Crystal Structures of Yellowtail Ascites Virus VP4 Protease
TRAPPING AN INTERNAL CLEAVAGE SITE TRANS ACYL-ENZYME COMPLEX IN A NATIVE SER/LYS DYAD ACTIVE SITE*

Ivy Yeuk Wah Chung and Mark Paetzel
From the Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

Received for publication, January 30, 2013, and in revised form, March 14, 2013 Published, JBC Papers in Press, March 19, 2013, DOI 10.1074/jbc.M112.386953

Yellowtail ascites virus (YAV) is an aquabirnavirus that causes ascites in yellowtail, a fish often used in sushi. Segment A of the YAV genome codes for a polyprotein (pVP2-VP4-VP3), where processing by its own VP4 protease yields the capsid protein precursor pVP2, the ribonucleoprotein-forming VP3, and free VP4. VP4 protease utilizes the rarely observed serine-lysine catalytic dyad mechanism. Here we have confirmed the existence of an internal cleavage site, preceding the VP4/VP3 cleavage site. The resulting C-terminally truncated enzyme (ending at Ala716) is active, as shown by a trans full-length VP4 cleavage assay and a fluorometric peptide cleavage assay. We present a crystal structure of a native active site YAV VP4 with the internal cleavage site trapped as the acyl-enzyme complexes and trans acyl-enzyme complexes. The acyl-enzyme complexes confirm directly the role of Ser633 as the nucleophile. A crystal structure of the lysine general base mutant (K674A) reveals the acyl–enzyme and empty binding site states of VP4, which allows for the observation of structural changes upon substrate or product binding. These snapshots of three different stages in the VP4 protease reaction mechanism will aid in the design of anti-birnavirus compounds, provide insight into previous site-directed mutagenesis results, and contribute to understanding of the serine-lysine dyad protease mechanism. In addition, we have discovered that this protease contains a channel that leads from the enzyme surface (adjacent to the substrate binding groove) to the active site and the deacylating water.

Birnaviruses have single-shelled non-enveloped capsids of 60–70 nm in diameter and carry bisegmented (segment A and B) double-stranded RNA genomes (1). Within the family of Birnaviridae, members are categorized into four genera according to the host they infect: Avibirnavirus, Aquabirnavirus, Entomobirnavirus, and Blosnavirus (2). In segment A, the larger open reading frame codes for the polyprotein NH2-pVP2-VP4-VP3-COOH (Fig. 1A). An exception is found in blotched snakhead virus (3) and tellina virus 1 (4), where polypeptide X is inserted between pVP2 and VP4. VP4 is the protease that processes the polyprotein into capsid protein precursor pVP2, free VP4, and ribonucleoprotein-forming VP3 (5–8). The pVP2 is further processed to release capsid protein VP2 and a series of peptides (9).

VP4 is a serine endopeptidase that utilizes a serine-lysine catalytic dyad mechanism rather than the classical serine-histidine-aspartate catalytic triad mechanism (10). The serine-lysine catalytic dyad is also found in bacterial enzymes: type 1 signal peptidase (11), UmuD (12), LexA (13), Lon protease (14), and signal peptide peptidase (SppA) (15). These enzymes utilize the Nδ of the lysine general base to activate the Oγ of the serine nucleophile, which then attacks the scissile carbonyl carbon of the substrate. The nucleophile attacks from the si-face of the scissile bond as opposed to the re-face, which is more commonly seen in classical serine proteases.

Yellowtail ascites virus (YAV) and YAV-like viruses are members of marine birnaviruses, which infects both fish and shellfish (16). YAV infection leads to ascites in yellowtail fish (Seriola quinqueradiata), which is popular in sushi, thus causing significant losses to the fish-farming industry (17). YAV-like viruses have been isolated from many fish species, including Japanese flounder (Paralichthys olivaceus), red sea bream (Pagrus major), and tiger puffer (Takifugu rubripes) (18–20). Marine birnavirus infections have been reported in shellfishes, such as the jackknife clam (Seriola quinqueradiata) and Japanese pearl oyster (Pinctada fucata) (21–23).

YAV VP4 protease is encoded as part of the 972-residue-long segment A polyprotein of YAV and cleaves at the C-terminal side of residues 508 and 734 to yield pVP2, VP4, and VP3 (24) (Fig. 1A). The cleavage sites at the pVP2/VP4 and VP4/VP3 junctions can be described by the consensus sequence Ser/Thr- X-Ala ↓ Ser-Gly (24). The serine-lysine catalytic dyad is shared

* This work was supported in part by the Canadian Institutes of Health Research (to M. P.), the National Science and Engineering Research Council of Canada (to M. P.), the Michael Smith Foundation for Health Research (to M. P.), and the Canadian Foundation of Innovation (to M. P.).

The atomic coordinates and structure factors (codes 4IZJ and 4IZK) have been deposited in the Protein Data Bank (http://wwwpdb.org).

1 To whom correspondence should be addressed: Dept. of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada. Tel.: 778-782-4230; Fax: 778-782-5583; E-mail: mpaetzel@sfu.ca.

2 The abbreviations used are: YAV, yellowtail ascites virus; CSO, 5-hydroxycysteine; PDB, Protein Data Bank.
among all VP4s in the family Birnaviridae. Sequence alignment analysis predicts Ser<sup>633</sup> to be the nucleophile and Lys<sup>674</sup> to be the general base in the YAV VP4 mechanism. Results from previous site-directed mutagenesis studies support the prediction of Lys<sup>674</sup> acting as the general base but fail to confirm Ser<sup>633</sup> as the nucleophile (24).

