Expression, purification and crystallization of a birnavirus-encoded protease, VP4, from blotched snakehead virus (BSNV)

Jaeyong Lee,[a] Anat R. Feldman,[a] Bernard Delmas[b] and Mark Paetzel[a]*

[a]Department of Molecular Biology and Biochemistry, Simon Fraser University, South Science Building, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada, and [b]Unité de Virologie et Immunologie Moléculaires, Institut National de la Recherche Agronomique, F-78350 Jouy-en-Josas, France

Correspondence e-mail: mpaetzel@sfu.ca

Blotched snakehead virus (BSNV) is a member of the *Birnaviridae* family that requires a virally encoded protease known as VP4 in order to process its polyprotein into viral capsid protein precursors (pVP2 and VP3). VP4 belongs to a family of serine proteases that utilize a serine/lysine catalytic dyad mechanism. A mutant construct of VP4 with a short C-terminal truncation was overexpressed in *Escherichia coli* and purified to homogeneity for crystallization. Using the sitting-drop vapour-diffusion method at room temperature, protein crystals with two distinct morphologies were observed. Cubic crystals grown in PEG 2000 MME and magnesium acetate at pH 8.5 belong to space group $I23$, with unit-cell parameters $a = b = c = 143.8$ Å. Trigonal crystals grown in ammonium sulfate and glycerol at pH 8.5 belong to space group $P321/P312$, with unit-cell parameters $a = b = 158.2$, $c = 126.4$ Å.

1. Introduction

Blotched snakehead virus (BSNV) was first identified in cell lines derived from the blotched snakehead (*Chana lucius*), a warm-water fish of south-east Asia (John & Richards, 1999). It is a member of the family *Birnaviridae*, which includes three genera, *Aquabirnavirus*, *Avibirnavirus* and *Entomobirnavirus*, that infect fish, birds and insects, respectively (Delmas *et al.*, 2005; http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/index.htm). The prototype members of this virus family, infectious pancreatic necrosis virus (IPNV) for *Aquabirnavirus* and infectious bursal disease virus (IBDV) for *Avibirnavirus*, are well known causative agents of economically important diseases affecting fish farming (Hill & Way, 1995) and the poultry industry (Muller *et al.*, 2003), respectively.

Viruses of the *Birnaviridae* family are characterized by their bisegmented double-stranded RNA genome (segments A and B) within a single-shell non-enveloped icosahedral particle of 60–70 nm diameter (Delmas *et al.*, 2005). The larger genome segment A encodes a polyprotein (NH2-pVP2-VP4-VP3-COOH) which is co-translationally processed by the VP4 (viral protein 4) protease activity, releasing pVP2 and VP3, the major capsid protein precursors. VP4 is a serine protease that has been shown to utilize a Ser-Lys catalytic dyad in which the lysine functions as the general base to activate the serine nucleophile (Birghan *et al.*, 2000; Lejal *et al.*, 2000; Petit *et al.*, 2000). The BSNV polyprotein (1069 amino acids) contains an additional polypeptide (X; 71 amino acids) which is sandwiched between the pVP2 and VP4 domains. Correspondingly, the BSNV VP4 domain (234 amino acids) is defined by its primary processing sites at the X/VP4 (557/558, PQA/ADL) and VP4/VP3 (791/792, CGA/ADE) interdomain junctions (Da Costa *et al.*, 2003). While the essential catalytic residues Ser692 and Lys729 are conserved in BSNV, there is a moderate level of sequence identity between BSNV and the other VP4 proteases (20% identity to IPNV VP4 and 23% identity to IBDV VP4).

