Chymotrypsinogen: 2.5-Å Crystal Structure, Comparison with α-Chymotrypsin, and Implications for Zymogen Activation*

S. T. Freer, J. Kraut, J. D. Robertus, H. T. Wright,† and Ng. H. Xuong

ABSTRACT: A 2.5-Å electron density map of bovine chymotrypsinogen A was calculated from phases determined with the aid of four heavy-atom derivatives. A preliminary model constructed from this map was compared with the reported 2.0-Å structure of α-chymotrypsin (Matthews, B. W., Sigler, P. B., Henderson, R., and Blow, D. M. (1967), *Nature 214*, 652; Sigler, P. B., Blow, D. M., Matthews, B. W., and Henderson, R. (1968), *J. Mol. Biol. 35*, 143). The α-chymotrypsin model used in the comparison was assembled from the coordinates of Birktoft *et al.* (Birktoft, J. J., Matthews, B. W., and Blow, D. M. (1969), *Biochem. Biophys. Res. Commun. 36*, 131). The overall folding of the chymotrypsinogen polypeptide chain is very similar to that of α-chymotrypsin. With the exception of a ten-residue segment of α helix at the C terminus, the backbone chain is more or less fully extended and often doubles back on itself to form large sections of distorted anti-parallel pleated sheet. Comparison of the structures of the zymogen and α-chymotrypsin reveals the essential changes associated with the primary activation process, since the π, δ, γ, and α forms of the enzyme are all nearly identical in structure. The results of this comparison can be classified under three general headings: (1) Ion Pair Formation: Upon tryptic scission of the Arg-15–Ile-16 peptide bond in chymotrypsinogen, the main chain rotates 180° about an axis roughly defined by the α-carbon atoms of residues 19 and 21. In this process the formerly exposed side chains of Ile-16 and Val-17 are buried, and the α-amino group of Ile-16 approaches the buried side-chain carboxyl of Asp-194 to form the internal ion pair previously found in α-chymotrypsin. Asp-194 is also buried in chymotrypsinogen, but its side chain is hydrogen bonded to the side chain of His-40. During activation the main chain rotates 180° about the carbonyl carbon to α-carbon bond of Asp-194, shifting the side chain of Asp-194 from the vicinity of His-40 into the vicinity of the new position of Ile-16. This rotation is also responsible for the large movement of Met-192 described below. (2) Catalytic Site: The hydrogen-bond network linking Ser-195, His-57, and Asp-102 previously found at the catalytic site of α-chymotrypsin is preformed in chymotrypsinogen. The "switching on" of catalytic activity during activation may, therefore, result from relatively small structural modification of this region of the molecule. Three such small changes are observed, one involving a slight movement of the His-57 side chain, the second involving the side chain of Ser-214, and the third involving the side chain of Ile-99, but it is not clear how, or even whether, they are responsible for genesis of enzymatic activity. (3) Specificity Cavity: The specificity cavity near Ser-195 is only partially formed in the zymogen. The cavity is completed during activation by a conformational change in the backbone chain which also shifts Met-192 from a completely buried position out to the surface of the molecule.

The crystal structure of bovine chymotrypsinogen A has been under investigation in this laboratory for 12 years. The early stages of the work were frustrated by lack of heavy-atom derivatives; a survey of over 200 possible heavy-atom reagents resulted in only one potentially useful heavy-atom derivative. This problem was solved in 1961 with the discovery of a new crystalline form of the zymogen (type F crystals) which quickly yielded four heavy-atom derivatives and a 5-Å electron density map (Kraut *et al.*, 1962). The perils of low-resolution model building and real space refinement were then explored while the search for further derivatives was continued and the resolution was extended to 4 Å (Kraut *et al.*, 1964). Further work on this structure was hampered by the fragility of the crystals and their extreme susceptibility to slipping during the measurement of X-ray intensity data. It was only when aquisition of an automatic diffractometer and the development of screenless precession photography had increased our speed of data collection to the point where entire sets of three-dimensional 2.5-Å data could be discarded with impunity, that extension of the resolution of chymotrypsinogen to 2.5 Å became practical.

Bovine chymotrypsinogen A, the inactive precursor of the π, δ, γ, and α family of chymotrypsins A, is a pancreatic protein composed of 245 amino acid residues arranged in a single polypeptide chain. It is cross-linked by five disulfide bridges, one of which includes the N-terminal residue. The complete amino acid sequence has been worked out by Hartley *et al.* (Hartley, 1964; Hartley and Kauffman, 1966; Blow *et al.*, 1969) and by Meloun *et al.* (1966).

Three distinct chymotrypsinogens, designated A, B, and C, have been isolated and characterized from bovine and porcine pancreas: bovine chymotrypsinogen A (Kunitz and Northrop, 1935), porcine chymotrypsinogen A (Charles *et al.*, 1967), bovine chymotrypsinogen B (Brown *et al.*, 1948), porcine chymotrypsinogen B (Gratecos *et al.*, 1969), subunit II of bovine procarboxypeptidase A (Brown *et al.*,

* From the Department of Chemistry, University of California at San Diego, La Jolla, California 92037. *Received February 9, 1970.* Ng. H. X. is also a member of the Departments of Biology and Physics. This research was supported by grants from the U. S. Public Health Service and the National Science Foundation.

† Present address: Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge, England.

1963), and porcine chymotrypsinogen C (Folk and Schirmer, 1965). The amino acid sequence of bovine chymotrypsinogen B (Smillie *et al.*, 1968), which also contains 245 residues, is very similar to the sequence of chymotrypsinogen A. Porcine chymotrypsinogen C, on the other hand, has a slightly higher molecular weight (Gratecos *et al.*, 1969) and its amino acid composition and N-terminal sequence closely resemble those of bovine subunit II (Peanasky *et al.*, 1969). All of these chymotrypsinogens are activated by tryptic cleavage of the first peptide bond in the N-terminal sequence that has its carbonyl group contributed by a basic residue. Upon activation, the A, B, and C zymogens give rise to enzymes which have similar, but not identical, catalytic activity.

Bovine chymotrypsinogen A, hereafter referred to simply as chymotrypsinogen, can be activated to yield several chymotrypsins A: α-chymotrypsin (Kunitz and Northrop, 1935), β- and γ-chymotrypsin (Kunitz, 1938), and π- and δ-chymotrypsin (Jacobsen, 1947). The zymogen can also be converted into enzymatically inert neochymotrypsinogens which, in turn, may be activated to the α form of the enzyme (Rovery *et al.*, 1957). The specific peptide-bond cleavages associated with the activation of chymotrypsinogen were worked out independently in the laboratories of Neurath and of Desnuelle (Dreyer and Neurath, 1955; Rovery *et al.*, 1955; Rovery *et al.*, 1957). Chymotrypsinogen is activated by tryptic cleavage of the Arg-15–Ile-16 peptide bond to give π-chymotrypsin; autocatalytic hydrolysis of bonds between residues 13–14, 146–147, and 148–149 gives rise to the δ, γ, and α forms of the enzyme, or to various neochymotrypsinogens, depending upon the order in which the four bonds are split. Wright *et al.* (1968) have proposed an activation scheme relating these proteins. The 2.0-Å structure of the well-known α-chymotrypsin, which differs from chymotrypsinogen by deletion of the two dipeptides Ser-14–Arg-15 and Thr-147–Asn-148, has been determined by the Cambridge group (Matthews *et al.*, 1967; Sigler *et al.*, 1968; Birktoft *et al.*, 1969).

All crystallographic results to date indicate that the conformations of the π, γ, and δ forms of chymotrypsin A are essentially identical with each other and with the α form of the enzyme (Kraut *et al.*, 1967; Wright *et al.*, 1968; Matthews *et al.*, 1968; Cohen *et al.*, 1969; Davies *et al.*, 1969). Therefore, the essential features of the activation process should be apparent from a comparison of the chymotrypsinogen and α-chymotrypsin structures. One striking change which accompanies activation, the formation of an ion pair between the N terminus of Ile-16 and the side-chain carboxyl group of Asp-194, was already evident from examination of the α-chymotrypsin structure (Matthews *et al.*, 1967). The role of the N-terminal group in stabilizing the enzyme conformation was noted by Hess and coworkers (Oppenheimer *et al.*, 1966; Hess, 1968).

