Structure of a hybrid squash inhibitor in complex with porcine pancreatic elastase at 1.8 Å resolution

**Jacqueline Aÿ,[a]† Kai Hilpert,[a]†
Norbert Krauss,[a] Jens Schneider-
Mergener[b] and Wolfgang
Höhne[a]***

[a]**Institut für Biochemie, Universitätsklinikum
Charité, Humboldt-Universität zu Berlin,
Abteilung Proteinstrukturforschung,
Monbijoustrasse 2, 10117 Berlin, Germany, and
[b]Institut für Medizinische Immunologie,
Universitätsklinikum Charité, Humboldt-
Universität zu Berlin, Schumannstrasse 20/21,
10098 Berlin, Germany

† These authors contributed equally to this
paper.

Correspondence e-mail:
wolfgang.hoehne@charite.de

The crystal structure of porcine pancreatic elastase in complex with a hybrid squash inhibitor (HEI-TOE I; 28 amino acids) has been determined to a resolution of 1.8 Å. To construct the hybrid inhibitor, the trypsin-binding loop of the squash inhibitor from *Ecballium elaterium* was substituted by the sequence of a peptide that was derived from the third domain of the turkey ovomucoid inhibitor and was optimized to inhibit porcine pancreatic elastase. This modification of the squash inhibitor changed its specificity for trypsin to a specificity for porcine pancreatic elastase. Specific interactions of this hybrid inhibitor with porcine pancreatic elastase and the differences from the interactions of the ovomucoid inhibitor with human leukocyte elastase are discussed. The binding loop of the inhibitor adopts a 'canonical' conformation and the scissile bond Leu-Glu remains intact.

1. Introduction

Designing potent synthetic protease inhibitors has great importance for many human diseases. Moreover, the protease–inhibitor interaction is an interesting example and a useful model system for studying protein–protein interactions in detail in order to gain a better understanding of molecular-recognition processes.

Several natural protease inhibitors are small but very stable proteins consisting of only 25–60 amino acids, meaning that they are within the scope of peptide synthesis. Known examples include the Kunitz inhibitory family, the domains of the Kazal-type inhibitors and the squash inhibitors.

The ovomucoid inhibitor from turkey egg white consists of three homologous Kazal-type domains. The third domain of this inhibitor (OMTKY3; 56 amino acids) is a potent serine-protease inhibitor which binds to a broad spectrum of serine proteases, such as chymotrypsin, porcine pancreatic and human leukocyte elastase, *Streptomyces griseus* proteinases A and B and subtilisin, with K_i values between 10^{-10} and 10^{-12} M (Bigler *et al.*, 1993).

Squash-type inhibitors from the seeds of the Curcubitacea (squash) plant family usually comprise only 27–34 amino-acid residues (Otlewski & Krowarsch, 1996). More than 40 inhibitors of this family have been characterized. Squash inhibitors have a rigid structure stabilized by three 'knotted' disulfide bonds and are resistant to heat or acid denaturation. Many of these inhibitors show strong inhibition of trypsin, with a K_i value of about 10^{-12} M, but they can also inhibit other proteases, *e.g.* plasmin, kallikrein, cathepsin G and blood-clotting factors Xa and XIIa (for reviews, see Otlewski, 1990; Otlewski & Krowarsch, 1996). Several structures are available from the Protein Data Bank (PDB) for free and

complexed squash inhibitors derived from NMR and X-ray crystallographic measurements, respectively.

In a previous investigation, an inhibitory peptide derived from the protease-binding region of OMTKY3 (P5–P4′; PACTLEYRC; the C-terminal cysteine was added for cyclization) was optimized by substitutional and length-variation analysis to specifically inhibit porcine pancreatic elastase (PPE) with a K_i of 6.7×10^{-7} M (Hilpert *et al.*, 2000). In the present work, we used this optimized ovomucoid inhibitor-derived peptide to substitute for the protease-binding region of a squash inhibitor and characterized this hybrid ovomucoid/squash-inhibitor protein by kinetic measurements and X-ray analysis of its complex with PPE. We could demonstrate that despite its small size the scaffold of this *Ecballium elaterium* squash inhibitor is stable enough to accommodate deviating sequences in its binding loop without loss of the native fold.

2. Materials and methods

2.1. Peptide and protein synthesis

The peptides and HEI-TOE I were synthesized according to standard Fmoc machine protocols using a multiple peptide synthesizer (Abimed, Langenfeld, Germany). Disulfide formation was performed with charcoal as described previously (Volkmer-Engert *et al.*, 1998). All peptides were purified by HPLC and characterized by MALDI-TOF mass spectrometry. HPLC purification of HEI-TOE I was performed twice: prior to and after the disulfide-formation step. Details of the procedures and yields have been published elsewhere (Hilpert *et al.*, 2000). The degree of disulfide formation was characterized by determining the amount of free cysteine with 5,5′-dithiobis-(2-nitrobenzoic acid).

2.2. Kinetic measurements

The activity of PPE was measured by monitoring the hydrolysis of the substrate suc-Ala-Ala-Ala-pNA (Serva, Heidelberg, Germany) on a UV-160A recording spectrophotometer (Shimadzu, Duisburg, Germany). By means of varying the substrate (0.25–1.3 m*M*) and inhibitor concentrations (0.9–0.3 m*M*), the K_i values were determined with 0.18 units (in 200 µl assay volume) PPE in 0.1 *M* Tris buffer pH 8.5 at 298 K. The same protocol was applied to HLE (Serva, Heidelberg, Germany) activity determination using solutions of 0.25–1.3 m*M* suc-Ala-Ala-Val-pNA (Serva, Heidelberg, Germany) as a substrate and 4.5 units HLE in 200 µl. The error in the K_i determination was usually between 10 and 20%.

