Structure of the Complex of *Streptomyces griseus* Protease B and the Third Domain of the Turkey Ovomucoid Inhibitor at 1.8-Å Resolution†

Randy J. Read, Masao Fujinaga, Anita R. Sielecki, and Michael N. G. James*

ABSTRACT: The structure of the complex between the serine protease *Streptomyces griseus* protease B (SGPB) and the third domain of the Kazal-type ovomucoid inhibitor from turkey has been solved at 1.8-Å resolution and refined to a conventional R factor of 0.125. As others have reported previously for analogous complexes of proteases and protein inhibitors, the inhibitor binds in a fashion similar to that of a substrate; it is not cleaved, but there is a close approach (2.7 Å) of the active site nucleophile Ser-195 O^{γ} to the carbonyl carbon of the reactive peptide bond of the inhibitor. Contrary to the structural reports regarding the other enzyme–inhibitor complexes, we conclude that there is no evidence for a significant distortion of this peptide bond from planarity. The mechanism of inhibition can be understood in terms of the equilibrium thermodynamic parameters K_a, the enzyme–inhibitor association constant, and K_{hyd}, the equilibrium constant for inhibitor hydrolysis. These thermodynamic parameters can be rationalized in terms of the observed structure.

Protein inhibitors of serine proteases are well studied and much is known about the interactions between an inhibitor and its cognate enzyme (Laskowski & Kato, 1980). Most of these inhibitors act by a common mechanism; they bind very tightly to the enzyme (low K_m) but are hydrolyzed very slowly, if at all (low k_{cat}). The inhibitors have at least one peptide bond called the reactive site. This is the bond that interacts with the enzyme's catalytic site and is the one that is cleaved if and when hydrolysis occurs.

There are several families of homologous serine protease inhibitors; ovomucoid inhibitors belong to the pancreatic secretory trypsin inhibitor (Kazal) family (Laskowski & Kato, 1980). The structure of the third domain of the ovomucoid inhibitor of Japanese quail (OMJPQ3) has been reported at 2.8-Å resolution (Weber et al., 1981) and has subsequently been refined at 1.8-Å resolution (Papamokos et al., 1982). The homologous third domain of the ovomucoid inhibitor of turkey (OMTKY3), used in our study, differs from OMJPQ3 only at five residues (Figure 1). Our numbering origin for OMTKY3 corresponds to residue 131 of the complete ovomucoid inhibitor. An I follows the sequence numbers of the residues of the inhibitor in order to distinguish them from those of the enzyme.

Streptomyces griseus protease B (SGPB) is a serine protease obtained from the extracellular culture filtrate, Pronase. Its structure has been described at 2.8-Å resolution (Delbaere et al., 1975) and is now refined at 1.7-Å resolution (L. Sawyer, A. R. Sielecki, and M. N. G. James, unpublished results). The three-dimensional structure has been compared with those of other serine proteases and shows a high degree of topological equivalence (James et al., 1978). Kinetic studies have also been carried out to determine the specificity of its binding sites (Bauer, 1978; James et al., 1980a). The numbering of the SGPB molecule is given in Table I and follows an alignment with chymotrypsinogen (James et al., 1978).

Aside from the present study, the structures of several other complexes between a serine protease and a protein inhibitor have been determined. Most of the work has been done with the pancreatic trypsin inhibitor (PTI). The complexes of this inhibitor with trypsin (Huber et al., 1974), anhydrotrypsin (Huber et al., 1975), and trypsinogen, as well as the ternary complex with trypsinogen and Ile-Val (Bode et al., 1978), have been refined at high resolution. Recently, the structure of the complex between trypsinogen and pancreatic secretory trypsin inhibitor (PSTI) has also been refined at high resolution (Bolognesi et al., 1982). In addition, the structure of *Streptomyces* subtilisin inhibitor (SSI) with subtilisin BPN′ (Hirono et al., 1979) and soybean trypsin inhibitor (STI) with porcine trypsin (Sweet et al., 1974) have been reported, but these structures are not refined yet and are not suitable for detailed analysis.

Both the 2.6-Å resolution structure of STI–trypsin (Sweet et al., 1974) and the initial 2.8-Å resolution structure of PTI–trypsin (Ruhlmann et al., 1973) had indicated that a covalent tetrahedral adduct is formed between the enzyme and the inhibitor. It was concluded that this tetrahedral adduct is isolated because the enzyme functions by stabilizing the transition state (Sweet et al., 1974; Pauling, 1946). However, upon refinement, PTI and trypsin were seen to form a noncovalent complex with a pyramidally distorted carbonyl carbon at the reactive site. The refined structure of OMTKY3–SGPB also indicates that there is a short noncovalent interaction between the two molecules, but the reactive site peptide bond is not significantly distorted from planarity.

A brief report on the refinement of the OMTKY3–SGPB complex has been given elsewhere (Fujinaga et al., 1982). We present here a detailed description and discussion of our results.

Experimental Procedures

Intensity Data Collection. The experimental details are summarized in Tables II and III. The 1.8-Å resolution data set consisted of a total of 20 227 reflections, of which 18 082 were unique. Of these reflections, 13 937 satisfied the condition $I \geq 2\sigma_I$. The program ORESTES (Thiessen & Levy, 1973) was used to obtain the absolute scale of 7.1 and the mean isotropic temperature factor (B) of 13.1 Å^2. This program was also used to calculate the normalized structure factor amplitudes ($|E|$) used in the structure determination.

Structure Determination. The phase problem for the structure of OMTKY3–SGPB was solved by using the technique of molecular replacement. The search model was the

† From the Medical Research Council of Canada Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7. *Received March 18, 1983.* This research was generously supported by a group grant from the Medical Research Council of Canada. R.J.R. and M.F. acknowledge the Medical Research Council of Canada and the Alberta Heritage Foundation for Medical Research for financial support.

0006-2960/83/0422-4420$01.50/0

Table I: Sequence Numbering of SGPB[a]

ILE	SER	GLY	GLY	ASP	ALA	ILE	TYR	SER	SER	THR	GLY	ARG	CYS	SER	LEU	GLY	PHE	ASN	VAL	ARG
16	17	18	19	29	30	31	32	33	34	39	40	41	42	43	44	45	46	47	48	48A
SER	GLY	SER	THR	TYR	TYR	PHE	LEU	THR	ALA	GLY	HIS	CYS	THR	ASP	GLY	ALA	THR	THR	TRP	TRP
48B	48C	48D	49	50	51	52	53	54	55	56	57	58	59	60	62	63	64	65	66	67
ALA	ASN	SER	ALA	ARG	THR	THR	VAL	LEU	GLY	THR	THR	SER	GLY	SER	SER	PHE	PRO	ASN	ASN	ASP
68	78	79	80	81	82	83	84	85	86	87	88	89	90	91	93	94	99A	100	101	102
TYR	GLY	ILE	VAL	ARG	TYR	THR	ASN	THR	THR	ILE	PRO	LYS	ASP	GLY	THR	VAL	GLY	GLY	GLN	ASP
103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123
ILE	THR	SER	ALA	ALA	ASN	ALA	THR	VAL	GLY	MET	ALA	VAL	THR	ARG	ARG	GLY	SER	THR	THR	GLY
124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	156
THR	HIS	SER	GLY	SER	VAL	THR	ALA	LEU	ASN	ALA	THR	VAL	ASN	TYR	GLY	GLY	GLY	ASP	VAL	VAL
157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177
TYR	GLY	MET	ILE	ARG	THR	ASN	VAL	CYS	ALA	GLU	PRO	GLY	ASP	SER	GLY	GLY	PRO	LEU	TYR	SER
178	179	180	181	182	183	184	190	191	192	192A	192B	193	194	195	196	197	198	199	200	201
GLY	THR	ARG	ALA	ILE	GLY	LEU	THR	SER	GLY	GLY	SER	GLY	ASN	CYS	SER	SER	GLY	GLY	THR	THR
202	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226
PHE	PHE	GLN	PRO	VAL	THR	GLU	ALA	LEU	VAL	ALA	TYR	GLY	VAL	SER	VAL	TYR				
227	228	229	230	231	232	233	234	235	235A	236	237	238	239	240	241	242				

[a] This numbering is based on an alignment of the bacterial serine proteinases with chymotrypsinogen (James et al., 1978).

Table II: Crystal Data for OMTKY3–SGPB Complex

unit-cell dimensions	$a = 45.35\ (4)$ Å
	$b = 54.52\ (5)$ Å
	$c = 45.65\ (4)$ Å
	$\beta = 119.2\ (1)°$
space group	$P2_1$
unit-cell content (Z)	2
growth conditions	0.75 M NaH_2PO_4 or KH_2PO_4, pH 6.3

```
          1                5                10                15
OMTKY3  Leu-Ala-Ala-Val-Ser-Val-Asp-Cys-Ser-Glu-Tyr-Pro-Lys-Pro-Ala-
OMJPQ3  Leu-Ala-Ala-Val-Ser-Val-Asp-Cys-Ser-Glu-Tyr-Pro-Lys-Pro-Ala-
                                                 *
                         ↓
                        20                25                30
OMTKY3  Cys-Thr-Leu-Glu-Tyr-Arg-Pro-Leu-Cys-Gly-Ser-Asp-Asn-Lys-Thr-
OMJPQ3  Cys-Pro-Lys-Asp-Tyr-Arg-Pro-Val-Cys-Gly-Ser-Asp-Asn-Lys-Thr-
          —   —   —                 —

                        35                40                45
OMTKY3  Tyr-Gly-Asn-Lys-Cys-Asn-Phe-Cys-Asn-Ala-Val-Val-Glu-Ser-Asn-
OMJPQ3  Tyr-Gly-Asn-Lys-Cys-Asn-Phe-Cys-Asn-Ala-Val-Val-Glu-Ser-Asn-
          **

                    50                56
OMTKY3  Gly-Thr-Leu-Thr-Leu-Ser-His-Phe-Gly-Lys-Cys
OMJPQ3  Gly-Thr-Leu-Thr-Leu-Asn-His-Phe-Gly-Lys-Cys
                        —

        * cis-Proline
        ** Polymorphic (Gly/Ser) in OMJPQ3 (Bogard et al., 1980)
```

FIGURE 1: Comparison of the sequences of the third domain of the ovomucoid inhibitor from turkey (OMTKY3) and Japanese quail (OMJPQ3) (Kato et al., 1978). The differences in the sequences are underlined. The arrow indicates the reactive site.

native SGPB structure (Delbaere et al., 1975), which, at the time of this work, had been refined to an R factor[1] of 0.177 at 1.7-Å resolution.

The rotational parameters were determined from the fast rotation function (Crowther, 1973) by using normalized structure factors ($|E|$) calculated for the model of SGPB in an orthogonal cell of symmetry $P1$, with a unit-cell edge of 60 Å. For a maximum radius of integration of 21 Å, this cell axis length should be sufficient to minimize the contribution of intermolecular vectors for the model molecule SGPB, which has a maximum diameter of about 40 Å.