Whereas the 'classical' Ser-His-Asp catalytic triad is found frequently in serine proteases of cellular and viral origin (Dodson & Wlodawer, 1998), proteases utilizing the Ser-Lys catalytic dyad are relatively rare and predominantly of prokaryotic origin (Paetzel & Dalbey, 1997). These proteases are categorized into the evolutionary clans SE, SF, SJ and SK in the MEROPS peptidase database (Rawlings *et al.*, 2006; http://merops.sanger.ac.uk/index.htm). Ser-Lys
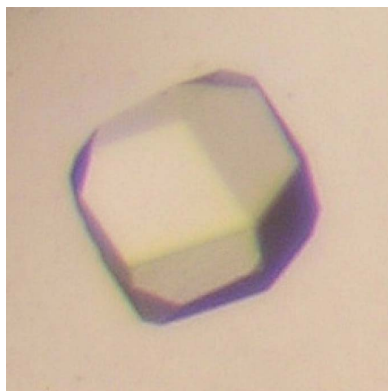
proteases whose crystal structures have been solved include signal peptidase (Paetzel *et al.*, 1998, 2002, 2004), UmuD′ (Peat *et al.*, 1996), LexA (Luo *et al.*, 2001), λ repressor (Bell *et al.*, 2000), D-Ala-D-Ala carboxypeptidase (Thunnissen *et al.*, 1995), Lon protease (Botos *et al.*, 2004, 2005; Im *et al.*, 2004) and the C-terminal processing peptidase (Liao *et al.*, 2000).

In this report, we describe the DNA manipulation, overexpression and crystallization of a C-terminal truncated version of BSNV VP4 (amino acids 558–773). The VP4 construct did not encode an affinity tag for purification and therefore a series of purification steps were necessary to obtain a homogeneous protein sample for crystallization. Upon screening for initial crystallization conditions, two different crystal forms were identified and optimized for X-ray diffraction. The structure of the *Birnaviridae* viral proteases has not been studied previously by crystallographic means and to our knowledge this is the first crystallization report of a Ser-Lys protease of viral origin. The BSNV VP4 structure is expected to reveal the configuration of a Ser-Lys catalytic dyad that may be structurally conserved in other genera of the *Birnaviridae*.

2. Materials and methods

2.1. DNA manipulation and expression

An expression construct encoding the entire open reading frame of BSNV VP4 (amino acids 558–791) in pET-28b+ *via* NcoI and BamHI sites was created as previously described (Da Costa *et al.*, 2003). A deletion mutant construct that removed 18 amino acids at the C-terminus was created by converting the codon for Pro774 of the original plasmid template to a termination codon. The required site-directed mutagenesis was carried out using the Quick Change kit (Stratagene) with forward primer 5′-GGAGTCAACAGTCACCCG-CTGAGAGAGCGTCGGACACCAG-3′ and reverse primer 5′-CT-GGTGTCCGACGCTCTCTCAGCGGGTGACTGTTGACTCC-3′. This new construct was confirmed by DNA sequencing. For overexpression, the new plasmid was transformed into *Escherichia coli* strain Tuner (DE3) and the transformants were selected on LB agar plates containing 0.05 mg ml^{-1} kanamycin. From a single colony, an overnight culture (100 ml) was grown at 310 K in LB media supplemented with 0.05 mg ml^{-1} kanamycin. The overnight culture was used to inoculate 1 l of the same media in a 1/100 dilution and the culture was induced with 0.5 mM IPTG at mid-log growth phase and grown for an additional 3 h at 310 K. The cells were harvested by centrifugation at 5000g for 15 min and stored frozen at 193 K.
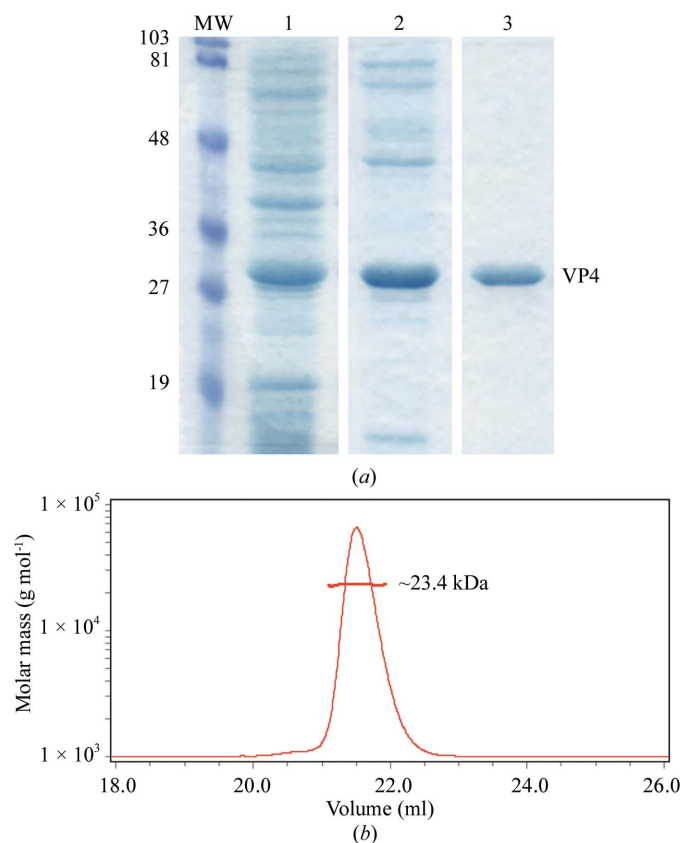
2.2. Protein purification

