γ-Chymotrypsin Is a Complex of α-Chymotrypsin with Its Own Autolysis

ARTICLE *in* BIOCHEMISTRY · JUNE 1991

Impact Factor: 3.02 · DOI: 10.1021/bi00235a015 · Source: PubMed

5 AUTHORS, INCLUDING:

Felix Frolow
Tel Aviv University
207 PUBLICATIONS **9,789** CITATIONS
SEE PROFILE

Joel Sussman
Weizmann Institute of Science
359 PUBLICATIONS **22,845** CITATIONS
SEE PROFILE

Hansen, J. P., & McDonald, I. R. (1976) *Theory of Simple Liquids*, Academic Press, London.
Harvey, S. C. (1989) *Proteins 5*, 78–92.
Jackson, J. D. (1975) *Classical Electrodynamics*, Wiley, New York.
Klapper, I., Hagstrom, R., Fine, R., Sharp, K., & Honig, B. (1986) *Proteins 1*, 47–59.
Kördel, J., Forsén, S., & Chazin, W. J. (1989) *Biochemistry 28*, 7065–7074.
Linse, P. (1986) *J. Phys. Chem. 90*, 6821–6828.
Linse, S., Brodin, P., Drakenberg, T., Thulin, E., Sellers, P., Elmdén, K., Grundström, T., & Forsén, S. (1987) *Biochemistry 26*, 6723–6735.
Linse, S., Brodin, P., Johansson, C., Thulin, E., Grundström, T., & Forsén, S. (1988) *Nature 335*, 651–652.
Linse, S., Johansson, C., Brodin, P., Grundström, T., Drakenberg, T., & Forsén, S. (1991) *Biochemistry 30*, 154–162.
Linse, S., Teleman, O., & Drakenberg, T. (1990) *Biochemistry 29*, 5925–5934.
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953) *J. Chem. Phys. 21*, 1087–1092.
Moews, P. C., & Kretsinger, R. H. (1975) *J. Mol. Biol. 91*, 201–228.
Rasmussen, H. (1986a) *N. Engl. J. Med. 314*, 1094–1101.
Rasmussen, H. (1986b) *N. Engl. J. Med. 314*, 1164–1170.
Sharp, K., & Honig, B. (1989) *Chem. Scr. 29A*, 71–74.
Skelton, N. J., Forsén, S., & Chazin, W. J. (1990a) *Biochemistry 29*, 5752–5761.
Skelton, N. J., Kördel, J., Forsén, S., & Chazin, W. J. (1990b) *J. Mol. Biol. 213*, 593–598.
Stern, O. (1924) *Z. Elektrochem. Angew. Phys. Chem. 30*, 508–516.
Svensson, B., & Woodward, C. (1988) *Mol. Phys. 64*, 247–259.
Svensson, B., Jönsson, B., & Woodward, C. E. (1990) *Biophys. Chem.* (in press).
Szebenyi, D. M. E., & Moffat, K. (1986) *J. Biol. Chem. 261*, 8761–8777.
Szebenyi, D. M. E., Obendorf, S. K., & Moffat, K. (1981) *Nature 294*, 327–332.
Tanford, C., & Kirkwood, J. G. (1957) *J. Am. Chem. Soc. 79*, 5333–5339.
Teleman, O., Svensson, B., & Jönsson, B. (1990) *Comput. Phys. Commun.* (in press).
Warwicker, J., & Watson, H. C. (1982) *J. Mol. Biol. 157*, 671–679.
Wendt, B., Hofmann, T., Martin, S. R., Bayley, P., Brodin, P., Grundström, T., Thulin, E., Linse, S., & Forsén, S. (1988) *Eur. J. Biochem. 175*, 439–445.
Wennerström, H., & Jönsson, B. (1988) *J. Phys. (Paris) 49*, 1033–1041.
Widom, B. (1963) *J. Chem. Phys. 39*, 2808–2812.

γ-Chymotrypsin Is a Complex of α-Chymotrypsin with Its Own Autolysis Products[†,‡]

M. Harel,[§] C.-T. Su,[∥] F. Frolow,[⊥] I. Silman,[*,∥] and J. L. Sussman[*,§]
Departments of Structural Chemistry, Neurobiology, and Chemical Services, Weizmann Institute of Science, Rehovot 76100, Israel

Received June 19, 1990; Revised Manuscript Received December 11, 1990

ABSTRACT: The determination of three separate γ-chymotrypsin structures at different temperatures and resolutions confirmed the presence of electron density in the active site, which could be interpreted as an oligopeptide as had previously been suggested by Dixon and Matthews [(1989) *Biochemistry 28*, 7033–7038]. HPLC analyses of the enzyme before and after crystallization demonstrated the presence of a wide variety of oligopeptides in the redissolved crystal, most with COOH-terminal aromatic residues, as expected of the products of chymotrypsin cleavage, which appeared to arise from extensive autolysis of the enzyme under the crystallization conditions. The refined structures agree well with the conformation of both γ-chymotrypsin and α-chymotrypsin. The electron density in the active site is thus interpreted as arising from a repertoire of autolysed oligopeptides produced concomitantly with crystallization. The COOH-terminal carbons of the polypeptide(s) display short contact distances (1.97, 2.47, and 2.13 Å, respectively) to Ser195 O^{γ} in all three refined structures, but the electron density is not continuous between these two atoms in any of them. This suggests that some sequences are covalently bound as enzyme intermediates while others are noncovalently bound as enzyme–product complexes.

The serine protease γ-chymotrypsin (γ-Cht)[1] was identified as a crystalline form distinct from α-chymotrypsin (α-Cht) by Kunitz (1938) although it has a very similar conformation (Matthews et al., 1967; Cohen et al., 1981). Although α-Cht and γ-Cht are believed to be identical in primary sequence (Desnuelle, 1960), the question of whether the two crystalline forms originate from two distinct conformational species in solution has remained a matter for discussion (Corey et al., 1965; Dixon & Matthews, 1989). γ-Cht is crystallized at higher pH (e.g., 5.6 vs 4.2) than α-Cht and packs as a mo-

[†] This work was supported by U.S. Army Medical Research and Development Command Contracts DAMD17-GC-7037 and DAMD17-89-C-9063, by the Minerva Foundation, Munich, Germany, and by the U.S.–Israel Binational Science Foundation.
[‡] Coordinates (code name 8GCH) for γ-chymotrypsin (Nat-1) were deposited in the Protein Data Bank, Brookhaven National Laboratory (Bernstein et al., 1977).
* Correspondence should be addressed to these authors.
[§] Department of Structural Chemistry.
[∥] Department of Neurobiology.
[⊥] Department of Chemical Services.

[1] Abbreviations: α-Cht, α-chymotrypsin; Cht, chymotrypsin; Chtph7, γ-Cht at pH 7.0; γ-Cht, γ-chymotrypsin; OMTKY3, turkey ovomucoid inhibitor third domain; TFA, trifluoroacetic acid.

0006-2960/91/0430-5217$02.50/0

nomer, while α-Cht crystallizes as a dimer. In γ-Cht, there is ready access to the active site, whereas that of α-Cht is only accessible to very small substrates and analogues [e.g., Henderson (1970)]. γ-Cht is thus the preferred crystal form for binding and inactivation studies of serine proteases (Ringe et al., 1985, 1986).

