PROTEINS: Structure, Function, and Genetics 8:82–101 (1990)

The Three-Dimensional Structure of Recombinant Bovine Chymosin at 2.3 Å Resolution

Gary L. Gilliland,[1,2] Evon L. Winborne,[3] Joseph Nachman,[2] and Alexander Wlodawer[2,4]

[1]Center for Advanced Research in Biotechnology of the Maryland Biotechnology Institute, University of Maryland, Shady Grove, Rockville, Maryland, 20878, [2]Center for Chemical Technology, National Institute of Standards and Technology, Gaithersburg, Maryland, 20899, [3]Genex Corporation, 10620 Industrial Drive, Gaithersburg, Maryland, 20877, and [4]Crystallography Laboratory, NCI-Frederick Cancer Research and Development Center, ABL-Basic Research Program, P.O. Box B, Frederick, Maryland, 21701

ABSTRACT **The crystal structure of recombinant bovine chymosin (EC 3.4.23.4; rennin), which was cloned and expressed in *Escherichia coli*, has been determined using X-ray data extending to 2.3 Å resolution. The crystals of the enzyme used in this study belong to the space group *I*222 with unit cell dimensions $a = 72.7$ Å, $b = 80.3$ Å, and $c = 114.8$ Å. The structure was solved by the molecular replacement method and was refined by a restrained least-squares procedure. The crystallographic *R* factor is 0.165 and the deviation of bond distances from ideality is 0.020 Å. The resulting model includes all 323 amino acid residues, as well as 297 water molecules. The enzyme has an irregular shape with approximate maximum dimensions of 40 × 50 × 65 Å. The secondary structure consists primarily of parallel and antiparallel β-strands with a few short α-helices. The enzyme can be subdivided into N- and C-terminal domains which are separated by a deep cleft containing the active aspartate residues Asp-34 and Asp-216. The amino acid residues and waters at the active site form an extensive hydrogen-bonded network which maintains the pseudo 2-fold symmetry of the entire structure. A comparison of recombinant chymosin with other acid proteinases reveals the high degree of structural similarity with other members of this family of proteins as well as the subtle differences which make chymosin unique. In particular, Tyr-77 of the flap region of chymosin does not hydrogen bond to Trp-42 but protrudes out in the P1 pocket forming hydrophobic interactions with Phe-119 and Leu-32. This may have important implications concerning the mechanism of substrate binding and substrate specificity.**

Key words: chymosin, acid proteinase, rennin, X-ray structure, structure comparison, catalytic site, crystal packing

INTRODUCTION

Chymosin (EC 3.4.23.4; rennin), the main enzymatic component of calf rennet, is one of the primary milk clotting enzymes used for cheese production. The enzyme specifically cleaves the Phe-105–Met-106 peptide bond of κ-casein[1] to initiate the milk clotting process.[2] The chymosin molecule contains a single chain of 323 amino acid residues with three disulfide bridges, and its molecular weight is 35,000.[3] At least two isozymes of this enzyme, chymosin A and chymosin B, are known; A has Asp and B has Gly at position 244. As with other pancreatic enzymes, chymosin is secreted as an inactive zymogen which is later activated by the removal of 42 N-terminal residues.

Chymosin is a member of a family of enzymes known as aspartic proteinases for which two subclasses of proteins have been identified. They can be distinguished by their relative sizes and their source of origin, being either eukaryotic or retroviral. The eukaryotic enzymes have a single polypeptide chain with two structurally similar domains which pack to form a deep active site cleft. In contrast, the retroviral enzymes are dimers of two identical polypeptide chains each of which is somewhat smaller than one half of the eukaryotic enzyme. The dimers of the retroviral proteinases pack to form the active site in a manner which is similar to the arrangement of the two domains of the eukaryotic enzyme. The second-

Received July 18, 1989; revision accepted January 30, 1990.

Address reprint requests to Dr. Gary L. Gilliland, Center for Advanced Research in Biotechnology of the Maryland Biotechnology Institute, University of Maryland, Shady Grove, 9600 Gudelsky Drive, Rockville, MD 20878.

Evon L. Winborne's current address is Department of Biochemistry, Biological Science, West, University of Arizona, Tucson AZ 85721.

Joseph Nachman's current address is Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Quebec H4P 2R2, Canada.

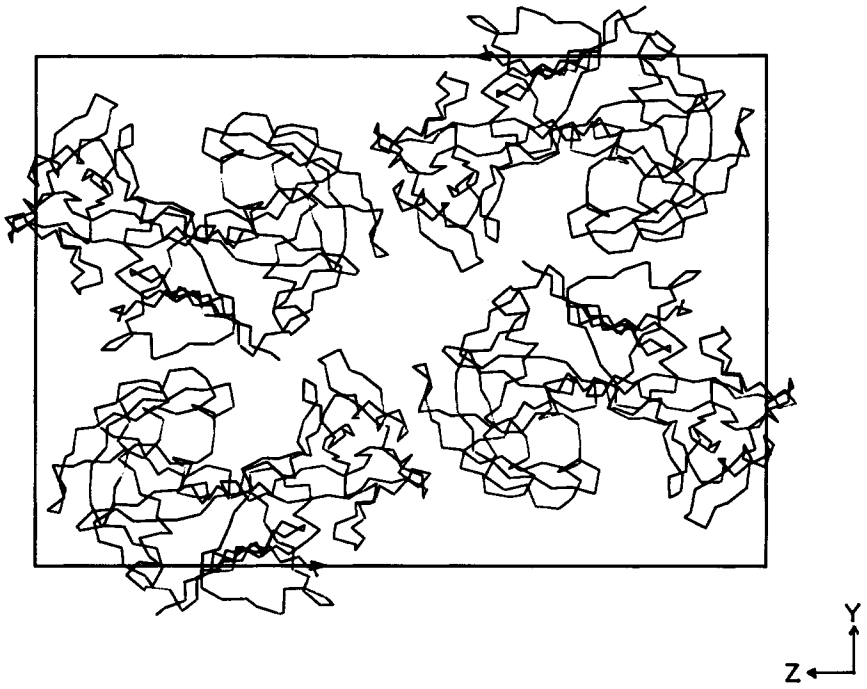

Fig. 1. The crystallographic packing of the recombinant chymosin molecule shown in projection looking down the crystallographic *x* axis. There is a difference in the assignment of the crystallographic axes between this work and that shown in Figure 1 of Safro et al.[26]

ary structure of both subclasses of enzymes is composed almost entirely of β-sheets. This family of enzymes is further characterized by the presence of two aspartyl groups, one contributed from each domain or each monomer (in the case of the retroviral proteinases), in the large active site cleft which is situated between the domains or monomers.

Several structures of single-chain proteinases of mammalian and fungal origin have already been solved and/or refined to high resolution (for a current review see James and Sielecki[4]). They include porcine pepsin,[5–7] porcine pepsinogen,[8] penicillopepsin,[9,10] endothiapepsin,[11–15] rhizopuspepsin,[12,16–18] and human renin.[19] All of these enzymes have a high degree of sequence and structural homology with one another (e.g., Subramanian et al.[12]) and display internal symmetry which is indicative of gene duplication.[20,21] Two reports of retroviral proteinase structures, the Rous sarcoma virus proteinase[22] and HIV-1 proteinase,[23] have recently been reported.

Crystallographic studies of chymosin isolated from calf rennet have been underway for a number of years. Berridge[24] crystallized chymosin from solutions containing high concentrations of sodium chloride. Crystals of that form were characterized and an X-ray multiple isomorphous replacement structure determination was undertaken by Bunn and co-workers,[25] but it was never completed. More recently, the structure solution using crystals of the natural enzyme has been attempted using molecular replacement with endothiapepsin[11] and bovine pepsin[26] as probes. In this report we describe the molecular replacement solution and the structure of recombinant chymosin B at 2.3 Å resolution, and compare the structure with other members of the aspartic proteinase family.

EXPERIMENTAL PROCEDURES

Enzyme and Reagents

Purified recombinant chymosin was supplied by scientists at Genex Corporation (C. Vaslet, D. Anderson, D. Scandella, D. Robinson, and J. McGuire, unpublished data). Prior to the crystallization trials, the enzyme was concentrated to 8.6 mg/ml in 0.05 M 2-[*N*-morpholino]ethanesulfonate (MES) buffer at pH 6.0 using an Amicon ultrafiltration, stirred cell with a 10,000 molecular-weight cut-off membrane. All chemicals used in the crystallization experiments were reagent grade.

Crystallization

Crystals were grown by the vapor diffusion method using the hanging-drop procedure.[27] The crystals were grown in Linbro plastic tissue culture plates (Stock. No. 76-033-05) in which 10 μl droplets of protein solution suspended from 22 mm in diameter, round microscope-slide coverslips were equilibrated with 1.0 ml of reservoir solution. Cargille type-B immersion oil for microscopy was used to seal

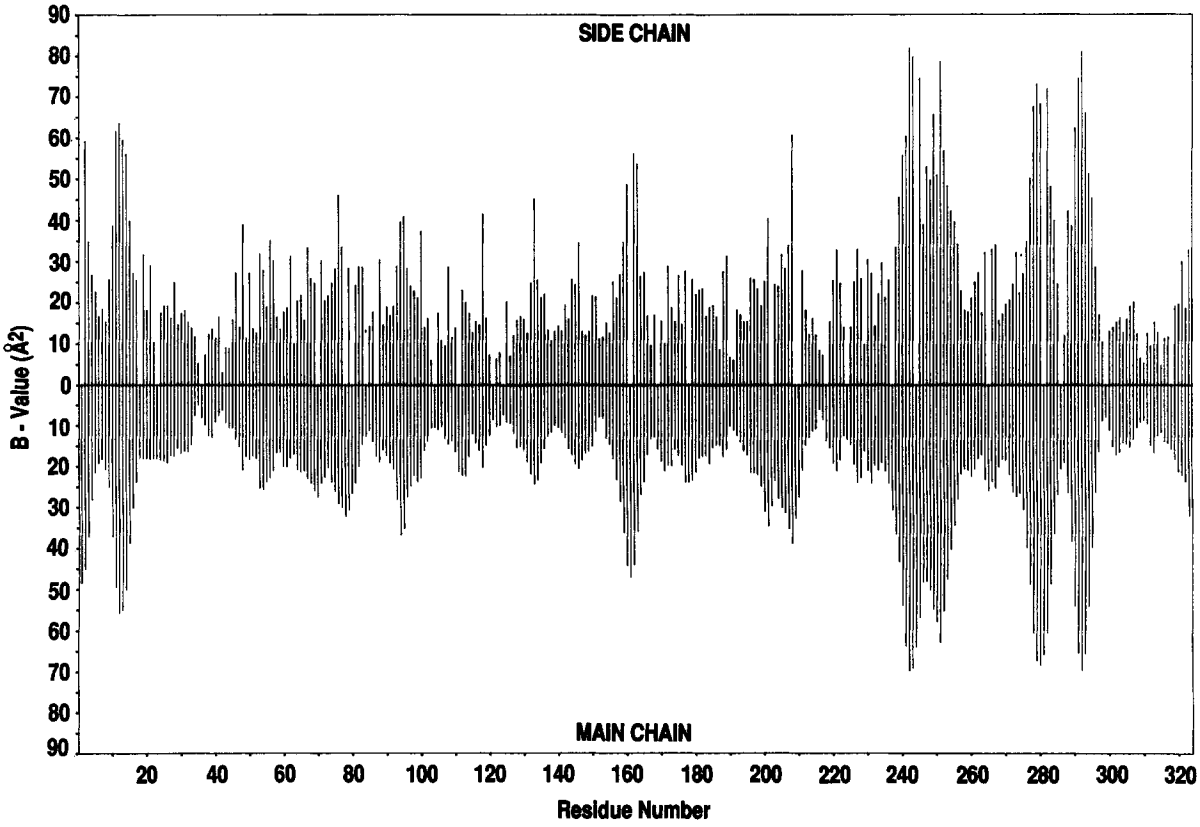

Fig. 2. Temperature factors averaged for the side chain (above center line) and for the main chain (below center lines) for the model of recombinant chymosin.

the tissue culture plate chambers. Crystallization trials were performed at room temperature, 18 to 20°C. All glass surfaces used in the crystallization and seeding experiments were siliconized using SCM Specialty Chemicals Prosil 28. Small well-formed crystals were obtained in droplets equilibrated against 25–35% saturated sodium chloride.

