Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of *Locusta migratoria* protease inhibitor 3 (LMPI-3) in complex with *Fusarium oxysporum* trypsin

Philippe Leone,[a] Alain Roussel[b] and Christine Kellenberger[b]*

[a]AFMB UMR6098 Case 932, 163 Avenue de Luminy, 13288 Marseille CEDEX 9, France, and [b]CBM UPR4301 Rue Charles Sadron 45071 Orléans CEDEX 2, France

Correspondence e-mail: kellen@cnrs-orleans.fr

Received 29 July 2008
Accepted 22 September 2008

PDB Reference: LMPI-3–trypsin complex, 2vu8, r2vu8sf.

Previous studies have shown that the trypsin inhibitors LMPI-1, LMPI-3 and SGTI from locusts display an unusual species selectivity. They inhibit locust, crayfish and fungal trypsins several orders of magnitude more efficiently than bovine trypsin. In contrast, the chymotrypsin inhibitors from the same family, LMPI-2 and SGCI, are active towards mammalian enzymes. The crystal structures of a variant of LMPI-1 and of LMPI-2 in complex with bovine chymotrypsin have revealed subtle structural differences between the trypsin and chymotrypsin inhibitors. In a previous report, it was proposed that Pro173 of bovine trypsin is responsible for the weak inhibitory activity of LMPI-1 and LMPI-3. A fungal trypsin from *Fusarium oxysporum* contains Gly173 instead of Pro173 and has been shown to be strongly inhibited by LMPI-1 and LMPI-3. To explore the structural features that are responsible for this property, the crystal structure of the complex between LMPI-3 and *F. oxysporum* trypsin was determined at 1.8 Å resolution. This study indicates that this small inhibitor interacts with the protease through the reactive site P3–P4′ and the P10–P6 residues. Comparison of this complex with the SGTI–crayfish trypsin and BPTI–bovine trypsin complexes reinforces this hypothesis on the role of residue 173 of trypsin in species selectivity.

1. Introduction

During the last decade, a new family of tight-binding serine protease inhibitors has emerged. This family, which has been named the ‘pacifastin family’ (Gaspari *et al.*, 2004), is composed of small peptides isolated from the locusts *Locusta migratoria* (LMPIs) and *Schistocerca gregaria* (SGTI and SGCI) and of domains of the pacifastin protein itself (Liang *et al.*, 1997). The *L. migratoria* peptides were initially named PMP-D2, PMP-C (Nakakura *et al.*, 1992) and HI (Kellenberger *et al.*, 1995) and were subsequently renamed LMPI-1 to LMPI-3, while the *S. gregaria* peptides were named either SGTI and SGCI or SGPI-1 and SGPI-2. These peptides are 35–36 residues long, display approximately 40% sequence identity and possess six conserved cysteines that are involved in three disulfide bonds. Although these peptides exhibit similarities, they have been classified into two separate groups according to their functional and structural properties (Roussel *et al.*, 2001). The peptides in group I (SGTI, LMPI-1 and LMPI-3) are trypsin inhibitors, while those in group II

(SGCI and LMPI-2) are chymotrypsin inhibitors. The most striking feature of this family is that only members of group I display species selectivity, inhibiting mammalian trypsins only weakly (K_i of around 10^{-7} *M* for bovine trypsin and K_i of $>10^{-5}$ *M* for porcine trypsin; Kellenberger *et al.*, 1995; Malik *et al.*, 1999). In contrast, the peptides from group II are potent inhibitors of mammalian chymotrypsins (Kellenberger *et al.*, 1995; Malik *et al.*, 1999). The nature of the P1 residue in the active site (Leu30; Schechter & Berger, 1967) agrees with their activity towards chymotrypsin. By analogy, the P1 residue of group I was expected to be Arg29 and this hypothesis was confirmed experimentally (Kellenberger *et al.*, 1995). Such a P1 residue accounts for trypsin specificity and is in apparent contradiction with the weak activity towards mammalian trypsins. More recently, LMPI-1, LMPI-3 and SGTI have been shown to be highly potent towards trypsins extracted from crustaceans (crayfish and shrimp; Patthy *et al.*, 2002) and from an insect (Kellenberger *et al.*, 2003). It thus appears that these trypsin inhibitors display novel species selectivity.

The elucidation of the structure of LMPI-1 (Mer *et al.*, 1994) and subsequently of LMPI-2 (Mer *et al.*, 1996) by NMR allowed the classification of these inhibitors into a novel family. Their structure consists of a twisted β-sheet composed of three antiparallel strands and stabilized by three disulfide bridges. The structure of LMPI-3 was subsequently solved using the same technique (Kellenberger *et al.*, 2003). The structures of SGTI and SGCI from *S. gregagria* have also been determined by NMR (Gaspari *et al.*, 2002). All these inhibitors show very similar overall conformations. The structures of LMPI-Iv, a variant of LMPI-1 (with Arg29 mutated to Leu and Lys30 mutated to Met), and of LMPI-2, both in complex with bovine chymotrypsin, have been solved by X-ray crystallography (Roussel *et al.*, 2001). Finally, a recent study described the complex between a crustacean trypsin (from the crayfish *Pontastacus leptodactylus*) and SGTI from *S. gregagria* (Fodor *et al.*, 2005).