In the course of a comparative crystallographic study of "aged" and "nonaged" organophosphoryl conjugates of γ-Cht (Harel et al., 1990), a high-resolution study of the native enzyme was carried out as a control. Surprisingly, this study revealed an area of continuous electron density within the active site which could be modeled as an oligopeptide. As this work was in progress, we became aware of the high-resolution (1.6-Å) structure determination of γ-Cht at pH 7.0 (Chtph7) of Dixon and Matthews (1989), which showed an oligopeptide bound at the active site. The electron density map could be satisfactorily modeled by the structure X-Pro-Gly-Ala/Val-Tyr. Yapel et al. (1966) earlier reported that commercial preparations of α-Cht contained contaminants, including presumed autolysis products, as well as a tightly bound moiety which prevented substrate binding. In order to further characterize the electron density within the active site of the native γ-Cht structure, we have now collected X-ray diffraction data sets from three separate crystals, at different resolutions (1.6, 1.8, and 1.9 Å) and temperatures (−183 and 25 °C), and refined all three structures independently. Concomitantly, we separated by HPLC, and subsequently sequenced, oligopeptides present in γ-Cht solutions obtained from redissolved crystals.

Experimental Procedures

Crystals of γ-Cht were grown, following the procedure of Cohen et al. (1981), from twice crystallized and lyophilized γ-Cht (Sigma lot no. C-4754) at pH 5.6. One set of X-ray intensity data (Nat-1) was collected from a single crystal, to 1.6-Å resolution, on a Rigaku AFC5-R rotating anode diffractometer at −183 °C, according to the cryotemperature technique developed by Hope (1988). The X-ray intensity data of the second set (Nat-2) were collected from two crystals, the first up to 2.5-Å resolution and the second from 2.5- to 1.9-Å resolution. Data were collected on a Rigaku AFC5-R rotating anode diffractometer at −183 °C. The third data set (Nat-3) was collected from a single crystal to 1.8-Å resolution, at room temperature, on a Siemens/Xentronics rotating anode area detector. The data collection parameters are summarized in Table I.

Refinement of the Nat-1 structure started from the 1.9-Å refined room temperature γ-Cht coordinates of Cohen et al. (1981), with all the water molecules excluded and all atomic temperature factors set to 10.0 Å^2. The refinement was carried out by using the modified restrained least-squares refinement program PROLSQ (Hendrickson, 1985). The version of the program used for these studies incorporates a fast Fourier algorithm to speed up calculations (Finzel, 1987), and it restrains intermolecular contacts (Sheriff, 1987). All reflections with $I > 1\sigma$ at a resolution range of 6–1.6 Å were used in the refinement. Electron density maps with coefficients $(2F_o - F_c)$ and $(F_o - F_c)$ were examined and fitted, on an interactive color computer graphics system (Evans and Sutherland PS390), using the program FRODO (Jones, 1978; Pflugrath et al., 1984).

Refinement of the Nat-2 cryotemperature structure started, as for Nat-1, from the 1.9-Å refined room temperature γ-Cht coordinates of Cohen et al. (1981) and was carried out as above. The resolution range used was increased gradually from 6–2.5 to 6–1.9 Å. All reflections with $F > 3\sigma$ at a resolution range of 6–1.9 Å were used in the refinement. At the final stages of the refinement, weaker reflections were added and 14893 reflections with $F > 0$ were used.

Table I: Crystal Data for Native γ-Cht

	Nat-1	Nat-2[a]	Nat-3
temperature (°C)	−183 (±0.5)	−183 (±0.5)	ca. 25
unit cell dimensions (Å)			
a	69.0	68.7; 68.8	69.8
c	95.3	95.6; 95.0	98.1
data collection	diffractometer	diffractometer	area detector
resolution (Å)	1.6	1.9	1.8
no. of reflections measured	33900	11192; 11986	100675
theoretical no. of unique reflections	29204	16854	19958
no. of reflections used for refinement	25266[b]	15578[c]	15469[d]
R_{sym}	0.024[e]	0.041; 0.023[e]	0.056[f]
length of data collection (days)	20	6; 8	2

[a] Data from two crystals, one measured to 2.5-Å resolution and the other between 2.5- and 1.9-Å resolution; $R_{merge} = 0.068$. [b] $I > 1\sigma$. [c] $F > 0.0$. [d] $F > 8.0$. [e] R_{sym} is $\sum|I - \langle I\rangle|/\sum\langle I\rangle$. [f] R_{sym} is the unweighted squared R factor on intensity.

The refinement of the Nat-3 room temperature structure started from the cryotemperature 1.9-Å refined γ-Cht structure (Nat-2), rotated so as to obtain the best molecular fit to the refined room temperature structure of Cohen et al. (1981). The initial model included the tripeptide Gly-Ala-Trp in the active site, 276 water molecules, and 3 sulfate ions. The refinement was carried out as above. Weak reflections were not included in the refinement, and ca. 75% of the total number collected (15469 of 20758) were used. The resolution range used was increased gradually from 6–2.5 to 6–1.8 Å.

During the iterative refinement process, solvent molecules were added by assigning peaks in the electron density maps when their density was greater than 3σ in the $(F_o - F_c)$ maps or greater than 1.5σ in the $(2F_o - F_c)$ maps and when their location was within hydrogen-bonding distance of proton donors or acceptors.

In order to establish the chemical identity of the polypeptide(s) in the active site, γ-Cht was subjected to HPLC analysis on an Aquapore Bu-300 (C4) column (dimensions 4.6 mm × 25 cm), on a Spectra Physics Model SP 8800 modular chromatograph Chrom-A-Scope, equipped with a BarSpec Model 1937. Samples before and after crystallization were dissolved in 0.1% trifluoroacetic acid (TFA) to a concentration of ca. 2 mg/mL. Volumes of 250 μL were injected and eluted by a gradient of 0–80% acetonitrile containing 0.1% TFA. The tripeptide Gly-Ala-Trp was synthesized and subjected to HPLC elution in order to establish its retention time on the column as a reference point. The eluted peaks were collected and were sequenced by using an Applied Biosystems Model 470 protein sequencer equipped with an online Model 120A PTH analyzer.

Results

(a) Refinement Results. (1) Nat-1. The initial R factor of the room temperature structure (Cohen et al., 1981) was 30.3%. As the refinement progressed, a region of continuous density was seen adjacent to Ser195 O^γ of the active site. Initially, we tried to assign it to solvent molecules in the same way as was done in the refined structure of *Streptomyces griseus* protease A (Sielecki et al., 1979), where solvent molecules are found close to the carboxylate at P_1 and within the hydrophobic pocket designed for aromatic rings of P_1. This

Table II: Results of PROLSQ Refinement of Native γ-Cht

	Nat-1	Nat-2	Nat-3
Crystal Data			
R factor	0.193	0.193	0.156
resolution (Å)	6–1.6	6–1.9	6–1.8
temperature (°C)	−183	−183	25
no. of water molecules	347	255	156
no. of sulfate ions	4	3	3
rms Deviation from Restraints (σ Target Values in Refinement)			
bond length (Å) (σ = 0.025)	0.014	0.014	0.014
angle distance (Å) (σ = 0.035)	0.027	0.028	0.027
planar 1–4 distance (Å) (σ = 0.050)	0.031	0.031	0.031
plane restraints (Å) (σ = 0.020)	0.010	0.009	0.010
chiral-center restraints (σ = 0.150)	0.121	0.127	0.136
nonbonded contact restraints (Å)			
single torsion contacts (σ = 0.5)	0.179	0.180	0.181
multiple torsion contacts (σ = 0.5)	0.217	0.223	0.204
possible hydrogen bonds (σ = 0.5)	0.271	0.268	0.211
conformational torsion planar angle (deg) (σ = 2.8)	1.8	1.7	2.1

Table III: rms Deviation in C^{α} Positions (Å) of Cht Structures

	Nat-1	Nat-2	Nat-3	Chtph7[a]	α-Cht[b]
Nat-1		0.15	0.31	0.29	0.48
Nat-2			0.31	0.29	0.49
Nat-3				0.17	0.44