Macroseeding

The macroseeding procedure employed the following protocol, which is similar to that described by others.[28,29] A single seed selected from a protein droplet containing many small protein crystals was transferred to a depression slide well containing 100 μl of 60% saturated sodium chloride at pH 6.0. After soaking for 60–90 seconds in this solution, the seed crystal was transferred to a second depression containing a solution of 50% saturated sodium chloride. After the same waiting period, the crystal was transferred to a 40% saturated sodium chloride solution. The seed crystal was finally transferred to the droplet containing the enzyme dissolved in 45–50% saturated sodium chloride and 0.05 M MES at pH 6.0. Droplets were suspended over well solutions containing 45–55% saturated sodium chloride in 0.05 M MES at pH 6.0. Crystal transfer utilized small siliconized-glass capillary pipettes. The amount of liquid associated with each transfer was kept to a minimum, 0.5–2.0 μl. All operations were viewed through a dissecting microscope.

X-Ray Data Collection

All data used in this study were collected from a single crystal which was shaped as a parallelpiped with the dimensions of 0.20 × 0.15 × 0.1 mm. Diffraction data were collected using a Siemens electronic area detector. This instrument was mounted on a Supper oscillation camera controlled by a Cadmus 9000 microcomputer. X Rays were generated with an Elliot GX-21 rotating anode, operated at 40 kV and 70 mA, with a 0.3 × 3 mm focal spot. A graphite monochromator followed by a 0.3 mm collimator were utilized. During data collection, the area detector chamber was mounted 10 cm from the crystal. The carriage angle was set at 20° intercepting data from infinity to 2.3 Å. Diffraction data collected with the area detector were recorded as a series of discrete frames or electronic images, each comprising a 0.25° oscillation counted for 2 minutes. The individual frames were contiguous in that the beginning of each small oscillation range coincided with the end of the previous range. The raw data

frames were transferred from the Cadmus 9000 to a Digital Equipment Corporation VAX 11/780 for final processing. All data collection was done at well-controlled room temperature (19–20°C).

X-Ray Data Processing

The determination of crystal orientation and the integration of reflection intensities were performed with the XENGEN program system.[30] A total of 34,846 observations of 15,628 unique reflection intensities were measured. A weighted least-squares R factor on intensity for symmetry-related observations ($R_\omega = \Sigma[I_{ij} - G_{ij} < I >_j)/\sigma_{ij}]^2/\Sigma(I_{ij}/\sigma_{ij})^2$, where $G_{ij} = g_i + A_i S_j + B_i s_j^2$; $s = \sin\theta/\lambda$; g, A, and B are scaling parameters) was 0.048, while the unweighted absolute R factor on intensities ($R_u = \Sigma(I_{ij} - G_{ij} < I >_j)/\Sigma \mid I_{ij} \mid$) was 0.068, 12,448 of the measured unique reflections had significant intensity [$I > 1.5\sigma(I)$]. A statistical summary of the final processing results of the crystallographic data is presented in Table I. The space group and unit cell parameters were determined using only the area detector data and the XENGEN program system.[31] The crystals of recombinant chymosin are isomorphous with crystals of the natural enzyme.[25] The space group is $I222$ with $a = 72.7$ Å, $b = 80.3$ Å, and $c = 114.8$ Å. The unit cell parameters reported here are actually a permutation of those reported earlier.

Structure Solution

The structure of recombinant chymosin was solved by using molecular replacement methods. The orientation of the molecule within the unit cell was found by independently performing two basically different rotation functions: Crowther's fast rotation function[32] as implemented in the molecular replacement program package MERLOT,[33] and the real space version of the Patterson search technique[34] as implemented in the program package PROTEIN.[35]

In both cases the search model was based on the atomic coordinates of the partially refined crystal structure of bovine pepsinogen (J. Remington, unpublished data). In order to minimize the errors due to the differences between the structure of pepsinogen and that of chymosin, residues 48–49, 114, 130–132, 158–160, 187–188, 200–209, 230–251, 280–282, and 293–299 were excluded from the model. These residues were omitted after a comparison of the three-dimensional structures of pepsinogen and the acid proteinase of *Rhizopus chinensis*[18] using Hendrickson's program TOSS[36] to align the molecules, and aligning the sequences of these two molecules with that of chymosin.

The fast rotation function searches were performed using data within various resolution limits, the upper one being 4.0 Å, and in steps of 2.5° in α, and 5.0° in β and γ, where α, β, and γ are Euler angles as defined by Crowther.[32] None of the searches produced very clear maps. However, careful examination of the corresponding lists of peaks with intensities higher than 75% of the maximum peak height revealed that only one peak, located at $\alpha = 100.0°$, $\beta = 78.0°$, and $\gamma = 177.0°$, appeared in all searches. This peak is the highest one in the search using data between 5.0 and 8.0 Å resolution, where its intensity is 4.0 rms units above background. This result was further confirmed by repeating the searches with two modified models: one, using only the C^α coordinates of the pepsinogen fragment, and the second, using the main chain atoms and C^β coordinates. Although in both cases the correct solution was a relatively high peak in the map, it was not the most significant.

The real space rotation search, performed in steps of 5.0° in all three Euler angles using data between 3.5 and 15.0 Å resolution, produced a map with a peak of 7.0 rms units above average at $\alpha = 99.2°$, $\beta = 77.8°$, and $\gamma = 180.0°$. The intensity of the next highest peak was 5.0 rms units, i.e., about 70% of the maximum peak height. Obviously, this result is practically identical to that obtained with the fast rotation function.

The position in the unit cell of the model resulting from the rotation search was found by using a "brute force" translation program TF (J. Remington, unpublished program). The search model was built from the coordinates of the pepsinogen fragment, rotated through the set of Euler angles (99.0°, 78.0°, and 180.0°). The translation search was performed on a grid of 1.0 Å in each direction, using data between 5.0 and 10.0 Å resolution. The resulting map had a maximum of 8.8 rms units, with a correlation coefficient of 0.32, located at $x = 0.234$, $y = 0.361$, and $z = 0.235$ (in fractions of unit cell dimensions); the next highest peak was 4.3 rms units.

To test this result, a search using Crowther and Blow's translation function,[37] as implemented in the program package MERLOT,[33] was performed using resolution data between 5.0 and 10.0 Å in grid steps of 0.02 fractions of unit cell dimensions in all three directions; this corresponds to steps of 1.45 Å in x, 1.61 Å in y, and 2.30 Å in z. The test model was the same as used with the program TF. In space group $I222$ the intermolecular vectors occur on the Harker sections $\mathbf{u} = 0.0$, $\mathbf{v} = 0.0$, and $\mathbf{w} = 0.0$; hence, the search was performed only over these three sections, rather than over the whole asymmetric unit. The search yielded three mutually consistent sets of vectors, corresponding to $2x = 0.53$, $2y = 0.29$, and $2z = 0.48$. Since the origin of each axis can be chosen arbitrarily at 0.0 or 0.5, there are 8 molecular positions that satisfy this relationship; one of them, at $x = 0.765$, $y = 0.645$, and $z = 0.24$, is related by crystallographic symmetry to the highest peak found using the program TF.

When the packing of molecules in the unit cell shown in Figure 1 is compared with that seen in

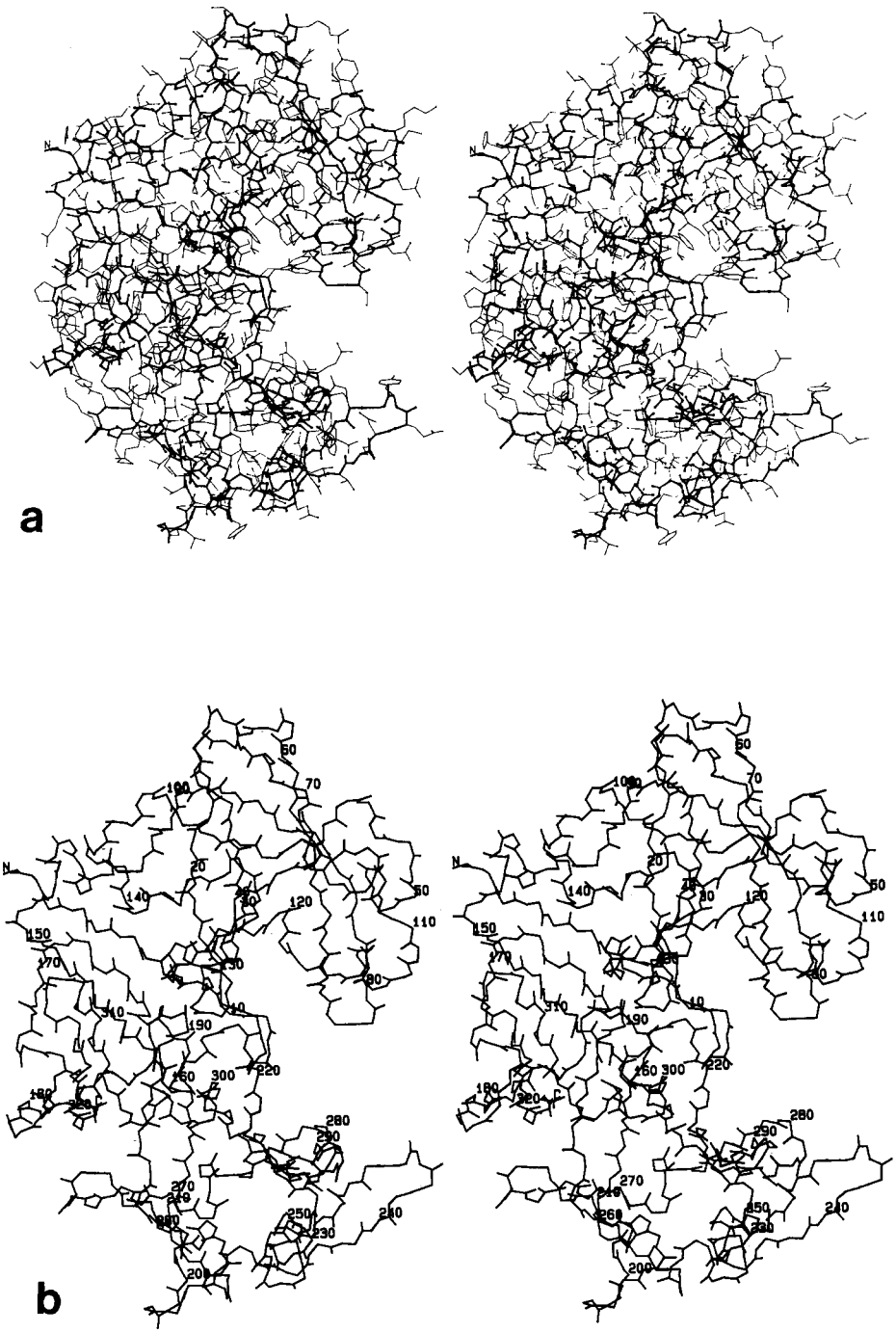

Fig. 3. **(a)** A stereoscopic view of all 2511 atoms of the recombinant chymosin molecule. The main chain atom bonds between N, C^{α}, C, and O are drawn in thick lines, and the side chain bonds are drawn with thin lines. **(b)** A stereoscopic view of the backbone atoms, N, C^{α}, C, and O of the recombinant chymosin molecule with the first (1), last (323), and every tenth residue labeled. **(c)** A *RASTER 3D* image[54] of the recombinant chymosin molecule: green arrows indicate β-strands, blue barrels represent helices, grey 'stove pipe' represents connecting backbone, and the two active site aspartates are shown in red.

Figure 1 of Safro and co-workers,[26] it is clear that the rotation is very similar, while the translation differs by a shift of about 0.25 along *b*. This may explain why this earlier model could not be refined.