Crystal structures of VP4 protease from blotched snakehead virus (25), infectious pancreatic necrosis virus (26), and tellina virus 1 (27) have been elucidated previously. The blotched snakehead virus VP4 structure revealed a native active site with an empty binding groove. The crystal structures of infectious pancreatic necrosis virus VP4 showed an empty binding groove, intermolecular acyl-enzyme complexes, and product-bound complexes in mutant active sites. The tellina virus 1 structure presented an intramolecular acyl-enzyme complex in a native active site but with a sulfate bound near the lysine general base.

Here we have demonstrated that YAV contains an internal cleavage site within the VP4 region of the polyprotein. Cleavage at this site produces a truncated yet proteolytically active enzyme that ends at Ala<sup>716</sup>. In one of the YAV VP4 crystal structures presented here, the C-terminal region (residues 711–716) of each molecule in the asymmetric unit is bound in the active site of an adjacent molecule, either in the form of a trans product complex or a trans acyl-enzyme complex. The acyl-enzyme complex clearly demonstrates that Ser<sup>633</sup> acts as the nucleophile in the YAV VP4 hydrolytic reaction. This is the first VP4 structure to show an intermolecular acyl-enzyme complex and product complex in the presence of a native active site and native substrate. The second crystal structure presented here is that of the YAV VP4 active site mutant (K674A). There are two molecules in the asymmetric unit, with one forming an intermolecular acyl-enzyme complex and the other showing an empty active site. This structure directly reveals changes that occur within the enzyme upon substrate binding. These crystal structures of YAV VP4 will be of value in the design of anti-birnavirus compounds and provide insights into previous site-directed mutagenesis studies. In addition, these structures reveal a channel that runs from the surface of VP4 to the proposed deacylating water.

**EXPERIMENTAL PROCEDURES**

**YAV VP4 Constructs**—The cDNA for segment A of YAV was generously provided by Dr. Syunichirou Oshima. The full-length YAV VP4 construct (residues 509–734, from the pVP2/VP4 cleavage site to the VP4/VP3 cleavage site) was PCR-amplified using a forward primer with the sequence 5′-GGAGCTCATGGCAGCGGCAGCAGCACATGGTGTTCTCTCCGC-3′ and a reverse primer with the sequence 5′-CTGGCCCTGAGTGCACGACGTCCTTCTCTGATCCTCC-3′. Vent DNA polymerase (New England BioLabs) was used. The amplicon was cloned into the restriction sites NcoI and Xhol of plasmid pET28b<sup>+</sup> (Novagen) using T4 DNA ligase (Fermentas). The ligation mix was transformed into *E. coli* strain NovaBlue (Novagen) for plasmid isolation. This creates a VP4 construct that contains residues 509–734 from segment A of YAV, with two additional residues at the N terminus (Met-Ala) and eight additional residues at the C terminus (Leu-Glu and a His<sub>6</sub> affinity tag) (Fig. 1A). The truncated YAV VP4 construct was amplified using the same forward primer as mentioned above along with a reverse primer with sequence 5′-CTG GCC TCG AGT GCT TTC TGC ACT GGT AGT GCT CC-3′. This DNA region was amplified using the polymerase chain reaction with Pfu DNA polymerase (Fermentas). The amplicon was cloned into the restriction sites Ncol and Xhol of plasmid pET28b<sup>+</sup> (Novagen) using T4 DNA ligase (Fermentas). The ligation mix was transformed into *E. coli* strain NovaBlue (Novagen) for plasmid isolation. This creates a VP4 construct that contains residues 509–716 from segment A of YAV, with the same additional residues at the termini as described for the full-length construct. To generate the active site mutant (K674A) constructs, site-directed mutagenesis reactions using Pfu DNA polymerase (Fermentas) were performed using primers with the following sequences: 5′-TGC GGT GTA GAC ATC GCA GCC ATC GCC GCC CAT-3′ and 5′-ATG GGC GGC GAT GGC TGC GAT GTC TAC ACC GCA-3′. The NCBI reference sequence for segment A of YAV indicates an asparagine at position 616 (a surface residue), but the sequencing result indicates an aspartic acid residue at this position. We used site-directed mutagenesis to change the sequence to match the reference sequence (Asn<sup>616</sup>). PfuUltra<sup>TM</sup> polymerase (Stratagene) along with the primers of sequences 5′-AAA GAGATT ACC TGG AGC AAG AAC GGA AAC ATC GTG GTG-3′ and 5′-CAC GAT GTC TCC GTT CTT CTG CAT GTC TCT TT-3′ were used in the reaction. The sequences of all VP4 constructs were verified by DNA sequencing (GenBank<sup>™</sup> accession number NC_004168, UniProt accession number P89521).

