THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 18, pp. 13068–13081, May 3, 2013

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**

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

Ivy Yeuk Wah Chung and Mark Paetzel[1]

From the Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

Background: YAV VP4 is a Ser/Lys dyad protease that processes the birnavirus polyprotein: pVP2-VP4-VP3.
Results: An acyl-enzyme is formed between Ser^{633} (nucleophile) and Ala^{716}, confirming the existence of an internal cleavage site.
Conclusion: VP4 terminating at Ala^{716} is active, and Ser^{633} serves as the nucleophile.
Significance: Structural insights into three stages of the VP4 reaction will promote anti-birnavirus compound development.

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 Ala^{716}) 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 *trans* product complexes and *trans* acyl-enzyme complexes. The acyl-enzyme complexes confirm directly the role of Ser^{633} 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 NH_2-pVP2-VP4-VP3-COOH (Fig. 1*A*). An exception is found in blotched snakehead 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 endoprotease 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 Nζ 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)[2] and YAV-like viruses are members of marine birnavirus, 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 (*Sinonovacula constricta*) 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. 1*A*). 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://wwpdb.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, *S*-hydroxycysteine; PDB, Protein Data Bank.

among all VP4s in the family Birnaviridae. Sequence alignment analysis predicts Ser^{633} to be the nucleophile and Lys^{674} to be the general base in the YAV VP4 mechanism. Results from previous site-directed mutagenesis studies support the prediction of Lys^{674} acting as the general base but fail to confirm Ser^{633} 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^{716}. 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^{633} 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 the serine-lysine dyad catalytic mechanism. These structures also 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′-GGA CTC CAT GGC CAG CGG CAC AGA CAC TGG G-3′ and a reverse primer with the sequence 5′-CTG GCC TCG AGT GCA GTT GTT CTC ATT AGT TCC CC-3′. Vent DNA polymerase (New England BioLabs) was used. The amplicon was cloned into the restriction sites NcoI and XhoI of plasmid pET28b$^+$ (Novagen) using T4 DNA ligase (Fermentas). The ligation mix was transformed into *Escherichia coli* strain *NovaBlue* (Novagen) for plasmid isolation. This creates a full-length 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_6 affinity tag) (Fig. 1*A*). 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 NcoI and XhoI of plasmid pET28b$^+$ (Novagen) using T4 DNA ligase (Fermentas). The ligation mix was transformed into *E. coli* strain *NovaBlue* (Novagen) for plasmid isolation. This creates a VP4 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^{616}). PfuUltra^{TM} polymerase (Stratagene) along with the primers of sequences 5′-AAA GAG ATC AAG AAG AAC GGA AAC ATC GTG GTG-3′ and 5′-CAC CAC GAT GTT TCC GTT CTT CTT GAT CTC TTT-3′ were used in the reaction. The sequences of all VP4 constructs were verified by DNA sequencing (GenBank^{TM} 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 M isopropyl β-D-1-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 buffer, pH 8.0, 10% glycerol, 1 mM dithiothreitol (DTT), 7 mM magnesium acetate, 0.1% Triton X-100, and 1 unit/ml benzonase). 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-nitrilotriacetic acid metal affinity column (Qiagen). 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_6-tagged protein. Fractions positive for VP4 were pooled, concentrated to 5 ml, and loaded onto a size exclusion chroma-

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 Prime™ 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 benzyloxycarbonyl-PVQKA-4-methylcoumaryl-7-amide was synthesized by LifeTein. The peptide was dissolved in 100% DMSO at a 50 mM stock concentration. Every reaction had a final DMSO concentration of 1% (v/v). 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 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 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-aminobenzeneazobenzoic 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 peptide was dissolved in 100% DMSO at a 50 mM stock concentration, and every reaction had a final DMSO concentration of 1% (v/v). Each 100-μl reaction contained 67 μM VP4 protease (YAV VP4 509–716) and 500 μM fluorometric peptide in reaction buffer (20 mM MES, pH 6.5) with or without 10 mM DTT. In the control reaction, the native active site enzyme was replaced with the inactive mutant (YAV VP4 509–716, K674A). The assay was run 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 M 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 M MES, pH 6.5, and 0.3 M magnesium chloride. These monoclinic crystals belonged to space group $P12_11$ 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 M 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 $F4_132$.

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 cryo-solution for the native active site VP4 crystals was 30% 2-methyl-2,4-pentanediol in mother liquor (35% PEG 2000, 0.1 M MES, pH 6.5, and 0.3 M magnesium chloride), whereas 20% glycerol in the mother liquor (25% PEG 2000, 0.1 M 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 0.4° 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 His_6 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. 1*C*). In contrast, no processing was observed when the active site mutant (K674A) version of YAV VP4 509–716 was used in the reaction (Fig. 1*C*). In addition, we used a fluorometric 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. Upon cleavage of the amide bond between the 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. 1*D*). In contrast, no activity was

