Crystal structure of an engineered subtilisin inhibitor complexed with bovine trypsin
(protein engineering/protein–protein interaction)

YASUO TAKEUCHI*, TAKAMASA NONAKAt, KAZUO T. NAKAMURA†, SHUICHI KOJIMA‡, KIN-ICHIRO MIURA‡‡, AND YUKIO MITSUI††

*Pharmaceutical Research Center, Meiji Seika Kaisha, Ltd., Morooka, Kohoku, Yokohama 222, Japan; †Faculty of Engineering, Nagaoka University of Technology, Nagaoka, Niigata 940-21, Japan; and ‡Faculty of Engineering, University of Tokyo, Hongo, Tokyo 113, Japan

Communicated by Frederic M. Richards, January 10, 1992 (received for review June 18, 1991)

ABSTRACT Proteinase specificity of a proteinaceous inhibitor of subtilisin (SSI; Streptomyces subtilisin inhibitor) can be altered so as to strongly inhibit trypsin simply by replacing P1 methionine with lysine (with or without concomitant change of the P4 residue) through site-directed mutagenesis. Now the crystal structure of one such engineered SSI (P1 methionine converted to lysine and P4 methionine converted to glycine) complexed with bovine trypsin has been solved at 2.6 Å resolution and refined to a crystallographic R factor of 0.173. Comparing this structure with the previously established structure of the native SSI complexed with subtilisin BPN', it was found that (i) P1 lysine of the mutant SSI is accommodated in the S1 pocket of trypsin as usual, and (ii) upon complex formation, considerable conformation change occurs to the reactive site loop of the mutant SSI. Thus, in this case, flexibility of the reactive site loop seems important for successfully changing the proteinase specificity through mere replacement of the P1 residue.

Hypervariability of reactive site amino acid residues during evolution of proteinaceous inhibitors of serine proteinases is one of the major mysteries of protein evolution (1). In contrast, in most other biologically active proteins, the replacement of active site residues, even by closely related ones, leads to a complete loss of activity or a dramatic decrease in activity. Among the reactive site residues of serine proteinase inhibitors, the P1 residue [after the Schechter and Berger notation (2)] most significantly affects proteinase specificity. For example, strong trypsin inhibitors most frequently have lysine or arginine at P1, while strong chymotrypsin inhibitors have tyrosine, phenylalanine, leucine, or methionine at P1 (1). Thus, the hypervariability of P1 residues inevitably causes a frequent change in proteinase specificity during gradual evolution of the same series of inhibitor proteins. It has been known that one can also invoke such a dramatic change in specificity by semisynthetic replacement of the P1 residue (3, 4) or by recombinant DNA techniques. Thus, the replacement of P1 methionine in α1-antitrypsin (a potent inhibitor of neutrophil elastase) with arginine converted the inhibitor into an efficient thrombin inhibitor (5). The replacement of P1 methionine in a proteinaceous inhibitor [SSI; Streptomyces subtilisin inhibitor (6), an almost exclusive subtilisin inhibitor] with lysine or arginine converted it to a potent inhibitor of both trypsin and subtilisin BPN' (7).

Despite rather abundant observations of the dramatic change in proteinase specificity, however, the structural mechanisms of these phenomena are by no means obvious considering that (i) very good complementarity between the inhibitor and proteinase surfaces is almost always found in the crystallographically established structures of the inhibitor–proteinase complexes (8–14) and thus such good complementarity seems essential for the tight binding (15), and yet (ii) the surface geometry of the active sites of proteinases is considerably varied even among enzymes of the trypsin family, much less between enzymes of the trypsin family and those of the subtilisin family having a completely different polypeptide chain folding. In this respect, the chance to compare the three-dimensional structure of a proteinaceous inhibitor complexed with its natural target proteinase and that of a variant of the inhibitor complexed with a distinct target enzyme has long been awaited. Here we report the crystal structure of an engineered SSI, in which P1 methionine and P4 methionine were genetically converted to lysine and glycine, respectively, complexed with bovine trypsin. Comparing this structure with the previously established structure of the native SSI complexed with subtilisin BPN' (16, 17), many intriguing points as to the structural mechanisms of the change in proteinase specificity have emerged.