**Protein Expression and Purification**—The expression vectors containing the YAV VP4 constructs were transformed into *E. coli* strain Tuner (DE3), followed by selection on Luria-Bertani (LB) agar plates supplemented with 50 μg/ml kanamycin. Six liters of cultures were grown for each batch of protein purified. A 10-ml aliquot of overnight culture was inoculated into each liter of LB/kanamycin medium. The cultures were incubated at 37 °C while shaking for 4 h before induction with 0.5 ml of 1 mM isopropyl β-D-thiogalactopyranoside. The induced cultures grew overnight at 25 °C. Cells were harvested by centrifugation at 9,110 × g for 7 min. To facilitate cell lysis, the cell pellet was stored at −80 °C for 15 min. The frozen cell pellets were then completely resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 10% glycerol, 1 mM dithiothreitol (DTT), 7 mM magnesium acetate, 0.1% Triton X-100, and 1 unit/ml benzamidine). The cells were sonicated three times at 30% amplitude for 5 s with a 10-s rest between each of the pulses (Fisher sonic dismembrator model 500). The sonicated cells were then lysed using an Avestin Emulsiflex-3C cell homogenizer (1,250 bars for 3 min). Cell debris was removed by centrifugation at 28,964 × g for 20 min. The resulting supernatant was applied to a 5–ml nickel-nitritolatriacetic acid metal affinity column (Qia-gen). The column was equilibrated with 5 column volumes of standard buffer (20 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10% glycerol, and 1 mM DTT). A step gradient containing 100, 300, and 600 mM imidazole in standard buffer was used to elute the His<sub>6</sub>-tagged protein. Fractions positive for VP4 were pooled, concentrated to 5 ml, and loaded onto a size exclusion chroma-
Structure of a Native Ser-Lys Protease trans Acyl-Enzyme

tography column (HiPrep 26/60 Sephacryl S-100 HR) equilibrated with crystallization buffer (20 mM Tris–HCl, pH 8.0, 100 mM NaCl, 10% glycerol, and 1% β-mercaptoethanol). The flow rate of the buffer was controlled by an ÄKTA PrimeTM system (GE Biosciences) set at 0.7 ml/min. Fractions with pure VP4 were pooled and concentrated to 30 mg/ml using a Millipore centrifugal filter with a molecular mass cut-off of 10 kDa. Protein concentration was measured with a Nanodrop UV spectrophotometer (Thermo Scientific) using an extinction coefficient of 9,970 m–1 cm–1 as calculated by the ProtParam server based on the amino acid sequence (28).

**YAV VP4 Full-length Self-cleavage Assay**—A 1.5-μl aliquot of truncated VP4 (YAV VP4 509–716 or YAV VP4 509–716 K674A, 30 mg/ml) was added to 120 μl of the reaction buffer (20 mM MES, pH 6.5; the same buffer and pH as the crystallization conditions). A 1.5-μl aliquot of full-length YAV VP4 (residues 509–734, 30 mg/ml) with a mutant active site (K674A) was added to this mixture as a substrate. The reaction was carried out at 23 °C and 20-μl aliquots were taken at times 0, 4, and 7 h and overnight. The samples were mixed with an equal volume of 2× sample buffer and loaded onto a 15% SDS-polyacrylamide gel. The gel was then stained with Coomassie Blue for visualization.

**Fluorometric Peptide Cleavage Assays**—The fluorometric peptide substrate benzoyloxycarbonyl-PVQKA-4-methylcoumar-7-amide was synthesized by LifeTein. The peptide was dissolved in 100% DMSO at a 50 mM stock concentration. Each 100-μl reaction contained 100 μM VP4 protease (YAV VP4 509–716 or YAV VP4 509–716 K674A) and 500 μM fluorometric peptide in reaction buffer (20 mM MES, pH 6.5). The control reaction contained the fluorometric peptide in reaction buffer. The assay was performed in triplicate on a 96-well black plate from Greiner Bio-One. The fluorescence (in relative fluorescence units) was monitored at 37 °C using a SpectraMax® M5 multimode microplate reader (Molecular Devices) with an excitation wavelength at 380 nm and emission wavelength at 460 nm. Data reduction was performed to overlay the graphs from three different conditions.

The YAV VP4 fluorometric peptide cleavage assay with and without DTT was performed with the substrate N,N-Dimethyl-p-aminobenzenesazobenzoic acid–KALPVQKAQG-ASE-5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid that was synthesized by ChinaPeptides. It contains residues P7–P4′ (residues 710–720) of the VP4 internal cleavage site. The assay was performed in triplicate on a 96-well μClear black plate from Greiner Bio-One. The fluorescence (in relative fluorescence units) was monitored at 37 °C using a SpectraMax® M5 multimode microplate reader (Molecular Devices) with an excitation wavelength at 340 nm and emission wavelength at 540 nm.

**Crystallization**—All crystallization trials were carried out at room temperature (~296 K). The crystals used in the diffraction studies were grown using the sitting-drop vapor diffusion method. For the native active site VP4, a 1-μl protein sample (30 mg/ml) was mixed with 1 μl of reservoir reagent and allowed to reach vapor equilibrium with 1 ml of reservoir reagent in a tape-sealed chamber. The initial crystallization condition was obtained from the Hampton Research Crystal Screen I condition 23 (30% PEG 400, 0.1 mM HEPES, pH 7.5, and 0.2 M magnesium chloride). The optimized condition gave rod-shaped crystals that grew out of precipitation in 35% PEG 2000, 0.1 mM MES, pH 6.5, and 0.3 M magnesium chloride. These monoclinic crystals belonged to space group P21; and had unit cell dimensions of 41.6 × 64.3 × 187.7 Å (β = 95.8°) with five molecules in the asymmetric unit (45% solvent, Matthews coefficient of 2.24). For the active site mutant, 1 μl of the native VP4 (30 mg/ml) was added to a 30-μl aliquot of the active site mutant (30 mg/ml) immediately prior to crystal plating. Three microliters of this VP4 mixture was mixed with 3 μl of the mother liquor and plated as described for the native active site VP4. The active site mutant crystals used for diffraction data collection grew out of 25% PEG 2000, 0.1 mM MES, pH 6.5, and 0.45 M magnesium chloride. These cubic crystals had unit cell dimensions of 273.5 × 273.5 × 273.5 Å with two molecules in the asymmetric unit (solvent content 74%, Matthews coefficient of 4.79) and belonged to space group F41,32.