Restrained Least-Squares Refinement

Structure refinement was performed with the restrained least-squares procedure of Hendrickson,[38]

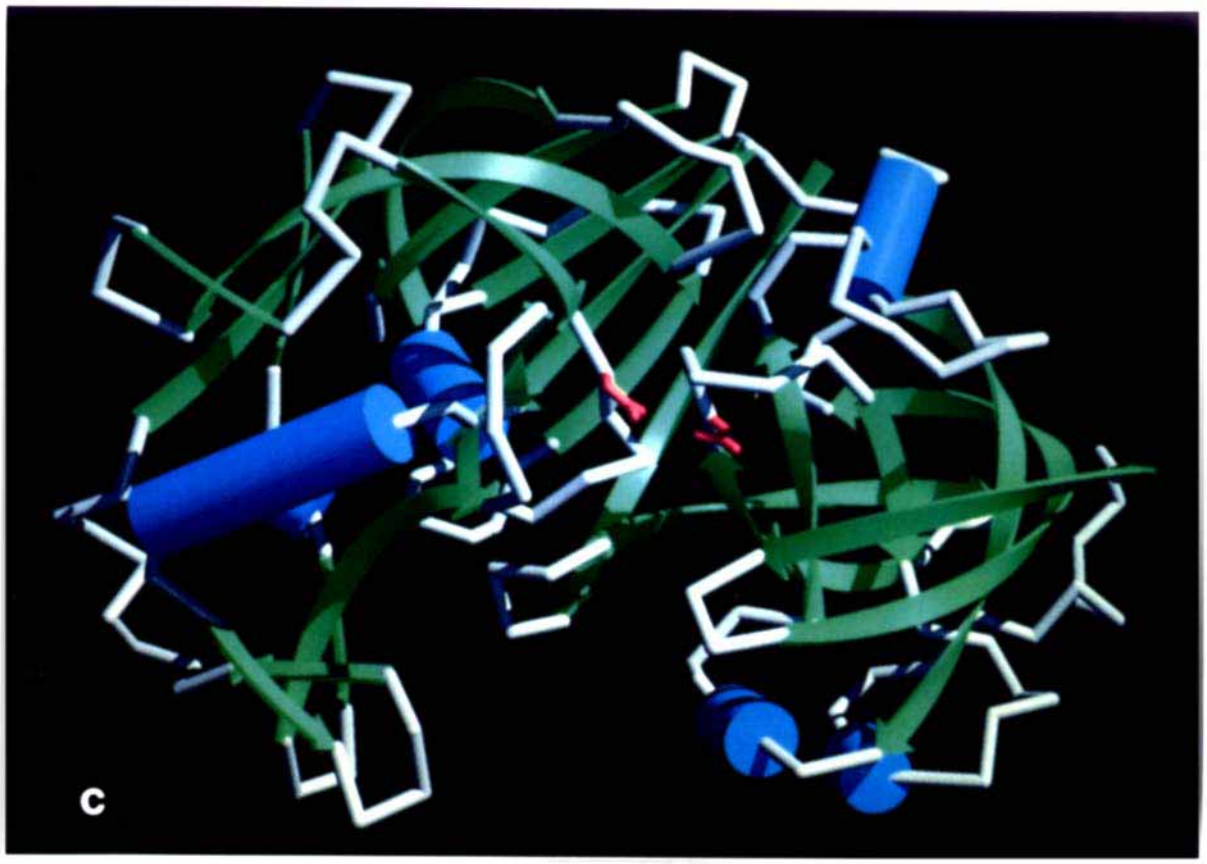

Fig. 3c.

TABLE I. Statistical Summary of Crystallographic Data for Recombinant Bovine Chymosin

Shell lower limit (Å)	Average reflection I	Average $I/\sigma(I)$	Number of Bragg reflections				R_w*	R_{uw}†
			Possible	Collected	Observed	$>2\sigma(I)$		
4.12	2,880	54.0	2,794	2,791	7,067	2,674	3.71	3.20
3.27	2,077	27.1	2,672	2,672	6,824	2,476	5.00	4.49
2.85	945	11.2	2,662	2,662	6,154	2,189	8.67	9.23
2.59	529	5.9	2,618	2,603	5,518	1,910	14.85	16.48
2.41	401	4.0	2,629	2,529	5,067	1,631	19.13	20.56
2.27	372	3.4	2,609	2,471	4,216	1,558	20.34	23.14
Totals	1,291	19.3	15,984	15,628	34,846	12,448	4.75	6.83

*R_w is the weighted-squared R factor on intensity × 100 (full definition in the text).
†R_{uw} is the unweighted absolute value R factor on intensity × 100 (full definition in the text).

which has been modified by Finzel[39] to incorporate the fast Fourier transform algorithms of Ten Eyck[40] and Agarwal[41] and by Sheriff[42] to restrain contacts between atoms belonging to different molecules in the unit cell. Diffraction data with $I > 1.5\sigma(I)$ and between 20.0 and 2.3 Å resolution were included in the calculations. The program FRODO[43] was used on an Evans and Sutherland PS300 graphics system to examine $2F_o - F_c$ and $F_o - F_c$ difference maps, as well as complete omit maps,[4] to adjust the model, and to add solvent molecules.

Solvent structure was analyzed in the difference maps, and waters with high temperature factors ($B > 38.0$ Å^2), low occupancies ($occ > 0.25$), or those not clearly visible in the $F_o - F_c$ maps were removed from the model during the course of the refinement, while new waters were added where indicated by map peaks. At the end of the refinement waters were renumbered based on the square of the occupancies divided by the temperature factors (occ^2/B).[10] The waters were renumbered sequentially from the largest to the smallest values starting with the number 401 and ending with 697.

Altogether 356 cycles of least-squares refinement were performed interspersed with 19 calculations of electron density maps followed by model adjust-

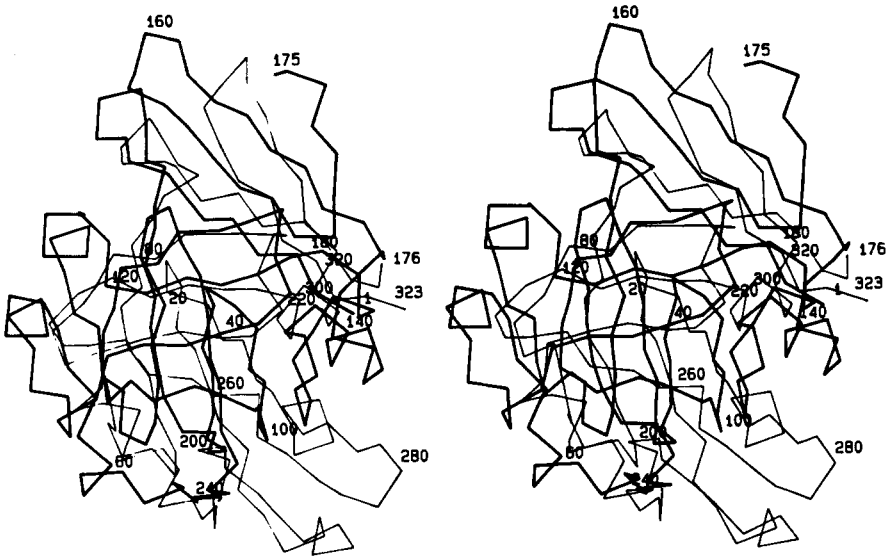

Fig. 4. A stereoscopic view of the C^{α} carbons of the C-terminal domain (amino acid residues 1–175) superimposed on those of the N-terminal domain (amino acid residues 176–323). The bonds connecting the C^{α} carbons of the N-terminal domain are drawn with thick lines, and the bonds connecting the C^{α} carbons of the C-terminal domain are drawn with thin lines.

ment. The final model consists of 2511 nonhydrogen atoms, which includes atoms of all 323 amino acid residues and of 297 water molecules (oxygen atoms only). The final crystallographic R factor of the structure of recombinant chymosin resulting from this work is 0.165 at 2.3 Å resolution, with a root-mean-square deviation of bonded distances from ideality of 0.020 Å. A summary of the refinement statistics is presented in Table II. Although we have not attempted to estimate the error in coordinates, it is greater than the 0.12 and 0.16 Å estimated for the 1.8 Å structures of penicillopepsin[10] and rhizopuspepsin,[18] respectively. This is a result of the lower resolution data obtained from the recombinant chymosin crystals. The coordinates have been deposited as data set 1CMS with the Brookhaven Protein Data Bank.[45]

TABLE II. Summary of Least-Squares Refinement Parameters for Recombinant Bovine Chymosin at 2.3 Å Resolution

	Target values	Final model
Reflections to 2.3 Å [$I > 1.5\ \sigma(I)$]	12410	
Crystallographic R factor*	0.165	
rms deviations from ideal distance (Å)		
Bond distance	0.020	0.020
Angle distance	0.030	0.044
Planar 1–4 distance	0.050	0.051
rms deviation from planarity (Å)	0.020	0.018
rms deviation from ideal chirality (Å^3)	0.150	0.217
Thermal parameter correlations (mean / ΔB /)		
Main chain bond	1.000	1.324
Main chain angle	2.000	2.168
Side chain bond	2.000	2.060
Side chain angle	3.000	3.187

*R factor $= \sum_{hkl} ||F_o| - |F_c|| / \sum_{hkl} |F_o|$.

RESULTS AND DISCUSSION

The Three-Dimensional Structure

The model of recombinant chymosin resulting from these diffraction studies contains the complete polypeptide chain of chymosin B with three disulfide bridges formed between Cys-47 and Cys-52, Cys-207 and Cys-211, and Cys-251 and Cys-283. The model contains a single *cis*-proline, Pro-25, and includes 297 water molecules. However, there are five regions of the polypeptide chain which have weak or poorly defined electron density in the final $2F_o - F_c$ map. These regions include residues 11–14, 159–163, 241–249, 278–282, and 288–294. All five of these polypeptide segments are in exposed loops on the surface of the protein and correspond to the regions of the molecule with the highest temperature factors (Fig. 2). Only two of these regions, 241–249 and 278–282, contain amino acid residues which are considered to be involved in crystal packing contacts. The interpretation of the model for these regions must be considered as tentative.

The chymosin molecule is kidney shaped and is composed of two domains with approximate overall dimensions of 40 × 60 × 65 Å. The folding pattern of the enzyme with its pseudo 2-fold symmetry[20] is typical for members of the acid proteinase family. The molecule is a predominantly β-structure protein with the N- and C-terminal domains divided by a

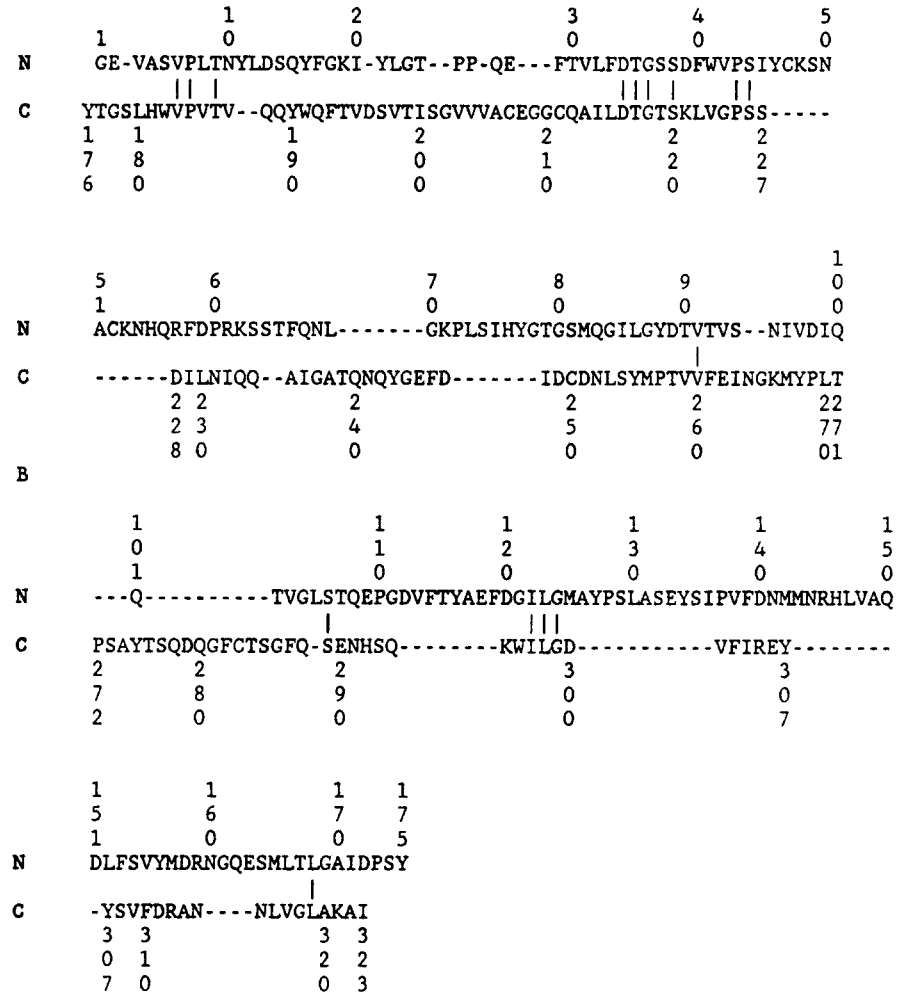

Fig. 5. Sequence alignment of the N-terminal (amino-acid residues 1–175) and C-terminal (amino-acid residues 176–323) domains of chymosin based on the least squares alignment of the corresponding regions of the three-dimensional structure of recombinant chymosin. A | indicates a topographically equivalent position with identical amino acid residues in both domains.

deep active-site cleft. The complete model for all nonhydrogen atoms of the enzyme, excluding the water molecules, is shown in Figure 3.

