Active-site geometry of proteinase K

Crystallographic study of its complex with a dipeptide chloromethyl ketone inhibitor

Christian Betzel, Gour P. Pal, Mark Struck, Klaus-Dieter Jany+ and Wolfram Saenger*

Institut für Kristallographie, Freie Universität Berlin, Takustr. 6, D-1000 Berlin 33, Germany and +*Institut für Organische Chemie, Biochemie und Isotopenforschung der Universität Stuttgart, Pfaffenwaldring 55, D-7000 Stuttgart, FRG*

Received 17 December 1985

Proteinase K (EC 3.4.21.14) from the fungus *Tritirachium album* Limber is the most active known serine endopeptidase. The sequence of its 275-residue long polypeptide chain and its three-dimensional folding show a high degree of homology with the bacterial subtilisin proteases. Using difference Fourier methods, the binding mode of the synthetic carbobenzoxy-Ala-Ala-chloromethyl ketone inhibitor to the active site of proteinase K was determined. In several cycles of restrained least-squares, the enzyme-inhibitor complex was refined to a current $R = 22\%$ for 9400 X-ray diffraction data between 2.2 and 5.0 Å resolution. The inhibitor is attached to proteinase K by two covalent bonds: one between the methylene carbon of the inhibitor and Nε2 of the catalytic His 68, the other between the ketone carbon atom of the inhibitor and Oγ of the catalytic Ser 221. In addition, two hydrogen bonds donated by the peptide NH of Ser 221 and by the side chain NH_2 of Asn 160 hold the hemiketal O^- in the oxyanion hole. The peptide inhibitor is further hydrogen bonded to the proteinase polypeptide chain in a three-stranded antiparallel pleated sheet.

Polypeptide inhibitor *Proteinase K* *Serine proteinase* *Active-site geometry* *Difference Fourier method* *Crystallography*

1. INTRODUCTION

Because of their importance in biological systems, proteases of the trypsin and subtilisin families have been extensively studied [1–7]. Although the overall 3-dimensional structures of the enzymes of these two families are very different, the functional groups in the catalytic site, His, Ser, Asp, are identical and are nearly in the same geometrical arrangement. These findings suggest that the trypsin and subtilisin families have evolved indepedently and converged upon the same, or at least a very closely related catalytic mechanism [8] which has been the subject of controversial views in the past decade [1–4]. In order to obtain more insight into the evolutionary and functional aspects of the subtilisin-type enzymes of the prokaryotic (subtilisin) and eukaryotic (proteinase K) varieties, the 3-dimensional structure of proteinase K has been determined by X-ray diffraction analysis [9] and is being refined at high resolution using synchrotron radiation [10] (refinement at 1.8 Å resolution is in progress).

Here we report the crystallographic study of a complex of proteinase K with the covalently bound synthetic inhibitor, Z-Ala-Ala-chloromethyl ketone, where Z denotes carbobenzoxy. Because the covalent and hydrogen-bonding mode of this inhibitor to the active site of proteinase K closely resembles that of substrates in their transition state, information is obtained on the stereo-

Dedicated to Professor Georg Manecke on the occasion of his 70th birthday

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

chemistry and specificity of polypeptide substrate binding.

2. MATERIALS AND METHODS

The dipeptide inhibitor Z-Ala-Ala-chloromethyl ketone was synthesized in Stuttgart as described in [11]. Proteinase K was purchased from Merck, Darmstadt, and further purified by gel filtration according to [12]. Crystals of suitable size (~1 mm diameter) were grown by microdialysis of a 10% protein solution in water against 50 mM Tris, 10 mM $CaCl_2$, 1 M $NaNO_3$ and 0.02% NaN_3 adjusted to pH 6.5 [9]. Space group and cell constants are given in table 1.

Proteinase K is active in the crystalline state. Using this advantage, the enzyme-inhibitor complex was prepared by soaking the crystals in inhibitor solution for 8 days. Since the inhibitor is insoluble in water, a 33 mM solution was prepared in 100% methanol and diluted to 8 mM by addition of crystallization buffer and methanol, so that the final concentration of methanol was 50%. After this treatment, the inhibitor-soaked crystals of proteinase K were no longer active towards the substrate succinyl-Ala-Ala-*p*-nitroanilide, in contrast to the crystals of the native enzyme.

