Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary crystallographic characterization of the N-terminal Kunitz domain of boophilin

Tatiana B. Cereija,[a] Ana C. Figueiredo,[a] Daniele de Sanctis,[b] Aparecida S. Tanaka[c] and Pedro José Barbosa Pereira[a]*

[a]IBMC – Instituto de Biologia Molecular e Celular, Universidade do Porto, 4150-180 Porto, Portugal, [b]Structural Biology Group, ESRF, 6 Rue Jules Horowitz, 38043 Grenoble CEDEX, France, and [c]Departamento de Bioquímica, Escola Paulista de Medicina, Universidade Federal de São Paulo, 04044-020 São Paulo-SP, Brazil

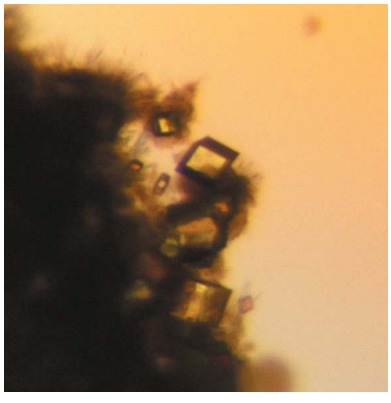
Correspondence e-mail: ppereira@ibmc.up.pt

Received 27 January 2012
Accepted 7 February 2012

Boophilin is a tight-binding thrombin inhibitor composed of two canonical Kunitz-type domains in a tandem arrangement. Thrombin-bound boophilin can inhibit a second trypsin-like serine proteinase, most likely through the reactive loop of its N-terminal Kunitz domain. Here, the crystallization and preliminary crystallographic analysis of the isolated N-terminal domain of boophilin is reported. The crystals belonged to the orthorhombic space group $P2_12_12_1$ and diffracted to beyond 1.8 Å resolution using a sealed-tube home source and to 0.87 Å resolution at a synchrotron source.

1. Introduction

Boophilin is a tight-binding (K_i = 57 p*M*; Soares *et al.*, 2012) thrombin inhibitor from the cattle tick *Rhipicephalus* (*Boophilus*) *microplus* (Macedo-Ribeiro *et al.*, 2008) and the only structurally characterized natural anticoagulant with two tandem canonical Kunitz-fold domains (Fuentes-Prior *et al.*, 1997; Koh *et al.*, 2011; Macedo-Ribeiro *et al.*, 2008; Richardson *et al.*, 2000; Rydel *et al.*, 1990; van de Locht *et al.*, 1995, 1996; Corral-Rodríguez *et al.*, 2009). The N-terminal Kunitz domain of boophilin has a lysine residue (Lys31) at the P1 position which is ideally suited for inhibiting trypsin-like enzymes. A basic residue at this position is conserved in the boophilin-like haemalin from *Haemaphysalis longicornis* (Liao *et al.*, 2009) and in the closely related inhibitors from *Dermacentor variabilis* (Anderson *et al.*, 2008) and *Amblyomma maculatum* (UniProt entry G3MNQ5; Fig. 1). However, thrombin inhibition by boophilin occurs by a noncanonical mechanism reminiscent of that of ornithodorin (van de Locht *et al.*, 1996), in which the N-terminal segment of the inhibitor inserts into the active-site cleft of thrombin while the C-terminal Kunitz domain interacts with exosite I of the proteinase mostly through hydrophobic interactions (Macedo-Ribeiro *et al.*, 2008). Strikingly, and in contrast to all other natural thrombin inhibitors characterized to date, when in complex with thrombin boophilin retains the ability to interact with an additional (nonthrombin) trypsin-like serine proteinase molecule (Macedo-Ribeiro *et al.*, 2008). This second interaction seems to only involve the N-terminal domain of boophilin (Soares *et al.*, 2012), presumably through its conserved and canonical reactive loop. *In*



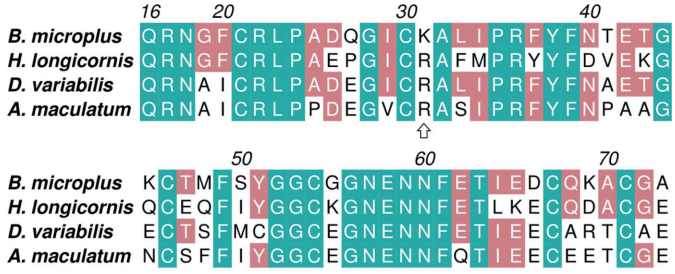


Figure 1
Amino-acid sequence alignment of boophilin D1 with the corresponding Kunitz domains of related inhibitors from *H. longicornis*, *D. variabilis* and *A. maculatum*. Strictly conserved amino acids are highlighted in green and those conserved between boophilin D1 and the other inhibitors are shown in salmon. The numbers above the alignment correspond to mature boophilin. The putative basic P1 residue is indicated by an arrow. This figure was prepared with *ALINE* (Bond & Schüttelkopf, 2009).