The rotation function was calculated on a 5° grid for an asymmetric unit of rotation function space (Rao et al., 1980). The radius of integration was from 3 to 21 Å, and data from

[1] $R = \sum||F_o| - |F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively.

Table III: Data Collection and Data Processing

diffractometer	Nonius CAD4
incident radiation	Ni-filtered Cu Kα, 40 kV, 26 mA, 1.3 mm diameter collimator
diffracted beam	60-cm crystal to counter distance, He-filled beam tunnel
peak scan	0.5° in ω at 0.67°/min
background scan	0.125° in ω on either side of the peak scan
background correction	bicubic spline fitting of the observed backgrounds as a function of θ and ϕ
absorption correction	empirical function of ϕ (North et al., 1968), maximum correction factor = 1.63
decay correction	function of θ and time (Hendrickson, 1976)
geometrical correction	Lorentz and polarization
intensity variance	$\sigma_I^2 = I + c^2I^2 + (t_I/t_{Bk})^2(\Sigma Bk + c^2\Sigma Bk^2)$[a]

[a] I = total intensity; Bk = background counts; t_I = time taken for intensity measurement; t_{Bk} = time taken for background measurement; c = instrument instability constant (=0.01).

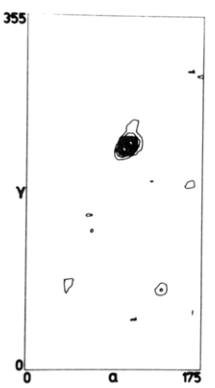

FIGURE 2: Section of the rotation-function map containing the highest peak. In this section, the Euler angle β is 108°. Contours are in increments of 1 standard deviation from the mean value of the map and start at 2 standard deviations.

10- to 3.5-Å resolution were included. The 875 $|E|$'s greater than 1.0 for the complex and the 2018 $|E|$'s greater than 2.0 for the search model were accepted. The two highest peaks in the rotation function were 10.9 and 3.9 standard deviations

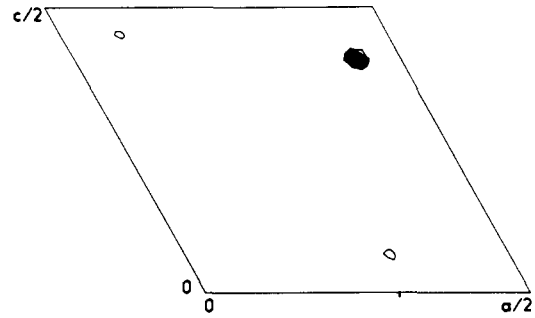

FIGURE 3: R factor translational search. Contours are in decrements of 1 standard deviation below the mean R factor and start at 2 standard deviations.

above the mean. A finer scan with intervals of 1° in the Euler angle β was calculated in the region of the highest peak. The section containing the correct peak (11.4 standard deviations above the mean) is shown in Figure 2. The appropriately rotated SGPB coordinates became the search model for the determination of the translational parameters.

An R factor search was used to solve the translation problem. Since $P2_1$ has a polar y axis, the search was carried out in the xz plane only. Intensity data in the 5–4-Å resolution range were used to calculate R factors as the search model was translated in 1-Å increments over an area from (0,0) to $(a/2,c/2)$. A contour map of this array of R factors is shown in Figure 3. The minimum R factor obtained was 0.35, 8 standard deviations lower than the average of 0.44. The second lowest value was 0.41. An additional search on a 0.25-Å grid located the minimum more accurately.

The R factor for the rotated and translated SGPB molecule was 0.39 for the data to 2.8-Å resolution. This model was used to obtain calculated structure factor amplitudes and phases for the computation of ΔF and F_o maps. These are electron density maps with coefficients $|F_o| - |F_c|$ and $|F_o|$, respectively, and calculated phases, α_c.

The MMS-X interactive graphics system (Barry et al., 1976) with the macromolecular modeling system, M3, developed by C. Broughton (Sielecki et al., 1982), was used for map interpretation and model fitting. The model fitting consisted of taking short segments of about five residues of the known OMTKY3 sequence and fitting them to the electron density. The SGPB molecule was not adjusted at this point. The N-terminal region of the inhibitor showed very little electron density, and the first six residues from Leu-1I to Val-6I were never located even after refinement at 1.8-Å resolution. The possibility of proteolytic removal of these residues is eliminated since the amino acid analysis of the complex shows the composition expected for an intact molecule (M. Laskowski, Jr., personal communication).

Refinement. The refinement of the structure was started with 50 out of the 56 residues of the inhibitor and the rotated and translated structure of the refined native SGPB molecule. The initial R factor for this model was 0.31 for the data in the 8.0–2.8-Å resolution shell. We used the restrained least-squares refinement program (Hendrickson & Konnert, 1980), which was modified by Furey (Furey et al., 1982) and one of us (M.F.) for the FPS 190L array processor.

At various stages during the refinement, electron density maps with coefficients $2|F_o| - |F_c|$ and $|F_o| - |F_c|$, with calculated phases, were produced. The maps were superimposed on the model and displayed on the graphics system, and corrections to the model were made where necessary. At later stages of refinement, difference maps were also used to locate possible solvent atom positions.

Table IV: Refinement Parameters and Results for the Last Cycle

no. of cycles	58
R ($=\Sigma \|\|F_o\| - \|F_c\|\| / \Sigma \|F_o\|$)	0.125 [0.145]
no. of reflections (resolution range)	13 457 (6–1.8 Å, $I \geq 2\sigma_I$) [16 245 (10–1.8 Å, $I \geq \sigma_I/2$)]
no. of protein atoms	1690
no. of solvent atoms	165
no. of variable parameters	7586
rms deviations from ideal values[a]	
distance restraints	
bond distance	0.017 (0.008) Å
angle distance	0.037 (0.016) Å
planar 1–4 distance	0.042 (0.016) Å
plane restraint	0.021 (0.012) Å
chiral-center restraint	0.192 (0.080) Å^3
nonbonded contact restraints	
single torsion contact	0.259 (0.400) Å
multiple torsion contact	0.155 (0.400) Å
possible hydrogen bond	0.201 (0.400) Å
conformational torsion angle restraint	
planar (ω)	3.8 (2.8)°

[a] The values of σ, in parentheses, are the input estimated standard deviations that determine the relative weights of the corresponding restraints [see Hendrickson & Konnert (1980)].

Different structure factor weighting schemes were used. In the initial cycles, weight $= 1/\sigma_F^2$, where $\sigma_F \approx (1/2)\langle\Delta F\rangle$, was used. At later stages, this was replaced by $\sigma_F = C_1 + C_2[(\sin\theta)/\lambda - 1/6]$, to give higher weights to higher resolution data. For the last cycle, the values of C_1 and C_2 were 10.0 and −55.0, respectively. Table IV lists the parameters and results of the last cycle.

After 58 cycles of refinement, the R factor has dropped to 0.125 for the data from 6.0- to 1.8-Å resolution. The R factor becomes 0.145 for the 16 245 structure factors in the range 10.0–1.8-Å resolution with $I \geq \sigma_I/2$. The model has a total of 1855 atoms, all with individual temperature factors. The occupancies of the 165 solvent molecules included were also refined. All solvent sites were interpreted as water molecules and refined as neutral oxygen atoms. The refined parameters of the OMTKY3–SGPB complex have been deposited with the Brookhaven Protein Data Bank (Bernstein et al., 1977).

Results and Discussion

Estimation of Error. At present, there is no accurate way of determining the errors in the atomic coordinates of a molecule refined by a restrained least-squares procedure. We looked at various ways of estimating coordinate errors.

The method of Luzzati (1952) gives an estimate of 0.13 Å for the overall mean error, $\langle|\Delta r|\rangle$, from the variation of the R factor with resolution. This value corresponds to a root mean square (rms) error of 0.14 Å for a normal distribution of Δr [$\langle|\Delta r|\rangle \text{rms} = (3\pi/8)^{1/2}\langle|\Delta r|\rangle$]. It should be noted, however, that this method is not strictly applicable in our case since, among other reasons, we have weighted the structure factors as a function of $\sin\theta$ in the refinement of the structure. In any event, an overall value is not very useful since there will be regions in the molecule with errors greatly exceeding this value.

The errors in the individual atomic positions may be estimated from the formula given by Cruickshank (1949, 1954, 1967). For a monoclinic space group, $\sigma x_i = \sigma y_i = \sigma z_i$, and σy_i is given by

$$\sigma y_i = \frac{b}{2\pi} \frac{[\sum_{hkl} k^2(|F_o| - |F_c|)^2]^{1/2}}{\sum_{hkl} \frac{m}{2} k^2 f_{oi} \exp[-B_i(\sin^2\theta)/\lambda^2]} \quad (1)$$

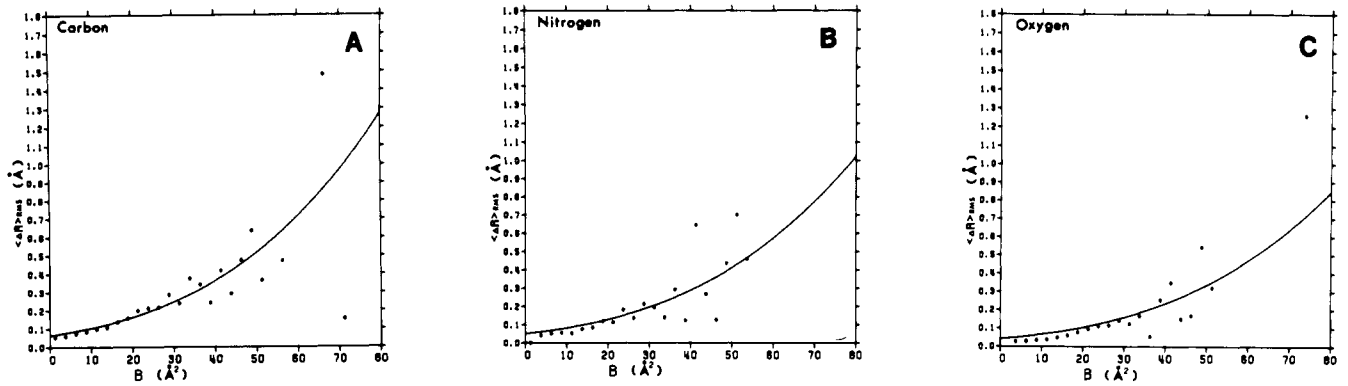

FIGURE 4: rms difference between restrained and unrestrained structures as a function of isotropic temperature factor. The solvent molecules were omitted from the calculation. The curve in each plot is the error given by the Cruickshank formula.

