Molecular Cell 22, 339–349, May 5, 2006 ©2006 Elsevier Inc. DOI 10.1016/j.molcel.2006.03.015

# Crystal Structures of the Vaccinia Virus Polyadenylate Polymerase Heterodimer: Insights into ATP Selectivity and Processivity

Carmen M. Moure,[1] Brian R. Bowman,[1,3] Paul D. Gershon,[2] and Florante A. Quiocho[1,*]
[1] Verna and Marrs McLean Department of Biochemistry and Molecular Biology
Baylor College of Medicine
Houston, Texas 77030
[2] Department of Molecular Biology and Biochemistry
University of California, Irvine
Irvine, California 92697

## Summary

**Polyadenylation of mRNAs in poxviruses, crucial for virion maturation, is carried out by a poly(A) polymerase heterodimer composed of a catalytic component, VP55, and a processivity factor, VP39. The ATP-γ-S bound and unbound crystal structures of the vaccinia polymerase reveal an unusual architecture for VP55 that comprises of N-terminal, central or catalytic, and C-terminal domains with different topologies and that differs from many polymerases, including the eukaryotic poly(A) polymerases. Residues in the active site of VP55, located between the catalytic and C-terminal domains, make specific interactions with the adenine of the ATP analog, establishing the molecular basis of ATP recognition. VP55's concave surface docks the globular VP39. A model for RNA primer binding that involves all three VP55 domains and VP39 is proposed. The model supports biochemical evidence that VP39 functions as a processivity factor by partially enclosing the RNA primer at the heterodimer interface.**

## Introduction

The mRNAs of poxviruses, like those of eukaryotes, are 5′ capped and 3′ polyadenylated (Moss et al., 1975). The enzymes involved in mRNA processing in poxviruses, which replicate in the cytoplasm of infected cells, are encoded by the viral genome. In vaccinia, the prototypical poxvirus for laboratory studies, and the material used in the vaccine responsible for the eradication of smallpox, mRNA polyadenylation is carried out by a heterodimeric poly(A) polymerase (PAP) composed of a catalytic 55 kDa subunit, VP55, and its 39 kDa processivity factor VP39 (Moss et al., 1975; Gershon et al., 1991; Gershon and Moss, 1992, 1993a). Using ATP and $Mg^{2+}$, VP55 transfers AMP to the mRNA 3′ OH end in a template-independent phosphoryl transfer reaction. Kinetic studies have shown that VP55 is highly specific for ATP in the presence of $Mg^{2+}$, although specificity is relaxed in the presence of $Mn^{2+}$ (Gershon and Moss, 1992; Shuman and Moss, 1988). In the absence of VP39, VP55 adds 24–30 nucleotide (nt) poly(A) tails in a rapid manner before dissociating from the RNA substrate. In the presence of VP39, VP55 does not dissociate from the RNA, allowing tail elongation to 200–300 adenylates in a semiprocessive manner (40-fold rate increase).

Poxvirus VP55 orthologs share greater than 95% sequence identity (BLAST search, data not shown). The poxvirus PAPs belong to the pol X superfamily of nucleotidyltransferases (NT) by virtue of their NT signature motif GSX(9-13)DXD (Aravind and Koonin, 1999). This superfamily includes the founding member polymerase β (pol β), bacterial PAP, eukaryotic PAP, CCA-adding enzyme, pol X (Shimizu et al., 1993), chromatin-associated TRF (Castano et al., 1996), and antibiotic nucleotidyl transferases among others. The known protein structures of the superfamily bear no resemblance to one another other than the catalytic domain, which consists, at a minimum, of a mixed five-stranded β sheet with two connecting helices. A helix turn bearing the NT signature motif is inserted within the β sheet. Three conserved aspartates are present at topologically equivalent positions in the NT proteins and other polymerases (Delarue et al., 1990). In addition, the NT motif plays a major role in binding the incoming nucleotide.

Besides acting as a processivity factor in poly(A) elongation at the 3′ end, VP39 also functions as a cap 0 ($m^7$G[5′]pppN···)-specific nucleoside-2′-*O*-methyltransferase at the mRNA 5′ end (Schnierle et al., 1992). The tertiary structure and function of the bifunctional VP39 have been extensively investigated. The determination in our laboratory of over a dozen crystal structures for VP39 alone and in complexes with various nucleotides and Ado-Met coenzyme or its Ado-Hcy product have revealed in atomic detail the mechanisms of $m^7$G cap recognition, the sequence-independent binding of single-stranded RNA substrates, and the nature of the methyltransferase active site (Hodel et al., 1996, 1997, 1998; Hu et al., 1999, 2002; Quiocho et al., 2000). However, the role of VP39 as a processivity factor in polyadenylation is not as well understood. Mutagenesis and protein footprinting studies have indicated that the heterodimerization interface of VP39 occurs on the reverse face from its methyltransferase active site so that both activities, 5′ methyl transfer and 3′ polyadenylation, can occur independently (Shi et al., 1997; Gershon et al., 1998).

The 3′ ends of vaccinia mRNAs are uridylate rich. Electrophoretic mobility shift assay (EMSA) studies have demonstrated that the heterodimer binds oligonucleotides at least 45 nt in length, possessing at least three uridylate (rU) residues positioned at −10, −36, and −37 with respect to the 3′ end (Gershon and Moss, 1993b; Deng et al., 1997; Deng and Gershon, 1997). Photocrosslinking studies demonstrated that the side chain of VP39's Arg107, a residue essential for heterodimerization (Shi et al., 1997), is involved in base-specific interactions with the downstream −10 rU (Deng et al., 1999), suggesting that the RNA primer binds at the heterodimer interface. Further evidence for this was obtained in more comprehensive photocrosslinking studies and a protein thiol-accessibility assay (Deng et al., 1999; Johnson and Gershon, 1999; Oguro et al., 2002). Functional studies have indicated that in order to extend the poly(A) tail

*Correspondence: faq@bcm.tmc.edu
[3] Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138.

beyond 3–7 nt, VP55 requires, in addition to the three rUs mentioned above, a cluster of six rUs at the 3′ end of the starting RNA primer (Johnson et al., 2004) and a ribose at −1 (Deng and Gershon, 1997).