SSI is one of the few well-characterized (for a detailed review, see ref. 6) microbial protein proteinase inhibitors and is a stable dimer (I2) composed of two identical subunits, each of Mr 11,500 (18). It strongly inhibits a microbial serine proteinase subtilisin BPN' (E), (Ki = 10–11 M) forming an E2–I2 complex. SSI virtually does not inhibit trypsin (Ki = 10–4 M; see ref. 19 for a review). The crystal structure of free SSI was solved (20) and refined at 2.05 Å resolution (Y.T. and Y.M., unpublished data). The crystal structure of the SSI–subtilisin complex was also solved (16, 20) and refined at 1.8 Å resolution (17). Compared with other similar complexes containing serine proteinases of the trypsin family (8–14, 21), the SSI–subtilisin complex is unique in several respects (16): (i) In addition to the usual intermolecular antiparallel β-sheet interaction involving P1–P3 residues of the inhibitor, there is another antiparallel β-sheet formed between P4–P6 residues and a previously unnoticed chain segment [the “S4-6 site” (16)] of subtilisin. (ii) Unlike various trypsin inhibitors, the P4–P9 residues of SSI are very flexible in the free state (22) and they undergo extensive rigidification and transconformation upon complex formation. The SSI gene has been cloned (23) and the expression system has been established (24). Genetic conversion of P1 Met-73 to lysine (or arginine but no other amino acid residues) proved to convert SSI into a potent inhibitor of both trypsin (Ki = 4.4 × 10–9 M) and subtilisin BPN' (Ki = 6 × 10–11 M) (25). Later, additional conversion of P4 Met-70 to glycine was found to enhance the inhibitory effect against trypsin (Ki = 2.1 × 10–9 M; against

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; SSI, Streptomyces subtilisin inhibitor.
†Present address: Institute for Biomolecular Science, Gakushuin University, Mejiro, Tokyo, Japan.
†To whom reprint requests should be addressed.
subtilisin BPN’, $K_i = 10^{-11}$ M) (26), providing the best candidate for isolation of the complex with trypsin. Hereafter the (Met-73 to lysine, Met-70 to glycine) double mutant of SSI is referred to simply as ‘mutant SSI.’

**MATERIALS AND METHODS**

**Crystalization and Data Collection.** Site-directed mutagenesis, DNA construction, and purification of the mutant SSI were performed as described (25). Bovine trypsin was purchased from Worthington and was used without further purification. The method of complex formation is similar to that described (27). Bovine trypsin and mutant SSI were dissolved separately in 0.05 M phosphate buffer (pH 7.2). The enzyme solution was added to the inhibitor solution in a 2:1:2 ratio (mol/mol) and the mixture was stirred gently at 37°C for 30 min. For isolation of the complex, the reaction mixture was applied to a Biogel P-200 column previously equilibrated with 0.05 M phosphate buffer (pH 7.2) and eluted with the same buffer. The fractions containing the complex were concentrated by ultrafiltration on a Millipore C3-TK. Hanging drop-lets of this complex solution (3 µl) in 0.05 M Tris-HCl buffer (pH 7.0) containing 40% saturated MgSO$_4$ were equilibrated with the reservoir solutions of 0.05 M Tris-HCl buffer (pH 7.0) containing 80% saturated MgSO$_4$. Crystals, including a single exceptionally large one (0.2 × 0.2 × 0.6 mm), were grown for a few months. The crystals belong to the space group I222 with the unit cell dimensions $a = 110.9$ Å, $b = 116.8$ Å, $c = 64.3$ Å. One asymmetric unit contains half the complex molecule (E-I). Hereafter, the mutant SSI–bovine trypsin complex is referred to simply as ‘SSI-trypsin complex.’ For X-ray data collection at ~15°C, a macromolecule-oriented Weissenberg camera (28) installed at the synchrotron radiation source at the National Institute of Standards and Technology was used. The wavelength was set to 1.04 Å. The single exceptionally large crystal mentioned above was used with the a axis as the rotation axis. The data sets recorded on a pack of Fuji imaging plate were processed as described (17). The $R_{sym}$ ($\Sigma_{hk}||(I) - \langle I\rangle/\Sigma_{hk}$) was 6.3% for 23,106 observations up to 2.6 Å resolution, $I > 2\sigma (I)$, which were reduced to 7579 independent reflections.