One unit of PPE activity corresponds to the hydrolysis of 1 µmol *N*-acetyl-(L-Ala)$_3$-methyl ester per minute at 298 K and pH 8.5. One unit of HLE releases 1 µmol *p*-nitroanilide per minute from MeO-suc-Ala-Ala-Pro-Val-pNA at 298 K and pH 8.0.

2.3. Crystallization

Porcine pancreatic elastase was purchased from Serva (Heidelberg, Germany) and was used without further purifi-

Table 1
X-ray diffraction data and refinement statistics for the PPE–inhibitor complex.

Values in parentheses are for the highest resolution shell.

Crystal data	
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 56.33$, $b = 56.44$, $c = 72.76$
Contents of asymmetric unit	1 elastase–inhibitor complex
Data statistics	
Resolution limits (Å)	14.0–1.8 (1.86–1.80)
No. of observations	79096
No. of unique reflections	22112 (1941)
Completeness (%)	98.3 (87.8)
R_{merge}† (%)	3.7 (11.6)
$I/\sigma(I)$	21.0 (4.5)
Reflections with $I/\sigma(I) > 3$ (%)	90.0 (72.2)
Refinement and model statistics	
Resolution limits in refinement (Å)	14–1.8
No. of protein non-H atoms	2023
No. of solvent molecules	353
No. of reflections	21938
R_{cryst}‡ (%)	18.2
R_{free}§ (%)	22.1
Average B factor (Å^2)	
All atoms (including solvent)	15.3
Elastase only	12.6
Inhibitor only	14.6
Residues in most favoured regions of Ramachandran plot (%)	85.5
Residues in additional allowed regions (%)	13.6
Residues in generously allowed regions (%)	0.9
Ramachandran outliers (%)	0
R.m.s. deviations from parameters of Engh & Huber (1991)	
Bond lengths (Å)	0.00460
Bond angles (°)	1.26

† $R_{merge} = \sum_{hkl} \sum_j |I_{hkl}(j) - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl}(j)$, where $\langle I_{hkl} \rangle$ is the mean intensity of the reflection (hkl). ‡ $R_{cryst} = \sum_{hkl} ||F_o| - |F_c|| / \sum_{hkl} |F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor amplitudes, respectively. § R_{free} is the same as the R factor but is calculated with 5% of the reflections that were excluded from refinement.

cation. Crystals of the complex of PPE and HEI-TOE I were usually obtained within 2 d by means of the hanging-drop method at 293 K with a 0.02 *M* citrate buffer pH 6.0 as the precipitant. Details of the crystallization procedure have been published elsewhere (Hilpert *et al.*, 2002).

2.4. Data collection and structure determination

Diffraction data were collected at 100 K to 1.8 Å resolution at synchrotron beamline X11 (EMBL Outstation, DESY, Hamburg) using a flash-frozen crystal soaked in 20% glycerol as cryoprotectant. The X-ray data were processed and merged with the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The structure was determined by molecular replacement using the program *AMoRe* (Navaza, 1994) and the coordinates for PPE (PDB code 1qnj; Würtele *et al.*, 2000) as a search model. After the orienting and positioning of the elastase molecule, continuous electron density for the inhibitor peptide was visible and the inhibitor was built into this density.

2.5. Model building and refinement

The refinement was performed using the maximum-likelihood method of the program *CNS* (Brünger *et al.*, 1998), followed by manual adjustments after each refinement cycle using the program *O* (Jones *et al.*, 1991). Solvent molecules were added as indicated by σ_A-weighted $F_o - F_c$ electron-

OMTKY3, binding region   . . .YPKPACTLEYRPLCG. . .(56 amino acids)

Optimized inhibitory peptide   PMTLEYR

EETI-II   GCPRILMRCKQDSDCLAGCVCGPNGFCG

HEI-TOE I, binding region   PCTLEYMRCK. . . (28 amino acids)

Figure 1
Comparison of the sequence of the inhibitory proteins and the peptide used in this study (P1 position underlined).
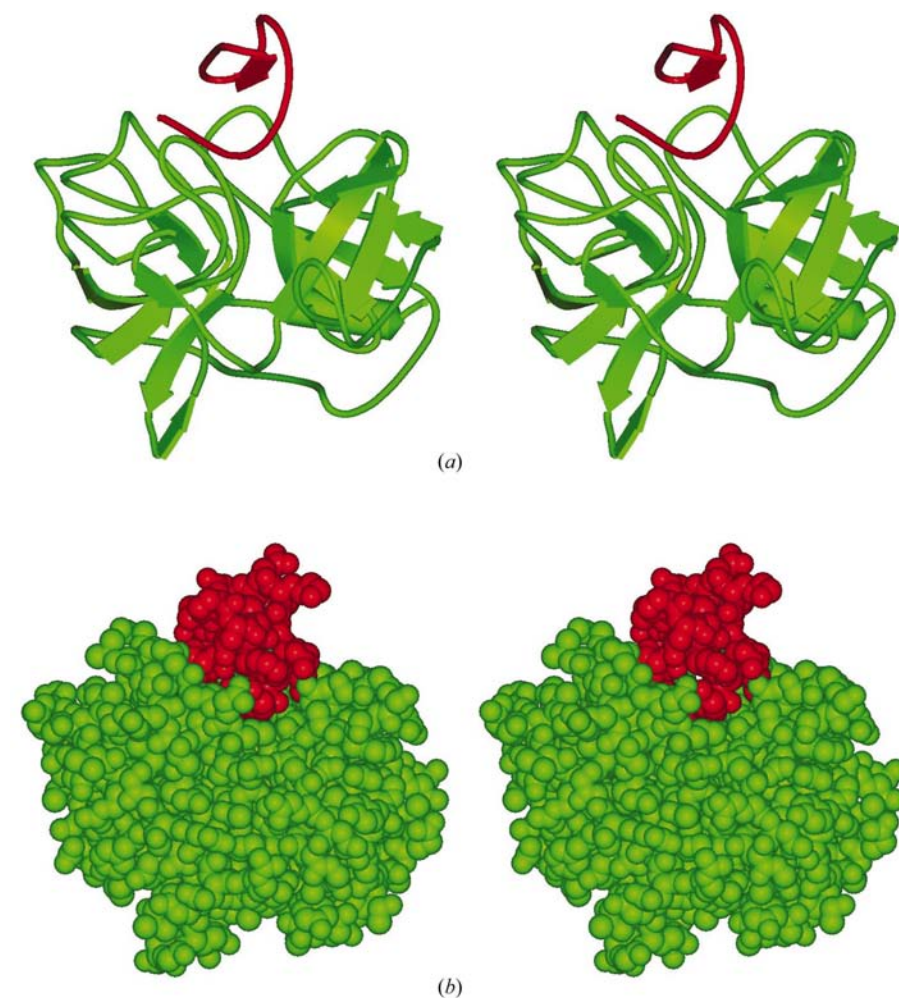