Induction of biological activity by limited proteolysis is widespread in nature, having been observed in bacterial, yeast, plant, invertebrate, and vertebrate systems. This general topic has been reviewed by Ottesen (1967) and Neurath (1957). The tryptic activation of the pancreatic proenzymes trypsinogen, chymotrypsinogen, and procarboxypeptidase are well-known examples of this process: in the conversion of chymotrypsinogen into π-chymotrypsin only a single peptide bond is broken, while the autocatalytic activation of trypsin is accompanied by the release of a small peptide from the N terminus of the zymogen (Davie and Neurath, 1955), and activation of procarboxypeptidase A is a complex process involving disaggregation of three subunits, only one of which gives rise to carboxypeptidase A (Neurath, 1964). Trypsin also activates pancreatic prephospholipase A_2 with the concomitant release of a heptapeptide from the N terminus (De Haas *et al.*, 1968). The isolation of two other pancreatic proenzymes, proelastase and a propeptidase, has been described by Hercz (1969). An example of activation by limited proteolysis in primitive organisms is found in group A streptococci. These organisms elaborate an extracellular zymogen that is transformed into an inactive enzyme by a limited proteolysis in which the molecular weight drops from 42,000 to 32,000. The inactive enzyme can then be activated by exposure to reducing agents (Liu and Elliott, 1965). There also exist enzyme precursors which are activated either by autocatalysis or by exposure to low pH: prorennin (Foltman, 1966), the gastric protein pepsinogen (Herriott, 1938), and the proproteinase C of baker's yeast (Hayashi *et al.*, 1968) are three examples. It is interesting to note that we have found no examples in the literature of proteolytic activation caused by hydrolysis of a peptide bond near the C terminus of a proenzyme. As was pointed out by Abita *et al.* (1969), such an activation mechanism is unlikely because protein biosynthesis proceeds from the N terminus of the molecule, and enzymatic activity could appear prior to completion of synthesis of such a proenzyme molecule.

Activation by limited proteolysis is not restricted to enzymes. The hormone insulin is also activated by partial proteolysis. Insulin is synthesized within the β cells of the pancreas as a single-chain precursor, proinsulin, which is activated, prior to secretion, by the splitting-out of a polypeptide connecting the C terminus of the B chain with the N terminus of the A chain (Steiner *et al.*, 1969).

Finally, two very complex examples of activation by limited proteolysis are the clotting of blood and the clotting of milk. Blood clotting involves activation of at least eight distinct factors in a cascade sequence that ultimately results in the conversion of soluble fibrinogen into an insoluble stabilized fibrin clot (Davie and Ratnoff, 1965). The clotting of milk is initiated by the conversion of prorennin into rennin followed by the rennin-catalyzed degradation of the κ-casein fraction. The degradation of this fraction allows the α_s-casein fraction to precipitate which, in turn, destabilizes the entire casein micelle (Ottesen, 1967).

The above examples obviously are not intended to cover all known cases, but merely to emphasize the complex and diverse nature of the phenomenon of activation by limited proteolysis. Comparison of the two structures, chymotrypsinogen and α-chymotrypsin, affords our first glimpse of the actual alteration of molecular architecture for one example of this general phenomenon.

Experimental Section

Preparation of Crystals and Heavy-Atom Derivatives. Bovine chymotrypsinogen A was purchased from Worthington Biochemical Corp. (chymotrypsinogen A, five-times crystallized) and from Calbiochem (α-chymotrypsinogen, A grade). Large type F crystals (Kraut *et al.*, 1962) suitable for X-ray intensity measurement were obtained from the former without further purification, but the latter was first

purified by repeated recrystallization from ammonium sulfate.

Type F crystals were grown by the following procedure. Chymotrypsinogen (150 mg) was dissolved in 0.5 ml of water, adjusted to pH 8–9 by addition of ammonium hydroxide, and desalted by passing through a Sephadex G-25 column (1 × 15 cm). The protein concentration of the eluent was reduced, when necessary, to 35 mg/ml by addition of water, and the pH was adjusted to 6.3. Next, the solution was cooled in an ice–salt bath and chilled 95% ethanol was added slowly, with constant stirring, until a faint turbidity appeared. An ethanol concentration of approximately 10% by volume was usually required. The solution was then allowed to warm up to room temperature, the turbidity disappeared, and 1 μl of 0.25 M diphenylcarbamyl fluoride in methanol was added to inhibit any contaminating chymotrypsin. Finally, 1-ml aliquots were placed in crystallizing dishes, the dishes were lightly seeded, sealed with glass cover slips, and allowed to stand in the dark at room temperature. Crystallization occurred within 1–7 days. The crystals obtained in this way were usually of sufficient size and quality to be used for collection of intensity data. A typical data crystal measured 0.2 × 0.4 × 0.5 mm and yielded precession photographs showing diffraction spots out to a Bragg spacing of 2 Å.

Whenever the crystals were transferred to a medium other than their own mother liquor, they would either severely crack or dissolve completely. Therefore, heavy-atom derivatives were made by adding the heavy-atom reagent directly to dishes in which the crystals had been grown. Two moles of reagent per mole of total protein was added by first withdrawing nearly all of the mother liquor, mixing the solution with the heavy-atom reagent, and pipetting the mixture back over the crystals. This procedure minimized crystal damage caused by local high concentrations of heavy-atom reagent. The crystals were allowed to soak anywhere from 7 to 30 days before being mounted for intensity measurement.

X-Ray Crystallographic Data. The space group, unit cell parameters, and crystal morphology are identical with those previously reported for chymotrypsinogen type F crystals (Kraut *et al.*, 1962): space group $P2_12_12_1$ with $a = 52.0$, $b = 63.9$, and $c = 77.1$ Å; there are four molecules per unit cell, one per asymmetric unit. The maximum apparent variation in cell parameters between different crystals of the parent protein was 0.6%. Within these limits, unit-cell parameters for crystals of the four heavy-atom derivatives used in the structure determination were indistinguishable from those for the parent crystals.

X-Ray intensity data were collected by parallel operation of a Hilger and Watts automatic diffractometer and screenless precession photography (Xuong *et al.*, 1968). This was desirable because, although screenless precession photography is very rapid, it is impractical to use it alone to measure the intensity of each and every unique reflection within any given limiting sphere of resolution. On the other hand, the diffractometer is able to examine every reciprocal lattice point, but is relatively slow because it can see only one reflection or, suitably modified, a very few reflections, at any one time. The data collection strategy was designed so that the same calendar time was required for measurement of the film and diffractometer portions of each data set. All intensity measurements were made with nickel-filtered Cu Kα radiation on crystals mounted in glass capillaries. Considerable difficulty was experienced during data collection because of unusually frequent movement of the crystals within the capillaries. As a result, more than 30% of the intensity measurements had to be discarded. In spite of this, we were able to obtain 225,000 usable reflection intensities in 6 months.

Intensity data films were exposed on Supper precession cameras with the layer line screen and holder removed, using two cameras on a single G. E. CA7H Cu target X-ray tube. All films were measured with a Tech-Ops rotating-drum scanning densitometer interfaced to an IBM 1800 computer (Xuong, 1969). The computer was programmed to index reflections, compensate for spot splitting, and discard reflections which were either obscured by other reflections or only partially recorded because of edge effects (Ng. H. Xuong, to be published). Integrated intensities were punched onto column-binary cards. Typically, 60% of all reflections recorded on a film yielded usable integrated intensities.

A complete three-dimensional set of 2.5-Å intensity data films was obtained for each crystal type, *i.e.*, parent and four heavy-atom derivatives. A set consisted of 32 exposures and required 8 crystals, 4 exposures/crystal. Exposure times were approximately 4 hr, the casette contained two films, and the precession angle was 2°. The films in each set were divided into four series. For the first, the crystal was mounted with the *b* axis along the camera spindle axis (horizontal on the film) and with the *c* axis vertical; the respective horizontal–vertical axes for the remaining three series were: *b*–*a*, *c*–*a*, and *c*–*b*. The exact initial orientation of each crystal was established by measurement of 30-min lineup pictures. In actual practice crystal alignment was much less critical than that required for normal precession photography because of the absence of a layer line screen. Within each series, eight exposures were made with the spindle dial incremented by 3, 6, 9, 12, 15, 18, 21, and 24°.

The quality of the film data was controlled by limiting crystal radiation damage and by monitoring crystal slippage. The susceptibility of chymotrypsinogen type F crystals to X-ray damage was established experimentally before any data films were exposed. Measurable radiation damage appeared only after 24 hr. Accordingly, crystals were discarded after four exposures (about 17 hr, total). For each crystal, a final lineup picture was also taken after the four data films had been exposed. Films were discarded whenever excessive crystal movement was revealed by comparison of the initial and final lineup pictures. For a typical data set, 32,000 reciprocal lattice points were measured with film.

Diffractometer data were collected with a step-scanning method (Wyckoff *et al.*, 1967). Reflections were scanned in 11 steps of 0.01° on ω and counts were accumulated for 2 or 3 sec per step. The intensity was taken to be proportional to the largest sum of four contiguous counts. Background corrections were applied from a curve of background scatter *vs.* 2θ which was empirically determined for each crystal. Empirical absorption corrections were also applied (Furnas, 1957). If the crystal moved excessively during data collection, both the background curve and the absorption correction curve were redetermined.

Three standard reflections were monitored automatically for all crystals after every 100 measurements, and, in addition, a few reflections which showed large intensity changes were periodically monitored for each of the heavy-atom derivatives.