**Structure Determination and Refinement.** The structure was determined by the molecular replacement method (29) making use of (i) the known structure of bovine trypsin in the bovine pancreatic trypsin inhibitor (BPTI)–trypsin complex (21) [Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (reference 30, code 2PCT)] and (ii) that of mutant SSI in the mutant SSI–subtilisin complex (31). The fast-rotation function program of Crowther (32) and the program BRUTE (33) were used. Final refinement of the orientation of the complex molecule using a routine of the program BRUTE (33) resulted in a correlation coefficient of 0.60.

The model was refined against 8–3 Å resolution data by a simulated annealing method using the program XFLOR (34). The final crystallographic $R$ factor was 0.225 for 6166 independent reflections in the 8–2.6 Å shell, with the rms deviation from identity being 0.020 Å, 4.3°, and 28.7° for bond length, bond angle, and dihedral angle, respectively. The final model had 2961 protein atoms and 200 solvent atoms. Crystallographic details will be published elsewhere.$^#$

**RESULTS AND DISCUSSION**

**The Mutant SSI–Bovine Trypsin Interface.** The secondary structures and relevant notations are schematically shown in Fig. 1. The region of contact of the SSI–trypsin complex is compared with the corresponding region of the SSI–subtilisin complex (Fig. 2). The corresponding whole views of the two complex molecules are shown in Fig. 3. The intermolecular contacts found in the SSI–trypsin complex are summarized in Table 1. There are a total of 145 contacts of <4.0 Å between the inhibitor and the enzyme. This is comparable to the corresponding numbers—126 found in the BPTI–trypsin complex (21) and 143 found in the SSI–subtilisin complex (16) [see tables 8.1(a) and -b of ref. 6 for counting details]—showing that the extent of loss of ‘accessible surface area’ (35) upon complex formation is roughly the same among the three complexes.

Of the two separate antiparallel β-sheet interactions present in the SSI–subtilisin complex (16), the second one, involving the P4–P6 residues of the inhibitor, does not exist, reflecting the fact that the required S4–S6 site is lacking in trypsin (16). Even the first β-sheet interaction involving the P1–P3 residues is only poorly preserved, the only remaining hydrogen-bond being between P3 and S3 (see Table 1). Inspecting the location of the S1–S3 segment shown in Fig. 2, the reason for the poorer interaction is easily seen: in the SSI–trypsin complex, the S1–S3 segment of the enzyme is considerably more distant from the P1–P3 segment of the inhibitor, making direct P1–S1 interactions impossible. As for the contacts involving the secondary contact region (see legend to Fig. 1), the Q4 (Ser-981) and Q3 (Asn-991) residues are in direct contact with trypsin. In particular, the O$_\alpha$ of Asn-991 forms a hydrogen bond with N$_\epsilon_1$ of Gln-192 (see Table 1). Thus, as in the case of the SSI–subtilisin complex (17), the secondary contact region contributes to intermolecular interaction in addition to its obvious role as a supporting device for the primary contact region.

---

$^9$The atomic coordinates have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (reference 30, code 2TLD).
The P1 residue (Lys-73I) of mutant SSI is accommodated in the S1 pocket with the hydrogen-bond distances from Ne to three enzymatic atoms, carbonyl O (Ser-190), Oτ (Ser-190), and Oδ2 (Asp-189), being 2.95, 2.84, and 2.56 Å, respectively. A similar scheme was found in the BPTI-trypsin complex (21) and in the pancreatic secretory trypsin inhibitor-trypsinogen complex (10). In contrast, in the complex of the same double-mutant SSI [(Met-73 to Lys, Met-70 to Gly) SSI] with subtilisin BPN’ (31), only one weak electrostatic interaction occurred between Ne of Lys-73I and Oε1 of Glu-156 (3.6 Å). Thus, the mutant SSI–trypsin complex seems to have a much more favorable scheme of P1–S1 interactions than the mutant SSI–subtilisin complex.

The distance between the carbonyl carbon [C’ (Lys-73I)] of the scissile peptide bond in mutant SSI and Oτ of the active Ser-195 in trypsin is 2.7 Å, so that there cannot be a covalent bond between the two atoms. Thus, the mutant SSI–trypsin complex can only be the Michaelis complex as in all other highly refined crystal structures of proteinaceous inhibitors complexed with proteinases of the trypsin or subtilisin family (for details, see table 4 of ref. 17). The carbonyl oxygen of the scissile bond is in the “oxyanion hole” receiving two hydrogen bonds from the peptide NH of Ser-195 and the side chain NH of Gly-193 as in the BPTI–trypsin complex (21), pancreatic secretory trypsin inhibitor–trypsinogen complex (10), and ovomucoid inhibitor from turkey domain 3–Streptomyces griseus protease B complex (12).