(a)

(b)

Figure 2
Stereoview of porcine pancreas elastase (green) with hybrid inhibitor HEI-TOE I (red): (*a*) schematic representation, (*b*) space-filling representation. All figures, except Fig. 3(*b*),were drawn with *ViewerPro*4.2 (Accelrys, San Diego, USA).

density maps calculated using *CNS*. The quality of the model was checked with the program *PROCHECK* (Laskowski *et al.*, 1993). Detailed statistics of the refinement and model are listed in Table 1. In the final model, the R and R_{free} (5%) converged at 18.3 and 22.3%, respectively. Nine disordered residues (alternative positions) were included in the final model. The program *NACCESS* (Hubbard & Thornton, 1996) was used to calculate the buried solvent-accessible surface area of the complex. Generally, all superpositions shown in the figures are least-squares superpositions of C$^\alpha$ atoms only.

3. Results and discussion

The inhibitory peptide (PMTLEYR), derived from OMTKY3 and optimized for PPE binding, was previously shown to undergo slow proteolytic degradation, making it unsuitable for X-ray structural analysis in a complex with PPE (Hilpert *et al.*, 2000). Soaking PPE crystals in a corresponding peptide solution resulted in an 1.1 Å high-resolution structure of PPE (Würtele *et al.*, 2000), but with no difference electron density for the peptide. Therefore, we decided to insert this peptide into the inhibitory loop of the squash-type trypsin inhibitor EETI-II in such a way that conserved positions were not altered, *i.e.* the cysteine at position 2, which is necessary to form a disulfide bridge, and the largely conserved methionine at position 7 (Otlewski & Krowarsch, 1996), which was kept to avoid two adjacent positive charges (Fig. 1). The resulting hybrid inhibitor (subsequently called HEI-TOE I: H for hybrid, E for elastase, I for inhibitor, T for turkey, O for ovomucoid and E for *E. elaterium*) was considerably more resistant to proteolytic attack than the free peptide (see below and data to be published). It was then possible to crystallize a complex of the modified squash inhibitor with PPE and to solve its structure to 1.8 Å resolution. Details of the conditions for crystal growth are given in Hilpert *et al.* (2002).

Crystals of about $0.1 \times 0.05 \times 0.6$ mm in size were flash-frozen and a data set was measured at a synchrotron beamline with 88% completeness in the highest resolution shell. The crystals belong to the space group $P2_12_12_1$. There is one elastase–inhibitor complex in the asymmetric unit, with a V_M value of 1.98 Å^3 Da^{-1}, corresponding to a solvent content of 37.8%. More details are given in Table 1.

A stereoview of the overall structure of the complex is given in Figs. 2(*a*) and 2(*b*). The inhibitor fits closely *via* its

Table 2
Contacts between hybrid squash inhibitor HEI-TOE I and PPE.

Residues with closest distances < 4.2 Å. The atoms involved (HEI-TOE I–PPE) are given in parentheses; hydrogen bonds are indicated.

Pro1	Val99	4.17 (CB–CG1)
	Trp172	3.80 (CG–CZ3)
	Thr175	3.95 (CG–CG2)
	Phe215	3.41 (CB–CD2)
	Val216	3.37 (CA–O)
	Arg217a	4.11 (N–CA)
Cys2	Gln192	3.32 (SG–NE2)
	Phe215	3.35 (O–CB)
	Val216	3.00 (O–N, hydrogen bond)
	Ser217	3.95 (CB–O)
	Arg217a	4.18 (CB–CA)
Thr3	His57	3.51 (CB–NE2)
	Val99	4.15 (CG2–CG2)
	Gln192	3.10 (O–NE2, hydrogen bond)
	Ser195	4.04 (C–OG)
	Ser214	3.46 (CA–O)
	Phe215	4.05 (CA–CB)
Leu4	His57	3.81 (N–NE2)
	Gly190	3.61 (CD1–CA)
	Cys191	3.45 (CD2–C)
	Gln192	3.48 (O–CA)
	Gly193	2.68 (O–N, hydrogen bond)
	Asp194	3.21 (O–N, hydrogen bond)
	Ser195	2.87 (O–N, hydrogen bond)
	Ser195	2.82 (N–OG, hydrogen bond)
	Ser195	2.63 (C–OG)
	Thr213	3.40 (CD1–CG2)
	Ser214	3.32 (N–O)
	Val216	3.46 (CD2–CG2)
	Cys220	4.13 (CD2–SG)
	Thr226	4.01 (CD1–CG2)
Glu5	Thr41	3.57 (CA–O)
	Cys42	3.62 (CB–SG)
	His57	3.62 (OE1–CD2)
	Arg61	2.94 (OE2–NH1, salt bridge)
	Gly193	3.76 (C–N)
	Ser195	3.02 (N–OG)
Tyr6	Tyr35	4.17 (O–CE1)
	His40	3.75 (CD2–O)
	Thr41	2.87 (N–O, hydrogen bond)
	Thr41	2.79 (O–OG1, hydrogen bond)
	Leu143	3.85 (CE1–CD2)
	Leu151	3.54 (CD2–CG)
	Gln192	4.12 (CB–C)
	Gly193	3.66 (N–N)
Met7	Tyr35	3.56 (CE–CE1)
	Thr41	3.48 (CE–OG1)
	Arg61	3.16 (CG–NE)
	Leu63	3.83 (CE–CD1)
Arg8		
First conformation	Tyr35	3.92 (N–OH)
Second conformation		3.04 (NH1–CE2)
Asp14	Arg61	3.05 (O–NH2, hydrogen bond)
Cys15	Arg61	4.03 (CA–NH2)
Leu16	His57	3.66 (CD1–CB)
	Asp60	3.63 (CD1–OD2)
	Arg61	3.93 (CG–NH2)
	Thr96	3.91 (CD2–OG1)
Cys27	Arg61	4.10 (SG–NH2)
	Gln192	3.59 (O–CG)
Gly28	Gln192	3.84 (CA–CD)

