

# Crystal Structure of the VP4 Protease from Infectious Pancreatic Necrosis Virus Reveals the Acyl-Enzyme Complex for an Intermolecular Self-cleavage Reaction\*

Received for publication, February 21, 2007, and in revised form, April 25, 2007. Published, JBC Papers in Press, June 6, 2007, DOI 10.1074/jbc.M701551200

Jaeyong Lee<sup>‡</sup>, Anat R. Feldman<sup>‡</sup>, Bernard Delmas<sup>§1</sup>, and Mark Paetzel<sup>‡2</sup>

From the <sup>‡</sup>Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada and <sup>§</sup>Unité de Virologie et Immunologie Moléculaires, Institut National de la Recherche Agronomique, F-78350 Jouy-en-Josas, France

Infectious pancreatic necrosis virus (IPNV), an aquatic birnavirus that infects salmonid fish, encodes a large polyprotein ( $\text{NH}_2$ -pVP2-VP4-VP3-COOH) that is processed through the proteolytic activity of its own protease, VP4, to release the proteins pVP2 and VP3. pVP2 is further processed to give rise to the capsid protein VP2 and three peptides that are incorporated into the virion. Reported here are two crystal structures of the IPNV VP4 protease solved from two different crystal symmetries. The electron density at the active site in the triclinic crystal form, refined to 2.2-Å resolution, reveals the acyl-enzyme complex formed with an internal VP4 cleavage site. The complex was generated using a truncated enzyme in which the general base lysine was substituted. Inside the complex, the nucleophilic Ser<sup>633</sup>Oγ forms an ester bond with the main-chain carbonyl of the C-terminal residue, Ala<sup>716</sup>, of a neighboring VP4. The structure of this substrate-VP4 complex allows us to identify the S1, S3, S5, and S6 substrate binding pockets as well as other substrate-VP4 interactions and therefore provides structural insights into the substrate specificity of this enzyme. The structure from the hexagonal crystal form, refined to 2.3-Å resolution, reveals the free-binding site of the protease. Three-dimensional alignment with the VP4 of blotched snakehead virus, another birnavirus, shows that the overall structure of VP4 is conserved despite a low level of sequence identity (~19%). The structure determinations of IPNV VP4, the first of an acyl-enzyme complex for a Ser/Lys dyad protease, provide insights into the catalytic mechanism and substrate recognition of this type of protease.

Birnaviruses are non-enveloped viruses of ~60–70 nm in diameter, which replicate in the cytoplasm of their host cells (1) and are characterized by their two double-stranded RNA genomic segments (A and B) (2). Segment A displays two overlapping reading frames, the larger one encoding a polyprotein ( $\text{NH}_2$ -pVP2-VP4-VP3-COOH; Fig. 1). The polyprotein is processed through the proteolytic activity of VP4 to generate the proteins pVP2 and VP3 as well as VP4 (3). During virus assembly, pVP2 is further processed by VP4 to generate the capsid protein VP2 and several structural peptides. The crystal structure of the infectious bursal disease virus particle (an avian birnavirus) solved at a resolution of 7 Å revealed the VP2 protein incorporated into an icosahedral viral capsid with triangulation  $T = 13$  (4).

The Birnaviridae virus family includes three genera, *Aquabirnavirus*, *Avibirnavirus*, and *Entomobirnavirus*, which infect fish, birds, and insects, respectively (5, 6). Infectious pancreatic necrosis virus (IPNV)<sup>3</sup> is a well known pathogen in salmonid fish. It is responsible for infectious pancreatic necrosis, a disease characterized by severe damage to the internal organs and tissues (7). Because of the high mortality rate it induces and its widespread distribution, this virus is a major economical and ecological threat to the aquaculture and sea farming industry worldwide (8).

The IPNV polyprotein is 972 residues in length (6) (Fig. 1). The IPNV VP4 protease cleaves its own N and C termini in the polyprotein after the amino acids 508 and 734, respectively. Three cleavage sites within pVP2 are located after amino acid residues 442, 486, and 495. The resulting products are VP2 (aa 1–442 of the polypeptide), p1 (aa 443–486), p2 (aa 487–495), p3 (aa 496–508), VP4 (aa 509–734), and VP3 (aa 735–972). The peptides p1, p2, and p3 are present in the virions but may also be further processed into additional cleaved products (9). There is a reported internal cleavage site near the C-terminal end of VP4 itself, which differs from the consensus cleavage sequence (3). The main cleavage sites and those involved in the processing of pVP2 have been described by the consensus (Ser/Thr)-X-Ala↓(Ser/Ala)-Gly motif (Fig. 1).

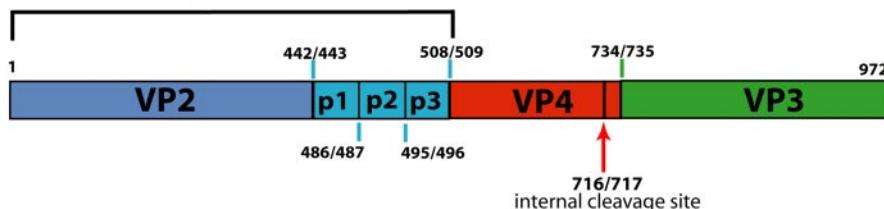
\* This work was supported in part by a Canadian Institute of Health Research operating grant, a National Science and Engineering Research Council of Canada discovery grant, a Michael Smith Foundation for Health Research Scholar award, and a Canadian Foundation of Innovation grant (to M. P.) and by a Canadian Cystic Fibrosis Foundation postdoctoral fellowship award (to A. R. F.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2PNL and 2PNM) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

<sup>1</sup> Supported by the ACI "Microbiologie" from the French Ministère de la Recherche et de la Technologie (MRT).

<sup>2</sup> To whom correspondence should be addressed: Dept. of Molecular Biology and Biochemistry, Simon Fraser University, South Science Bldg., 8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada. Tel.: 604-291-4230; Fax: 604-291-5583; E-mail: mpaetzel@sfu.ca.

<sup>3</sup> The abbreviations used are: IPNV, infectious pancreatic necrosis virus; BSNV, blotched snakehead virus; VP4, viral protein (protease) 4 in the polyprotein (pVP2-VP4-VP3) coded for in the birnavirus; VP4<sup>tri</sup>, the IPNV VP4 protease construct (VP4\_514–716,K674A) that produced the triclinic crystal form; VP4<sup>hex</sup>, the IPNV VP4 protease construct (VP4\_524–716,K674A) that produced the hexagonal crystal form; aa, amino acid(s); PEG, polyethylene glycol.

A. **pVP2**

B. 524 **VP4<sup>hex</sup>** 716 (hexagonal crystals, P6<sub>1</sub>22)  
514 **VP4<sup>tri</sup>** 716 (triclinic crystals, P1)

## C.

## IPNV cleavage sites

VP2/p1	(442/443)	P6	P5	P3	P1	Scissile bond	P2'
p1/p2	(486/487)	SSD	LPT	SKA		WGWRDI	
p2/p3	(495/496)	I	G	D	LTKTNA	AGGRYH	
p3/VP4	(508/509)	A	G	R	YHSMA	AGGRYK	
internal (716/717)		Y	K	D	VLESWA	SGGPDG	
VP4/VP3	(734/735)	T	G	T	LPVQRA	KGSNKR	
		L	E	I	MASNA	SGMDEE	

FIGURE 1. A, the IPNV gene segment A polyprotein is made up of the capsid protein VP2 (blue), the propeptides of VP2 (p1, p2, p3) (cyan), VP4 protease (red), and the structural protein VP3 (green). The P1/P1' residues for the cleavage sites are labeled including the internal cleavage site within the C terminus of VP4. B, schematic of the VP4<sup>hex</sup> (VP4\_524–716, K674A) and VP4<sup>tri</sup> (VP4\_514–716, K674A) constructs of IPNV VP4 that gave hexagonal crystals (P6<sub>1</sub>22) and triclinic crystals (P1), respectively. C, alignment of the IPNV polyprotein cleavage sites.

Site-directed mutagenesis studies on IPNV VP4 protease show that the conserved catalytic residues Ser<sup>633</sup> and Lys<sup>674</sup> are essential for polyprotein processing (3). Similar results are found for the VP4 protease in the blotched snakehead virus (BSNV; Ser<sup>692</sup> and Lys<sup>729</sup>) (10) and infectious bursal disease virus (IBVD; Ser<sup>652</sup> and Lys<sup>692</sup>) (11, 12). The VP4 proteases of the Birnaviridae family, therefore, are proposed to utilize a Ser/Lys catalytic dyad mechanism to catalyze the processing of the polyprotein (3, 11, 12). This property distinguishes VP4 proteases from other viral proteases.