**Data Collection**—Although the two crystal forms grew out of very similar crystallization conditions, the cryo-solvent conditions for the two crystals were significantly different. The cryosolution for the native active site VP4 crystals was 30% 2-methyl-2,4-pentanediol in mother liquor (35% PEG 2000, 0.1 mM MES, pH 6.5, and 0.3 M magnesium chloride), whereas 20% glycerol in the mother liquor (25% PEG 2000, 0.1 mM MES, pH 6.5, and 0.45 M magnesium chloride) was used for the active site mutant crystals. The crystal of the native active site VP4 was transferred into cryo-solution immediately prior to mounting in a loop with a copper base. Diffraction data were collected at the macromolecular x-ray diffraction data collection facility in the Molecular Biology and Biochemistry Department of Simon Fraser University, which includes a MicroMax-007 Microfocus x-ray generator, osmic confocal VariMax high flux optics, an R-AXIS IV++ image plate, and an X-stream 2000 cryosystem. The program CrystalClear was used for data collection. The crystal-to-detector distance was 220 mm. The crystal was exposed for 3 min with an oscillation angle of 0.5° and an x-ray wavelength of 1.5418 Å. The crystal of the active site mutant was transferred into cryo-solution and immediately mounted in a loop with a copper base and then flash-cooled with liquid nitrogen prior to transport to the Canadian Light Source for data collection at beamline 08ID-1. The crystal-to-detector distance was 310 mm. The crystal was exposed for 1 s with an oscillation angle of 4° and an x-ray wavelength of 0.97949 Å.

**Structure Solution and Refinement**—The indexing and integration for both data sets were carried out using MOSFLM (29), and scaling steps were performed with SCALA (30) from the CCP4i suite (31). The program POINTLESS was used to confirm the space group assignment (30). Molecular replacement was employed to obtain phase estimates. For the native active
site VP4, the solution was found using the program MOLREP (32) utilizing the infectious pancreatic necrosis virus VP4 structure (chain A of PDB entry 2PNM) as the search model. Rigid body and restrained refinement were performed using REFMAC5 (33–35). For the active site mutant, the coordinates from the native active site YAV VP4 were used as a molecular replacement search model in the program PHASER (36). Rigid body and restrained refinements were performed using REFMAC5 (33–35). Editing and fitting of atomic coordinates were performed using the program Coot (37).

**Structural Analysis**—The STRIDE server was used to assign secondary structure elements (38). The protein–protein interactions were analyzed using the PROTORP server (39). The channel diameter was calculated using CAVER 3.0 (40). Validation analysis was performed with Coot, Procheck, and the PDB validation server (37, 41–44).

**Figure Preparation**—PyMOL was used to generate the figures (45). The electron density map figures were prepared using REFMAC5 (33–35), the program FFT (46–48), and PyMOL. The tunnel figure was generated using the programs CAVER 3.0 (40) and PyMOL. The accessible surface area was calculated using the program Surface Racer (49).

**RESULTS AND DISCUSSION**

**Self-cleavage at the YAV VP4 Internal Cleavage Site**—Based on sequence similarity with infectious pancreatic necrosis virus VP4, it was previously proposed that YAV VP4 (Fig. 1, A and B) contains an internal cleavage site (between residues 716 and 717) preceding the VP4/VP3 cleavage site, which is between residues 734 and 735 (24). Consistent with this proposal, we have observed that a truncated YAV VP4 construct (residues 509–716) with a C-terminal His6 tag self-cleaved the tag upon incubation.

We were unable to isolate the full-length wild-type YAV VP4 enzymes (residues 509–734) due to a low expression level. However, the same construct with the active site mutation K674A was purified with a high yield. Therefore, we were able to test the ability of a truncated YAV VP4 with a native active site to cleave the full-length YAV VP4 protein at the internal cleavage site in a trans fashion. We found that the truncated YAV VP4 (residues 509–716) was active and fully processed the full-length YAV VP4 (residues 509–734, K674A) after an overnight incubation (Fig. 1C). In contrast, no processing was observed when the active site mutant (K674A) version of YAV VP4 509–716 was used in the reaction (Fig. 1C). In addition, we used a fluorescent peptide assay to show that the YAV VP4 truncated at the internal cleavage site (residues 509–716) is active. The peptide sequence used (PVQKA) within the substrate corresponds to residues P5–P1 of the internal cleavage site (residues 712–716 of the YAV polyprotein). This peptide is modified at the C terminus with a 4-methylcoumaryl-7-amide group and C-terminal residue, the fluorophore 7-amino-4-methylcoumain is released, which results in a fluorescence emission signal at 460 nm when excited at 380 nm (50). We found that the truncated YAV VP4 (residues 509–716) was active because it recognized and cleaved at the internal cleavage site sequence (Fig. 1D). In contrast, no activity was observed when the active enzyme was substituted with either the truncated active site mutant (residues 509–716, K674A) or the reaction buffer control.

**Self-cleavage at the YAV VP4 Internal Cleavage Site Promotes Crystallization**—As described above, the native active site version of YAV VP4 with a C-terminal affinity tag slowly cleaved itself after Ala716, the P1 residue for the internal cleavage site. The resulting processed protein produced monoclinic crystals.
with five molecules in the asymmetric unit that diffracted to 2.5 Å resolution. An active site mutant version (K674A) of YAV VP4 was produced to add to the native active site structure. The five molecules in the asymmetric unit that diffracted to 2.3 Å resolution. Interestingly, crystals of the active site mutant only appeared when a small amount of native VP4 was added to the protein fold. The refined crystal structures display electron density for residues 515–716 in each molecule of the asymmetric unit. The crystallographic statistics and validation are listed in Table 1.