(a)

(b)

Figure 3
The binding loop (PCTLEYMR) of hybrid inhibitor HEI-TOE I in the binding pocket of porcine pancreatic elastase. (a) Representation of PPE with solvent-accessible surface, (b) the same part of the inhibitor within the $2F_o - F_c$ electron-density map contoured at 1.0 r.m.s. deviations above the mean density and drawn with *SETOR* (Evans, 1993). The inhibitor-binding loop is in red; the alternative Arg side-chain position is in brown.

binding loop into the substrate-binding cleft of the protease. There is clear electron density for all the hybrid inhibitor amino-acid side chains, including the residues interacting with the elastase substrate-binding site as shown in Fig. 3. Several alternative ('disordered') side-chain positions were clearly
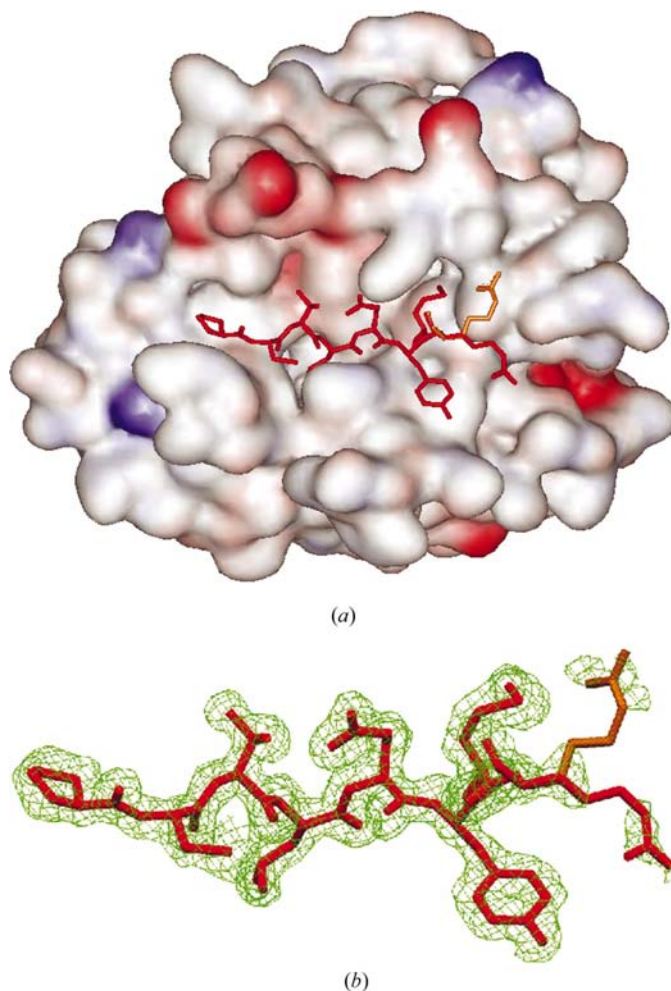
identified within the electron-density map (Gln49, Gln86, Ser116, Gln119, Arg145, Ser170b, Asn178, Ser236 of the protease and Arg8 of the inhibitor). Well defined interactions exist between the seven N-terminal residues Pro1–Met6 of the inhibitor and the elastase subsites S4–S'3 (Table 2). Pro1, Leu4, Tyr6 and Met7[1] undergo hydrophobic interactions with surrounding residues from the elastase substrate subsites. A salt bridge is established between the Glu5 residue of the inhibitor and the Arg61 residue of elastase. Leu4 (P1) of the inhibitor stacks deeply in the S1 subsite and is almost completely buried from solvent access. The carbonyl oxygen of P1 occupies the oxyanion hole, with hydrogen bonds to Gly193 N, Asp194 N and Ser195 N. Five more residues of the inhibitor (Asp14, Cys15, Leu16, Cys27 and Gly28) are in

---

[1] Met7 in HEI-TOE I corresponds to Arg in OMTKY3 and also in the optimized peptide inhibitor PMTLEYR.

contact, but weakly, with the protease (Table 2). The Arg8 side chain of the inhibitor shows two alternate positions (Fig. 3).