Comparison of the crystal structures shows that the LMPI-1 variant exhibits an additional interaction site between residues 20–24 (loop P6–P10) and the loop consisting of residues 172–175 of α-chymotrypsin (Roussel *et al.*, 2001). The corresponding loop in bovine trypsin contains a proline instead of a glycine at position 173. The docking of LMPI-1 to bovine trypsin reveals steric hindrance in this region. Our hypothesis is that this unfavourable interaction may contribute to the species selectivity. A search of a sequence database showed that the proline at position 173 is highly conserved among mammalian trypsins (MEROPS database). In contrast, insect, crustacean and fungal trypsins display a larger loop that is devoid of proline (Figs. 1 and 5*b*). In a previous study, we showed that the docking of LMPI-1 and LMPI-3 to *Fusarium oxysporum* trypsin occurs without any obvious steric hindrance. This model was validated by the measurement of a potent inhibitory activity towards the fungal trypsin ($K_i \leq 2 \times 10^{-11}$ *M*; Kellenberger *et al.*, 2003). In the present study, we crystallized and solved the structure of locust LMPI-3 in complex with *F. oxysporum* trypsin. The position of loop P6–P10 of the inhibitor in the crystal structure confirms our hypothesis. Our results are discussed in comparison to the structural study on the complex between crayfish trypsin and SGTI inhibitor.

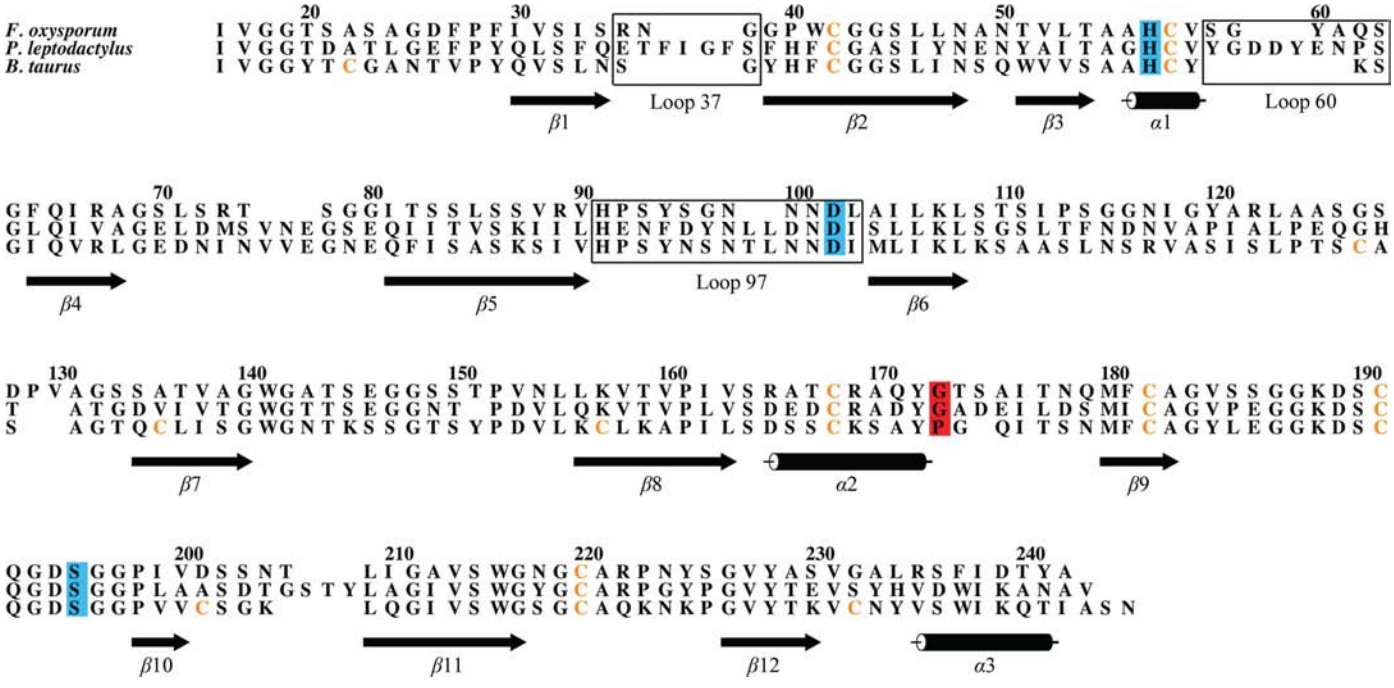

Figure 1
Sequence alignment of trypsins. The fungal (*F. oxysporum*), crayfish (*P. leptodactylus*) and bovine (*B. taurus*) trypsins were aligned based on superimposition of the three-dimensional structures using *ALSCRIPT* (Barton, 1993). The three loops (37, 60 and 97) are boxed, secondary-structure elements, *i.e.* helices and strands, are represented by cylinders and arrows, the cysteines are shown in orange, the amino acids of the catalytic triad His57, Asp102 and Ser195 are shown on a blue background and residue 173 is shown on a red background.