### Table 1: Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>Native active site, PDB 4IZJ</th>
<th>Mutant active site (K674A), PDB 4IZK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crystal parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P12,1</td>
<td>F4,32</td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>41.6, 64.3, 187.7</td>
<td>273.5, 273.5, 273.5</td>
</tr>
<tr>
<td>β (degrees)</td>
<td>95.8</td>
<td></td>
</tr>
<tr>
<td><strong>Data collection statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.5418</td>
<td>0.9795</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>64.3–2.5 (2.6–2.5)</td>
<td>55.8–2.3 (2.4–2.3)</td>
</tr>
<tr>
<td>Total reflections</td>
<td>122,204 (13.215)</td>
<td>343,613 (45.664)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>33,009 (4,043)</td>
<td>39,388 (5.637)</td>
</tr>
<tr>
<td>Rmerge (all I and I—)</td>
<td>0.088 (0.278)</td>
<td>0.105 (0.249)</td>
</tr>
<tr>
<td>Mean(I)/σ(I)</td>
<td>10.5 (3.8)</td>
<td>13.8 (6.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95.9 (80.7)</td>
<td>100.0 (100.0)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.7 (3.3)</td>
<td>8.7 (8.1)</td>
</tr>
<tr>
<td><strong>Refinement statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein molecules (chains)</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Residues</td>
<td>1014</td>
<td>405</td>
</tr>
<tr>
<td>Water molecules</td>
<td>271</td>
<td>320</td>
</tr>
<tr>
<td>Total number of atoms</td>
<td>7,898</td>
<td>3,356</td>
</tr>
<tr>
<td>Rmerge (all I and I—)</td>
<td>17.8/24.4</td>
<td>17.3/19.7</td>
</tr>
<tr>
<td>Average B-factor (Å²)</td>
<td>26.8</td>
<td>35.2</td>
</tr>
<tr>
<td>Wilson plot estimated B-factor (Å²)</td>
<td>40.0</td>
<td>36.3</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>0.94</td>
<td>0.96</td>
</tr>
<tr>
<td>Root mean square deviation, angles (degrees)</td>
<td>0.900</td>
<td>1.428</td>
</tr>
<tr>
<td>Root mean square deviation, bonds (Å)</td>
<td>0.007</td>
<td>0.012</td>
</tr>
</tbody>
</table>

### Overall Architecture and Active Site of YAV VP4 Protease—Consistent with the theoretical isoelectric point of 4.6, the surface of YAV VP4 is predominantly negatively charged. The protease has an α/β protein fold consisting of 13 β-strands, four α-helices, and two β-sheet flanked by three α-helices (α2, α3, and α4) and a β-hairpin (β9 and β10). This region hosts the catalytic dyad with the serine nucleophile (Ser633) found immediately preceding α-helix 2 (α2) and the lysine general base (Lys674) arriving from α-helix 3 (α3). There is clear electron density for the Lys674 side chain in all five molecules in the asymmetric unit of the native active site structure. The Nζ for Lys674 has an average B-factor of 21.3 Å². The Nζ for Lys674 is highly coordinated by the Oγ1 of Thr655, which resides between β-strands 8 and 9, the Oγ of
Ser633 (the nucleophile), the main chain carbonyl oxygen of Cys669, the carboxylate oxygen of Ala716 (C-terminal residue for the product bound complexes), and a buried water (a proposed deacylating water). For the acyl-enzyme structures (chains C and D, Figs. 3A and 4A) the ester carbonyl oxygen (Ala716 O) is pointed toward the oxyanion hole (Ser633 NH), away from the general base lysine N/H9256. The buried water, which we propose is deacylating, is coordinated by: Thr655 O/H9253, Ser633 O/H9253, Lys674 N/H9256, and Pro544 O.

Cysteine Residues—YAV VP4 contains five cysteine residues at positions 588, 604, 634, 669, and 688. VP4 from other Birnavirus species have on average just two cysteine residues. Residues 588 and 604 form a disulfide bridge (Cys588-Cys604) that links β-strands 5 and 6 (Fig. 2A). This is the first disulfide bond observed within a VP4 protease. Despite being purified and crystallized in the presence of reducing agents, some of the molecules in the asymmetric unit for each structure show electron density off the thiol group of Cys634 and Cys669, which is consistent with them being oxidized to S-hydroxycysteine (CSO), a sulfenic acid. Conversely, the electron density for Cys688 always appears unmodified. The modification state for each cysteine is summarized in Table 2.

Different C-terminal Conformations—In the native active site structure, each protein chain in the asymmetric unit has a C terminus that is in an extended conformation (Fig. 2A). In contrast, the active site mutant (K674A) structure reveals two sep-
arate C-terminal conformations, one that is extended and another where the C terminus packs against the surface of the enzyme (Fig. 2B).