Crystal packing contacts are provided by elastase and inhibitor molecules, but do not involve any inhibitor–inhibitor interactions. There are 353 water molecules, one calcium ion and a sulfate ion of 50% occupancy refined in the final structure. The calcium ion occupies the usual metal-binding site of elastase, with O atoms from the side chains of Glu70, Asp77, Glu80 and from the main chains of Asn72 and Gln75 as the ligands. In addition, there is a water molecule within coordinating distance as observed with the free PPE structure (PDB code 1qnj; Wurtele *et al.*, 2000) or other porcine pancreatic elastase–inhibitor complexes (*e.g.* PDB codes 1haz and 1btu; Wilmouth *et al.*, 1998, 2001). Of the 353 water molecules in the present model, 60%[2] are conserved compared with those in the uncomplexed elastase structure at 1.1 Å resolution (Wurtele *et al.*, 2000).

Of a total of 49 water molecules (14%) buried or partially buried in the interior of the elastase structure, 43 are conserved (88%). 21 water molecules and a sulfate ion are displaced from the elastase substrate site by inhibitor binding. 46 water molecules are associated with the inhibitor and 21 of these are in contact with both the protease and the inhibitor molecules. Of the 21 water molecules bridging the elastase and inhibitor, 12 are conserved. Another three water molecules described in the complex of CMTI-I with bovine trypsin as stabilizing the binding loop within the inhibitor structure by a hydrogen-bond network (Bode *et al.*, 1989) are also present in the HEI-TOE I–PPE complex (waters 29, 39 and 47).

The solvent-accessible surfaces buried by inhibitor binding are 779.5 Å^2 for the elastase and 941.5 Å^2 for the inhibitor molecule, giving a total of 1721.0 Å^2 for the complex. This is comparable with the area buried in the complex of ovomucoid fragment OMTKY3 with human leukocyte elastase (HLE; PDB code 1ppf; 1324.4 Å^2) or that for the squash inhibitor CMTI-I bound to bovine trypsin (PDB code 1ppe; 1687.8 Å^2).

Comparison of the free and complexed elastase structure (Fig. 4) generally shows a very similar conformation, with only minor differences in the position of two loop regions adjacent to the bound inhibitor. In the loop region Arg145–Gln150 there is a maximal shift with respect to the C$^\alpha$ position of Asn148 of 1.56 Å, and in the other loop region Ser217–Gly219 the maximal shift in C$^\alpha$ with respect to Arg217a is 0.76 Å. The largest side-chain differences are also observed in this region, with Asn148 shifted outwards in the case of inhibitor binding by 2.3 Å (C$^\gamma$) and Arg217a shifted outwards by 3.7 Å (C$^\zeta$). Another significant outward shift near the inhibitor-binding site is that of Glu62, with a shift of 1.20 Å for the C$^\alpha$ position and 2.7 Å for the side-chain C$^\delta$ position. In addition, there is another remarkable shift of a side chain for Arg61 (3.6 Å for C$^\zeta$), which forms a salt bridge

[2] Water considered as conserved if $d < 1.5$ Å.

with Glu5 of the inhibitor in the complex. Slight changes in side-chain position also occur in the active site for His57 and Ser195 of the catalytic triad.

There is no X-ray structure available for the original EETI-II squash inhibitor complexed with a protease, but there is a model structure for free EETI-II (Chiche *et al.*, 1989; PDB code 2eti) and a set of NMR structures for an EETI-II I5L mutant (Nielsen *et al.*, 1994; PDB code 2let). Neither can be easily compared with the structure of the hybrid inhibitor in our complex. The model structure shows large differences compared with our inhibitor structure and also compared with the mutant NMR structures. The set of NMR structures exhibits a very different course of the peptide chain in the binding region, indicating some flexibility in this region of the uncomplexed inhibitor. There is a crystal structure in the PDB of a complex between CMTI-I and bovine trypsin (Bode *et al.*,
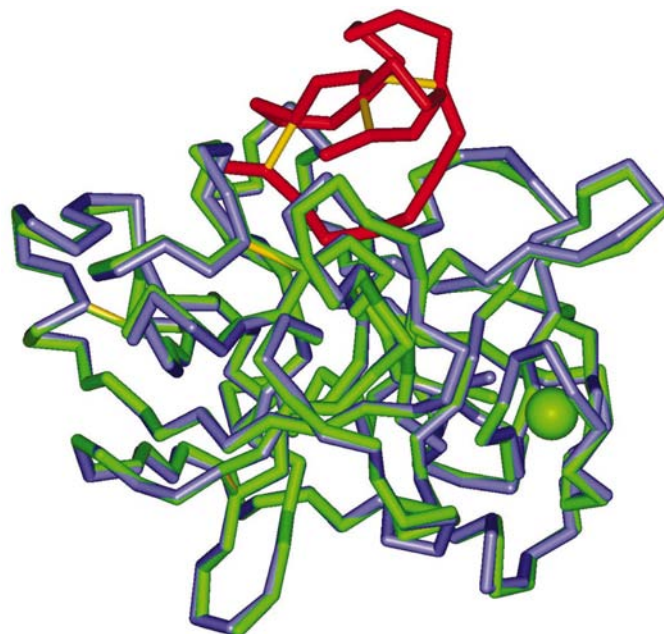

Figure 4
Comparison of the backbone conformation (C$^\alpha$ positions) of free porcine pancreatic elastase (blue; PDB code 1qnj) and in complex with the hybrid inhibitor HEI-TOE I (green). The sphere is the bound Ca^{2+} ion.
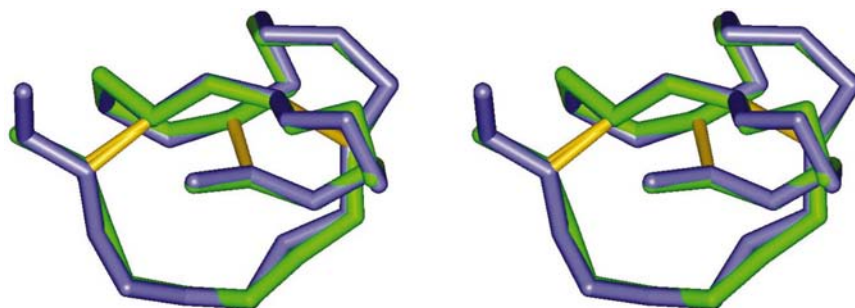

Figure 5
Comparison of hybrid inhibitor HEI-TOE I in complex with PPE (green) and CMTI-I in complex with bovine trypsin (blue; PDB code 1ppe), with protease structures omitted for clarity. Only the C$^\alpha$ positions of the inhibitors were used for superposition. Stereoview; disulfide bridges are in yellow.

