interior main chain, respectively, of CHT. Although these authors did not discuss molecular asymmetry, the present work indicates that a careful analysis of the TIM dimer might well reveal additional examples of variability in protein oligomeric structure.

On the other hand, the structure determination of CHT might be unique in its quality. From the start, great care was taken in intensity measurements, and innovative diffractometric techniques were developed to make the measurements. Furthermore, the multiple isomorphous replacement phases were calculated from six derivatives of high quality, and the 2.8 Å resolution map thus obtained was extended to 1.8 Å resolution through careful, well-monitored and documented application of a mild density modification technique. More than usual care was taken in fitting the CHT model to this

electron density map, and an innovative, very accurate and precise technique and protocol of coordinate measurement was developed in order to transform the model thus built into a reliable set of coordinates. There is no doubt that such requirements of accuracy, precision, and reliability are an important reason that such analyses of protein structural variability have not been attempted previously.

## REFERENCES

## REFERENCES

1. B.W. Matthews, P.B. Sigler, R. Henderson, and D.M. Blow, *Nature* 214, 652-656 (1967).
2. J.J. Birktoft and D.M. Blow, *J. Mol. Biol.* 68, 187-240 (1972).
3. A. Tulinsky, R.L. Vandlen, C.N. Morimoto, N.V. Mani, and L.H. Wright, *Biochemistry* 12, 4185-4192 (1973).
4. A. Tulinsky, N.V. Mani, C.N. Morimoto, and R.L. Vandlen, *Acta Crystallogr.* B29, 1309-1322 (1973).
5. N.V. Raghavan and A. Tulinsky, *Acta Crystallogr.* B35, 1776-1785 (1979).
6. A. McPherson, I. Molineux, and A. Rich, *J. Mol. Biol.* 106, 1077-1081 (1976).
7. P.R. Evans, P.J. Hudson, *Nature* 279, 500-504 (1979).
8. T.J. Mercolino, H.D. Bellamy, and F.S. Mathews, *J. Mol. Biol.* 139, 557-560 (1980).
9. R.L. Vandlen, D.L. Ersfeld, A. Tulinsky, and W.A. Wood, *J. Biol. Chem.* 248, 2251-2253 (1973).
10. M.J. Adams, G.C. Ford, R. Koekoek, P.J. Lentz, Jr., A. McPherson, Jr., M.G. Rossmann, I.E. Smiley, R.W. Schevitz, and A.J. Wonacott, *Nature* 227, 1098-1103 (1970).
11. C.-I. Brändén, H. Eklund, B. Nordström, T. Boiwe, G. Söderlund, E. Zeppezauer, I. Ohlsson, and Å. Åkeson, *Proc. Nat. Acad. Sci. USA* 70, 2439-2442 (1973).
12. J.S. Richardson, K.A. Thomas, B.H. Rubin, and D.C. Richardson, *Proc. Nat. Acad. Sci. USA* 72, 1349-1353 (1975).

13. J.W. Campbell, H.C. Watson, and G.I. Hodgson, *Nature* 250, 301-303 (1974).
14. M.S. Weininger and L.J. Banaszak, *J. Mol. Biol.* 119, 443-449 (1978).
15. D.W. Banner, A.C. Bloomer, G.A. Petsko, D.C. Phillips, and I.A. Wilson, *Biochem. Biophys. Res. Comm.* 72, 146-155 (1976).
16. P.C. Weber, R.G. Bartsch, M.A. Cusanovich, R.C. Hamlin, A. Howard, S.R. Jordan, M.D. Kamen, T.E. Meyer, D.W. Weatherford, Ng. h. Xuong, and F.R. Salemme, *Nature* 286, 302-304 (1980).
17. I. Mavridis and A. Tulinsky, *Biochemistry* 15, 4410-4417 (1976).
18. The Molecular Replacement Method, ed. M.G. Rossmann, Gordon and Breach: New York (1972).
19. M.R.N. Murthy, R.M. Garavito, J.E. Johnson, and M.G. Rossmann, *J. Mol. Biol.* 138, 859-872 (1980).
20. G. Biesecker, J.I. Harris, J.C. Thierry, J.E. Walker, and A.J. Wonacott, *Nature* 266, 328-333 (1977).
21. G. Bricogne, *Acta Crystallogr.* A32, 832-847 (1976).
22. C. Abad-Zapatero, S.S. Abdel-Meguid, J.E. Johnson, A.G.W. Leslie, I. Rayment, M.G. Rossmann, D. Suck, and T. Tsukihara, *Nature* 286, 33-39 (1980).
23. T. Unge, L. Liljas, B. Strandberg, I. Vaara, K.K. Kannan, K. Fridborg, C.E. Nordman, and P.J. Lentz, Jr. , *Nature* 285, 373-377 (1980).
24. A.C. Bloomer, J.N. Champness, G. Bricogne, R. Staden, and A. Klug, *Nature* 276, 362-368 (1978).
25. S.C. Harrison, A.J. Olson, C.E. Schutt, F.K. Winkler, and G. Bricogne, *Nature* 276, 368-373 (1978).
26. M. Buehner, G.C. Ford, D. Moras, K.W. Olsen and M.G. Rossmann, *J. Mol. Biol.* 82, 563-585 (1974).
27. A.V. Xavier, E.W. Czerwinski, P.H. Bethge, and F.S. Mathews, *Nature* 275, 245-247 (1978).
28. E.T. Adman, R.E. Stenkamp, L.C. Sieker, and L.H. Jensen, *J. Mol. Biol.* 123, 35-47 (1978).