TABLE 1: Statistics from the Last Cycle of Refinement.[a]

Derivative	Site No.	x/a	y/b	z/c	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}	Z	k	c	Root-Mean-Square f_H	Root-Mean-Square E	R_1	R_2
Mercuric bromide	1	0.58	0.37	0.03	22.1	23.5	9.2	6.4	0.6	7.7	88	0.935	−2.582	118.6	64.4	0.095	0.54
	2	0.10	0.13	0.00	25.6	13.5	2.0	−5.2	−2.9	0.7	29						
	3	0.54	0.39	0.04	17.6	7.4	5.7	−3.4	−4.0	−3.3	21						
	4	0.61	0.34	0.04	11.3	17.4	3.5	2.8	−2.5	−0.6	32						
	5	0.63	0.44	0.05	−8.9	34.6	0.9	21.4	4.6	6.3	8						
Uranyl fluoride	1	0.62	0.35	0.04	43.0	14.2	7.0	17.1	−6.1	−1.7	33	1.238	−0.062	84.2	60.5	0.093	0.57
	2	0.01	0.37	0.41	49.1	27.3	10.1	−5.7	2.8	1.5	58						
	3	0.58	0.38	0.03	36.5	38.1	6.8	5.7	6.3	9.9	34						
	4	0.96	0.37	0.01	71.4	50.8	32.5	8.6	2.6	14.4	24						
	5	0.64	0.43	0.05	1.3	68.4	15.4	−19.5	−6.2	20.2	18						
	6	0.68	0.36	0.05	36.2	13.7	11.2	5.7	11.1	−3.9	19						
Uranyl pyrophosphate	1	0.64	0.44	0.06	15.1	17.8	14.4	−5.5	3.3	2.9	66	1.102	0.172	96.1	57.0	0.090	0.55
	2	0.66	0.49	0.06	103.0	56.7	26.6	−14.7	11.3	−6.0	43						
	3	0.70	0.35	0.06	32.1	−0.8	5.6	−4.8	2.1	−0.8	17						
	4	0.78	0.37	0.02	17.9	−1.8	7.3	−2.5	−1.8	−2.2	16						
	5	0.67	0.42	0.04	32.9	11.2	19.2	−2.8	2.1	−7.2	29						
	6	0.74	0.44	0.00	25.5	7.4	6.5	1.0	1.6	6.0	16						
Mersalyl	1	0.97	0.02	0.48	41.1	23.1	14.3	−6.5	−4.2	−0.1	80	1.522	0.866	79.2	57.8	0.094	0.58
	2	0.40	0.30	0.25	25.5	21.1	19.0	2.8	0.7	−0.4	29						

[a] Mean figure of merit = 0.70. β as given here is defined by: $T = \exp\{-10^{-4}(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + 2\beta_{12}hk + 2\beta_{13}hl + 2\beta_{23}kl)\}$. Z is the occupancy in electrons. k is the scale factor to put the structure factor for a derivative on the same scale as the parent. c is the factor that compensates for any difference in the temperature factor between the derivative and the parent: $k' = k \exp c(\sin\theta/\lambda)^2$. f_H is the calculated heavy atom contribution to the structure factor. E is the closure error, $|k'F_H(\text{obsd}) - F_H(\text{calcd})|$. $R_1 = \Sigma|k'F_H(\text{obsd}) - F_H(\text{calcd})|/\Sigma k'F_H(\text{obsd})$ sums over all hkl. $R_2 = \Sigma||k'F_H(\text{obsd}) \mp F(\text{obsd})| - f_H|/\Sigma|k'F_H(\text{obsd}) \mp F(\text{obsd})|$ sums over centric zone reflections only.

A crystal was discarded if any of these decreased by more than 15%. From one to three crystals were needed to obtain a set of diffractometer data which, typically, consisted of 13,000 observations.

Intensities were scaled together by the method of Monahan *et al.* (1967) using all reflections of nonzero intensity that occurred in two or more different scaling groups. Since the scaling program required that these reflections be stored in the computer core memory, the maximum number of intensities which could be used for scaling a particular data set was approximately 24,000. If there were more than this number, only the 24,000 most intense scaling reflections were automatically selected for computation of scale constants. For the parent protein and each heavy-atom derivative, the film and diffractometer data were first scaled separately, the obvious errors were deleted, and finally, the two groups of data were combined, scaled, and edited. On the average, each unique intensity was measured from 2.3 to 4.0 times for the heavy-atom derivatives, and 5.1 times for the parent. Overall agreement of replicated intensity measurements was about 8%. The agreement was determined from the expression

$$\left\{\sum_{ijk}\frac{1}{\sigma_{ijk}^2}\left[I_{ijk} - K_j\langle I_i\rangle\right]^2 \Big/ \sum_{ijk}\frac{1}{\sigma_{ijk}^2}I_{ijk}^2\right\}^{1/2}$$

where the index k represents the kth observation on film j (or scale group j for diffractometer data) of reflection i, and where I_{ijk} and σ_{ijk} are, respectively, the intensity and estimated standard deviation of a particular observation. The symbol $\langle I_i\rangle$ represents the mean intensity for the corresponding reflection, and K_j represents the scale constant for film or diffractometer scale group j.

Approximately 88% of all unique reflection intensities within 2.5-Å resolution were obtained by screenless precession photography. The automatic diffractometer was used (1) to make at least two measurements of all intensities not recorded on films; (2) to make one measurement of all intensities recorded only once on films; and (3) to measure the intensities of 2000 scaling reflections distributed uniformly over all films in a data set.

Heavy-atom derivatives were evaluated for three-dimensional high-resolution data collection with the aid of 2.5-Å projection difference-Patterson and difference-Fourier maps. The projection intensity data were obtained by scanning conventional 21° precession photographs. The following heavy-atom derivatives produced interpretable maps: mercuric bromide, uranyl fluoride, uranyl pyrophosphate, mersalyl, mercuric idodide, potassium chloroplatinite, potassium chloroplatinate, iridium trichloride, and platinum tetranitrite. The first four heavy-atom derivatives in this list gave the cleanest difference-Patterson maps and so were selected for high-resolution phasing in three dimensions.

Phase Refinement. The intensities of Bijvoet-related reflections were averaged for phase calculation; no use was made of Bijvoet differences for phasing. Refinement of phases was carried out on a CDC 3600 computer by the usual method of minimizing $\Sigma[k'F_H(\text{obsd}) - F_H(\text{calcd})]^2$ in alternating cycles of phase calculation and parameter refinement (Dickerson *et al.*, 1968). Parameters refined in the last cycle, and their final values, are listed in Table I. Although the temperature factor matrices of four minor sites (mercuric bromide site 5, uranyl fluoride site 5, and uranyl pyrophosphate sites 3 and 4) are nonpositive definite, no attempt was made to correct them. This is not an unusual occurrence with sites of low occupancy, and experience has shown that little or no improvement of the final electron density map is achieved by slight adjustment of minor site parameters. Heavy-atom site occupancies in Table I are given in electrons. Conversion from an initially arbitrary occupancy scale into a scale approximately based on electrons was made on the assumption that the mersalyl site 1 was fully occupied. The centroid, or figure of merit, decreased monotonically from 0.87 for the ∞–5-Å shell to 0.55 for the 2.6–2.5-Å shell.

Two final checks were made on the quality of heavy-atom phasing: (1) phases computed using only the single major site of each derivative yielded difference-Fourier maps containing positive regions at all major and minor sites; (2) shell difference-Fourier maps computed for each heavy-atom derivative, using only terms between 2.8 and 2.5 Å, displayed significantly positive regions for all major sites.

Electron Density Map and Its Interpretation. The electron-density map was calculated on a 1.1 × 1.1 × 1.3 Å grid from centroid phases, but without centroid weighting factors. Forty contoured sections from $z = -{}^{7}/_{40}$ to ${}^{32}/_{40}$ were drawn by a CalComp plotter to a scale of 0.368 cm/Å, with the lowest contour at 0.25 e/Å^3 and subsequent contours at intervals of 0.25 e/Å^3. The contoured sections were then transferred onto 8.5 × 11 in. transparent plastic sheets with a Xerox 914 copier and the sheets were affixed to ${}^{3}/_{16}$-in. thick Plexiglass plates in order to produce vertical spacing on the same scale as that in the horizontal plane. The new map was centered about a complete zymogen molecule which had been picked out earlier at 5-Å resolution (Kraut *et al.*, 1962).

Interpretation of the 2.5-Å electron density map began at four readily recognizable regions: the helical segment at the C terminus, and three of the five disulfide bridges (residues 42–58, 136–201, and 168–182). We started matching the known sequence (Hartley, 1964; Hartley and Kauffman, 1966; Meloun *et al.*, 1966; Blow *et al.*, 1969) to the map at these regions and then followed the remainder of the backbone chain by checking especially prominent side chains. We finally constructed a Kendrew–Watson skeletal model of the complete molecule on a scale of 2 cm/Å. Because of uncertainties in the interpretation of the map, the positions of some atoms in the backbone chain may be in error by as much as 1.5 Å, and the conformation of two segments of chain, residues 8–13 and 149–153, is still tentative. The chymotrypsinogen model was built independently (in so far as was humanly possible) from a model of α-chymotrypsin, which we simultaneously assembled from coordinates supplied by the Cambridge group (Birktoft *et al.*, 1969).