**trans Acyl-Enzyme and Enzyme-Product Complexes in a Native Active Site**—In the asymmetric unit of the native active site structure, each substrate binding groove is occupied by the last six residues (residues 711–716) of the adjacent molecule, thus stabilizing the extended conformation for each C terminus (Figs. 2A, 3A, and 4A). The average buried surface area between the enzyme and the neighboring molecules is 894 Å$^2$. Analysis of the electron density reveals that two of the five VP4 molecules in the asymmetric unit form acyl-enzyme complexes with

<table>
<thead>
<tr>
<th>TABLE 2 Cysteine modification in YAV VP4</th>
<th>Native active site structure</th>
<th>Mutant active site structure (K674A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chain A (empty)</td>
<td>Chain B (acyl-enzyme)</td>
<td></td>
</tr>
<tr>
<td>Cys$^{614}$</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Cys$^{669}$</td>
<td>×</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 4. trans acyl-enzyme complexes and trans enzyme-product complexes revealed in a crystal structure of YAV VP4 protease with a native active site.** A, the acyl-enzyme complex is shown with the substrate colored in cyan (molecule E) and the enzyme colored orange (molecule D). B, the enzyme-product complex is shown with the product colored in magenta (molecule A) and the enzyme colored in cyan (molecule E). The surface diagrams for the respective molecules are shown in the left-hand corner. Water molecules are shown as cyan spheres. The hydrogen bonds are shown as dashed lines. The black-colored hydrogen bonds represent those that are conserved in each molecule in the asymmetric unit. The red-colored hydrogen bonds are those that vary from molecule to molecule in the asymmetric unit, depending on the acylation state. The hydrogen bonding distances are given in Å.
neighboring VP4 molecules (Fig. 3A). The Oγ of the serine nucleophile (Ser633) is covalently linked to the carbonyl carbon of Ala716, the P1 residue for the internal cleavage site. The electron density is consistent with a trigonal planar geometry that is expected for an ester bond (Fig. 3A).

The P6–P1 residues (residues 711–716) are stabilized by hydrogen bonding interactions in an anti-parallel fashion with residues 546–550 on one side of the binding groove and interact in a parallel fashion with residues 627–630 on the other side of the binding groove (Fig. 4A). Five of the hydrogen bonds involved in linking the C terminus of one molecule to the binding groove of the neighboring molecule are conserved in both the native active structure and the mutant active site structure. These are Thr548 O–Val713 N, Thr548 N–Val713 O, Gly627 N–Pro712 O, Pro628 O–Gln714 N (via water), and Met630 N–Gln714 O (depicted in black dotted lines in Fig. 4). The acyl-enzyme complexes reveal hydrogen-bonding interactions between the main chain of the P2 and P1 residues of the internal cleavage site and residues that line the binding groove near the catalytic residues. These hydrogen bonds are Ala716 O–Ser633 N, Ala716 N–Met630 O, Lys715 O–Val546 N, and Lys715 N–Val546 O (Fig. 4). The hydrogen bond between Ala716 O (P1 residue) and Ser633 N (main chain nitrogen of the nucleophile) is the only hydrogen bond observed in the oxynion hole. There are no other hydrogen bond donors near the carbonyl oxygen of Ala716. It is possible that the positive dipole of α-helix 2, on which the nucleophilic Ser633 resides, contributes to oxynion stabilization.

Several water molecules participate in the hydrogen bonding network within the substrate binding groove. The water that bridges the main chain amide nitrogen of Gln714 (P3) to the carbonyl oxygen of Pro628 (Fig. 4) is seen in each of the observed reaction steps. A water that bridges Gln550 N to Leu711 O is observed in all of the acyl-enzyme complexes captured. In addition, chain B, which is involved in the product complex also has this water. In some of the acyl-enzyme and enzyme-product complexes, a water molecule bridges the Oγ of Thr548 to the carbonyl oxygen of Val546 and the amide nitrogen of Lys715 (Fig. 4). The most interesting water seen in the enzyme-product complexes is the one that resides within the oxynion hole (Fig. 4B). This water, which is displaced in the acyl-enzyme complex, makes hydrogen bonding interactions with the carbonylate oxygen of Ala716 (P1 residue), the Oγ and main chain nitrogen of Ser633 (nucleophile), and the main chain carbonyl oxygen of Met630 (Fig. 4B).

The major substrate specificity pockets are the S1 and S3 (Figs. 4–6). Residues Ile633, Pro634, Val635, Leu636, Ile636, Ile692, Ala629, Met630, Gly631, Pro692, Ser633, and Cys634 form a continuous surface (the S1 pocket) for the binding of the P1 residue side chain (Ala716). Residues Val635, His634, Ser635, Leu660, Leu661, Gln676, Gln676, Gly691, Ala697, Pro690, Val724, Ala626, Gly627, Pro628, and Ala629 form a continuous surface (the S3 pocket) for the binding of the P3 residue side chain (Gln714). A cleft that is open on one side accommodates the P5 residue (Pro712). The side chain of Leu711 (the P6 residue) points into a shallow pocket near the end of the substrate binding groove. There are no binding pockets for the P2 and P4 residues because their side chains are pointing toward the solvent, away from the substrate binding groove.

Substrate Binding Groove, Empty versus Bound—The crystal structure of the active site mutant of YAV VP4 (K674A) reveals an acyl-enzyme complex and an empty active site in the same asymmetric unit (Fig. 3B). This gives us the opportunity to observe structural differences within the substrate binding groove during these distinctly different enzyme states. One difference is seen in the loop, between β-strands 2 and 3, that resides adjacent to the substrate binding groove (Fig. 2). In the absence of a bound peptide (P6–P1 from the neighboring molecule), this loop is more disordered, with clear density observed only for the main chain atoms. With bound peptide, this turn becomes ordered and shows clear density for both main chain and side chain atoms. Another clear difference between the bound and unbound states of VP4 is seen within the S3 specificity binding pocket. The side chains of His547 and Val545 both show a change in rotamer conformation, which results in a larger S3 binding pocket in the bound state (Fig. 6). The main chain atoms that make up the rim of the S3 pocket have also moved to widen the opening in the bound state.