Frozen cell pellets from 2 l induced culture (4 g wet weight) were completely resuspended in 50 mM Tris–HCl buffer pH 8.0 containing 10% glycerol, 1 mM DTT, 7 mM magnesium acetate and 0.1% Triton X-100. The cells were lysed by addition of lysozyme (0.2 mg ml^{-1}) and benzonase (1 U ml^{-1}) and incubation overnight at 277 K. The lysate was then centrifuged at 27 000g for 30 min to remove cell debris. As a first purification step, the clear supernatant was subjected to ammonium sulfate fractionation at 30, 40, 50 and 60%(w/v) salt concentrations. After addition of the salt, the supernatant was incubated on ice for 30 min and then centrifuged at 27 000g for 30 min. Pellets were dissolved in a small volume of buffer A (20 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA, 10% glycerol, 1 mM DTT) and fractions containing VP4 were pooled and dialyzed against the same buffer. This sample was then applied onto a Q-Sepharose anion-exchange column containing 1.5 ml of matrix equilibrated in buffer A. The column was eluted with twice the column volume of buffer A

with increasing NaCl concentrations (0.05, 0.1, 0.2, 0.3, 0.35, 0.4, 0.5, 1.0 M). The elution fractions were analyzed by SDS–PAGE and the fractions containing the most VP4 were pooled and applied onto a size-exclusion column (HiPrep 16/60 Sephacryl S-100 HR) pre-equilibrated in buffer B (20 mM Tris–HCl pH 8.0, 100 mM NaCl, 10% glycerol, 1% β-mercaptoethanol) on an ÄKTA Prime system (Pharmacia). The column was run at a flow rate of 1 ml min^{-1} in buffer B. Analysis of the elution fractions by SDS–PAGE demonstrated that the protein had been purified to homogeneity (Fig. 1a). The gel-filtration fractions that contained the purified VP4 were pooled and concentrated using a Millipore centrifugal filter (5 kDa cutoff) and the protein concentration was measured by UV absorption at 280 nm (ε = 24 750 M^{-1} cm^{-1}). The protein concentration used for crystallization in buffer B was 24 mg ml^{-1}.

2.3. Multi-angle light-scattering analysis

Purified BSNV VP4 (10 mg ml^{-1}) was loaded onto a Shodex protein KW-803 size-exclusion column equilibrated with 20 mM Tris–HCl, 100 mM NaCl, 1% β-mercaptoethanol and connected in line with a Dawn 18-angle light-scattering detector coupled to an Optilab interferometric refractometer and a WyattQELS Quasi-Elastic Light-Scattering instrument (Wyatt Technologies). Data analysis was performed with the *ASTRA V* software (Wyatt Technologies) and

(a)

(b)