The recently solved crystal structure of VP4 protease from BSNV supports previous findings suggesting that VP4 is a Ser/Lys protease. It also revealed that, despite almost no sequence similarity, VP4 is structurally similar to the Lon proteases (13–15). Evolutionarily, VP4 protease (S50.001) belongs to the clan SJ and family S50 in the MEROPS protease data bank (16). Other characterized proteases with a Ser/Lys catalytic dyad mechanism include bacterial signal peptidase (17–19), LexA (20), phage repressor cI $\lambda$  (21), UmuD' (22), and bacterial Lon protease (15).

Here we report x-ray crystal structures of the IPNV VP4 protease in the unbound form (free binding site) and the product bound form (acyl-enzyme complex). The acyl-enzyme complex corresponds to an intermolecular (*trans*-cleavage) event at an internal cleavage site within VP4. The structure determinations allow for the identification of many of the molecular interactions important for substrate recognition.

## EXPERIMENTAL PROCEDURES

**Cloning and Mutagenesis**—In an attempt to generate crystals suitable for structure determination, many different truncated

and mutant forms of IPNV VP4 protease have been generated (23). Two of these IPNV VP4 protease constructs are described here. A pET-28b plasmid encoding a truncated form of the IPNV VP4 gene (residues 514–734) was used as the initial DNA template (3). The codon for Lys<sup>674</sup> was changed to a codon for alanine using site-directed mutagenesis by PCR. The codon for Lys<sup>717</sup> was mutated to a stop codon. The expression product from this plasmid VP4\_514–716, K674A (VP4<sup>tri</sup>) is 204 residues long (including initial methionine) with a molecular mass of 21.7 kDa and a theoretical isoelectric point of 4.5.

The plasmid for VP4<sup>tri</sup> was then used as a template in a PCR reaction to amplify a DNA segment encoding VP4\_524–716, K674A (VP4<sup>hex</sup>). It was cloned into pET-24a (Novagen) via the NdeI and SalI restriction sites. The resulting expressed protein is 194 residues long (including

the initial methionine) with a molecular mass of 20.5 kDa and a theoretical isoelectric point of 4.3. Sequencing was performed to confirm each DNA construct.

**Expression and Purification of Selenomethionine-incorporated Proteins**—Both VP4 plasmid constructs were transformed into *Escherichia coli* Tuner (DE3) cells for overexpression of the VP4 proteases. An overnight culture in M9 minimal media (10 ml) was used to inoculate 1 liter of M9 minimal media. All growth media were supplemented with 0.05 mg/ml kanamycin for selection. Cultures were grown for 8 h at 37 °C with shaking, and a mixture of the following L-amino acids was added directly: 100 mg of lysine, phenylalanine, and threonine; 50 mg of isoleucine, leucine, and valine; 60 mg of selenomethionine. After 15 min, the protein expression was induced by adding 0.5 ml of 1 M isopropyl 1-thio- $\beta$ -D-galactopyranoside. The cells were grown for another 3 h and then harvested by centrifugation at 4500 rpm for 15 min. The harvested cell pellets were frozen and then lysed by resuspending the frozen pellet in the lysis buffer (50 mM Tris-HCl (pH 8.0), 10% glycerol, 1 mM dithiothreitol, 7 mM magnesium acetate, 0.2 mg/ml lysozyme, 1 unit/ml benzonase, 0.1% Triton X-100) and incubating overnight at 4 °C with gentle agitation. The lysate was centrifuged at 14,000 rpm for 30 min to remove cell debris and obtain a clear supernatant.

The clear supernatant was then subjected to ammonium sulfate fractionation at 30, 40, 50, and 60% (w/v) ammonium sulfate concentrations. The ammonium sulfate precipitates were pelleted by centrifugation at 14,000 rpm for 30 min and then dissolved in 1 ml of buffer A (20 mM Tris-HCl (pH 8), 10% glycerol, 1 mM EDTA, 50 mM NaCl, 1 mM dithiothreitol). The

fractions containing VP4 were pooled and dialyzed against buffer A.

The dialyzed protein was then applied to a Q-Sepharose FF anion-exchange column (2 ml bed volume) equilibrated with buffer A. After washing with 2 ml of buffer A, the protein was eluted with a stepwise NaCl gradient (0.1 M increments up to 0.5 M) in buffer A.

The ion-exchange elution fractions containing the VP4 were pooled and subjected to gel-filtration chromatography using a Hiprep 16/60 Sephadryl S-100 HR column equilibrated in buffer B (20 mM Tris-HCl (pH 8), 10% glycerol, 100 mM NaCl, 1% (v/v)  $\beta$ -mercaptoethanol). The column was run at a flow rate of 1 ml/min, and 3-ml fractions were collected. The fractions containing the pure VP4 were then pooled and concentrated using a Millipore centrifugal filter (5000-Da MW cutoff). The proteins VP4<sup>tri</sup> and VP4<sup>hex</sup> were concentrated to 50 and 30 mg/ml, respectively, for crystallization.

**Crystallization**—The crystals used for data collection were grown by the sitting-drop vapor diffusion method. The crystallization drops were prepared by mixing 1  $\mu$ l of protein with 1  $\mu$ l of reservoir solution and then equilibrating the drop against 1 ml of reservoir solution. The VP4<sup>tri</sup> protein produced triclinic crystals (P1) with unit cell dimensions of  $41.7 \times 69.3 \times 191.4 \text{ \AA}$  ( $\alpha = 93.1^\circ$ ,  $\beta = 95.0^\circ$ ,  $\gamma = 97.6^\circ$ ) with 10 molecules in the asymmetric unit and a Matthews coefficient of  $2.5 \text{ \AA}^3 \text{ Da}^{-1}$  (51.3% solvent). The VP4<sup>hex</sup> protein produced hexagonal crystals (P6<sub>1</sub>22) with unit cell dimensions of  $77.1 \times 77.1 \times 136.4 \text{ \AA}$  with 1 molecule in the asymmetric unit and a Matthews coefficient of  $2.9 \text{ \AA}^3 \text{ Da}^{-1}$  (57.3% solvent).

For the triclinic crystals, the optimized crystallization reservoir condition was 0.1 M Tris-HCl (pH 8.5), 35% PEG 4000, 0.4 M Li<sub>2</sub>SO<sub>4</sub>, 0.4 M guanidine-HCl. The crystallization was conducted at 4 °C. The cryosolvent condition was 0.1 M Tris-HCl (pH 8.5), 35% PEG 4000, 0.45 M guanidine-HCl, 10% glycerol.

For the hexagonal crystals, the optimized crystallization reservoir condition was 0.1 M Tris-HCl (pH 8.5), 22% PEG 2000 MME, 0.45 M calcium acetate. The crystallization was conducted at 18 °C. The cryosolvent condition was 0.1 M Tris-HCl (pH 8.5), 22.5% PEG 2000 monomethyl ether, 0.3 M calcium acetate, 15% glycerol. Both crystals were incubated under their respective cryosolvent conditions for ~20 min before being flash-cryocooled in liquid nitrogen.

**Data Collection**—Diffraction data were collected on seleno-methionine-incorporated crystals at beamline X8C of the Brookhaven National Laboratory, National Synchrotron Light Source, using a Q4 CCD detector and a Nonius CAD4 Goniometer. The crystal-to-detector distance was 200 mm for the P1 crystal and 175 mm for the P6<sub>1</sub>22 crystal. Data were collected with 1° oscillations, and each image was exposed for 30 s with the hexagonal crystal and 3 s for the triclinic crystals. The P6<sub>1</sub>22 crystal diffraction data were processed with the program HKL2000 (24), and the P1 crystal diffraction data were processed with MOSFLM (25). See Table 1 for data collection statistics.