Results

Description of the Chymotrypsinogen Molecule and Comparison with α-Chymotrypsin. The following description of the chymotrypsinogen molecule is based on a Kendrew–Watson skeletal model which, in turn, represents our present interpretation of the electron density map. Comparison of the chymotrypsinogen and α-chymotrypsin structures was performed by examining this model and a similar model of α-chymotrypsin assembled from the coordinates of Birktoft *et al.* (1969). The two electron density maps have not yet

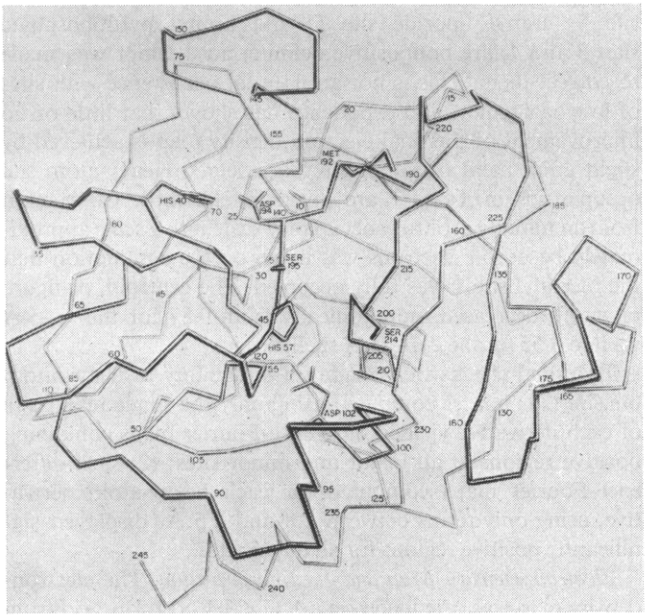

FIGURE 1: Chymotrypsinogen simplified backbone chain linking the α-carbon atoms of each amino acid residue, drawn from a vantage point looking toward the latent catalytic site region. The side chains of some important amino acid residues are also shown. The drawing was prepared from a computer-plotted perspective projection of all α-carbon atoms.

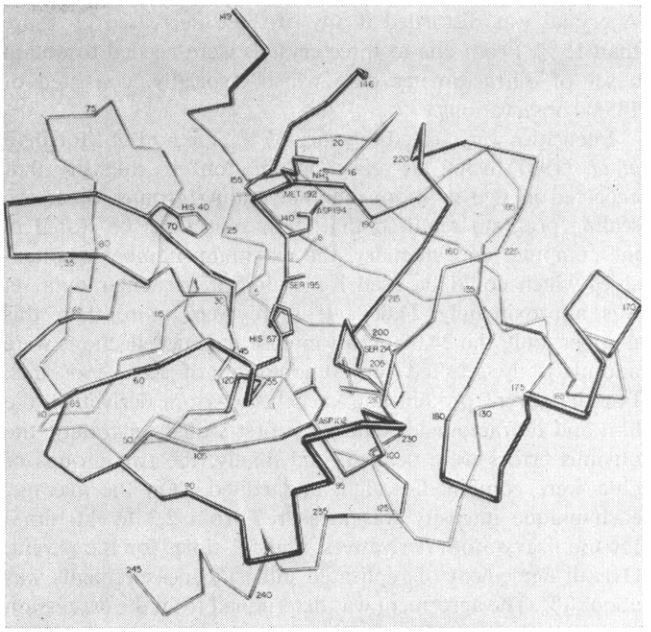

FIGURE 2: α-Chymotrypsin simplified backbone chain linking the α-carbon atoms of each amino acid residue, drawn from a vantage point equivalent to that of Figure 1. The side chains of some important amino acid residues are also shown. This drawing and all subsequent α-chymotrypsin drawings were prepared from the coordinates of tosyl-α-chymotrypsin. Accordingly, the side chain of Ser-195 is shown in a position appropriate to the inhibited enzyme. In native α-chymotrypsin this side chain is rotated slightly about its α–β bond and assumes a position close to that shown for the zymogen in Figure 1. The drawing was prepared from a computer-plotted perspective projection of the α-carbon atoms. A portion of the chain, from residues 9 to 13, is not shown because of the uncertainty in position of these residues (Birktoft *et al.*, 1969).

been compared directly. A more detailed interpretation of the chymotrypsinogen electron density map and a list of atomic coordinates will be published after further refinement. It should be clearly understood that the results reported here are therefore preliminary, but are unlikely to be appreciably altered by further work now in progress.

The overall folding of the polypeptide chain of chymotrypsinogen (Figure 1) is very similar to that of α-chymotrypsin (Figure 2). The general description of the conformation of α-chymotrypsin reported by the Cambridge group (Matthews *et al.*, 1967; Sigler *et al.*, 1968; Blow, 1969) is therefore applicable to the zymogen as well. With the exception of a ten-residue segment of α helix at the C terminus, which is identical both in zymogen and enzyme, both molecules are composed almost entirely of more or less fully extended polypeptide chain which often folds back on itself to form large sections of distorted antiparallel pleated sheet (Blow, 1969). The additional segment of α helix involving residues 161–173 of α-chymotrypsin (Blow, 1969) is poorly formed in the zymogen. As was expected, the two dipeptides Ser-14–Arg-15 and Thr-147–Asn-148, which are split out upon conversion into α-chymotrypsin, lie on the surface of the zymogen and are therefore easily accessible to enzymatic attack.

In order to make comparison of the two proteins as objective as possible, the atomic coordinates of α-chymotrypsin (Birktoft *et al.*, 1969) were first approximately transformed into the crystallogrphic coordinate system of chymotrypsinogen, and a computer program was written to refine the transformation by minimizing the sum of squares of distances between α-carbon atoms of identical residues in the two molecules. The final transformation was determined in two steps: first, coordinates for all available α-carbon atoms were used (coordinates were not available for residues 9–13 of α-chymotrypsin); and second, all α-carbon atoms which had moved by more than 3 Å (32 atoms in all) were eliminated from the minimization process. It is noteworthy that, although some α-carbon atoms move by as much as 11 Å upon activation, the second step of the transformation refinement produced very little change in the rotation matrix and translation vector. How little the two molecules differ in conformation can be seen by noting that the mean displacement between *all* equivalent α-carbon atoms is only 1.8 Å; the most constant region is the C-terminal α helix, where the mean displacement is 0.6 Å. A list of α-carbon atom displacements is given in Table II for all α-carbon atoms which differ in position by more than 3.6 Å between the two structures.

There are six segments of backbone chain which exhibit relatively large conformational differences between the zymogen and enzyme. The segments are composed of from two to seven residues and have been arbitrarily labeled with Roman numerals starting from the N terminus. These conformational differences are of two general types: (1) those which involve only the repositioning of a segment of chain on the surface of the molecule (segments I, III, IV, and V); and (2) those in which residues move either from the surface to the interior or from the interior to the surface of the molecule (segments II and VI). It should be noted that some surface-to-surface differences may not be essential to the activa-

tion process since they may occur in flexible regions of the chain that are susceptible to alteration by crystal packing forces. Specifically, segment IV, residues 72–77, where the α-carbon atoms differ by 5–10 Å between zymogen and enzyme, is known to assume different conformations in the two independent molecules of the α-chymotrypsin asymmetric unit (Birktoft *et al.*, 1969). However, one of the surface-to-surface differences (segment V) and both of the interior-to-surface differences (segments II and VI) are probably involved with the activation process. These "significant" conformational changes are described below.

The changes in segment II involve the repositioning of residues Ile-16–Val-17, and result in the formation of the internal ion pair between the α-amino group of Ile-16 and the buried side-chain carboxyl group of Asp-194. The residues involved in the segment II changes are shown in Figures 3 and 4. Referring to these figures, residues 16 and 17 of the zymogen are above residues 18–21, while in the enzyme the two residues are below the 18–21 segment. Apparently, upon tryptic scission of the Arg-15–Ile-16 peptide bond, the N-terminal residues of the newly formed B chain swing out through the solvent as the main chain rotates 180° about an axis roughly defined by the α-carbon atoms of residues 19 and 21. In this process, the formerly exposed hydrophobic side chains of Ile-16 and Val-17 are buried beneath the surface of the molecule and the N terminus of Ile-16 approaches the side-chain carboxyl group of Asp-194. The side-chain carboxyl group of Asp-194 is buried in chymotrypsinogen as well as in α-chymotrypsin, but in the zymogen it is hydrogen bonded to the Nε2[1] of His-40. During activation this group undergoes an approximately 4-Å change in position toward the new position of Ile-16. The finding that Asp-194 is buried in the zymogen is in agreement with the chemical modification studies of Carraway *et al.* (1969) and Abita and Lazdunski (1969).

The most dramatic result of the segment V conformational changes is the change in position of the Arg-145 side chain. The changes associated with segment V may best be visualized from examination of Figures 1, 2, 5, and 6: the main-chain repositioning is clearly evident from Figures 1 and 2, while the change in position of the Arg-145 side chain can be seen in Figures 5 and 6. In the zymogen, the guanidinium group of Arg-145 may be close enough to the buried side-chain carboxyl of Asp-194 to permit appreciable electrostatic interaction between these two charged groups. In the enzyme, on the other hand, the backbone carrying Arg-145 has moved by 9 Å (segment V) and the side chain of Arg-145 has swung up out of the vicinity of the specificity site to become fully extended into the solvent. We must add a cautionary note regarding the position of the Arg-145 side chain in the zymogen, since it is also possible, with some modification to the main chain, to construct the chymotrypsinogen model with this side chain extending out into the solvent. However, the electron density for this alternate position is much weaker than that at the present position, whereas the map is very clear for the side chains of the other three arginine residues. It is possible that the structure is disordered in this region with the Arg-145 side chain occupying the first position in some molecules and the alternate position in others; in the alternate position, however, the guanidinium group could not electrostatically interact with Asp-194.

