

Ivy Yeuk Wah Chung and
Mark Paetzel*Department of Molecular Biology and
Biochemistry, Simon Fraser University,
South Science Building, 8888 University Drive,
Burnaby, British Columbia V5A 1S6, Canada

Correspondence e-mail: mpaetzel@sfu.ca

Received 6 November 2010

Accepted 22 November 2010

Expression, purification and crystallization of VP4 protease from *Tellina* virus 1

Tellina virus 1 is an aquabirnavirus that was isolated from the sand-dwelling marine bivalve mollusc *Tellina tenuis*. The self-encoded protease viral protein 4 (VP4) processes its own polyprotein to yield the individual proteins VP2 and VP3 that are required for viral assembly. VP4 protease utilizes a serine–lysine catalytic dyad in its mechanism. A full-length VP4 construct was overexpressed in *Escherichia coli* and purified to homogeneity using nickel-affinity chromatography. Ion-exchange and size-exclusion chromatographic steps were utilized to isolate a monomeric fraction of the protein. The purified monomeric VP4 was subjected to limited proteolysis to yield crystallizable protein. Crystal growth was performed using the hanging-drop vapour-diffusion method and was carried out at room temperature (~296 K). Hexagonal crystals grew in the presence of PEG 8000, ammonium sulfate and urea. These crystals diffracted to beyond 2.1 Å resolution and belonged to space group $P6_422$, with unit-cell parameters $a = 59.1$, $b = 59.1$, $c = 208.1$ Å, one molecule in the asymmetric unit and a solvent content of 42%.

1. Introduction

Viruses of the *Birnaviridae* family are categorized into three genera: *Avibirnavirus*, *Aquabirnavirus* and *Entomobirnavirus*, which infect birds, aquatic organisms and insects, respectively (Delmas *et al.*, 2005; Dobos, 1995). Well studied birnaviruses include the aquabirnavirus infectious pancreatic necrosis virus (IPNV), which infects salmonid fish, and the avibirnavirus infectious bursal disease virus (IBDV), which infects chickens. *Tellina* virus 1 (TV-1) is an aquabirnavirus that infects the sand-dwelling mollusc *Tellina tenuis* (Hill, 1976). Signs of TV-1 infection include a thinner chalkier shell and a pale yellow digestive gland.

The double-stranded RNA genome of birnaviruses is enclosed within an icosahedral capsid of 60–70 nm in diameter (Espinoza *et al.*, 2000; Dobos *et al.*, 1979). Its genomic DNA is divided into two segments (segments A and B). Most of its genes are encoded in segment A, which has the conserved arrangement NH₂-pVP2-VP4-VP3-COOH. One exception is the insertion of gene X between pVP2 and VP4 in TV-1 and blotched snakehead virus (BSNV; Da Costa *et al.*, 2003; Nobiron *et al.*, 2007). The polyprotein-encoded VP4 protease is responsible for the cleavage of the polyprotein into its substituent components (Petit *et al.*, 2000). Protein pVP2 is further processed to yield capsid protein VP2 and additional peptides (Coulibaly *et al.*, 2005). Protein VP3 associates with the genome to form a ribonucleoprotein complex (Hjalmarsson *et al.*, 1999).

The known cleavage sites within the TV-1 segment A polyprotein have the consensus Ala-X-Ala↓ (Nobiron *et al.*, 2007). Site-directed mutagenesis studies have identified Ser738 as the nucleophile and Lys777 as the general base in the TV-1 VP4 protease mechanism (Nobiron *et al.*, 2007). This serine–lysine catalytic dyad mechanism is shared among all birnavirus VP4s. All VP4s are grouped under the evolutionary clan SJ according to the MEROPS protease database (Rawlings *et al.*, 2010). The VP4s from IPNV, IBDV and BSNV



belong to protease family S50, whereas that from TV-1 belongs to family S69 owing to a high level of sequence divergence (Rawlings *et al.*, 2010).