Figure 1
(*a*) SDS–PAGE of the BSNV VP4 protease purification steps. The left lane (MW) shows the molecular-weight standards (kDa). Lane 1 is the cell lysate. Lane 2 is the sample after ammonium sulfate fractionation and ion exchange. Lane 3 is the sample after size-exclusion chromatography and is the sample used for the crystallization experiments. (*b*) The multi-angle light-scattering (MALS) analysis of purified BSNV VP4 protease is represented by a molar mass (g mol^{-1}) *versus* volume (ml) plot overlaid with a size-exclusion elution profile. The measured molecular weight of 23.4 kDa shows that the purified BSNV VP4 is monodisperse and monomeric in solution at a concentration of 10 mg ml^{-1}.

Table 1
Data-collection statistics.

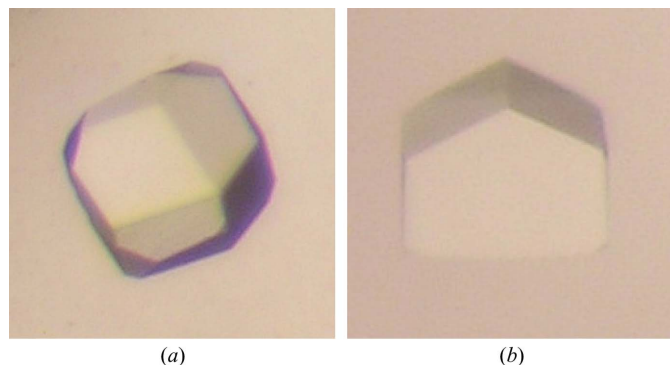| Crystal shape	Polyhedral	Hexagonal
No. of crystals used	1	1
Wavelength (Å)	1.5418	1.5418
Oscillation angle (°)	0.5	0.5
Frames collected	360	270
Exposure time (min)	4	4
Crystal-to-detector distance (mm)	210	210
Space group	$I23$	$P321/P312$
Unit-cell parameters		
$\quad a$ (Å)	143.8	158.2
$\quad b$ (Å)	143.8	158.2
$\quad c$ (Å)	143.8	126.4
Resolution range (Å)	38.43–2.52 (2.61–2.52)	34.25–2.52 (2.61–2.52)
Total reflections	327334	452061
Unique reflections	16866	61846
Redundancy	19.41 (16.83)	7.31 (7.24)
Completeness (%)	100.0 (99.8)	100.0 (100.0)
R_{merge}	0.136 (0.410)	0.132 (0.470)
$\langle I/\sigma(I)\rangle$	17.6 (7.2)	11.1 (4.1)

molecular weights were calculated using the Debye fit method (Andersson *et al.*, 2003).

2.4. Crystallization

Crystallization was carried out by the sitting-drop vapour-diffusion method at 294 K. 1 µl protein sample and 1 µl reservoir were mixed thoroughly and equilibrated against 1 ml of reservoir solution. The initial crystallization conditions were screened using commercial sparse-matrix (Jancarik & Kim, 1991) and PEG/Ion screen kits (Hampton Research). The protein was found to crystallize under several different conditions within 4–7 d. Two distinctly different conditions that generated crystals with different morphology were optimized for size and uniformity. Crystals with a polyhedral appearance (Fig. 2*a*) were obtained in 0.1 *M* Tris–HCl pH 8.5, 25% PEG 2000 MME (polyethylene glycol monomethyl ether) and 40 m*M* magnesium acetate. Hexagonal shaped crystals (Fig. 2*b*) were obtained in 0.1 *M* Tris–HCl pH 8.5, 1.65 *M* ammonium sulfate and 20% glycerol.