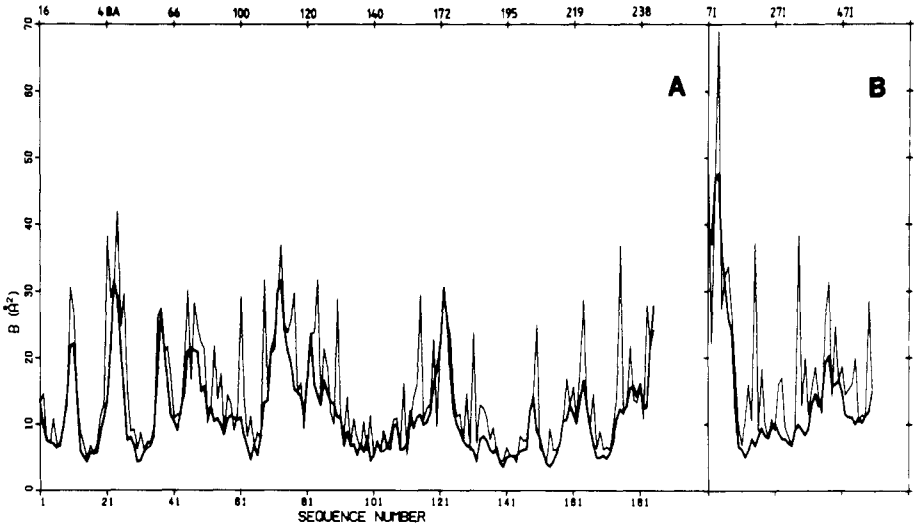

FIGURE 5: Variation of isotropic temperature factor, B, with position along the polypeptide chain. The thick lines show the mean B of the main-chain atoms while the thin lines show the mean B of the side-chain atoms. (A) SGPB molecule. Chymotrypsinogen numbering is shown at the top. (B) OMTKY3 molecule. The plot begins at 7I because the first six residues were not located.

where b is the axial length, f_{oi} is the atomic scattering factor, and $m = 2$ or 1, depending on whether the reflection is centric or noncentric, respectively. Formula 1 gives the standard deviation as a function of atom type and individual temperature factor. In its derivation, well-resolved spherically symmetric atoms were assumed. The radial error is given by $\sigma r_i = 3^{1/2}\sigma y_i$. The Cruickshank formula leads to an rms error of 0.14 Å for all atoms in the OMTKY3–SGPB structure.

Alternatively, the errors in the individual atomic positions may be estimated by comparing the refined structure with that refined without structural restraints (Chambers & Stroud, 1979). The refined structure was used as the starting model in seven cycles of unrestrained refinement. This refinement converged with an R factor of 0.103. The rms difference between the coordinates of the restrained and unrestrained structure is 0.14 Å for all atoms; Figure 4 shows these rms differences for different atom types plotted against the temperature factor B, as well as the errors predicted by the Cruickshank formula. These plots show fairly good agreement between the two methods. Despite the fact that the assumptions used in deriving the Cruickshank formula are not completely satisfied, our results indicate that reasonable estimates of errors can be obtained by the method.

Contrary to our results, Chambers & Stroud (1979) concluded that the Cruickshank formula underestimated the error in their structure. They interpreted the formula as estimating the rms shift during a single cycle of difference Fourier refinement. Thus, for a structure that is still far from convergence, the formula underestimates the errors. The agreement that we obtain could indicate that our structure is closer to convergence. The similarity of the results from the three different methods of error estimation gives us some confidence in them. The Cruickshank formula (see Figure 4) is the most useful for estimating the errors of different parts of the molecule as a function of B. The variation of the mean isotropic temperature factor along the polypeptide chain is shown in Figure 5.

Overall Description. Figure 6 shows the α-carbon representation of the OMTKY3–SGPB complex (see paragraph at end of paper regarding supplementary material). The structure of SGPB in the complex closely resembles that of the native molecule. The similarities in the three-dimensional structure of the native SGPB molecule with the bacterial serine proteases *S. griseus* protease A (SGPA) and α-lytic protease, as well as with the pancreatic enzymes chymotrypsin and elastase, have been shown before (Delbaere et al., 1975; James et al., 1978).

The structure of OMTKY3 in the complex is shown in Figure 7, and a listing of the hydrogen-bonding distances in the inhibitor is given in Table V. The folding of OMTKY3 (Fujinaga et al., 1982) is like that of the homologous protein of the Japanese quail, OMJPQ3 (Weber et al., 1981). The similarity of the two proteins is expected since their sequences differ in only five positions.

The regions of the active site of SGPB and the reactive site of OMTKY3 are shown, superimposed on the $2|F_o| - |F_c|$ map,

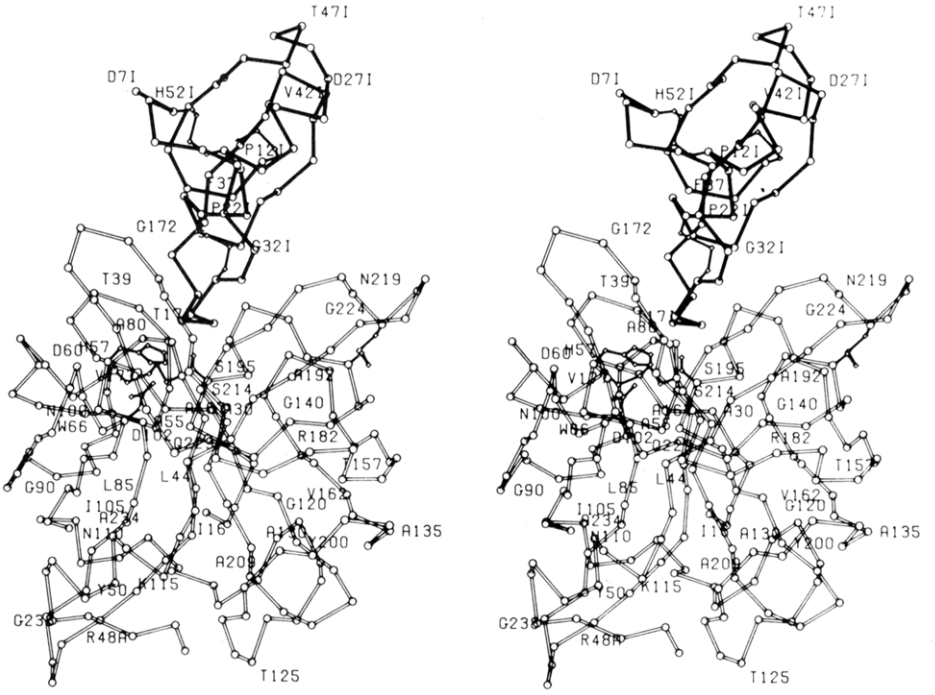

FIGURE 6: An α-carbon representation of the molecular complex between SGPB (open bonds) and OMTKY3 (filled bonds). Every fifth amino acid residue in each molecule is labeled with the residue type and sequence number. Disulfide bridges and side-chain atoms of His-57, Asp-102, and Ser-195 are also shown.

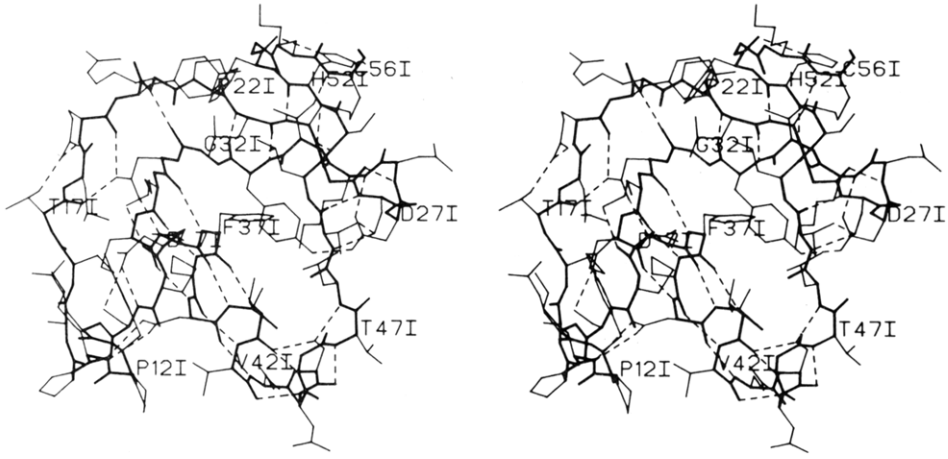

FIGURE 7: Structure of OMTKY3 molecule in the complex. The main chain is indicated in thick lines, and the hydrogen bonds are shown in dashed lines. The following criteria were used to determine hydrogen bonds. If the position of the hydrogen atom could be deduced from stereochemistry, then it was required that the distance from the donor atom to the acceptor atom be less than 3.4 Å, the distance from the acceptor to the hydrogen atom be less than 2.4 Å, and the angle at the hydrogen be greater than 135°. If the hydrogen atom position was ambiguous, then it was placed in the closest possible position to the acceptor atom, and the same criteria were applied.

in parts A and B of Figure 8, respectively. The three terminal atoms of the Arg-21I side chain have *B* factors of about 50 Å^2, and no associated density can be seen at the contour level used in the figure. As mentioned before, the N-terminal region of the inhibitor is disordered, and Figure 8C shows that there is very little electron density in this area. The extent of disorder can also be appreciated from the very high *B* factors of these residues (Figure 5).

Binding Sites. The small contact area between the inhibitor and the enzyme can be seen in Figure 6. A close-up view of the active site region is shown in Figure 9, and a summary of the 108 intermolecular hydrogen bonds and van der Waals contacts less than 4.0 Å is given in Table VI.

Studies of small peptide substrates with SGPB (Bauer, 1978) and of many protease–inhibitor pairs (Laskowski & Kato, 1980) have shown that the amino acid occupying the primary specificity pocket is the most important to binding. The optimal P_1 residues[2] for substrates of SGPB are phenylalanine, tyrosine, and, to a lesser extent, leucine (Bauer, 1978). The relative importance of Leu-18I in the P_1 position of OMTKY3 is indicated by the 29 contacts that it makes to the active site of the enzyme (Table VI).

In addition, Thr-17I and Tyr-20I at positions P_2 and P_2', respectively, make substantial contributions to the binding. The relatively small number of contacts made by Glu-19I at position P_1' is due, in part, to its unusual side-chain conformation ($\chi^1 = 61°$, $\chi^2 = -62°$, in contrast with the most highly

[2] The nomenclature introduced by Schechter & Berger (1967) is used to facilitate discussion about the interactions between a protease and bound peptides. Amino acid residues of substrates are numbered P_1, P_2, P_3, etc. toward the N-terminal direction and P_1', P_2', etc. in the C-terminal direction from the scissile bond. The complementary subsites of the enzyme binding region are numbered S_1, S_2, etc. and S_1', S_2', etc.