**Structure Determination and Refinement**—The hexagonal crystal form (VP4<sup>hex</sup>) was solved by single-wavelength anomalous dispersion (SAD) using a data set collected at the peak wavelength (0.9792  $\text{\AA}$ ) and the program autoSHARP (26). Values for  $f'$  and  $f''$  were found to be -7.26 and 4.36, respectively. All four of the

**TABLE 1**  
Data collection, phasing, and refinement statistics

	VP4 <sup>hex</sup>	VP4 <sup>tri</sup>
<b>Crystal parameters</b>		
Space group	P6 <sub>1</sub> 22	P1
$a, b, c$ ( $\text{\AA}$ )	$77.0 \times 77.0 \times 136.4$	$41.7 \times 69.3 \times 191.4$
$\alpha, \beta, \gamma$ (°)	$90.0 \times 90.0 \times 120.0$	$93.1 \times 95.0 \times 97.6$
<b>Data collection statistics</b>		
Wavelength	0.9792	0.9794
Resolution ( $\text{\AA}$ )	47.7–2.3 (2.4–2.3)	50.0–2.2 (2.3–2.2)
Total reflections	129,151 (18805)	403,600 (33041)
Unique reflections	11,248 (1584)	103,935 (9718)
$R_{\text{merge}}^a$	0.076 (0.186)	0.082 (0.280)
$R_{\text{rim}}^b$	0.077 (0.216)	0.079 (0.223)
$R_{\text{pim}}^c$	0.020 (0.057)	0.022 (0.063)
Mean ( $I$ )/ $\sigma(I)$	23.5 (9.8)	20.7 (4.5)
Completeness	99.9 (99.8)	97.5 (91.5)
Redundancy	11.5 (11.9)	3.9 (3.4)
Anomalous completeness	100 (100)	
Anomalous multiplicity	6.3 (6.4)	
<b>Phasing statistics</b>		
Number of sites	4	
Acentric reflections	8,920	
Overall FOM acentric <sup>d</sup>	0.37443	
Centric reflections	2314	
Overall FOM centric <sup>d</sup>	0.1174	
<b>Refinement statistics<sup>e</sup></b>		
Protein molecules (chains) in AU	1	10
Residues	181	2,030
Water molecules	159	1,216
Total number of atoms	1,496	16,426
$R_{\text{crys}}/R_{\text{free}}^g$ (%)	18.2/23.5	19.4/26.7
Average B-factor ( $\text{\AA}^2$ ) (all atoms)	23.9	24.9
r.m.s. deviation on angles (°)	1.476	1.754
r.m.s. deviation on bonds ( $\text{\AA}$ )	0.015	0.018

<sup>a</sup>  $R_{\text{merge}} = \frac{\sum \sum |I_i(hkl) - \bar{I}(hkl)|}{\sum \sum I_i(hkl)}$ , where  $I_i(hkl)$  is the observed intensity and  $\bar{I}(hkl)$  is the average intensity obtained from multiple observations of symmetry-related reflections after rejections.

<sup>b</sup>  $R_{\text{rim}} = \text{redundancy-independent merging } R\text{-factor (62). } R_{\text{rim}} = \sum_{hkl} [N/(N-1)]^{1/2} \sum_i |I_i(hkl) - \bar{I}(hkl)| / \sum \sum I_i(hkl)$ .

<sup>c</sup>  $R_{\text{pim}} = \text{precision-indicating merging } R\text{-factor (62). } R_{\text{pim}} = \sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \bar{I}(hkl)| / \sum \sum I_i(hkl)$ .

<sup>d</sup> FOM = figure of merit =  $\langle |\sum P(\alpha) e^{i\alpha}| / \sum P(\alpha) \rangle$ , where  $\alpha$  is the phase angle and  $P(\alpha)$  is the phase probability distribution.

<sup>e</sup> AU, asymmetric unit; r.m.s., root mean square.

<sup>f</sup>  $R_{\text{cryst}} = \sum |F_o| - |F_c| / \sum |F_o|$ , where  $F_o$  and  $F_c$  are the observed and calculated structure factors, respectively.

<sup>g</sup>  $R_{\text{free}}$  is calculated using 5% of the reflections randomly excluded from refinement. The data collection statistics in parentheses are the values for the highest resolution shell.

selenium sites were found. The program SOLOMON (27) was used to perform density modification. The program Arp/Warp (28) automatically built ~70% of the polypeptide chain. The rest of the model was built and fit using the program Coot (29). The triclinic crystal form (VP4<sup>tri</sup>) was solved by molecular replacement using the program Phaser (30) and the coordinates from the VP4<sup>hex</sup> structure as a search model. The structures were refined using the programs CNS (31) and Refmac5 (32). The final models were obtained by running restrained refinement in REFMAC5 with TLS (translation/libration/screw-rotation) restraints obtained from the TLS motion determination server (33). The PRODRG server (34) was used to define the molecular topologies and parameters for the ester linkage between the Ser<sup>633</sup>Oγ and carbonyl carbon of Ala<sup>716</sup> in the VP4<sup>tri</sup> structure. The data collection, phasing, and refinement statistics are summarized in Table 1.

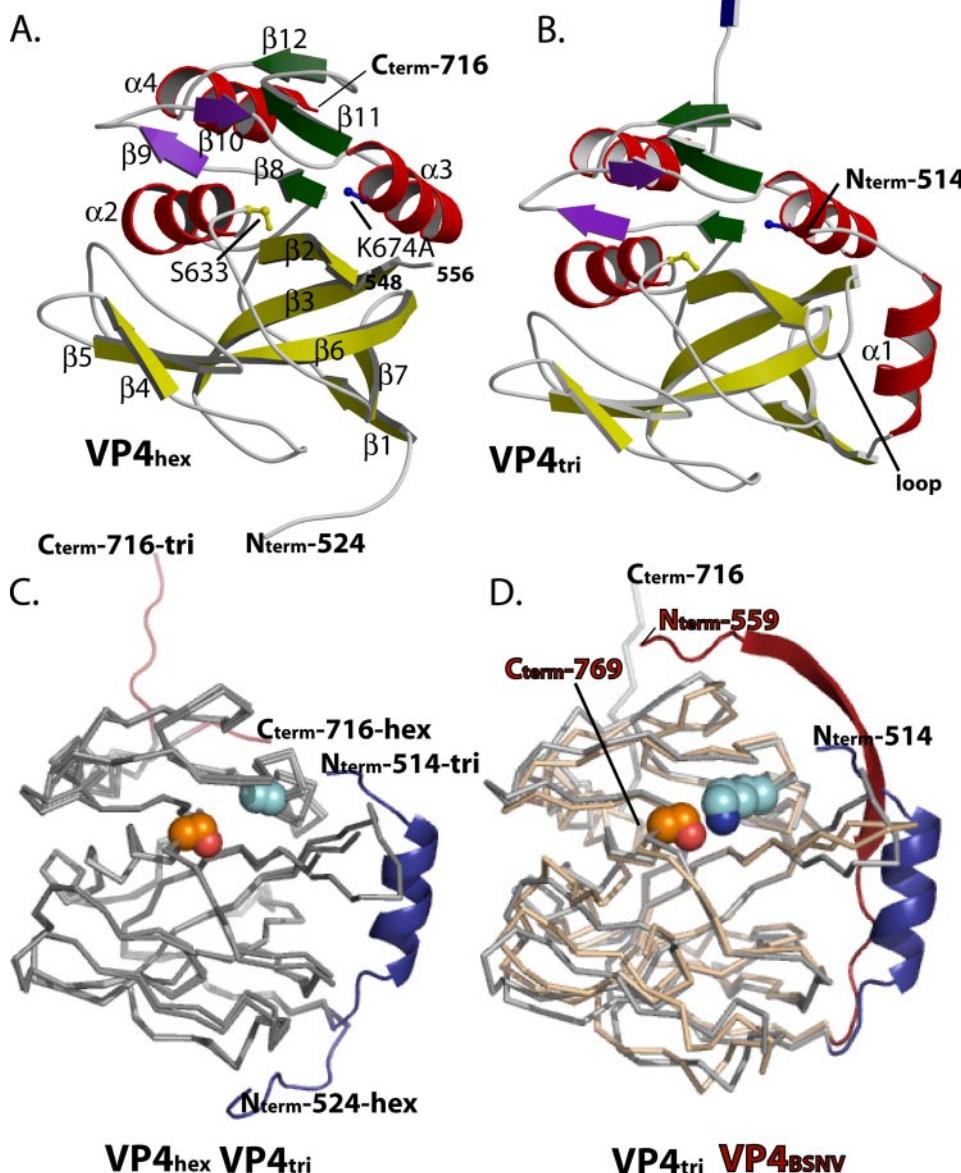
A probe radius of 1.4 Å was used in the calculations. The protein-protein interaction server (41, 42) was used to analyze the interactions between the molecules in the asymmetric unit in the VP4 crystal. The stereochemistry of the structures were analyzed with the program PROCHECK (43). The coordinates have been deposited in the Protein Data Bank (61).

**Figure Preparation**—Figures were prepared using the programs Molscript (44), Raster3D (45), and PyMol (46). The alignment figure was prepared using the programs ClustalW (47) and ESPript (48).