TABLE II: List of All Residues for Which α-Carbon Atoms Differ in Position by More Than 3.6 Å between the Chymotrypsinogen and α-Chymotrypsin Structures.[a]

Segment	Residue	Displacement (Å)
I	Gln-7	4.8
	Pro-8	10.0*
II	Ile-16	11.3
	Val-17	6.6
III	Thr-37	4.0
	Gly-38	6.6
IV	Asp-72	5.6
	Gln-73	9.6
	Gly-74	9.1
	Ser-75	6.2
	Ser-76	10.1
	Ser-77	5.6
V	Thr-144	5.9
	Arg-145	8.7
	Tyr-146	4.6
	Ala-149	4.7
	Asn-150	6.7*
	Thr-151	4.7*
	Pro-152	4.6*
VI	Met-192	8.4
	Gly-193	6.6

[a] Chymotrypsinogen coordinates for residues marked * are only tentative.

The segment VI changes involve the relocation of Met-192 from a position buried within the zymogen molecule out to the surface of the enzyme molecule. This 8-Å migration, although apparent in all of the figures, is most evident from comparison of Figures 5 and 6. That the Met-192 side chain moves from the molecular interior out to the surface, however, is best seen in Figures 7 and 8. In a somewhat oversimplified sense, the segment VI changes can be thought of as resulting from a 180° rotation of the main chain about the carbonyl carbon to α-carbon bond of Asp-194. This rotation also causes the side chain of Asp-194 to move (toward the top of the page in Figure 3) and its side-chain carboxyl to shift position from the vicinity of His-40 to the vicinity of Ile-16. During this process the imidazole ring of His-40 rotates slightly. An interesting change in hydrogen bonding accompanies these contortions. The original hydrogen bond between the side-chain carboxyl of Asp-194 and the Nε2 of His-40 in the chymotrypsinogen molecule is broken during activation, and the Asp-194 side-chain carboxyl group then forms the internal ion pair with the α-amino group of Ile-16 previously alluded to, while the Nε2 of His-40 forms a hydrogen bond with the carbonyl oxygen of Gly-193.

Obviously, the classification of the activation conformational changes according to various segments is only a convenience to aid in their description. The burying of the N

[1] Notation and numbering of amino acid side chains is that given by Edsall *et al.* (1966).

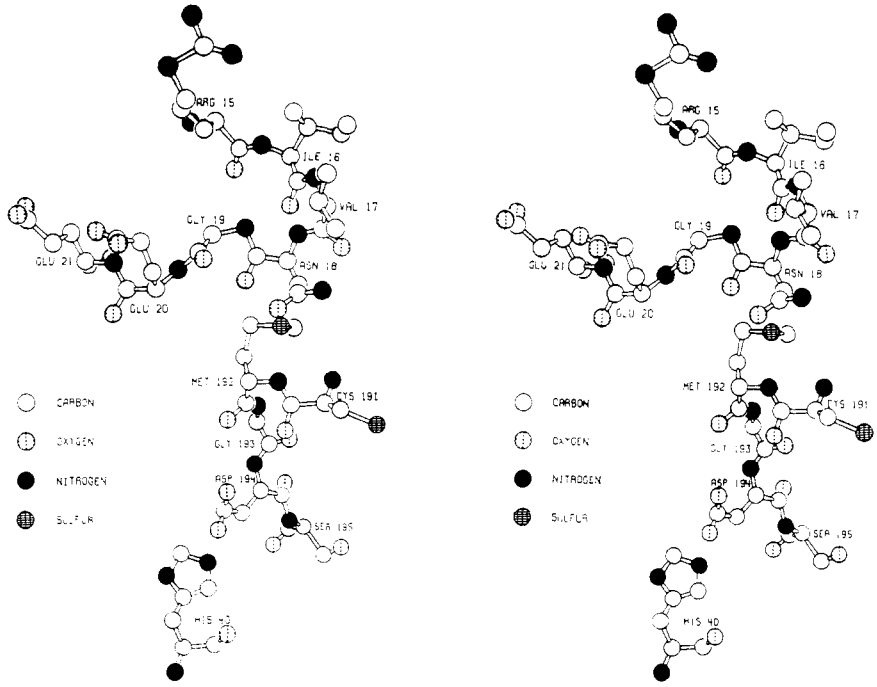

FIGURE 3: Chymotrypsinogen stereoscopic projection of amino acid residues involved in the formation of the Ile-16–Asp-194 ion pair during activation. The figure is drawn from a vantage point above the surface looking toward the interior of the molecule (from top to bottom on Figure 1). Figures 3–8 were prepared from computer-plotted stereoscopic projections.

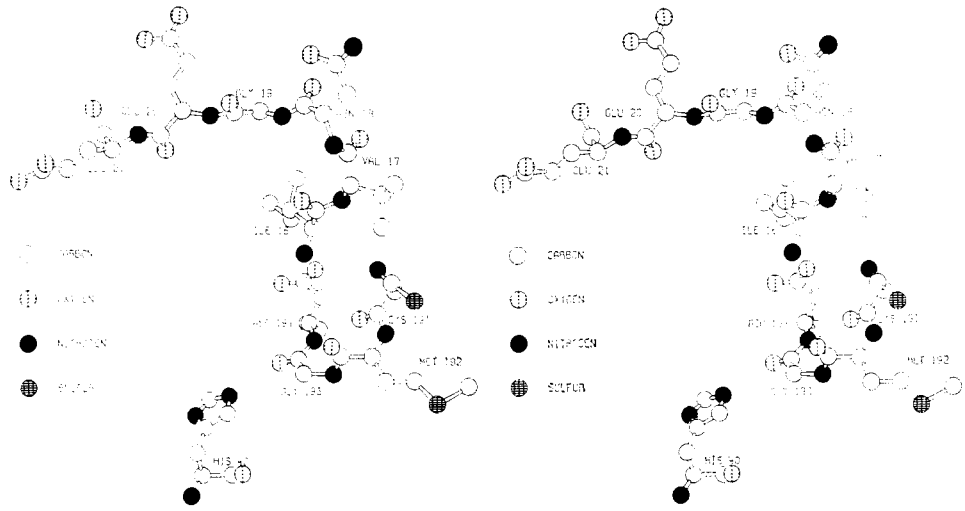

FIGURE 4: α-Chymotrypsin stereoscopic projection of amino acid residues involved in the formation of the Ile-16–Asp-194 ion pair during activation. The figure is drawn from a vantage point equivalent to that of Figure 3.

terminus of the newly formed B chain, the rotation of the Asp-194 and His-40 side chains, the surfacing of Met-192, the positional shift of Arg-145 and all of the other more subtle changes which accompany activation of the zymogen clearly occur in concert. To assign cause and effect relationships among these molecular reorganizations would require a much more detailed balancing of the energy changes involved than our present knowledge permits.

Discussion

The Activation Process. The question arises as to whether all of the conformational differences we observe between chymotrypsinogen and α-chymotrypsin are significant to the activation process or are merely nonrelated changes accumulated during the progressive steps in the conversion of π-chymotrypsin into α-chymotrypsin. We conclude from evidence summarized below, that all forms of the active enzyme are essentially identical in structure, and hence that the conformational changes of the polypeptide backbone segments listed in Table II, with the possible exception of the flexible segment IV, must take place when the zymogen is initially converted into the π form of the enzyme by tryptic hydrolysis of the Arg-15–Ile-16 peptide bond. Three crystal-

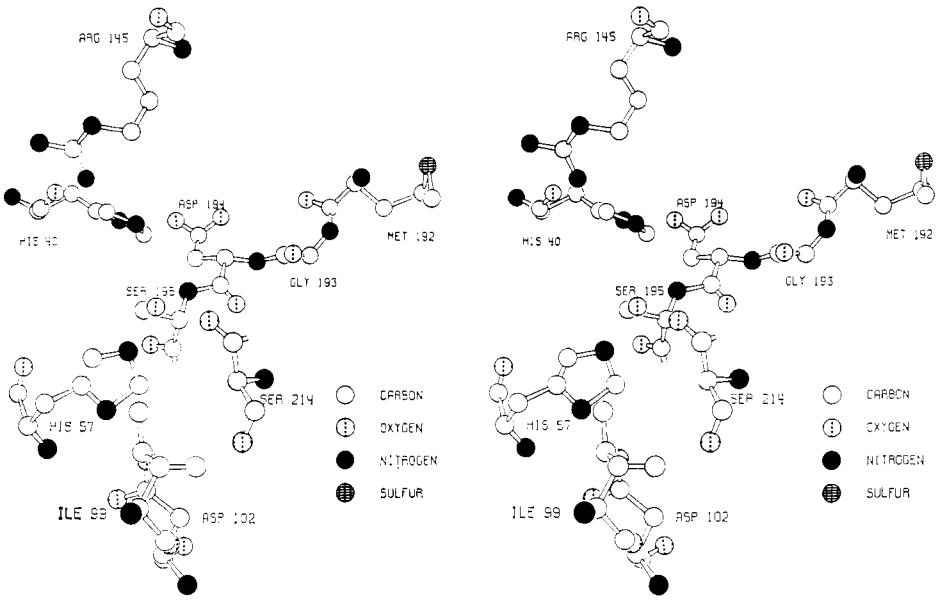

FIGURE 5: Chymotrypsinogen stereoscopic projection of the latent catalytic site region drawn from a vantage point looking into the page (see Figure 1) at an angle inclined about 20° to the right of perpendicular.