The crystal structures of VP4 protease from the aquabirnaviruses IPNV and BSNV have previously been determined. The structure of BSNV VP4 was the first VP4 protease structure to be elucidated (Feldman *et al.*, 2006). An intermolecular acyl-enzyme intermediate was seen in an active-site mutant of IPNV VP4 (Lee *et al.*, 2007). The TV-1 VP4 protease shares 20% sequence identity with BSNV VP4 and only 12% sequence identity with IPNV VP4. Here, we report the DNA manipulation, protein overexpression, purification and crystallization of a chymotrypsin-treated full-length construct of TV-1 VP4 (residues 619–830) with a wild-type active site. The sequential use of nickel-affinity, cation-exchange and size-exclusion chromatography successfully selected a monomeric form of VP4. This monomeric form of VP4 failed to yield crystals in initial crystallization screens. Limited proteolysis treatment of TV-1 VP4 produced a slightly smaller and stable protein that was crystallizable. These crystals diffracted to beyond 2.1 Å resolution.

2. Material and methods

2.1. DNA manipulation and expression

The DNA coding region of *Tellina virus 1* VP4 (residues 619–830; Swiss-Prot accession No. Q2PBR5) was amplified using Vent DNA polymerase (New England Biolabs) by means of polymerase chain reaction (PCR) using the forward primer 5'-AGGCCCATATGGC-CGACAGGCCCATGATC-3' and the reverse primer 5'-CTTGTG-AAGGCGGCCGCTCATGCTGCGCCACGTTCTTCCGGAGG-AGAA3'. The PCR amplicon was digested overnight with restriction enzymes *NdeI* and *NotI* from Fermentas. The digested PCR product was then ligated into the respective restriction sites on vector pET28b+ using T4 DNA ligase (Fermentas). The use of pET28b+ allows the incorporation of a histidine tag at the amino-terminus to aid in protein purification. The ligation mix was transformed into NovaBlue Cells (Novagen) and spread on an LB plate supplemented with 0.05 mg ml⁻¹ kanamycin to select for transformants. DNA sequencing was performed to confirm the sequence of the insert. There are 21 extra amino acids (MGSSHHHHHHSSGLVPRGSHM) at the amino-terminus, which include the histidine-tag and linker regions, when this construct is expressed. This 233-amino-acid TV-1 VP4 construct has a theoretical molecular mass of 24 745 Da and a theoretical pI of 9.9.

2.2. Protein purification

The TV-1 VP4-containing plasmid was transformed into *Escherichia coli* strain Tuner (DE3) for protein overexpression. The method for overexpressing the selenomethionine-labelled (SeMet-labelled) TV-1 VP4 construct is described. 100 ml overnight culture grown in LB/kanamycin was spun down to yield a cell pellet for use as an inoculum for each litre of M9 minimal medium. 4 l of culture was used for each batch of protein. The culture was allowed to grow for 8 h at 310 K with shaking. 15 min prior to induction, a mixture of L-amino acids (100 mg lysine, phenylalanine and threonine and 50 mg isoleucine, leucine and valine) and 60 mg selenomethionine were added to each litre of culture. Each litre of culture was then induced using 0.5 ml 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated overnight on an orbital shaker at 298 K. Overexpression of the unlabelled TV-1 VP4 construct was performed in a similar manner except that the culture was grown in LB and was incubated for 4 h before induction. The culture was centrifuged at 9110g for

Table 1

Data-collection statistics for TV-1 VP4 protease crystals.

Values in parentheses are for the highest resolution shell.