vivo, it is likely that in addition to inhibiting thrombin boophilin can target another serine proteinase of the coagulation cascade, such as factor Xa, in the context of the prothrombinase complex (Macedo-Ribeiro *et al.*, 2008). This would result not only in thrombin inhibition but also in impairment of thrombin generation, making boophilin a good model for the design of new therapeutic anticoagulants. In this paper, we report the purification, crystallization and preliminary crystallographic characterization of the isolated N-terminal domain (D1) of boophilin. The D1 crystals diffracted to 1.8 Å resolution using a sealed-tube home source and to 0.87 Å resolution at a third-generation synchrotron source and were suitable for structure determination.

2. Materials and methods

2.1. Expression and purification of recombinant boophilin D1

The N-terminal domain of boophilin (GenBank accession No. CAC82582.1; residues Gln16–Ala58) was expressed in *Pichia pastoris* KM71H cells transformed with pPICZαB (Invitrogen) harbouring an ORF coding for boophilin D1 fused to an N-terminal signal peptide (α-mating factor) as described previously (Soares *et al.*, 2012). The cell-free culture medium was loaded onto an immobilized-trypsin affinity column (AminoLink Plus Coupling Resin, Thermo Scientific Pierce) previously equilibrated with 20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl (buffer *A*). The column was extensively washed with buffer *A* and bound proteins were eluted with 0.2 *M* glycine pH 3.0. The collected fractions were immediately neutralized with 0.05 volumes of 1 *M* Tris (non-adjusted pH), pooled and dialysed against buffer *A*. Purified boophilin D1 was concentrated to 10 mg ml^{-1} on a centrifugal concentration device with a 3 kDa molecular-weight cutoff membrane (Millipore).

2.2. Crystallization of boophilin D1

Initial crystallization conditions were screened at 293 K using the sitting-drop method with commercial sparse-matrix crystallization screens. The drops consisted of equal volumes (2 µl) of protein (at 10 mg ml^{-1} in buffer *A*) and precipitant solution and were equilibrated against a 300 µl reservoir. Crystals were obtained after 34 d using 1.0 *M* lithium sulfate monohydrate, 2%(*w*/*v*) PEG 8000 as precipitant. The crystals were sequentially transferred to solutions of precipitant supplemented with 5, 10, 15, 20 and 25% glycerol as cryoprotectant and then flash-cooled in liquid nitrogen.

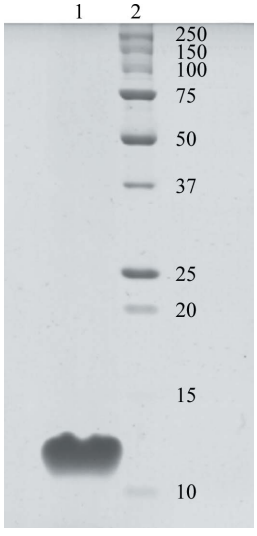


Figure 2
SDS–PAGE (15%) analysis of recombinant boophilin D1. Lane 1, affinity-purified boophilin D1; lane 2, protein molecular-weight standards (labelled in kDa).

Table 1
Statistics of data collection and processing.

Values in parentheses are for the outermost resolution shell.

Data set	In-house	ESRF ID29
Wavelength (Å)	1.54056	0.68878
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 24.8, b = 30.7, c = 55.0	a = 24.7, b = 30.4, c = 54.8
Resolution range (Å)	14.7–1.80 (1.90–1.80)	27.4–0.87 (0.92–0.87)
Reflections (measured/unique)	21141/4127	122468/34459
Completeness (%)	97.9 (87.6)	99.1 (99.1)
Multiplicity	5.1 (2.5)	3.6 (3.5)
$R_{\rm merge}$†	0.058 (0.150)	0.037 (0.570)
$R_{\rm p.i.m.}$‡	0.025 (0.104)	0.022 (0.344)
$\langle I/\sigma(I)\rangle$	7.1 (5.1)	15.2 (2.0)
Monomers per asymmetric unit	1	1
Matthews coefficient (Å^3 Da^{-1})	1.64	1.62
Solvent content (%)	25.2	24.2

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations of symmetry-related reflections. ‡ $R_{\rm p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations of symmetry-related reflections.

