The crystal structure of the transition state analog complex formed covalently between proteinase K and methoxysuccinyl-Ala-Ala-Pro-Ala-chloromethyl ketone was determined by x-ray diffraction methods at a resolution of 2.2 Å and refined by constrained least squares to an R factor of 19.8% for the 11864 structure amplitudes greater than 1e-5. The chloromethyl ketone group is covalently linked with the active site functional groups His<sup>68</sup>(N<sub>y</sub>) and Ser<sup>224</sup>(O<sub>y</sub>). The former has substituted for chlorine and the latter has attacked the tetrahedral carbon atom of the transition state analog. The peptide part of the inhibitor is in an extended conformation and fills subsites S<sub>1</sub> to S<sub>5</sub> of the substrate recognition site. Its backbone hydrogens bond with the bulky proline side chain. The methoxysuccinyl group is stacked on the phenolic side chain of Tyr<sup>104</sup> that is a part of the bottom of the recognition site. Biochemical studies show that shorter inhibitors of this type are less effective than the longer one, because there are fewer hydrogen bonding and van der Waals/stacking interactions.

The inhibition of proteases by specifically designed molecules is of particular interest in view of the involvement of these enzymes in biological processes which frequently are of medical importance. One class of inhibitors, the peptide chloromethyl ketones, has been more intensively studied for two reasons: (i) they bind covalently to the catalytically active histidine and serine residues in the ubiquitous serine proteases and are good transition state analogs, and (ii) their peptide sequence can be synthesized according to the substrate specificity of the target enzyme (1-3). For the subtilisin-related proteinase K, the binding of two such inhibitors, carbobenzox-Ala-Ala-chloromethyl ketone (4) and its Ala-Phe derivative (5), were investigated by crystallographic methods to locate the active site of the enzyme and to characterize the substrate binding site.

In the present study, we elongated the inhibitor to map and to saturate the substrate binding site of proteinase K. The search for an optimum inhibitor was guided by manual model building using FRodo (6) and an Evans and Sutherland graphics unit and subsequent energy minimization studies using the BIOSYM program package (7). These studies indicate that a chloromethyl ketone inhibitor, bound to the active site residues His<sup>68</sup> and Ser<sup>224</sup> in a way similar to the already known shorter inhibitors (4, 5), fits completely into the substrate binding site even if the amino acid in the subsite S<sub>2</sub> (according to the Schechter and Berger nomenclature (8)) is a proline, followed by two to four amino acids with small side chains. To test this suggestion obtained from model building studies, the inhibitory power against proteinase K of the commercially available N-terminal protected methoxysuccinyl-Ala-Ala-Pro-Ala-chloromethyl ketone was investigated by kinetic assays. Since it strongly inhibits proteinase K, the two molecules were cocrystallized and the structure of the complex determined by x-ray diffraction methods.

**MATERIALS AND METHODS**

Proteinase K was obtained from SERVA, Heidelberg, and purified by Sephadex-G50 gel filtration in 50 mM Tris-HCl, pH 7.5, containing 1 mM CaCl<sub>2</sub> to remove autolytic fragments, as previously described (9), and lyophilized.

The inhibitors Ala-Ala-COOH (I), Ala-Ala-Phe-COOH (II), Phe-Pro-Arg-COOH (III), methoxysuccinyl-Ala-Ala-Pro-Ala-COOCH<sub>2</sub>Cl (IV), were purchased from SERVA, Heidelberg, and from Calbiochem, Bubendorf (Switzerland), and used without further purification. The latter inhibitor is soluble in water, and the others were first dissolved in a minimum volume of methanol and then added to aqueous, buffered solutions of proteinase K. For inhibition studies, inhibitors I-III were dissolved 0.5 µg/µl in 50 mM Tris-HCl, 5 mM CaCl<sub>2</sub>, pH 8.0, containing 60% methanol. Stock solutions of proteinase K (0.2 µg/µl) and of the substrate succinyl-(Ala)-CO-NH-(CH<sub>2</sub>)-NO<sub>2</sub> (1 mM) were prepared in the same buffer without methanol; in all experiments the molar ratio of inhibitor to enzyme was 100:1. Assays contained 0.7 mM of proteinase K, 70 mM of inhibitor, 0.75 µmol of substrate, 0.04-4% methanol. After enzyme and inhibitor were incubated for 15 min at 25 °C, the substrate succinyl-(Ala)-CO-NH-(CH<sub>2</sub>)-NO<sub>2</sub> was added and reacted 1 h. The reaction was then stopped with glacial acetic acid and the released p-nitrophenol detected using 40 nm with a Beckman DU6 spectrophotometer. Relative inhibition rates are given in Table I.