2. Experimental procedure

2.1. Crystallization and X-ray diffraction

Trypsin from *F. oxysporum* was kindly provided by Professor W. R. Rypniewsky at a concentration of 9.7 mg ml^{-1} in a buffer solution containing 0.1 *M* boric acid, 2 m*M* $CaCl_2$, 0.1 *M* NaCl and 10 m*M* 3,3-dimethylglutaric acid pH 6.0. In crystallization trials, the protein was used at a concentration of 2.9 mg ml^{-1}. The peptide LMPI-3 was synthesized in the solid phase using the Fmoc strategy and refolded by air oxidation of the six cysteines as described previously (Kellenberger *et al.*, 1995). The enzyme and the inhibitor were mixed together in a molar ratio of 1:1.3 (enzyme:inhibitor) and incubated for 1 h before crystallization.

Crystals of the complex were obtained at 293 K using the hanging-drop vapour-diffusion method. The protein drops consisted of 1.5 µl complex mixed with an equal volume of reservoir solution containing 0.1 *M* sodium citrate pH 4.9 and 24% PEG 4000.

X-ray diffraction data were collected at 100 K to 1.8 Å resolution using a MAR CCD imaging plate on beamline EH1 at the ESRF (Grenoble, France). The data were processed with *MOSFLM* and *SCALA* (Powell, 1999). The statistics are given in Table 1. Specific volume calculations showed one molecule per asymmetric unit, with a solvent content of 60%. Despite interpretable diffraction patterns, the cumulative intensity distribution curve given by *TRUNCATE* (Collaborative Computational Project, Number 4, 1994) was sigmoidal, reflecting a twinned crystal. Using *CNS* (Brünger *et al.*, 1998), the twin fraction of the crystal was estimated to be 0.461–0.455 based on $\langle H \rangle$ and $\langle H^2 \rangle$, which is considered to be perfect hemihedral twinning.

2.2. Structure determination

The position of *F. oxysporum* trypsin (PDB code 1try) was determined by molecular replacement using the *auto-AMoRe* program (Navaza, 2001) with data between 3.5 and 15 Å resolution. The correlation coefficient and the *R* factor were 0.284 and 0.467, respectively. After rigid-body refinement in *CNS*, the structure of the variant LMPI-1v in complex with bovine chymotrypsin (PDB code 1gl0) was placed in the $F_o - F_c$ electron-density map by superimposition of the enzymes. The residues corresponding to the sequence of LMPI-3 were then built manually with *TURBO-FRODO* (Roussel & Cambillau, 1991). R_{work} and R_{free} were very high at this stage (0.370 and 0.406, respectively) and strong electron density remained in the solvent channel (Fig. 2*a*). The strategy for refinement was to first carry out several cycles of slow cooling, followed by individual *B*-factor refinements with *CNS* (with the 'Refinement with hemihedral twinning' option) using the original data and taking into account the hemihedral twinning (twinning operation $-h, -k, l$; twinning fraction 0.5). For further refinement, the data were then detwinned in *CNS* using the last model obtained and were used in refinement with *REFMAC* (Murshudov *et al.*, 1997). After several cycles of refinement, a lot less density remained outside the protein region (Fig. 2*b*) and R_{work} and R_{free} fell to 0.197 and 0.240, respectively. During refinement, modifications of the residues and positioning of the water molecules were made manually with *TURBO-FRODO*. It is noteworthy that the number of unique reflections decreased after each detwinning step to reach 27 186 at the end of the refinement. According to the Ramachandran plot (see Table 1), one residue [Asp13(I) of LMPI-3] is located in a disallowed region. This is also the case for SGTI in complex with crayfish trypsin (Fodor *et al.*, 2005). The quality of the final electron-density map is shown in Fig. 2(*c*).

Table 1
Crystallographic data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell (1.86–1.8 Å).

Data collection	
Space group	$P3_12_1$
Unit-cell parameters	$a = b = 70.35$, $c = 110.27$
Resolution limits (Å)	30–1.8
Completeness (%)	99.6 (99.6)
Redundancy	5.4 (5.4)
$I/\sigma(I)$	9.3 (2.4)
R_{merge}†	7.9 (30.9)
Refinement	
No. of reflections in working set	27186
No. of reflections in test set	1301
No. of non-H atoms	
Enzyme	1558
Inhibitor	245
Water	191
R_{work}‡	0.197
R_{free}§	0.240
B factor (Å^2)	
Enzyme	10.0
Inhibitor	14.5
Water	22.6
R.m.s.d. bond lengths (Å)	0.009
R.m.s.d. bond angles (°)	1.33
Ramachandran plot, residues in	
Most favoured regions (%)	87.8
Additional allowed regions (%)	11.7
Generously allowed regions (%)	0
Disallowed regions (%)	0.5
PDB code	2vu8

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{work} = \sum_h ||F_{obs}| - |F_{calc}|| / \sum_h |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes of the working data set, respectively. § R_{free} was calculated with 5% of the diffraction data that were not used during the refinement.