## RESULTS

**Overall Protein Architecture**—The final refined structure of the hexagonal crystal form of IPNV VP4 (VP4<sup>hex</sup>) shows electron density for all residues except for those in a disordered loop region between  $\beta$ -strands  $\beta_2$  and  $\beta_3$  (Val<sup>549</sup>–Glu<sup>555</sup>) as well as the last three residues at the C terminus (Fig. 2A). In contrast, the refined structure for the triclinic crystal form of IPNV VP4 (VP4<sup>tri</sup>) displays electron density for all residues including those in the disordered loop region missing in VP4<sup>hex</sup> (Fig. 2B). The core of the two structures are very similar and can be superposed with a root mean square deviation of 0.65 Å over 180 aligned residues (residues 524–548 and 556–710). IPNV VP4 protease has an  $\alpha/\beta$  protein fold that is composed of 13  $\beta$ -strands ( $\beta_1$ – $\beta_{13}$ ) and four  $\alpha$ -helices ( $\alpha_1$ – $\alpha_4$ ) as well as two  $\eta$ -helices ( $\eta_1$  and  $\eta_2$ ) (Figs. 2 and 3). Seven  $\beta$ -strands ( $\beta_1$ – $\beta_7$ ) assemble into a mixed, but mostly anti-parallel, twisted  $\beta$ -sheet that provides the scaffold for the substrate binding

Downloaded from www.jbc.org at University of British Columbia on June 11, 2008



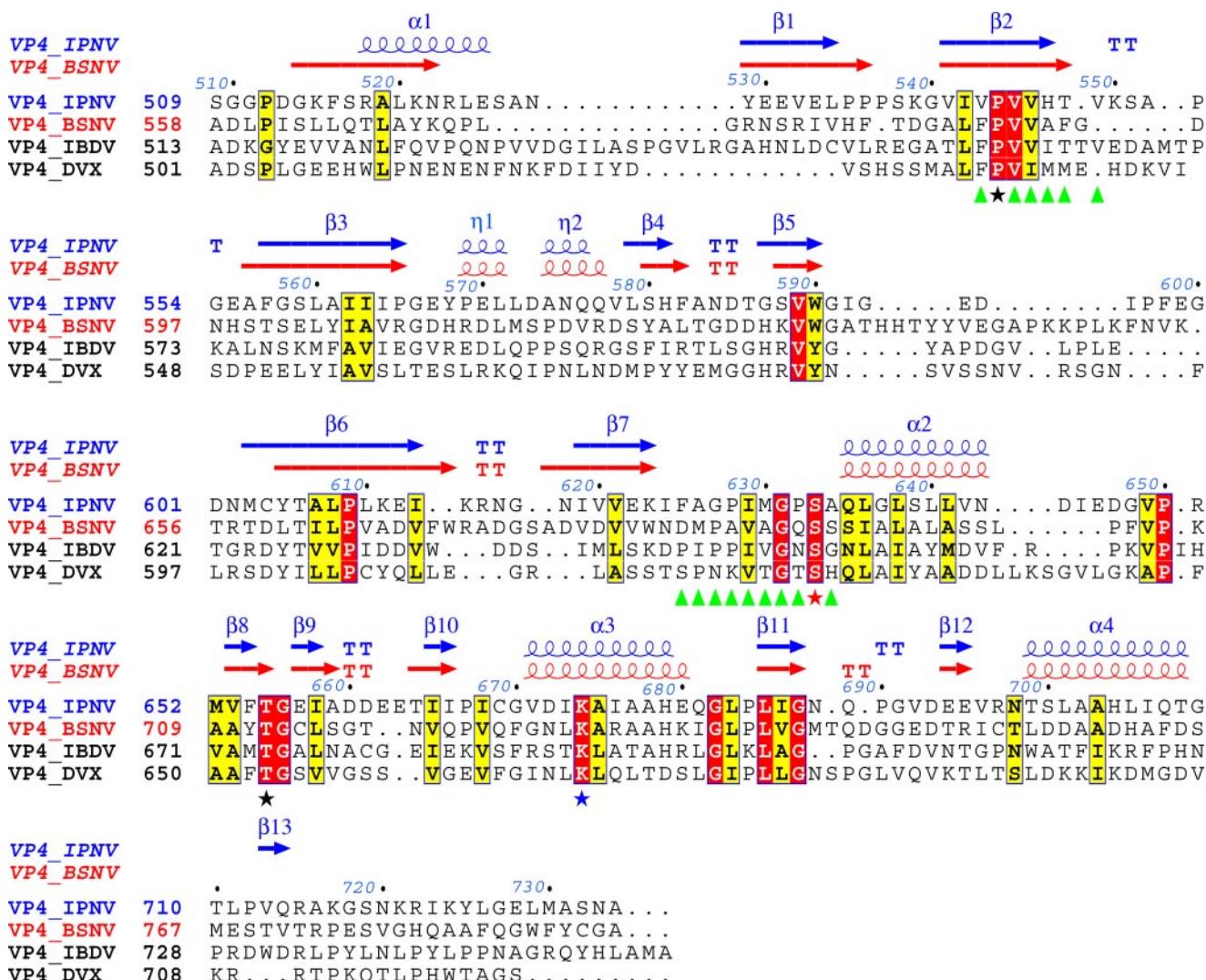
**FIGURE 2.** The structural fold of IPNV VP4 protease. *A*, ribbon diagram of IPNV VP4<sup>hex</sup>. The  $\alpha$ -helices (red) and  $\beta$ -strands (yellow, green, purple, and blue) are numbered sequentially. The  $\beta$ -sheets are colored yellow, green, and purple. The nucleophile Ser<sup>633</sup> and general base Lys<sup>674</sup> (which is mutated to alanine) are shown in stick format. The N and C termini residues are labeled. *B*, ribbon diagram of IPNV VP4<sup>tri</sup>. The secondary structural elements ( $\alpha_1$ ,  $\beta_{13}$ ), the disordered loop that was not observed in VP4<sup>hex</sup>, and the N and C termini are labeled. *C*, a superpositional comparison of the two IPNV VP4 crystal forms revealing differences in the orientation and structure of the N termini (blue) and C termini (orange). *D*, a superposition of IPNV VP4<sup>tri</sup> and BSNV VP4 (Protein Data Bank code 2GEF (13)). IPNV VP4 is light gray, and its N-terminal region (residues 514–531) is blue. BSNV VP4 is light orange, and its N-terminal region (residues 559–578) is red. The nucleophile Ser<sup>692</sup> and general base Lys<sup>729</sup> of BSNV VP4 are rendered as van der Waals spheres.

**Structural Analysis**—The secondary structural analysis was performed with the program DSSP (35). The programs SUPERIMPOSE (36) and SUPERPOSE (37) were used to overlap coordinates for structural comparison. The program CONTACT within the program suite CCP4 (38) was used to measure the hydrogen bonding and van der Waals contacts. The program CASTp (39) was used to analyze the molecular surface and measure the substrate binding site. The program SURFACE RACER 1.2 (40) was used to measure the solvent-accessible surface of the protein and individual atoms within the protein.

site. Three  $\alpha$ -helices ( $\alpha_2$ ,  $\alpha_3$ ,  $\alpha_4$ ) pack against a parallel  $\beta$ -sheet composed of three strands ( $\beta_8$ ,  $\beta_{11}$ ,  $\beta_{12}$ ), and the remaining two strands ( $\beta_9$ – $\beta_{10}$ ) form a short  $\beta$ -hairpin. The nucleophile Ser<sup>633</sup> resides just before  $\alpha$ -helix 2 ( $\alpha_2$ ) and the general base Lys<sup>674</sup> (which has been mutated to an alanine in these structures) is part of  $\alpha$ -helix 3 ( $\alpha_3$ ).

The alignment of the VP4 proteases across the different genera of the Birnavirus family reveals relatively little sequence identity and moderate sequence similarity (Fig. 3). For example, IPNV VP4 shares only 19% identity with BSNV VP4 over 225

## Structure of IPNV VP4 Protease



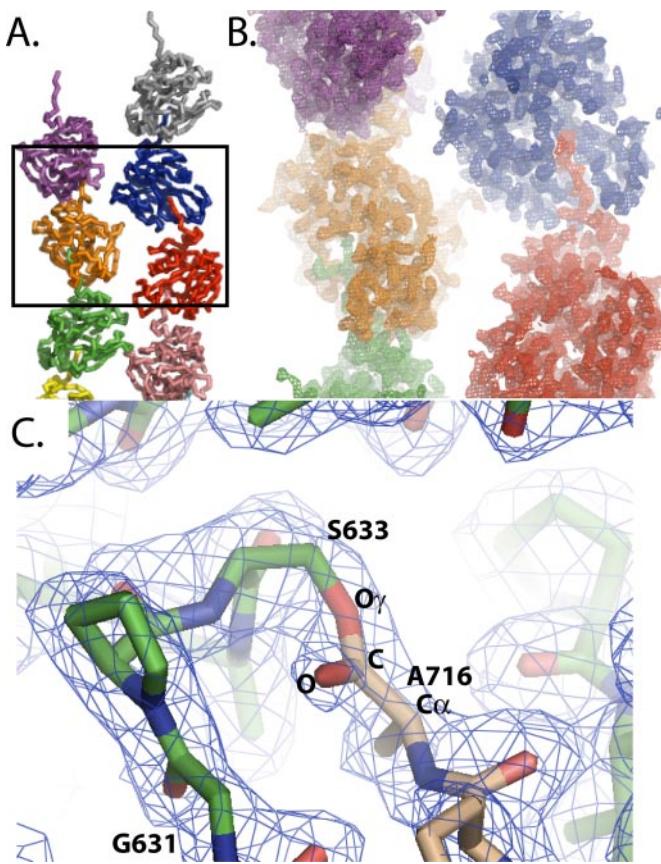
**FIGURE 3. A structure based sequence alignment of birnavirus VP4 proteases.** The VP4 sequences for IPNV (P90205), BSNV (Q8AZM0), infectious bursal disease virus (IBDV; P15480), and *Drosophila* X virus (DXV; Q96724) were obtained from the Swiss-Prot data bank (the corresponding accession numbers are in parentheses). The N and C termini of the VP4 sequences are defined by the pVP2/VP4 cleavage site and the VP4/VP3 cleavage site, respectively (see Fig. 1). The secondary structural elements of IPNV (blue) and BSNV (red) are shown aligned to their respective sequences. The secondary structural elements are numbered sequentially and are consistent with the ribbon diagrams shown in Fig. 2. The residues found in the catalytic site are denoted by stars: nucleophile Ser<sup>633</sup>, red star; general base Lys<sup>674</sup>, blue star; Pro<sup>544</sup> and Thr<sup>655</sup>, black stars. The residues that flank the substrate binding site and contribute main-chain hydrogen bonding interactions with the substrate are denoted by triangles. Residues identically conserved or similar in all four birnavirus sequences are highlighted in red and yellow, respectively.