[a] γ-Cht at 1.6 Å, pH 7 (Dixon & Matthews, 1989). [b] Molecule 1 of the dimer of α-Cht at 1.67 Å (Tsukada & Blow, 1985). Residues 75–79, which differ by a maximum of 3.8 Å between all γ-Cht structures and α-Cht, were omitted from the calculation. The rms deviations with this loop included are 0.59, 0.59, and 0.54 Å, respectively.

assignment failed both since the peaks in the active site region appeared to have very short intermolecular distances (<1.8 Å) and because the stretch of continuous density was long. Since the density was shaped like a short polypeptide, we attempted to superimpose the complex of the turkey ovomucoid inhibitor third domain (OMTKY3) bound to α-Cht (Fujinaga et al., 1987) on the refined γ-Cht structure. The root-mean-square deviation of the C^{α} coordinates between the two structures is 0.44Å. The fold of part of the OMTKY3 polypeptide chain in the active site region (residues Cys16I–Leu18I) fitted both the observed and the unaccounted for electron density very well. The side chains were changed to fit the observed density from Cys-Thr-Leu in OMTKY3 to Gly-Ala-Trp. The refinement converged to an *R* factor of 19.3%. No major main-chain changes from the room temperature structure were needed.

(*2*) *Nat-2.* The initial *R* factor of the room temperature structure (Cohen et al., 1981) was 25.2%. The density of the tripeptide in the active site could be fitted best by the sequence Gly-Ala-Trp. The refinement converged to an *R* factor of 19.3%.

(*3*) *Nat-3.* The refinement started at an *R* factor of 34.5%. As the refinement progressed, the density of the polypeptide in the active site appeared longer than that of a tripeptide and could be fitted best by the sequence of the pentapeptide Ala-Ala-Ala-Ala-Tyr. The refinement converged to an *R* factor of 15.6%.

The final parameters for the three refinements are shown in Table II.

(*b*) *Agreement between Refined Structures.* All three refined structures agree well with each other as can be seen from Table III and Figure 1. It is interesting that the best agreement in the rms deviation of the C^{α} positions is between Nat-1 and Nat-2, both of whose diffraction data were collected at cryotemperature, and between Nat-3 and Chtph7 (Dixon & Matthews, 1989), for both of which the diffraction data

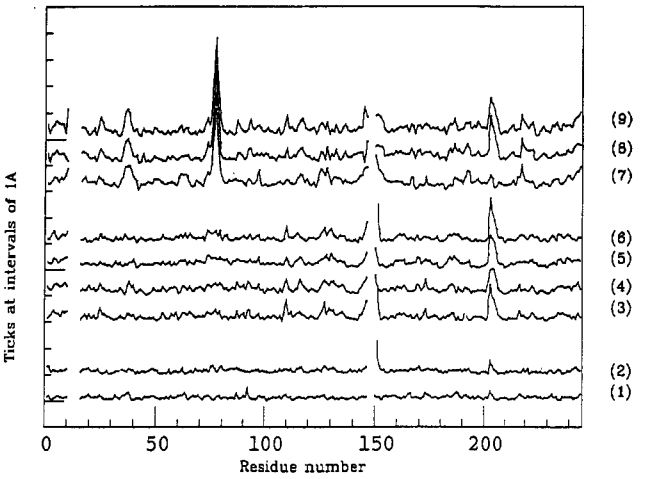

FIGURE 1: Root-mean-square deviation of C^{α} positions between pairs of refined structures. (Lower group) Pairs of γ-Cht (data collected at the same temperature): (1) Nat-1 vs Nat-2 both at cryotemperature; (2) Nat-3 vs Chtph7 both at room temperature. (Middle group) Pairs of γ-Cht (data collected at different temperatures): (3) Nat-2 vs Nat-3; (4) Nat-2 vs Chtph7; (5) Nat-1 vs Nat-3; (6) Nat-1 vs Chtph7. (Top group) Pairs of γ-Cht vs α-Cht: (7) Nat-3 vs α-Cht (both at room temperature); (8) Nat-1 vs α-Cht (at different temperatures); (9) Nat-2 vs α-Cht (at different temperatures).

were collected at room temperature. For the latter pair of refined structures a single deviation was noted at Asn150 (see below). The deviation between structures at cryotemperature vs those of room temperature is seen mainly in loop 202–204 where the C^{α} position of residue 203 differs by ca. 1 Å. A comparison of the γ-Cht structures to the refined room temperature structure of α-Cht (Tsukada & Blow, 1985) (see Table III and top group in Figure 1) shows a slightly larger rms deviation. The largest difference is found in the position of loop 75–79 which in γ-Cht is held by two hydrogen bonds between Ser76 and a symmetry-related molecule whose nearby surface does not allow the loop freedom of motion. Comparisons of the C^{α} positions with loop 75–79 omitted decrease the deviations by ca. 0.1 Å. There is better agreement between α-Cht and γ-Cht room temperature structures than between similar pairs collected at different temperatures, with loop 202–204 again being the principal factor in the deviation (see top group in Figure 1).

(*c*) *Attempts at Identification of the Polypeptide in the Active Site.* HPLC analysis of γ-Cht before crystallization (Figure 2a) shows four peaks above noise level besides the major protein peak at 35.5 min. These four peaks all eluted later than 20 min after injection. A similar analysis of a dissolved crystal of γ-Cht (Figure 2b) shows a large number of peaks in addition to the protein peak at 36.9 min. Peaks appeared as early as 3 min after injection and can be attributed to oligopeptides. The synthetic tripeptide Gly-Ala-Trp elutes at 12.7 min. All three refined structures of γ-Cht contain a polypeptide terminating with an aromatic residue in their active site. Accordingly, peaks that exhibited a high absorption at 280 nm (see Figure 2c), and should thus contain an aromatic residue, were sequenced. The results of sequencing are shown in Table IV. It can be seen that, out of 11 oligopeptides sequenced, 10 possess sequences corresponding to known sequences in α-Cht. One such peptide, identified as Gly-Ala-Trp, eluted at the same position as the authentic synthetic tripeptide.

(*d*) *Active Site Conformation.* The active site in the three refined crystals shows hydrogen bonding in the charge relay system similar to that reported for α-Cht (Tsukada & Blow, 1985) and for Chtph7 (Dixon & Matthews, 1989). The conformation is summarized in Table V. A careful analysis