Table V: Hydrogen-Bonding Distances (Å) in OMTKY3

Asp-7I O- - -Lys-34I N^{ζ}	3.3
Asp-7I $O^{\delta 1}$ - - -Ser-9I N	3.0
Asp-7I $O^{\delta 1}$ - - -Ser-9I O^{γ}	2.9
Lys-13I N- - -Asn-39I $O^{\delta 1}$	3.1
Lys-13I O- - -Asn-39I $N^{\delta 2}$	3.2
Thr-17I O- - -Asn-33I $N^{\delta 2}$	2.9
Thr-17I $O^{\gamma 1}$ - - -Glu-19I $O^{\epsilon 1}$	2.5
Glu-19I O- - -Asn-33I $N^{\delta 2}$	3.0
Glu-19I $O^{\epsilon 1}$ - - -Glu-19I N	2.9
Arg-21I N- - -Gly-32I O	2.7
Leu-23I N- - -Tyr-31I O	2.9
Leu-23I O- - -Tyr-31I N	2.8
Cys-24I N- - -His-52I O	2.8
Cys-24I O- - -Ser-51I N	2.8
Cys-24I O- - -His-52I N	3.4
Gly-25I N- - -Lys-29I O	2.9
Gly-25I O- - -Asn-28I N	2.9
Ser-26I N- - -Thr-49I O	2.8
Ser-26I O^{γ}- - -Thr-49I N	3.0
Ser-26I O^{γ}- - -Thr-49I $O^{\gamma 1}$	3.4
Asp-27I $O^{\delta 1}$ - - -Lys-29I N	3.1
Asp-27I $O^{\delta 2}$ - - -Tyr-31I O^{η}	2.6
Thr-30I $O^{\gamma 1}$ - - -Cys-56I N	2.9
Asn-33I O- - -Phe-37I N	3.0
Asn-33I $O^{\delta 1}$ - - -Asn-36I N	2.8
Lys-34I O- - -Cys-38I N	2.8
Cys-35I O- - -Asn-39I N	3.1
Asn-36I O- - -Ala-40I N	3.1
Phe-37I O- - -Val-41I N	2.9
Cys-38I O- - -Val-42I N	3.0
Asn-39I O- - -Glu-43I N	3.0
Val-41I O- - -Ser-44I O^{γ}	2.7
Val-41I O- - -Gly-46I N	3.1
Val-42I O- - -Asn-45I N	2.9
Ser-44I O- - -Thr-47I N	3.3
Ser-44I O^{γ}- - -Leu-48I N	3.1
His-52I $N^{\delta 1}$ - - -Phe-53I O	3.0

populated conformation of −60° and 180°; Janin et al., 1978). The side-chain carboxyl group of Glu-19I forms a hydrogen bond to the main-chain NH (Figure 9). Since the S_1' site on the enzyme binds hydrophobic residues (Bauer, 1978), the glutamate side chain is not likely to interact with it.

The binding interactions in OMTKY3–SGPB are similar to those in PTI–trypsin (Huber et al., 1974), as well as to those in tetrapeptide complexes with the homologous bacterial serine protease SGPA (James et al., 1980b). The hydrogen bonds involved in binding OMTKY3 to SGPB are listed in Table VII, and their lengths are compared with the corresponding distances in the structures of the other complexes. Peptides I and II are hydrolysis products, Ac-Pro-Ala-Pro-Phe-OH and Ac-Pro-Ala-Pro-Tyr-OH, respectively, bound noncovalently to SGPA. Peptide IV is an aldehyde, Ac-Pro-Ala-Pro-Phe-H, bound covalently to SGPA through a hemiacetal linkage to Ser-195. This latter structure is taken as an analogue of the tetrahedral intermediate. Figure 10 shows the superposition of the active site regions of the various complexes.

The P_3 residue, Cys-16I, makes two hydrogen bonds in an antiparallel fashion to Gly-216, as in the complexes of SGPA. In PTI, the P_3 residue is proline; therefore, only one hydrogen bond can be made (Table VII). Antiparallel hydrogen bonds are observed also between the P_2' Tyr-20I and Arg-41 in SGPB. In trypsin, the conformation of the polypeptide chain at the homologous Phe-41 precludes the formation of a hydrogen bond to the amide nitrogen of the P_2' residue in PTI.

The carbonyl oxygen of the P_1 residue, Leu-18I, is located in the oxyanion hole (Robertus et al., 1972) and accepts one strong hydrogen bond from Gly-193 N and a weaker one from Ser-195 N. The same pattern of hydrogen bonding is seen in all of the complexes presented in Table VII except for the

Table VI: Intermolecular Contacts Less Than 4.0 Å for SGPB–OMTKY3[a]

		SGPB														
OMTKY3	site	T-39	G-40	R-41	C-42	H-57	V-169	Y-171	A-192	E-192A	P-192B	G-193	D-194	S-195	S-214	G-215	G-216	S-217	Σ
K13I	P_6							5 (1)											5
P14I	P_5																	2	2
A15I	P_4						1	2											8
C16I	P_3							2									5		9
T17I	P_2					8		6								3	4 (2)		16
L18I	P_1					1			1	2	4	4 (1)	1	11 (1)	2	3	1		29
E19I	P_1'			2	1	1					1	1		3	1				9
Y20I	P_2'		2	11 (2)							1	3							17
R21I	P_3'	5																	5
G32I	P_{14}'										4								4
N36I	P_{18}'										4								4
Σ		5	2	13	1	10	1	15	1	2	14	8	1	14	3	6	10	2	108

[a] The numbers in parentheses are the number of hydrogen bonds included in the total.

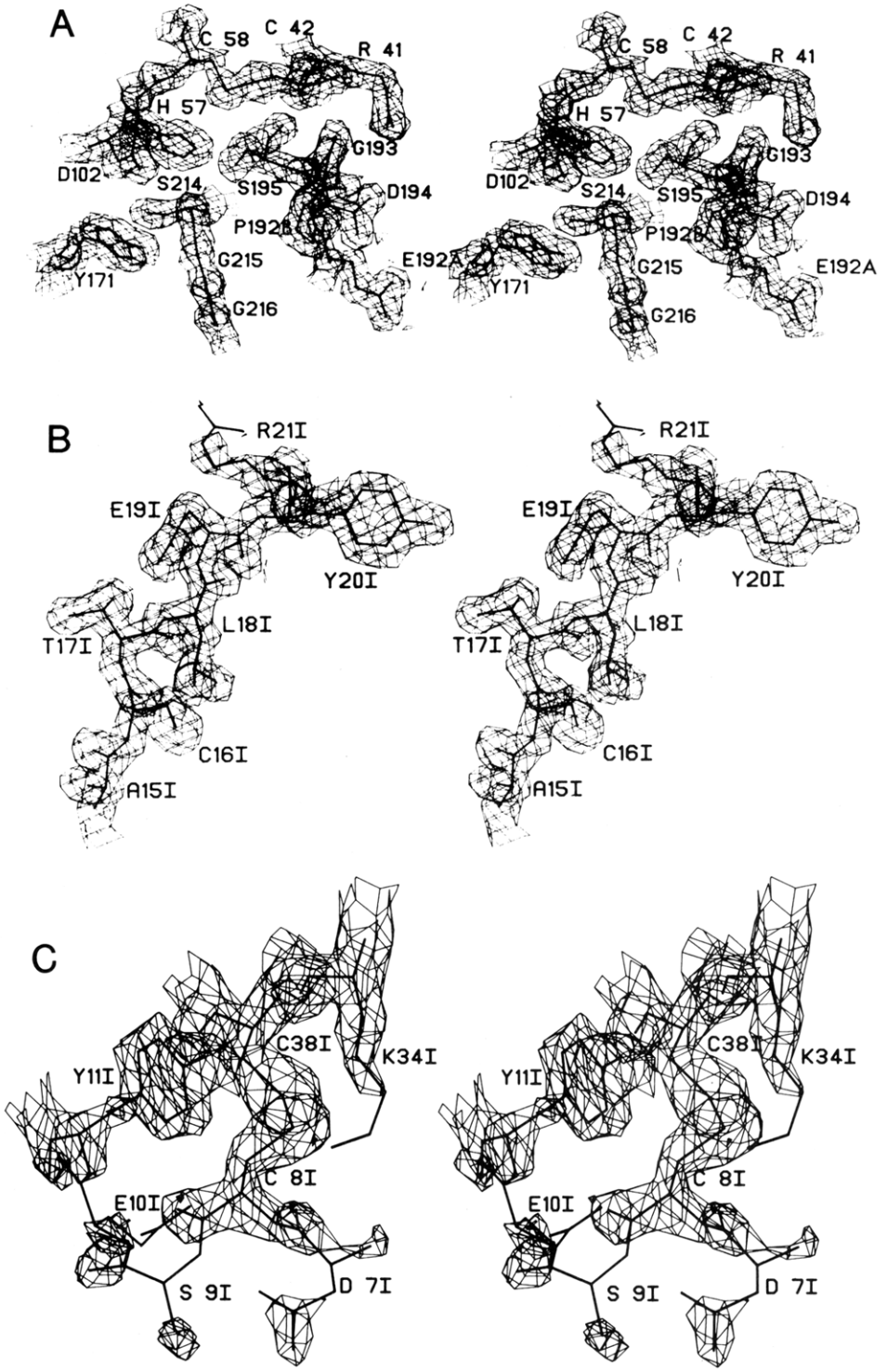

FIGURE 8: Electron density maps of different regions in the complex. The maps were calculated with coefficients $2|F_o| - |F_c|$ and calculated phases. Contour surfaces are drawn at 0.40 e Å^{-3}. (A) Catalytic site of SGPB. (B) Reactive site of OMTKY3. The side chain of Arg-21I is highly mobile and is associated with low density at its terminus. (C) N-Terminal region of OMTKY3. This region of the molecule is highly disordered and has little associated density.

aldehyde structure. The two long hydrogen bonds in this complex can be attributed to the protonation of the aldehyde carbonyl oxygen and the formation of a covalent bond to SGPA (James et al., 1980b).

The interaction between the amide NH of the P_1 residue and Ser-214 O is too long and too poorly oriented to be considered a hydrogen bond in OMTKY3–SGPB and PTI–trypsin (Table VII). It has been suggested that the tetrahedral intermediate is stabilized in part by the formation of a good hydrogen bond in this position (Robertus et al., 1972). Hydrogen bonds are formed to Ser-214 O in the tetrapeptide complexes with SGPA. This interaction is strongest in the aldehyde structure, as shown by the short distance of 2.8 Å. Previous model-building experiments (James et al., 1980b) had indicated that this hydrogen bond to Ser-214 O had to be relaxed in order to accommodate a P_1' residue.