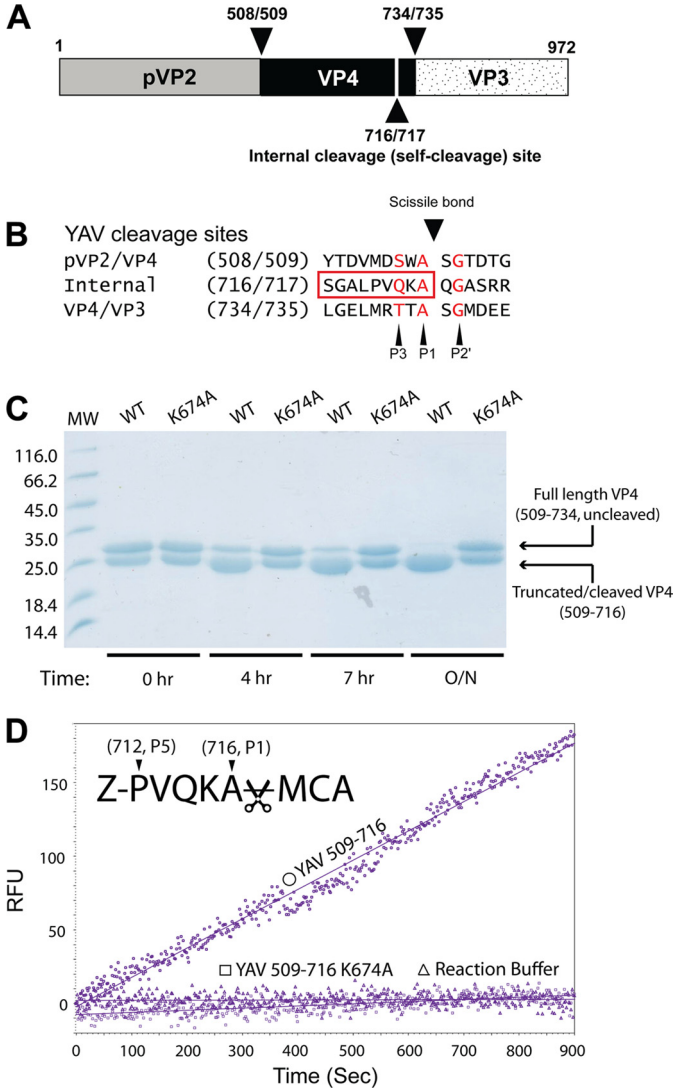

FIGURE 1. **YAV VP4 protease truncated at the internal cleavage site is active.** *A*, the genomic arrangement of segment A in YAV is shown. The major sites of cleavage are denoted by *arrowheads* with the P1 and P1′ residue numbers listed. *B*, a list of known major YAV VP4 protease cleavage sites. The residues preceding P1′ in the internal cleavage site are visible in the crystal structures and are *boxed* in *red*. *C*, to demonstrate processing at the internal cleavage site (after residue 716), truncated YAV VP4 (residues 509–716) was incubated with full-length YAV VP4 with an active site mutation (residues 509–734, K674A) for 0, 4, and 7 h and overnight at 23 °C. As a control, the same experiment was repeated with an enzyme with an active site mutation (residues 509–716, K674A). A 20-μl aliquot of the reaction was run on an SDS-polyacrylamide gel and visualized by Coomassie stain. *D*, a fluorometric peptide (benzyloxycarbonyl-PVQKA-4-methylcoumaryl-7-amide (*Z-PVQKA-MCA*)) corresponding to the sequence for residues P5–P1 of the YAV VP4 internal cleavage site (residues 712–716 of the YAV polyprotein) was incubated with truncated YAV VP4 (○; residues 509–716), truncated YAV VP4 active site mutant (□; residues 509–716, K674A) and reaction buffer (△; 20 mM MES, pH 6.5) alone. The relative fluorescence units (*RFU*; reduced values) are plotted over time.

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 Ala^{716}, the P1 residue for the internal cleavage site. The resulting processed protein produced monoclinic crystals

TABLE 1
Data collection and refinement statistics

	Native active site, PDB 4IZJ	Mutant active site (K674A), PDB 4IZK
Crystal parameters		
Space group	$P12_11$	$F4_132$
a, *b*, *c* (Å)	41.6, 64.3, 187.7	273.5, 273.5, 273.5
β (degrees)	95.8	
Data collection statistics		
Wavelength (Å)	1.5418	0.9795
Resolution (Å)	64.3-2.5 (2.6-2.5)[a]	55.8-2.3 (2.4-2.3)[a]
Total reflections	122,204 (13,215)	343,613 (45,664)
Unique reflections	33,009 (4,043)	39,388 (5,637)
R_{merge}[b]	0.088 (0.278)	0.105 (0.249)
R_{pim}[c] (all I+ and I−)	0.052 (0.178)	0.037 (0.093)
Mean $(I)/\sigma(I)$	10.5 (3.8)	13.8 (6.5)
Completeness (%)	95.9 (80.7)	100.0 (100.0)
Redundancy	3.7 (3.3)	8.7 (8.1)
Refinement statistics		
Protein molecules (chains) in asymmetric unit	5	2
Residues	1014	405
Water molecules	271	320
Total number of atoms	7,898	3,356
R_{cryst}[d]/R_{free}[e] (%)	17.8/24.4	17.3/19.7
Average *B*-factor (Å^2) (all atoms)	26.8	35.2
Wilson plot estimated *B*-factor (Å^2)[f]	40.0	36.3
Correlation coefficient[g]	0.94	0.96
Root mean square deviation, angles (degrees)	0.900	1.428
Root mean square deviation, bonds (Å)	0.007	0.012
Ramachandran plot[h]		
Core region, no. of residues (%)	887 (87.5)	367 (90.6)
Allowed region, no. of residues (%)	127 (12.5)	38 (9.4)
Outliers, no. of residues (%)	0 (0)	0 (0)

[a] The data collection statistics in parentheses are the values for the highest resolution shell.
[b] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual reflection, and $\langle I(hkl)\rangle$ is the mean intensity of that reflection.
[c] R_{pim} (precision-indicating merge) $= \Sigma_{hkl}(1/n_{hkl} - 1)^{1/2}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where n is the number of observed hkl reflections.
[d] $R_{cryst} = \Sigma_{hkl}||F_o| - |F_c||/\Sigma_{hkl}|F_o|$, where F_o and F_c are the observed and calculated structure-factor amplitude, respectively.
[e] R_{free} is calculated using 5% of the reflections randomly excluded from refinement.
[f] Wilson plot estimated *B*-factor was calculated using program Truncate from the CCP4 program suite (31, 61).
[g] Correlation coefficient was calculated using Refmac5 from the CCP4 program suite (31, 35).
[h] Ramachandran plot was calculated using Procheck from the CCP4 program suite (41).

with five molecules in the asymmetric unit that diffracted to 2.5 Å resolution. An active site mutant version (K674A) of YAV VP4 produced cubic crystals with two 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 promote cleavage at Ala^{716}. 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.

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 3_{10}-helices (Fig. 2*A*). β-Strands 1–7 form a platform on which the binding pockets for substrate recognition are constructed. Most of these β-strands interact in an anti-parallel fashion. Two 3_{10}-helices (η1 and η2) reside on a loop between β3 and β4. Strands β8, β11, and β12 form a parallel 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 (Ser^{633}) found immediately preceding α-helix 2 (α2) and the lysine general base (Lys^{674}) arriving from α-helix 3 (α3). There is clear electron density for the Lys^{674} side chain in all five molecules in the asymmetric unit of the native active site structure. The Nζ for Lys^{674} has an average *B*-factor of 21.3 Å^2. The Nζ for Lys^{674} is partially buried and has an average accessible surface area of 6.6 Å^2. The average accessible surface area for the nine other lysine Nζ atoms in chain A is 50.6 Å^2. The Nζ of Lys^{674} is highly coordinated by the Oγ1 of Thr^{655}, which resides between β-strands 8 and 9, the Oγ of

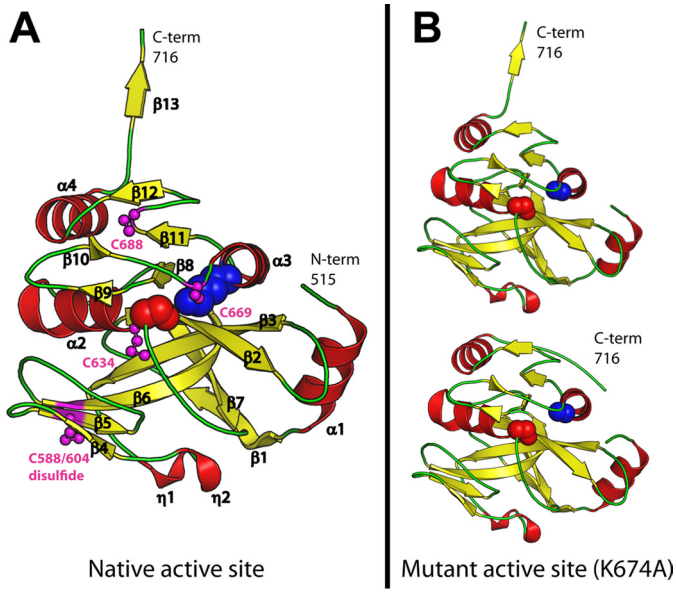

FIGURE 2. **The YAV VP4 protein fold.** *A*, the protein fold for a representative protein chain in the native active site structure. The β-strands are depicted as *yellow arrows*, and α-helices are shown as *red coils*. The catalytic active site residues Ser^{633} (*red*) and Lys^{674} (*blue*) are shown as *spheres*. The cysteine residues are shown as *magenta ball-and-stick models*. *B*, the protein fold for the two different VP4 conformations observed in the mutant active site (K674A) structure. The C terminus is labeled.