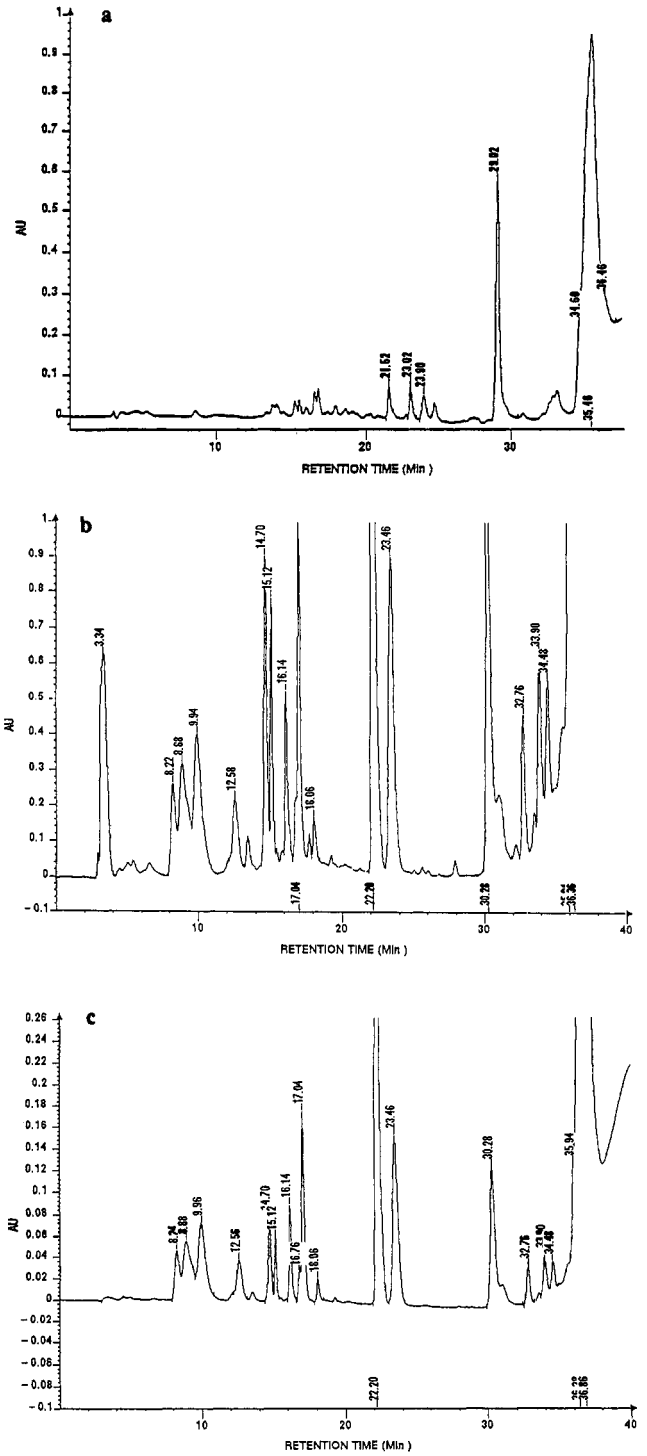

FIGURE 2: HPLC of γ-Cht before and after crystallization. (a) γ-Cht before crystallization, monitored at 214 nm. (b) Dissolved γ-Cht crystal, monitored at 214 nm. (c) Same as (b) but monitored at 280 nm. The numbers above the peaks represent their retention times in minutes.

of distances and dihedral angles of the independently determined structures shows that in α-Cht the distance Ser195 O^{γ}···His57 $N^{\epsilon 2}$ differs by 3σ (σ = 0.08 Å) from the mean and, likewise, Ser195 χ_1 differs by 2.5σ (σ = 1.6°) from the mean. This indicates a small but significant conformational change in residue Ser195 between the α-Cht and γ-Cht structures.

As mentioned above, the observed density at the active site could be fitted by a polypeptide which superimposes well on the backbone of the ovomucoid polypeptide at the three positions preceding the scissile carboxyl (P_1, P_2, and P_3) of α-Cht:OMTKY3, as can be seen in Figure 3. The side chains of the polypeptide were replaced by those of Gly-Ala-Trp, Gly-Ala-Trp, and Ala-Ala-Ala-Ala-Tyr, fitting the observed density in Nat-1, Nat-2, and Nat-3, respectively (see Figure 4, panels a, b, and c). The hydrogen bonding of the polypeptides in the active site resembles that in other complexes of serine proteases and polypeptides. Comparison with a nonbonded tetrapeptide and a bonded tetrapeptide terminating with an aldehyde in the active site of protease A (James et al., 1980), as well as with Chtph7 (Dixon & Matthews, 1989), is shown in Table VI. The COOH-terminal carbon displays a short contact distance to Ser195 O^{γ} in all three refined native structures (1.97, 2.47, and 2.13 Å, respectively), but the electron density is not continuous between these two atoms in any of the structures (see Figure 4). In the refinement, all van der Waals restraints have been removed from the terminal carboxyl atom in order to let it refine unhindered. A further test of the validity of the short contacts was carried out by performing seven cycles of refinement, starting with a bonding distance of 1.8 Å between the active site Ser195 O^{γ} and the polypeptide terminal carboxyl C. The refinement resulted in the movement of the O^{γ} back to its previously refined position. The oligopeptide terminal carboxyl group density appeared planar although one of the terminal oxygens had a low occupancy (0.6).

(*e*) *Missing Residues.* In the process of activation of chymotrypsinogen, in vivo, two dipeptides, Ser14-Arg15 and Thr147-Asn148, are excised (Desnuelle, 1960). Several residues adjacent to the excision sites (Ser11, Gly12, Leu13, Ala149, Asn150) could not be traced in the 1.9-Å resolution map of γ-Cht (Cohen et al., 1981). In the −183 °C map of Nat-1, residue Asn150 was traced and refined to an average *B* factor of 19.3 $Å^2$, somewhat higher than the average *B* factor for the whole protein, which is 11.7 $Å^2$. In all three refined structures, an electron density peak is seen ca. 1.4 Å from the carbonyl C atom of residue Leu10. This peak can be attributed either to a terminal oxygen of a planar carboxyl group, indicating cleavage of residues 11–13, or to the N atom of the following disordered residue—Ser11.

(*f*) *Solvent Structure.* Water molecules were ascribed to electron density peaks located less than 3.2 Å from proton donors or acceptors in the protein. In the cryotemperature Nat-1 structure, 347 such peaks were found, while in Nat-2 (at the same temperature but at 1.9-Å resolution), only 255 were found and in Nat-3, at room temperature and 1.8-Å resolution, just 156 were found. A comparison of the positions of the Nat-1 water peaks with those in Nat-3 showed 81% of the molecules in common, i.e., within 1 Å of each other, while of the 255 putative water molecules in the Nat-2 structure, 85% were within 1 Å of the Nat-1 water positions.

After refinement of Nat-1, four of the solvent molecules still showed a $(F_o - F_c) > 3\sigma$ peak and their occupancy tended to increase beyond 1.0. The electron density of these peaks clearly showed a tetrahedral shape, and they were accordingly assigned to be sulfate ions. Analysis of the solvent structure of α-Cht at 1.67-Å resolution shows that two of its four sulfate ions are in the vicinity of Arg154 and one is near Lys93 (Blevins & Tulinsky, 1985), while in the Nat-1 structure Sul299 interacts with Arg154 and Sul300 with Ser92. In the structures of Nat-2 and Nat-3, refined independently, only three sulfate ions were seen, while the fourth position was assigned as water.

(*g*) *Discrete Disorder.* In the refined structure of Nat-1, after almost all solvent molecules had been assigned, the highest peaks with $(F_o - F_c) > 3\sigma$ were all located at bonding distances near the termini of side chains and were assigned

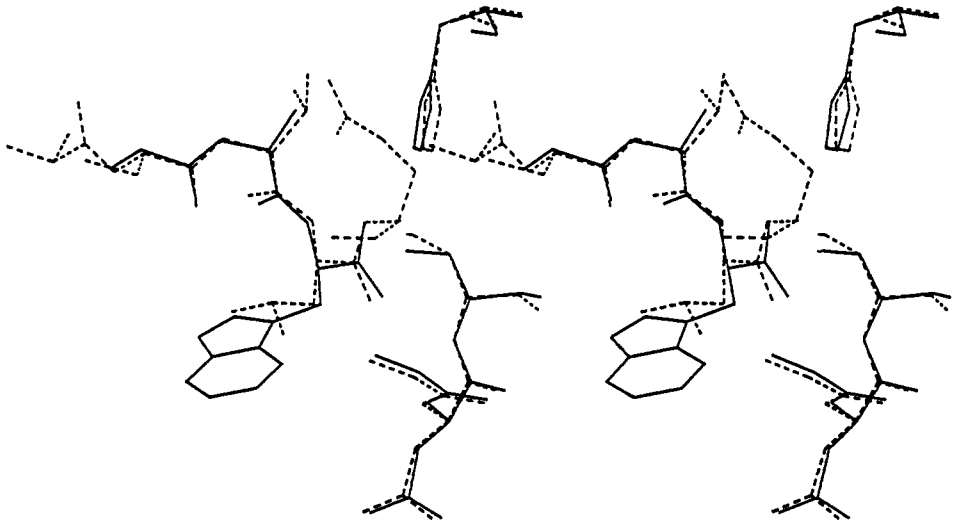

FIGURE 3: Stereo superposition of the γ-Cht active site residues His57, Gly193-Asp194-Ser195, and the tripeptide Gly-Ala-Trp of Nat-1 (solid line) on the same α-Cht residues and the bound inhibitor OMTKY3 residues at P_4, P_3, P_2, P_1, and S_1 (dashed line).