	SeMet derivative (with urea)		Native (without urea)	
Wavelength (Å)	0.97893		1.5418	
Source	Canadian Light Source		Rotating anode	
Beamline	08ID-1		SFU	
Space group	P6 ₄ 22		P6 ₄ 22	
Unit-cell parameters (Å)	a = b = 59.1, c = 208.1		a = b = 59.0, c = 207.2	
Resolution limits (Å)	52.0–2.1 (2.2–2.1)	52.0–2.0 (2.2–2.0)	51.8–2.4 (2.5–2.4)	51.8–2.1 (2.2–2.1)
Total reflections	154167 (9938)	163831 (6280)	163860 (21601)	198558 (9523)
Unique reflections	13466 (1841)	16125 (1870)	8990 (1258)	13242 (1727)
R _{merge} †	0.107 (0.350)	0.109 (0.535)	0.062 (0.368)	0.071 (0.789)
Completeness (%)	99.8 (98.5)	96.5 (80.5)	100.0 (100.0)	99.0 (93.0)
Multiplicity	11.4 (5.4)	10.2 (3.4)	18.2 (17.2)	15 (5.5)
Mean I/σ(I)	14.5 (4.3)	12.5 (2.2)	39.5 (9.1)	28.2 (2.1)

† $\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections after rejections.

7 min to obtain the cell pellet. To facilitate cell lysis, the cell pellet was kept at 193 K for 15 min prior to resuspension in lysis buffer [50 mM Tris–HCl buffer pH 8.0, 10% glycerol, 1 mM dithiothreitol (DTT), 7 mM magnesium acetate, 0.1% Triton X-100, 1 U ml⁻¹ benzonase and 0.2 mg ml⁻¹ lysozyme]. This lysis step was carried out at 277 K overnight with gentle agitation and was followed by centrifugation at 28 964g for 20 min to remove cell debris. The clear supernatant was then applied onto a 5 ml nickel–nitrilotriacetic acid (Ni–NTA) metal-affinity column (Qiagen) equilibrated with standard buffer (20 mM Tris–HCl pH 8.0, 50 mM NaCl, 10% glycerol and 1 mM DTT). The column was then washed with five column volumes of the standard buffer and the histidine-tagged VP4 was eluted using a step gradient of 100, 300 and 600 mM imidazole in standard buffer. The VP4 started to elute at 100 mM imidazole. The fractions containing VP4 were pooled and purified by size-exclusion chromatography followed by ion exchange and then a second size-exclusion step (Fig. 1). The ion-exchange chromatography step was performed on a 5 ml SP-Sepharose FF cation-exchange column equilibrated with standard buffer. The protein was applied onto the column, washed with five column volumes of standard buffer and then eluted stepwise using 100, 300 and 500 mM NaCl in standard buffer. The VP4 started to elute at 100 mM NaCl. The size-exclusion chromatography was performed on a HiPrep 16/60 Sephacryl S-100 HR column equilibrated with crystallization buffer (20 mM Tris–HCl pH 8.0, 100 mM NaCl, 10% glycerol and 1% β-mercaptoethanol) pumped at a flow rate of 0.7 ml min⁻¹ using an ÄKTAprime system from Pharmacia. The column was calibrated using Pharmacia molecular-weight standards. The full-length TV-1 VP4 eluted at 160 ml. The purified VP4 was then subjected to limited proteolysis with α-chymotrypsin overnight at 277 K with a 1:1250 chymotrypsin:VP4 ratio. The VP4 was then applied onto the same size-exclusion column as mentioned above and run under the same conditions. The chymotrypsin-treated TV-1 VP4 eluted at 170 ml. All protein-concentration steps were performed using a Millipore centrifugal filter (10 kDa cutoff).

2.3. Crystallization

Crystals were grown at room temperature (~296 K) using the hanging-drop method. On a cover slip, 1 μl VP4 (40 mg ml⁻¹ in crystallization buffer) was mixed with 1 μl reservoir reagent and the drop was equilibrated over 1 ml reservoir reagent in a grease-sealed chamber. From the initial crystal screen, half-hexagon-shaped crystals appeared using the reservoir conditions 30% PEG 8000, 0.2 M

ammonium sulfate (Hampton Research Crystal Screen condition No. 30). The optimized reservoir conditions for the native crystals were 30% PEG 8000, 0.35 M ammonium sulfate. To obtain the optimal SeMet-labelled crystals, the drop was seeded with 1 µl of SeMet-labelled crystals from an older drop and 1 µl 0.2 M urea was added to the drop as an additive. The optimized reservoir conditions for the SeMet-labelled crystals were 21% PEG 8000, 0.55 M ammonium

sulfate. The pH of the drop was determined to be approximately 5 using ColorpHast pH-indicator strips from EM Reagents.