Reactive Site. The distance between Ser-195 O^γ and Leu-18I C, the carbonyl carbon at the reactive site, is 2.7 Å, too long to be a covalent bonding distance and too short for a favorable van der Waals interaction. This close contact is

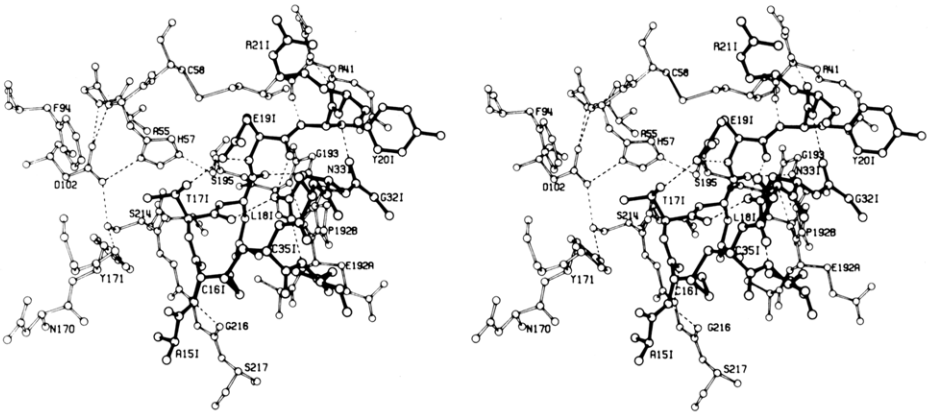

FIGURE 9: An ORTEP (Johnson, 1965) diagram showing interaction between SGPB and OMTKY3 (side chain of Lys-34I omitted for clarity). There is a close contact of 2.7 Å between Ser-195 O^{γ} and Leu-18I C that seems to involve an attractive interaction between the atoms.

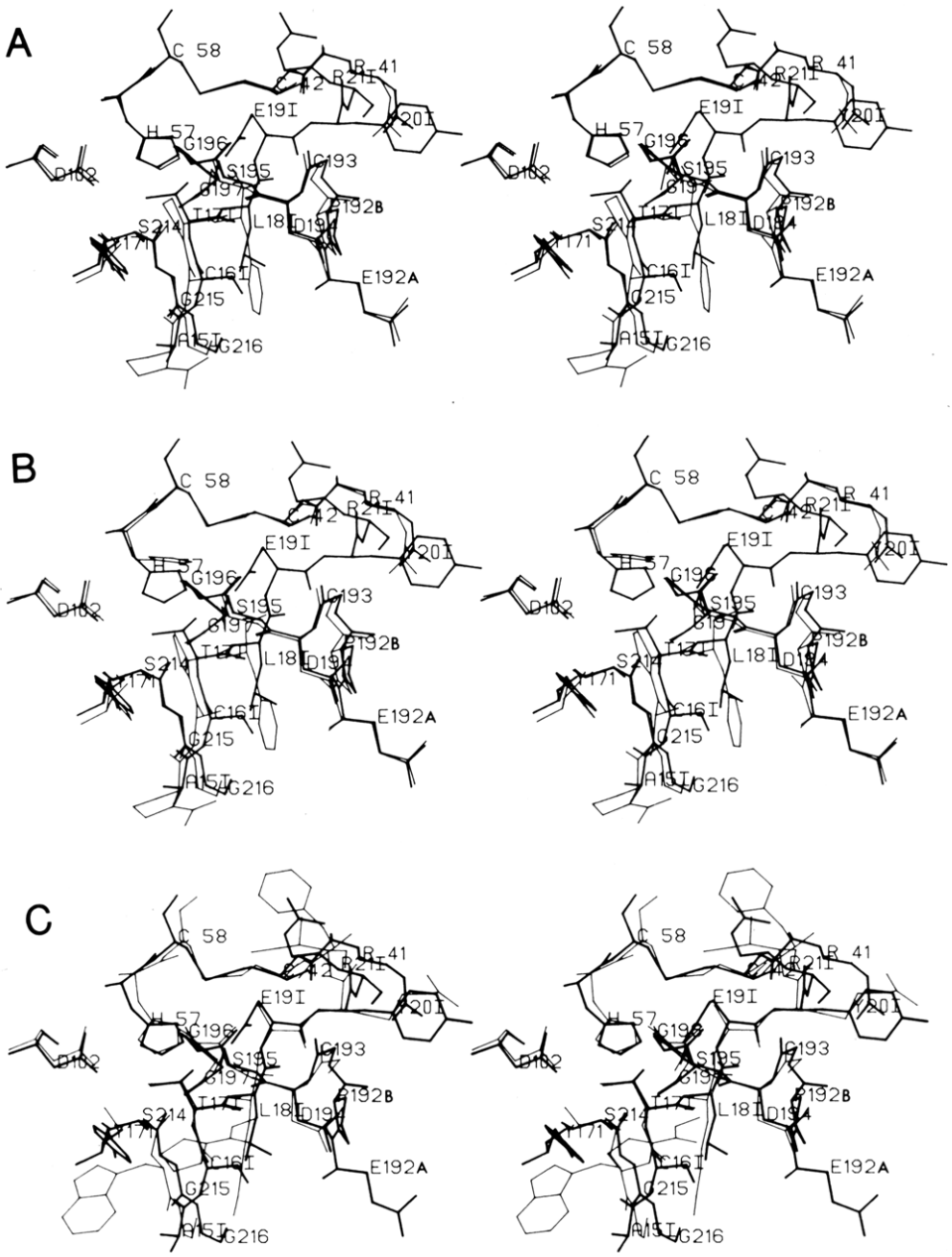

FIGURE 10: Comparison of active site regions of various complexes against OMTKY3–SGPB. The structures were superimposed by a least-squares procedure (program of W. Bennett) using α-carbon atoms of equivalent residues. The OMTKY3–SGPB complex is shown in thick lines with its residues labeled. Superimposed structures are shown in thin lines. (A) Tetrapeptide product Ac-Pro-Ala-Pro-Phe-OH bound to SGPA. (B) Tetrapeptide aldehyde Ac-Pro-Ala-Pro-Phe-H bound covalently to SGPA. (C) PTI–trypsin. The coordinates for this complex were obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977).

Table VII: Hydrogen-Bond and Short-Contact Distances (Å) in the Active Sites of Several Serine Protease Complexes

site		OMTKY3-SGPB	peptide I-SGPA[a]	peptide II-SGPA[b]	peptide IV-SGPA[c]	PTI-trypsin[d]
P_3	Cys-16I N- - -Gly-216 O	3.0	3.0	2.8	3.0	
	Cys-16I O- - -Gly-216 N	2.9	3.0	3.0	3.1	3.3
P_1	Leu-18I N- - -Ser-214 O	3.6	3.1	3.0	2.8	3.4
	Leu-18I O- - -Gly-193 N	2.6	2.8	2.8	3.1	2.9
	Leu-18I O- - -Ser-195 N	3.1	3.2	3.2	3.2	3.1
	Leu-18I C- - -Ser-195 O^{γ}	2.7	2.7	2.6	1.7	2.6
P_2'	Tyr-20I N- - -Arg-41 O	2.8				
	Tyr-20I O- - -Arg-41 N	3.2				3.0
P_6	Lys-13I N^{ζ}- - -Tyr-171 O	2.9				

[a] Peptide I = virtual substrate Ac-Pro-Ala-Pro-Phe-OH. [b] Peptide II = virtual substrate Ac-Pro-Ala-Pro-Tyr-OH. [c] Peptide IV = aldehyde inhibitor Ac-Pro-Ala-Pro-Phe-H. [d] Calculated from coordinates obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977).

of similar length in all of the noncovalent complexes (Table VII). We have concluded that there is an attractive force between Ser-195 O^{γ} and Leu-18I C (Fujinaga et al., 1982).

A pyramidal distortion of the carbonyl group of the reactive site has been reported in the structures of complexes of the various forms of trypsin with PTI (Huber & Bode, 1978). Such a distortion is not observed in the OMTKY3–SGPB complex. The discrepancy between our result and those for the various complexes of PTI may be due to the differences in the refinement procedures. In the refinement of the PTI structures, an extra parameter, θ^4 at the reactive site, was added. This parameter is the out-of-plane bend angle that the carbonyl oxygen atom makes with the plane defined by the C^{α}, C, and N atoms. The extra flexibility given to this group may be the cause of the distortion. In our refinement, we have not treated any group in a special way; however, the stereochemical restraints may force the structure to adopt an ideal conformation.

Two experiments were done to determine whether the geometrical restraints applied during refinement were preventing any real deviations from ideality from being observed. First, planar restraints on the atoms of the inhibitor were relaxed during several cycles of refinement. The resulting mean θ^4 of the main-chain peptide groups in the inhibitor was −1°, with a standard deviation of 8°. The maximum deviation from 0° was at Ser-44I with a θ^4 of 23°. The value for Leu-18I at the reactive site was −10°, as compared with −5° obtained in the restrained refinement.

The second approach was to determine whether real deviations from ideality could be observed under the conditions of restrained refinement. To do this, we modified our refined structure so that θ^4 at Leu-18I was −40°, similar to the value of −35° reported for this angle in PTI (Huber & Bode, 1978). These modified coordinates were used to calculate structure factors that were then used as "observed" structure factors for refinement. The starting model was the unmodified structure with θ^4 at Leu-18I of −5°. The weights were those specified in Table IV. After six cycles of refinement, the mean θ^4 was 0° with a standard deviation of 2°; the value at Leu-18I was −14°. When this experiment was repeated with relaxed planar restraints, the mean θ^4 remained 0° with a standard deviation of 5°, and the value at Leu-18I became −28°. Thus, even though the refinement procedure damps actual deviations from ideality, the presence of the deviation is still apparent. It should be noted that in this experiment we have an ideal case with error-free data and a nearly perfect model. In a normal refinement, one has the added complication of errors in the data.

In addition to the above arguments, it is possible to calculate the uncertainty in θ^4 due to coordinate errors. With a conservative value of 0.1 Å and the assumption of independent atoms, the estimated standard deviation in θ^4 is 8°, at $\theta^4 \approx 0°$. Thus, within the accuracy of our coordinates, the carbonyl carbon of Leu-18I has a planar configuration. This planarity is not imposed by the stereochemical restraints used in refinement, and it should be possible to observe a large distortion.

Comparison of SGPB Structures. The crystal of native SGPB and the present crystal of its complex with OMTKY3 differ in crystallization conditions, in packing interactions, and in the presence or absence of the inhibitor molecule, all of which are expected to affect the conformation to some degree. A comparison of the two structures can indicate the nature and extent of these effects.

The level at which two structures can be compared depends on their accuracy. Both structures of SGPB are well-refined, high-resolution structures with estimated rms coordinate errors of 0.1–0.2 Å, so that comparisons can be quite detailed. The native SGPB structure has been refined at 1.7-Å resolution to an R factor of 0.149 (L. Sawyer, A. R. Sielecki, and M. N. G. James, unpublished results).