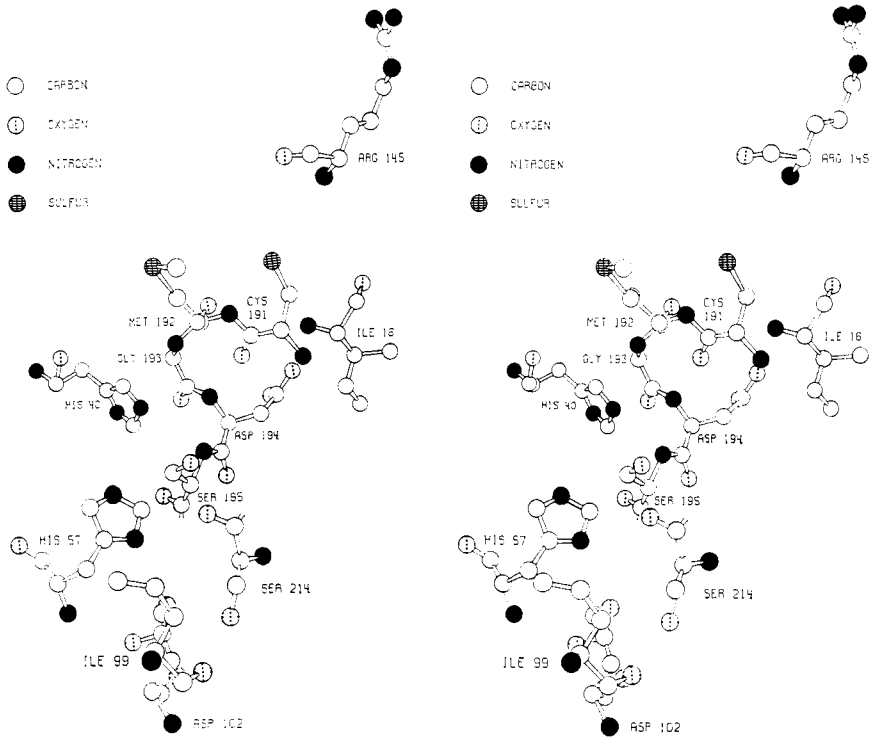

FIGURE 6: α-Chymotrypsin stereoscopic projection of the catalytic site region drawn from a vantage point equivalent to that of Figure 5.

lographic studies support this conclusion: first, 5-Å difference-Fourier maps calculated with δ-chymotrypsin phases and coefficients $F(\pi) - F(\delta)$ and $F(\delta) - F(\gamma)$ showed positive regions only in the vicinity of the autocatalytically excised dipeptides (Kraut *et al.*, 1967); second, comparison of the 5-Å structure of δ-chymotrypsin with the 2-Å structure of α-chymotrypsin failed to reveal any significant differences other than excision of the Thr-147–Asn-148 dipeptide (Wright *et al.*, 1968); and third, comparison of the 5.5- and 2.7-Å structures of γ-chymotrypsin with the 2-Å structure of α-chymotrypsin indicates that these molecules are essentially identical (Matthews *et al.*, 1968; Cohen *et al.*, 1969; Davies *et al.*, 1969). It should be emphasized, however, that except for α-chymotrypsin, only low-resolution structures or pre-

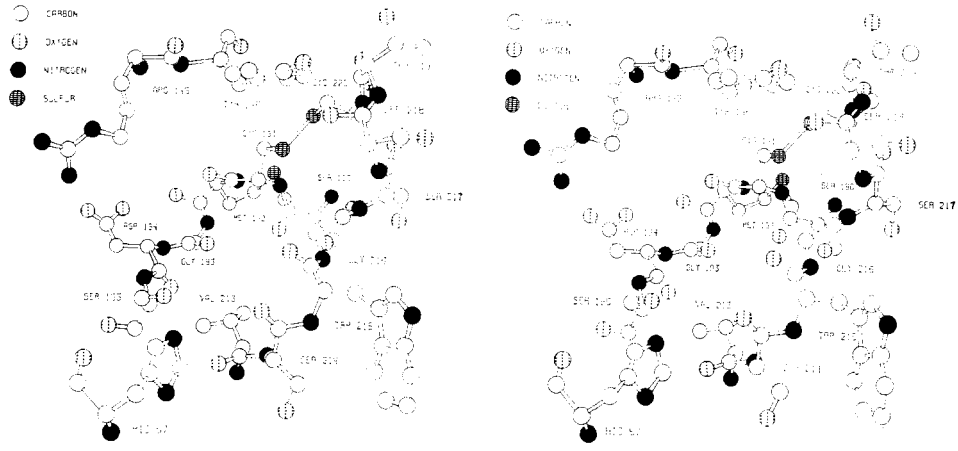

FIGURE 7: Chymotrypsinogen stereoscopic projection of the latent specificity cavity drawn from roughly the same vantage point as Figure 1.

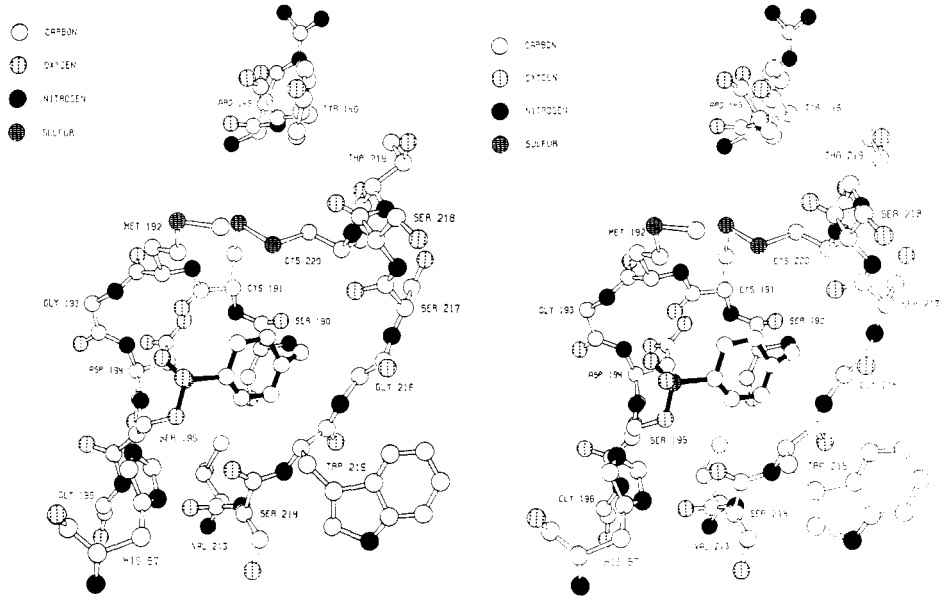

FIGURE 8: Tosyl-α-chymotrypsin stereoscopic projection of the specificity cavity drawn from a vantage point equivalent to that of Figure 7. The tosyl group is drawn with solid bonds to contrast with the bonds of the amino acids forming the specificity cavity in the enzyme.

liminary results were used to make these comparisons. The comparisons do suggest, though, that once the enzyme is formed, its conformation remains essentially rigid and that the π, δ, γ, and α forms of the enzyme are almost identical in structure. Accordingly, we take this view in the following discussion and assume that at least certain of the conformation differences we observe between chymotrypsinogen and α-chymotrypsin represent significant events in the genesis of enzyme activity. We shall discuss only changes for which such a role seems plausible, but clearly the possibility does exist that there are additional significant changes which are not conspicuous from comparison of our present chymotrypsinogen model with α-chymotrypsin.

Catalytic Site Conformational Changes. The catalytic activity of the chymotrypsin family of serine proteases has been shown by direct chemical evidence to depend on the participation of a serine (Balls and Wood, 1956; Schaffer *et al.*, 1957; Oosterbaan and van Adrichem, 1958) and a histidine (Ong *et al.*, 1965) residue. Based on the three-dimensional structure of α-chymotrypsin, Blow *et al.* (1969) have proposed a catalytic mechanism which involves a "charge relay system" composed of a hydrogen-bond network between the side chains of Ser-195, His-57, and Asp-102. It was, therefore, very surprising to find that the spatial arrangement of these three residues in the zymogen is, to within the limits of our ability to interpret the electron density map, extremely similar to that in the catalytically active enzyme. Indeed, except for an apparent slight rotation of the plane of the imidazole ring of His 57 (compare Figures 1 with 2, 6 with 7, or 8 with 9), these residues are indistinguishable in the zymogen and enzyme at our present resolution. Thus, the structural similarity of zymogen and enzyme extends, not only to their

overall molecular conformaiton, but even to the arrangement of side-chain residues at the catalytic site. It would appear, then, that the "switching-on" of catalytic activity during activation of the zymogen must involve only small structural modification of this region of the molecule.