amino acids. Yet, three-dimensional alignment of the IPNV VP4 structure with the recently solved BSNV VP4 structure (13) (Protein Data Bank code 2GEF) reveals an overall conservation of the protein architecture (Fig. 2D). The VP4 structures superimposed with a root mean square deviation of 1.8 Å over 170 aligned residues. The most significant difference is seen at the N terminus where the IPNV VP4<sup>tri</sup> contains an  $\alpha$ -helix, whereas the BSNV VP4 has an extended  $\beta$ -strand that forms an anti-parallel  $\beta$ -sheet interaction with a neighboring molecule in the BSNV crystal structure (13) (Fig. 2D).

The most significant differences between the two IPNV VP4 structures occur at the N and C termini (Fig. 2C). In the crystal structure of VP4<sup>tri</sup>, residues 518–526, near the N terminus, form an  $\alpha$ -helix ( $\alpha$ 1) that flanks the loop region (residues 548–557) between strands  $\beta$ 2 and  $\beta$ 3. This loop region lays adjacent to the substrate binding site. Two residues just before helix  $\alpha$ 1

(Lys<sup>515</sup> and Ser<sup>517</sup>) make hydrogen-bonding interactions with residues in the loop region (Gly<sup>554</sup>, Glu<sup>555</sup>, and Ala<sup>556</sup>), helping to stabilize it. In addition, the Arg<sup>518</sup> side chain is within hydrogen bonding distance to the Glu<sup>680</sup> side chain in helix  $\alpha$ 3, further stabilizing this region. In the VP4<sup>hex</sup> crystal structure, the shorter N terminus does not form an  $\alpha$ -helix ( $\alpha$ 1) and is positioned away from the loop (residues 548–557) and the binding site. The lack of stabilizing interaction is consistent with the missing density for the loop (residues 548–557) in the VP4<sup>hex</sup> structure. Both VP4<sup>hex</sup> and VP4<sup>tri</sup> end with residue 716, yet the C termini are very different in the two structures because of an interesting intermolecular interaction.

*The Acyl-Enzyme Intermediate for a trans-Cleavage Reaction Illuminates the Substrate Binding Sites*—The differences in the C-terminal position between the two IPNV VP4 structures stems from an intimate intermolecular interaction between the



**FIGURE 4. The acyl-enzyme intermediate for an intermolecular (trans) cleavage reaction by the Ser/Lys protease IPNV VP4.** *A*, C $\alpha$  trace of the molecules in the asymmetric unit of the triclinic crystal form of IPNV VP4. The C terminus of each VP4 molecule resides in the substrate binding site of the neighboring VP4 molecule. *B*, 2F<sub>o</sub> – F<sub>c</sub> electron density map (contoured at 1 sigma) for the molecules that are boxed in *A*. *C*, 2F<sub>o</sub> – F<sub>c</sub> electron density map (contoured at 1 sigma) near the nucleophile Ser<sup>633</sup> in IPNV VP4 protease. The atoms that form the ester linkage between Ser<sup>633</sup>O $\gamma$  in one VP4 molecule (green, carbon; blue, nitrogen; red, oxygen) and the Ala<sup>716</sup>C from a neighboring molecule (tan, carbon; blue, nitrogen; red, oxygen) are labeled.

VP4 molecules in the asymmetric unit of the triclinic crystal form (VP4<sup>tri</sup>). The C terminus of each VP4<sup>tri</sup> molecule in the asymmetric unit is bound in an anti-parallel  $\beta$ -sheet fashion into the substrate binding site of its neighboring VP4<sup>tri</sup> molecule (Fig. 4, *A* and *B*). Close inspection of 2F<sub>o</sub> – F<sub>c</sub> and omit electron density maps near the nucleophile Ser<sup>633</sup> is consistent with 6 of the 10 molecules in the asymmetric unit forming a covalent linkage between the Ser<sup>633</sup> side-chain O $\gamma$  and the C-terminal Ala<sup>716</sup> main-chain carbonyl carbon (Fig. 4C). The acyl-enzyme ester bond (O $\gamma$ –C–O–C $\alpha$ ) appears to be planar, consistent with a previously reported serine protease acyl-enzyme complex with a short peptide (49) (Fig. 4C). To determine the absolute degree of planarity will require a higher resolution structure of VP4<sup>tri</sup>.

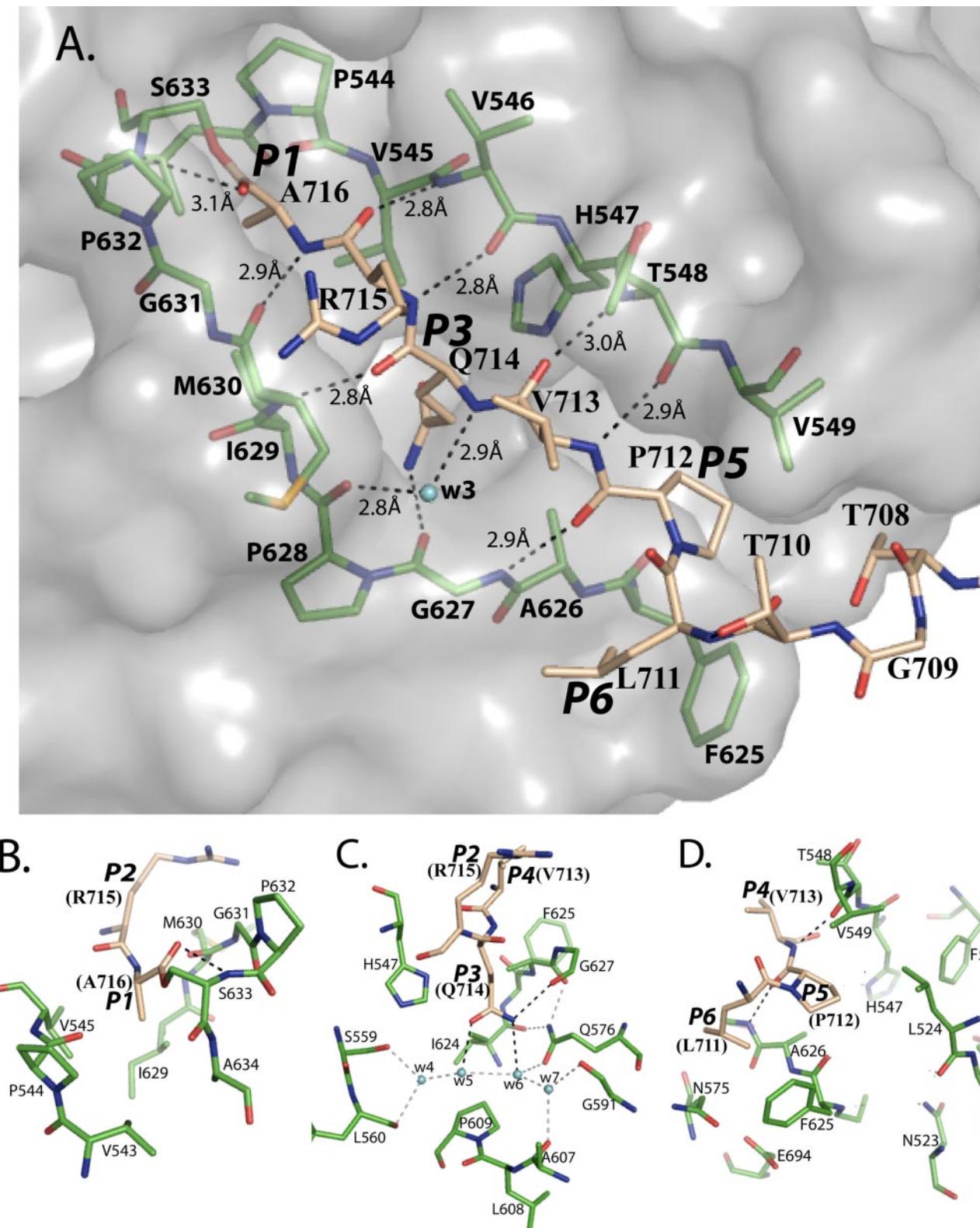
The C-terminal Ala<sup>716</sup> represents the P1 residue (50) of the previously described internal cleavage site within the C-terminal end of the VP4 molecule (3) (Fig. 1). Therefore the VP4<sup>tri</sup> structure with the C terminus bound into the binding site of its neighboring VP4<sup>tri</sup> molecule allows us to describe the substrate/protease binding site interactions for this intermolecular (trans) self-cleavage reaction.

