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

Crystallization and preliminary X-ray crystallographic analysis of a highly specific serpin from the beetle *Tenebrio molitor*

Sun Hee Park, Shunfu Piao, Hyun-Mi Kwon, Eun-Hye Kim, Bok Luel Lee and Nam-Chul Ha*

College of Pharmacy and Research Institute for Drug Development, Pusan National University, Jangjeon-dong, Geumjeong-gu, Busan 609-735, Republic of Korea

Correspondence e-mail: hnc@pusan.ac.kr

Received 3 December 2009
Accepted 10 December 2009

The Toll signalling pathway, which is crucial for innate immunity, is transduced in insect haemolymph *via* a proteolytic cascade consisting of three serine proteases. The proteolytic cascade is downregulated by a specific serine protease inhibitor (serpin). Recently, the serpin SPN48 was found to show an unusual specific reactivity towards the terminal serine protease, Spätzle-processing enzyme, in the beetle *Tenebrio molitor*. In this study, the mature form of SPN48 was overexpressed in *Escherichia coli* and purified. The purified SPN48 protein was crystallized using 14% polyethylene glycol 8000 and 0.1 *M* 2-(*N*-morpholino)ethanesulfonic acid pH 6.0 as the precipitant. The crystals diffracted X-rays to 2.1 Å resolution and were suitable for structure determination. The crystals belonged to space group $P2_1$. The crystal structure will provide information regarding how SPN48 achieves its unusual specificity for its target protease.

1. Introduction

Serine protease inhibitors (serpins), which exhibit multiple conformations, act as suicide substrates by binding covalently to their target proteases (Gettins, 2002). Serpins play essential roles in the down-regulation of extracellular signals exerted by serine protease cascades (Gooptu & Lomas, 2009). In particular, the serpins involved in mammalian blood coagulation and complement reactions have been well studied (Gooptu & Lomas, 2009). In insects, the Toll signalling pathway is triggered by the recognition of peptidoglycan and β-1,3-glucan, the major cell-wall components of bacteria and fungi, in bacterial and functional defence (Khush *et al.*, 2001). The initial recognition signals are amplified in haemolymph (insect blood) *via* a three-step proteolytic cascade (Kim *et al.*, 2008), leading to the processing of Spätzle, a ligand of the cell-surface receptor Toll (Lemaitre & Hoffmann, 2007). This pathway finally induces antimicrobial peptide (AMP) gene expression through activation of the NF-κB transcription factor; AMP then kills the invading pathogens (Khush *et al.*, 2001). To date, four *Drosophila* serpins related to innate immunity have been analyzed using genetic approaches (Levashina *et al.*, 1999; Ligoxygakis *et al.*, 2002). Only the crystal structure of serpin 1K from the hornworm moth *Manduca sexta* is available among the insect serpins (Li *et al.*, 1999), but the function of serpin 1K remains to be elucidated. Thus, no crystal structure of a serpin that is involved in insect immunity has been determined.

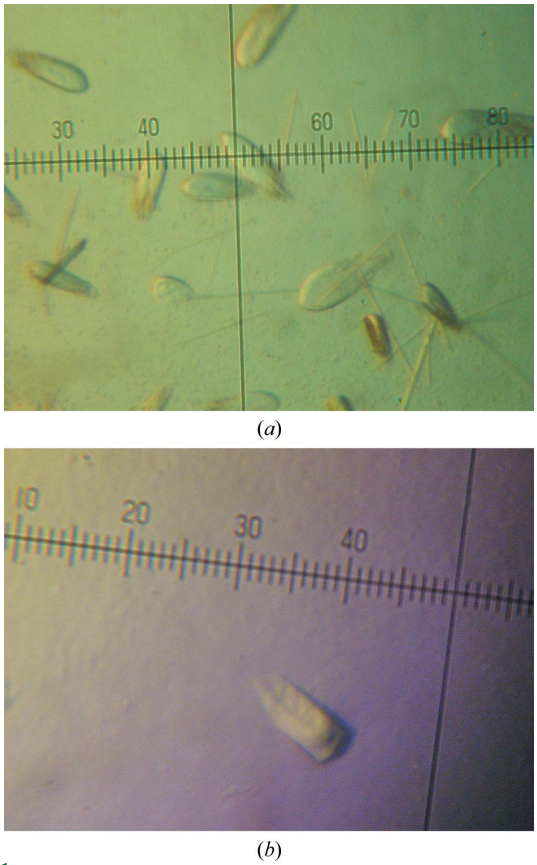
Our research group has identified and analyzed serine proteases involved in insect immunity (Piao, Kim *et al.*, 2005; Piao, Song *et al.*, 2005; Piao *et al.*, 2007; Kim *et al.*, 2008). Recently, we purified and identified three serpins from *Tenebrio molitor* haemolymph that regulate the Toll signalling pathway (Jiang *et al.*, 2009). The three serpins target the three serine proteases to generate three protease–serpin complexes (Jiang *et al.*, 2009). Of these serpins, SPN48 is highly specific for the terminal serine protease, Spätzle-processing enzyme (SPE). SPN48 shows unusual characteristics compared with the serpins found in mammalian blood. The reactive-centre loop (RCL)