2.4. X-ray diffraction analysis

The cryosolution consisted of 70% reservoir solution and 30% glycerol. The crystal was first transferred into the cryosolution and then flash-cooled in liquid nitrogen prior to diffraction analysis. Data collection from the SeMet-labelled crystals was conducted on beamline 08ID-1 at the Canadian Light Source (CLS). The wavelength was set to 0.97893 Å. Data were collected with 0.5° oscillations and each image was exposed for 5 s. The distance between the detector and the crystal was 350 mm. The initial screening and data collection were performed with the Macromolecular Crystallography Data Collection (MXDC) graphical user interface. The indexing and integration of reflections were carried out using the program *iMOSFLM* and the data were scaled to 2.1 Å resolution using the program *SCALA* (Collaborative Computational Project, Number 4, 1994). The data set for the native crystal was collected at the Macromolecular X-ray diffraction Data Collection Facility in the

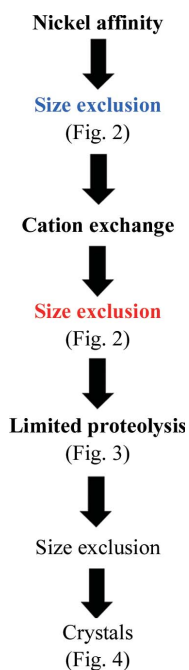


Figure 1
A flow diagram of TV-1 VP4 purification.

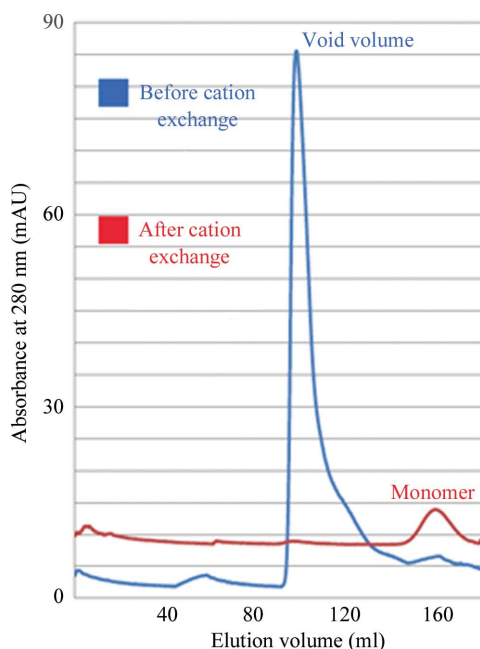


Figure 2
Size-exclusion chromatography of TV-1 VP4 protease before and after cation exchange. The elution fractions collected from the nickel-affinity column were subjected to size-exclusion chromatography (blue). The void volume fractions were collected and applied onto an SP-Sepharose FF cation-exchange column. The elution fractions from the cation-exchange column were again subjected to size-exclusion chromatography (red) and the fractions that eluted within a volume corresponding to a monomeric protein (~25 kDa) were collected.

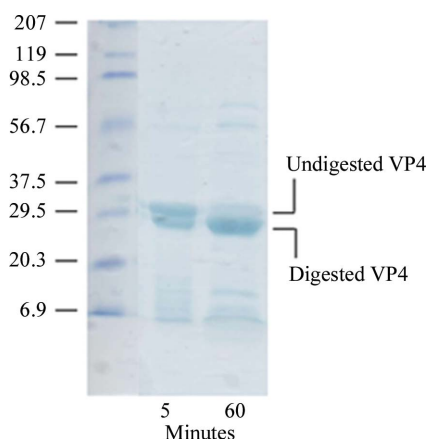


Figure 3
Limited proteolysis of monomeric TV-1 VP4 protease using chymotrypsin. SDS-PAGE analysis stained with Coomassie Blue stain (15% gel). Left lane, molecular-weight markers (kDa); middle lane, TV-1 VP4 + chymotrypsin after 5 min; right lane, TV-1 VP4 + chymotrypsin after 1 h.