For crystallization of the complex between proteinase K and methoxysuccinyl-Ala-Ala-Pro-Ala-COOCH<sub>2</sub>Cl, the lyophilized enzyme was dissolved at 5% (w/v) in 50 mM Tris-HCl, 1 mM CaCl<sub>2</sub>, pH 7.5, with 2-fold molar excess of the inhibitor. 20-µl drops of this solution were equilibrated in the sitting drop vapor diffusion method against 0.75 M Na<sub>2</sub>SO<sub>4</sub> in the same buffer. Single crystals of size 0.8 × 0.5 × 0.6
mm$^3$ grew within 1 day at room temperature. The crystals of diameter ~0.8 mm are isomorphous to crystals of native proteinase K (10, 11). X-ray data collection to 2.2-Å resolution was carried out by film methods using an Arndt-Wonacott rotation camera installed on an Elliott GX20 rotating anode x-ray generator (see Table II). The x-ray intensities on the films were digitized with an Optronics PlOOO film scanner and processed with the MOSCO program system (12).

The coordinates of all the protein and two Ca$^{2+}$ atoms of the 1.5-Å resolution structure of proteinase K (11) were used as a starting model in the structure determination. For the preliminary stages of restrained least-squares refinement (13), only the 7809 data $\sigma$ > 2σ and within the 5-2.2-Å resolution shell were used, with an initial $R$ factor of 0.283. After 8 cycles of refinement, the $R$ factor converged at 0.198 and a 2$F_o$ - $F_c$ electron density map was calculated using the CCP4 suite of protein crystallographic programs and interpreted with FRODO (8) by means of an Evans and Sutherland Ps350 graphics display. The map clearly revealed the Ala-Ala-Pro-Ala portion of the inhibitor and the covalent bonds of the -CO-$\text{CH}_2$- end group with His$^{60}(N_{\text{E}})$ and Ser$^{224}(O_{\text{S}})$, but the methoxysuccinyl group could not be located. A further 5 cycles of refinement with restrained covalent enzyme-inhibitor bonds led to a map in which electron density for the protecting group could be interpreted (Fig. 1). The subsequent refinement was carried out with the reciprocal space least-squares program TREET (14) because it allows for better definition of the restraints within the inhibitor binding site. At this stage all the 11864 x-ray data above the 1σ level were successively included and the resolution range extended to 10 Å. Inspection of 2$F_o$ - $F_c$ maps not only allowed for improvement of the position of the inhibitor, but also to locate 170 water molecules and the terminal six atoms of the Arg$^{176}$ side chain which were not defined in the structure of native proteinase K. The strong Ca$^{2+}$ binding site Ca2 is fully occupied but the weak binding site Ca2 appears to be filled by a water molecule, as indicated by the occupation/thermal parameters for this site.

The final $R$ factor for the refined structure containing 2022 protein atoms, 31 inhibitor atoms, 170 water oxygen atoms, and one Ca$^{2+}$ atom is 0.198, based on the 11864 reflections with $F$ > 2σ. The atomic coordinates of the complex will be deposited with the Brookhaven protein data file.