On precession photographs, the binding of the inhibitor was indicated by slight changes in intensities when compared with photographs of crystals of the native enzyme. The unit cell constants were the same (within 0.5%) for both crystals. Using an Arndt-Wonacott camera (Nonius, Delft) mounted on an X-ray generator with rotating anode (Marconi-Elliott), X-ray intensity data were recorded on films to a resolution of 2.2 Å. The photographs were digitized with an Optronics drum scanner and processed by means of the Mosco film evaluation system [13] (see table 1).

Table 1

Some details of X-ray data collection for the proteinase K-inhibitor complex

Space group	$P4_32_12$
Cell constants	$a = b = 68.1(1)$ Å, $c = 108.7(1)$ Å
X-ray source	Elliott GX20 rotating anode operated at 40 kV, 70 mA (2.8 kW), focal size 0.2×2 mm^2
Radiation	CuKα/graphite monochromator/1.54182 Å
Collimation	0.6 mm ∅
Detector	film, Kodak DEF 2
Film cassettes	flat plate, maximum radius 60 mm
Crystal to film distance	70 mm
Crystal mounting	along *c*-axis as rotation axis
Maximum resolution	2.2 Å
Total rotation for a data set	50° (45° is required for tetragonal space group; this gives 4 equivalent observations for each unique term)
for each exposure	2°
Number of films per data set	75 (25 three-film packs)
Time for each exposure	4000 s/degree; 2.2 h (10 oscillations)
Total number of observed reflections	41796
Number of unique reflections	11705
Number of unique reflections above the 3σ level in the range 2.2–5.0 Å used in the refinement	9400
Overall merging R(I) factor	7.9%

Difference electron density maps were calculated at 2.5 Å resolution using the coefficients ($Fi - Fn$) and ($2Fi - Fn$), where Fi and Fn represent the structure factors of proteinase K-inhibitor complex and native proteinase K, respectively. Phases were derived from structure factor calculations based on the crystal structure of proteinase K (fig.1) presently refined by restrained-parameter least-squares methods (PROLSQ [14]) to an R factor of 24% at 2.2 Å resolution [10]. ($R = \Sigma \, ||F_o| - |F_c|| / \Sigma \, |F_o|$ where F_o and F_c are the observed and calculated structure factors, respectively). The interpretation of the electron density map and the fitting of the inhibitor (fig.2) were carried out on an interactive color graphics display (Evans & Sutherland) using the program FRODO [15]. The proteinase K-inhibitor model was refined using the same methods [14] as for the native enzyme. After 23 cycles of refinement the R factor was reduced from 0.301 (5.0–2.5 Å resolution) to 0.224 (5.0–2.2 Å resolution). The root-mean-square (rms) deviations from expected stereochemical parameters are 0.025 Å

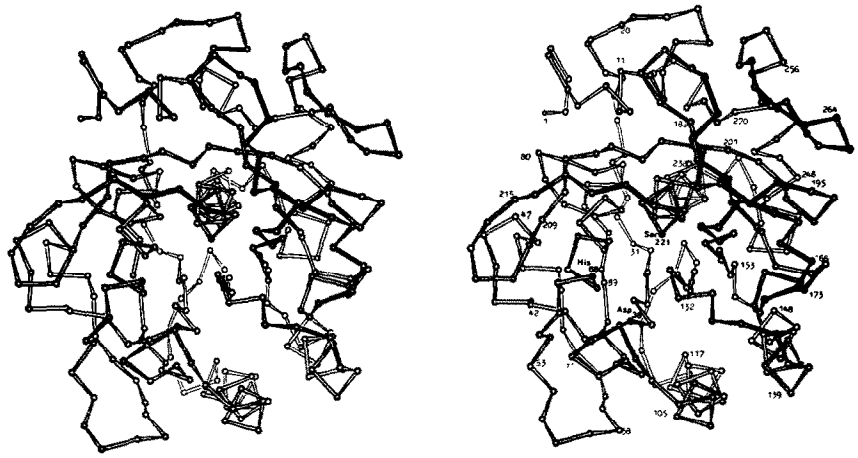

Fig.1. Stereo diagram showing the α-carbon chain of proteinase K viewed along the central helix. Residues of the active site are labelled Asp 38, His 68, Ser 221, and some other residues are indicated by numbers.