2.3. Rigid-body docking

Fast and simple docking of LMPI-3 onto various trypsins was performed by superimposing the structures of the trypsins onto that of the fungal trypsin from the present study. The superimposition was performed using the program *TURBO-FRODO*. An initial positioning of each trypsin was performed using the C^α atoms of the amino acids of the catalytic triad. The superimposition was then refined by taking into account all the C^α pairs at a distance of less than a given cutoff distance. In our strategy, the initial 1.5 Å cutoff distance was decreased to 1 Å in a second step. The quality of the superimposition could then be estimated by the number of C^α pairs at a distance of less than 1 Å. As no further molecular-dynamics calculations were applied, the procedure resulted in

rigid-body positioning of the inhibitor in the various trypsin active sites.

3. Results

3.1. Overall structure of LMPI-3 in complex with *F. oxysporum* trypsin

The crystals of the complex between LMPI-3 and the fungal trypsin belonged to space group $P3_12_1$ and diffracted to 1.8 Å resolution. Despite perfect hemihedral twinning, structure determination was possible by molecular replacement using the structure of the same trypsin (PDB code 1try) as a search model. Several cycles of refinement yielded an R_{work} and an R_{free} of 0.197 and 0.240, respectively (Table 1).

The model of the complex accounts for the 224 residues of the fungal trypsin (16–242) and for 33 residues of the inhibitor (Fig. 3*a*). The two first residues [Ala1(I) and Gly2(I)] of LMPI-3 were not built because of a lack of electron density. The structure of *F. oxysporum* trypsin exhibits the classical fold of serine proteases and does not show any large differences when compared with the complex with phosphoryl-isopropane (Rypniewski *et al.*, 1993; r.m.s.d. of 0.369 Å on 219 of 224 C^{α} atoms). Only a few slight deviations occur at residue 150 and in the regions 92–99 and 201–204. Minor differences are described in §3.3. The structure of LMPI-3 (Fig. 3*a*) is similar to that previously determined for LMPI-1 (Roussel *et al.*, 2001). The X-ray structure of bound LMPI-3 was compared with the free form (minimized average structure; PDB code 1wo9) solved in solution by ^{1}H NMR (Kellenberger *et al.*, 2003). Using a cutoff of 1 Å, the two structures superimpose on 14 residues with an r.m.s.d. of 0.56 Å. Although the third strand of the β-sheet (residues 25–29) is slightly shifted, the overall shape of the molecule is conserved. The main