Table IV: Sequences of HPLC Peaks Obtained from Crystalline γ-Cht Showing High Absorption at 280 nm

retention time (min)	position 1	2	3	4	5	6	7	8	9	10	possible sequence in Cht
8.88	A, G, S	E, W, G	F								GEF_{71}; SW_{215}
9.96	V, S, G, A	G, F, A, W	G	P	L						
12.56	G	A	W								GAW_{207}
12.72[a]	G	A	W								
14.70	S, T	T	P	G	V	Y					$STPGVY_{228}$
15.12	T	S, G	T	P	G	V	Y				$TSTPGVY_{228}$
16.14	V	N	W								VNW_{237}
17.04	S, A	V, L	P	G	S	W					$AVPGSW_{27}$
22.20	S	W	P	W							$SWPW_{29}$
23.46	A	V	P	G	S	W	P	W			$AVPGSWPW_{29}$
30.28	I	V	N	G	E	E	A	V	P	G...	$IVNGEEAVPG_{25...}$[b]

[a] Synthetic tripeptide Gly-Ala-Trp. [b] Sequence determined only this far.

Table V: Conformation of Charge Relay System of Cht

	Nat-1	Nat-2	Nat-3	Chtph7[a]	α-Cht[b]
	Hydrogen Bonds (Å)				
Ser195 O^{γ}–His57 $N^{\epsilon 2}$	2.99	2.92	3.16	3.09	2.78
Asp102 $O^{\delta 2}$–His57 $N^{\delta 1}$	2.70	2.81	2.74	2.56	2.63
Asp102 $O^{\delta 1}$–His57 N	2.96	2.79	2.91	2.82	2.84
Asp102 $O^{\delta 1}$–Ala56 N	2.82	2.80	2.98	2.79	2.77
	Torsion Angles (deg)				
Ser195 χ_1	−77	−73	−60	−77	−82
His57 χ_1	84	86	86	82	79
His57 χ_2	−105	−93	−103	−98	−94
Asp102 χ_1	−168	−163	−166	−171	−166
Asp102 χ_2	−165	−173	−165	−167	−166

[a] γ-Cht at 1.6 Å, pH 7 (Dixon & Matthews, 1989). [b] α-Cht at 1.67 Å; average of molecules 1 and 2 of the dimer (Tsukada & Blow, 1985).

as arising from discrete disorder. Two serine side chains displayed disorder in their terminal O^{γ} atom (Ser32, Ser159), and two valine side chains displayed disorder in their terminal atoms $C^{\gamma 1}$ and $C^{\gamma 2}$ (Val9, Val65). The refined room temperature structure of Nat-3 showed disorder only in the O^{γ} atom of Ser159, while in Nat-2 no disorder could be seen.

(*h*) *Side-Chain Electron Density.* In the refined structure of Nat-1, the electron densities of all side chains but two were clearly visible. The density for Ser113 does not show the location of O^{γ} (see Figure 5a), and the $(F_o - F_c)$ map, omitting the Ser113 side chain, shows only the C^{β} position. All positions tried for the O^{γ} resulted in a peak of negative density. Since the electron density around the C^{β} is slightly mushroom-shaped, we conclude, considering the quality of the map elsewhere, that the O^{γ} atom is disordered. No electron density was seen, either, for Ser113 O^{γ} in the refined structures of Nat-3. In the Glu49 side chain, no electron density was seen for the carboxyl atoms in Nat-1. All attempts at placing them resulted in a large negative density peak, and omitting the side chain from the refinement did not produce any density beyond C^{γ} (see Figure 5b). In contrast, both Nat-2 and Nat-3 show density for all atoms of Glu49.

Discussion

The three independent structure determinations were carried out in order to reconcile the results of the HPLC analysis, which showed that autolysis of γ-Cht results in a variety of oligomers under the crystallization conditions employed, with the identification of a seemingly "unique" oligomer in the active site of the refined γ-Cht X-ray structure.

The three refined structures, obtained at cryotemperature and room temperature, look very similar. Nat-1 and Nat-2, however, show the expected shrinkage of the unit cell volume by ca. 6% as a consequence of freezing the crystal at −183 °C. Although the average temperature factor of the structures did not differ much between the room temperature and cryotemperature structures, the fluctuation of the *B* factors along the chain was considerably reduced in the cryotemperature structures. The different *R* factors obtained for the structures reflect the different resolutions and weak reflection cutoffs. It became clear, when looking at the electron density maps of Nat-1, that if the counting statistics are sufficiently good, it is possible to use virtually all reflections, resulting in a significantly improved map, although this clearly has an ad-