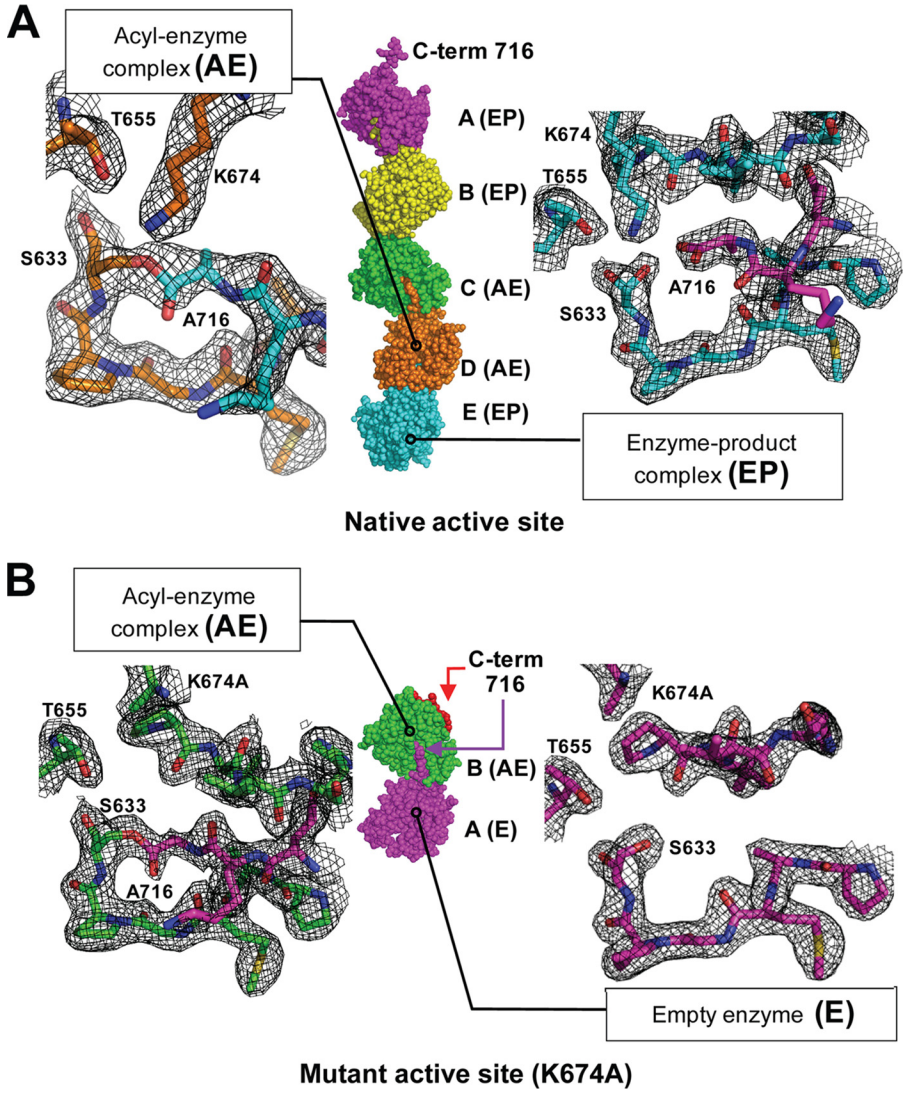

FIGURE 3. **The YAV VP4 protease intermolecular (*trans*) acyl-enzyme complex, product complex, and empty active site.** YAV VP4 protease is able to cleave at an internal cleavage site (after Ala^{716}) near its own C terminus (after Ala^{734}), the VP4/VP3 junction. Two crystal structures of YAV VP4 (with and without an active site mutation) reveal the internal cleavage site bound within the active site of a neighboring VP4 molecule. Analysis of the electron density within each molecule of the asymmetric unit reveals three different enzyme states of the protease reaction cycle (empty active site, acyl-enzyme, and product complex). The packing of YAV VP4 protease in the asymmetric unit for each structure is shown along with representative $2F_o - F_c$ electron density maps (contoured at 1σ) at the active site for representative types of complexes. *A*, in the monoclinic crystal of native active site YAV VP4, five molecules of VP4 are in the asymmetric unit with the C terminus of each molecule bound in the active site of its neighbor. The C-terminal carbonyl carbons of molecules D and E form an intermolecular acyl-enzyme complex (*AE*) with the active site nucleophilic serine Oγ of molecules C and D, respectively. An enzyme-product (*EP*) complex is formed between molecule pairs E/A, A/B, and B/C. *B*, in the cubic crystal of mutant YAV VP4 (Lys^{674} general base mutated to alanine), two molecules are in the asymmetric unit. The C-terminal carbonyl carbon of molecule A forms an intermolecular acyl-enzyme complex (*AE*) with the nucleophilic serine Oγ of molecule B. The active site of molecule A remains in the unbound state (*E*) because the C terminus of molecule B (shown as *red spheres*) folds back onto itself instead of being in an extended conformation.

Ser^{633} (the nucleophile), the main chain carbonyl oxygen of Cys^{669}, the carboxylate oxygen of Ala^{716} (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. 3*A* and 4*A*) the ester carbonyl oxygen (Ala^{716} O) is pointed toward the oxyanion hole (Ser^{633} NH), away from the general base lysine Nζ. The buried water, which we propose is deacylating, is coordinated by: Thr^{655} Oγ1, Ser^{633} Oγ, Lys^{674} Nζ, and Pro^{544} 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 (Cys^{588}-Cys^{604}) that links β-strands 5 and 6 (Fig. 2*A*). 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 both structures show electron density off the thiol group of Cys^{634} and Cys^{669}, which is consistent with them being oxidized to *S*-hydroxycysteine (CSO), a sulfenic acid. Conversely, the electron density for Cys^{688} Sγ 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. 2*A*). In contrast, the active site mutant (K674A) structure reveals two sep-

TABLE 2
Cysteine modification in YAV VP4
×, modification to CSO.

	Native active site structure					Mutant active site structure (K674A)	
	Chain A (product complex)	Chain B (product complex)	Chain C (acyl-enzyme)	Chain D (acyl-enzyme)	Chain E (product complex)	Chain A (empty)	Chain B (acyl-enzyme)
Cys^{634}	×	×	×		×		
Cys^{669}	×			×	×	×	