**RESULTS AND DISCUSSION**

In the previous crystallographic studies on native proteinase K at 1.5-Å resolution (11) and on proteinase K complexed with the inhibitor carbobenzoxy-Ala-Ala-COCH$_2$Cl (4), the catalytic site was identified as the triad Asp$^{22}$His$^{60}$Ser$^{224}$, with the free Cys$^{41}(SH)$ near the imidazole ring of His$^{60}$. The substrate recognition site is formed by the two peptide chains, 99-104 and 132-134, respectively (Fig. 2, a and b).

These peptide chains are oriented approximately parallel and directly connected only by a hydrogen bond formed between the Tyr$^{104}(O_{\text{S}})$ and Gly$^{132}(O)$ (Fig. 2a). The previously studied peptide inhibitors (4, 5) and the one presently investigated are inserted into the recognition site such that an antiparallel three-stranded pleated sheet is formed. The sheet has a commonly observed overall left-handed twist.

In the reaction between proteinase K and the inhibitor methoxysuccinyl-Ala$^{261}$-Ala$^{262}$-Pro$^{263}$-Ala$^{264}$-COCH$_2$Cl, the active site residues His$^{60}(N_{\text{E}})$ and Ser$^{224}(O_{\text{S}})$ react with the terminal -COCH$_2$Cl group, to form a covalently bonded complex with proteinase K as shown in Fig. 2b. This complex mimics the transition state with tetrahedral C$^{\delta+}$ atom, except for the -CH$_2$- group which, in the transition state, would be the -NH of the leaving group (the product). The -CH$_2$-His$^{60}(N_{\text{E}})$ bond restricts the orientation of this group so that the C$^{\delta+}(O)$ oxygen cannot fully move into the "oxyanion hole" which is formed by Ser$^{224}(N)$, Asn$^{146}(N)$, and Asn$^{146}(N)$ in the native proteinase K, this hole is filled by a water molecule (11). Instead, the Ala$^{264}(O_{\text{S}})$ atom forms hydrogen bonds only to Ser$^{224}(N)$, and because it cannot move further into the "oxyanion hole" to approach and to bind to Asn$^{146}(N)$ and Asn$^{146}(N)$, the resulting gap is filled by a well defined water molecule Wat$^{265} (B = 23.4 \AA^2)$ in hydrogen bonding distance, see Table III.

The covalent binding geometry parameters within the residues His$^{60}$ Ser$^{224}$, and Ala$^{264}$ were refined with constraints. As expected, the bond angles around atom Ala$^{264}(C)$ are in the range 107-112°, the angle at Ser$^{224}(O)$ is 107°, and that at the -CH$_2$- atom Ala$^{264}(C)$ is 116°; the CH$_2$-N bond distance...
is 1.4 Å, the C-O distances are 1.35–1.36 Å, and the C-C distances fall in the range 1.50–1.55 Å. The inhibitor adopts an extended conformation with trans Pro$^{285}$, and all Ψ angles correspond to the β-pleated sheet region in a Ramachandran plot (Ala$^{283}$: ϕ = −100°, Ψ 178°; Ala$^{282}$: ϕ = −146°, Ψ 163°; Pro$^{285}$: ϕ = −80°, Ψ 151°; Ala$^{284}$: ϕ = −125°). In addition to the covalent bonds with active site residues Ser$^{224}$ and His$^{39}$, the inhibitor is anchored in the substrate recognition site by several hydrogen bonds, details are given in Table III.