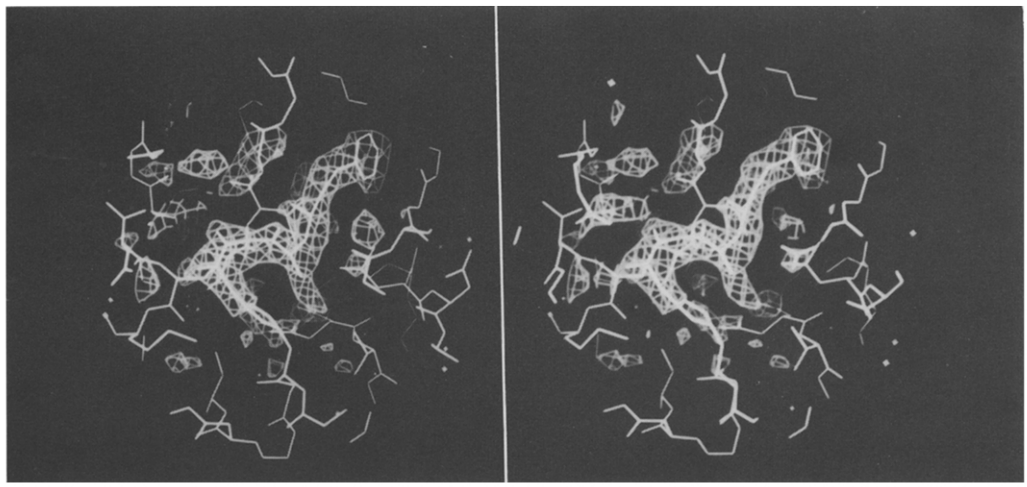

Fig.2. Stereo view showing electron density differences at 2.5 Å resolution between the proteinase K-inhibitor complex and native proteinase K. Superimposed is the skeleton structure of the inhibitor, carbobenzoxy-Ala-Ala-chloromethyl ketone, with chlorine replaced by Nϵ2 of His 68.

for the 2013 covalent bond distances; 0.058 Å for the 2732 inter-bond angle distances; 0.018 Å for the 351 planar groups of the complex; and 5.1° for the 258 peptide bond torsion angles. No water molecules have been included in the refinement process.

3. RESULTS AND DISCUSSION

3.1. *Location of inhibitor in the active site of proteinase K*

The (Fi – Fn) difference electron density map displays a well-defined continuous chain of positive electron density running from the catalytic site near His 68 nearly parallel to the proteinase K polypeptide backbone chain from Ser 131 to Gly 133 (fig.3). The (2Fi – Fn) map shows density corresponding to His 68 and Ser 221 and confirms the interpretation of a covalent bond between the methylene carbon atom of the inhibitor and Nϵ2 of His 68, requiring a slight rotation of the histidine ring about the Cα–Cβ and Cβ–Cγ bonds (fig.2). The bond between Oγ of catalytic Ser 221 and the ketone carbonyl carbon atom of the inhibitor is