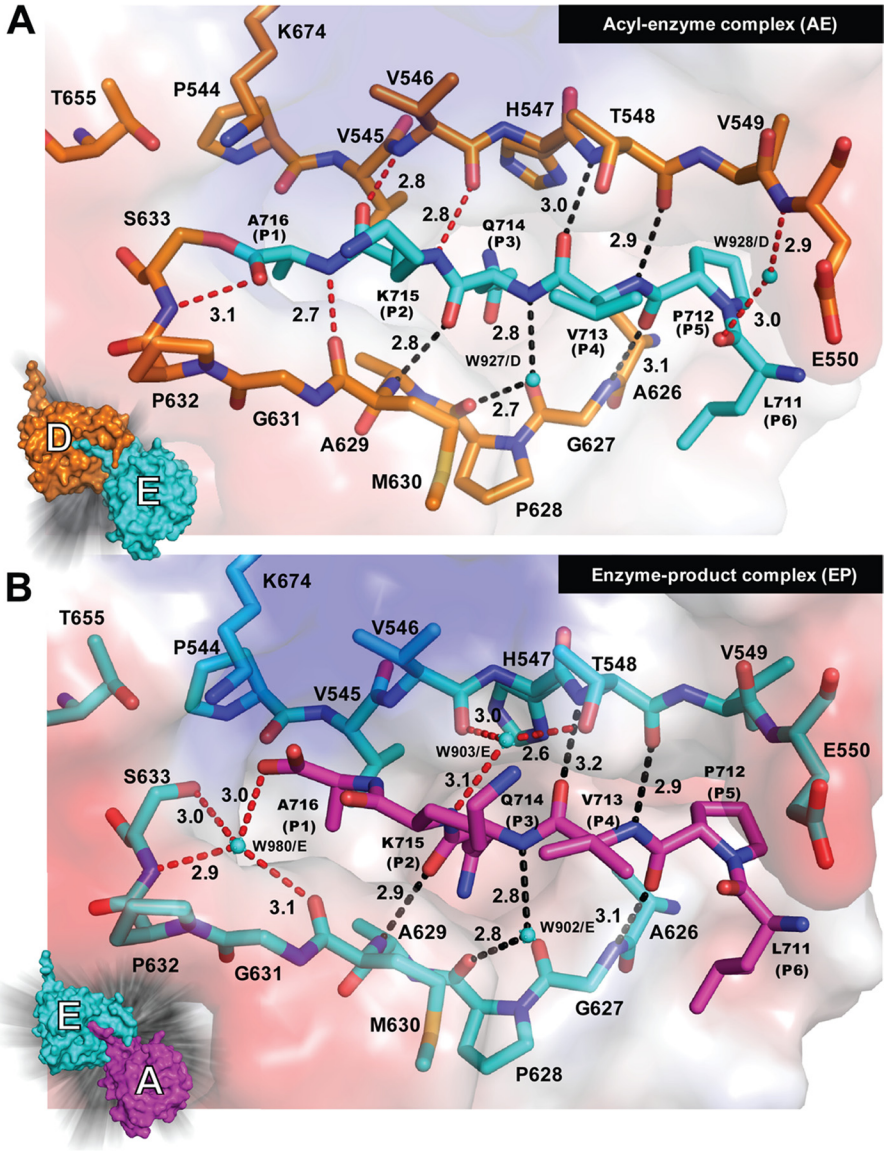

FIGURE 4. ***trans*** **acyl-enzyme complexes and** ***trans*** **enzyme-product complexes revealed in a crystal structure of YAV VP4 protease with a native active site.** *A*, the acyl-enzyme complex is shown with the substrate *colored* in *cyan* (molecule E) and the enzyme *colored orange* (molecule 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.

arate C-terminal conformations, one that is extended and another where the C terminus packs against the surface of the enzyme (Fig. 2*B*).

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. 2*A*, 3*A*, and 4*A*). 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

neighboring VP4 molecules (Fig. 3*A*). The Oγ of the serine nucleophile (Ser^{633}) is covalently linked to the carbonyl carbon of Ala^{716}, 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. 3*A*).

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. 4*A*). 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 Thr^{548} O–Val^{713} N, Thr^{548} N–Val^{713} O, Gly^{627} N–Pro^{712} O, Pro^{628} O–Gln^{714} N (via water), and Met^{630} N–Gln^{714} 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 Ala^{716} O–Ser^{633} N, Ala^{716} N–Met^{630} O, Lys^{715} O–Val^{546} N, and Lys^{715} N–Val^{546} O (Fig. 4). The hydrogen bond between Ala^{716} O (P1 residue) and Ser^{633} N (main chain nitrogen of the nucleophile) is the only hydrogen bond observed in the oxyanion hole. There are no other hydrogen bond donors near the carbonyl oxygen of Ala^{716}. It is possible that the positive dipole of α-helix 2, on which the nucleophilic Ser^{633} resides, contributes to oxyanion 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 Gln^{714} (P3) to the carbonyl oxygen of Pro^{628} (Fig. 4) is seen in each of the observed reaction steps. A water that bridges Glu^{550} N to Leu^{711} 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 Thr^{548} to the carbonyl oxygen of Val^{546} and the amide nitrogen of Lys^{715} (Fig. 4). The most interesting water seen in the enzyme-product complexes is the one that resides within the oxyanion hole (Fig. 4*B*). This water, which is displaced in the acyl-enzyme complex, makes hydrogen bonding interactions with the carboxylate oxygen of Ala^{716} (P1 residue), the Oγ and main chain nitrogen of Ser^{633} (nucleophile), and the main chain carbonyl oxygen of Met^{630} (Fig. 4*B*).

The major substrate specificity pockets are the S1 and S3 (Figs. 4–6). Residues Ile^{543}, Pro^{544}, Val^{545}, Leu^{561}, IIe^{563}, Ile^{592}, Ala^{629}, Met^{630}, Gly^{631}, Pro^{632}, Ser^{633}, and Cys^{634} form a continuous surface (the S1 pocket) for the binding of the P1 residue side chain (Ala^{716}). Residues Val^{545}, His^{547}, Ser^{559}, Leu^{560}, Leu^{561}, Leu^{572}, Gln^{576}, Gly^{591}, Ala^{607}, Pro^{609}, Val^{624}, Ala^{626}, Gly^{627}, Pro^{628}, and Ala^{629} form a continuous surface (the S3 pocket) for the binding of the P3 residue side chain (Gln^{714}). A cleft that is open on one side accommodates the P5 residue (Pro^{712}). The side chain of Leu^{711} (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. 3*B*). 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 His^{547} and Val^{545} 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: Ile^{592}, Glu^{594}, Asp^{595}, Ile^{596}, Pro^{597}, Ala^{629}, Met^{630}, Gly^{631}, Pro^{632}, CSO^{634} or Cys^{634}, Gln^{635}, and Leu^{638}. 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 Glu^{594}, Asp^{595}, Ala^{629}, Gly^{631}, and Gln^{635}. 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 Cys^{634} or CSO^{634}, Ile^{592}, Ala^{629}, Met^{630}, and Gly^{631} (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.

Ser^{633} is the YAV VP4 Nucleophile, but Could Thr^{655} Oγ Function as a Nucleophile in the Absence of Ser^{633} Oγ?—The serine/lysine catalytic dyad mechanism is conserved among all birnavirus VP4 proteases studied thus far. Therefore, Ser^{633} and Lys^{674} were predicted to be the nucleophile and general base, respectively, in YAV VP4 based on sequence alignments