In order to obtain insight into the polyadenylation mechanism of vaccinia, a process that is essential for viral replication, we determined the crystal structure of the vaccinia PAP heterodimer with and without a bound nonhydrolyzable ATP analog, ATP-γ-S. The structure reveals an unusual architecture for a polymerase. It identifies active site residues, shedding light on the mechanism of ATP selectivity. The structure further delineates a groove at the VP55-VP39 interface that can accommodate the downstream portion of the RNA primer. A model for mRNA primer binding, which takes into account this finding as well as available biochemical data, is presented.

## Results and Discussion

### Structure Determination

Crystallization experiments employed VP55 cloned from the vaccinia Western Reserve strain that differed from the wild-type protein in having a Leu36 to Ser substitution and missing the N-terminal ten amino acids (see Experimental Procedures). The modified protein was indistinguishable from the wild-type protein in PAP activity. The fully active AS11 mutant of VP39 was used, as in previous crystallographic studies (Hodel et al., 1996). The VP55-VP39 complex was crystallized at pH 8.7, which is optimum for polymerase activity (Moss et al., 1975). The structure of the heterodimer complex was determined by MAD phasing to 3 Å resolution from a selenomethionine variant of VP39 and refined at 2.4 Å resolution (Table 1). Data for the complex with the ATP analog ATP-γ-S were obtained from a crystal soaked in solution containing the analog and used in the refinement at 2.3 Å resolution (Table 1). VP55 and VP39 each possessed a pair of disordered segments of surface polypeptide (Figure 1A), one of which was also disordered in the monomeric VP39 structures (Hodel et al., 1996; Hu et al., 1999). The structures of the heterodimer in the presence and absence of the ATP analog were virtually identical with conformational differences localized to only a few residues that contact the ATP at the active site (further discussed below). Unless otherwise indicated, the structure described in this report is that of the VP55-VP39-ATP analog ternary complex.

Table 1. Data Collection, Phasing, and Refinement Statistics

| | | |
|---|---|---|
| Seleno-MAD (Ligand-Free) Data Collection | | |
| Space group | $P2_12_12_1$ | |
| Unit cell dimensions (Å) | a = 69.08, b = 91.26, c = 133.46 | |
| Resolution (Å) | 50–2.26 | |
| Wavelength | Peak | Inflection |
| | 0.9793 | 0.9792 |
| Completeness % | 83.4 (68.1) | 83 (63.7) |
| $R_{sym}$[a] % (outer shell) | 7.1 (48.6) | 7.7 (65.3) |
| Redundancy (outer shell) | 5.5 (4.8) | 5.5 (4.3) |
| I/σ | 10.8 (3.0) | 9.9 |
| Phasing (50–3 Å Resolution Range) | | |
| Figure of merit | 0.52 | |
| ATP-γ-S Soaked Data Collection | | |
| Space group | $P2_12_12_1$ | |
| Cell parameters (Å) | a = 70.05, b = 91.69, c = 133.7 | |
| Z[b] | 1 | |
| Resolution (Å) | 50–2.3 | |
| Number of Reflections | | |
| Total | 172,749 | |
| Unique | 38,285 | |
| Completeness % (outer shell) | 99.2 (97.9) | |
| $R_{sym}$ % (outer shell) | 8.6 (43.6) | |
| I/σ | 10.2 (3.5) | |
| Refinement | ATP-γ-S Soaked | Ligand Free |
| Resolution (Å) | 50.0–2.3 | 50.0–2.4 |
| R factor[c]/R free (%) | 22.1/26.0 | 24.3/28.8 |
| Number of Atoms/B Factor | | |
| Protein | 6003/33.04 | 6008/40.78 |
| Solvent | 260/34.22 | 140/43.94 |
| Metal ions | 2/54.03 | |
| ATP-γ-S | 62/58.71 | |
| Rmsd | | |
| Bond lengths (Å) | 0.008 | 0.008 |
| Angles (°) | 1.60 | 1.42 |
| Ramachandran (%) | | |
| Allowed | 99.1 | 98.1 |
| Generously allowed | 0.9 | 1.9 |

[a] $R_{sym} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I(h)$.
[b] Number of molecules in the asymmetric unit.
[c] R factor = $\sum (|F_o| - k|F_c|) / \sum |F_o|$. R free was calculated by using a random 5% of the reflection data that was omitted in the refinement.

### An Unusual Fold for VP55 PAP

The structure of VP55 is shaped roughly like a kidney bean (Figure 1A), with the concave surface docking the nearly globular single domain of VP39 (discussed further below). The VP55 structure is composed of three domains: the N-terminal or N domain (residues 12–117), the central or catalytic domain (residues 161–278 and 465–479), and the C-terminal or C domain (residues 130–149 and 279–464) (Figure 1A). All three domains have distinct topologies. The N domain has an all α topology, with six of the eight helices forming three helix-hairpin-helix (HhH) motifs that are packed to create an elongated ensemble. The catalytic domain has an α/β topology. Its core comprises of a mixed nine-stranded twisted β sheet (Figure 1B), with the first seven strands being folded from one long (118 residue) polypeptide segment that follows the N domain and the last two strands (β16 and β17) originating from the extreme C-terminal-most 15 residues of the entire polypeptide chain. Of the four α helices in the catalytic domain, the two longest (J and L) near the surface pack against one side of the β sheet in a nearly parallel orientation to the β strands. The C domain has essentially an α + β topology, consisting of a six-stranded β sheet with one connecting helix (Q) and a separate cluster of six α helices (I, N to P, and R to S) (Figure 1A). Helix I of the helix cluster is linked to the N domain H helix and the catalytic