In addition to the apparent slight rotation of the His-57 imidazole ring, two other small changes are observable in the neighborhood of the charge relay system: (1) the side chain of Ile-99 moves away from the imidazole ring of His-57, changing the environment of the charge relay system in such a way as to permit limited access to the surrounding aqueous environment; and (2) the hydroxyl group of Ser-214 apparently moves closer to the side-chain carboxyl of Asp-102 enabling a second hydrogen bond to be formed to the Oδ2 of Asp-102. Whether these subtle changes contribute toward bringing about enzymatic activity is not clear, but it would perhaps be unwise to regard them as merly coincidental to the activation process. We will, therefore, consider the possible significance of these two changes in more detail.

The first conformational change in the vicinity of the catalytic site, the displacement of the Ile-99 side chain, is very noticeable. In the zymogen the Ile-99 side chain is in van der Waals contact with the imidazole ring of His-57 and, in conjunction with neighboring side chains, completely blocks the access of solvent to the His-57–Asp-102 portion of the catalytic site hydrogen-bond network. During activation, the Ile-99 side chain rotates approximately 90° about its α-carbon to β-carbon bond and moves away from the His-57 ring into a position more exposed to the solvent (compare Figures 5 and 6). Curiously, the conformation of the main chain in this vicinity does not change dramatically. Finally, it may be noted that conservation of a hydrophobic side chain at the Ile-99 position (either Ile, Leu, or Val) in five closely related proteins—bovine trypsinogen, bovine chymotrypsinogens A and B, porcine elastase (Dayhoff, 1969), and bovine thrombin (Magnusson, 1969)—is also consistent with the possibility that this side chain plays some role in the activation process.

The second small conformational change in the neighborhood of the catalytic site, the movement of Ser-214, is much more subtle and, indeed, may not even exist. However, in view of the possible role of this second serine in the catalytic site hydrogen-bond network, which will be discussed below, it is intriguing to note that the hydrogen bond between Oγ of Ser-214 and Oδ2 of Asp-102 is apparently the only bond in the catalytic site hydrogen-bond network that is formed during activation. In our present model of the zymogen, the other two bonds of the network are preformed, whereas the distance between the potential donor and acceptor of the third hydrogen bond is too great, about 4 Å, for the bond to exist (compare Figures 5 and 6). It must be emphasized that, strictly speaking, this distance is actually not known to better than perhaps 1.5 Å, but the obvious change in conformation of the segment of main chain bearing Ser-214 also suggests that this hydrogen bond is formed during activation: segment 214–217 of the zymogen main chain is slightly folded in such a way as to restrict the approach of Ser-214 Oγ to the Oδ2 of Asp-102, but, in the enzyme, this segment is fully extended and the two oxygen atoms approach to within 2.6 Å of each other.

There are several observations suggesting that a second serine (Ser-214 in chymotrypsin) is involved in the activity of serine proteases generally. (1) The hydrogen-bond network at the catalytic site of α-chymotrypsin includes a hydrogen bond between the side chains of Ser-214 and Asp-102, in addition to the hydrogen bonds involving Ser-195, His-57, and Asp-102. (2) Even though serine is the most mutationally variable amino acid (Dayhoff *et al.*, 1969), Ser-214 is invariant in the sequence of five related proteolytic enzymes—bovine trypsinogen, bovine chymotrypsinogens A and B, porcine elastase (Dayhoff, 1969), and bovine thrombin (Magnusson, 1969). In the case of elastase, serine is the only invariant residue in the immediate vicinity. (3) A nearly identical hydrogen-bond network of active serine, histidine, aspartic acid, and second serine is also present at the catalytic site of the bacterial protease subtilisin BPN′ (Alden *et al.*, 1970), and the related protease subtilisin Carlsberg contains all components of this hydrogen-bond network, except that a threonine residue is substituted for the second serine. The utter dissimilarity in conformation and amino acid sequence between chymotrypsin and subtilisin implies that these proteins represent an example of the convergent evolution of a functionally identical complex hydrogen-bond network. These observations, taken together, strongly suggest that all three hydrogen bonds in the network play a significant role in the enzymatic activity of these proteins.

In light of the proposal of Steitz *et al.* (1969) that the carbonyl oxygen of Ser-214 is involved in productive binding of substrates, it might be thought that the sole function of the second serine hydrogen bond is to maintain this carbonyl group in the required orientation. This seems unlikely, however, because in subtilisin the orientation of the carbonyl oxygen of the corresponding second serine is such that it could not be involved in substrate binding.

Specificity Cavity Conformational Changes. Near the active serine, Ser-195, the main chain and the side chains of several residues of the enzyme form the "specificity cavity," a cavity that holds substrate side chains during enzymatic hydrolysis. When the zymogen is activated, the changes occurring at this portion of the molecule are much more pronounced than the slight molecular rearrangement at the catalytic site noted above. A relatively large movement of the backbone chain from residue 187 to residue 194, including region VI noted in Table II, results in an appreciable conformational change in this area, with the formation of a definite cavity which exists only partially in the zymogen. An especially dramatic result of this rearrangement is the migration of Met-192 from a completely buried position in the interior of the zymogen out to the surface of the enzyme molecule. This process may be visualized by comparing Figures 7 and 8. As may be seen in these pictures, Met-192 forms the lid of the specificity cavity in the enzyme; this portion of the cavity does not exist in the zymogen. The Cambridge group established that this cavity was involved in substrate binding by showing, in a 2.5-Å difference-Fourier map, that the side chain of the virtual substrate *N*-formyl-L-tryptophan is bound there (Steitz *et al.*, 1969). Further, the tryptophan residue is bound in the same general area of the enzyme as the tosyl group shown in Figure 8. No change in enzyme conformation accompanies the binding of formyl-L-tryptophan, but when dioxan is bound to the enzyme, again in this same cavity, there is a slight repositioning of Met-192 (Steitz *et al.*, 1969). The significance of Met-192 in substrate binding was demon-

strated earlier by chemical modification studies (Koshland *et al.*, 1962; Knowles, 1965; Lawson and Schramm, 1965).

Another activation change which may be significant to substrate binding is the displacement of Arg-145. As was mentioned above, in our present chymotrypsinogen model the side chain of Arg-145 is in a position where it could, possibly, stabilize the charge on the buried Asp-194 side-chain carboxyl group. One of the major activation changes is the repositioning of the guanidinium group of this residue some 15 Å away from its original position in the zymogen. Perhaps this movement helps to drive the side chain of Asp-194 away from His-40 and toward the NH_3^+ of Ile-16, which in turn results in the specificity cavity being completed in the manner already described. The conservation of a basic residue, either arginine or lysine, at the Arg-145 position in bovine trypsinogen, bovine chymotrypsinogens A and B, porcine elastase (Dayhoff, 1969), and bovine thrombin (Magnusson, 1969) is consistent with these conjectures.

Obviously, at least as many questions are raised by these observations as are answered. What detailed balance of driving forces is responsible for the conformation effects induced by the splitting of a single peptide bond in the zymogen? Is the absence of the specificity cavity in the zymogen an adequate explanation for its inactivity? If not, can the inactivity of the zymogen be accounted for on the basis of the small alterations already observed in the arrangement of side chains within and in the neighborhood of the catalytic site, or would improved electron density maps reveal still more subtle but essential changes in structure. What is the role of the "second serine" in the enzyme's activity...., and so on. If nothing else, such questions serve to emphasize the rudimentary state of our present understanding of how chymotrypsin in particular, and enzymes in general, do their work.

Acknowledgments

We are indebted to Dr. Richard A. Alden for assistance with all phases of this work, Dr. Jens Birktoft for helpful and critical discussions, Dr. Oliver Seely for help with crystal preparation and precession photography, Mr. David Fehr and Mrs. Sheryl Glasser for assistance with construction of the models, and Mr. Richard Buller for performing calculations associated with production of the figures and transformation of the α-chymotrypsin coordinates.

References

Abita, J. P., Delaage, M., Lazdunski, M., and Savrda, J. (1969), *European J. Biochem. 8*, 314.

Abita, J. P., and Lazdunski, M. (1969), *Biochem. Biophys. Res. Commun. 35*, 707.

Alden, R. A., Wright, C. S., and Kraut, J. (1970), *Phil. Trans. Roy. Soc. (London) B257*, 119.

Balls, A. K., and Wood, H. N. (1956), *J. Biol. Chem. 219*, 245.

Birktoft, J. J., Matthews, B. W., and Blow, D. M. (1969), *Biochem. Biophys. Res. Commun. 36*, 131.

Blow, D. M. (1969), *Biochem. J. 112*, 261.

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

Brown, J. R., Greenshields, R. N., Yamasaki, M., and Neurath, H. (1963), *Biochemistry 2*, 867.

Brown, K. D., Shupe, R. E., and Laskowski, M. (1948), *J. Biol. Chem. 173*, 99.

Carraway, K. L., Spoerl, P., and Koshland, D. E. (1969), *J. Mol. Biol. 42*, 133.