Conformation Changes in Trypsin and SSI on Complex Formation. We first compared the highly refined structure of trypsin in the BPTI–trypsin complex (21) with that of trypsin in the present complex. A global least-squares superposition of the two α-carbon chains gave a rms deviation of 0.77 Å. Six α-carbons deviate in position by >1.5 Å: residues 25–26, 29, 147–148, and 174. All of them are on the exposed loops. These sites are localized and there seems to be no significant global conformation change upon complex formation.

The crystal structure of free SSI in a trigonal crystal form has recently been refined at 2.05 Å resolution to an R factor of 0.194 (Y.T. and Y.M., unpublished data). Initial global least-squares superposition of the α-carbon chain of free SSI
Table 1. Contacts of <4.0 Å between trypsin and mutant SSI

<table>
<thead>
<tr>
<th>Residue</th>
<th>Site</th>
<th>No. of contacts</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg-65I</td>
<td>P9</td>
<td>23</td>
<td>Close contacts of side chain with Ser-217, Gly-219, Ala-221 side chain</td>
</tr>
<tr>
<td>Gly-66I</td>
<td>P8</td>
<td>1</td>
<td>Close contact with Ser-217 side chain</td>
</tr>
<tr>
<td>Glu-67I</td>
<td>P7</td>
<td>13</td>
<td>Close contacts with Ser-217 Ala-221, Lys-224 side chain</td>
</tr>
<tr>
<td>Asp-68I</td>
<td>P6</td>
<td>11</td>
<td>Close contacts with Ser-217 Lys-224 side chain</td>
</tr>
<tr>
<td>Val-69I</td>
<td>P5</td>
<td>8</td>
<td>Close contacts with Ser-217</td>
</tr>
<tr>
<td>Gly-70I</td>
<td>P4</td>
<td>3</td>
<td>Close contacts with Gly-216</td>
</tr>
<tr>
<td>Cys-71I</td>
<td>P3</td>
<td>7</td>
<td>Carbonyl O and peptide NH forming two hydrogen bonds with NH and O, respectively, of Gly-216, close contacts with Trp-215 side chain, Gly-216</td>
</tr>
<tr>
<td>Pro-72I</td>
<td>P2</td>
<td>2</td>
<td>Close contacts with His-57, Gln-192 side chain</td>
</tr>
<tr>
<td>Met-73I</td>
<td>P1</td>
<td>47</td>
<td>Carbonyl O hydrogen-bonded to NHs of Gly-193 and Ser-195 (oxyanion hole); side chain accommodated in the S1 binding pocket</td>
</tr>
<tr>
<td>Val-74I</td>
<td>P1</td>
<td>9</td>
<td>Close contacts with Cys-42, His-57, Gln-192, Ser-195 side chain</td>
</tr>
<tr>
<td>Tyr-75I</td>
<td>P2</td>
<td>9</td>
<td>Close contacts with Tyr-39 side chain, His-40, Phe-41</td>
</tr>
<tr>
<td>Asp-76I</td>
<td>P3</td>
<td>2</td>
<td>Close contacts with Tyr-39, Lys-60 side chain</td>
</tr>
<tr>
<td>Pro-77I</td>
<td>P4</td>
<td>2</td>
<td>Close contacts with Tyr-39 side chain</td>
</tr>
<tr>
<td>Ser-98I</td>
<td>Q4</td>
<td>5</td>
<td>Close contacts with Gln-192 side chain</td>
</tr>
<tr>
<td>Asn-99I</td>
<td>Q3</td>
<td>3</td>
<td>Close contacts with Gln-192 side chain, hydrogen bond of 3.01 Å from O6 to Gln-192 Ne</td>
</tr>
</tbody>
</table>

For the meaning of primary and secondary contact regions (P and Q notation) see Fig. 1.