The β-structure of each domain is composed of both parallel and antiparallel β-strands. There are a few short segments of α-helices bridging β-strands in each domain. In general the β-strands of one domain have topographically equivalent strands in the other, but this is not true in the case of the α-helices. The secondary structure assignments, made according to Kabsch and Sander,[46] are listed in Table III and illustrated in Figure 3c.

The N- and C-terminal domains are composed of residues 1–175 and 176–323, respectively. The two domains have been aligned by a least-squares superposition using the program ALIGN.[47] The results indicate 87 pairs of topographically equivalent residues with an rms deviation of 2.3 Å between corresponding C^{α} carbons. The superposition of the C-terminal domain on the N-terminal domain is shown in Figure 4. An alignment of the sequences of the N- and C-terminal domains based on this structural alignment is shown in Figure 5.

The Catalytic Site

The two active site aspartates, Asp-34 and Asp-216, are found at the base of the deep cleft separating the N- and C-terminal domains. The side chains of these two aspartates are oriented in the same general way as those found in the active sites of other acid proteinases.[10,13,18] The side chain carboxylate groups are almost in the same plane and are oriented toward each other with solvent molecule Wat-411 disposed between them (see Fig. 6). The nearly exact 2-fold symmetry of the acid proteinase active site is maintained except that the plane of the carboxylate group of Asp-216 is not parallel to the peptide bond between residues Gly-218 and Thr-219 as

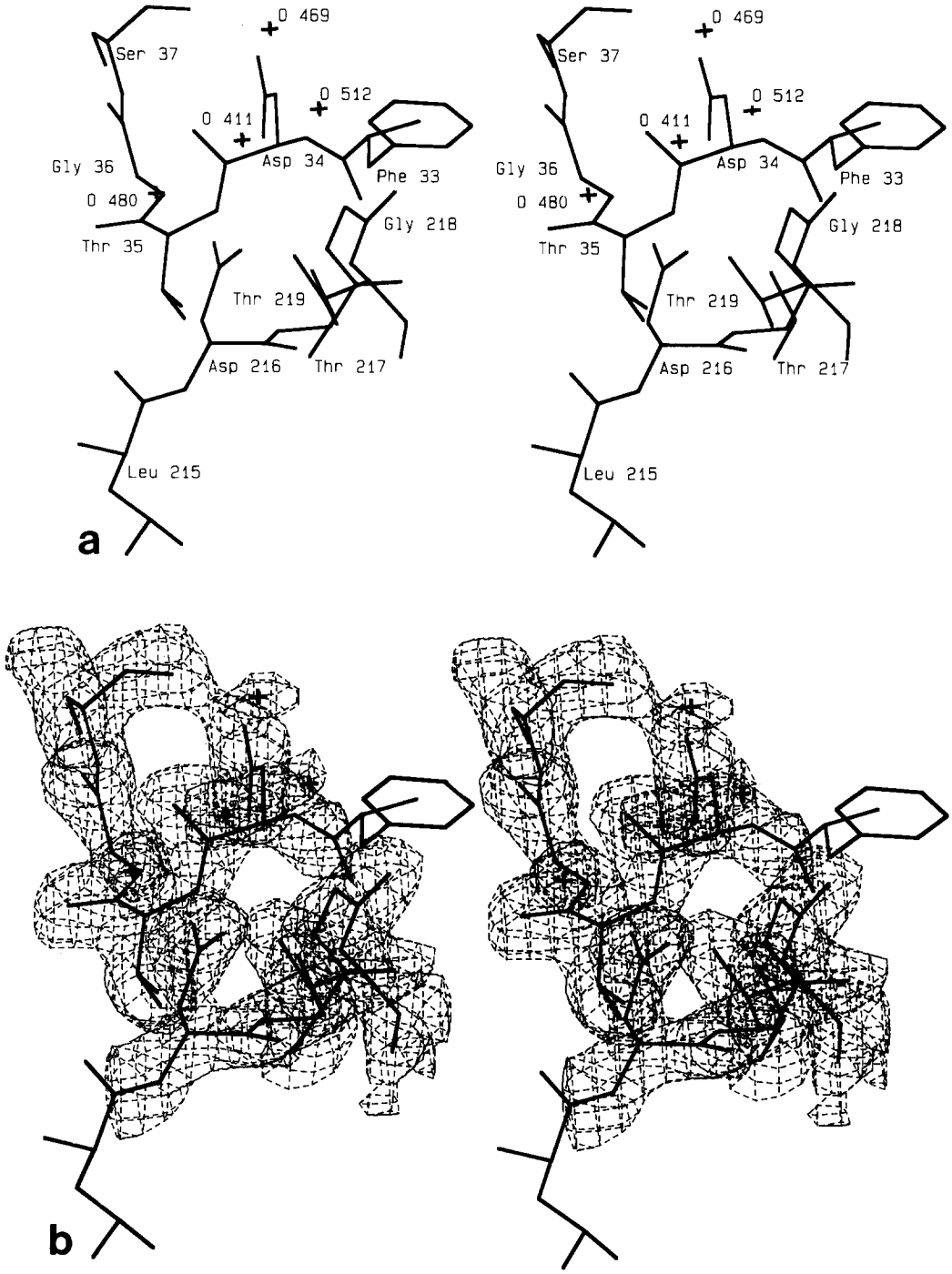

Fig. 6. Two different representations of the recombinant chymosin active site: **(a)** all residues and solvent molecules shown are labeled, and **(b)** the electron density of the $2F_o-F_c$ map contoured at 1σ with the model superimposed (the labels are omitted for clarity).

it is with the corresponding carboxylate group of Asp-34 and the peptide bond between residues Gly-36 and Ser-37. The primary interactions of Asp-34 and Asp-216, described in Table IV, are very similar. Asp-34 OD1 is within 2.4 Å of the OG atom of Ser-37 and Asp-216 OD2 is within 2.9 Å of Thr-219 OG. Asp-34 and Asp-216 are symmetrically disposed about Wat-411; however, Wat-411 is associated closer to Asp-216 with distances of 3.3 and 2.9 Å to OD1 and OD2, respectively. The corresponding distances to OD1 and OD2 of Asp-34 are 3.1 and 3.8 Å.

The positioning of the two active site aspartate

residues is the result of the interaction of two structurally homologous segments of the polypeptide chain composed of residues 33–37 with the sequence Phe-Asp-Thr-Gly-Ser for the N-terminal domain and residues 215–219 with the sequence Leu-Asp-Thr-Gly-Thr for the C-terminal domain. These polypeptide chain segments both form Type I β-turns.[48] The interaction of these two loops forms the "Fireman's grip," previously described for endothiapepsin by Pearl and Blundell.[13] This structure is formed by the hydrogen bonding of two threonine residues, one from each domain, with the main chain atoms associated or adjacent to the threonine in the opposing domain.

The two threonines, Thr-35 and Thr-217, follow in sequence the two active site aspartates and are positioned right around the pseudo-2-fold axis of the molecule. The main chain carbonyl oxygen atoms of these two residues hydrogen bond to solvent molecules in the active-site cleft; Wat-415 is 2.8 Å from the Thr-35 O atom, and Wat-531 is 2.4 Å from Thr-217 O atom. However, the side chains of these two residues are virtually buried within the hydrophobic interior of the chymosin molecule; Thr-35 has a completely inaccessible side chain, and Thr-217 has an accessible surface area of only 7 Å^2. These two side chains project into a hydrophobic pocket formed by Phe-153, Val-155, Met-157, Trp-191, Leu-215, Ile-303, and Phe-310. There are no solvent interactions with the peptide backbone or of the OG1's of these two residues. Hydrogen bonds are formed between the two chain segments by the OG1 hydrogen atoms of Thr-35 and Thr-217 interacting with the carbonyl oxygens of Leu-215 and Phe-33, the corresponding

TABLE III. Secondary Structure Assignments of Recombinant Chymosin*

```
Sequence #          10          20          30          40          50          60
Sequence    GEVASVPLTN  YLDSQYFGKI  YLGTPPQEFT  VLFDTGSSDF  WVPSIYCKSN  ACKNHQRFDP
Summary      B EEEE EE  ETTTEEEEEE  EETTTTEEEE  EEEETT   E  EE BTT  SH  HHHTS  B G
3-Turn                  >33<                    >33<        >33< >3  3X33<   >>
4-Turn                  >444<                               >>  44<<
5-Turn                              >5555<
Bend                    SSSS         SSS           SS          SS  SS  SSSSS    S
Chirality   ----------  --++-+--+-  -+-+---+--  ----+-+-+-  ----++++-+  ++++-----+
Bridge 1     a BBBB GG  G   GGG     II    HHHH  HH        K  K  o                o
Bridge 2                       HHHH  HH          jjjj        11
Sheet        A BBBB CC  C   CCCCCC  CC    CCCC  CCCC      C  CC D                D

Sequence #          70          80          90          100         110         120
Sequence    RKSSTFQNLG  KPLSIHYGTG  SMQGILGYDT  VTVSNIVDIQ  QTVGLSTQEP  GDVFTYAEFD
Summary     GG SS EEEE  EEEEEEETTE  EEEEEEEEEE  EEETTEEEEE  EEEEEEEE    SHHHHH SSS
3-Turn      3<<                 >33<            >33<                    >33<
4-Turn                                                                  >>>4<<<
5-Turn
Bend        SS SS               SS                SS    S       S       SSSSSS SSS
Chirality   ++-+++---+  -----+--+-  --+--+-+--  -+--+---++  -----++++  -++++++-+--
Bridge 1          MM     MMMMMM  M  MMMMM MM    II          11
Bridge 2                             NNNNNNN   NNN  NNN    NNNNNN N
Sheet             CC     CCCCCC  C  CCCCCCCCC  CCC  CCC    CCCCCC C

Sequence #          130         140         150         160         170         180
Sequence    GILGMAYPSL  ASEYSIPVFD  NMMNRHLVAQ  DLFSVYMDRN  GQESMLTLGA  IDPSYYTGSL
Summary     EEEE S GGG  S TT   HHH  HHHHTT BSS  SEEEEE  TT  SS  EEEES    GGGEEEEE
3-Turn           >>3<  <>33<  >33  < >33<              >33  <           >>3<<
4-Turn           >444  <      >>>>  X<<<<
5-Turn                              >5555<
Bend          S SSS  S SS   SSS  SSSSS  SSS  S        SS  SS      SS     SSS SS
Chirality   ----+--+++  --++---+++  ++++-+-+-  +--------+-  ++++----+-  --++++-+--
Bridge 1    jjjj                          a    CCCC          BBBB         EE
Bridge 2     KK                               DDDD          DDDD
Sheet       CCCC                          A    BBBBB         BBBB         BB

Sequence #          190         200         210         220         230         240
Sequence    HWVPVTVQQY  WQFTVDSVTI  SGVVVACEGG  CQAILDTGTS  KLVGPSSDIL  NIQQAIGATQ
Summary     EEEE SSBTT  EEEEEEEEES  SSSEEE TT   EEEEE TT S  SEEEEEHHHHH  HHHHHHT EE
3-Turn                              >33<        >33<        >33  <  >33<
4-Turn                                                      >>>>XX  XXX<<<<
5-Turn           >555  5<                                            >5555<
Bend           SSSSS          S  SSS  S SS        SS S  S    SSSSS  SSSSSS
Chirality   -----+-+++-  ++------+--  -+-+-+-+--  -+--+-++-+  --+--+++++  +++++-+---
Bridge 1    EEEE   P    P   R RRR     S S      QQQQQ         uuuu              XX
Bridge 2                 QQQQQ   SS              ttt          VV
Sheet       BBBB   E    EEEEE EEE     E E      EEEEE         EEEE              FF
```