1989; PDB code 1ppe). CMTI-I is a squash inhibitor from *Cucubita maxima*, with 20 out of 28 residues indentical to the EETI-II sequence and 16 out of 28 residues identical to the hybrid EETI-II-derived inhibitor. The backbone conformations of both inhibitors are very similar (Fig. 5), even in the binding region, where the sequence of HEI-TOE I is closer to the corresponding sequence of the ovomucoid inhibitor. This is not so surprising if one supposes that the conformation of

the inhibitor region interacting with the protease active-site cleft is determined largely by the geometry of this cleft and only partly by the remainder of the inhibitor structure itself, leading to the so-called 'canonical' conformation. This becomes more obvious if the structure of the complex between PPE and the HEI-TOE I inhibitor is compared with that of HLE and the OMTKY3 inhibitor (Bode *et al.*, 1986; PDB code 1ppf), where the corresponding parts of the peptide chain bound in the protease substrate cleft superpose well (Fig. 6a). Also, the binding loop of free OMTKY3 is quite flexible compared with the rest of the structure, as concluded from sets of NMR structures of OMTKY3 in solution (PDB codes 1tur, 1omu). The φ and ψ angles of the positions P1, P2, P'1 and P'2 in the binding loop of HEI-TOE I, when complexed to PPE, correspond well to those reported for other inhibitors with 'canonical' binding-loop conformations (Bode & Huber, 1992), whereas those for positions P3 and P'3 fall slightly outside the reported regions (Table 3). The two hydrogen bonds of the inhibitor P3 N to 216 C=O and P3 C=O to 216 N are also described for other 'canonical' loop conformations (Bode & Huber, 1992).

The distance between the PPE active-site Ser195 side-chain O atom (O$^\gamma$) and the HEI-TOE I leucine-4 carbonyl O atom in the oxyanion hole is 2.98 Å, which is somewhat shorter than that for the corresponding atoms in the HLE–OMTKY3 complex (3.22 Å), but larger than in the complex of PPE with a chymotrypsin/elastase inhibitor from the parasitic roundworm *Ascaris suum* (PDB code 1eai; 2.71 Å). In the trypsin–CMTI-I complex this distance is 2.88 Å, close to that of the PPE–HEI-TOE I complex which thus seems to be within a typical range for this class of inhibitors.

An important feature of the OMTKY3-derived optimized peptide and the HEI-TOE I squash inhibitor hybrid is the high specificity towards PPE resulting from the affinity-optimization process (Hilpert *et al.*, 2000), in contrast to the ovomucoid inhibitor fragment OMTKY3 (Table 4). The affinity differences between PPE and HLE increase from a factor of 6.7 for OMTKY3 or 1.5 for the OMTKY3-derived free peptide to a factor of 164 for the peptide optimized for PPE binding and even further to a factor of
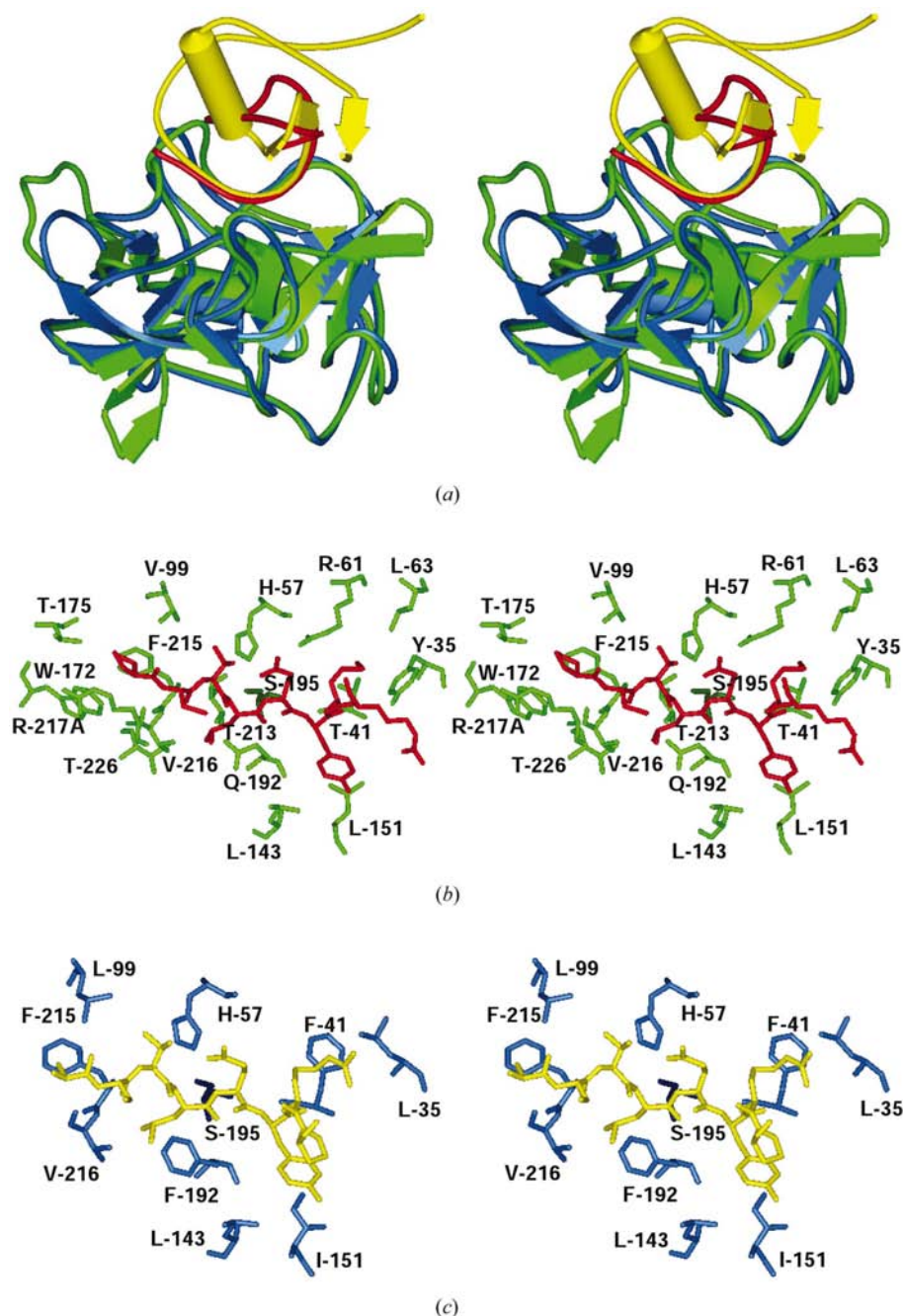