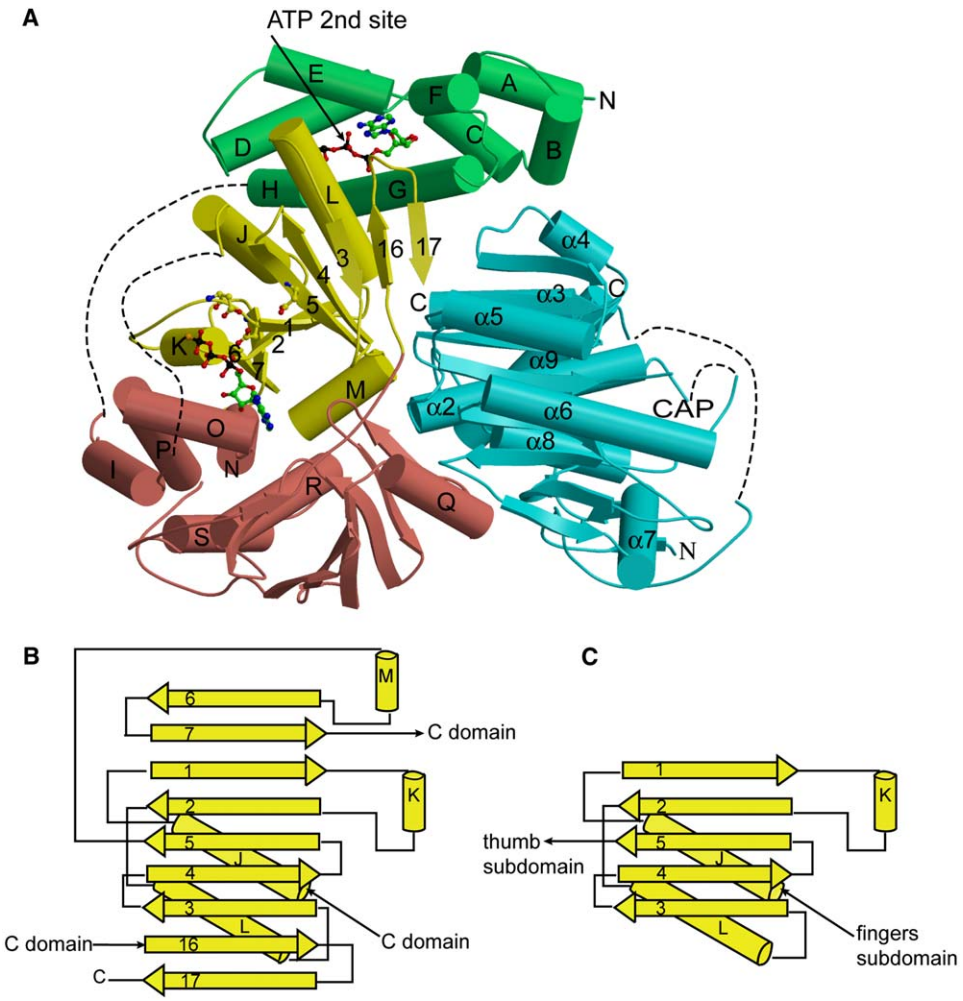


Figure 1. Structure of the Poly(A) Polymerase Heterodimer

(A) Ribbon diagram. This view represents the "front face" of the heterodimer. The three domains of VP55 are the N domain (green), the catalytic or central domain (yellow), and the C domain (salmon). VP39 is colored in cyan. Disordered polypeptide segments in VP55 (residues 118–129 and 150–160) and VP39 (residues 27–32 and 142–144) are represented by dashed lines. Helices in the entire VP55 are identified by letters, and β strands in the catalytic domain are identified by numbers. Helices in VP39 are identified by numbers preceded by α. The three active site aspartates and the two bound molecules of ATP-γ-S (one in the active site and the other [ATP second site] on the protein surface) in VP55 are shown as ball-and-stick models. "CAP" marks the location of the $m^7G$ cap binding site in VP39, which is adjacent to the methyltransferase active site.

(B) Topology diagram of the catalytic domain of VP55.

(C) Topology diagram of the canonical palm motif of DNA polymerase β.

Note that in both (B) and (C) strands 1–5 are equivalent.

domain J helix through two disordered segments, residues 118–129 and 150–160, respectively (Figure 1A).

A search for overall structural similarities of the entire VP55 molecule in the DALI database (Holm and Sander, 1996) revealed no significant matches. Neither did searches with the N or C domains alone. A portion of the ~135 residue catalytic domain shows low level homology to the catalytic domains of other NT proteins, confirming their evolutionary relationship. For example, DALI detected homology to the catalytic domains of the CCA-adding tRNA nucleotidyl transferase of *Aquifex aeolicus* (PDB code 1VFG, [Tomita et al., 2004]; Z = 4.4, rmsd = 3.1 Å over just 79 residues) and pol β (PDB code 1BPY, [Sawaya et al., 1997]; Z = 3.5, rmsd = 4.9 Å for only 80 residues). With a total of nine β strands (Figures 1A and 1B), the sheet fold of the VP55 catalytic domain is the most extended version of the canonical five-stranded β sheet topology of the NT superfamily (e.g., see Figure 1C) (Davies et al., 1994; Sawaya et al., 1994). Interestingly, helix K in both VP55 (Figure 1B) and pol β (Figure 1C) is part of the helix-turn motif that contains the NT signature. The CCA-adding tRNA nucleotidyl transferase (Tomita et al., 2004) also contains an extended version of the β sheet topology with seven β strands that correspond to strands 1–7 in VP55.

Based strictly on the domain composition, VP55 resembles the fragment of pol β that, in addition to the central or catalytic subdomain (the "palm"), also contains a flanking N subdomain (the "fingers") and a C subdomain (the "thumb") (Davies et al., 1994; Pelletier et al., 1994; Sawaya et al., 1994), all three named as such because the spatial arrangement of the subdomains resembles a right hand (Ollis et al., 1985). The overall kidney bean shape and the juxtaposition of the three

domains of VP55 do not, however, resemble a hand, as in other eukaryotic PAPs (Martin et al., 2000), and thus, we have not retained the hand analogy.

### Structure of the VP55-VP39 Complex

The VP55-VP39 heterodimer structure looks approximately like a thick triangular slab (Figure 1A; see also Figure 4), with the front side containing the polymerase active site and the back side harboring a second ATP binding site. The region of VP39 involved in dimer formation is opposite from the 5′ cap mRNA binding site (Figure 1A). VP39 fits snugly in the concave surface of VP55 and interacts with elements of all three domains of VP55. Most of the intersubunit interactions at the interface are electrostatic in nature, formed either direct or through water molecules, which could explain the salt dependency of the dimerization (Shi et al., 1997).