of serpins is known to interact with the target protease and acts as a bait to capture the target serine protease (Johnson *et al.*, 2006). Thus, the RCL is usually a good substrate of the target substrate (Johnson *et al.*, 2006). However, the RCL of SPN48 does not contain the canonical sequence motif recognized by SPE. Although the target protease SPE has proteolytic activity after an arginine or lysine residue (Kim *et al.*, 2008), a glutamic acid residue is present at the cleavage site of the SPN48 RCL (Jiang *et al.*, 2009).

In order to gain insight into how this serpin tightly controls the innate immune response with its noncanonical RCL, we expressed, purified and crystallized SPN48 from the insect *T. molitor*.

2. Cloning, expression and purification

The mature form (residues 17–389) of *T. molitor* SPN48 (Jiang *et al.*, 2009) was inserted into the *Nde*I and *Hin*dIII sites of a modified pPROEX-HTa plasmid (Invitrogen, USA) in which the *Nco*I enzyme site was substituted by an *Nde*I site. The resulting plasmid, called pPROEX-HTa (*Nde*I)-SPN48, encodes a hexahistidine tag and TEV protease cleavage site at the N-terminus followed by SPN48 (Jiang *et al.*, 2009). This vector was transformed into *Escherichia coli* BL21-CodonPlus(DE3)-RIL (Stratagene, USA) for expression of SPN48. Cultures were grown in Luria–Bertani medium with 100 µg ml^{-1} ampicillin and 34 µg ml^{-1} chloramphenicol at 310 K until the mid-exponential growth phase and the gene was then induced by incubation with 0.5 m*M* isopropyl β-D-1-thiogalactopyranoside for 6 h at 303 K.



(*a*)

(*b*)

Figure 1
Crystals of SPN48 from *T. molitor*. (*a*) Crystals grown using the initial crystallization conditions; their approximate dimensions are 0.02 × 0.02 × 0.06 mm. Eight digits represent 0.1 mm. (*b*) A crystal grown using the optimal crystallization conditions; its approximate dimensions are 0.03 × 0.03 × 0.08 mm.

Table 1
Diffraction statistics.

Values in parentheses are for the highest resolution shell. A 0σ cutoff filter was applied during the scaling process.

X-ray source	Beamline 4A, PAL
Wavelength (Å)	1.000
Resolution (Å)	30–2.1 (2.14–2.10)
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 66.3, b = 43.0, c = 66.0, β = 105.1
Completeness (%)	93.7 (69.9)
$R_{\rm merge}$† (%)	5.3 (31.0)
Redundancy	3.2 (2.2)
Average $I/\sigma(I)$	18.0 (2.0)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the observed intensity and the mean intensity of related reflections, respectively.

The recombinant SPN48 protein was purified using a modification of a previously described procedure (Jiang *et al.*, 2009). Cells were collected by centrifugation (4520*g*, 30 min at 277 K) and resuspended in lysis buffer containing 20 m*M* Tris–HCl pH 8.0 and 150 m*M* NaCl. Resuspended cells were sonicated on ice with an output energy of 560 W for a total of 5 min. The supernatant after centrifugation (18 700*g*, 60 min at 277 K) was applied onto an Ni–NTA affinity chromatographic column (GE Healthcare, USA) equilibrated with lysis buffer. After washing with buffer supplemented with 20 m*M* imidazole, the column was eluted with lysis buffer supplemented with 200 m*M* imidazole. To remove the hexahistidine tag, the eluted fraction was treated with recombinant TEV protease (2 µg ml^{-1}) at 287 K for 12 h and the treated fraction was then concentrated using Centriprep (Amicon, USA). The sample was applied onto a gel-filtration chromatographic column (HiLoad 16/60 Superdex 200 prep-grade, GE Healthcare) equilibrated with lysis buffer. The elution peak was observed at 86 ml, indicating a monomeric state of SPN48. The highly pure peak fractions were then concentrated to 3 mg ml^{-1}. During purification, the presence of the protein was confirmed by SDS–PAGE.

3. Crystallization

SPN48 (373 amino acids) was crystallized using the sitting-drop vapour-diffusion method. Initial crystallization conditions were identified using a screening kit from Hampton Research (Riverside, USA; Jancarik *et al.*, 1991). Initial crystals of SPN48 were obtained in 50 µl reservoir solution [12% polyethylene glycol 20 000, 0.1 *M* 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.5] at 295 K (Fig. 1*a*). After optimization of the crystallization conditions, monoclinic crystals were obtained at 295 K by hanging-drop vapour diffusion, mixing equal volumes (1 µl) of protein (3 mg ml^{-1}) and reservoir solution (14% polyethylene glycol 8000, 0.1 *M* MES pH 6.0). The best crystals grew within 4 d to final dimensions of about 0.03 × 0.03 × 0.08 mm (Fig. 1*b*).