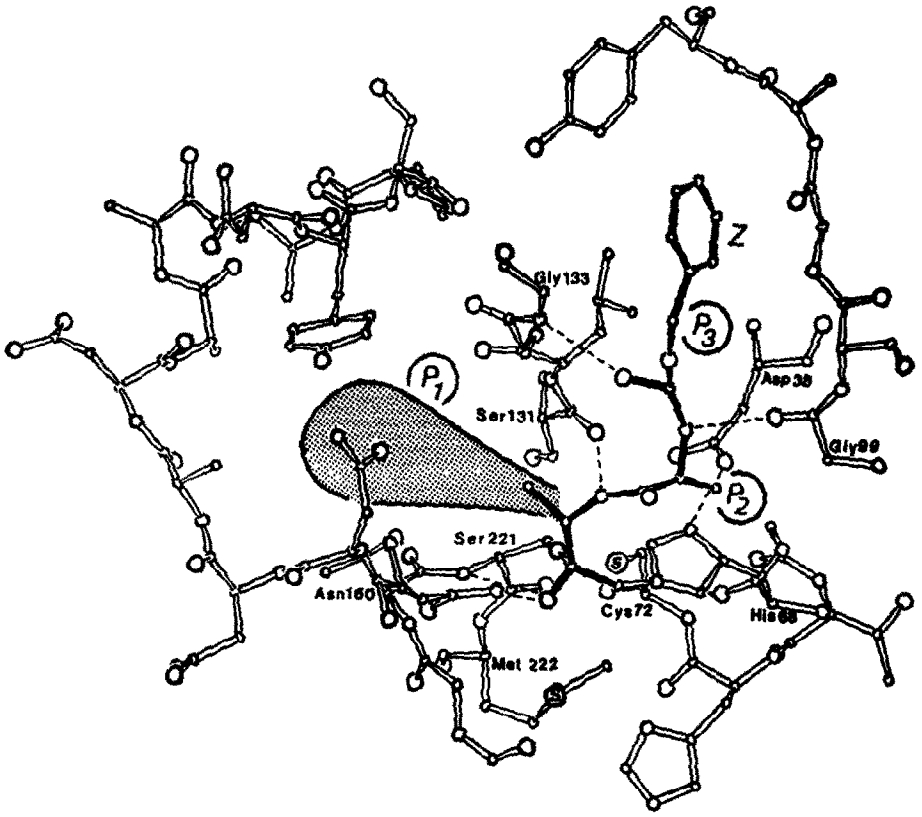

Fig.3. The active-site region of proteinase K with the inhibitor carbobenzoxy-Ala-Ala-chloromethyl ketone covalently bound to Nϵ2 of His 68 and to Oγ of Ser 221. The inhibitor is shown by solid lines, hydrogen bonds being indicated by dashed lines. P_1, P_2, P_3 denote substrate (inhibitor) sites given in fig.4. The P_1 site is shown shadowed to illustrate that the P_1 amino acid side chain can be of any size to be accommodated in the active site; hence the non-specific nature of proteinase K.

clearly indicated by positive density. This bond requires the carbonyl carbon atom to be in the hemiketal form with tetrahedral configuration, which is also in agreement with the electron density. The hemiketal form is stabilized by two hydrogen bonds between the oxyanion and the side chain NH_2 of Asn 160 and the peptide NH of Ser 221, which are placed in near-ideal tetrahedral positions in the oxyanion hole. In addition to these interactions involving the reactive chloromethyl ketone group, there are several hydrogen bonds between the peptide groups of the inhibitor and of proteinase K. These are part of an antiparallel, 3-stranded β-pleated sheet with the central strand formed by the Z-Ala-Ala moiety and the 2 other strands by proteinase K segments Ser 131 to Gly 133, and Gly 99 (see figs 3,4).

The overall geometry of the structure of proteinase K is unchanged upon binding of the Z-Ala-Ala-chloromethyl ketone inhibitor. There are, however, some minor residual peaks and holes in the vicinity of the inhibitor which suggest that the enzyme active site has suffered some slight local adjustments (fig.2).

If the binding geometries of oligopeptide chloromethyl ketone inhibitors in the active sites of proteinase K and of subtilisin [5,16,17] are compared, it is found that there are clear and well-defined similarities. Subtilisin has a chymotrypsin-like specificity because it preferentially attacks peptide bonds following hydrophobic or aromatic residues. The recognition of these side chains occurs in the S1 specificity crevice which is lined by hydrophobic residues and therefore will repel charged or