(continued)

TABLE III. Secondary Structure Assignments of Recombinant Chymosin (Continued)*

```
Sequence #          250         260         270         280         290         300
Sequence    NQYGEFDIDC DNLSYMPTVV FEINGKMYPL TPSAYTSQDQ GFCTSGFQSE NHSQKWILGD
Summary      TTS EEE T TGGGG   EE EEETTEEEEE HHHHEEEET  TEEEESEEEE SSS  EEE H
3-Turn      >33<    >3 >X><<<    >33<          >3 3<                    >3
4-Turn                                      >>44<<                       >>
5-Turn
Bend         SSS     S S  SS        SS       SSSS    S S    S      SSS      S
Chirality   -+-+------+ +++++---  -+--+-----  -++++-+-- +----+---- -++-----+-+
Bridge 1        XX             R RRR  WWWWW       ZZZZ    YY  uuuu       ttt
Bridge 2         YY          WW WWW                     ZZZZ               VV
Sheet           FFF          EE EEE  EEEEE       FFFF   FFFF EEEE       EEE

Sequence #          310         320
Sequence    VFIREYYSVF DRANNLVGLA KAI
Summary     HHHTTEEEEE ETTTTEEEEE EE
3-Turn      3X33<
4-Turn      4><<4<     >444<
5-Turn                 >5555<
Bend        SSSSSS      SSS
Chirality   ++++------ +++-+---+- -
Bridge 1          CCC C    EEEE  EE
Bridge 2       FFFFF  F    FFFFF F
Sheet          BBBBB  B    BBBBB BB
```

*The structural assignments were done with the Kabsch and Sander program (Kabsch, W., Sander, C. Biopolymers. 22:2577–2637, 1983). Each of the rows in this table are interpreted as described below. The Sequence rows: the sequence is numbered from the N-terminus (1) and is given in 1-letter code. The Summary row: H, 4-helix (α-helix); B, residue in an isolated β-bridge; E, extended strand, participates in β-ladder; G, 3-helix (3_{10}-helix); I, 5-helix (π-helix); T, hydrogen bonded turn; S, bend. In cases where there is a structural overlap, priority is given to the first occurrence in the above list. The 5-Turn, 4-Turn, and 3-Turn rows: these rows indicate the hydrogen bonding pattern for turns and helices: >, backbone CO of this residue makes hydrogen bond (i, $i+n$); <, backbone NH of this residue makes hydrogen bond ($i-n$, n); X, both CO and NH make an hydrogen bond; 3, 4, and 5 indicate the residues bracketed by hydrogen bonds. The Bend row: S indicates the center of 5 residues with a curvature of at least 70°. The Chirality row: the sign, + or -, of the dihedral angle defined by C^{α}_{i-1} to C^{α}_{i-1}. The Bridge 1 and Bridge 2 rows: the one letter name of β-ladders in which residue i participates. A capital letter indicates an antiparallel β-ladder; a lowercase letter indicates a parallel β-ladder. The Sheet row: a letter indicating which sheet this residue participates.

TABLE IV. Hydrogen Bond Environment of Active Site Asp-34 and Asp-216 and Wat-401

	Donor–acceptor	(Å)
Asp-34	OD2–Asp-216 OD1	3.5
	OD2–Thr-35 N	3.5
	OD2–Gly-36 N	3.1
	OD2–Wat-411 O	3.1
	OD1–Ser-37 OG	2.4
	OD1–Wat-469 O	2.4
Asp-216	OD1–Asp-34 OD2	3.5
	OD1–Gly-36 N	3.3
	OD1–Thr-217 N	3.1
	OD1–Gly-218 N	2.6
	OD1–Thr-219 N	3.3
	OD1–Wat-411 O	3.3
	OD1–Thr-219 OG	2.9
	OD1–Wat-411 O	2.9
	OD1–Wat-480 O	2.7
Wat-411	O–Wat-512 O	2.5
	O–Wat-469 O	3.2
	O–Asp-34 OD2	3.1
	O–Asp-216 OD1	3.3
	O–Asp-216 OD2	2.9

distances being 2.8 and 3.1 Å, respectively. The peptide N atoms of Thr-35 and Thr-217 can also potentially interact with the OG1 atoms of Thr-217 and Thr-35. The corresponding distances for these pairs

TABLE V. Root-Mean-Square Differences (Å) Between the Aligned Structures of Acid Proteinases

	2 APR	2 APP	4 APE	
1 CMS	1.21	1.39	1.71	rms(Å)
	295	292	295	pairs
	104	92	84	identities
2 APR		1.21	1.22	rms(Å)
		294	294	pairs
		125	121	identities
2 APP			1.10	rms(Å)
			300	pairs
			167	identities

of atoms are 3.1 and 3.0 Å, respectively. These distances are somewhat larger than found for other acid proteinases, and may reflect errors in the chymosin model due to lower resolution of the X-ray data than what was used for the fungal enzymes.

Chymosin A vs. Chymosin B

Whether amino acid residue at position 244 is a glycine or aspartate determines whether the isozyme is chymosin A or chymosin B. Early activity measurements indicated that chymosin A had a significantly higher specific activity than chymosin B.[49] This residue is positioned in a loop on the sur-

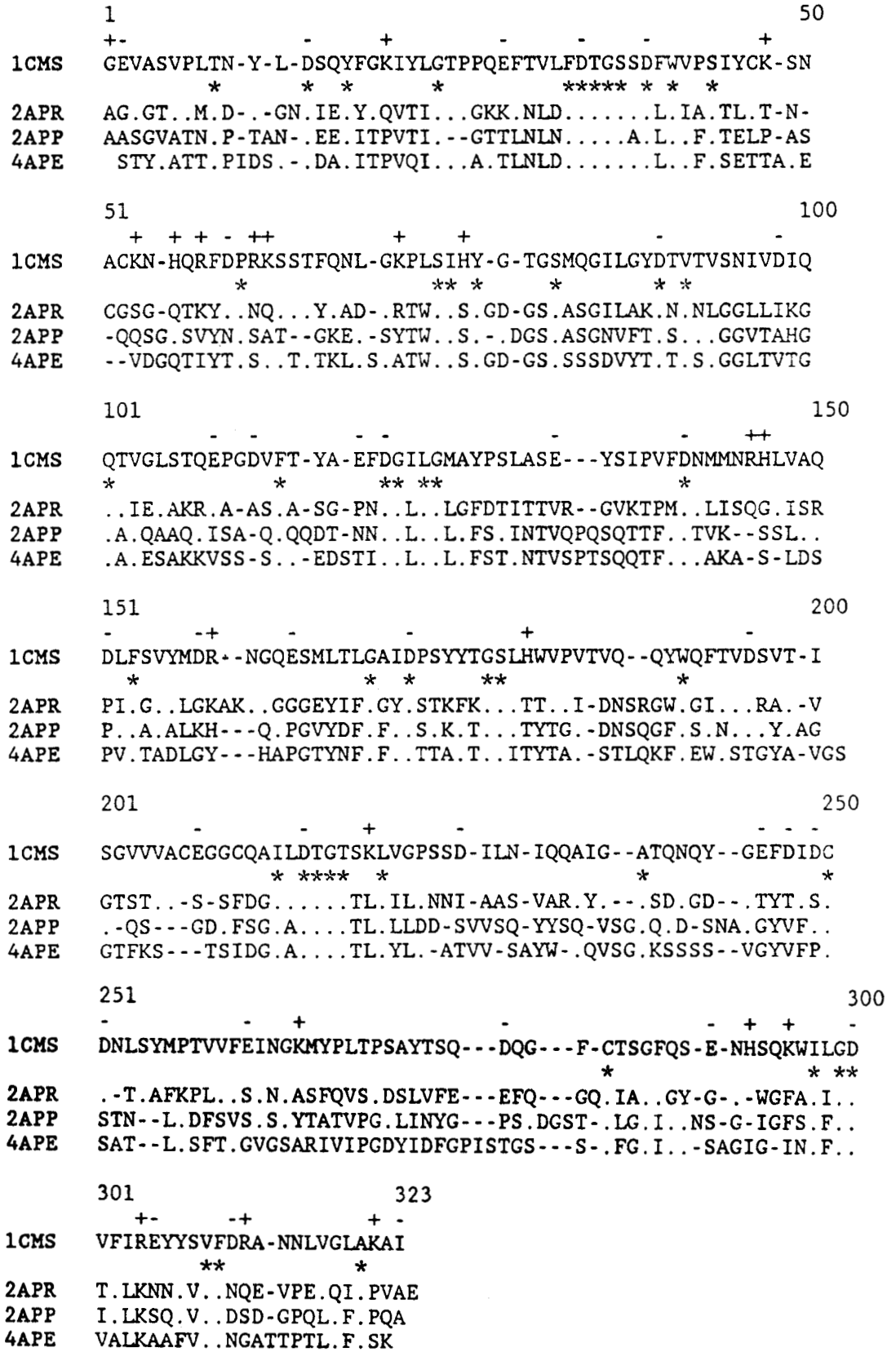

Fig. 7. Sequence alignment of chymosin (1CMS) with rhizopuspepsin (2APR), penicillopepsin (2APP), and endothiapepsin (4APE) based on the least-squares alignment of their three-dimensional structures. A "•" indicates an amino acid residue which is identical to that found in the chymosin sequence. An "*" indicates a topographically equivalent position with the identical amino acid residue in all four structures. Amino acid residues which can be charged are indicated with a "+" or "−" above the appropriate one letter code.

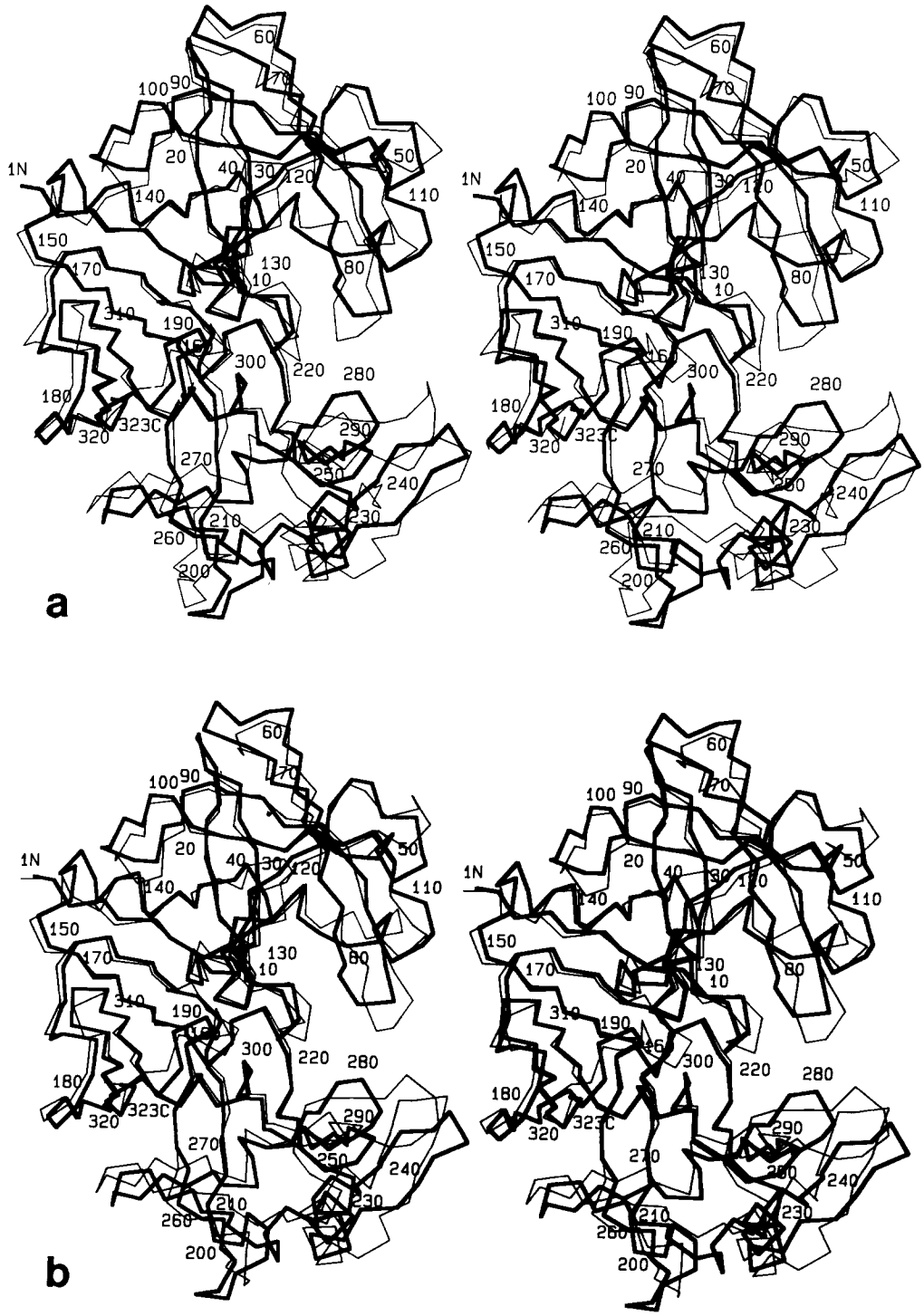

Fig. 8. A stereoscopic view of recombinant chymosin (heavy line) with **(a)** endothiopepsin, **(b)** penicillopepsin, and **(c)** rhizopuspepsin. The recombinant chymosin molecule is represented with a heavy line, the fungal proteinases with thin lines. The N- and C-termini and every tenth residue of the recombinant chymosin molecule are numbered.