4. X-ray diffraction study

For X-ray data collection, a single crystal was briefly soaked in a cryoprotectant solution containing 12% polyethylene glycol 8000, 0.08 *M* MES pH 6.0 and 25% glycerol. The crystal was taken directly from the cryoprotectant solution and flash-cooled in a nitrogen stream. X-ray diffraction data from the crystal were collected using an ADSC Q310 CCD detector on the 4A beamline of Pohang Accelerator Laboratory (PAL), Republic of Korea. One set of 210 images (0–210°) was collected using a 1° oscillation width and a 40 s

exposure time at 1.00 Å wavelength. Indexing and integration of the images was performed using *DENZO* and the intensity data were scaled using *SCALEPACK* from the *HKL*-2000 package (Otwinowski & Minor, 1997).

The SPN48 crystal diffracted to 2.1 Å resolution and belonged to space group $P2_1$, with unit-cell parameters $a = 66.3$, $b = 43.0$, $c = 66.0$ Å, $\beta = 105.1^\circ$. The diffraction data set exhibited a completeness of 95.2% and an $R_{\rm merge}$ of 7.6% and analysis of the diffraction along the k axis revealed that the crystal b axis was a screw axis (Table 1). Assuming the presence of one SPN48 molecule per asymmetric unit, the Matthews parameter ($V_{\rm M}$; Matthews, 1968) was 2.17 Å^3 Da^{-1} and the solvent content was 43.4%. To solve the structure, molecular replacement was performed using the structure of antitrypsin (23% sequence similarity) as a search model with the program *MOLREP* (Collaborative Computational Project, Number 4, 1994). A full structural determination is currently under way. We hope that we will be able to further clarify the unusual characteristics of SPN48 in the biological context based on the crystal structure of SPN48.

We would like to thank the support staff of Beamline 4A at Pohang Accelerator Laboratory (Pohang, Republic of Korea). This work was supported for two years by a Pusan National University Research Grant.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Gettins, P. G. (2002). *Chem. Rev.* **102**, 4751–4804.
Gooptu, B. & Lomas, D. A. (2009). *Annu. Rev. Biochem.* **78**, 147–176.
Jancarik, J., Scott, W. G., Milligan, D. L., Koshland, D. E. Jr & Kim, S.-H. (1991). *J. Mol. Biol.* **221**, 31–34.
Jiang, R., Kim, E.-H., Gong, J.-H., Kwon, H.-M., Kim, C.-H., Ryu, K.-H., Park, J.-W., Kurokawa, K., Zhang, J., Gubb, D. & Lee, B.-L. (2009). *J. Biol. Chem.* **284**, 35652–35658.
Johnson, D. J., Langdown, J., Li, W., Luis, S. A., Baglin, T. P. & Huntington, J. A. (2006). *J. Biol. Chem.* **281**, 35478–35486.
Khush, R. S., Leulier, F. & Lemaitre, B. (2001). *Trends Immunol.* **22**, 260–264.
Kim, C.-H., Kim, S.-J., Kan, H., Kwon, H.-M., Roh, K.-B., Jiang, R., Yang, Y., Park, J.-W., Lee, H.-H., Ha, N.-C., Kang, H. J., Nonaka, M., Soderhall, K. & Lee, B. L. (2008). *J. Biol. Chem.* **283**, 7599–7607.
Lemaitre, B. & Hoffmann, J. (2007). *Annu. Rev. Immunol.* **25**, 697–743.
Levashina, E. A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J. A. & Reichhart, J. M. (1999). *Science*, **285**, 1917–1919.
Li, J., Wang, Z., Canagarajah, B., Jiang, H., Kanost, M. & Goldsmith, E. J. (1999). *Structure*, **7**, 103–109.
Ligoxygakis, P., Pelte, N., Ji, C., Leclerc, V., Duvic, B., Belvin, M., Jiang, H., Hoffmann, J. A. & Reichhart, J. M. (2002). *EMBO J.* **21**, 6330–6337.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Piao, S., Kim, D., Park, J. W., Lee, B. L. & Ha, N.-C. (2005). *Biochim. Biophys. Acta*, **1752**, 103–106.
Piao, S., Kim, S., Kim, J. H., Park, J. W., Lee, B. L. & Ha, N.-C. (2007). *J. Biol. Chem.* **282**, 10783–10791.
Piao, S., Song, Y.-L., Kim, J. H., Park, S. Y., Park, J. W., Lee, B. L., Oh, B.-H. & Ha, N.-C. (2005). *EMBO J.* **24**, 4404–4414.