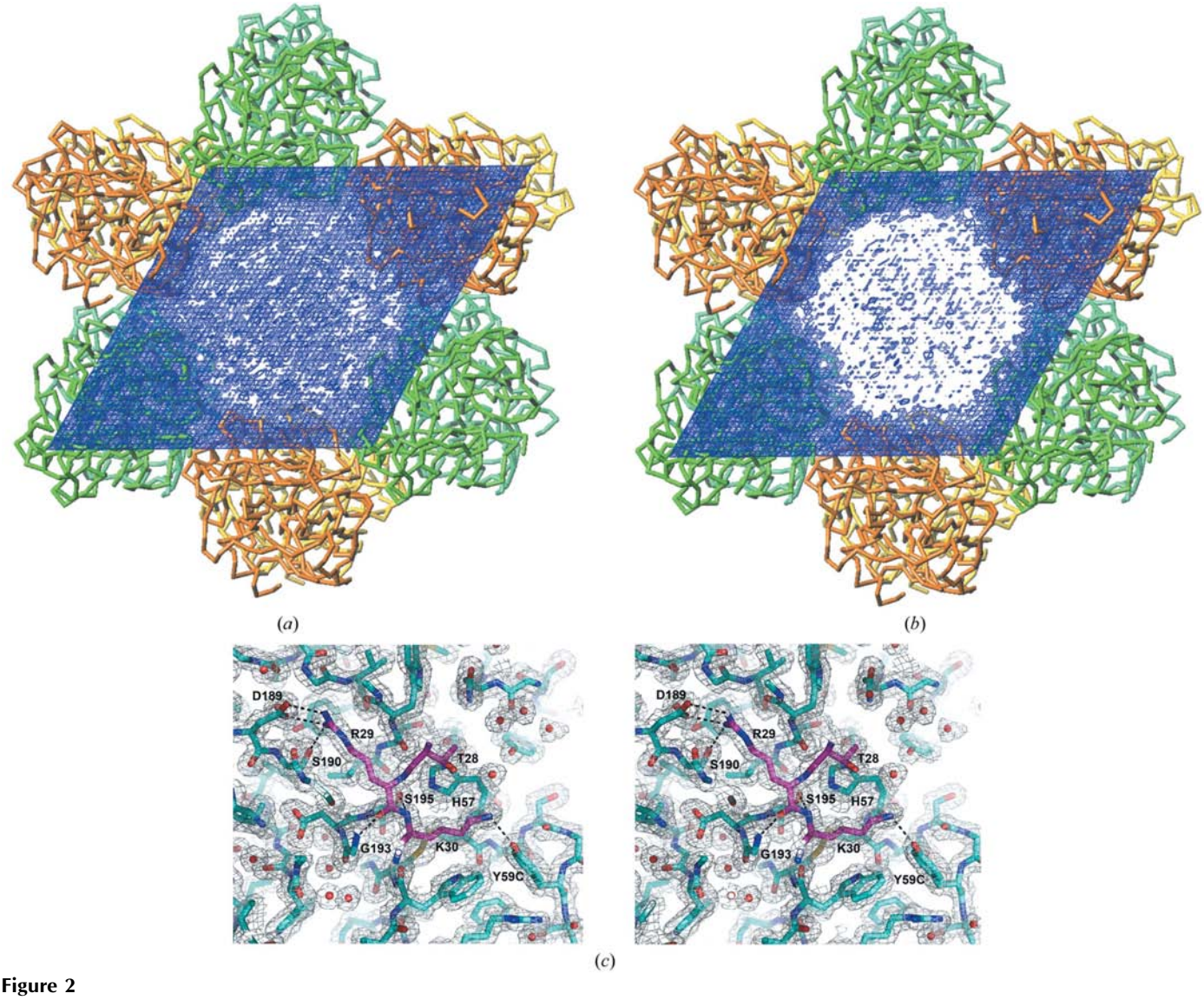

Figure 2
(*a*) and (*b*) Electron density in the solvent channel. Symmetric molecules were generated with *TURBO-FRODO* (Roussel & Cambillau, 1991) to visualize the crystal packing and the solvent channel along the *c* axis (*a*) before refinement and (*b*) after refinement. (*c*) Stereoview of the final ($2F_o - F_c$) electron-density map (contoured at 1σ) in the interface region between trypsin (C atoms in blue) and LMPI-3 (C atoms in purple).

difference arises from the N- and C-terminal tails (residues 3–6 and 30–35), which completely diverge. The superimposition of 33 residues gives an r.m.s.d. of 3.19 Å. The 16 NMR structures display an average overall r.m.s.d. of 1.98 Å, while the value is 0.62 Å for residues 3–28. As the r.m.s.d. values from NMR studies are smaller than the deviations resulting from NMR/X-ray structure superimposition, the latter may be considered as conformationally different in the N- and C-terminal regions as well as in the active loop. Asn15(I), which plays a central role in the stabilization of P1′ and P2 through a hydrogen-bond network, does not interact with P1′ and P2 in the NMR structure, this contributes to the flexibility of the active loop.