These hydrogen bonds involve, in subsite P1 (for definition

<table>
<thead>
<tr>
<th>Hydrogen bond D–H...A between inhibitor, water molecules, and proteinase K</th>
<th>H atom positions were calculated with MOLEDT (7).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atoms involved</td>
<td>D...A</td>
</tr>
<tr>
<td>MSU$^{280}$(O)–Tyr$^{104}$(N)$^*$</td>
<td>3.31</td>
</tr>
<tr>
<td>MSU$^{280}$(O)–Wat$^{463}$</td>
<td>2.99</td>
</tr>
<tr>
<td>Gly$^{134}$(N)–Wat$^{448}$</td>
<td>3.10</td>
</tr>
<tr>
<td>Ala$^{281}$(N)–Gly$^{134}$(O)</td>
<td>2.72</td>
</tr>
<tr>
<td>Ala$^{281}$(O)–Gly$^{134}$(N)</td>
<td>3.47</td>
</tr>
<tr>
<td>Ala$^{281}$(N)–Gly$^{134}$(O)</td>
<td>3.07</td>
</tr>
<tr>
<td>Ala$^{281}$(O)–Gly$^{134}$(N)</td>
<td>3.14</td>
</tr>
<tr>
<td>Ala$^{281}$(O)–Wat$^{448}$</td>
<td>3.17</td>
</tr>
<tr>
<td>Pro$^{285}$(O)–Wat$^{448}$</td>
<td>3.24</td>
</tr>
<tr>
<td>Ala$^{284}$(N)–Ser$^{211}$(O)</td>
<td>2.86</td>
</tr>
<tr>
<td>Ala$^{284}$(N)–Ser$^{211}$(O)</td>
<td>2.93</td>
</tr>
<tr>
<td>Ala$^{284}$(O)–Ser$^{211}$(N)</td>
<td>2.71</td>
</tr>
<tr>
<td>Ala$^{284}$–Wat$^{465}$</td>
<td>2.84</td>
</tr>
<tr>
<td>Asx$^{161}$(N3)–Wat$^{455}$</td>
<td>3.24</td>
</tr>
</tbody>
</table>

Additional hydrogen bonds stabilizing the active site geometry

Gly$^{134}$(O)–Tyr$^{104}$(O) | 2.79 |
| Asp$^{283}$(O2)–His$^{224}$(N3) | 2.69 |
| Asp$^{283}$(O2)–His$^{224}$(N3) | 2.91 |

$^*$ MSU, methoxysuccinyl.

Wat, water.
no systematic and significant conformational changes were found. Close to the active site and just "below" the plane of functionally important His\textsuperscript{69} imidazole, there is the free SH group of Cys\textsuperscript{79} (Fig. 2a), the role of which is not yet understood. Its sulfur atom is engaged in several close contacts with active site residues (His\textsuperscript{69}(O), Ser\textsuperscript{24}(O), Ser\textsuperscript{24}(O)) which do not change significantly their positions after inhibitor binding except for the distance to His\textsuperscript{69}(O), which increases by 0.7 Å.

This increase in distance is caused by a movement of His\textsuperscript{69} which is necessary to permit formation of the covalent bond between the \textsuperscript{His}\textsuperscript{68}(N) and the inhibitor \textsuperscript{CH}_2 group by a nucleophilic displacement of the chlorine atom. The movement of the His\textsuperscript{68} imidazole is accompanied by a similar movement of the Asp\textsuperscript{39} carboxylate so that the hydrogen bonds between these two side chains are not disrupted; the movement of the imidazole is also transmitted to the C atom so that the positions of the His\textsuperscript{68} main chain atoms and of the associated amino acids Ala\textsuperscript{70} and Gly\textsuperscript{71} are also affected.

In contrast, binding of the inhibitor does not perturb significantly the conformation around the active Ser\textsuperscript{24}, although the Ser\textsuperscript{24} is covalently bonded with the inhibitor ketone group to form a structure reminiscent of the tetrahedral intermediate. Since this reaction also occurs with a "real" substrate and is indispensable for catalysis, it obviously can proceed smoothly, with no major conformational rearrangement.