Figure 6
Stereoview of (*a*) the superposition of the complexes PPE–HEI-TOE I (green/red) and HLE–OMTKY3 (blue/yellow), schematic representation, (*b*) the hybrid inhibitor HEI-TOE I (red) bound to PPE (green; active-site Ser195 in dark green), important side-chain contacts for the binding-loop sequence PCTLEYMR, (*c*) the OMTKY3 inhibitor (yellow) bound to human leukocyte elastase (blue; active site Ser195 dark blue), important side-chain contacts for the binding-loop sequence ACTLEYRP.

Table 3
φ and ψ angles for the 'canonical' conformation of the HEI-TOE I binding loop in a complex with PPE and reference data from Bode & Huber (1992).

Position	Residue	HEI-TOE I–PPE		Reference data	
		φ (°)	ψ (°)	φ (°)	ψ (°)
P1	Leu	−103.2	28.7	−120 to −95	9–50
P2	Thr	−67.15	158.6	−100 to −60	139–180
P3	Cys	−141.3	123.7	−140 to −120	140–170
P′1	Glu	−87.05	157.9	−100 to −60	139–180
P′2	Tyr	−111.3	113.7	−140 to −99	70–120
P′3	Met	−150.3	143.3	−140 to −99	70–120

Table 4
Specificity of the inhibitor peptides and proteins [inhibition constants K_i (M)].

	PPE	HLE
OMTKY3	2.4×10^{-11}†	1.6×10^{-10}†
PACTLEYRC	1.7×10^{-4}	2.5×10^{-4}
PMTLEYR	6.7×10^{-7}	1.1×10^{-4}
Squash (EETI II)	8.2×10^{-4}	No inhibition
HEI-TOE I	9.8×10^{-8}	4.5×10^{-4}

† From Bigler *et al.* (1993).

about 4600 with the HEI-TOE I inhibitor. More details concerning the binding behaviour and protease specificity will be published elsewhere, together with thermodynamic data.

The differences in the specificity of the OMTKY3 ovomucoid fragment compared with the HEI-TOE I inhibitor must mainly be attributable to differences in the interactions of the residues Pro1 (Ala15 in OMTKY3) and Met7 (Arg in OMTKY3), since these are the only altered side chains of the binding loop with contacts to the protease active-site cleft. By inspecting these interactions in the PPE–HEI-TOE I inhibitor complex structure (Fig. 6*b*) and comparing them with the HLE–OMTKY3 complex (Fig. 6*c*), it becomes obvious that there are many more contacts for Pro1 with the S3 subsite than there are for Ala15: Pro1 shows several close contacts to Val99, Trp172, Phe215 and Arg217a in PPE, whereas Ala15 establishes contact with only two residues (Leu99 and Phe215) in the corresponding subsite of HLE (the importance of the P4 position for inhibitor specificity together with P1 has also been reported for other serine protease inhibitors with a 'canonical' binding mode; Bode & Huber, 1992). Elsewhere, for Met8 in the S′3 subsite of PPE there are van der Waals contacts with Leu63, Tyr35, Thr41 and Arg61, whereas the corresponding Arg21 in OMTKY3 interacts only with Leu35 and Phe41 in S′3 of HLE. In addition, a salt bridge is formed between Glu5 of the HEI-TOE I inhibitor and Arg61 in PPE, which does not exist in HLE since there is an Asn instead of Arg at the corresponding position. However, the fact that there is also a Glu in this position in OMTKY3 (Glu19), where the affinities towards HLE and PPE are not so different, indicates that this salt bridge does not contribute significantly, at least to the high affinity of OMTKY3, which may be accounted for by other interactions.

Whereas it is possible to give a sufficiently convincing explanation for the differences in specificity, it is not easy to understand why the affinity of OMTKY3 towards both elastases is several orders of magnitude higher than that of the HEI-TOE I inhibitor towards PPE. The minor outward shift of two loops in the PPE structure adjacent to the bound HEI-TOE I molecule (see above) may indicate some steric hindrance on inhibitor binding. If the structure of the complex of HLE with the ovomucoid fragment OMTKY3 (PDB code 1ppf) is compared with those of complexes of HLE with small peptide inhibitors (PDB code 1hne or 1ppg, not shown) then minor shifts in loop positions also become obvious, but in the

opposite direction, leading to a closer contact of HLE with the OMTKY3 inhibitor molecule.