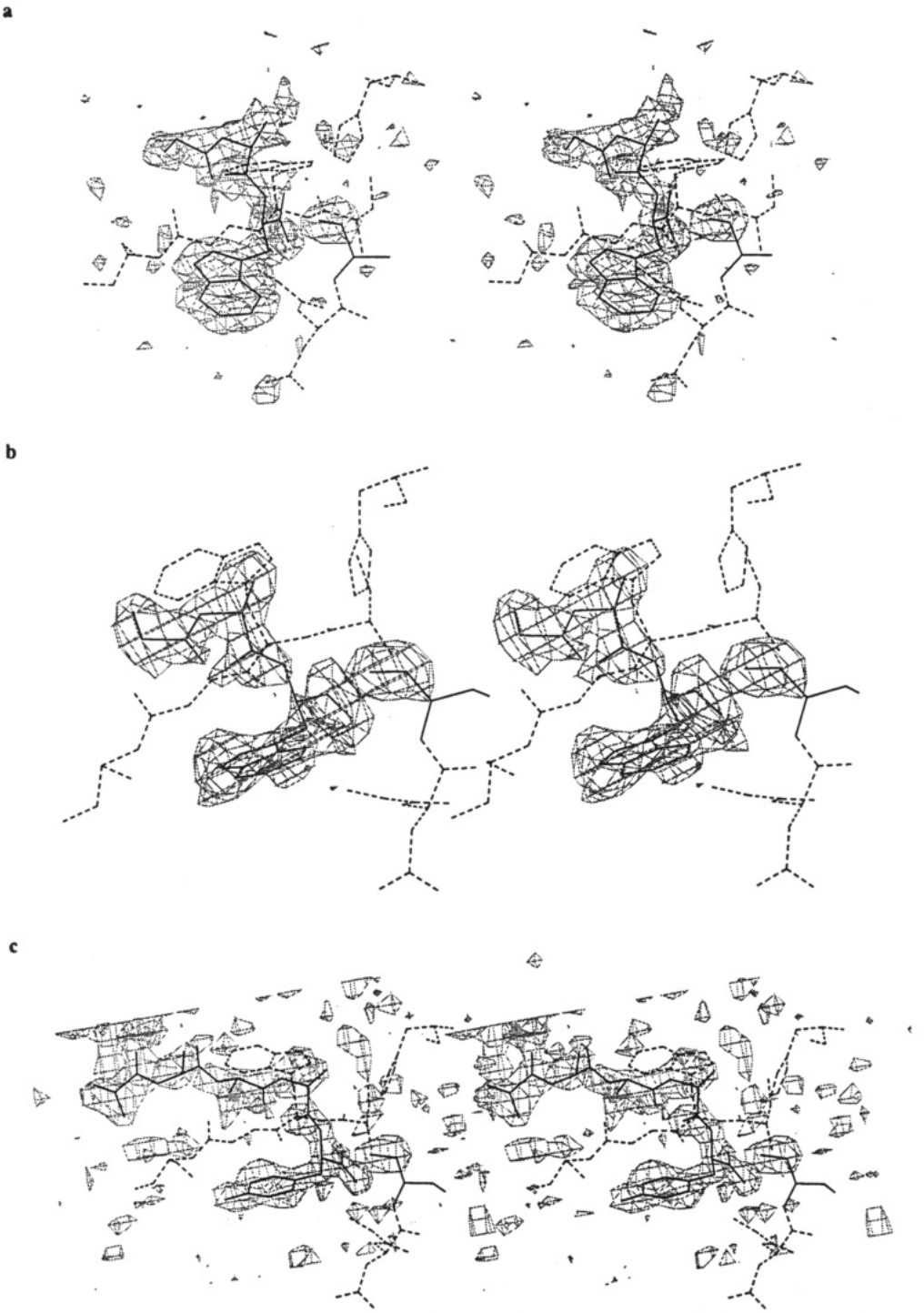

FIGURE 4: "Omit" difference electron density maps drawn in stereo, showing the bound oligopeptide (solid line) in the active site of the enzyme (dashed line). The maps were calculated with amplitudes of $(F_o - F_c)$ and phases from γ-Cht after seven additional cycles of refinement omitting the putative oligopeptide and the side chain of Ser195. Contours drawn at 2σ. (a) Omit map for Nat-1 showing the putative tripeptide Gly-Ala-Trp and the neighboring residues. (b) Omit map for Nat-2 showing the putative tripeptide Gly-Ala-Trp and the neighboring residues as above. (c) Omit map for Nat-3 showing the putative pentapeptide Ala-Ala-Ala-Ala-Tyr and the neighboring residues as above.

verse effect on the *R* factor. However, for Nat-2 and Nat-3, where the counting statistics were poorer, including the weakest reflections was deleterious to the map. The *R* factor for the three structures, as a function of resolution, is shown in Figure 6.

All three refined structures of γ-Cht, regardless of temperature or resolution, revealed, similarly to Chtph7 (Dixon & Matthews, 1989), the presence of an oligopeptide in their active site. The fold of the polypeptide follows that of the residues at the sites preceding the scissile bond in a complex of α-Cht with a natural polypeptide inhibitor (α-Cht–OMTKY3; Fujinaga et al., 1987) (see Figure 3). The electron density of the polypeptide could be best fitted by the sequences Gly-Ala-Trp, Gly-Ala-Trp, and Ala-Ala-Ala-Ala-Tyr in Nat-1, Nat-2, and Nat-3, respectively, as compared to Pro-Gly-Ala-Tyr seen by Dixon and Matthews (1989). Two features stand out as common to these sequences: (1) They all terminate in an aromatic residue. (2) The preceding side chains are all very short.

Comparison of the HPLC profiles of γ-Cht, before and after crystallization, showed that the enzyme undergoes extensive autolysis under the conditions of crystallization. The tendency

Table VI: Hydrogen Bonds and Short Contacts (Å) in the Active Site of γ-Cht

site	(atom 1)...(atom 2)	Nat-1[a]	Nat-2[b]	Nat-3[c]	Chtph7[d]	Pr-NB[e]	Pr-B[f]
P_3	O...Gly216 N	3.16	3.00	3.32	3.3	3.0	3.1
	N...Gly216 O	2.92	3.46	3.13	2.8	2.8	3.0
	N...Wat530 O[g]	2.52	2.45	–[i]	–	na	na
P_2	O...Wat487 O[g]	2.93	–	–	–	na	na
P_1	O...Ser195 N	2.69	2.97	2.67	3.5	3.2	3.2
	O...Gly193 N	3.03	3.03	2.84	2.8	2.8	3.1
	O...Wat615 O[g]	2.85	–	–	–	na	na
	O...Ser195 O^{γ}	2.20	2.43	2.04	2.6	na	na
	C...Ser195 O^{γ}	1.97	2.47	2.13	1.7	2.6	1.7
	N...Ser214 O	3.00	3.13	3.00	2.9	3.0	2.8
	N...Ser195 O^{γ}	3.02	3.14	2.96	3.1	na	na
	OT...His57 $N^{\epsilon 2}$	3.13	–	–	–	na	–
	OT...Wat647 O[g]	1.94	–	–	–	na	–
	OT...Wat487 O[g]	2.77	–	2.87	–	na	–
	Trp $N^{\epsilon 1}$...Ser217 O	3.12	3.22	–	–	–	–
	Tyr OH...Ser217 O	–	–	2.83	3.4	–	–
	Tyr OH...Ser217 N	–	–	2.92	3.3	–	–
	Tyr OH...Wat435 O[h]	–	–	3.09	–	–	–

[a] γ-Cht with peptide Gly-Ala-Trp in the active site. [b] γ-Cht with peptide Gly-Ala-Trp in the active site. [c] γ-Cht with peptide Ala-Ala-Ala-Ala-Tyr in the active site. [d] γ-Cht with peptide Pro-Gly-Ala-Tyr in the active site (Dixon & Matthews, 1989). [e] Nonbonded complex of protease A with Ac-Pro-Ala-Pro-Tyr (James et al., 1980) [f] Bonded complex of protease A with Ac-Pro-Ala-Pro-Phe(aldehyde) (James et al., 1980) [g] Water numbering is from Nat-1. [h] Water numbering is from Nat-2. [i] (–) indicates that one or both atoms are not seen in the structure. [j] na indicates nonavailable distance.