The heterodimer interface buries a large surface area (3,435 Å$^2$). Although a large proportion (25%) of the total accessible surface of VP39, mostly from the semicircular edge composed of termini of many helices and β strands (Figure 1A), is buried at the interface with VP55, the structure of VP39 in the heterodimer is almost identical to the unbound VP39 (PDB code AV6; [Hodel et al., 1998]) with an rmsd for Cα atoms of 0.6 Å. The absence of large structural changes is not a surprising result given the compactness of the single domain VP39 structure and the lack of involvement of flexible regions at the interface. However, to achieve a better fit, the side chains of 11 almost entirely polar residues (Lys52, Arg55, His56, Tyr83, Lys90, Arg107, Lys132, Asn194, Lys195, Arg220, and Arg268) of VP39 contacting VP55 undergo localized conformational changes in going from the unbound to the bound structure. Some of these residues (Arg55, His56, and Arg107) have been previously identified in mutagenesis studies as critical for heterodimerization (Shi et al., 1997). The most notable is Arg107 in the α4 of VP39, whose solvent-exposed guanidinium group forms a salt link with the carboxylate of Asp99 (G helix) of the N domain of VP55. Arg107 has been implicated to interact with the downstream nt of the RNA primer (further discussed below).

### VP55 Polymerase Active Site

The polymerase active site is situated in the cleft at the interface of the catalytic and the C domains (Figure 1A). It is defined by the helix-turn motif that contains the NT signature and the antiparallel β sheet of the catalytic domain and helices N, O, and P of the C domain. The linker connecting the C domain to the last two strands of the β sheet in the catalytic domain delimits the cleft on one side. The disordered linker connecting helix I in the C domain to the catalytic domain likely demarcates the active site cleft on the other side. Three aspartates, Asp202 (the first aspartate of the motif located in the turn of the NT helix-turn motif), Asp204 (the second aspartate and last residue of the NT motif located at the beginning of the β2 strand that follows the helix-turn motif), and Asp253 (in β5) (Figures 1A and 2A) are topologically located at equivalent positions to the conserved aspartates in other polymerase active sites (Delarue et al., 1990). Two divalent metals (A and B) are observed bound to the carboxylates of Asp202 and Asp204 (Figures 2A and 2B). The metals, which are 4.0 Å apart, were assigned as $Ca^{2+}$ ions because the crystallization solution contains high concentrations of $CaCl_2$. The metals were absent in the ATP analog-free structure (data not shown). The presence of two active site metals is consistent with a two-metal ion phosphoryl transfer mechanism that is common to all polymerases (Steitz, 1997). Metals A and B overlap with the corresponding metal sites in the structure of pol β, an NT enzyme (Sawaya et al., 1997). Metal coordinations (distances from 2.3 to 2.5 Å) are depicted in Figure 2B. Metal A is coordinated to the carboxylates of Asp202 and Asp204 and a water molecule. Metal B is coordinated to the carboxylate oxygens of Asp202 and Asp204, a water molecule, and the nucleotide bridging phosphate oxygens O3α and O3β, making the α-β bond labile to hydrolysis. Whereas the coordination sphere of metal B is complete, the coordination sphere for metal A is not. In pol β, the equivalent of metal A is further coordinated to the 3′ OH end of the primer and the Asp residue that is equivalent to Asp253 of VP55. However, in VP55, the carboxylate oxygens of Asp253 are 4.0 Å from metal A in VP55, insufficiently close for coordination within the primer unbound form of VP55 shown here. Similar observations regarding the lack of metal coordination for the third aspartate have been made for other polymerases (Steitz, 1997).

### ATP Binding

To understand the features that govern the specificity of the ATP binding site, we determined the structure of the VP55-VP39 heterodimer complex with ATP-γ-S bound at the active site. The bound nucleotide makes numerous specific contacts with active site residues (Figures 2A and 3A) that are invariant throughout the poxvirus PAP family. The triphosphate of the ATP analog is anchored by hydrogen-bonding and charge-coupling interactions with side chains of Ser188 (K helix) and Asp202 (via a water molecule and a metal) of the catalytic domain and Lys304 (P helix) of the C domain. S-γ-1, which substitutes for an oxygen in the native ATP, accepts a hydrogen bond from the OG of Ser188.

The specific recognition of ribose and adenine rings of the nucleotide is achieved by several interactions with residues solely from the C domain, predominantly from the N and O helices. For example, the 2′ OH of the ribose is engaged in "cooperative hydrogen bonds" by acting simultaneously as a donor to the main chain carbonyl oxygen of Gln281 (N helix) and an acceptor from the backbone NH of Met285 (N helix). This finding provides the atomic basis for the biochemical data indicating preference for ATP over 2′ deoxy-ATP (Shuman and Moss, 1988; Gershon and Moss, 1992).

The involvement of the C domain residues Met288 (N helix) and Arg294 (O helix) in adenine recognition is accompanied by conformational changes. In the ATP bound structure, the side chain of Met288 moves to the space occupied by the Arg294 side chain in the unbound structure (data not shown) in order to form hydrophobic interactions with all five carbons of the adenine ring. In turn, the side chain of Arg294 positions itself to hydrogen bond to the adenine endocyclic N7 group. Finally, and more importantly, the OD1 and ND2 groups of Asn402 (located in the loop between β13 and β14 in the C domain) are engaged in bidentate hydrogen bonds