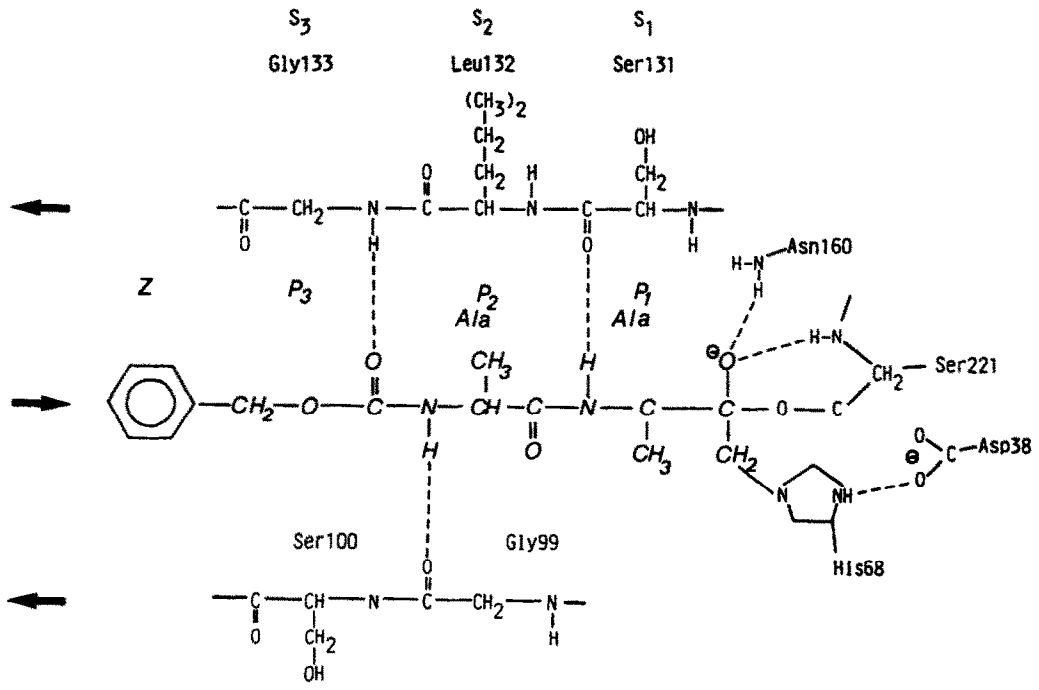

Fig.4. A schematic representation of the binding of the carbobenzoxy-Ala-Ala-chloromethyl ketone inhibitor to the active site of proteinase K. Atoms belonging to the inhibitor are given in italics, hydrogen bonds being indicated by dashed lines. S_1, S_2, S_3 are the binding positions of the substrate on the enzyme and P_1, P_2, P_3 the corresponding sites on the substrate (and product). Z symbolizes the carbobenzoxy protecting group. The arrows on the left margin give the directions of the strands of the triple antiparallel pleated sheet formed by proteinase K and the inhibitor.

hydrophilic amino acid side chains in the P1 position of the substrate. In proteinase K the S1 crevice is also of hydrophobic character, and it is very wide so that there are virtually no steric limitations to the size of the amino acid side chains in P1 position. This suggests that proteinase K is unspecific, as verified in the cleavage pattern of proteinase K towards the B chain of insulin [18].

3.2. *Structures of proteinase K and of subtilisin are closely related: the free cysteine*

Amino acid sequence data indicate that there is about 35% sequence homology between proteinase K and the subtilisins [19]. Therefore, the fungal, eukaryotic proteinase K is evolutionarily related to the bacterial, prokaryotic subtilisins of type BPN, Novo, Carlsberg and DY. As expected, the closest similarities in sequence are observed in the region around the active-site residues Asp 38, His 68 and Ser 221, and in the substrate-binding subsites 131–133. A least-squares superposition of structurally equivalent α-carbon atoms of subtilisin BPN and proteinase K gives an rms deviation of 1.1 Å. If only 27 atoms of the amino acids involved in the catalytic mechanism are considered, the rms deviation reduces to 0.6 Å.

This high homology in 3-dimensional structure suggests that proteinase K and the subtilisins belong to the same family of enzymes. There is, however, a distinct difference in the amount and position of cysteine residues. Proteinase K contains 5 cysteines of which 4 are involved in disulfide bridges and the remaining Cys 72 is free. The two other known proteinases of the subtilisin family, thermitase [20] an thermomycolin [21], each contain one cysteine in a homologous position whereas the subtilisins lack this amino acid. It is striking that the free Cys 72 in proteinase K is located in the immediate vicinity of the active site and could influence the substrate cleavage. This view is supported by the finding that proteinase K is inactivated by the addition of the S-H reagent $HgCl_2$