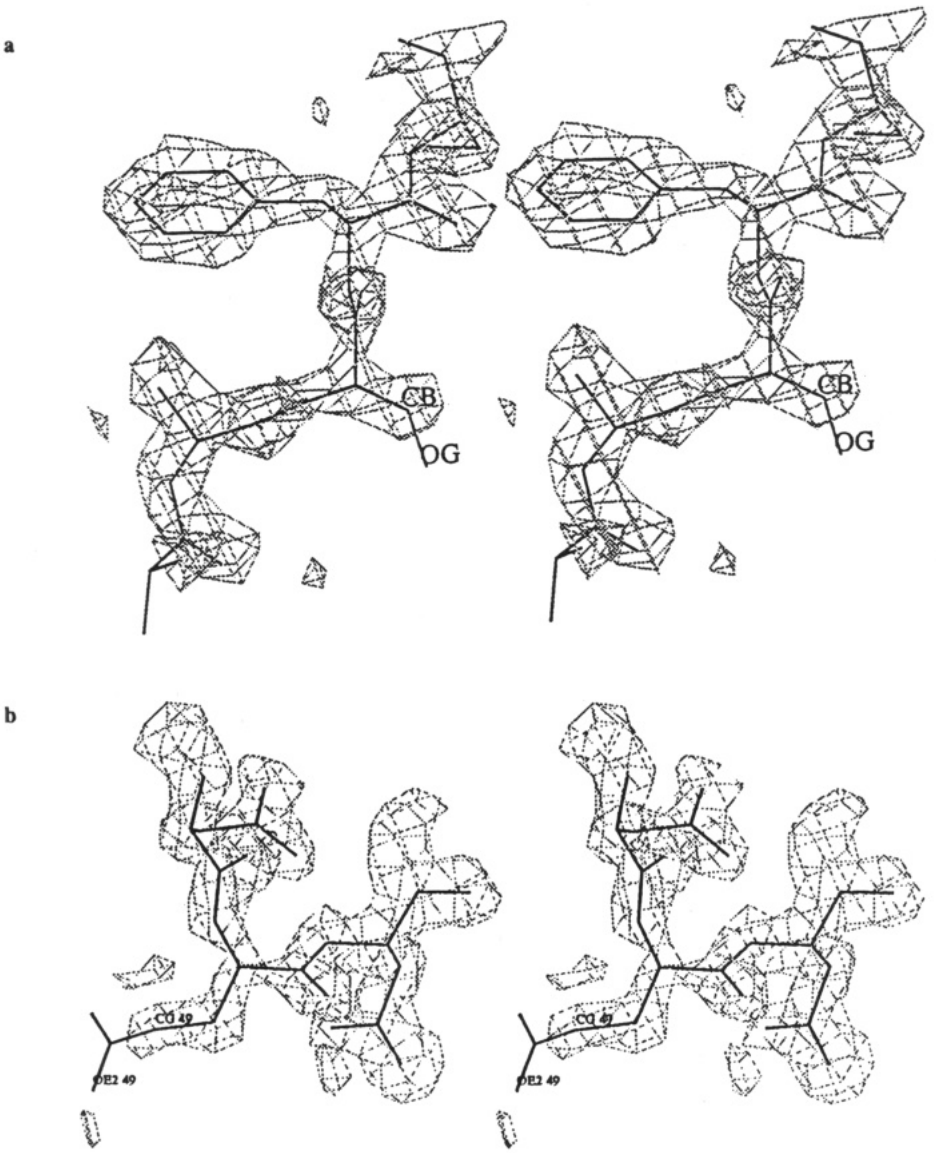

FIGURE 5: Stereo electron density map with coefficients $(2F_o - F_c)$ showing the lack of density for side chains in Nat-1. Contours drawn at 2σ. (a) Ser113 O^{γ}. (b) Carboxyl groups of Glu49.

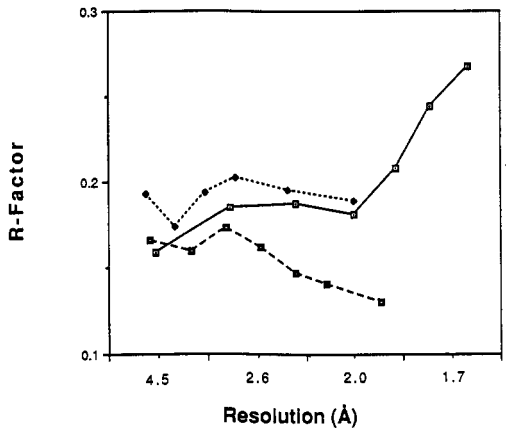

FIGURE 6: *R* factor as a function of resolution. Nat-1 (solid line); Nat-2 (dotted line); Nat-3 (dashed line).

Table VII: Average Main-Chain Temperature Factors (Å^2) in Polypeptides Found in γ-Cht Active Site

residue	1–245	P_1	P_2	P_3	P_4	P_5
Nat-1	9.7	14.8	18.7	21.4		
Nat-2	10.6	19.2	24.2	26.2		
Nat-3	8.8	19.4	21.4	24.0	26.9	27.6
Chtph7[a]	13.5	28.2	21.9	32.7	45.4	

[a] γ-Cht at 1.6 Å, pH 7 (Dixon & Matthews, 1989).

of α-Cht to autolysis was earlier noted by Yapel et al. (1966). The known specificity of Cht displays a preference for cleaving on the carboxyl side of aromatic residues (Desnuelle, 1960). This tendency can be seen in the list of HPLC peaks sequenced (see Table IV), most of which correspond to known sequences in Cht. It is possible, therefore, that this repertoire of autolysed sequences binds to the enzyme active site and that the refined structures show the density of an averaged bound sequence. Accordingly, the scissile residue shows the preponderance of aromatic side chains at this position while the residues preceding it all appear to have very short side chains (i.e., Ala or Gly). The apparent short side chains may be explained as being due to the density at these positions representing an average of densities of various side chains. Hence only the main-chain densities, and to a lesser degree, the C^β density which is common to all, are seen in the maps. The bound peptides are of varying lengths, and this is reflected in the gradual increase of the main-chain *B* factors (see Table VII). As the residues recede from the P_1 position, there will, on the average, be less occupancy due to the distribution of lengths of the bound polypeptide. It is worth noting that two of the peptides separated by HPLC, STPGVY, and TSTPGVY (corresponding to residues 223–228 and 222–228 in Cht) are actually compatible with the refined structure of Dixon and Matthews (1989) if, as these authors propose, the $C^{\gamma 1}$ and $C^{\gamma 2}$ atoms of the valine are disordered (as in fact is the case for Val9 and Val65 in Nat-1) and the NH_2-terminal residues of the bound oligopeptide are also disordered. Furthermore, another of the peptides separated and sequenced, Gly-Ala-Trp, is identical with the refined structures for the peptide bound in the active site of Nat-1 and Nat-2.

It should be noted that although native γ-Cht, in its crystalline state, appears to be an oligopeptide–protein complex, Cht can also crystallize in the γ-Cht habit when modified with diisopropyl phosphorofluoridate to produce bis(isopropyloxy)phosphoryl-Cht [Corey et al., 1965; for a discussion of this issue see Dixon and Matthews (1989)]. In contrast, although bis(isopropyloxy)phosphoryl-Cht can also crystallize in the α form, in native crystals of α-Cht, the active site was found to be unoccupied (Tsukada & Blow, 1985), and this may be ascribed to the difference in the dissociation constant of the product at the different pH values used in crystallization, 4.2 and 5.6 for α-Cht and γ-Cht, respectively (Kunitz, 1938). Due to the nature of the oxyanion hole, it is ideal for a carboxylate ion (present at pH 5.6), while rather poor for an uncharged carboxyl group (present at pH 4.2) whose ionization would involve a severe energy penalty.

The bonding distance of the carbonyl carbon to the Ser195 O^γ may also represent an average of a repertoire of bonding distances of peptides, some of which are bound to Ser195 O^γ covalently as enzyme intermediates at a distance of ca. 1.7 Å—typical examples being 1.73 Å between a tetrapeptide aldehyde and protease A (James et al., 1980) and 1.7 Å in Chtph7 (Dixon & Matthews, 1989)—while others are bound noncovalently as enzyme products at a distance of ca. 2.7 Å—e.g., 2.6 Å in the PTI–trypsin complex (Huber & Bode, 1978) and 2.7 Å in α-Cht–OMTKY3 (Fujinaga et al., 1987). Indeed, all three refined structures show different bonding distances (see Table VI). The bonding distances may thus reflect an average value for polypeptides bound covalently and noncovalently to the enzyme molecules in each crystal measured. Earlier studies (Cohen et al., 1969; Segal et al., 1971) showed that covalent oligopeptide inhibitors can be soaked into γ-Cht crystals. This occurs, presumably, with concomitant displacement of the putative endogenous polypeptide.

It is of interest that our proposal of a repertoire of oligopeptides at the active site of crystalline γ-Cht resembles the model proposed for the human class I histocompatibility antigen, HLA-A2, by Bjorkman et al. (1987).