2.5. X-ray diffraction analysis

Both crystal forms were transferred to cryosolutions and incubated for 30 min before cryocooling. For polyhedral shaped crystals, the reservoir condition was supplemented with 20%(*v/v*) glycerol. For hexagonal shaped crystals, the cryocondition was 1.8 *M* ammonium sulfate, 0.1 *M* Tris–HCl pH 8.5 and 30%(*v/v*) glycerol. The crystals

(*a*)   (*b*)

Figure 2
Two different morphologies and symmetries of BSNV VP4 protease crystals. (*a*) Polyhedral shaped crystal with space group *I*23; (*b*) hexagonal shaped crystal with space group *P*321/*P*312.

were mounted in a 0.2–0.3 mm loop with a copper base (Hampton Research), flash-cooled in a 100 K nitrogen-gas stream and subjected to diffraction analysis. The X-ray diffraction data-collection system (Rigaku-MSC, USA) included a MicroMax-007 Microfocus X-ray generator running at 40 kV and 20 mA using a Cu anode, Osmic Confocal VariMax High Flux optics, an R-AXIS IV^{++} image plate and an X-stream2000 cryosystem set at 100 K. The data were collected with an oscillation angle of 0.5°, an exposure time of 4 min and a crystal-to-detector distance of 210 mm. The data collection, indexing, integration and scaling were carried out using the program *Crystal-Clear* (Pflugrath, 1999). The data-collection statistics are shown in Table 1.

3. Results and discussions

Initial attempts to obtain crystallization conditions using full-length BSNV VP4 were unsuccessful. However, since the terminal regions of the full-length VP4 protease are defined by the proteolytic cleavage sites and these regions are likely to be flexible for substrate presentation during self-cleavage, it was hypothesized that they may be hindering the protein from crystallizing. Therefore, recombinant BSNV VP4 protein was expressed as a truncated protein with 18 amino acids removed at the C-terminus. The polypeptide has a predicted weight of 23 441 Da and starts with a methionine preceding Ala558 and terminates at Arg773. Interestingly, the band corresponding to the overexpressed protein migrates between the molecular-weight markers at 36 and 27 kDa on SDS–PAGE (Fig. 1*a*). However, its weight measured by multi-angle light-scattering analysis at a concentration of 10 mg ml^{-1} is 23.4 kDa, corresponding almost exactly to the calculated molecular weight (Fig. 1*b*). The size of the purified protein was also verified by MALDI–TOF mass spectrometry (data not shown).

When the crystallization characteristics were closely analyzed using custom-designed screens, the purified protein was found to crystallize readily in a minimal reservoir condition of 25–30% PEG 2000 MME and 0.1 *M* Tris–HCl pH 8–9. These crystals, resembling multiple thin plates, were unsuitable for X-ray diffraction analysis. However, the presence of divalent salts above specific threshold concentrations such as 40 m*M* Mg^{2+} and 6 m*M* Ca^{2+} caused a substantial change in morphology. With Mg^{2+}, the drop initially displayed a heavy precipitation from which protein crystals nucleated within 3 d. Crystal growth coincided with the disappearance of the precipitate. These protein crystals had cubic or polyhedral appearance with an average length of 0.2 mm in each dimension. Analysis of the diffraction data indicates that the crystals are cubic and to belong to space group *I*23, with unit-cell parameters $a = b = c = 143.8$ Å. When analyzed by the program *MATTHEWS_COEF* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994), the data suggest a Matthews coefficient value of 2.6 Å^3 Da^{-1} and a solvent content of 53.1% with two molecules in the asymmetric unit (Matthews, 1968; see Table 1 for data-collection statistics).

The presence of Ca^{2+}, on the other hand, produced crystals that resembled hexagonal disks when fully formed but diffracted poorly. However, better diffracting crystals with similar morphology could be produced in conditions containing ammonium sulfate and glycerol at pH 8.5. These hexagonal-shaped crystals appeared within 4 d from clear drops and had maximum dimensions of 0.2 × 0.2 × 0.1 mm. Analysis of the diffraction data suggests that these crystals are trigonal and belong to space group *P*321 or *P*312, with unit-cell parameters $a = b = 158.2$, $c = 126.4$ Å. Calculation of the Matthews coefficient using the unit-cell parameters and molecular weight of

VP4 is consistent with seven, eight or nine molecules per asymmetric unit as equally plausible values, with V_M values of 2.8, 2.4 and 2.2 $\text{Å}^3\,\text{Da}^{-1}$, respectively (Kantardjieff & Rupp, 2003; http://ruppweb.dyndns.org/mattprob).