An important reason why we incorporated the OMTKY3-derived affinity-optimized peptide into a squash-inhibitor scaffold was to improve the stability of this peptide. The free inhibitory peptide is also a substrate, but a poor one, meaning that it undergoes slow hydrolysis by elastase, as indicated by a loss of inhibitory power during prolonged preincubation with the protease (data to be published elsewhere). In contrast, the HEI-TOE I inhibitor with the peptide incorporated in the squash-inhibitor scaffold does not alter its inhibitory power on a comparable time scale. An inspection of the electron density of the scissile bond (Leu-Glu) in the complex clearly indicates that this peptide bond remains uncleaved. This means that the peptide conformation in the context of the squash-inhibitor fold is stabilized in such a way that it is prevented from undergoing hydrolysis, *e.g.* by freezing a Michaelis complex conformation unfavourable for formation of the proper transition state necessary for effective catalysis of the peptide-bond cleavage by the protease (Bode & Huber, 1992). Alternatively, if treated as an equilibrium process between bond cleavage and resynthesis, it may be that resynthesis is very effective, as has been shown for the CMTI–salmon trypsin complex (Helland *et al.*, 1999).

In conclusion, the following observations and information can be drawn from the X-ray structure of the HEI-TOE I–PPE complex.

(i) Only minor changes in main-chain or side-chain positions in PPE are induced by inhibitor binding.

(ii) The conformation of the binding loop in the inhibitor is fixed in a manner very similar to otherwise structurally different inhibitor proteins. This allocates HEI-TOE I to the group of 'canonical' inhibitors with substrate-like binding behaviour.

(iii) The marked specificity of the HEI-TOE I inhibitor for PPE over HLE, in contrast to the ovomucoid fragment OMTKY3 from which the inhibitory peptide was derived, is mainly a consequence of improved interactions in the subsite positions P4 and P′3. These positions are comparatively distant from the important P1 and P2 subsites. A salt bridge formed in P′1 of PPE but not present with HLE may contribute to this specificity, but obviously does not play a major role in the high affinity of OMTKY3, which is comparable for PPE and HLE.

(iv) The incorporation of an inhibitory peptide into the squash inhibitor mini-protein scaffold leads to a considerable

increase in affinity and especially in the stability of this peptide towards proteolytic attack. This constitutes an important advantage if these inhibitors are to be developed for use, for example, in therapeutic applications.

(v) The overall fold of the small squash trypsin inhibitor protein is sufficiently stable to allow considerable amino-acid exchange in its protease-binding loop as shown previously by Christmann *et al.* (1999), making it a valuable scaffold for the production of binding compounds with specificities deviating from the original protease-inhibition function.

References

Bigler, T. L., Lu, W., Park, S. J., Tashiro, M., Wieczorek, M., Wynn, R. & Laskowski, M. (1993). *Protein Sci.* **2**, 786–799.

Bode, W., Greyling, H. J., Huber, R., Otlewski, J. & Wilusz, T. (1989). *FEBS Lett.* **242**, 285–292.

Bode, W., Wei, A. Z., Huber, R., Meyer, E., Travis, J. & Neumann, S. (1986). *EMBO J.* **5**, 2453–2458.

Bode, W. & Huber, R. (1992). *Eur. J. Biochem.* **204**, 433–451.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Chiche, L., Gaboriaud, C., Heitz, A., Mornon, J. P., Castro, B. & Kollman, P. A. (1989). *Proteins Struct. Funct. Genet.* **6**, 405–417.

Christmann, A., Walter, K., Wentzel, A., Kratzner, R. & Kolmar, H. (1999). *Protein Eng.* **12,** 797–806.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Evans, S. V. (1993). *J. Mol. Graph.* **11**, 134–138.

Helland, R., Berglund, G. I., Otlewski, J., Apostoluk, W., Andersen, O. A., Willassen, N. P. & Smalas, A. O. (1999). *Acta Cryst.* D**55**, 139–148.

Hilpert, K., Schneider-Mergener, J. & Aÿ, J. (2002). *Acta Cryst.* D**58**, 672–674.

Hilpert, K., Wessner, H., Schneider-Mergener, J. & Hohne, W. (2000). *J. Biochem. (Tokyo)*, **128**, 1051–1057.

Hubbard, S. & Thornton, J. (1996). *NACCESS v.2.1.1 – Atomic Solvent Accessible Area Calculations*, http://wolf.bms.umist.ac.uk/naccess/.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Nielsen, K. J., Alewood, D., Andrews, J., Kent, S. B. & Craik, D. J. (1994). *Protein Sci.* **3**, 291–302.

Otlewski, J. (1990). *Biol. Chem. Hoppe-Seyler*, **371**, 23–28.

Otlewski, J. & Krowarsch, D. (1996). *Acta Biochim. Pol.* **43**, 431–444.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Volkmer-Engert, R., Landgraf, C. & Schneider-Mergener, J. (1998). *J. Pept. Res.* **51**, 365–369.

Wilmouth, R. C., Edman, K., Neutze, R., Wright, P. A., Clifton, I. J., Schneider, T. R., Schofield, C. J. & Hajdu, J. (2001). *Nature Struct. Biol.* **8**, 689–694.

Wilmouth, R. C., Westwood, N. J., Anderson, K., Brownlee, W., Claridge, T. D., Clifton, I. J., Pritchard, G. J., Aplin, R. T. & Schofield, C. J. (1998). *Biochemistry*, **37**, 17506–17513.

Würtele, M., Hahn, M., Hilpert, K. & Hohne, W. (2000). *Acta Cryst.* D**56**, 520–523.