Aqueous Channel Leads to Active Site—Interestingly, the S1 binding pocket in YAV VP4 is not fully enclosed within the acyl-enzyme or product complex structures; rather it is open to the surface of the enzyme via a channel (Fig. 7). This fenestration leads from the protein surface to an area near the proposed deacylating water within the active site. The channel is made up of atoms from the following residues: Ile592, Glu594, Asp595, Ile660, Pro692, Ala629, Met630, Gly631, Pro692, Cso634 or Cys634, Gly635, and Leu638. The interior of the channel is filled with hydrogen bond donors and acceptors. In both the native active site and mutant active site structures, ordered water molecules are coordinated at and near the entrance of the channel by Glu594, Asp595, Ala629, Gly631, and Gly635. In the active site mutant structure, a water molecule is also found near the narrowest region of the channel (bottleneck) in the VP4 with the empty binding groove. However, such a water is absent in the acyl-enzyme-containing VP4. Superposition of these molecules reveals no significant change in the position of the residues that form the channel. The average diameter at the narrowest point (bottleneck diameter) in the channel for the five molecules in the native active site structure is 2.1 Å. The residues that form the narrowest point of the channel are Cys634 or Cso634, Ile592, Ala629, Met630, and Gly631 (Fig. 8). Although the diameter of a water molecule is ~2.8 Å, there are examples of water channels with narrow passageways of approximately the same dimensions as seen in the VP4 channel. For example, a selectivity filter of around 2 Å in diameter has been reported in E. coli aquaporin Z, a water channel protein (51, 52). It is possible that a deacylating (catalytic) water could utilize such a channel to position itself in the correct trajectory for an attack on the ester bond of the acyl-enzyme.

Ser633 is the YAV VP4 Nucleophile, but Could Thr655 Oγ Function as a Nucleophile in the Absence of Ser633 Oγ?—The serine/lysine catalytic dyad mechanism is conserved among all birnavirus VP4 proteases studied thus far. Therefore, Ser633 and Lys674 were predicted to be the nucleophile and general base, respectively, in YAV VP4 based on sequence alignments.
Consistent with these predictions, the processing of the YAV segment A polyprotein was completely abolished when Lys$^{674}$, the proposed general base, was mutated to an aspartic acid in a previous mutagenesis study (24). We show here that mutating Lys$^{674}$ to alanine also produces an inactive enzyme. However, previous mutagenesis experiments did not conclusively show that Ser$^{633}$ plays the role of the nucleophile. Both pVP2-VP4 and VP4-VP3 fragments from the YAV segment A polyprotein were detected when Ser$^{633}$ was mutated to a proline. Interestingly, similar results were also seen in infectious bursal disease virus mutagenesis experiments (53, 54).

Our crystal structures of the YAV VP4 trans acyl-enzyme complex clearly support the initial hypothesis of Ser$^{633}$ being the nucleophile and Lys$^{674}$ serving as the general base. The acyl-enzyme complexes directly reveal electron density linking the $\text{O}\text{H}_{9253}$ of Ser$^{633}$ to the carbonyl carbon of Ala$^{716}$ (the P1 residue of the internal cleavage site) (Fig. 3).

Besides the nucleophilic Ser$^{633}$, Thr$^{655}$ is the only titratable residue within the vicinity of the general base Lys$^{674}$ (Figs. 3 and 4). A serine or threonine is observed in a similar position in all other known proteases utilizing a Ser/Lys catalytic dyad mechanism (55). If it were possible for the hydroxyl of Thr$^{655}$ to function as the nucleophile in the absence of a Ser$^{633}$ hydroxyl, this would put the attack on the scissile bond in the opposite
direction (re- rather than si-). This would probably affect the deacylation step of the reaction, resulting in a single cleavage event, thus providing a possible explanation for the presence of the pVP2-VP4 and VP4-VP3 precursors in the S633P mutant experiment.

Interpreting Non-active Site Residue Mutagenesis Results—Site-directed mutagenesis experiments show that changes to YAV VP4 non-active site residues can affect polyprotein processing (24). A cleavage product that was slightly smaller than the YAV segment A polyprotein was observed when Ile543 was mutated to a glycine. The side chain of Ile543 makes up part of the S1 binding pocket (Fig. 9). Replacing Ile543 with a smaller residue like glycine would enlarge the S1 pocket and allow bulkier side chains to bind. Because S1 and S3 are the major specificity pockets present in the binding groove, such a mutation might significantly alter the substrate specificity. Thus, the enzyme is likely to be functional but with the ability to cleave the polyprotein at more locations. In a V686Q mutant, both the polyprotein and pVP2 were observed (24). This suggested that the cleavage at the VP4/VP3 junction was affected. Residue Val686 does not lie in the active site, nor does it reside near the substrate specificity pockets. Instead, it is located adjacent to the C-terminal α-helix, with its main chain carbonyl oxygen forming a hydrogen bond with the amide nitrogen of Thr655, a residue that coordinates the Nc of the general base (Fig. 9). The main chain interaction at this position is observed in all other VP4 structures solved thus far (25–27). Because this mutant still generated pVP2 by cleaving at the VP2/VP4 junction, it is unlikely that it affected the catalytic dyad. Amino acid sequence alignment analysis revealed that uncharged residues, such as valine, isoleucine, leucine, and alanine, are found at this position in other birnaviruses. Replacement of Val686 with a more polar residue like glutamine could affect the packing around the C terminus, the substrate at the VP4/VP3 junction. Moreover, Leu705, which resides on the conserved α-helix 4 near the C terminus, is located across from Val686. The packing interactions of this last α-helix are likely to be required for the proper presentation of the VP4/VP3 cleavage site.