29. R.L. Vandlen and A. Tulinsky, *Acta Crystallogr.* B27, 437-442 (1971).
30. D.M. Collins, M.D. Brice, T.F.M. laCour, and M.J. Legg, *Crystallographic Computing Techniques*, ed. F.R. Ahmed, pp. 330-335, Munksgaard: Copenhagen (1976).
31. A. Tulinsky, I. Mavridis, and R.F. Mann, *J. Biol. Chem.* 253, 1074-1078 (1978).
32. A. Tulinsky, *Biomolecular Structure, Conformation, Function, and Evolution*, vol. 1, ed. R. Srinivasan, pp. 183-199, Pergamon Press: Oxford and N.Y. (1980).
33. R.L. Vandlen and A. Tulinsky, *Biochemistry* 12, 4193-4200 (1973).
34. A. Mavridis, A. Tulinsky, and M.N. Liebman, *Biochemistry* 13, 3661-3666 (1974).
35. L.S. Hibbard and A. Tulinsky, *Biochemistry*, 17, 5460-5468 (1978).
36. A. Tulinsky and L.H. Wright, *J. Mol. Biol.* 81, 47-56 (1973).
37. J.H. Konnert, *Acta Crystallogr.* A32, 614-617 (1976).
38. I. Mavridis, A. Tulinsky, M.H. Hatada, L. Lebioda, in preparation.
39. G.H. Cohen, E.W. Silverton, and D.R. Davies, *J. Mol. Biol.* 148, 449-479 (1981).
40. G.H. Cohen, B.W. Matthews, and D.R. Davies, *Acta Crystallogr.* B26, 1062-1069 (1970).
41. D.A. Matthews, R.A. Alden, J.J. Birktoft, S.T. Freer, and J. Kraut, *J. Biol. Chem.* 252, 8875-8883 (1977).
42. F.M. Richards, *J. Mol. Biol.* 37, 225-230 (1968).
43. D. Tsernoglou, G.A. Petsko, J.E. McQueen, Jr., and J. Hermans, *Science* 197, 1378-1381 (1977).
44. R. Diamond, *Acta Crystallogr.* 21, 253-266 (1966).
45. F.R. Salemme and D.G. Fehr, *J. Mol. Biol.* 70, 697-700 (1972).
46. R.M. Burnett, G.D. Darling, D.S. Kendall, M.E. LeQuesne, S.G. Mayhew, W.W. Smith, and M.L. Ludwig, *J. Biol. Chem.* 249, 4383-4392 (1974).

47. M.A. Frentrup and A. Tulinsky, *J. Appl. Cryst.*, in press.
48. F.A. Saul, L.M. Amzel, and R.J. Poljak, *J. Biol. Chem.* 253, 585-597 (1978).
49. M.F. Perutz, *J. Mol. Biol.* 13, 646-668 (1965).
50. J.C. Kendrew, R.E. Dickerson, B.E. Strandberg, R.G. Hart, D.R. Davies, D.C. Phillips, and V.C. Shore, *Nature* 185, 422-427 (1960).
51. A.R. Sielecki, W.A. Hendrickson, C.G. Broughton, L.T.J. Delbaere, G.D. Brayer, and M.N.G. James, *J. Mol. Biol.* 134, 781-804 (1979).
52. C.M. Venkatachalam, *Biopolymers* 6, 1425-1436 (1968).
53. B.W. Matthews, *J. Mol. Biol.* 33, 491-497 (1968).
54. S.E.V. Phillips, *J. Mol. Biol.* 142, 531-554 (1980).
55. K.D. Watenpaugh, T.N. Margulis, L.C. Sieker, and L.H. Jensen, *J. Mol. Biol.* 122, 175-190 (1978).
56. G.M. Ramachandran, C. Ramakrishnan and V. Sasisekharan, *J. Mol. Biol.* 7, 95-99 (1963).
57. S.T. Freer, R.A. Alden, C.W. Carter, Jr., and J. Kraut, *J. Biol. Chem.* 250, 46-54 (1975).
58. M.G. Rossmann and P. Argos, *J. Biol. Chem.* 250, 7525-7532 (1975).

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03056 6634