ACKNOWLEDGMENTS

We thank Prof. M. Fridkin from the Organic Chemistry Department for synthesis of Gly-Ala-Trp, Dr. S. Weinstein for help in the HPLC analysis, and Ms. I. Lehavot for the sequencing data. We are grateful to Dr. A. Wlodawer and Dr. D. Davies for valuable discussions and to Prof. B. Matthews for making available to us results of his work before publication and his refined coordinates for comparison.

Registry No. α-Chymotrypsin, 9004-07-3.

REFERENCES

Bernstein, F. C., Koetzel, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Schimanouchi, T., & Tasunmi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1987) *Nature 329*, 512–518.

Blevins, R. A., & Tulinsky, A. (1985) *J. Biol. Chem. 260*, 8865–8872.

Cohen, G. H., Silverton, E. W., Matthews, B. W., Braxton, H., & Davies, D. R. (1969) *J. Mol. Biol. 44*, 129–141.

Cohen, G. H., Silverton, E. W., & Davies, D. R. (1981) *J. Mol. Biol. 148*, 449–479.

Corey, R. B., Battfay, O., Brueckner, D. A., & Mark, F. G. (1965) *Biochim. Biophys. Acta 94*, 535–545.

Desnuelle, P. (1960) in *The Enzymes* (Boyer, P. D., Lardy, H., & Myrback, K., Eds.) 2nd ed., Vol. 4, pp 93–118, Academic Press, New York.

Dixon, M. M., & Matthews, B. W. (1989) *Biochemistry 28*, 7033–7038.

Finzel, B. C. (1987) *J. Appl. Crystallogr. 20*, 53–55.

Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 397–418.

Harel, M., Su, C.-T., Frolow, F., Ashani, Y., Silman, I., & Sussman, J. L. (1990) Abstracts of the 40th Meeting, American Crystallography Association, New Orleans, April 1990, p 61.
Henderson, R. (1970) *J. Mol. Biol. 54*, 341–354.
Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.
Hope, H. (1988) *Acta Crystallogr. B44*, 22–26.
Huber, R., & Bode, W. (1978) *Acc. Chem. Res. 11*, 114–121.
James, M. N. G., Sielecki, A. R., Brayer, G. D., & Delbaere, L. T. J. (1980) *J. Mol. Biol. 144*, 43–88.
Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.
Kunitz, M. (1938) *J. Gen. Physiol. 22*, 207–237.
Matthews, B. W., Sigler, P. B., Henderson, R., & Blow, D. M. (1967) *Nature 214*, 652–656.
Pflugrath, J. W., Saper, M. A., & Quiocho, F. A. (1984) in *Methods and Applications in Crystallographic Computing* (Hall, S., & Ashida, T., Eds.) pp 404–407, Clarendon Press, Oxford.
Ringe, D., Seaton, B. A., Gelb, M. H., & Abeles, R. H. (1985) *Biochemistry 24*, 64–68.
Ringe, D., Mottonen, J. M., Gelb, M. H., & Abeles, R. H. (1986) *Biochemistry 25*, 5633–5638.
Segal, D. M., Powers, J. C., Cohen, G. H., & Davies D. R. (1971) *Biochemistry 10*, 3728–3738.
Sheriff, S. (1987) *J. Appl. Crystallogr. 20*, 55–57.
Sielecki, A. R., Hendrickson, W. A., Broughton, C. G., Delbaere, L. T. J., Bayer, G. D., & James, M. N. G. (1979) *J. Mol. Biol. 134*, 781–804.
Tsukada, H., & Blow, D. M. (1985) *J. Mol. Biol. 184*, 703–711.
Yapel, A., Han, M., Lumry, R., Rosenberg, A., & Shiao, D. F. (1966) *J. Am. Chem. Soc. 88*, 2573–2584.

Identification of the Disulfide Bond Pattern in Albolabrin, an RGD-Containing Peptide from the Venom of *Trimeresurus albolabris*: Significance for the Expression of Platelet Aggregation Inhibitory Activity†

Juan J. Calvete,‡ Wolfram Schäfer,‡ Tomasz Soszka,§ Weiqi Lu,§ Jacquelynn J. Cook,§ Bradford A. Jameson,∥ and Stefan Niewiarowski*,§

Max Planck Institut für Biochemie, D-8033 Martinsried/München, Federal Republic of Germany, and Department of Physiology, Thrombosis Research Center, School of Medicine, and Fels Research Institute, Temple University, Philadelphia, Pennsylvania 19140

Received November 28, 1990; Revised Manuscript Received February 14, 1991

ABSTRACT: Albolabrin is a 73 amino acid peptide isolated from the venom of *Trimeresurus albolabris*. It contains an RGD sequence and 12 cysteines and is a potent inhibitor of both platelet aggregation and fibrinogen binding to the GPIIb/IIIa complex. This protein shows a high degree of analogy (primarily due to the alignment of all cysteines and the RGD) with a number of other viper venom proteins which inhibit cell adhesion and platelet aggregation and are referred to as disintegrins: rhodostomin, trigramin, flavoridin, applagin, elegantin, and batroxostatin. In this study, we found that the reduction and vinylpyridylethylation of albolabrin and flavoridin decreased their platelet aggregation inhibitory activity approximately 40–50 times. It can be postulated that the higher potency of native and reduced flavoridin as compared to albolabrin depends on the substitution of the Asp of albolabrin with a Phe at the C-terminal end of the RGD in flavoridin. The activity of a synthetic C-terminal peptide derived from flavoridin (residues 35–65) containing four cysteines was about 75 times lower than that of the original flavoridin. The substitution of a pair of cysteine residues with alanines in this peptide resulted in further loss of activity. In order to identify the disulfide bonds in albolabrin, the molecule was digested consecutively by trypsin and porcine pancreatic elastase. Peptides resulting from this digestion were isolated by reverse-phase HPLC and identified by amino acid composition and mass spectrometry. Direct evidence for the existence of a linkage between cysteine-10 and cysteine-12 and between cysteine-6 and cysteine-11 was obtained, and indirect evidence suggested links between cysteines-1–3, -2–4, -5–8, and -7–9. On the basis of this information, we propose a model for albolabrin which may be appropriate for the structure of the entire disintegrin family.

It is well established that short RGDX peptides, representing the cell recognition site of a number of adhesive proteins such as fibronectin, fibrinogen, vitronectin, von Willebrand factor, and thrombospondin, prevent cell adhesion and platelet aggregation (Ruoslahti & Pierschbacher, 1987). The concentrations of these peptides required to inhibit platelet aggregation range from 10 to 200 μM. Huang et al. (1987a, 1989) isolated a potent platelet aggregation inhibitor from the snake venom of the *Trimeresurus gramineus* snake. This peptide, referred to as trigramin, is composed of a single polypeptide chain of 72 amino acid residues; it contains the RGD sequence and 6 disulfide bridges. Trigramin blocks platelet aggregation and binding of ^{125}I-fibrinogen to ADP-stimulated platelets ($K_i = 2 \times 10^{-8}$ M). It binds to the glycoprotein IIb/IIIa complex on the platelet surface (Huang et al. 1987a), and it blocks the adhesion of cultured human melanoma cells and fibroblasts

† This work was supported by NIH Grants HL15226 and HL36579 (to S.N.) and BMFT (Bonn) Grant PTB 038824 (to W.S.). The provision of a Max-Planck-Gesellschaft fellowship to Juan J. Calvete is gratefully acknowledged. B.A.J. is a scholar of the American Foundation for Aids Research.

* Correspondence should be addressed to this author at the Department of Physiology, Temple University School of Medicine, Philadelphia, PA 19140.
‡ Max Planck Institut für Biochemie.
§ Department of Physiology, Temple University.
∥ Fels Research Institute, Temple University.

0006-2960/91/0430-5225$02.50/0