When the inhibitor binds to proteinase K, there is insertion into the recognition site and hydrogen bonds are formed. There may be two reasons why the position of strand 100-104 of the substrate recognition site is not significantly affected. First, it is tightly anchored with the bulk of the proteinase K structure (Fig. 3); its N terminus is linked with a \textbeta\-strand through strand \textbeta\textsubscript{II}A (11), its C terminus with \alpha\-helix a4, and its segment 134-137 forms an antiparallel \beta\-sheet (\textbeta\textsubscript{III} (11)) with amino acids 168-170. Second, the Pro\textsuperscript{83} side chain pushes the segment 100-104 while segment 132-136 is in the correct position to smoothly hydrogen bond to the inhibitor. Segment 100-104, although connected at its N terminus through \beta\-turn t9 with strand \textbeta\textsubscript{II}B of the pleated sheet and, at its C terminus, with \alpha\-helix a3 (11), is not engaged in the secondary structure hydrogen bonding but exposed to solvent, and can move according to sterical requirements. When the methoxysuccinyl-Ala-Ala-Pro-Ala-CH\textsubscript{2}-Cl inhibitor binds into the recognition site, strand 100-104 is consequently pushed away by ~1 Å to open the binding cleft, probably caused by the tight interaction with the side chain of the inhibitor proline, Pro\textsuperscript{83}(C\textsubscript{\alpha}) · · · (O)Gly\textsuperscript{100}, 2.98 Å.

The inhibition of proteinase K by the four different peptide chloromethyl ketone inhibitors I to IV is described in Table I. The data indicate that the inhibition (binding of the inhibitors) is better with longer peptide chains of the inhibitors. The sample of inhibitors given in Table I is not sufficient to propose a correlation between inhibitor sequence and binding affinity, but it is clear from this x-ray study that proline interferes with peptide binding to proteinase K. It suggests that more effective inhibitors would have the proline substituted by another amino acid to avoid the steric interference with strand 100-104 of the recognition site.

CONCLUSIONS

Complementary to our earlier publications on the crystal structures of complexes between proteinase K and dipeptide inhibitors carbobenzoxy-Ala-Ala-C(O)CH\textsubscript{2}Cl and carbobenzyloxy-Ala-Phe-C(O)CH\textsubscript{2}Cl, we have now investigated a tetrapeptide which, in contrast to the earlier, shorter inhibitors saturates the whole substrate binding site. There is antiparallel \beta\-pleated sheet formation with all the amino acids of the inhibitor engaged in hydrogen bonding except for Pro\textsuperscript{83}. If this proline were replaced by another amino acid, one more hydrogen bond Pro\textsuperscript{83}(N) · · · Gly\textsuperscript{100}(O) should be possible after minor conformational changes around position 283 of the inhibitor and/or of the protein segment Gly\textsuperscript{100}-Tyr\textsuperscript{104}.

The contact surface between the inhibitor and the substrate binding site is extensive. Besides the hydrogen bonds, it involves van der Waals interactions between the side chains of the Ala\textsuperscript{70}, Pro\textsuperscript{83}, and the substrate recognition site, and the stacking between the succinyl moiety and the Tyr\textsuperscript{104} phenolic group. As the "bottom" of the proteinase K substrate recognition site is predominantly hydrophobic due to the side chains of Leu\textsuperscript{133} and Tyr\textsuperscript{104}, it is not the sequence of the substrate that is of importance in the recognition but only the actual size of the side chains. Even larger side chains can be tolerated due to the flexibility of the segment Gly\textsuperscript{100}-Tyr\textsuperscript{104}, as shown with Pro\textsuperscript{83} of the peptide inhibitor used in this study. In this light, the observation that the affinity of proteinase K correlates with the length of the inhibitor is readily...
explained as the binding to the substrate recognition site will be the stronger the more peptide-peptide hydrogen bonds are formed between inhibitor and enzyme, and the more hydrophobic contacts are made. Any sequence dependence should be of minor influence, as actually shown by the unspecified nature of proteinase K which has only some preference for aromatic, bulky amino acid side chains.

Acknowledgments—W. M. W. and T. P. S. thank the Alexander von Humboldt Stiftung for a fellowship. The help of Dirk Kostrewa with initial stages of the refinement is gratefully acknowledged.

REFERENCES