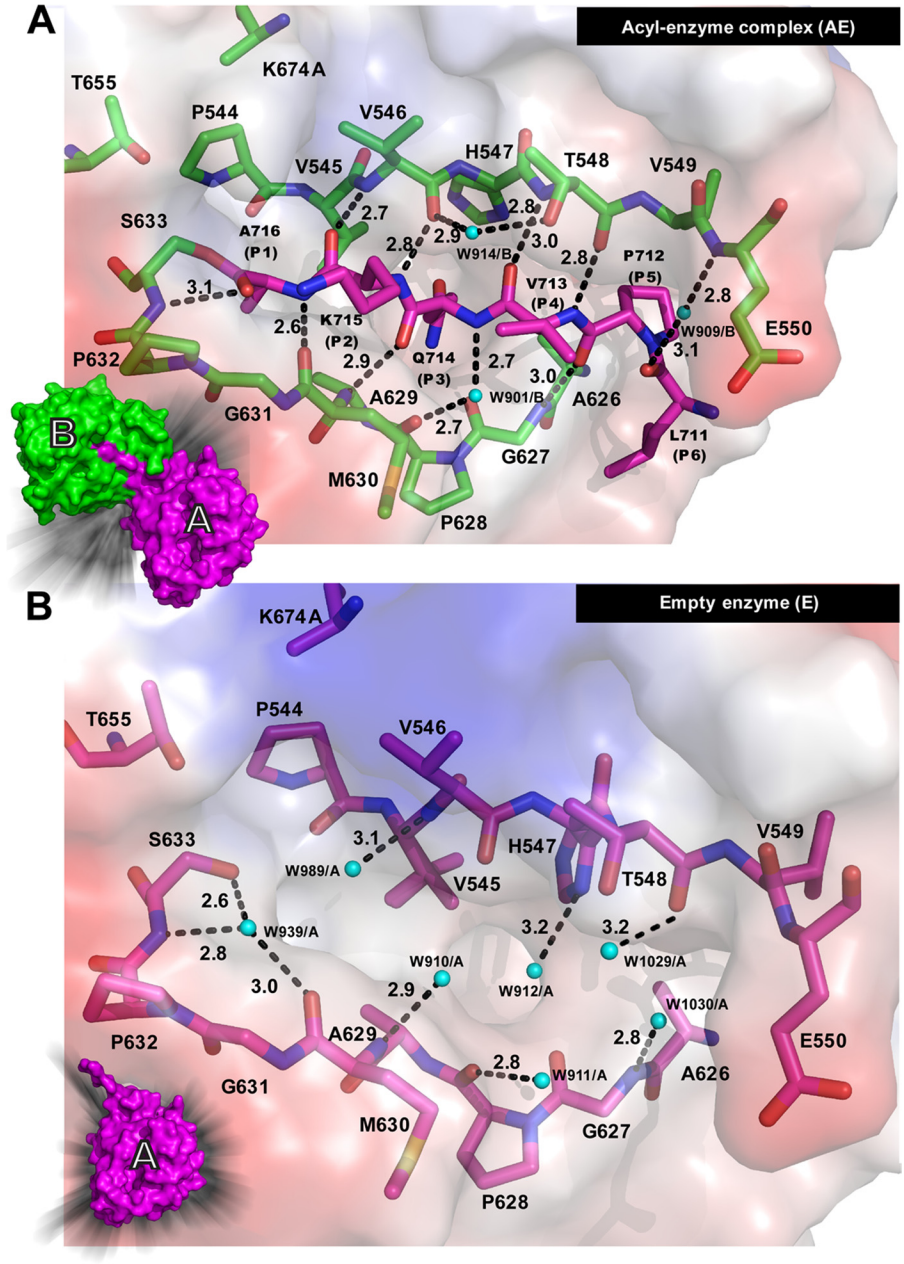

FIGURE 5. **A *trans* acyl-enzyme complex and an empty substrate binding groove revealed in a crystal structure of a YAV VP4 protease active site mutant (K674A).** *A*, the acyl-enzyme complex is shown with the substrate *colored* in *magenta* (molecule A) and the enzyme *colored* in *green* (molecule B). *B*, the molecule with an empty substrate binding groove (molecule A, *magenta*) is shown. 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 *black dashed lines*. The hydrogen bonding distances are given in Å.

(5, 24, 27). 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 a 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 Oγ 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

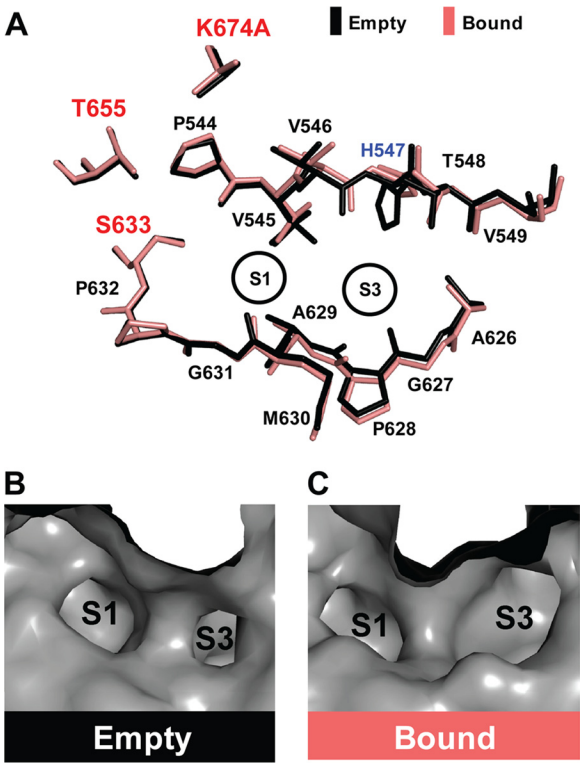

FIGURE 6. **Empty active site *versus* acyl-enzyme; changes observed in the substrate specificity binding pockets.** The active site mutant of YAV protease crystallized with two molecules in the asymmetric unit: one molecule in the bound state (acyl-enzyme) and the other with an empty active site. *A*, 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.

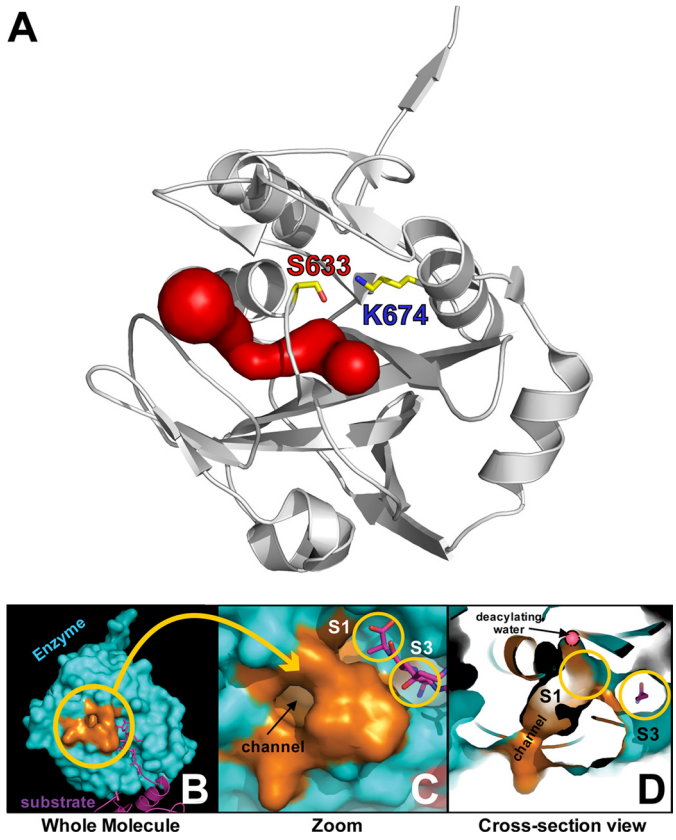

FIGURE 7. **A channel leads to the active site of YAV VP4 protease.** *A*, a channel (*red surface*), found adjacent to the substrate binding groove, leads from the enzyme surface to a region near the catalytic residues (shown in a *stick representation*) and the proposed deacylating water. *B*, the 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*).