When the two SGPB structures are compared by at least-squares superposition of all 1310 atoms (program of W. Bennett), the rms deviation in atomic positions is 0.58 Å (for the 741 main-chain atoms, the rms deviation is 0.52 Å; for the 569 side-chain atoms, it is 0.66 Å). The relative orientation derived from this superposition has been used in a detailed structural comparison.

(*i*) *Effect of Crystal Packing.* In order to study the effect of crystal-packing contacts on structure, we have applied the symmetry operations for each crystal to the appropriately oriented SGPB molecule from the other crystal. In this way, contacts that are present can be compared with those that would exist if the conformation were that of SGPB in the other crystal.

When such comparisons are made, one can see that differences in crystal-packing environment are often accompanied by changes in van der Waals and hydrogen-bonding interactions. It is generally evident that changes are necessary to avoid unfavorable contacts, but it is sometimes difficult to ascribe cause and effect.

Some of the largest conformational differences observed involve the C-terminal residues of SGPB. In the crystal of the native enzyme there is an intermolecular contact between the segment Gly-172-Gly-173 and the C-terminal residues Val-241 and Tyr-242 (Figure 11). In the crystal of the complex, both segments are essentially exposed to solvent, although Tyr-171 is in contact with the bound inhibitor (Table VI). If SGPB with the conformation observed in the complex were placed in the crystal of the native enzyme, unfavorable van der Waals contacts would exist (Figure 11). The conformation in the native structure avoids these unfavorable interactions and, in addition, allows Val-241 O to form hy-

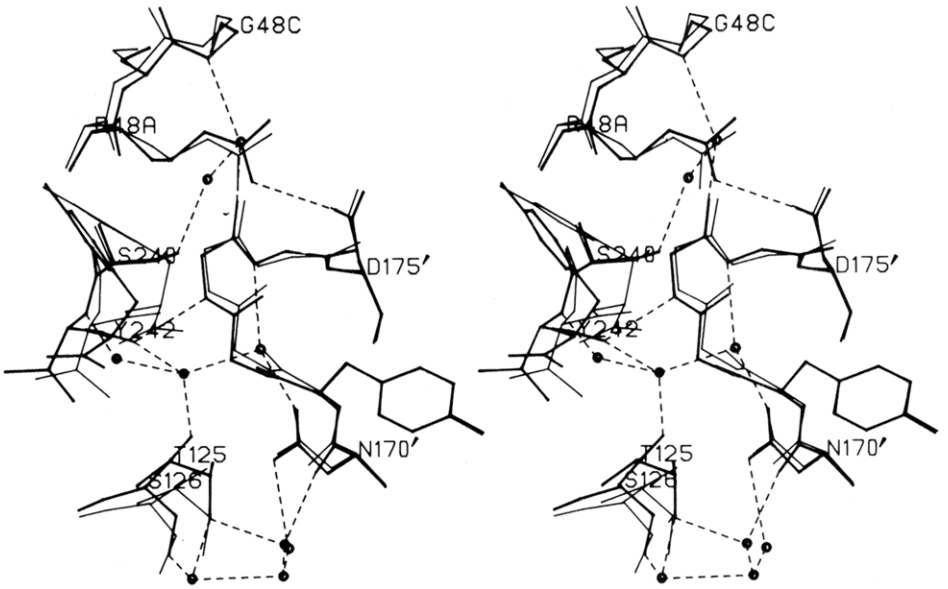

FIGURE 11: Intermolecular contact interface in the packing of crystals of native SGPB. The segment of the molecule denoted N-170′ to D-175′ is related to the molecule at x, y, z by the following coordinate transformation: $x, y, -1 + z$. Thick lines in the representation indicate the structure observed in the native enzyme. Thin lines correspond to the superimposed structure of SGPB in the conformation of its complex with OMTKY3. Had the C-terminal residues Val-241 and Tyr-242 retained their orientation observed in the crystals of the complex, unacceptably short nonbonded contacts with Gly-172′ and Gly-173′ would occur (Table VIII). Only those hydrogen bonds (dashed lines) and solvent molecules (circles) involved in bridging the interface are shown.

Table VIII: Intermolecular Contacts in the Crystal of Native SGPB

atom 1	atom 2[a]	obsd distance (Å)	potential contact (Å)[b]	shift of atom 1 (Å)	shift of atom 2 (Å)
Val-241 C	Gly-173 C^{α}	4.3	3.4	1.1	0.5
Val-241 C	Gly-173 N	3.8	2.8	1.1	0.4
Val-241 O	Gly-172 C	3.7	3.0	1.4	0.5
Val-241 O[c]	Gly-173 N	2.9	1.9	1.4	0.4
Val-241 O	Gly-173 C^{α}	3.7	2.6	1.4	0.5
Tyr-242 C^{α}	Gly-172 C^{α}	4.3	3.2	2.0	0.5
Tyr-242 C^{α}	Gly-172 C	4.5	3.0	2.0	0.5
Tyr-242 C^{α}	Gly-173 N	4.1	2.8	2.0	0.4
Tyr-242 C^{α}	Gly-173 C^{α}	4.8	3.5	2.0	0.5
Tyr-242 C^{β}	Gly-172 C	3.9	3.2	1.9	0.5
Tyr-242 C^{β}	Gly-173 N	3.6	2.9	1.9	0.4
Tyr-242 C^{β}	Gly-173 C^{α}	4.0	3.0	1.9	0.5

[a] Coordinates generated by the symmetry operation $x, y, -1 + z$. [b] Distances calculated with coordinates of both atoms from the superimposed structure of SGPB in the complex. [c] Hydrogen-bonded interaction in the crystal of native SGPB.

drogen bonds both to Gly-173 N and to a water molecule involved in the solvent structure of the interface (Table VIII). The small shift of Gly-172 and Gly-173 toward the contacting protein molecule might be attributed to the inhibitor binding interaction at Tyr-171 (see below). It is not clear why the C terminus, instead of the diglycyl β-bend, adjusts to avoid the potential bad contacts.

The shift of the C-terminal residues propagates by intramolecular contacts to the segment from Lys-115 to Ala-127 (Figure 12). Without concerted shifts of Lys-115 and, especially, Gly-117, prohibitively close contacts would result (Table IX). These shifts lead to a reorganization of the whole segment, even those residues relatively far from the C terminus. In fact, the largest main-chain difference between the two SGPB structures (2.8 Å for Gly-120 C) occurs in the β-bend from Val-119 to Gln-122. As can be seen in Figure 12, as well as from the data in Table X, the pattern of hydrogen bonding involving the residues Lys-115 to Ala-127 is also somewhat altered. One of these hydrogen bonds, that from the carboxyl group of Tyr-242 to Gly-117 O in native SGPB, is perhaps a contributing factor in the reorganization. The existence of this hydrogen bond implies that the carboxy terminus is protonated, which is possible for the native SGPB structure,

Table IX: Contacts Leading to Propagation of Shifts from the C Terminus to the Segment Lys-115 to Ala-127 in the Native Structure

atom 1	atom 2	obsd distance (Å)	potential distance (Å)[a]	shift of atom 2
Tyr-242 C	Gly-117 C^{α}	3.7	2.9	1.0
Tyr-242 C	Gly-117 C	3.9	3.1	0.8
Tyr-242 C	Gly-117 O	3.3	2.9	0.5
Tyr-242 O	Gly-117 C^{α}	3.2	2.6	1.0
Tyr-242 O	Gly-117 C	3.3	2.7	0.8
Tyr-242 O	Gly-117 O	2.6	2.1	0.5
Tyr-242 C^{ζ}	Lys-115 O	3.5	2.9	0.7
Tyr-242 C^{ϵ_2}	Lys-115 O	3.3	2.8	0.7
Tyr-242 C^{ϵ_2}	Asp-116 C	4.5	3.5	0.8
Tyr-242 C^{ϵ_2}	Gly-117 N	3.8	3.0	0.9
Tyr-242 C^{ϵ_2}	Gly-117 C^{α}	3.7	3.1	1.0
Tyr-242 C^{δ_2}	Gly-117 C^{α}	3.6	2.8	1.0
Tyr-242 O^{η}	Lys-115 C	3.7	3.1	0.6

[a] Distances calculated with coordinates of atom 2 from the superimposed structure of SGPB in the complex.

determined at a pH of 4.2, but not for the present structure (pH 6.3). It would appear from Figure 12 that the large movement of the β-bend, Val-119 to Glu-122, is in turn

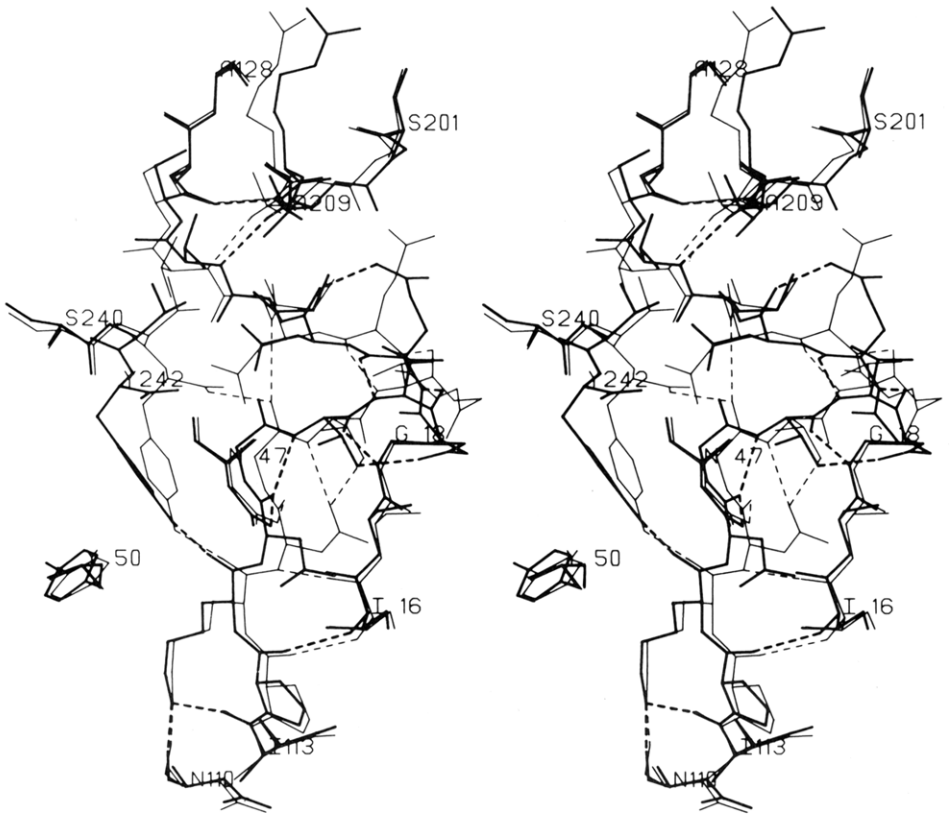

FIGURE 12: Propagation of shift of C-terminal residues Val-241 and Tyr-242 to the peptide chain Lys-115 to Ser-126. Thick lines represent the conformation of SGPB in the crystal of its complex with OMTKY3; thin lines indicate the native structure. Dashed lines represent hydrogen bonds. The major differences are seen for the β-bend Val-119–Glu-122, where conformational differences up to 2.8 Å occur. This β-bend makes few intramolecular contacts and is thus relatively free to move.