 cis versus trans Cleavage at the YAV VP4 Internal Cleavage Site—All of the molecules in the crystal structures presented here reveal intermolecular (trans) associations between the cleavage site and active site, but is this what happens in solution? Sequence and structural analysis suggests that in solution, the internal cleavage site processing probably only occurs via an intermolecular (trans) reaction. The previous tellina virus 1 VP4 structure, which does not have an internal cleavage site, reveals that the VP4/VP3 cleavage site (position 734/735 in the YAV polyprotein sequence) is just long enough to form the intramolecular (cis) acyl-enzyme complex (27). Therefore, the

FIGURE 6. Empty active site versus acyl-enzyme; changes observed in the substrate specificity binding pockets. The active site mutant of YAV protease crystallized with two molecules in the asymmetric unit: one molecule in the bound state (acyl-enzyme) and the other with an empty active site. A, a superposition of the residues that make up the substrate binding groove of YAV VP4 protease in the empty (black) and bound states (salmon) is shown. The positions of the S1 and S3 binding pockets are denoted by open black circles. The solvent-accessible surfaces for the empty (B) and the bound states (C) of the VP4 substrate binding pockets are shown.

FIGURE 7. A channel leads to the active site of YAV VP4 protease. A, a channel (red surface), found adjacent to the substrate binding groove, leads from the enzyme surface to a region near the catalytic residues (shown in a stick representation) and the proposed deacylating water. B, the residues surrounding the channel are shown in copper color. The neighboring protein chain forming a product complex is shown in a magenta schematic. C, a close-up view of the channel with the S1 and S3 pockets circled in orange. D, a cross-section view of the channel reveals that it is continuous with the S1 binding pocket (orange circle) and the pocket where the deacylating water resides (pink).
internal cleavage site (position 716/717) in YAV would not be accessible for an intramolecular (cis) cleavage event without significant unfolding of the secondary structure. The intermolecular processing of the internal cleavage site is further supported by our trans-cleavage assay of the full-length YAV VP4 and our observation that the active site mutant enzyme (K674A) would not crystallize until a small amount of the active native active site enzyme was added to cleave the inactive mutant at the internal cleavage site.

How do these intermolecular acyl-enzyme complexes in VP4 protease form, and why are they stable enough to be seen in the crystal structure? With the mutant enzyme, we had to add a small amount of native active site enzyme to cleave at the internal cleavage site before crystals would form. This suggests that the internal cleavage site product was presented to the VP4 active site, driving the reaction in the reverse direction to form the acyl-enzyme. With the general base absent, there is no suitable deprotonated functional group in the vicinity to activate the would-be deacylating water, thus stabilizing the acyl-enzyme complex. It is not so easy to explain how the acyl-enzyme is stabilized in the native active site enzyme. The crystals were formed at pH 6.5, which is probably well below the pKₐ of the lysine general base; thus, the low pH could be slowing down the deacylation step. It is also possible that the internal cleavage site plays a regulatory role in the VP4 protease. The glutamine found at the P3 position of the internal cleavage site is significantly longer and more polar than the residue seen at the P3 position of the pVP2/VP4 and VP4/VP3 cleavage sites (Ser⁵⁰⁶ and Thr⁷³², respectively). This allows Gln⁷¹⁴ to make multiple hydrogen bonding interactions with residues at the bottom of the S3 pocket. It is possible that these additional interactions slow down the release of the internal cleavage site from the VP4...
binding groove, thereby acting as an internal inhibitor that functions after the main cleavage sites are processed. Interestingly, there have been previous reports of VP4 assembling into tubules (56).

**Can Cysteine Modification Have an Effect on YAV VP4 Activity?**—The thiol side chain of cysteine can be oxidized by biological oxidants to form CSO, a sulfenic acid. Increasingly, there have been previous reports of VP4 assembling into tubules (56).

**Structure of a Native Ser-Lys Protease trans Acyl-Enzyme**

In this study, we have shown that YAV VP4 contains an internal cleavage site (position 716/717) near its C terminus (residue 734). We have shown that the resulting truncated VP4 is proteolytically active. We present for the first time the structure of a serine-lysine protease trans acyl-enzyme complex with a native substrate in a native active site. The acyl-enzyme structures directly prove the role of Ser633 as the nucleophile in the YAV VP4 reaction mechanism. A structure of a YAV VP4 mutant with the lysine general base changed to alanine is also presented. Together, these two structures reveal the substrate binding site for this protease in three separate stages of the protease reaction cycle: empty active site, acyl-enzyme complex, and product complex. Analysis of the two structures has helped us to provide insights into previous mutagenesis results and will be of value in the design of anti-birnavirus compounds and in the basic understanding of the serine-lysine dyad catalytic mechanism. We have discovered that this protease contains a channel that leads from the enzyme surface to the proposed decaying water.

**Acknowledgments**—We thank the staff at beam line 08ID-1 at the Canadian Light Source, Saskatoon, Canada (especially Shaunivan Labiuk) for technical assistance with data collection. The Canadian Light Source is supported by the National Science and Engineering Research Council, National Research Council, Canadian Institutes of Health Research, and University of Saskatchewan. We thank Dr. Syunichirou Oshima for providing the YAV cDNA.
Structure of a Native Ser-Lys Protease trans Acyl-Enzyme

REFERENCES