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 Ile^{543} was mutated to a glycine. The side chain of IIe^{543} makes up part of the S1 binding pocket (Fig. 9). Replacing IIe^{543} 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 Val^{686} 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 Thr^{655}, a residue that coordinates the Nζ 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 Val^{686} with a more polar residue like glutamine could affect the packing around the C terminus, the substrate at the VP4/VP3 junction. Moreover, Leu^{705}, which resides on the conserved α-helix 4 near the C terminus, is located across from Val^{686}. 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

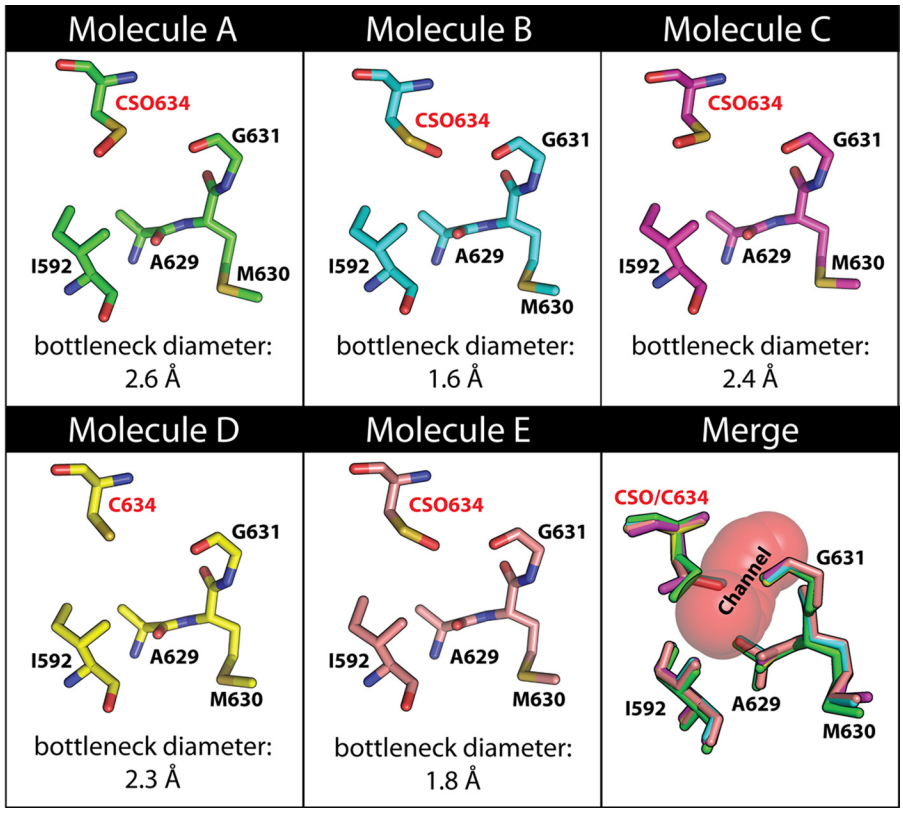

FIGURE 8. **Variations in the diameter of the YAV VP4 channel are not correlated with the modification of Cys^{634}.** The channel is made up of atoms from the following residues: Ile^{592}, Glu^{594}, Asp^{595}, Ile^{596}, Pro^{597}, Ala^{629}, Met^{630}, Gly^{631}, Pro^{632}, CSO^{634} or Cys^{634}, and Gln^{635}. The residues that form the narrowest point (bottleneck diameter) of the channel are Cys^{634} or CSO^{634}, Ile^{592}, Ala^{629}, Met^{630}, and Gly^{631}. These residues are shown for each molecule in the asymmetric unit in the native active site structure (molecules A–E). The carbon atoms for each molecule are rendered in a *different color*, all oxygen atoms are *red*, all nitrogen atoms are *blue*, and all sulfur atoms are *yellow*. The *merge panel* shows a superposition of these residues from molecule A to E, and the channel is shown as *red spheres*. The cysteine at position 634 is modified to CSO in four of the five molecules. The diameter of the channel at its narrowest point (bottleneck diameter) is shown in Å at the *bottom* of each *panel*. The average diameter for the five molecules in the native active site structure is 2.1 Å.

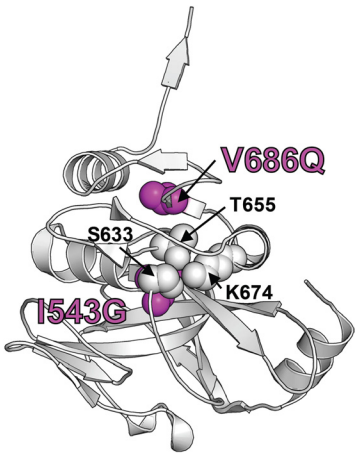

FIGURE 9. **Mapping non-active site mutations that affect catalysis.** The YAV VP4 active site residues are shown as *spheres* (Ser^{633}, Lys^{674}, and Thr^{655}). Polyprotein cleavage assays conducted by Imajoh *et al.* (24) have shown that non-active site mutations I543G and V686Q can significantly affect catalysis (shown as *magenta spheres*).

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 pK_a 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 (Ser^{506} and Thr^{732}, respectively). This allows Gln^{714} 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