face of the enzyme composed of residues 241–249, which is adjacent to a symmetry related molecule. As mentioned earlier, this region was difficult to fit because of weak electron density for this segment of the polypeptide, as indicated by the high B values (Fig. 2). This residue is located more than 20 Å (C^{α} to C^{α} distance) from the active site aspartates.

Safro and Andreeva (The 18th Linderstrom-Lang

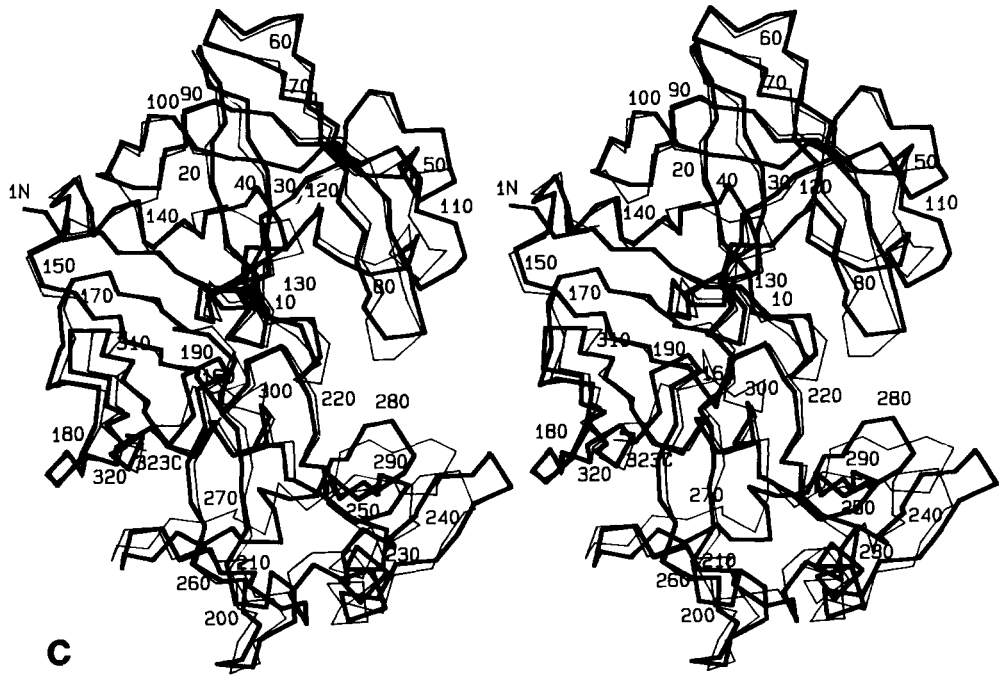

Fig. 8c.

Conference on Aspartic Proteinases, Abstracts, p. 80) have postulated, based on activity studies of Visser and co-workers[50] and their own model-building studies, that the difference in activity may be a result of the increased affinity of the natural substrate κ-casein. In addition to Asp-244, the loop contains three additional negatively charged residues, Asp-245, Asp-247, and Asp-249. The additional negative charge aids in the electrostatic stabilization of the substrate–enzyme complex.

Structural Comparison With Other Acid Protease Structures

The structure of recombinant chymosin has been compared with three other high-resolution acid proteinase structures that have been deposited in the Brookhaven Protein Data Bank.[45] They include rhizopuspepsin (2APR) at 1.8 Å resolution,[18] penicillopepsin (2APP) at 1.8 Å resolution,[10] and endothiapepsin (4APE) at 2.1 Å resolution.[13] A least-squares superposition of all permutations of these enzymes was performed using the program ALIGN.[47] A summary of the results of these comparisons is presented in Table V, and an alignment of the sequences based upon the structural homology of these fungal acid proteinases with chymosin is shown in Figure 7. Of the three structures compared to chymosin, the rhizopuspepsin molecule has the closest structural agreement with an rms of 1.21 Å for 295 structurally equivalent C^{α} positions, where 104 pairs of amino acid side chains are identical. This compared well with the penicillopepsin–endothiapepsin comparison which had an rms of 1.10 Å for the largest number of structurally equivalent C^{α} positions, 300, with the largest number of sequence identities, 167.

The structural superposition of each of the fungal proteinases with chymosin is shown in Figure 8. Upon exmination of the structural alignment, it is evident that in each case the structural similarity is greater for the N-terminal domains than it is for the C-terminal domains. The sequence alignment shown in Figure 4 also reveals the greater similarity of the N-terminal domain sequences over those of the C-terminal domains. There is a larger number of sequence identities for all four enzymes in the N-terminal half of the molecules. It is also quite evident from the comparison studies that the structural integrity of the active site is highly conserved. This was also found to be the case with the comparison of rhizopuspepsin with penicillopepsin.[18]

From the structural comparisons the largest rms deviations of the main chain atoms occur in loop regions which connect the secondary structure elements; these regions of homologous proteins typically have the preponderance of insertions, deletions, and substitutions of amino acid residues. Several of the major differences in the loops of these structures are in regions of the chymosin molecule which have poor or weak electron density, possibly due to disorder of these regions of the molecule.

One important difference between chymosin and the fungal proteinases is the position of Tyr-77. In

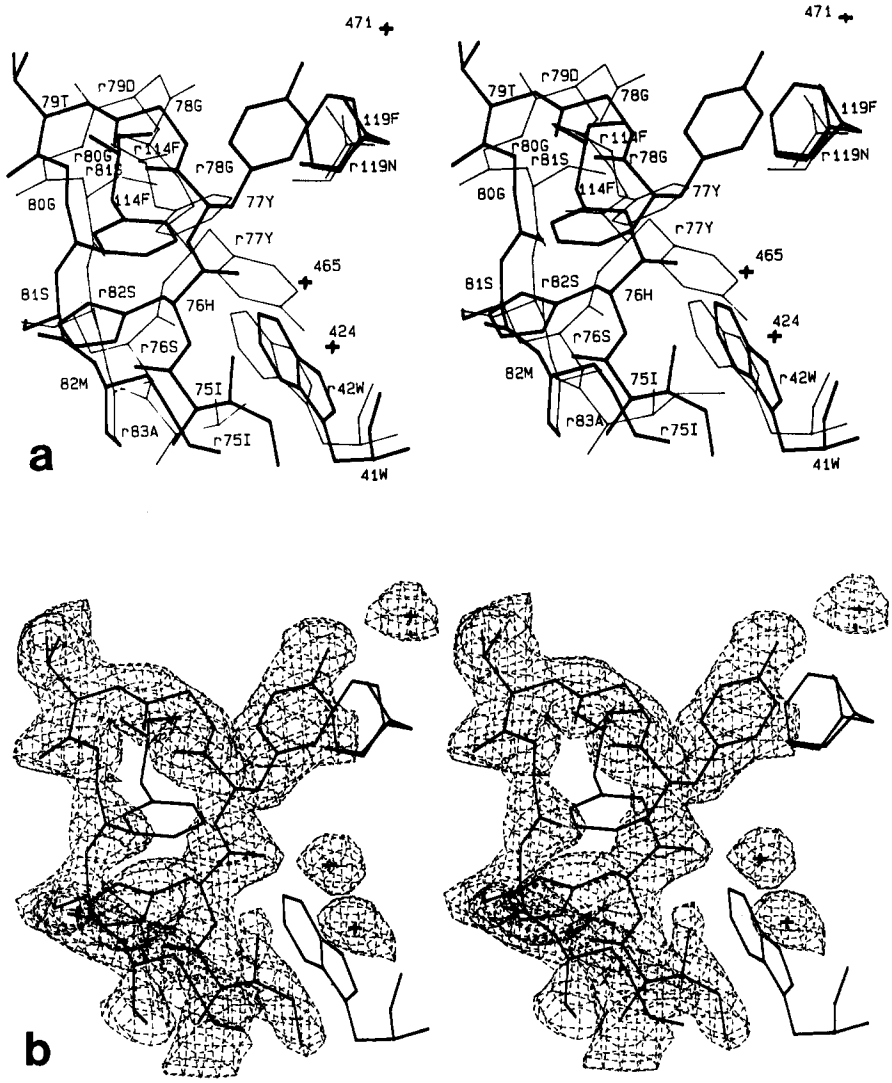

Fig. 9. A stereoscopic view of (**a**) the flap region of recombinant chymosin (heavy line) with the homologous structure of rhizopuspepsin (thin line) superimposed. The residues of both recombinant chymosin and rhizopuspepsin are labeled with the labels of rhizopuspepsin preceded with an "r." (**b**) The $2F_o - F_c$ electron density of the flap contoured at 1σ (the labels are omitted for clarity).

the fungal proteinases this residue is tucked under the flap with the tyrosine OH hydrogen bonding to a tryptophan NE1 (Trp-42 for chymosin). In chymosin this residue protrudes out into the P1 binding pocket with two water molecules, Wat-424 and Wat-465, occupying the region near the position of the tyrosine in the fungal proteinases. One of these, water Wat-424, is within hydrogen bonding distance to the tryptophan NE1. The flap region is pulled in (see Fig. 8) relative to what is found in the fungal proteinases. The position of Tyr-77 is stabilized by aromatic–aromatic interaction with Phe-119 (Fig. 9). The aromatic side chain packing of Tyr-77 with Phe-119 is similar to the interactions found in other protein structures.[51] The residue corresponding to Phe-119 of chymosin is asparagine in the rhizopuspepsin and penicillopepsin, and isoleucine in the endothiapepsin. In addition to the aromatic change in this region, chymosin has a leucine at position 32; the corresponding residue is aspartate in rhizopuspepsin and endothiapepsin and asparagine in penicillopepsin. These two substitutions, Phe-119 and Leu-32, near Tyr-77 make the local environment much more hydrophobic compared to that found in the fungal proteinases.

This dramatic difference was verified by both omitting the flap region during refinement and refining after moving Tyr-77 and the other residues of the flap to the same relative position as that found in rhizopuspepsin. In both cases the difference electron density map clearly indicated that the position of the flap and Tyr-77 reported here is correct.