The intermolecular interactions seen between the VP4 C terminus with the binding site of its neighboring VP4 molecule buries  $\sim$ 1000  $\text{\AA}^2$  of accessible surface area on each molecule (1070  $\text{\AA}^2$  on molecule A and 942  $\text{\AA}^2$  on molecule B); this is  $\sim$ 10% of the total accessible surface area on the VP4 molecule. Helix 4 ( $\alpha$ 4) near the C-terminal end of the VP4 molecule is positioned close to the opening of the substrate binding groove of the neighboring molecule. The C-terminal residues 711–716 extend into the substrate binding groove, ending with the terminal residue, Ala<sup>716</sup> placed in close proximity to the nucleophile Ser<sup>633</sup>. The residues Pro<sup>712</sup>, Val<sup>713</sup>, Gln<sup>714</sup>, Arg<sup>715</sup>, and Ala<sup>716</sup> are stabilized by antiparallel  $\beta$ -sheet hydrogen bonding interactions with the substrate binding groove, formed on one side by  $\beta$ -strand 2 ( $\beta$ 2) and on the other side by the loop that leads to  $\alpha$ -helix 2 ( $\alpha$ 2) from  $\beta$ -strand 7 ( $\beta$ 7) (Figs. 2A and 5A). Interestingly, a water molecule (w3) helps to complete the hydrogen bonding network in the substrate binding groove. This water is also seen in the apoenzyme structure (VP4<sup>hex</sup>). The residues that contribute atoms to the substrate binding groove in the VP4<sup>tri</sup> structure are denoted in the sequence alignment in Fig. 3.

**The Substrate Binding Pockets**—The VP4 acyl-enzyme complex binding site interactions are depicted in Fig. 5. The side chain of Leu<sup>711</sup>, the P6 residue in the internal cleavage site, lies against a shallow pocket (S6) near the entrance of the substrate binding groove. The Leu<sup>711</sup> side chain packs against the residues Asn<sup>575</sup> (C $\beta$ ), Phe<sup>625</sup>, and Ala<sup>626</sup> (Fig. 5D). The P5 residue (Pro<sup>712</sup>) fits into a deep cleft (S5) that is open on one side. The residues of S5 that are closest to the P5 residue include Leu<sup>524</sup>, Val<sup>549</sup>, Phe<sup>625</sup>, and Ala<sup>626</sup> (Fig. 5D). The side chain of Gln<sup>714</sup> (P3) points into the S3 binding pocket, made up of Val<sup>545</sup>, Val<sup>546</sup>, His<sup>547</sup>, Ser<sup>559</sup>, Gln<sup>576</sup>, Pro<sup>609</sup>, Ile<sup>624</sup>, Gly<sup>627</sup>, Pro<sup>628</sup>, and Ile<sup>629</sup>. The side-chain N82 of Gln<sup>714</sup> is within the hydrogen bonding distance (3.1  $\text{\AA}$ ) to the main-chain oxygen of Gly<sup>627</sup>. The N82 and O81 of Gln<sup>714</sup> are also hydrogen-bonded to two waters (w5 and w6) at the base of the S3 binding site (Fig. 5C). These waters are also seen in the apoenzyme structure. The methyl side chain of Ala<sup>716</sup> (the P1 residue) resides within a hydrophobic pocket (S1) adjacent to the nucleophile Ser<sup>633</sup>. Residues contributing to the S1 binding pocket include Val<sup>543</sup>, Pro<sup>544</sup>, Val<sup>545</sup>, Ile<sup>629</sup>, Met<sup>630</sup>, Gly<sup>631</sup>, Pro<sup>632</sup>, Ser<sup>633</sup>, and Ala<sup>634</sup> (Fig. 5B). The side chains of Val<sup>713</sup> (P4) and Arg<sup>715</sup> (P2) point toward the solvent and away from the VP4 molecular surface.

## DISCUSSION

**Recognition of the Internal Cleavage Site within the VP4 Polypeptide Sequence**—Based on our own IPNV VP4 purification studies and previous reports of self-cleavage at an internal VP4 cleavage site (3) (between residues Ala<sup>716</sup> and Lys<sup>717</sup>; Fig. 1), we designed a number of IPNV VP4 constructs with a C terminus ending at Ala<sup>716</sup> (23). The crystal structure of the triclinic crystal form of IPNV VP4 (VP4<sup>tri</sup>) reveals the C terminus (residues 712–716) of each VP4<sup>tri</sup> molecule lying in an extended  $\beta$ -conformation along the substrate binding groove of its neighboring VP4<sup>tri</sup> molecule (Fig. 5). In addition to the anti-parallel  $\beta$ -sheet hydrogen bonding pattern formed to stabilize the interaction, the appropriate side chains for the cleavage specificity residues (P1 (Ala<sup>716</sup>), P3 (Gln<sup>714</sup>), P5 (Pro<sup>712</sup>), and P6



**FIGURE 5. The substrate binding subsites of IPNV VP4 protease as revealed by the acyl-enzyme complex.** **A**, antiparallel  $\beta$ -sheet interactions between the IPNV VP4 protease substrate binding site and the N-terminal product for the internal cleavage site with VP4 attached by an acyl-ester bond to the Ser<sup>633</sup> nucleophile. Those residues that line each side of the substrate binding groove are shown in stick format behind a semitransparent molecular surface. The hydrogen bonds that are formed between the binding site residues of one molecule (green) and the C terminus of a neighboring molecule (tan) are shown by dashed lines. Water molecule w3 is shown in cyan and is labeled. **B**, molecular interactions observed between the P1 residue of the substrate (Ala<sup>716</sup>, internal cleavage site) and the S1 binding pocket. **C**, molecular interactions observed between the P3 residue of the substrate (Gln<sup>714</sup>, internal cleavage site) and the S3 binding pocket of IPNV VP4 protease. The residues Val<sup>546</sup>, His<sup>547</sup>, Pro<sup>628</sup>, and Ile<sup>629</sup>, which are part of the S3 pocket, are not shown for clarity of the figure. Dashed lines depict the hydrogen bonding network formed between the substrates P3 residue (Gln<sup>714</sup>, shown in tan color) and the residues (green) and waters (w4, w5, w6, and w7: cyan spheres) residing in the S3 binding pocket. **D**, molecular interactions observed between the P5 and P6 residues of the substrate (Pro<sup>712</sup>/Leu<sup>711</sup>, internal cleavage site) and the S5 and S6 binding pockets, respectively.

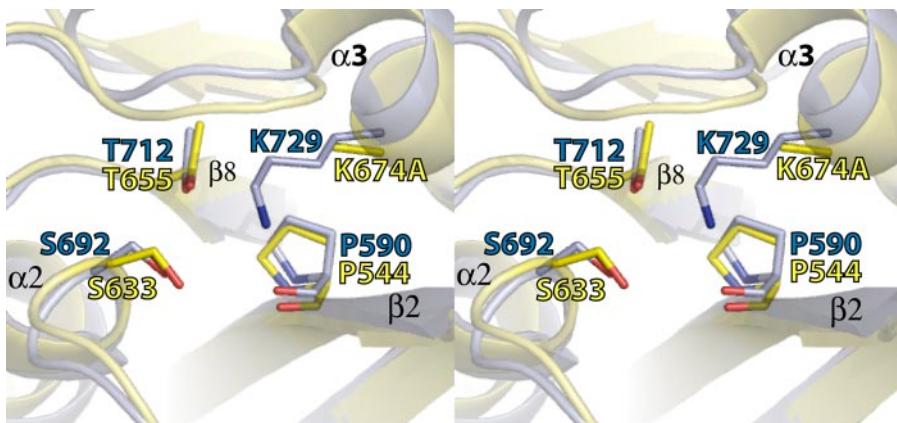


FIGURE 6. Conservation of the active site region of VP4 protease. Superposition of the IPNV VP4 protease (yellow) and the BSNV VP4 protease (blue) is shown.