Charles, M., Gratecos, D., Rovery, M., and Densuelle, P. (1967), *Biochim. Biophys. Acta 140*, 395.

Cohen, G. H., Silverton, E. W., Matthews, B. W., Braxton, H., and Davies, D. R. (1969), *J. Mol. Biol. 44*, 129.

Davie, E. W., and Neurath, H. (1955), *J. Biol. Chem. 212*, 515.

Davie, E. W., and Ratnoff, O. D. (1965), *Proteins 3*, 359.

Davies, D. R., Cohen, G. H., Silverton, E. W., Braxton, H. P., and Matthews, B. W. (1969), *Acta Cryst. A25*, S182.

Dayhoff, M. O. (1969), *Atlas Protein Sequence Struct. 4*, D224.

Dayhoff, M. O., Eck, R. V., and Park, C. M. (1969), *Atlas Protein Sequence Struct. 4*, 75.

De Haas, G. H., Postema, N. M., Nieuwenhuizen, W., and Van Deenen, L. L. M. (1968), *Biochim. Biophys. Acta 159*, 118.

Dickerson, R. E., Weinzierl, J. E., and Palmer, R. A. (1968), *Acta Cryst. B24*, 997.

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

Edsall, J. T., Flory, P. J., Kendrew, J. C., Liquori, A. M., Nemethy, G., Ramachandran, G. N., and Scheraga, H. A. (1966), *J. Mol. Biol. 15*, 399.

Folk, J. E., and Schirmer, E. W. (1965), *J. Biol. Chem. 240*, 181.

Foltmann, B. (1966), *Compt. Rend. Trav. Lab. Carlsberg 35*, 143.

Furnas, T. C. (1957), Single Crystal Orienter Instruction Manual, Milwaukee, Wis., General Electric Co., p 71.

Gratecos, D., Guy, O., Rovery, M., and Desnuelle, P. (1969), *Biochim. Biophys. Acta 175*, 82.

Hartley, B. S. (1964), *Nature 201*, 1284.

Hartley, B. S., and Kauffman, D. L. (1966), *Biochem. J. 101*, 229.

Hayashi, R., Oka, Y., Doi, E., and Hata, T. (1968), *Agr. Biol. Chem. 32*, 367.

Hercz, A. (1969), *J. Biol. Chem. 244*, 5556.

Herriott, R. M. (1938), *J. Gen. Physiol. 22*, 65.

Hess, G. P. (1968), *Brookhaven Symp. Biol. 21*, 155.

Jacobsen, C. F. (1947), *Compt. Rend. Trav. Lab. Carlsberg 25*, 325.

Kunitz, M. (1938), *J. Gen. Physiol. 22*, 207.

Kunitz, M., and Northrop, J. H. (1935), *J. Gen. Physiol. 18*, 433.

Knowles, J. R. (1965), *Biochem. J. 95*, 180.

Koshland, D. E., Strumeyer, D. H., and Ray, W. J., Jr. (1962), *Brookhaven Symp. Biol. 15*, 101.

Kraut, J., High, D. F., and Sieker, L. C. (1964), *Proc. Natl. Acad. Sci. U. S. 51*, 839.

Kraut, J., Sieker, L. C., High, D. F., and Freer, S. T. (1962), *Proc. Natl. Acad. Sci. U. S. 48*, 1417.

Kraut, J., Wright, H. T., Kellerman, M., and Freer, S. T. (1967), *Proc. Natl. Acad. Sci. U. S. 58*, 304.

Lawson, W. B., and Schramm, H. J. (1965), *Biochemistry 4*, 377.

Liu, T., and Elliott, S. D. (1965), *J. Biol. Chem. 240*, 1138.

Magnusson, S. (1969), *Proc. Symp. Struct. Function Relationships Proteolytic Enzymes, Copenhagen* (in press).

Matthews, B. W., Cohen, G. H., Silverton, E. W., Braxton, H.. and Davies, D. R. (1968), *J. Mol. Biol. 36*, 179.

Matthews, B. W., Sigler, P. B., Henderson, R., and Blow, D. M. (1967), *Nature 214*, 652.
Meloun, B., Kluh, I., Kostka, V., Moravek, L., Prusik, Z., Vanecek, J., Keil, B., and Sorm, F. (1966), *Biochim. Biophys. Acta 130*, 543.
Monahan, J. E., Schiffer, M., and Schiffer, J. P. (1967), *Acta Cryst. 22*, 322.
Neurath, H. (1957), *Advan. Protein Chem. 12*, 319.
Neurath, H. (1964), *Federation Proc. 23*, 1.
Ong, E. B., Shaw, E., and Schoellmann, G. (1965), *J. Biol. Chem. 240*, 694.
Oosterbaan, R. A., and van Adrichem, M. E. (1958), *Biochim. Biophys. Acta 27*, 423.
Oppenheimer, H. L., Labouesse, B., and Hess, G. P. (1966), *J. Biol. Chem. 241*, 2720.
Ottesen, M. (1967), *Ann. Rev. Biochem. 36*, 55.
Peanasky, R. J., Gratecos, D., Baratti, J., and Rovery, M. (1969), *Biochim. Biophys. Acta 181*, 82.
Rovery, M., Curnier, A., and Desnuelle, P. (1955), *Biochim. Biophys. Acta 17*, 565.
Rovery, M., Poilroux, M., Yoshida, A., and Desnuelle, P. (1957), *Biochim. Biophys. Acta 23*, 608.
Schaffer, N. K., Simet, L., Harshman, S., Engle, R. R., and Drisko, R. W. (1957), *J. Biol. Chem. 225*, 197.
Sigler, P. B., Blow, D. M., Matthews, B. W., and Henderson, R. (1968), *J. Mol. Biol. 35*, 143.
Smillie, L. B., Furka, A., Nagabhushan, N., Stevenson, K. J., and Parkes, C. O. (1968), *Nature 218*, 343.
Steiner, D. F., Clark, J. L., Nolan, C., Rubenstein, A. H., Margoliash, E., Aten, B., and Oyer, P. E. (1969), *Recent Progr. Hormone Res. 25*, 207.
Steitz, T. A., Henderson, R., and Blow, D. M. (1969), *J. Mol. Biol. 46*, 337.
Wright, H. T., Kraut, J., and Wilcox, P. E. (1968), *J. Mol. Biol. 37*, 363.
Wyckoff, H. W., Doscher, M., Tsernoglou, D., Inagami, T., Johnson, L. N., Hardman, K. D., Allewell, N. M. Kelly, D. M., and Richards, F. M. (1967), *J. Mol. Biol. 27*, 563.
Xuong, Ng. H. (1969), *J. Sci. Instrum. 2*, 485.
Xuong, Ng. H., Kraut, J., Seely, O., Freer, S. T., and Wright, C. S. (1968), *Acta Cryst. B24*, 289.

Environment of Copper in *Pseudomonas fluorescens* Azurin: Fluorometric Approach*

Alessandro Finazzi-Agrò, Giuseppe Rotilio, Luciana Avigliano, Pietro Guerrieri, Vittoriano Boffi, and Bruno Mondovì

ABSTRACT: The relationships between tertiary structure and copper binding in *Pseudomonas fluorescens* azurin have been studied by fluorescence, absorption, and electron paramagnetic resonance spectra. The fluorescence spectrum at neutral pH shows an emission maximum at 308 nm when excited at 275 nm. Below pH 2 or in the presence of 6 M guanidine hydrochloride at neutral pH the fluorescence maximum shifts toward longer wavelengths. The addition of sodium dodecyl sulfate at low pH restores the original wavelength of emission. The loss of the blue color at low pH is not due to the reduction of copper, whereas in the presence of guanidine hydrochloride it is concomitant with copper reduction. Above pH 9 copper is reduced without any apparent change in protein structure. The recombination of apoprotein with copper causes a quenching of protein fluorescence, without any increase of absorbance at the wavelengths of the emission spectrum of the protein. Mercuric ion resembles copper in its quenching effect. *p*-Mercuribenzoate reacts only with the copper-free protein but does not influence the fluorescence.

The results suggest the interaction of copper with tryptophan and sulfhydryl in a strongly hydrophobic site. Copper is reduced by SH when the site becomes exposed to the solvent (*i.e.*, in guanidine hydrochloride at neutral pH) or when the pH is sufficiently high to favor the oxidation of the sulfhydryl in the native site.

The nature of the amino acid residues which form the copper binding site in copper proteins has been only tentatively established in the proteins studied so far, on the basis of mostly indirect evidence or of a questionable use of protein reagents, since more direct and reliable methods did little to solve the question. Electron paramagnetic resonance spectroscopy—in the absence of ligand hyperfine structure, as in native copper proteins—can give only suggestions and working hypotheses; X-ray crystallography is in its initial stages in this field, because of the difficulties involved in crystallizing copper proteins. A great deal of the theory on the copper binding sites in proteins rests on information derived from spectroscopy (optical spectroscopy, optical rotatory dispersion, and electron paramagnetic resonance); this, however, mostly concerns the symmetry of the sites

* From the Institute of Biological Chemistry, C. N. R., Centre for Molecular Biology, University of Rome, and the Institute of Microbiology, University of Rome (V. B.), Rome, Italy. *Received November 4, 1969.*