2.3. Data collection and processing

In-house diffraction data were collected on an Onyx CCD detector using a Gemini PX Ultra diffractometer (Oxford Diffraction) with Cu $K\alpha$ radiation. A single cryocooled crystal was used and 291 frames were measured in 1° oscillation steps over four distinct ω scans with an 80 mm sample-to-detector distance and 6 min exposure per frame. A high-resolution data set was collected on a Pilatus 6M detector (Dectris) with synchrotron radiation on ESRF beamline ID29 (Grenoble, France) using 59% of the full X-ray beam intensity in 16-bunch mode operation. A single cryocooled crystal was used and 400 frames were measured in shutterless mode in 0.25° oscillation steps with a 160 mm sample-to-detector distance and 0.125 s exposure per frame. The data-collection strategy was calculated with *BEST* (Bourenkov & Popov, 2010). In-house diffraction data were integrated with *CrysAlis^{Pro}* (Oxford Diffraction) and synchrotron data were integrated with *XDS* (Kabsch, 2010). Both data sets were scaled with *SCALA* (Evans, 2006) and reduced with utilities from the *CCP*4 program suite (Winn *et al.*, 2011). Data-collection and processing statistics are summarized in Table 1.

2.4. Structure solution

The structure of boophilin D1 was solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007) as implemented in the *CCP*4 suite (Winn *et al.*, 2011) using the coordinates of the N-terminal domain of boophilin (residues Asn18–Cys71) from PDB entry 2ody (Macedo-Ribeiro *et al.*, 2008) as a search model with the diffraction data collected in-house.

3. Results and discussion

3.1. Expression and purification of recombinant boophilin D1

The N-terminal domain of boophilin, termed boophilin D1, was expressed in considerable amounts (8 mg purified protein per litre of culture) in *P. pastoris* as reported previously (Soares *et al.*, 2012). This

was a significant improvement over the rather low yields of the prokaryotic system previously used for the expression of boophilin (Macedo-Ribeiro *et al.*, 2008). The recombinant inhibitor was purified to homogeneity (Fig. 2) from the culture supernatant in a single chromatographic step on an immobilized-trypsin affinity column, concentrated and used for downstream crystallization experiments.

3.2. Crystallization of boophilin D1

Crystals of recombinant boophilin D1 could be grown by vapour-diffusion methods in sitting-drop geometry using a lithium sulfate/PEG 8000 mixture as precipitant (Fig. 3). A large number of small needle-like crystals appeared in the crystallization drops shortly after setup. Within a month, larger prismatic crystals formed from these needle clusters and were suitable for diffraction data collection. The orthorhombic crystals grew to maximum dimensions of 90 × 90 × 90 µm. They had an unusually low solvent content (Table 1) and diffracted to beyond 1.8 Å resolution using a sealed-tube X-ray source (Fig. 4) and to 0.87 Å at a third-generation synchrotron source. The data-collection and processing statistics are summarized in Table 1.

Figure 3
Crystals of boophilin D1 belonging to the orthorhombic space group $P2_12_12_1$. The scale bar is 200 µm in length.

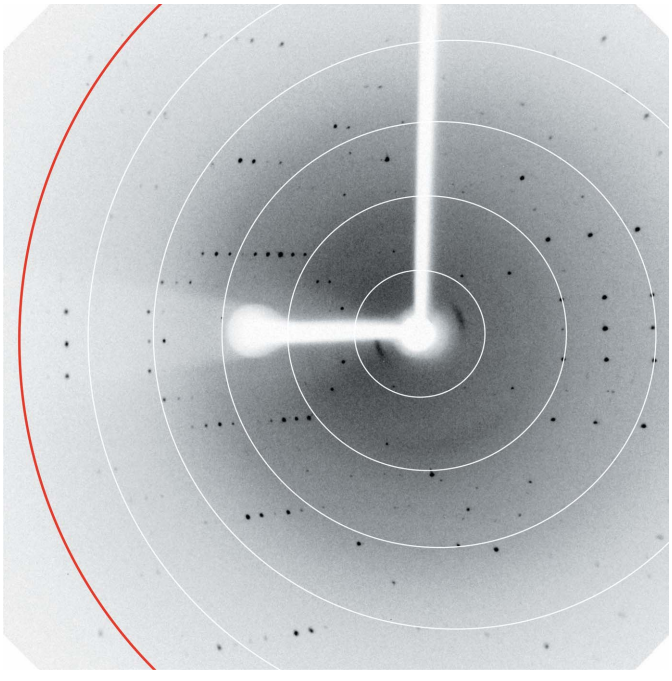


Figure 4
X-ray diffraction image of a crystal of boophilin D1 obtained using a sealed-tube X-ray source. The red line corresponds to a resolution limit of 1.94 Å.