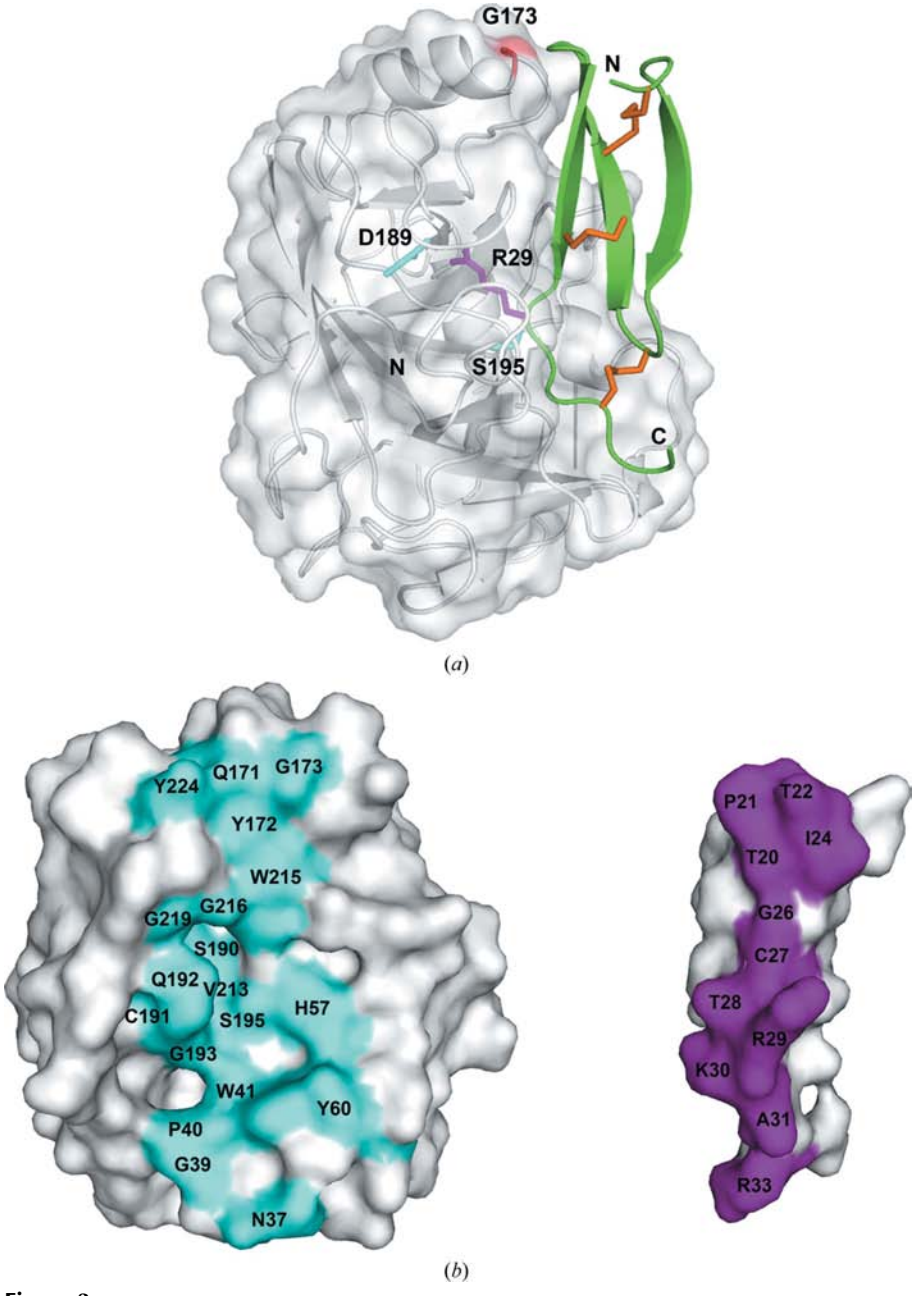

Figure 3
(*a*) General overview of the complex between trypsin from *F. oxysporum* and LMPI-3 from *L. migratoria*. The trypsin is represented as a ribbon diagram (coloured grey) and its surface is shown as half-transparent. Residue Gly173 is coloured red. The residues Asp189 (specificity pocket) and Ser195 (catalytic serine) are depicted as sticks and coloured cyan. The inhibitor LMPI-3 is in ribbon representation and is coloured green. The disulfide bridges of LMPI-3 are shown in orange. The residue Arg29 (P1) is depicted as sticks and coloured magenta. The figure was generated using *PyMOL* (DeLano, 2002). (*b*) Interaction surfaces of *F. oxysporum* trypsin and of LMPI-3. The two partners of the complex are represented as grey molecular surfaces and the amino acids that are buried upon binding are labelled and are coloured cyan (trypsin) and magenta (inhibitor). This figure was generated using *PyMOL*.

3.2. Interaction at the P1 active site

Residues P4–P4′ interact with the enzyme in accordance with the 'canonical' model of protease–tight-binding inhibitor interaction (Bode & Huber, 2000). In addition to the hydrogen bond to Cys14(I) mentioned by Fodor *et al.* (2005), the P1′ residue interacts with loop 60 (as described in §3.4). The P4′ residue is in contact with loop 37 (see below).

3.3. Secondary interaction site of LMPI-3: the P6–P10 loop

Upon binding to the trypsin, LMPI-3 buries 1007 Å^2 of its solvent-accessible surface (Fig. 3*b*). In comparison, BPTI buries only 763 Å^2 of its surface upon binding to bovine trypsin (PDB code 3btk). This difference mainly arises from the occurrence of a secondary interaction site for LMPI-3. This latter is composed of the region P6–P10 (residues 20–24), which contacts residues 171–173 of the fungal trypsin. The hydroxyl groups of Thr20(I) and Thr22(I) (P10 and P8) form hydrogen bonds to the Gln171 and Tyr172 backbones, which are reinforced by van der Waals interactions between the P9, P8, P6 and P4 residues and Gln171, Tyr172, Gly173, Trp215 and Tyr224.

Compared with the free trypsin (PDB code 1try), the C$^\alpha$ trace between residues 173 and 175 shows a slight deviation (about 0.5 Å) away from the inhibitor; this could correspond to a structural adaptation of this loop upon binding.

3.4. Comparison of trypsin–inhibitor complexes

A recent study explored the structural features responsible for species selectivity by solving the structure of SGTI, a member of the pacifastin family, in complex with a crayfish trypsin (Fodor *et al.*, 2005). This com-

plex (PDB code 2f91) was superimposed on the structure in the present study. SGTI and LMPI-3 superimpose very well over 32 residues (3–34) with an r.m.s.d. of 0.79 Å. Although the core of the trypsins superimposes well, crayfish trypsin displays several longer loops. Three loops (named loops 37, 60 and 97) are noteworthy with regard to inhibitor binding (Figs. 1 and 4). Two of them (loops 37 and 60) have previously been described by Fodor *et al.* (2005). Loop 37 (residues 37–39) of both fungal and crayfish trypsins interacts with P4′ of LMPI-3 and SGTI, although in a very different manner. One hydrogen bond occurs between the carbonyl O atom of Asn37 (*F. oxysporum*) and the guanidinium group of Arg33(I) of LMPI-3. In crayfish trypsin, a hydrophobic cluster of phenylalanines interacts with Pro33(I) of SGTI. Loop 60 (residues 59–61) has a rather divergent backbone trace. Tyr59C of fungal trypsin makes a hydrogen bond to the Lys30(I) (P1′) side chain. In crayfish, the loop takes another orientation that prevents any contact with SGTI, the P1′ side chain of which is maintained in a conformation similar to that of LMPI-3. Loop 97 (residues 94–99) of fungal trypsin is two residues shorter than that of crayfish. In crayfish, the main chain of Asn97 interacts with the P12 position of SGTI [side chain OD1 of Asn18(I)]. The small size of the fungal trypsin loop prevents any contact with LMPI-3.