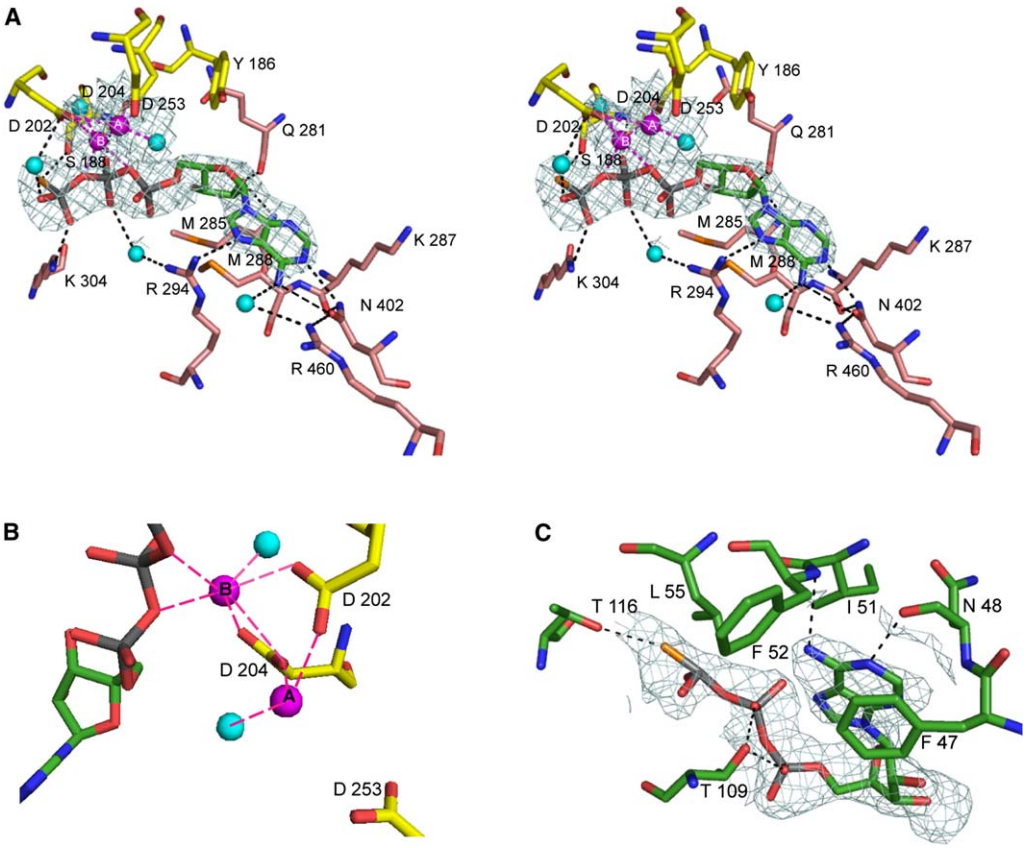


Figure 2. The Active Site of VP55

(A) Stereo view of the ATP analog bound in the active site. Residues in yellow are from the catalytic domain, and those in salmon are from the C domain (see Figure 1A). The $F_o - F_c$ electron density map (cyan) is a simulated annealing omit map, contoured at 2.5 $\sigma$. It was calculated at 2.3 Å from a model in which residues within a 5 Å sphere around the ATP-$\gamma$-S molecule were deleted. Salt links and hydrogen bonds (distances from 2.5 to 3.5 Å) between VP55 residues and the ATP analog are shown as black dotted lines. The two metal ions (magenta) are labeled "A" and "B." The metals coordination spheres are shown as magenta dotted lines. Cyan spheres represent water molecules.

(B) Close-up view of the coordination spheres of metals A and B.

(C) ATP-$\gamma$-S secondary binding site interactions. The electron density map (in blue) corresponds to an $F_o - F_c$ simulated annealing omit map contoured at 3 $\sigma$. This map was calculated by omitting residues lying in a sphere of 5 Å around the nucleotide.

with the exocyclic N6 proton (as the donor) and the endocyclic N1 lone pair (as the acceptor), respectively. These two key interactions by one residue for adenine recognition rely on the precise and stable orientation of the Asn402 side chain as evidenced by further hydrogen-bonding interactions between its OD1 and the NH2 of Arg460 (a residue located in the linker region between the C and the last two $\beta$ strands of the catalytic domain) and ND2 and the main chain carbonyl oxygen of Lys287 (N helix). The NH2 moiety of Arg460 is also hydrogen bonded to N6 via a water molecule.

In contrast to the templated polymerase pol $\beta$, the nontemplated VP55 does not seem to undergo large protein conformational changes upon binding the incoming nucleotide. Thus, structures of complexes of pol $\beta$ with incoming nucleotide and template/primer DNA revealed closure of the thumb subdomain in response to binding of the correctly base-paired incoming nucleotide to the template, allowing interactions of the thumb subdomain with the sugar and base of the nucleotide (Sawaya et al., 1997). In contrast, the C domain of VP55 in the ATP-free structure is already in a position to interact with the sugar and adenine rings of the ATP without the requirement of a large interdomain conformational change. This conformation of the C domain was also observed in the P1 space group structure of the heterodimer in the absence of the ATP analog (C.M.M. and F.A.Q., unpublished data), thus making it unlikely the result of crystal packing contacts. Moreover, because the C domain is buttressed heavily through extensive contacts with the catalytic domain and VP39 (Figure 1A), it is not expected to undergo a large conformational rearrangement.

Difference electron density maps further revealed the presence of another ATP analog molecule (Figure 2C) bound on the surface of VP55 between helices D and H of the N domain on the back face of VP55 (opposite from the active site) (Figure 1A). This ATP is more solvent exposed than the one bound in the active site (265 $Å^2$ accessible surface compared to 62 $Å^2$ for the ATP bound in the active site) and has fewer interactions with VP55 (Figure 2B). The residues involved in the interactions with the second ATP molecule are conserved throughout the poxvirus family. It is notable that none of the triphosphate groups is involved in charge-coupling interactions. Rather, the S$\gamma$-1 is hydrogen bonded to the

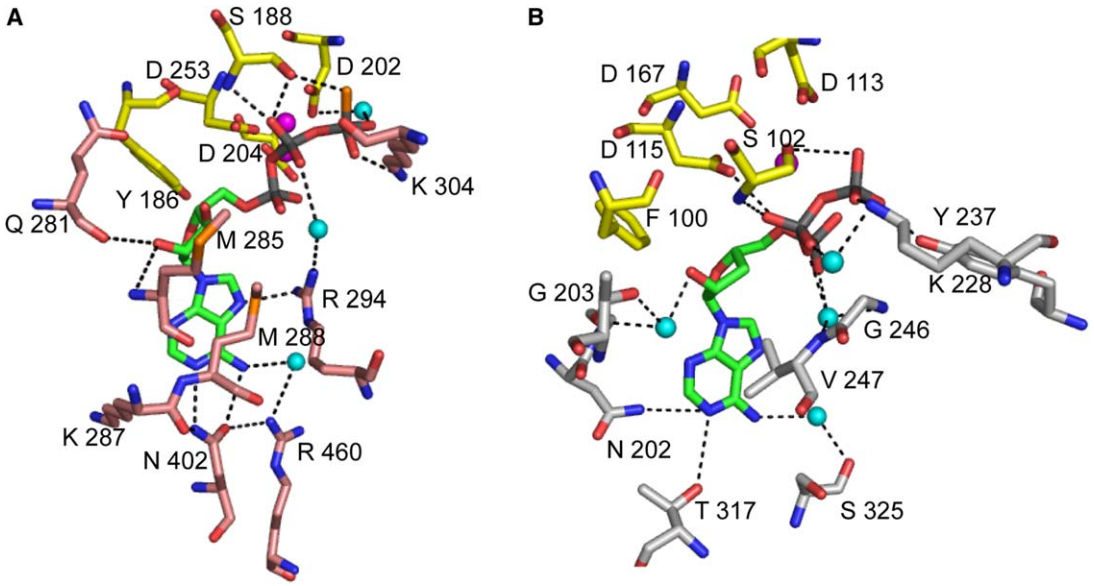


Figure 3. Side-by-Side Comparison of the Modes of Binding of ATP in the Active Site of VP55 and Bovine PAP