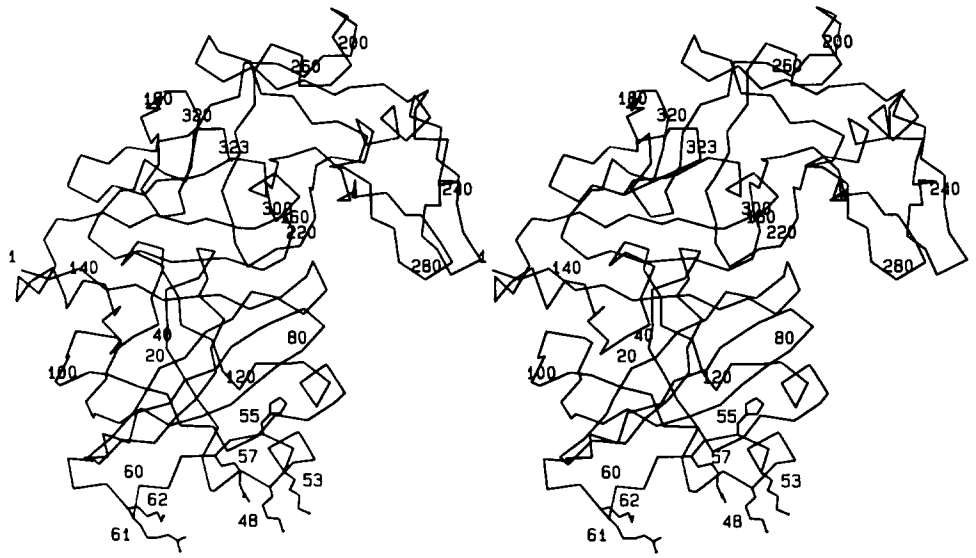

Fig. 10. A stereoscopic view of the N-terminal patch of positive charges amino acid residues. The recombinant chymosin C^{α} backbone is shown with every twentieth C^{α} position labeled. The atoms of side chains of the positive residues in the region from 48 to 62 in the amino acid sequence are shown and labeled.

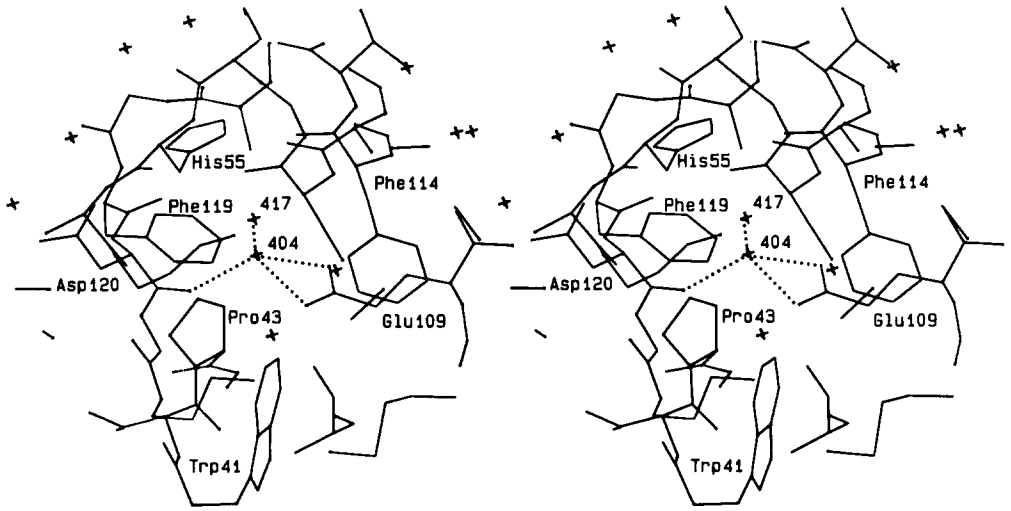

Fig. 11. A stereoscopic view of the environment of Glu-109. Possible hydrogen bonds are shown with dashed lines.

S_1 and S_1' Substrate Binding Subsites

Although attempts in our laboratory of obtaining crystals of enzyme–inhibitor complexes have not proven fruitful, the structural homology of chymosin with other acid proteinases allows the postulation that substrate binding, and thus inhibitor binding, is analogous to that found for the homologous enzymes. Detailed structural studies of a number of aspartic proteinase–inhibitor complexes have been reported.[7,17,52,53] For the fungal acid proteinases there is general agreement as to which residues are involved in interactions with the substrates at the S_1 and S_1' subsites. The residues interacting with the ligands at each subsite in several studies are summarized in Table VI with the corresponding residues for chymosin in parentheses.

The structurally equivalent residues of the S_1 and S_1' sites of chymosin are for the most part very similar to those found in the fungal proteinases. Several of the similarities occur in the region of the flap, residues 73–85. Even though the structurally equivalent residues are similar, as mentioned above, the position of the flap region is the unliganded chymosin is quite different from that found for other

TABLE VI. S and S′ Amino Acid Residue Correspondences Between Recombinant Chymosin and the Fungal Proteinases

Enzyme/complex	S	S′
Rhizopuspepsin/ pepstatin*	Phe-75 (Tyr-77) Gly-76 (—) Asp-77 (Gly-78) Phe-111 (Phe-114) Leu-120 (Ile-122)	Tyr-189 (Tyr-190) Ile-213 (Ile-214) Ile-301 (Ile-297)
Rhizopuspepsin/ D-His-Pro-Phe His-PheC[CH-NH] Phe-Val-Tyr†	Asp-33 (Leu-32) Asp-35 (Asp-34) Tyr-77 (Tyr-77) Asp-79 (Gly-78) Ser-81 (Gly-80) Phe-114 (Phe-114) Leu-122 (Ile-122) Asp-218 (Asp-216) Gly-220 (Gly-218)	Gly-37 (Gly-36) Tyr-77 (Tyr-77) Gly-78 (—) Ile-216 (Ile-214) Asp-218 (Asp-216) Trp-294 (Ser-293) Ile-298 (Ile-297)
Penicillopepsin/ Iva-Val-Val-Sta-OEt‡,§	Asn-31 (Leu-32) Asp-33 (Asp-34) Tyr-75 (Tyr-77) Asp-77 (—) Ser-79 (Gly-80) Leu-121 (Ile-122) Asp-213 (Asp-216) Gly-215 (Gly-218) Thr-216 (Thr-219)	Gly-35 (Gly-36) Thy-75 (Tyr-77) Gly-76 (Gly-78) Phe-190 (Tyr-190) Ile-211 (Ile-214) Asp-213 (Asp-216) Thr-216 (Thr-219)
Penicillopepsin/ Iva-Val-Val-Lys-Sta-OEt‡,§	Asp-33 (Asp-34) Gly-35 (Gly-36) Tyr-75 (Tyr-77) Asp-77 (—) Ser-79 (Gly-80) Phe-112 (Phe-114) Leu-121 (Ile-122) Asp-213 (Asp-216) Gly-215 (Gly-218) Thr-216 (Thr-219)	Gly-35 (Gly-36 Tyr-75 (Tyr-77) Gly-76 (Gly-78) Asp-213 (Asp-216) Thr-216 (Thr-219)

*Bott, R., Subramanian, E., Davies, D. R. Biochemistry 21:6956–6962, 1982.
†Suguna, K., Bott, R. R., Padlan, E. A. Subramanian, E., Sheriff, S., Cohen, G. H., Davies, D. R. J. Mol. Biol. 196:877–900, 1987.
‡James, M. N. G., Sielecki, A. R., Hofmann, T. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. New York: de Gruyter, 1985: 163–177.
§James, M. N. G., Sielecki, A. R. In: "Biological Macromolecules and Assemblies," Vol. 3. Jurnack, F., McPherson, A., eds. New York: Wiley, 1987: 413–482.

TABLE VII. Crystal Packing Contacts*

Symmetry operators	Number of residues for H bonds	van der Waals contacts	*x,y,z* molecule	Transformed molecule
$-x + 1/2, -y + 1/2, z + 1/2$	5	62	12	10
$-x + 1/2, -y + 1/2, z - 1/2$	5	62	10	12
$-x, -y + 1, z$	4	17	4	4
$-x, y, -z$	0	1	1	1
$-x + 1/2, y + 1/2, -z + 1/2$	3	47	11	10
$-x + 1/2, y - 1/2, -z + 1/2$	3	47	10	11
$-x, y - 1, -z$	0	7	1	1
$x + 1/2, -y + 1/2, -z + 1/2$	5	58	5	8
$x - 1/2, -y + 1/2, -z + 1/2$	5	58	8	5

*Contact distance was restricted to 4.0 Å; hydrogen bond distance restricted to 3.2 Å.

acid proteinases. The altered position of the flap is due to the reorientation of Tyr-77, as mentioned above, and a deletion of one amino acid residue in this loop. The fungal proteinases all have an aspartic acid residue in this loop. In chymosin the only charged residue in the flap is His-76. If it is assumed

that substrate of inhibitor binding is similar to that of other acid proteinases, the flap must move to accommodate ligand binding. Our failure to obtain enzyme–inhibitor complexes may be a direct result of the inability of this flap to move because of crystal packing constraints. It also cannot be ruled out that crystal packing constraints may contribute to the differences between chymosin and other acid proteinases in the flap region.

Another difference in the S_1 subsites of chymosin and the fungal proteinase is the substitution of Leu-32 in chymosin for the Asp-33 and Asn-31 of the rhizopuspepsin and penicillopepsin structures, respectively. These two differences provide a more hydrophobic S_1 subsite and undoubtedly contribute to the altered substrate specificity.

Surface Charge Distribution

In addition to the N- and C-termini, there are 54 negatively or positively charged residues on the surface of chymosin at neutral pH. The distribution of these residues is shown in Fig. 7. The chymosin molecule has a net charge of −12 with the N- and C-terminal domains contributing −5 and −7, respectively, to the overall charge. Of the two domains, the N-terminal domain has the largest fraction of positive charge to total charge for the domain, and six of these positive charges are located within residues 48–62 forming a positive patch on the surface of the molecule (Fig. 10). This sequence of positive residues is not found in other acid proteinase molecules.

This patch of positive residues effectively polarizes the chymosin molecule (J. Moult, M. Toner, D. Bacon, and G. Gilliland, unpublished data). This polarization may play a role in determining the efficiency of chymosin cleavage of the Phe-105–Met-106 peptide bond of κ-casein. The polarization of the molecule may predispose the correct orientation of chymosin as it nears the substrate; this may be particularly important if the micelle surface to which κ-casein is associated is negatively charged. Alternatively the patch of positive residues may interact favorably with other components of the casein micelles to facilitate attacking of the preferred peptide bond of κ-casein.

Near the positive patch of residues on the surface of the N-terminal domain there is one completely buried glutamate residue, Glu-109 (Fig. 11). This residue is conserved in the sequence of the mammalian gastric enzymes and several of the fungal enzymes, including rhizopuspepsin. This residue is hydrogen bonded to what has been modeled as a water molecule, Wat-404. The distances between Wat-404 and OE1 and OE2 of Glu-109 are 3.0 and 2.5 Å, respectively. This solvent molecule, which has a low B factor of 7.3 $Å^2$, has two other oxygen ligands, the peptide O of Asp-120 and Wat-417, with bond distances of 2.6 and 2.9 Å, respectively, in an almost planar trigonal arrangement. The average B value for the surrounding oxygen atoms of Wat-404 is 11.3 $Å^2$. Since Wat-404 interacts directly with the carboxylate oxygens of Glu-109, it may possibly be a positively charged ion, such as hydronium ion or a sodium ion. If it is a positive ion, the distances between atoms are rather long for the presence of a sodium ion, and the ligand geometry suggests an hydronium ion would be the more likely of these two candidates. Above and below the plane of these interactions are His-55 and Trp-41, both at a distance greater than 4.0 Å from Wat-404.

In acid proteinases other than the mammalian gastric enzymes, with the exception of rhizopuspepsin, the residue corresponding to Glu-109 in chymosin is usually hydrophobic and always uncharged. The substitutions at this position include leucine, methionine, isoleucine, valine, alanine, and glutamine (Foltmann, B. Aspartic proteinases: Alignment of amino acid sequences. The 18th Linderstrom-Lang Conference on Aspartic Proteinases, pp. 7–20, 1988). Since the gastric enzymes are active under acidic conditions, we would like to suggest that this residue may play a role in an additional method of pH activation. When the gastric enzymes are synthesized in the cell at neutral pH, the glutamate residue would carry a negative charge, and thus would be unlikely to bury itself in the hydrophobic interior of the protein. However, under acidic conditions this residue would be protonated and thus more likely to bury itself inside the protein.

Crystal Packing Interactions

Each chymosin molecule in the *I*222 crystal lattice is associated with nine other molecules. The interactions between the chymosin molecule and its symmetry mates are summarized in Table VII. There are six unique sets of interactions of which three have duplicate sets of interactions. The three unique sets of interactions are at symmetry positions which allow interactions of identical residues in each molecule of the pair.