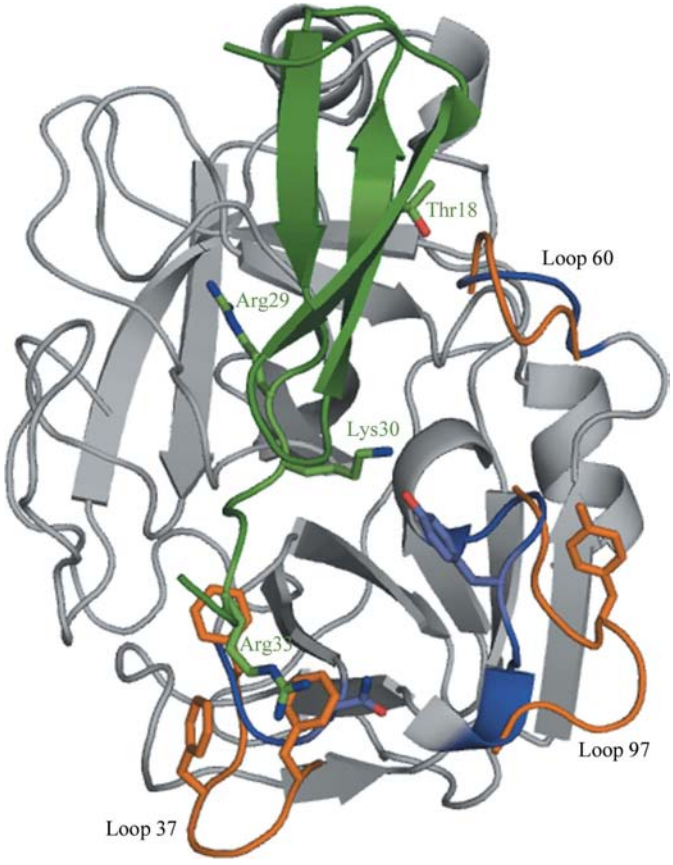

Figure 4
Comparison of the active-site environment between trypsins from *F. oxysporum* and from *P. leptodactylus*. The fungal trypsin is represented as a grey ribbon and its loops 37, 60 and 97 are coloured blue. The side chains of Asn37 and Tyr59C are represented as sticks, with the N atom in blue and the O atom in red. The loops 37, 60 and 97 of crayfish trypsin are represented as ribbons and coloured orange. The side chains of Phe37, Phe37C, Phe39 and Tyr60 are represented as sticks. The inhibitor LMPI-3 is shown as a green ribbon. The side chains of residues P1 (Arg29), P1′ (Lys30), P4′ (Arg33) and Thr18 are represented as sticks and coloured according to atom type, with C atoms in green.

3.5. Role of loop 173

In order to explain the weak inhibitory activity of LMPI-3 towards bovine trypsin (K_i of around 10^{-7} *M*), the inhibitor was docked onto this trypsin. Rigid-body docking is relevant because of the rigidity of the inhibitor. The docking was performed by superimposing the structures of the bovine trypsin–BPTI (PDB code 3btk) and the fungal trypsin–LMPI-3 complexes. The weak binding between LMPI-3 and bovine trypsin can be attributed to the contribution of unfavourable steric hindrances. They occur between the P6–P10 loop [Thr20(I), Pro21(I) and Thr22(I)] and loop 173 (residues 172–175), where very short distances (up to 1.3 Å) are found between Pro173 and Thr22(I) and no accommodation seems possible as the two loops are highly constrained. A proline at position 173 constrains the direction taken by the backbone at the end of helix 164–172, as previously mentioned by Fodor *et al.* (2005). In fact, the C atoms CB, CG and CD protrude toward the inhibitor (Fig. 5*a*). Thus, it appears that Pro173 does play a role in the species selectivity of inhibition.

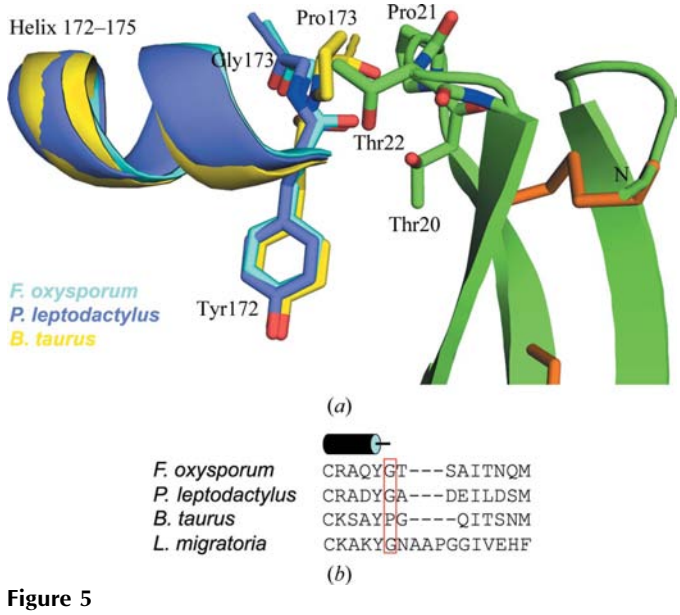

Figure 5
(*a*) Close-up view of the interaction between the trypsins and the inhibitor LMPI-3 in the region containing residue 173. For the trypsins, the end of the α2 helix is represented as a spiral and one colour is attributed to each trypsin: cyan for fungal, blue for crayfish and yellow for bovine. LMPI-3 is coloured green and shown as a ribbon diagram. The disulfide bonds are shown in orange. Residues 172 and 173 of the trypsins and residues 20, 21 and 22 of LMPI-3 are shown in stick format. O and N atoms are coloured red and blue, respectively. This figure was drawn using *PyMOL*. (*b*) Sequence alignment of trypsins in the region of loop 173. The sequence of fungal (*F. oxysporum*), crayfish (*P. leptodactylus*) and bovine (*B. taurus*) trypsins are aligned based on superposition of the three-dimensional structures. Three gaps were manually introduced in order to align the loop of *L. migratoria* trypsin. Residues 173 are boxed in red and an arrow indicates the α2 helix (C-terminal part).