(Leu<sup>711</sup>) point into complementary pockets on the surface of the neighboring VP4<sup>tri</sup> (Fig. 5). These intermolecular interactions orient the C-terminal Ala<sup>716</sup> main-chain carbonyl carbon adjacent to the nucleophilic Ser<sup>633</sup>O $\gamma$  and forms a planar covalent ester bond (in 6 of the 10 molecules in the asymmetric unit) as expected for an acyl-enzyme intermediate. These interactions collectively contain the structural hallmarks of a proteinaceous substrate bound within the active site of a protease, providing our first view of substrate binding and acyl-enzyme intermediate formation in the functionally diverse Ser/Lys protease superfamily. The complex captured in our crystal structures also clearly defines the self-cleavage (in *trans*) of a viral polyprotein by its inherent protease, a process essential to the replication and assembly of many viruses and, to the best of our knowledge, never previously observed directly by crystallographic methods.

Whereas the P1 (Ala) and P2' (Gly) are highly conserved across all IPNV cleavage sites (the latter we propose playing an essential role in redirecting the polyprotein substrate out of the active site), the P3 position is somewhat more variable (including Ser, Thr, and Gln; Fig. 1C). When serine or threonine are modeled at the P3 position (residue 714) in the VP4<sup>tri</sup> structure, the hydroxyl of the side chain comes in close proximity to the His<sup>547</sup> side chain in the S3 binding pocket (Fig. 5C). This polar interaction may explain the prevalence of a hydroxyl group at the P3 position of the cleavage site. In support of our structural observations, previous mutagenesis studies have shown that a H547S mutation affects cleavage and substrate specificity at the VP4/VP3 junction (3). A notable difference of the S3 subsite as compared with the fully hydrophobic S1 pocket is the deeper size and ring of highly ordered water molecules that line the pocket base. The presence of these waters appear to act as molecular putty to mold around the serine, threonine, or larger glutamine that are accommodated at this site (the latter observed directly in our structures; Fig. 5C).

Because of the  $\beta$ -sheet-type conformation of the substrate in the VP4 binding site, the side chains of the P2 and P4 residues are oriented away from the VP4 binding site surface, toward the solvent, and thus would not be a factor in the specificity, in line with the lack of conservation at these residues in the cleavage sites of the viral polyprotein. The residues at the P5 position are somewhat variable, but all have a common hydrophobic nature. In

our acyl-enzyme structure the P5 Pro<sup>712</sup> sits in a deep cleft capable of accepting hydrophobic side chains of varying sizes (Fig. 5D). By extending the alignment of the IPNV cleavage sites out to the P6 residue (Fig. 1C) one can see that four of the six residues at this position are leucine, the others being valine and arginine. The VP4 acyl-enzyme structure shows that the Leu<sup>711</sup> side chain lies in a shallow hydrophobic pocket, which would be consistent with the substrate residues at this position (Fig. 5). Because of the structural orientation of the substrate at the P6

position, the one cleavage site that has an arginine side chain at this position would be able to contribute the aliphatic portion of its side chain into the S6 pocket while having the charged guanidinium group still exposed to the solvent.

**VP4 Catalytic Machinery**—The most ordered crystals of IPNV VP4 protease obtained so far have been those that have formed from constructs where the lysine general base has been mutated to an alanine (23). Although this prevents us from observing the Ser/Lys catalytic dyad in this enzyme, we have determined previously the structure of VP4 from the BSNV virus (13), where the proposed general base Lys<sup>729</sup>N $\zeta$  is appropriately positioned within hydrogen bonding distance to the nucleophilic Ser<sup>692</sup>O $\gamma$ , with further stabilization by the conserved Thr<sup>712</sup>Oyl and Pro<sup>590</sup>O (Fig. 3). Despite the 19% sequence identity, our structure of the IPNV VP4 mutant presented here and that of the BSNV variant overlap very closely such that residues Ser<sup>633</sup>, Lys<sup>674</sup>, Thr<sup>655</sup>, and Pro<sup>544</sup> (IPNV VP4) match their active site counterparts in BSNV VP4 (Ser<sup>692</sup>, Lys<sup>729</sup>, Thr<sup>712</sup> and Pro<sup>590</sup>, respectively; Fig. 6). Hence in IPNV VP4, it is likely that the general base Lys<sup>674</sup>N $\zeta$  would be positioned in a similar fashion as that seen for Lys<sup>729</sup>N $\zeta$  in the BSNV VP4 structure. The environment for the lysine general base in the VP4 protease is also similar to that seen in other Ser/Lys proteases such as bacterial Lon protease (14, 15), bacterial signal peptidase (19, 51, 52), and bacterial LexA repressor (20). Based upon the position of the nucleophile Ser<sup>633</sup> relative to the location of the S1 binding pocket, it can be seen that the IPNV VP4, like the other Ser/Lys proteases, attacks the scissile bond of its substrate from the *si*-face rather than the *re*-face typical of other serine proteases such as the trypsin-like proteases.

The oxyanion hole in serine proteases is an important component of the catalytic machinery. It functions by neutralizing the charge that develops on the oxyanion tetrahedral intermediate. It is usually constructed from two main-chain amides providing hydrogen bond donors to the scissile carbonyl oxygen (53). Our structure of the acyl-enzyme intermediate of the IPNV VP4 protease presented here provides the first opportunity to directly define the oxyanion hole components of this class of enzyme. We observe the scissile carbonyl oxygen is within hydrogen bonding distance of the serine nucleophiles main-chain nitrogen, a common feature of many classes of serine hydrolases. A second electrostatic feature of the oxyanion hole

arises from the appropriate positioning of the carbonyl oxygen of the P1 residue (Ala<sup>716</sup>) relative to  $\alpha$ -helix 2 immediately following the nucleophile Ser<sup>633</sup> (Fig. 2). The structure thus suggests that the partial positive charge on the helix dipole may play a role in stabilizing the oxyanion, as suggested previously for other proteases such as subtilisin (54) and rhomboid protease (55). Based on the conservation of the residues in the binding site (Fig. 3), an additional residue that might be expected to aid in forming the oxyanion hole in IPNV VP4 would be Gly<sup>631</sup>. However, as can be seen from Fig. 5 the main-chain NH of Gly<sup>631</sup> is pointing away from the binding site, in an orientation incompatible with a role in oxyanion stabilization. The conformation of this conserved Gly<sup>631</sup> is maintained in both the apo- and acyl-enzyme crystal structures of the IPNV VP4 presented here as well as that of the earlier BSNV VP4 crystal structure (Gly<sup>690</sup> being the equivalent residue in BSNV), suggesting that its observed orientation is a feature of these enzymes and is not due to constraints imposed by crystallographic packing.

It is an interesting question as to how VP4<sup>tri</sup>, with the lysine general base mutated to an alanine, is able to activate the Ser<sup>633</sup>O $\gamma$  for attack on the carbonyl of the C-terminal Ala<sup>716</sup> to create the acyl-enzyme complex, the first step in this protease's reverse reaction. There are no titratable functional groups from the binding site in the vicinity of the Ser<sup>633</sup>O $\gamma$  that could function as the general base. Interestingly, Arg<sup>715</sup> at the P2 position of the substrate (Fig. 5A) has a guanidinium group in its side chain that could potentially be positioned close enough to the Ser<sup>633</sup> hydroxyl such that it could assist in the deprotonation. Similarly, guanidine is present in the crystal, although we did not observe a guanidine molecule in a position amenable for acting as the general base. In addition, the pH of the crystallization condition is not basic enough for a guanidinium to be significantly deprotonated at its usual  $pK_a$  of 12.5, unless it were buried within the protein such that it had a depressed  $pK_a$ . Another possibility is that the C-terminal Ala<sup>716</sup> carboxylate itself could be functioning to deprotonate the Ser<sup>633</sup> hydroxyl. The acylation may be driven by the high protein concentration in the crystal or by dehydrating conditions in the crystallization reagents. Preliminary studies suggest that this reaction only occurs to any significant degree in the crystal and not in solution (data not shown). Interestingly, short peptides ending with a C-terminal carboxylate have been observed previously to form acyl-enzyme complexes with the serine protease elastase (49, 56). Although in the case of elastase, the general base was present in the active site.

Despite the fact that the deacylation step is rate-limiting for many serine proteases, it is difficult to trap the acyl-enzyme complex in a crystal structure. As discussed by Radisky *et al.* (57), acyl-enzyme structures have been reported with serine proteases using: (i) short peptide substrates and low pH (49, 56), (ii) small ester substrates using flash cooling (58, 59), (iii) small molecule inhibitors such as  $\beta$ -lactams (19, 60), and (iv) small peptide substrates using pseudo-steady state conditions (57). In the VP4<sup>tri</sup> crystals with the general base Lys<sup>674</sup> mutated to alanine, the acyl-enzyme seems to be in equilibrium with the N-terminal product complex. The protein/protein acyl-enzyme complex in the VP4<sup>tri</sup> crystals reveals an extended interaction along the substrate binding groove, which has not been

observed in the previous serine protease acyl-enzyme structures with short peptides.