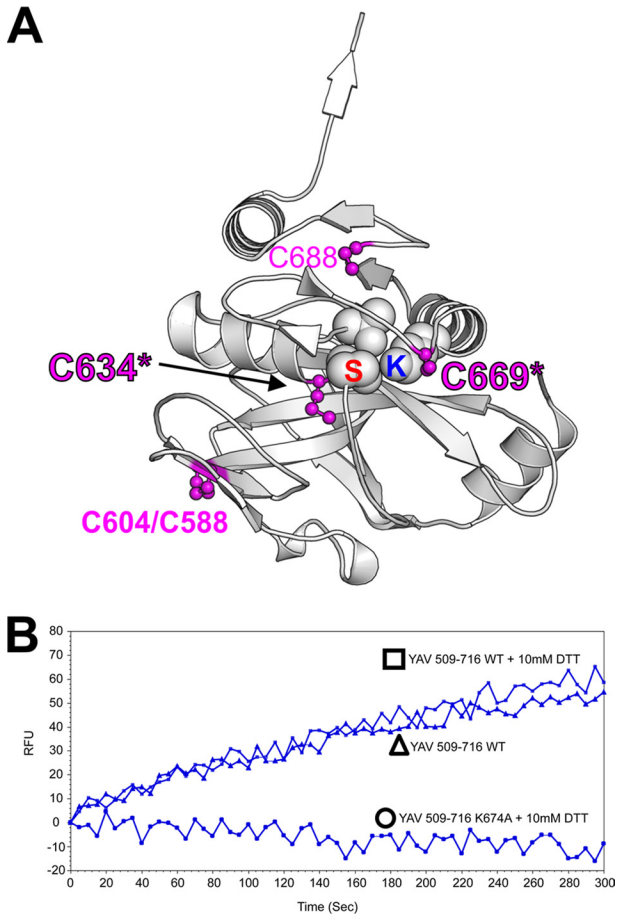

FIGURE 10. **Cysteine modification does not affect the YAV VP4 activity.** *A*, the position of cysteine residues in YAV VP4 is shown. The YAV VP4 is shown as a *white ribbon diagram*. The disulfide bond that is formed between Cys^{588} and Cys^{604} is shown (*magenta*). Additional cysteine residues are found at positions 634, 669, and 688 (*magenta*). A *star* and *larger font* denote the positions for the two cysteine residues (Cys^{634} and Cys^{669}) that are modified to CSO residues in some molecules in the asymmetric unit. *S*, Ser^{633}, the nucleophile (*red*); *K*, Lys^{674}, the general base (*blue*). *B*, incubation of YAV VP4 with excess reducing agent (DTT) does not affect activity, based on a fluorometric peptide cleavage assay. A fluorometric peptide, *N,N*-Dimethyl-p-aminobenzeneazobenzoic acid-KALPVQKAQGASE-5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid, carrying residues P7–P4′ of the YAV VP4 internal cleavage site (residues 710–720) was incubated with YAV VP4 509–716 (△), YAV VP4 509–716 with 10 mM DTT (□), and YAV VP4 509–716 K674A with 10 mM DTT (○). The reaction was carried out at 37 °C.

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, this posttranslational modification reaction is being investigated as a means of regulation (57–60). Interestingly, despite being purified in the presence of reducing agents, the electron densities for two of the five cysteine residues in YAV VP4 (Cys^{634} and Cys^{669}) are consistent with them being oxidized to CSO in some of the molecules in the asymmetric unit (Table 2 and Fig. 10*A*). This brings up the following question. Could the state of oxidation of these cysteine side chains affect the activity of YAV VP4? To address this question, we incubated VP4 in the presence and absence of an excess of reducing agent and tested its ability to cleave a fluorometric peptide substrate with a sequence matching that of the YAV VP4 internal cleavage site. The reducing environment did not affect the activity (Fig. 10*B*). Because Cys^{634} immediately follows the nucleophile (Ser^{633}), this prompted us to analyze the structures to see if there was a correlation between the oxidation state of Cys^{634} and the acylation state of Ser^{633}. Two of the five molecules in the native active site structure form an acyl-enzyme complex (molecules C and D). In molecule C, Cys^{634} (C) is modified as a CSO, and Ser^{633} (C) is covalently bonded to Ala^{716} (D), forming an acyl-enzyme, whereas in molecule D, Cys^{634} (D) is not modified, but Ser^{633} (D) is again covalently bonded to Ala^{716} (E), forming an acyl-enzyme (Table 2). Therefore, the electron density of the crystal structure does not support the suggestion that modification (oxidation to CSO) of Cys^{634} affects the activity of YAV VP4. By looking at the native active site structure, we can see that the main chain carbonyl oxygen of Cys^{669} is within hydrogen bonding distance to the Nζ of the general base lysine (Lys^{674}). However, no significant change in hydrogen bonding distance is seen between the Nζ of the lysine general base and the Oγ of the serine nucleophile (Ser^{633}) when molecules with different oxidization states at Cys^{669} were compared. In addition, no correlation appears to exist between the cysteine modification and the geometry of the YAV VP4 aqueous channel (Fig. 8).

CONCLUSION

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 Ser^{633} 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 deacylating 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.

REFERENCES

1. Dobos, P., Hill, B. J., Hallett, R., Kells, D. T., Becht, H., and Teninges, D. (1979) Biophysical and biochemical characterization of five animal viruses with bisegmented double-stranded RNA genomes. *J. Virol.* **32,** 593–605
2. Delmas, B., Mundt, E., Vakharia, V. N., and Wu, J. L (2011) in *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses* (King, A. M., Lefkowitz, E., Adams, M. J., and Carstens, E. B., eds) pp. 499–507, Elsevier Science, Amsterdam
3. Da Costa, B., Soignier, S., Chevalier, C., Henry, C., Thory, C., Huet, J. C., and Delmas, B. (2003) Blotched snakehead virus is a new aquatic birnavirus that is slightly more related to avibirnavirus than to aquabirnavirus. *J. Virol.* **77,** 719–725
4. Nobiron, I., Galloux, M., Henry, C., Torhy, C., Boudinot, P., Lejal, N., Da Costa, B., and Delmas, B. (2008) Genome and polypeptides characterization of Tellina virus 1 reveals a fifth genetic cluster in the Birnaviridae family. *Virology* **371,** 350–361
5. Petit, S., Lejal, N., Huet, J. C., and Delmas, B. (2000) Active residues and viral substrate cleavage sites of the protease of the birnavirus infectious pancreatic necrosis virus. *J. Virol.* **74,** 2057–2066
6. Hjalmarsson, A., Carlemalm, E., and Everitt, E. (1999) Infectious pancreatic necrosis virus. Identification of a VP3-containing ribonucleoprotein core structure and evidence for *O*-linked glycosylation of the capsid protein VP2. *J. Virol.* **73,** 3484–3490
7. Casañas, A., Navarro, A., Ferrer-Orta, C., González, D., Rodríguez, J. F., and Verdaguer, N. (2008) Structural insights into the multifunctional protein VP3 of birnaviruses. *Structure* **16,** 29–37
8. Birghan, C., Mundt, E., and Gorbalenya, A. E. (2000) A non-canonical lon proteinase lacking the ATPase domain employs the Ser-Lys catalytic dyad to exercise broad control over the life cycle of a double-stranded RNA virus. *EMBO J.* **19,** 114–123
9. Da Costa, B., Chevalier, C., Henry, C., Huet, J. C., Petit, S., Lepault, J., Boot, H., and Delmas, B. (2002) The capsid of infectious bursal disease virus contains several small peptides arising from the maturation process of pVP2. *J. Virol.* **76,** 2393–2402
10. Ekici, O. D., Paetzel, M., and Dalbey, R. E. (2008) Unconventional serine proteases. Variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci.* **17,** 2023–2037
11. Paetzel, M., Karla, A., Strynadka, N. C., and Dalbey, R. E. (2002) Signal peptidases. *Chem. Rev.* **102,** 4549–4580
12. Paetzel, M., and Woodgate, R. (2013) UmuD and UmuD′ proteins. in *Handbook of Proteolytic Enzymes* (Rawlings, N. D., and Salvesen, G., eds) pp. 3487–3492, Academic Press, Inc., New York
13. Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W., and Strynadka, N. C. (2001) Crystal structure of LexA. A conformational switch for regulation of self-cleavage. *Cell* **106,** 585–594
14. 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) The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J. Biol. Chem.* **279,** 8140–8148
15. Kim, A. C., Oliver, D. C., and Paetzel, M. (2008) Crystal structure of a bacterial signal peptide peptidase. *J. Mol. Biol.* **376,** 352–366
16. Hosono, N., Suzuki, S., Kusuda, R. (1996) Genogrouping of birnaviruses isolated from marine fish: A comparison of VP2/NS junction regions on genome segment A. *J. Fish Dis.* **19,** 295–302
17. Hirayama, T., Nagano, I., Shinmoto, H., Yagyu, K., and Oshima, S. (2007) Isolation and characterization of virulent yellowtail ascites virus. *Microbiol. Immunol.* **51,** 397–406
18. Kusuda, R., Nagato, K., Kawai, K. (1994) Characteristics of a virus isolated from red sea bream, *Pagrus major* showing exophthalmos. *Suisanzoshoku* **42,** 145–149
19. Kusuda, R., Kado, K., Takeuchi, Y., Kawai, K. (1989) Characteristics of two virus strains isolated from young Japanese flounder *Paralichthys olivaceus*. *Suisanzoshoku* **37,** 115–120
20. Kamakura, M., Suzuki, S., and Kusuda, R. (1995) Characterization of birnavirus isolated from diseased tiger puffer. Annual Meeting of the Japan Society of Fish Pathology, September 1995, Mie, Japan, abstract p 14 (in Japanese)
21. Suzuki, S., Utsunomiya, I., Kusuda, R. (1998) Experimental infection of marine birnavirus strain JPO-96 to Japanese pearl oyster *Pinctada fucata*. *Bull. Mar. Sci. Fish Kochi Univ.* **18,** 39–41
22. Suzuki, S., Nakata, T., Kamakura, M., Yoshimoto, M., Furukawa, Y., Yamashita, Y., Kusuda, R. (1997) Isolation of birnavirus from Agemaki (jack knife clam) *Sinonovacura constricta* and survey of the virus using PCR technique. *Fish Sci.* **63,** 563–566
23. Suzuki, S., Kamakura, M., Kusuda, R. (1998) Isolation of birnavirus from Japanese pearl oyster *Pinctada fucata*. *Fish Sci.* **64,** 342–343
24. Imajoh, M., Goto, T., and Oshima, S. (2007) Characterization of cleavage sites and protease activity in the polyprotein precursor of Japanese marine aquabirnavirus and expression analysis of generated proteins by a VP4 protease activity in four distinct cell lines. *Arch. Virol.* **152,** 1103–1114
25. Feldman, A. R., Lee, J., Delmas, B., and Paetzel, M. (2006) Crystal structure of a novel viral protease with a serine/lysine catalytic dyad mechanism. *J. Mol. Biol.* **358,** 1378–1389
26. Lee, J., Feldman, A. R., Delmas, B., and Paetzel, M. (2007) Crystal structure of the VP4 protease from infectious pancreatic necrosis virus reveals the acyl-enzyme complex for an intermolecular self-cleavage reaction. *J. Biol. Chem.* **282,** 24928–24937
27. Chung, I. Y., and Paetzel, M. (2011) Crystal structure of a viral protease intramolecular acyl-enzyme complex. Insights into cis-cleavage at the VP4/VP3 junction of *Tellina birnavirus*. *J. Biol. Chem.* **286,** 12475–12482
28. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. in *The Proteomics Protocols Handbook* (Walker, J. M., ed) pp. 571–607, Humana Press, New York
29. Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, number 26
30. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62,** 72–82
31. Collaborative Computing Project, Number 4 (1994) The CCP4 Suite. Programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
32. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 22–25
33. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
34. Murshudov, G., Vagin, A., and and Dodson, E. (1996) Application of maximum likelihood refinement. in *The Refinement of Protein Structures: Proceedings of the Daresbury Study Weekend*, Science and Engineering Research Council, Daresbury, UK
35. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 355–367
36. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
37. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
38. Frishman, D., and Argos, P. (1995) Knowledge-based protein secondary structure assignment. *Proteins* **23,** 566–579
39. Reynolds, C., Damerell, D., and Jones, S. (2009) ProtorP. A protein-protein interaction analysis server. *Bioinformatics* **25,** 413–414
40. Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J., and Damborsky, J. (2012) CAVER 3.0. A tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput. Biol.* **8,** e1002708
41. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK. A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291
42. Engh, R. A., and Huber, R. (2001) Structure quality and target parameters. in *International Tables for Crystallography* (Rossman, M. G., and Arnold,