upon that of mutant SSI in the present complex gave a rms deviation of 0.81 Å. Then we realized that the structurally most conserved region of the SSI subunit resided in the core of the β-sheet, the β1, β2, and β3-strands. A second least-squares superposition based on the 20 α-carbons belonging to this region gave a rms deviation of 0.4 Å (Fig. 4). Most conspicuous deviations are seen around the scissile bond (the reactive site) and along the flexible loop (see Fig. 1). These conformation changes seem to propagate through the α2-helix down to the C-terminal segment as indicated by the solid arrows in Fig. 1. The slight change in orientation of the β2-strand and the conformation change in the turn around residue 20 are also notable. Thus, the conformation change in SSI upon complex formation is a global one and appears to originate from a local conformation change in the reactive site and in the flexible loop, both of which, in turn, are obviously due to their direct interaction with the surface of the enzyme. It appears that orientation of the P3–P2' segment (the reactive site) relative to the core of the inhibitor molecule is considerably different in the case of SSI complexed with trypsin and that of SSI complexed with subtilisin BPN'. Thus, the two SSI subunits superimposed with respect to the reactive site α-carbons (see legend to Fig. 2) are seen tilted by a few degrees from each other (compare the yellow and blue SSIs in Fig. 3). We conclude that SSI is more liable to undergo local and global conformation change than a trypsin inhibitor BPTI, where no significant conformation change upon complex formation was observed (36). It appears that the presence of the flexible loop is most responsible for such a plasticity of the SSI molecule. Enhanced inhibition (see Introduction) of trypsin through the replacement of P4 methionine by glycine (in addition to the Met-73 to lysine conversion) may well be due to the increased flexibility of this loop. Detailed analysis of the conformation change in various liganded states will be published elsewhere.

The Reason for Easy Change in Proteinase Specificity. As mentioned in the Introduction, the apparently easy change in proteinase specificity, while maintaining the inhibitory activity, by simply replacing the P1 (and a few other) amino acid residues is a phenomenon widely observed among the proteinaceous inhibitors of serine proteinases. In contrast, similar changes produced in the functionally important amino acid residues of other kinds of proteins usually result in dramatic weakening of the activity even if the ligand specificity is successfully changed. Thus, in these proteins, functionally important residues are stubbornly conserved during evolution. In the case of catalytic residues of enzymes, for example, this tendency is understandable considering the extreme sensitivity of quantum chemical processes (such as proton transfer and nucleophilic attack) to geometrical as well as chemical environments. In contrast, in the case of proteinaceous inhibitors, the only requirements are perhaps the tight binding to the target enzymes and stability of the resultant complex. Thus, there seems to exist a fundamental reason to make functional modification of proteinaceous inhibitors generally easier than that of other proteins, especially the enzymes.

Still, even the simpler requirements for tight and stable binding to the target enzymes appear not necessarily easy to fulfill. Since strong (noncovalent) interactions are most probably largely dependent on good complementarity of the surface geometries of the two component molecules (15, 35), the surface geometries of relevant target proteinases must be more or less similar for an easy switch in proteinase specificity to occur. This would be relatively easily fulfilled if the switch is within proteinases of the trypsin family, since they are evolutionarily and structurally closely related to each other. For a switch to evolutionarily unrelated proteinases of the subtilisin family, however, a certain degree of flexibility in interacting one or both sides of the interacting molecules will have to be present. In the case of SSI, the flexible loop seems to play such a role. It remains to be seen whether or not such an easy and broad specificity change can occur for an inhibitor having no remarkable flexible parts.

Moreover, in some cases, the site remote from the reactive site can hamper such a broad specificity change. It has been known through extensive computer simulation studies (27) that the hypothetical complex of BPTI or soybean trypsin inhibitor with subtilisin BPN' would give rise to serious steric collisions at the sites (Cys-38 of BPTI or His-71 of soybean trypsin inhibitor) remote from the reactive sites. Thus, conversion of a typical trypsin inhibitor (BPTI or soybean trypsin inhibitor) into a good subtilisin inhibitor appears hopeless if the effort is restricted to simple replacements of the P1 residue (Lys-15 or Arg-63, respectively) by, for example, methionine.

We thank Prof. R. Huber for advice in the crystallization of the present complex. Data collection at the photon factory synchrotron radiation source (Takuba) was made possible through generous support by Prof. N. Sakabe and Dr. A. Nakagawa. We also thank Profs. S. Ishii, M. Kainosho, A. Takenaka, and Dr. Y. Miyake for discussion; Profs. D. M. Blow, S. Murao, K. Hiromi, M. Miwa, K.
Hamaguchi, and A. Wada for encouragement. Part of the present work was performed at the Faculty of Pharmaceutical Sciences, University of Tokyo.