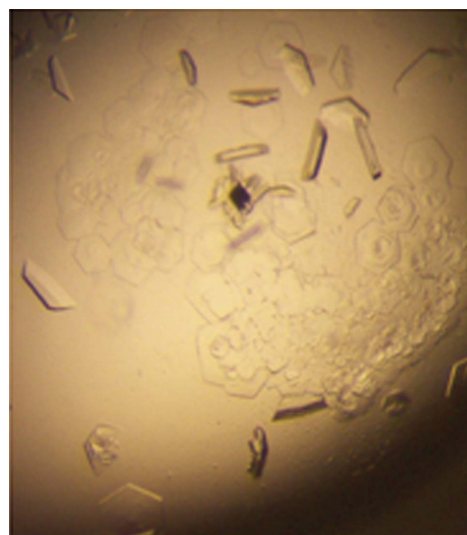


Figure 4
Native crystals of TV-1 VP4 protease. These crystals were grown in 30% PEG 8000, 0.35 M ammonium sulfate.

Molecular Biology and Biochemistry Department of Simon Fraser University, which consisted of a MicroMax-007 Microfocus X-ray generator, Osmic Confocal VariMax High Flux optics, an R-AXIS IV⁺⁺ image-plate detector and an X-stream 2000 cryosystem. The program *CrystalClear* (Rigaku) was used for data collection, indexing, integration and scaling. The crystal-to-detector distance was 200 mm and the crystal was exposed for 3 min at a wavelength of 1.5418 Å and with an oscillation angle of 0.5°. These native crystals diffracted to beyond 2.4 Å resolution. The crystal parameters and diffraction data statistics are given in Table 1.

3. Results and discussion

The full-length N-terminally His₆-tagged TV-1 VP4 construct was expressed in the soluble fraction and was initially purified on a nickel-affinity column, which removed most of the bacterial contaminants. However, size-exclusion chromatographic analysis revealed that TV-1 VP4 eluted in the void volume, where proteins greater than 100 kDa would be expected to elute (Fig. 2). Since the theoretical molecular mass of TV-1 VP4 is approximately 25 kDa, this suggested that TV-1 VP4 was either aggregated or in a homotetrameric state or larger. Crystallization screens using this protein did not produce any hits. To increase the purity of the sample, the protein eluted from the size-exclusion column was applied onto a cation-exchange column. The sodium chloride-eluted TV-1 VP4 was pooled and subjected to size-exclusion chromatography for a second time, which revealed that TV-1 VP4 now eluted within a volume consistent with a monomeric nature (Fig. 2). Initial attempts to crystallize the full-length monomeric TV-1 VP4 failed; thus, limited proteolysis was performed to remove flexible regions that might possibly hinder crystal packing. A variety of proteases and conditions were investigated. Chymotrypsin-treated TV-1 VP4 gave a single and slightly smaller major band on the SDS-PAGE gel after 1 h incubation (Fig. 3); therefore, chymotrypsin was chosen for proteolytic treatment of this protein. The TV-1 VP4 was purified by size-exclusion chromatography again, concentrated and set up for crystallization screens. The chymotrypsin-treated monomeric TV-1 VP4 native protein yielded hexagonal crystals with an optimized reservoir condition of 30% PEG 8000, 0.35 M ammonium sulfate (Fig. 4). The reproducibility of these crystals was low, with a success rate of 5–10%. The success rate was increased by seeding with crystals from an older drop. These crystals appeared approximately one week after crystal plating and continued to grow over a month. The native crystal diffracted to beyond 2.4 Å resolution using a rotating-anode X-ray source. To obtain phase information, a batch of SeMet-labelled TV-1 VP4 was purified. The optimized reservoir condition for the TV-1 VP4 SeMet-labelled protein was 21% PEG 8000, 0.55 M ammonium sulfate. The SeMet-labelled VP4 crystallization drops were seeded with native crystals. The resulting SeMet-labelled VP4 crystals were more reproducible than the native crystals, with a success rate of ~80–90%. To increase the size of these hexagonal crystals, the optimized condition was screened against Additive Screen from Hampton Research. The SeMet-labelled crystals grew larger in the presence of urea. These crystals appeared 2–3 d after crystal plating and had maximum dimensions of approximately 0.1 × 0.2 × 0.05 mm. However, these crystals