E., eds) pp. 382–392, Kluwer Academic Publishers, Dordrecht, Netherlands
43. Kleywegt, G. J., and Jones, T. A. (1996) Phi/psi-chology. Ramachandran revisited. *Structure* **4,** 1395–1400
44. Liebecq, C. (1992) *Compendium of Biochemical Nomenclature and Related Documents*, Portland Press, London
45. DeLano, W. L. (2002) *The PyMOL User's Manual*, DeLano Scientific, San Carlos, CA
46. Ten Eyck, L. F. (1973) Crystallographic fast Fourier transforms. *Acta Crystallogr. A* **29,** 183–191
47. Read, R. J., and Schierbeek, A. J. (1988) A phased translation function. *J. Appl. Crystallogr.* **21,** 490–495
48. Immirzi, A. (1966) *Crystallographic Computing Techniques* (Ahmed, F. R., ed) p. 399, Munksgaard, Copenhagen
49. Tsodikov, O. V., Record, M. T., Jr., and Sergeev, Y. V. (2002) Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *J. Comput. Chem.* **23,** 600–609
50. Zimmerman, M., Ashe, B., Yurewicz, E. C., and Patel, G. (1977) Sensitive assays for trypsin, elastase, and chymotrypsin using new fluorogenic substrates. *Anal. Biochem.* **78,** 47–51
51. Savage, D. F., O'Connell, J. D., 3rd, Miercke, L. J., Finer-Moore, J., and Stroud, R. M. (2010) Structural context shapes the aquaporin selectivity filter. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 17164–17169
52. Savage, D. F., Egea, P. F., Robles-Colmenares, Y., O'Connell, J. D., 3rd, and Stroud, R. M. (2003) Architecture and selectivity in aquaporins. 2.5 Å x-ray structure of aquaporin Z. *PLoS Biol.* **1,** E72
53. Rodríguez-Lecompte, J. C., and Kibenge, F. S. (2002) Site-directed mutagenesis of avibirnavirus VP4 gene. *Virology* **292,** 241–246
54. Lejal, N., Da Costa, B., Huet, J. C., and Delmas, B. (2000) Role of Ser-652 and Lys-692 in the protease activity of infectious bursal disease virus VP4 and identification of its substrate cleavage sites. *J. Gen. Virol.* **81,** 983–992
55. Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (2002) Crystal structure of a bacterial signal peptidase apoenzyme. Implications for signal peptide binding and the Ser-Lys dyad mechanism. *J. Biol. Chem.* **277,** 9512–9519
56. Granzow, H., Birghan, C., Mettenleiter, T. C., Beyer, J., Köllner, B., and Mundt, E. (1997) A second form of infectious bursal disease virus-associated tubule contains VP4. *J. Virol.* **71,** 8879–8885
57. Claiborne, A., Yeh, J. I., Mallett, T. C., Luba, J., Crane, E. J., 3rd, Charrier, V., and Parsonage, D. (1999) Protein-sulfenic acids. Diverse roles for an unlikely player in enzyme catalysis and redox regulation. *Biochemistry* **38,** 15407–15416
58. Kettenhofen, N. J., and Wood, M. J. (2010) Formation, reactivity, and detection of protein sulfenic acids. *Chem. Res. Toxicol.* **23,** 1633–1646
59. Poole, L. B., Karplus, P. A., and Claiborne, A. (2004) Protein sulfenic acids in redox signaling. *Annu. Rev. Pharmacol. Toxicol.* **44,** 325–347
60. Reddie, K. G., and Carroll, K. S. (2008) Expanding the functional diversity of proteins through cysteine oxidation. *Curr. Opin. Chem. Biol.* **12,** 746–754
61. French, S., and Wilson, K. (1978) On the treatment of negative intensity observations. *Acta Crystallogr. A* **34,** 517–525