The extensive interaction involved in crystal packing include 17 unique hydrogen bonds and 192 unique van der Waals contacts closer than 4.0 Å. There are no electrostatic interactions. A total of 62 residues on the surface of each molecule are involved in crystal packing. The number and kinds of interactions are similar to those found for rhizopuspepsin[18] and penicillopepsin.[10] The flap region of chymosin is associated quite closely with symmetry molecule $-x+1/2, y-1/2, -z+1/2$. This interaction coupled with the fact that residues of the flap must move as mentioned above probably accounts for the inability of chymosin crystals to bind inhibitors. This is also the case for penicillopepsin but not for rhizopuspepsin.[18]

ACKNOWLEDGMENTS

The crystallization and X-ray data collection for this project was performed at Genex Corporation while GLG was an employee. The authors would like to thank Genex Corporation for providing X-ray data for the structure determination which was carried out at CARB/NIST. Research sponsored in part by the National Cancer Institute, DHHS, under contract NO. NO1-CO-74101 with ABL. The contents of this publication do not necessarily reflect the views of policies of the Department of Health and Human Services or the National Institute of Standards and Technology. Certain commercial equipment, instruments, and materials are identified in this paper in order to specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology or National Cancer Institute, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

REFERENCES

1. Jolles, J., Alias, C., Jolles, P. The tryptic peptide with rennin-sensitive linkage of cow's κ-casein. Biochim. Biophys. Acta 168:591–593, 1968.
2. MacKinlay, A. G., Wake, R. G. κ-casein and its attack by rennin (chymosin). In: "Milk Proteins," Vol. 2. McKenzie, H. A., (ed). New York: Academic Press, 1971: 175–215.
3. Foltmann, B., Pedersen, V. B., Kaufmann, D., Wybrandt, G. The complete amino acid sequence of prochymosin. J. Biol. Chem. 254:8447–8456, 1979.
4. James, M. N. G., Sielecki, A. R. Aspartic proteinases and their catalytic pathway. In: "Biological Macromolecules and Assemblies," Vol 3. Jurnak, F., McPherson, A., eds. New York: Wiley, 1987: 413–482.
5. Andreeva, N. S., Gustchina, A. E., Federov, A. A., Shutzkever, N. E., Volnova, T. V. X-ray crystallographic studies of pepsin. In: "Acid Proteases, Structure, Function and Biology." Tang, J., ed. New York: Plenum, 1977: 23–31.
6. Andreeva, N., Zdanov, A., Gustchina, A., Federov, A. X-ray diffraction analysis of procine pepsin structure. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. New York: de Gruyter, 1985: 137–150.
7. Andreeva, N. S., Zdanov, A. S., Gustchina, A. E., Fedorov, A. A. Structure of ethanol-inhibited porcine pepsin at 2-Å resolution and binding of the methyl ester of phenylalanyl-diiodotyrosine to the enzyme. J. Biol. Chem. 259:11353–11365, 1984.
8. James, M. N. G., Sielecki, A. R. Molecular structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 Å resolution. Nature (London) 319:33–38, 1986.
9. Hsu, I.-N., Delbaere, L. T. J., James, M. N. G., Hofmann, T. Penicillopepsin: 2.8 Å structure, active site conformation and mechanistic implications. In: "Acid Proteases, Structure, Function and Biology." Tang, J., ed. New York: Plenum, 1977: 61–81.
10. James, M. N. G., Sielecki, A. R. Structure and refinement of penicillopepsin at 1.8 Å resolution. J. Mol. Biol. 163: 299–361, 1983.
11. Jenkins, J., Tickle, I., Sewell, T., Ungaretti, L., Wollmer, A., Blundell, T. X-ray analysis and circular dichroism of the acid proteinase from *Endothia parasitica* and chymosin. In: "Acid Proteases, Structure, Function and Biology." Tang, J., ed. New York: Plenum, 1977: 43–60.
12. Subramanian, E., Swan, I. D. A., Liu, M., Davies, D. R., Jenkins, J. A., Tickle, I. J., Blundell, T. L. Homology among acid proteinases: Comparison of crystal structures at 3 Å resolution of acid proteinases from *Rhizopus chinensis* and *Endothia parasitica*. Proc. Natl. Acad. Sci. U.S.A. 74:556–559, 1977.
13. Pearl, L., Blundell, T. The active site of acid proteinases. FEBS Lett. 174:96–101, 1984.
14. Blundell, T., Jenkins, J., Pearl, L., Sewell, T., Pederson, V. The high resolution structure of endothiapepsin. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. New York: de Gruyter, 1985: 151–161.
15. Wong, C.-H., Lee, T. J., Lee, T.-Y., Lu, T.-H., Yang, I.-H. Structure of acid proteinase from *Endothia parasitica* in cross-linked form at 2.45-Å resolution. Biochemistry 18: 1638–1640, 1979.
16. Subramanian, E., Liu, M., Swan, I. D. A., Davies, D. R. The crystal structure of an acid proteinase from *Rhizopus chinensis* at 2.5 Å resolution. In: "Acid Proteases, Structure, Function and Biology." Tang, J., ed. New York: Plenum, 1977: 33–41.
17. Bott, R., Subramanian, E., Davies, D. R. Three-dimensional structure of the complex of the *Rhizopus chinensis* carboxyl proteinase and pepstatin at 2.5-Å resolution. Biochemistry 21:6956–6962, 1982.
18. Suguna, K., Bott, R. R., Padlan, E. A., Subramanian, E., Sheriff, S., Cohen, G. H., Davies, D. R. Structure and refinement at 1.8 Å resolution of the aspartic proteinase from *Rhizopus chinensis*. J. Mol. Biol. 196:877–900, 1987.
19. Sielecki, A. R., Hayakawa, K., Fujinaga, M., Murphy, M. E. P., Fraser, M., Muir, A. K., Carilli, C. T., Lewicki, J. A., Baxter, J. D., James, M. N. G. Structure of recombinant renin, a target for cardiovascular-active drugs, at 2.5 Å resolution. Science 243:1346–1351, 1989.
20. Tang, J., James, M. N. G., Hsu, I. N., Jenkins, J. A., Blundell, T. L. Structural evidence for gene duplication in the evolution of the acid proteinases. Nature (London) 271: 618–621, 1978.
21. Blundell, T. L., Sewell, B. T., McLachlan, A. D. Four-fold structural repeat in the acid proteinases. Biochim. Biophys. Acta 580:24–31, 1979.
22. Miller, M., Jaskolski, M., Rao, J. K. M., Leis, J., Wlodawer, A. Crystal structure of retroviral proteinase proves relationship to aspartic proteinase family. Nature (London) 337:576–579, 1989.
23. Navia, M. A., Fitzgerald, P. M. D., McKeever, B. M., Leu, C.-T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L., Springer, J. P. Three-dimensional structure of aspartyl proteinase from human immunodeficiency virus HIV-1. Nature (London) 337:615–620, 1989.
24. Berridge, N. J. The purification and crystallization of rennin. Biochem. J. 39:179–186, 1945.
25. Bunn, C. W., Camerman, N., T'sai, L. T., Moews, P. C., Baumber, M. E. X-ray diffraction studies of rennin crystals. Phil. Trans. Roy. Soc. B257:253–158, 1970.
26. Safro, M. Andreeva, N., Zdanov A. The determination of the three-dimensional structure of chymosin. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. New York: de Gruyter, 1985: 183–187.
27. Gilliland, G. L., Davies, D. R. Protein crystallization: The growth of large-scale single crystals. In: "Methods in Enzymolology," Vol. 104. Jakoby, W., ed. New York: Academic Press, 1984: 370–381.
28. Smit, J. D. G., Winterhalter, K. H. Crystallographic data for haemoglobin from the lanceolate fluke *Dicrocoelium dendriticum*. J. Mol. Biol. 146:641–647, 1981.
29. Thaller, C., Weaver, L. H., Eichele, G., Wilson, E., Karlsson, R., Jansonius, J. N. Repeated seeding technique for growing large single crystals of proteins. J. Mol. Biol. 147: 465–469, 1981.
30. Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., Salemme, F. R. The use of an imaging proportional counter in macromolecular crystallography. J. Appl. Crystallogr. 20:383–387, 1987.
31. Gilliland, G. L., Howard, A. J., Winborne, E. L., Poulos, T. L., Steward, D. B., Durham, D. R. Crystallization and preliminary x-ray diffraction studies of subtilisin GX from *Bacillus sp.* GX6644. J. Biol. Chem. 262:4280–4283, 1987.
32. Crowther, R. A. The fast rotation function. In: "The Molecular Replacement Method." Rossmann, M. G., ed. New York: Gordon and Breach, 1972: 173–178.
33. Fitzgerald, P. A. M. MERLOT, an integrated package of computer programs for determination of crystal structures by molecular replacement. J. Appl. Crystallogr. 21:273–278, 1988.

34. Huber, R. Die automatisierte faltmolekulmethode. Acta Crystallogr. 19:353–356, 1965.
35. Steigemann, W. Die entwicklung und anwendung von rechenverfahren und rechenprogrammen zur struckturanalyse von proteinen am biespiel des trypsin-trypsininhibitor komplexes, des freien inhibitors und der L-asparaginase. Ph. D. Thesis, Technical University, Munich, 1974.
36. Hendrickson, W. A. Transformations to optimize the superposition of similar structures. Acta Crystallogr. A35: 158–163, 1979.
37. Crowther, R. A., Blow, R. W. A method of positioning a known molecule in an unknown crystal structure. Acta Crystallogr. 23:544–549, 1967.
38. Hendrickson, W. A. Stereochemically restrained refinement of macromolecular structures. In: "Methods of Enzymology," Vol. 115. Wyckoff, H. W., Hirs, C. H. W., Timasheff, S. N., eds. New York: Academic Press, 1985: 252–270.
39. Finzel, B. C. Incorporation of fast Fourier transforms to speed restrained least-squares refinement of protein structures. J. Appl. Crystallogr. 20:53–55, 1987.
40. Ten Eyck, L. F. Crystallographic fast Fourier transforms. Acta Crystallogr. A29:183–191, 1973.
41. Agarwal, R. C. A new least-squares refinement technique based on the fast Fourier transform algorithm. Acta Crystallogr. A34:791–909, 1978.
42. Sheriff, S. Addition of symmetry-related contact restraints to PROTIN and PROLSQ. J. Appl. Crystallogr. 20:55–57, 1987.
43. Jones, T. A. A graphics model building and refinement system for macromolecules. J. Appl. Crystallogr. 11:268–272, 1978.
44. Bhat, T. N. Calculation of an OMIT map. J. Appl. Crystallogr. 21:279–281, 1988.
45. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rogers, J. R., Kennard, O., Shimanouchi, T., Tasumi, M. The protein data bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
46. Kabsch, W., Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 22:2577–2637, 1983.
47. Satow, Y., Cohen, G. H., Padlan, E. A., Davies, D. R. Phosphocholine binding immunoglobulin Fab McPC603. An X-ray diffraction study at 2.7 Å. J. Mol. Biol. 190:593–604, 1986.
48. Crawford, J. L., Lipscomb, W. N., Schellman, C. G. The reverse turn as a polypeptide conformation in globular proteins. Proc. Natl. Acad. Sci. U.S.A. 70:538–542, 1973.
49. Foltmann, B. Chromatographic purification of prorennin. Acta Chem. Scand. 14:2247–2249, 1960.
50. Visser, S., Slangen, C. J., van Rooijen, P. J. Peptide substrates for chymosin (rennin). Biochem. J. 244:553–558, 1987.
51. Burley, S. K., Petsko, G. A. Aromatic-aromatic interaction: A mechanism of protein structure stabilization. Science 229:23–28, 1985.
52. James, M. N. G., Sielecki, A. R., Hofman, T. X-ray diffraction studies on penicillopepsin and its complexes: the hydrolytic mechanism. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. New York: de Gruyter, 1985: 163–177.
53. Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D., Davies, D. R. Binding of a reduced peptide inhibitor to the aspartic proteinase from *Rhizopus chinensis*: Implications for a mechanism of action. Proc. Natl. Acad. Sci. U.S.A. 84:7009–7013, 1987.
54. Bacon, D., Anderson, W.A. Fast algorithms for rendering space-filling molecular pictures. J. Mol. Graphics. 6:219–220, 1988.