3.3. Structure solution

The molecular coordinates of boophilin residues 18–71, as found in its complex with bovine α-thrombin (PDB entry 2ody; Macedo-Ribeiro *et al.*, 2008), were used as the search model to solve the structure of boophilin D1 by molecular replacement. A single molecular-replacement solution could be found with *Phaser* (McCoy *et al.*, 2007) in space group $P2_12_12_1$, with Z scores of 6.9 and 12.4 for the rotation and translation functions, respectively, corresponding to an overall log-likelihood gain of 223. The boophilin D1 N-terminal residues (16–17) and C-terminal residues (72–73) excluded from the molecular-replacement model could be readily located in the difference electron-density maps. The model was subjected to alternating cycles of manual building with *Coot* (Emsley *et al.*, 2010) and refinement with *PHENIX* (Adams *et al.*, 2010) against the low-resolution data set (data between 14.7 and 1.80 Å) until an R_{work} of 0.140 and an R_{free} of 0.169 were attained. Refinement of this model against the high-resolution data is currently under way.

We acknowledge the ESRF for provision of synchrotron-radiation facilities and the ESRF staff for assistance in using beamline ID29. This work was funded in part by Fundação para a Ciência e a Tecnologia (Portugal) through grant REEQ/564/BIO/2005 (EU-FEDER and POCI 2010), grant PTDC/BIA-PRO/70627/2006 to PJBP and postdoctoral fellowship SFR/BPD/46722/2008 to ACF. PJBP and AST were the recipients of a travel grant from GRICES (Portugal; Proc. 4.1.7 GRICES/CNPq) and CNPq (Brazil; 490574/2006-8).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Anderson, J. M., Sonenshine, D. E. & Valenzuela, J. G. (2008). *BMC Genomics*, **9**, 552.
Bond, C. S. & Schüttelkopf, A. W. (2009). *Acta Cryst.* D**65**, 510–512.
Bourenkov, G. P. & Popov, A. N. (2010). *Acta Cryst.* D**66**, 409–419.
Corral-Rodríguez, M. A., Macedo-Ribeiro, S., Barbosa Pereira, P. J. & Fuentes-Prior, P. (2009). *Insect Biochem. Mol. Biol.* **39**, 579–595.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Fuentes-Prior, P., Noeske-Jungblut, C., Donner, P., Schleuning, W. D., Huber, R. & Bode, W. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 11845–11850.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Koh, C. Y., Kumar, S., Kazimirova, M., Nuttall, P. A., Radhakrishnan, U. P., Kim, S., Jagadeeswaran, P., Imamura, T., Mizuguchi, J., Iwanaga, S., Swaminathan, K. & Kini, R. M. (2011). *PLoS One*, **6**, e26367.
Liao, M., Zhou, J., Gong, H., Boldbaatar, D., Shirafuji, R., Battur, B., Nishikawa, Y. & Fujisaki, K. (2009). *J. Insect Physiol.* **55**, 165–174.
Locht, A. van de, Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kröger, B., Höffken, W. & Bode, W. (1995). *EMBO J.* **14**, 5149–5157.
Locht, A. van de, Stubbs, M. T., Bode, W., Friedrich, T., Bollschweiler, C., Höffken, W. & Huber, R. (1996). *EMBO J.* **15**, 6011–6017.
Macedo-Ribeiro, S., Almeida, C., Calisto, B. M., Friedrich, T., Mentele, R., Stürzebecher, J., Fuentes-Prior, P. & Pereira, P. J. B. (2008). *PLoS One*, **3**, e1624.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Richardson, J. L., Kröger, B., Hoeffken, W., Sadler, J. E., Pereira, P., Huber, R., Bode, W. & Fuentes-Prior, P. (2000). *EMBO J.* **19**, 5650–5660.
Rydel, T. J., Ravichandran, K. G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C. & Fenton, J. W. (1990). *Science*, **249**, 277–280.
Soares, T. S., Watanabe, R. M. O., Tanaka-Azevedo, A. M., Torquato, R. J. S., Lu, S., Figueiredo, A. C., Pereira, P. J. B. & Tanaka, A. S. (2012). *Vet. Parasitol.*, doi:10.1016/j.vetpar.2012.01.027.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.