(A) and (B) show the active site in VP55 and PAP, respectively. The adenine ring of the bound ATPs is in a very similar orientation. Residues provided by the catalytic domains of VP55 and bovine PAP are shown in yellow. C domain residues of VP55 are colored in salmon, and central domain residues of bovine PAP are colored in gray. The protein-nucleotide interactions in the bovine structure were obtained from the Protein Data Bank (1Q78). Metal ions A and B in VP55 and the single metal ion found in the bovine PAP structure are shown as magenta spheres. Cyan spheres represent water molecules.

OG of Thr116, and the phosphate oxygens O1β and O2α are hydrogen bonded to the OG of Thr109. The adenine N1 is hydrogen bonded to the carbonyl oxygen of the amide side chain of Asn48, and N6 is hydrogen bonded to the main chain NH of Phe52. In addition, the side chains of Phe47 and Ile51 form hydrophobic contacts with the adenine ring. No metal ions were detected at this site.

### Comparison between VP55 and Eukaryotic PAPs

The structures of the eukaryotic 3′-deoxy-ATP-loaded bovine (PDB codes 1F5A and 1Q78; [Martin et al., 2000, 2004]) and yeast (PDB code 1FA0; Bard et al. 2000) PAPs, determined in the absence of processivity factors, are very similar and show a tripartite organization composed of an N-terminal or catalytic domain, a central domain, and a C-terminal domain whose topology resembles that of the RNA recognition motif (RRM) family (Varani and Nagai, 1998). However, they differ significantly from the structure of VP55. In fact, DALI revealed no structural homology between VP55 and the bovine or yeast PAPs. The eukaryotic PAP N-terminal catalytic domain, consisting of the canonical five strands and two connecting helices (e.g., Figure 1C), has very low structural homology (Z value below 2) with the central catalytic domain of VP55.

The locations of the active sites in VP55 and bovine PAP are also different. The active site in VP55 resides between the central (or catalytic) and C domains, whereas the active site in the eukaryotic PAPs resides between the N (or catalytic) and central domains. Thus, it is not unexpected to find that the recognition mechanisms for the adenosine of the ATP bound to VP55 and eukaryotic PAPs are also fundamentally different (Figure 3). Residues in the C domain of VP55, which are involved in ribose and adenine recognition (Lys304, Gln281, Met285, Met288, Arg294, and Asn402) do not match with those in the bovine PAP (or yeast PAP, data not shown). The recognition of the N6 of the adenine, which confers specificity for ATP, is achieved by interaction with Asn402, whereas the mechanism of ATP selectivity could not be ascertained in the bound bovine or yeast PAP structure. Only the residues in the catalytic domains of VP55 and bovine PAP (as well as Lys304 of the C domain of VP55 and Lys288 of the central domain of bovine PAP), which interact with the triphosphate group of the ATP bound in both proteins, have identical side chains (Figure 3). These types of interactions are preserved throughout the NT superfamily.

### RNA Binding Model

Electrostatic surface potential calculations show large areas of intensely positively charged surface on both sides of the VP55-VP39 structure, suggesting a possible role in RNA binding. One area of particular interest is further defined by a groove that begins at the polyadenylation site, tracks along the β sheet region of the catalytic domain, and extends to a segment at the interface between VP55 and VP39 on the front side of the heterodimer. Our modeling study indicates that the groove could accommodate ~13 nt (Figure 4A), including the six rUs at −1 to −5 and −7 that are critical for elongation of poly(A) tails beyond 3–7 nt (Johnson et al., 2004). The RNA binding model on the front face is consistent with crosslinking results, which indicate that the downstream half of the primer interacts with both subunits (Deng et al., 1997; Johnson and Gershon, 1999). One important result demonstrates that the guanidinium group of Arg107 of VP39, which is found in the groove (Figure 4A), forms a short-range (3.0 Å) photocrosslink