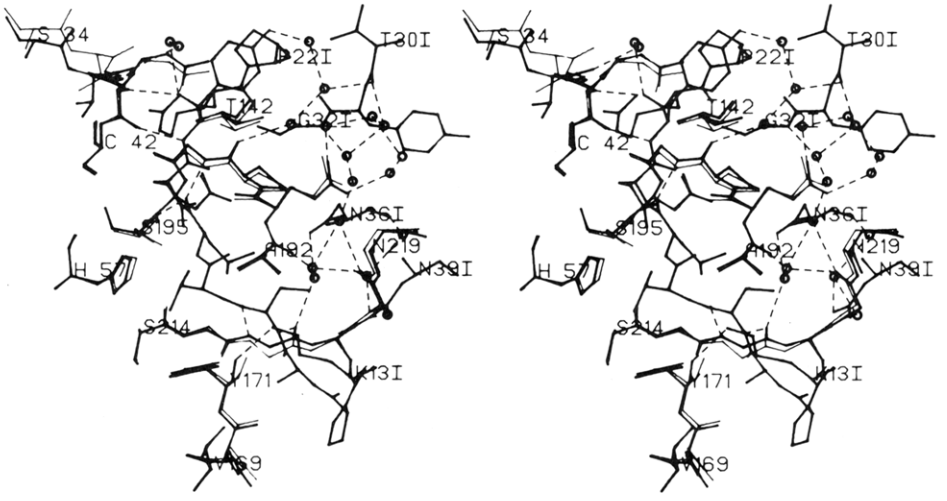

FIGURE 13: Intermolecular contacts at the interface of SGPB and OMTKY3. Thick lines correspond to the observed structure in the crystal of the complex; thin lines represent the suitably transformed coordinates of native SGPB for comparison. Conformational changes in SGPB induced upon inhibitor binding are relatively small (see text and Table XI). Only those hydrogen bonds (dashed lines) and solvent molecules (circles) involved in bridging the contact are included in the figure.

propagated through to Ser-126, as well as to the neighboring β-bend from Ser-201 to Arg-208.

The comparison of the two structures of SGPB shows that in the regions of intermolecular contact, the differences are not large and tend to be smaller than those discussed in the above example. For SGPB, there does not seem to be much propagation of the conformational changes from the surface residues to the more internal residues of the structural core.

(*ii*) *Effect of Inhibitor Binding.* The residues of SGPB that interact with OMTKY3 are mostly exposed to solvent in the crystal of the native enzyme. Therefore, their conformational changes should closely parallel the movements that occur on binding in solution. Figure 13 shows the contact region, with the native SGPB structure superimposed on that of the complex. Contact distances from OMTKY3 to the SGPB molecule in both conformations are given in Table XI.

The conformational changes of the catalytic residues, His-57 and Ser-195, are important in the hydrolytic mechanism. As discussed above, Ser-195 O^{γ} is only 2.7 Å away from Leu-18I C, in what appears to be an attractive interaction. Nonetheless, upon binding, this residue moves slightly away from the inhibitor. As well, the side chain of His-57 rotates away from close contacts with Thr-17I. The net result of these two movements is the formation of a strong hydrogen bond between

Table X: Hydrogen Bonds Involving Lys-115 to Ala-127 in the Crystal of the OMTKY3–SGPB Complex and in the Native Crystal

donor	acceptor	length (Å) OMTKY3–SGPB	length (Å) native SGPB
Gly-18 N	Thr-118 O	2.9	
Asn-47 $N^{\delta 2}$	Asp-116 O	2.6	2.6
Lys-115 N^{ζ}	Asn-110 O	2.8	2.9
Lys-115 N^{ζ}	Ile-113 O	2.8	2.6
Asp-116 N	Ile-16 O		2.9
Thr-118 N	Asp-116 O	3.0	
Thr-118 N	Asp-116 $O^{\delta 1}$		2.9
Thr-118 $O^{\gamma 1}$	Asp-116 $O^{\delta 1}$		2.6
Thr-118 $O^{\gamma 1}$	Gly-120 O	2.6	
Val-119 N	Gln-122 O	2.9	2.7
Gln-122 N	Val-119 O	2.8	3.1
Gln-122 $N^{\epsilon 2}$	Asp-123 O	2.9	
Ile-124 N	Gly-117 O		3.4
Thr-125 N	Thr-207 O	2.8	2.8
Ala-209 N	Ser-126 O	2.8	2.7
Tyr-242 O	Gly-117 O		2.6
Tyr-242 O^{η}	Lys-115 O	2.6	2.9

His-57 $N^{\epsilon 2}$ and Ser-195 O^{γ}. It has been proposed that the formation of this hydrogen bond is important in activating O^{γ} as a nucleophile and in allowing the transfer of the proton from Ser-195 via His-57 to the leaving-group amide during hydrolysis (Huber & Bode, 1978).

In the binding of OMTKY3 to SGPB, the conformational changes are mostly small, though significant. This observation favors the lock and key model of enzyme–substrate interaction (Fischer, 1894). On the other hand, the small conformational changes in the catalytic residues could have quite profound effects on the mechanism of hydrolysis.

Mechanism of Inhibition. Protein inhibitors of serine proteases have a common inhibitory mechanism (Laskowski & Kato, 1980). The inhibitor binds to the active site of the cognate enzyme in the manner of a good substrate, but very tightly, so that very little of the enzyme remains free. The observed structure of the complex of SGPB and OMTKY3 supports this general proposal for a mechanism. Figure 10a shows that the binding of the inhibitor to SGPB is very similar to that of a tetrapeptide product (virtual substrate) to the homologous enzyme SGPA.

Only free enzyme can catalyze peptide hydrolysis, and the inhibitor reduces the concentration of free enzyme by forming a stable complex with it. In outline, the reaction between enzyme and inhibitor can be represented as

$$E + I \rightleftharpoons E{\cdot}I \rightleftharpoons E + I^*$$

where I is the virgin inhibitor, and I* is the cleaved, or modified, inhibitor. E·I is the form of the enzyme–inhibitor complex that has the lowest free energy. In this scheme, the various catalytic intermediates along the path of peptide hydrolysis (Kraut, 1977), as well as the loose, noncovalent complexes that are predicted from kinetic data (Luthy et al., 1973), have been omitted. Individual rate constants derived according to such a scheme would thus be suspect, but equilibria are still valid. The amount of complex E·I that is present at equilibrium is governed by

$$K_{assoc} = [E{\cdot}I]/([E][I + I^*]) \quad (2)$$

(Finkenstadt et al., 1974). For SGPB and OMTKY3, K_{assoc} is 6×10^{10} M^{-1} (W. Ardelt and M. Laskowski, Jr., personal communication). The overall equilibrium for the hydrolysis of inhibitor is given by

$$K_{hyd} = [I^*]/[I] \quad (3)$$

Finally, the equilibrium constant for the association of virgin inhibitor with enzyme can be denoted as

Table XI: Intermolecular Contacts between SGPB and OMTKY3

SGPB	OMTKY3	obsd distance (Å)	potential contact (Å)[a]	shift of SGPB atom (Å)	
Thr-39 $C^{\gamma 2}$	Arg-21I $N^{\eta 1}$	4.1	2.9	2.7	main-chain and side-chain movements alleviate close contact; note, however, that the preceding residue, Ser-34, forms an H bond to a neighboring molecule of SGPB in the crystal of the complex
His-57 $C^{\epsilon 1}$	Thr-17I C^{β}	3.8	3.5	0.4	His-57 side-chain reorientation avoids close contact, as well as making H bonding to Ser-195 possible
His-57 $N^{\epsilon 2}$	Thr-17I C^{β}	3.6	3.3	0.5	
His-57 $C^{\delta 2}$	Thr-17I C^{β}	3.6	3.5	0.3	
Tyr-171 C	Lys-13I C^{δ}	4.0	3.6	0.5	main-chain shift relieves bad contacts, while allowing H-bond formation
Tyr-171 O	Lys-13I C^{δ}	3.5	2.9	0.7	
Tyr-171 O	Lys-13I C^{ϵ}	3.8	3.3	0.7	
Tyr-171 O^b	Lys-13I N^{ζ}	2.9	2.6	0.7	
Pro-192B C^{α}	Leu-18I C	3.9	3.7	0.4	several unfavorable contacts relieved by main-chain shift; however, the interpretation is complicated by the fact that this segment lies near a crystallographic 2-fold axis in the native crystal and there is a carboxyl–carboxylate interaction between Glu-192A and its 2-fold related mate (Sawyer & James, 1982); as well, Thr-142, which is H bonded to Pro-192B, shifts to avoid contacts with the 2-fold related molecule (Figure 13)
Pro-192B C	Leu-18I O	3.5	3.2	0.4	
Gly-193 N	Leu-18I C	3.6	3.2	0.5	
Gly-193 N^b	Leu-18I O	2.6	2.2	0.5	
Gly-193 C^{α}	Leu-18I O	3.4	2.9	0.6	
Gly-193 C	Leu-18I O	3.5	3.3	0.3	
Gly-215 C^{α}	Leu-18I C^{γ}	3.9	3.7	0.4	main-chain shift alleviates close contacts
Gly-215 C^{α}	Leu-18I $C^{\delta 1}$	3.8	3.5	0.4	
Gly-216 C^{α}	Ala-15I C^{β}	4.0	3.4	0.6	
Gly-216 C	Ala-15I C^{β}	3.8	3.5	0.4	
Gly-216 O	Ala-15I C^{α}	3.5	3.2	0.3	

[a] Distances calculated with coordinates of the superimposed native SGPB structure. [b] Hydrogen-bonded interaction in the complex.

$$K_a = [E{\cdot}I]/([E][I]) \quad (4)$$

From eq 2–4, we can derive

$$K_{assoc} = K_a/(1 + K_{hyd}) \quad (5)$$

The function of the inhibitor is to maximize K_{assoc}, thus minimizing [E], the concentration of free enzyme. With the insight provided by eq 5, this problem can be broken down into two parts: (1) maximizing K_a and (2) minimizing K_{hyd}. The first part involves the specific interactions between the enzyme and its inhibitor, while the second is a function only of the inhibitor structure.