4. Concluding remarks

Residue 172, a tyrosine that is strictly conserved in trypsins, has been shown to interact synergistically with residues of the S1 binding pocket and was thus identified as a substrate-specificity determinant in the trypsin family (Hedstrom *et al.*, 1994). The key role of Tyr172 implies a strictly conserved location of the helix 164–172 and thus the slightest variation at position 173 (Pro instead of Gly) is likely to provoke a dramatic effect.

Another question remains regarding the physiological targets of the pacifastin inhibitors. Recently, the cDNA of a trypsin-like protease originating from the same organism as the LMPIs was cloned (Wei *et al.*, 2007). Alignment with the trypsins in this study (Fig. 5*b*) highlights a divergence at loop 173 that is unusually large and involves a nonconservative replacement by a flexible Gly residue of a larger rigid Pro at position 173. The absence of a proline is likely to be favourable to inhibition by LMPI-1 and by LMPI-3; however, the unusual size of this loop may disturb the binding. Thus, further work appears to be necessary in order to evaluate the inhibition of this potential target trypsin by the locust inhibitors.

We thank Professor Rypniewski for providing *F. oxysporum* trypsin and the European Synchrotron Radiation Facility (ESRF) at Grenoble, in particular beamline ID214-EH1. We thank Fabienne Damblon for careful reading and for improving the English language of the manuscript.

References

Barton, G. J. (1993). *Protein Eng.* **6**, 37–40.
Bode, W. & Huber, R. (2000). *Biochim. Biophys. Acta*, **1477**, 241–252.
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.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://www.pymol.org.
Fodor, K., Harmat, V., Hetenyi, C., Kardos, J., Antal, J., Perczel, A., Patthy, A., Katona, G. & Graf, L. (2005). *J. Mol. Biol.* **350**, 156–169.
Gaspari, Z., Ortutay, C. & Perczel, A. (2004). *Bioinformatics*, **20**, 448–451.
Gaspari, Z., Patthy, A., Graf, L. & Perczel, A. (2002). *Eur. J. Biochem.* **269**, 527–537.
Hedstrom, L., Perona, J. J. & Rutter, W. J. (1994). *Biochemistry*, **33**, 8757–8763.
Kellenberger, C., Boudier, C., Bermudez, I., Bieth, J. G., Luu, B. & Hietter, H. (1995). *J. Biol. Chem.* **270**, 25514–25519.
Kellenberger, C., Ferrat, G., Leone, P., Darbon, H. & Roussel, A. (2003). *Biochemistry*, **42**, 13605–13612.
Liang, Z., Sottrup-Jensen, L., Aspan, A., Hall, M. & Soderhall, K. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 6682–6687.
Malik, Z., Amir, S., Pal, G., Buzas, Z., Varallyay, E., Antal, J., Szilagyi, Z., Vekey, K., Asboth, B., Patthy, A. & Graf, L. (1999). *Biochim. Biophys. Acta*, **1434**, 143–150.
Mer, G., Hietter, H., Kellenberger, C., Renatus, M., Luu, B. & Lefevre, J. F. (1996). *J. Mol. Biol.* **258**, 158–171.
Mer, G., Kellenberger, C., Koehl, P., Stote, R., Sorokine, O., Van Dorsselaer, A., Luu, B., Hietter, H. & Lefevre, J. F. (1994). *Biochemistry*, **33**, 15397–15407.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nakakura, N., Hietter, H., Van Dorsselaer, A. & Luu, B. (1992). *Eur. J. Biochem.* **204**, 147–153.
Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.
Patthy, A., Amir, S., Malik, Z., Bodi, A., Kardos, J., Asboth, B. & Graf, L. (2002). *Arch. Biochem. Biophys.* **398**, 179–187.
Powell, H. R. (1999). *Acta Cryst.* D**55**, 1690–1695.
Roussel, A. & Cambillau, C. (1991). *Silicon Graphics Geometry Partners Directory*. Mountain View, USA: Silicon Graphics.
Roussel, A., Mathieu, M., Dobbs, A., Luu, B., Cambillau, C. & Kellenberger, C. (2001). *J. Biol. Chem.* **276**, 38893–38898.
Rypniewski, W. R., Hastrup, S., Betzel, C., Dauter, M., Dauter, Z., Papendorf, G., Branner, S. & Wilson, K. S. (1993). *Protein Eng.* **6**, 341–348.
Schechter, I. & Berger, A. (1967). *Biochem. Biophys. Res. Commun.* **27**, 157–162.
Wei, Z., Yin, Y., Zhang, B., Wang, Z., Peng, G., Cao, Y. & Xia, Y. (2007). *Dev. Growth Differ.* **49**, 611–621.