dissolved in less than a week and required continuous crystal plating to provide seeds for further crystal growth. The SeMet-labelled crystals were plated 2 d prior to data collection on CLS beamline 08ID-1 and diffracted to beyond 2.1 Å resolution. These crystals belonged to space group *P6₄22*. The ambiguity between *P6₄22* and its enantiomorph *P6₂22* was resolved during the phasing and initial refinement steps. The unit cell has parameters $a = b = 59.1$, $c = 208.1$ Å with one molecule in the asymmetric unit, a Matthews coefficient of 2.12 Å³ Da⁻¹ and a solvent content of 42.1% (based on the theoretical mass of the uncleaved protein). The diffraction resolution limit for the native data set was 2.1 Å based on a data cutoff where the mean $I/\sigma(I)$ value was 2.1, but this resulted in an R_{merge} in the highest resolution bin of 0.789; therefore, we chose to cut the data at 2.4 Å resolution ($R_{\text{merge}} = 0.368$). The diffraction resolution limit for the SeMet-labelled crystal data set was 2.0 Å based on a data cutoff where the mean $I/\sigma(I)$ value was 2.2, but this resulted in an R_{merge} in the highest resolution bin of 0.535; therefore, we chose to cut the data at 2.1 Å resolution ($R_{\text{merge}} = 0.350$). Details of the crystal parameters and diffraction data statistics are given in Table 1.

This work was supported in part by the Canadian Institute of Health Research (to MP), the National Science and Engineering Research Council of Canada (to MP), the Michael Smith Foundation for Health Research (to MP) and the Canadian Foundation of Innovation (to MP). We thank the staff at the macromolecular beamline 08ID-1, Canadian Light Source, Saskatoon, Canada for their technical assistance with data collection. The Canadian Light Source is supported by NSERC, NRC, CIHR and the University of Saskatchewan. We thank Bernard Delmas for providing the initial TV-1 cDNA.

References

- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Coulibaly, F., Chevalier, C., Gutsche, I., Pous, J., Navaza, J., Bressanelli, S., Delmas, B. & Rey, F. A. (2005). *Cell*, **120**, 761–772.
- 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: The Eighth Report of the International Committee on Taxonomy of Viruses*, edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball, p. 561. Amsterdam: Elsevier.
- Dobos, P. (1995). *Virology*, **208**, 19–25.
- Dobos, P., Hill, B. J., Hallett, R., Kells, D. T., Becht, H. & Teninges, D. (1979). *J. Virol.* **32**, 593–605.
- Espinoza, J. C., Hjalmarsson, A., Everitt, E. & Kuznar, J. (2000). *Arch. Virol.* **145**, 739–748.
- Feldman, A. R., Lee, J., Delmas, B. & Paetzel, M. (2006). *J. Mol. Biol.* **358**, 1378–1389.
- Hill, B. J. (1976). *Wildlife Diseases*, edited by L. A. Page, pp. 445–452. New York, London: Plenum Press.
- Hjalmarsson, A., Carlemalm, E. & Everitt, E. (1999). *J. Virol.* **73**, 3484–3490.
- Lee, J., Feldman, A. R., Delmas, B. & Paetzel, M. (2007). *J. Biol. Chem.* **282**, 24928–24937.
- Nobiron, I., Galloux, M., Henry, C., Torhy, C., Boudinot, P., Lejal, N., Da Costa, B. & Delmas, B. (2007). *Virology*, **371**, 350–361.
- Petit, S., Lejal, N., Huet, J.-C. & Delmas, B. (2000). *J. Virol.* **74**, 2057–2066.
- Rawlings, N. D., Barrett, A. J. & Bateman, A. (2010). *Nucleic Acids Res.* **38**, D227–D233.