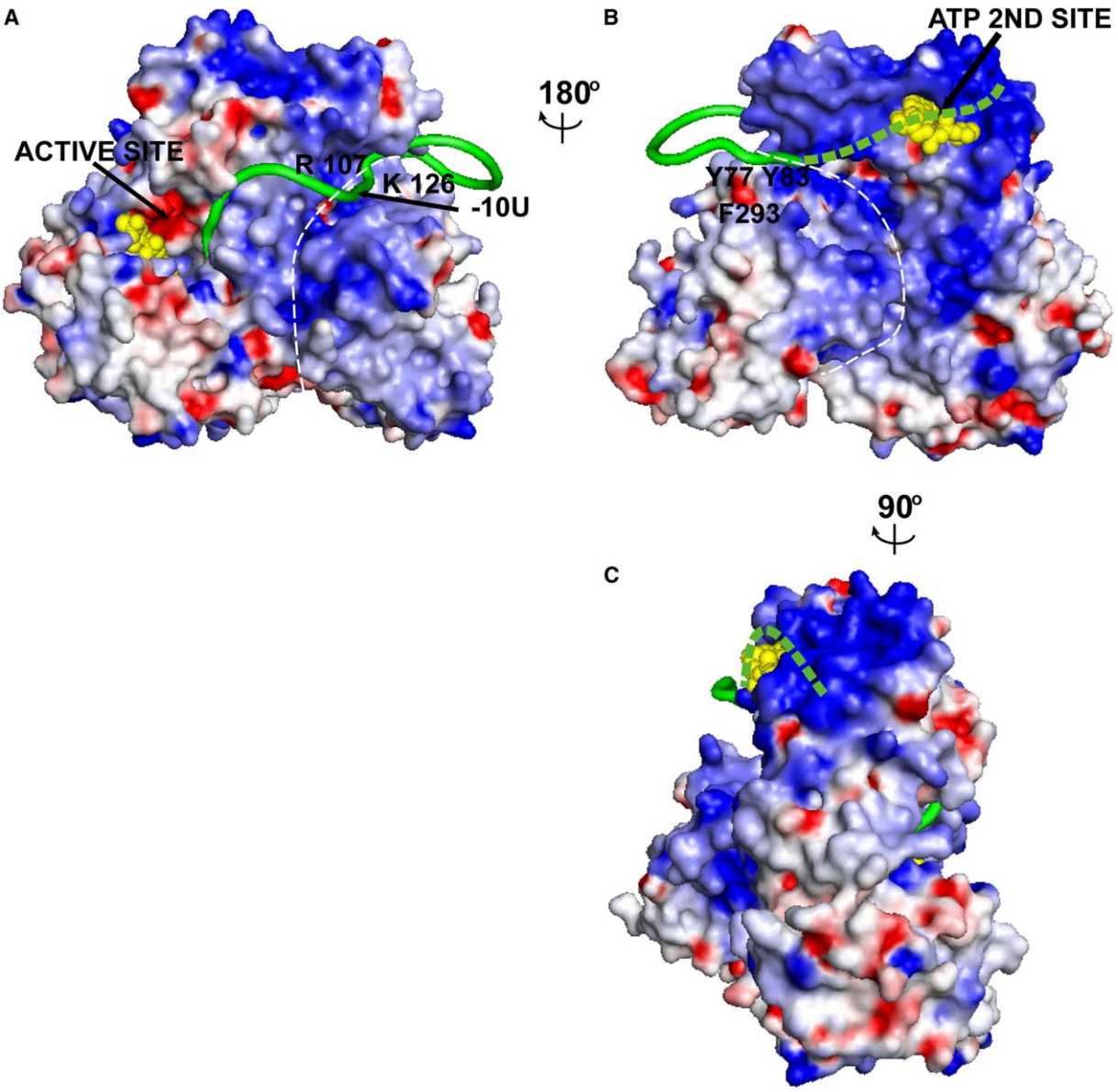


Figure 4. RNA Binding Model

Three views are shown: (A) the front face, which is in identical orientation as in Figure 1A, (B) the back face, and (C) one of the edges. GRASP (Nicholls et al., 1991) electrostatic potential surface is contoured at −10 kT (red) and +10 kT (blue). The white dash line demarcates the interface between VP55 and VP39. The ATP analog at the active site and the second site are depicted as yellow CPK models. A modeled downstream RNA primer binding on the front face (A) is shown in thick green ribbon. The proposed path of the upstream RNA primer along VP55, mostly on the back face, is shown as a green dotted line. Shown in (A) is the close proximity of VP39's Arg107 to the uracil at −10 of the modeled RNA primer. Arg107 has been shown to form a short-range photocrosslink to the uracil (see text). Also indicated are the locations of VP39 residues Lys126 on the front face (A) and Tyr77, Tyr83, and Phe293 on the back side (B) that photocrosslink to nt between −11 and −21.

with the rU at position −10 nt of the cognate RNA (Deng et al., 1999), indicating a possible protein-RNA interaction, and that Lys126, a VP39 residues also located in the groove on the front side, is proximal to the RNA primer according to another crosslinking experiment (Oguro et al., 2002). Further support for the model for binding downstream bases comes from mutagenesis and primer binding studies of bovine PAP (Martin et al., 2004), which indicated that residues from β4 in the β sheet region of the catalytic N domain, which is equivalent to β4 in VP55 (Figures 1B and 1C), interact with the RNA primer. Moreover, the involvement of residues of a β sheet region in polynucleotide binding has been seen in the structures of pol β with bound template/primer DNA (Pelletier et al., 1996) and its structural homolog, the terminal deoxynucleotidyltransferase (Delarue et al., 2002).

In order to accommodate the required ~45 nt length for a stably binding nucleic acid primer (Johnson and Gershon, 1999), the path may continue along the back side of the heterodimer. Photocrosslinking studies indicated that positions −11 to −21 track along the interface between the two subunits (Johnson and Gershon, 1999), and moreover, VP39 residues Tyr77, Tyr83, and Phe293,

located near the interface on the backside (Figure 4B), are proximal to the primer (Oguro et al., 2002). These findings indicate that the RNA is likely to bend within the −11 to −21 segment if it is to continue along the heterodimer interface. Beyond −21 nt, photocrosslinking studies indicate that the affinity-augmenting UU at −36 and −37 interacts solely with VP55 (Johnson and Gershon, 1999). Accordingly, in our modeling, the RNA upstream of ~−24 (represented by a green dotted line in Figures 4B and 4C) deviates from the interface to VP55, specifically to the N domain, which exhibits several features conducive to RNA binding. For example, the backside of the N domain has the highest concentration of surface-positive charges in the entire VP55-VP39 complex structure (Figures 4B and 4C). In addition, the N domain is composed almost entirely of HhH motifs, which have been demonstrated to be docking sites for nucleic acid binding in pol β (Pelletier et al., 1996). Basic helical hairpins have been further implicated in mRNA binding of the yeast protein Shep2 (Niessing et al., 2004), and basic amino acid-rich regions have been shown to be important for RNA binding in other viral proteins (Smith et al., 2000). Moreover, the N domain also harbors the second ATP site (Figures 1A and 4B), which is rich in aromatic residues that can provide stacking interactions with RNA bases.

Overall, our proposed path of the RNA primer along the heterodimer suggests a binding mode for the template-independent VP55-VP39 polymerase that is very different from the polynucleotide path observed in pol β (Pelletier et al., 1994; Sawaya et al., 1994) and other template-dependent polymerases (Doublie et al., 1998; Ollis et al., 1985). In these polymerases, the fingers and thumb subdomains form a binding channel that contacts the nucleotides concomitant with rotational movements of the fingers or thumbs. A large rearrangement of the N and C domains of VP55 that would form a channel similar to those formed by the finger and thumb of template-dependent polymerases while remaining associated to VP39 seems unlikely. Less dramatic conformational changes affecting the individual domains of VP55 as well as the geometry of the active site may occur upon RNA binding.