(K.-D. Jany, unpublished). In the crystal structure analysis of proteinase K, $HgCl_2$ was used for the preparation of a heavy atom derivative by the soaking method [9]. The obtained crystals showed the presence of a minor occupied Hg site near His 68 and too far away from Cys 72 to be covalently bound. It appears from these results and from inspection of the 3-dimensional structure model of proteinase K that the Cys 72 S-H group is hidden behind the imidazole of His 68 and therefore not accessible for reaction with $HgCl_2$. The inactivation of proteinase K by $HgCl_2$ in solution could indicate that the S-H group is more exposed due to conformational flexibility compared with the crystalline state and therefore reactive towards $HgCl_2$, or that $HgCl_2$ is, in fact, bound to the active-site His 68. Further studies are in progress to clarify this point.

As already suggested on the basis of the amino acid sequence analysis [20], it appears that the subtilisin-type proteases thermitase and thermomycolin, with one cysteine in the same position as Cys 72 in proteinase K, represent a subgroup of this family of enzymes. The subgroup character is indicated by the higher conformational stability of these enzymes and by the as yet unproven but feasible involvement of the free Cys 72 S-H group in the catalytic mechanism.

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 9, Teilprojekt A 7), Fonds der Chemischen Industrie, and the Bundesminister für Forschung und Technologie (FKZ: B72C050).

REFERENCES

[1] Blow, D.M. (1976) Acc. Chem. Res. 9, 145–152.
[2] Kraut, J. (1977) Annu. Rev. Biochem. 46, 331–358.
[3] Kossiakoff, A.A. and Spencer, S.A. (1980) Nature 288, 414–416.
[4] Moult, J. Sussman, F. and James, M.N.G. (1985) J. Mol. Biol. 182, 555–566.
[5] Wright, C.S., Alden, R.A. and Kraut, J. (1969) Nature 221, 235–242.
[6] Drenth, J., Hol, W.G.J., Jansonius, J.N. and Kökök, R. (1972) Eur. J. Biochem. 26, 177–181.
[7] Robertus, J.D., Kraut, J., Alden, R.A. and Birktoft, J.J. (1972) Biochemistry 11, 4293–4303.
[8] Kraut, J., Robertus, J.D., Birktoft, J.J., Alden, R.A. Wilcox, P.E. and Powers, J.C. (1977) Cold Spring Harbor Symp. 36, 117–123.
[9] Pähler, A., Banerjee, A., Dattagupta, J.K., Fujiwara, T., Lindner, K., Pal, G.P., Suck, D., Weber, G. and Saenger, W. (1984) EMBO J. 3, 1311–1314.
[10] Betzel, C. (1986) PhD Thesis, Freie Universität Berlin.
[11] Leary, R. (1977) Biochemistry 16, 5857.
[12] Pal, G.P., Betzel, C., Jany, K.D. and Saenger, W. (1986) FEBS Lett. 197, 111–114.
[13] Machin, P.A., Wonacott, A. and Moss, D. (1983) Daresbury Laboratory News 10, pp. 3–9.
[14] Hendrickson, W.A. and Konnert, J.H. (1981) in: International Symposium on Biomolecular Structure (Srinivasan, R. ed.) pp. 43–57, Pergamon, Oxford.
[15] Jones, T.A. (1983) in: Computational Crystallography (Sayre, D. ed.) pp. 303–317, Oxford University Press, PS 300 version by Pflugrath, J.W. and Saper, M.A., Rice University.
[16] Robertus, J.D., Alden, R.A. Birktoft, J.J., Kraut, J., Powers, J.C. and Wilcox, P.E. (1972) Biochemistry 11, 2439–2449.
[17] Poulos, T.L., Alden, R.A., Freer, S.T., Birktoft, J.J. and Kraut, J. (1976) J. Biol. Chem. 251, 1097–1103.
[18] Kraus, E., Kiltz, H.H. and Femfert, U.F. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 233–237.
[19] Jany, K.D. and Mayer, B. (1985) Biol. Chem. Hoppe-Seyler 366, 485–492.
[20] Meloun, B., Baudys, M., Kostka, V., Hausdorf, G., Frömmel, C. and Höhne, W.E. (1985) FEBS Lett. 183, 195–200.
[21] Gaucher, G.M. and Stevenson, K.J. (1976) Methods Enzymol. 45, 414–433.