References

Andersson, M., Wittgren, B. & Wahlund, K.-G. (2003). *Anal. Chem.* **75**, 4279–4291.

Bell, C. E., Frescura, P., Hochschild, A. & Lewis, M. (2000). *Cell*, **101**, 801–811.

Birghan, C., Mundt, E. & Gorbalenya, A. E. (2000). *EMBO J.* **19**, 114–123.

Botos, I., Melnikov, E. E., Cherry, S., Kozlov, S., Makhovskaya, O. V., Tropea, J. E., Gustchina, A., Rotanova, T. V. & Wlodawer, A. (2005). *J. Mol. Biol.* **351**, 144–157.

Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulova, F., Dauter, Z., Maurizi, M. R., Rotanova, T. V., Wlodawer, A. & Gustchina, A. (2004). *J. Biol. Chem.* **279**, 8140–8148.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Da Costa, B., Soignier, S., Chevalier, C., Henry, C., Thory, C., Huet, J. C. & Delmas, B. (2003). *J. Virol.* **77**, 719–725.

Delmas, B., Kibenge, F. S. B., Leong, J. C., Mundt, E., Vakharia, V. N. & Wu, J. L. (2005). *Virus Taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses*, edited by H. V. Van Regenmortel, D. H. L. Bishop, M. H. Van Regenmortel & C. M. Fauquet, pp. 561–569. London: Academic Press.

Dodson, G. & Wlodawer, A. (1998). *Trends Biol. Sci.* **23**, 347–352.

Hill, B. J. & Way, K. (1995). *Annu. Rev. Fish Dis.* **5**, 55–77.

Im, Y. J., Na, Y., Kang, G. B., Rho, S. H., Kim, M. K., Lee, J. H., Chung, C. H. & Eom, S. H. (2004). *J. Biol. Chem.* **279**, 53451–53457.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

John, K. R. & Richards, R. H. (1999). *J. Gen. Virol.* **80**, 2061–2065.

Kantardjieff, K. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.

Lejal, N., Da Costa, B., Huet, J. C. & Delmas, B. (2000). *J. Gen. Virol.* **81**, 983–992.

Liao, D. I., Qian, J., Chisholm, D. A., Jordan, D. B. & Diner, B. A. (2000). *Nature Struct. Biol.* **7**, 749–753.

Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W. & Strynadka, N. C. (2001). *Cell*, **106**, 585–594.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Muller, H., Islam, M. R. & Raue, R. (2003). *Vet. Microbiol.* **97**, 153–165.

Paetzel, M. & Dalbey, R. E. (1997). *Trends Biol. Sci.* **22**, 28–31.

Paetzel, M., Dalbey, R. E. & Strynadka, N. C. (1998). *Nature (London)*, **396**, 186–190.

Paetzel, M., Dalbey, R. E. & Strynadka, N. C. J. (2002). *J. Biol. Chem.* **277**, 9512–9519.

Paetzel, M., Goodall, J. J., Kania, M., Dalbey, R. E. & Page, M. G. (2004). *J. Biol. Chem.* **279**, 30781–30790.

Peat, T. S., Frank, E. G., McDonald, J. P., Levine, A. S., Woodgate, R. & Hendrickson, W. A. (1996). *Nature (London)*, **380**, 727–730.

Petit, S., Lejal, N., Huet, J. C. & Delmas, B. (2000). *J. Virol.* **74**, 2057–2066.

Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.

Rawlings, N. D., Morton, F. R. & Barrett, A. J. (2006). *Nucleic Acid Res.* **34**, D270–D272.

Thunnissen, M. M. G. M., Fusetti, F., Deboer, B. & Dijkstra, B. W. (1995). *J. Mol. Biol.* **247**, 149–153.