In the VP4<sup>tri</sup> structure we arrived at the acyl-enzyme complex by providing the N-terminal product at an equimolar concentration to the binding site in the crystal without the catalytic general base/acid present. It would be interesting to investigate whether this maybe a general method for revealing the structure of acyl-enzyme complexes in other viral serine or cysteine proteases that cleave in *trans*.

**Concluding Remarks**—The structures of IPNV VP4 protease have allowed for the identification of the substrate binding subsites (S1, S3, S5, and S6) and a structural explanation for the IPNV VP4 cleavage site specificity. Comparing the IPNV VP4 structures with the BSNV VP4 structure reveals that the Ser/Lys catalytic machinery, as well as the overall protein fold, is conserved among the birnavirus VP4 proteases. This first substrate complex for a Ser/Lys protease reveals important insights into the structure and function of this unique class of serine protease and provides the structural basis for the rational design of antiviral compounds for the birnaviruses.

**Acknowledgments**—We thank Dr. R. M. Sweet and all of the Rapi-Data 2006 staff at the Brookhaven National Laboratory NSLS beam line X8C. We thank Dr. Lawrence P. McIntosh for critical reading of the manuscript.

## REFERENCES

1. Espinoza, J. C., Hjalmarsson, A., Everitt, E., and Kuznar, J. (2000) *Arch. Virol.* **145**, 739–748
2. Dobos, P., Hill, B. J., Hallett, R., Kells, D. T., Becht, H., and Teninges, D. (1979) *J. Virol.* **32**, 593–605
3. Petit, S., Lejal, N., Huet, J. C., and Delmas, B. (2000) *J. Virol.* **74**, 2057–2066
4. Coulibaly, F., Chevalier, C., Gutsche, I., Pous, J., Navaza, J., Bressanelli, S., Delmas, B., and Rey, F. A. (2005) *Cell* **120**, 761–772
5. Delmas, B., Kibenge, F. S. B., Leong, J. C., Mundt, E., Vakharia, V. N., and Wu, J. L. (2005) in *Virus Taxonomy: The Eighth Report of the International Committee on Taxonomy of Viruses* (Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., and Ball, L. A., eds) pp. 561–569, Elsevier, Amsterdam
6. Dobos, P. (1995) *Virology* **208**, 19–25
7. Roberts, R. J., and Pearson, M. D. (2005) *J. Fish Dis.* **28**, 383–390
8. Reno, P. (1999) in *Fish Diseases and Disorders* (Woo, P. T. K., and Bruno, D. W., eds) Vol. 3, pp. 1–55, CABI Publishing, New York
9. Galloux, M., Chevalier, C., Henry, C., Huet, J. C., Costa, B. D., and Delmas, B. (2004) *J. Gen. Virol.* **85**, 2231–2236
10. Da Costa, B., Soignier, S., Chevalier, C., Henry, C., Thory, C., Huet, J. C., and Delmas, B. (2003) *J. Virol.* **77**, 719–725
11. Birghan, C., Mundt, E., and Gorbatenya, A. E. (2000) *EMBO J.* **19**, 114–123
12. Lejal, N., Da Costa, B., Huet, J. C., and Delmas, B. (2000) *J. Gen. Virol.* **81**, 983–992
13. Feldman, A. R., Lee, J., Delmas, B., and Paetzel, M. (2006) *J. Mol. Biol.* **358**, 1378–1389
14. Botos, I., Melnikov, E. E., Cherry, S., Kozlov, S., Makhovskaya, O. V., Tropea, J. E., Gustchina, A., Rotanova, T. V., and Wlodawer, A. (2005) *J. Mol. Biol.* **351**, 144–157
15. 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., and Gustchina, A. (2004) *J. Biol. Chem.* **279**, 8140–8148
16. Rawlings, N. D., and Barrett, A. J. (1999) *Nucleic Acids Res.* **27**, 325–331
17. Paetzel, M., and Dalbey, R. E. (1997) *Trends Biochem. Sci.* **22**, 28–31
18. Paetzel, M., and Strynadka, N. C. (1999) *Protein Sci.* **8**, 2533–2536
19. Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (1998) *Nature* **396**, 186–190

20. Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W., and Strynadka, N. C. J. (2001) *Cell* **106**, 1–10

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

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

23. Lee, J., Feldman, A. R., Chiu, E., Chan, C., Kim, Y. N., Delmas, B., and Paetzel, M. (2006) *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **62**, 1235–1238

24. Otwinowski, Z. and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326

25. Leslie, A. G. W. (1992) *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, No. 26, SERC, Daresbury Laboratory, Warrington, UK

26. Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2006) *Methods Mol. Biol.* **364**, 215–230

27. Abrahams, J. P., and Leslie, A. G. (1996) *Acta Crystallogr. Sect. D* **52**, 30–42

28. Morris, R. J., Perrakis, A., and Lamzin, V. S. (2003) *Methods Enzymol.* **374**, 229–244

29. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D* **60**, 2126–2132

30. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) *Acta Crystallogr. Sect. D* **61**, 458–464

31. Brunger, 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., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D* **54**, 905–921

32. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) *Acta Crystallogr. Sect. D* **57**, 122–133

33. Painter, J., and Merritt, E. A. (2006) *Acta Crystallogr. Sect. D* **62**, 439–450

34. van Aalten, D. M., Bywater, R., Findlay, J. B., Hendlich, M., Hooft, R. W., and Vriend, G. (1996) *J. Comput. Aided Mol. Des.* **10**, 255–262

35. Kabsch, W., and Sander, E. (1983) *Biopolymers* **22**, 2577–2637

36. Diederichs, K. (1995) *Proteins* **23**, 187–195

37. Maiti, R., Van Domselaar, G. H., Zhang, H., and Wishart, D. S. (2004) *Nucleic Acids Res.* **32**, W590–W594

38. Collaborative Computing Project No. 4 (1994) *Acta Crystallogr. Sect. D* **50**, 760–763

39. Liang, J., Edelsbrunner, H., and Woodward, C. (1998) *Protein Sci.* **7**, 1884–1897

40. Tsodikov, O. V., Record, M. T., Jr., and Sergeev, Y. V. (2002) *J. Comput. Chem.* **23**, 600–609

41. Jones, S., and Thornton, J. M. (1995) *Prog. Biophys. Mol. Biol.* **63**, 31–65

42. Jones, S., and Thornton, J. M. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13–20

43. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291

44. Kraulis, P. G. (1991) *J. Appl. Crystallogr.* **24**, 946–950

45. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277**, 505–524

46. DeLano, W. L. (2002) *The PyMOL Molecular User's Manual*, DeLano Scientific, San Carlos, CA

47. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680

48. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* **15**, 305–308

49. Katona, G., Wilmouth, R. C., Wright, P. A., Berglund, G. I., Hajdu, J., Neutze, R., and Schofield, C. J. (2002) *J. Biol. Chem.* **277**, 21962–21970

50. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27**, 157–162

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

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

53. Menard, R., and Storer, A. C. (1992) *Biol. Chem. Hoppe-Seyler* **373**, 393–400

54. Hol, W. G., van Duijnen, P. T., and Berendsen, H. J. (1978) *Nature* **273**, 443–446

55. Lemieux, M. J., Fischer, S. J., Cherney, M. M., Bateman, K. S., and James, M. N. G. (2007) *Proc. Natl. Acad. Sci. U. S. A.* **104**, 750–754

56. Wilmouth, R. C., Clifton, I. J., Robinson, C. V., Roach, P. L., Aplin, R. T., Westwood, N. J., Hajdu, J., and Schofield, C. J. (1997) *Nat. Struct. Biol.* **4**, 456–462

57. Radisky, E. S., Lee, J. M., Lu, C. J., and Koshland, D. E., Jr. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 6835–6840

58. Alber, T., Petsko, G. A., and Tsernoglou, D. (1976) *Nature* **263**, 297–300

59. Ding, X., Rasmussen, B. F., Petsko, G. A., and Ringe, D. (1994) *Biochemistry* **33**, 9285–9293

60. Wilmouth, R. C., Westwood, N. J., Anderson, K., Brownlee, W., Claridge, T. D., Clifton, I. J., Pritchard, G. J., Aplin, R. T., and Schofield, C. J. (1998) *Biochemistry* **37**, 17506–17513

61. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* **28**, 235–242

62. Weiss, M. G. (2001) *J. Appl. Crystallogr.* **34**, 130–135