The question of maximizing K_a can be examined by considering how the differences between the inhibitor and a small peptide substrate lead to tighter binding of the inhibitor. The most obvious proposal would be that the inhibitor makes a larger number of favorable contacts with the enzyme. However, almost all (100 out of 108) of the intermolecular contacts less than 4 Å (Table VI) are made by residues P_6 through P_3', i.e., relatively close to the scissile bond. Therefore, most of the interactions with SGPB that are made by OMTKY3 would also be available to a small peptide substrate. For example, there are 78 nonbonded contacts less than 4 Å between SGPA and the tetrapeptide product Ac-Pro-Ala-Pro-Tyr-OH (James et al., 1980b). Thus, the intrinsic binding energy (Jencks, 1975) of the inhibitor, the sum of the favorable free-energy terms, is expected to be similar to that of a good substrate. Normally, however, a significant portion of the intrinsic binding energy would be used to reduce the entropy of the substrate, locking it into its reactive conformation (Jencks, 1975). But, the inhibitor is conformationally much less mobile, so that much more of the intrinsic binding energy is realized as observed binding energy. This concept has generally been expressed in terms of a rigid enzyme meeting a rigid, and complementary, substrate (Blow, 1974). It is important to note that this is not an absolute inflexibility but a relative one, when compared to good substrates such as small peptides or floppy external loops of globular proteins. Indeed, in the native structure of the closely related OMJPQ3, the reactive site region shows considerable flexibility, as indicated by the thermal motion parameters (Weber et al., 1981; Papamokos et al., 1982).

K_{hyd} for inhibitors is generally close to unity near neutral pH (Finkenstadt et al., 1974; Laskowski & Kato, 1980). It can be seen from eq 5 that the value of K_{assoc} would not, in fact, be increased much by decreasing K_{hyd} below this value. The region of the reactive site in inhibitors must have considerably less freedom to relax after hydrolysis than regions containing hydrolyzable bonds in globular proteins, which undergo virtually complete hydrolysis (Finkenstadt et al., 1974). It is not difficult to rationalize this relative inflexibility in terms of the structure of OMTKY3. To begin with, the two sides of the reactive bond are fixed to the body of the inhibitor by disulfide bridges at Cys-16I (P_3) and Cys-24I (P_6'). The residues between these disulfides form six hydrogen bonds to the rest of the inhibitor (see Table V). Moreover, one can see in Figure 7 that the residues from Pro-22I to Cys-24I are sandwiched between the segments Lys-29I to Lys-34I and Leu-50I to Cys-56I. Leu-23I and Cys-24I are part of an antiparallel β-structure. These covalent and noncovalent interactions in the region of the scissile bond would reduce the conformational freedom of the cleaved inhibitor, thus making hydrolysis less favorable.

It is possible to explain the general properties of protein inhibitors of serine proteases on the basis of the thermodynamic parameters for their hydrolysis. In addition, it is possible to rationalize these equilibrium constants in terms of the molecular structure. A more complete understanding of the interactions between SGPB and OMTKY3 would require a knowledge of the structures of all of the intermediates and the individual rate constants for each of the steps. Unfortunately, detailed kinetic data for this specific enzyme–inhibitor pair are not yet available, and we know only the present structure of OMTKY3–SGPB. The next step on the pathway to hydrolysis from the observed complex would be the completion of the nucleophilic attack of Ser-195 O^{γ} on the carbonyl carbon of Leu-18I to form the tetrahedral intermediate. We have described previously the relative movements of enzyme and inhibitor that would be required to form the tetrahedral intermediate and some interactions that would hinder these movements (Fujinaga et al., 1982). Part of these interactions involve hydrogen bonds from the side-chain amide of Asn-33I to the carbonyl oxygen atoms of Thr-17I and Glu-19I that would oppose the movement of the carbonyl carbon of Leu-18I toward Ser-195 O^{γ}. The position of this amide group is stabilized, in turn, by a hydrogen bond from Asn-36I NH to Asn-33I $O^{\delta 1}$. Thus, the central α-helix serves as an anchor for the reactive bond.

The specific hindrance that we predict for the formation of the tetrahedral intermediate will not affect the equilibrium properties that make OMTKY3 an inhibitor. However, if the interactions we have described increase the activation energy barrier to a level comparable to or greater than the highest energy level on the reaction coordinate, the approach to the equilibrium between E·I and E + I* would be slowed. In fact, it has been observed that the formation of the complex from modified inhibitor is generally slower than from intact inhibitor (Empie & Laskowski, 1982). Prevention of the formation of I* can be advantageous if I* is more susceptible to additional peptide cleavage, as is the case for PSTI and other "temporary" inhibitors.

The proposal that protein inhibitors of serine proteases bind to their cognate enzymes in the manner of a good substrate (Laskowski & Kato, 1980) is strongly supported by the present analysis. There is no covalent interaction between the two in the minimum free energy complex that we have isolated. The very high value of the association constant seems to be due primarily to the exceptional complementarity of the enzyme and inhibitor surfaces.

Acknowledgments

We thank W. Ardelt and M. Laskowski, Jr., for the gift of purified OMTKY3 and for many stimulating discussions. Koto Hayakawa grew the crystals used for this work. C. G. Broughton has provided us with exceptional graphic-system programs. We thank R. Huber for sending us the manuscript on the refined OMJPQ3 structure prior to publication.

Supplementary Material Available

One table containing the residues of SGPB and OMTKY3 and their conformational angles ϕ, ψ, ω, and χ^i (4 pages). Ordering information is given on any current masthead page.

References

Barry, C. D., Molnar, C. E., & Rosenberger, F. U. (1976) *Technical Memo No. 229*, Computer Systems Lab, Washington University, St. Louis, MO.

Bauer, C.-A. (1978) *Biochemistry 17*, 375–380.

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

Blow, D. M. (1974) *Bayer-Symp. 5*, 677–678.
Bode, W., Schwager, P., & Huber, R. (1978) *J. Mol. Biol. 118*, 99–112.
Bogard, W. C., Jr., Kato, I., & Laskowski, M., Jr. (1980) *J. Biol. Chem. 255*, 6569–6574.
Bolognesi, M., Gatti, G., Menegatti, E., Guarneri, M., Marquart, M., Papamokos, E., & Huber, R. (1982) *J. Mol. Biol. 162*, 839–868.
Chambers, J. L., & Stroud, R. M. (1979) *Acta Crystallogr., Sect. B B35*, 1861–1874.
Crowther, R. A. (1973) in *The Molecular Replacement Method* (Rossmann, M. G., Ed.) International Science Review 13, pp 173–178, Gordon & Breach, New York.
Cruickshank, D. W. J. (1949) *Acta Crystallogr. 2*, 65–82.
Cruickshank, D. W. J. (1954) *Acta Crystallogr. 7*, 519.
Cruickshank, D. W. J. (1967) in *International Tables for X-ray Crystallography* (Kasper, J. S., & Lonsdale, K., Eds.) Vol. II, pp 318–340, Kynoch Press, Birmingham, England.
Delbaere, L. T. J., Hutcheon, W. L. B., James, M. N. G., & Thiessen, W. E. (1975) *Nature (London) 257*, 758–763.
Empie, M. W., & Laskowski, M., Jr. (1982) *Biochemistry 21*, 2274–2284.
Finkenstadt, W. R., Hamid, M. A., Mattis, J. A., Schrode, J., Sealock, R. W., Wang, D., & Laskowski, M., Jr. (1974) *Bayer-Symp. 5*, 389–411.
Fischer, E. (1894) *Chem. Ber. 27*, 2985–2993.
Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M., Jr., & James, M. N. G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4868–4872.
Furey, W., Jr., Wang, B. C., & Sax, M. (1982) *J. Appl. Crystallogr. 15*, 160–166.
Hendrickson, W. A. (1976) *J. Mol. Biol. 106*, 889–893.
Hendrickson, W. A., & Konnert, J. H. (1980) in *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., Ed.) Vol. I, pp 43–57, Pergamon Press, Oxford.
Hirono, S., Nakamura, K. T., Iitaka, Y., & Mitsui, Y. (1979) *J. Mol. Biol. 131*, 855–869.
Huber, R., & Bode, W. (1978) *Acc. Chem. Res. 11*, 114–122.
Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., & Steigemann, W. (1974) *J. Mol. Biol. 89*, 73–101.
Huber, R., Bode, W., Kukla, D., Kohl, U., & Ryan, C. A. (1975) *Biophys. Struct. Mech. 1*, 189–201.
James, M. N. G., Delbaere, L. T. J., & Brayer, G. D. (1978) *Can. J. Biochem. 56*, 396–402.
James, M. N. G., Brayer, G. D., Delbaere, L. T. J., Sielecki, A. R., & Gertler, A. (1980a) *J. Mol. Biol. 139*, 423–438.
James, M. N. G., Sielecki, A. R., Brayer, G. D., Delbaere, L. T. J., & Bauer, C.-A. (1980b) *J. Mol. Biol. 144*, 43–88.
Janin, J., Wodak, S., Levitt, M., & Maigret, B. (1978) *J. Mol. Biol. 125*, 357–386.
Jencks, W. P. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol. 43*, 219–410.
Johnson, C. K. (1965) *ORTEP*, Report ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, TN.
Kato, I., Kohr, W. J., & Laskowski, M., Jr. (1978) *Proc. FEBS Meet. 47*, 197–206.
Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331–358.
Laskowski, M., Jr., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593–626.
Luthy, J. A., Praissman, M., Finkenstadt, W. R., & Laskowski, M., Jr. (1973) *J. Biol. Chem. 248*, 1760–1771.
Luzzati, V. (1952) *Acta Crystallogr. 5*, 802–810.
North, A. C. T., Phillips, D. C., & Mathews, F. S. (1968) *Acta Crystallogr., Sect. A A24*, 351–359.
Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I., & Laskowski, M., Jr. (1982) *J. Mol. Biol. 158*, 515–537.
Pauling, L. (1946) *Chem. Eng. News 24*, 1375.
Rao, S. N., Jih, J.-H., & Hartsuck, J. A. (1980) *Acta Crystallogr., Sect. A A36*, 878–884.
Robertus, J. D., Kraut, J., Alden, R. A., & Birktoft, J. J. (1972) *Biochemistry 11*, 4293–4303.
Ruhlmann, A., Kukla, D., Schwager, P., Bartels, K., & Huber, R. (1973) *J. Mol. Biol. 77*, 417–436.
Sawyer, L., & James, M. N. G. (1982) *Nature (London) 295*, 79–80.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.
Sielecki, A. R., James, M. N. G., & Broughton, C. G. (1982) in *Crystallographic Computing, Proceedings of the International Summer School* (Sayre, D., Ed.) pp 409–419, Carleton University, Ottawa, Oxford University Press, Oxford.
Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H., & Blow, D. M. (1974) *Biochemistry 13*, 4212–4228.
Thiessen, W. E., & Levy, H. A. (1973) *J. Appl. Crystallogr. 6*, 309.
Weber, E., Papamokos, E., Bode, W., Huber, R., Kato, I., & Laskowski, M., Jr. (1981) *J. Mol. Biol. 149*, 109–123.