The 5′ capped RNA binding site in VP39 (shown in Figure 1A), which is sufficient to accommodate the $m^7G$ cap and at least three transcribed ribonucleotides (Hodel et al., 1998), does not overlap with the proposed path of the polyadenylation primer, consistent with evidence that methyl transfer and polyadenylation activities can occur independently (Gershon et al., 1998).

### Processivity and Translocation

Processivity in replicative polymerases is achieved by the binding of accessory proteins or by the involvement of domains that undergo conformational changes that stabilize the interaction of the polymerase to its nucleic acid template/primer, stimulating synthesis and preventing premature termination. The best-characterized processivity factors are the DNA sliding clamps of prokaryote and eukaryote DNA replicases, which oligomerize to form a channel that encircles the DNA duplex without directly interacting with the DNA (Johnson and O'Donnell, 2005; Shamoo and Steitz, 1999). In contrast, the mechanism by which monomeric processivity factors achieve their function is less well understood. The structure of T7 DNA polymerase in complex with its processivity factor, thioredoxin, and DNA primer/template (PDB code 1T7P) (Doublie et al., 1998) shows thioredoxin bound to a flexible region of the thumb domain of the polymerase without making contact with the DNA, suggesting that the complex may be in an open conformation. In another example, the crystal structure of the herpes simplex virus processivity factor UL42 bound to a C peptide of DNA polymerase (PDB code 1DML) (Zuccola et al., 2000) indicates that positive electrostatics on the surface of UL42 drive the DNA into the polymerase. Our structural and RNA modeling studies suggest a different processivity mechanism for VP39 by forming an extensive interface with the catalytic subunit VP55 that seems poised to accommodate the RNA. The model for RNA primer binding suggests that the heterodimer interface partly engulfs the RNA primer, at least the downstream portion that has been indicated by crosslinking experiments to bind between VP55 and VP39 (Johnson and Gershon, 1999; Oguro et al., 2002). This in turn would stabilize RNA binding, preventing its dissociation from the heterodimer.

The molecular mechanism of translocation of template-independent polymerases is not well understood. For VP55, functional and crosslinking studies with its RNA primer have thus far shown that the RNA translocates discontinuously toward the active site (Johnson et al., 2004) in a process that involves some flexibility of the catalytic domain of the protein and probably also of the RNA primer. The RNA primer would be initially anchored by the interactions of −10 rU with the interface VP39 residue Arg107 on the front side and −36 and −37 rUs with residues presumably on the backside of VP55. In addition, a patch of six rUs is required at the 3′ end of the RNA to allow for tail elongation. Crosslinking of the RNA primer to −4 was found to prevent the initiation of poly(A) tail, whereas anchoring to positions −11, −16, −21, and −28 did not prevent the initial extension of the tail. This suggests movement of the region of the protein in contact with the first 10 nt of the primer while upstream regions of the RNA, including the essential rUs at −36 and −37 remain anchored in the first stages of poly(A) tail elongation. The region contacting the first 10 nt in the model presented here corresponds to the β sheet of the catalytic domain, a platform known for docking RNA or DNA primer in polymerases.

### Concluding Remarks

To our knowledge, the VP55-VP39 crystal structure represents the first for a multiprotein complex from a poxvirus, and the first structure to be determined for a viral poly(A) polymerase or for a PAP complexed with its processivity factor. VP55 is composed of three domains with different motifs and functions. The central domain plays multiple key roles by providing the three catalytic aspartates that coordinate the two metals, interacting with the ATP triphosphate group and most certainly participating, along with VP39, in binding the first few downstream nucleotides of the RNA primer. The C domain deploys all the residues that recognize specifically the adenine group of the ATP. The N domain is the only domain lacking a well-defined role, but it has all the earmarks (HhH motif, a nucleotide binding site, and intense

electropositive surface interspersed with aromatic residues) for binding upstream primer nucleotides. Furthermore, portions of all three domains at the periphery of the concave side of VP59 dovetail extensively with the globular single domain processivity factor VP39. The heterodimer interface is defined by a long groove that may partially enclose the RNA, thus preventing its dissociation from the polymerase. Further studies based on the tertiary structure of the heterodimer are required to further deepen our understanding of the polymerase function, including the mechanism by which the mRNA is translocated at its 3′ end and positioned for catalysis after every additional cycle while maintaining the integrity of the heterodimer.

## Experimental Procedures

### Protein Purification and Crystallization

AS11 mutant VP39 and full-length VP55 were expressed in *E. coli* as GST- and N-terminal 6His-tagged forms, respectively, and purified as described (Gershon, 2004), except that VP39 was expressed in BL21 DE3 (Novagen). Due to heterogeneity of the N-terminal end of VP55 and its degradation during protein storage, VP55 was recloned as a stable Δ10N form with residue Pro 11 corresponding to the start of the first predicted helix in the structure. Selenomethionine substitution led to dramatically decreased yield and solubility of VP55. Therefore, the selenomethionine variant of VP39, prepared by inhibition of the methionine synthesis pathway (Doublie, 1997), was used for structure determination.

For crystal growth, individually purified VP55 and VP39 were mixed at 1:1 ratio in 75 mM NaCl and 10 mM Tris-HCl (pH 8.7) then concentrated to about 6 mg/mL by ultrafiltration. Crystals were grown in hanging drops at room temperature by mixing equal volume of the protein solution and the well solution containing 18%–21% PEG 4000 (w/v), 0.24 M $CaCl_2$, 5% glycerol (v/v), and 10 mM Tris-HCl (pH 8.7) over several days. Two crystals forms were obtained under these conditions: one in the $P2_12_12_1$ space group and the other in the P1 space group. Crystals were harvested in a cryoprotectant solution comprising the well solution plus 20% glycerol, then frozen in liquid nitrogen. ATP-γ-S (Sigma) was soaked into the crystals at 3 mM concentration for a period of 2 days.

### Data Collection, Structure Determination, and Refinement

A 2.4 Å resolution dataset from VP55-VP39 complex crystal in the P1 space group containing two complexes in the asymmetric unit was collected at the Advanced Photon Source (APS) 19-ID beamline. Although it was possible to obtain initial phases by molecular replacement using the structure of VP39 as a search model, the resulting electron density map for VP55, which amounts to two-thirds of the total protein scattering matter in the asymmetric unit, was difficult to interpret. Therefore, a two-wavelength MAD data set was collected on the orthorhombic crystal at beamline 17-ID at the APS. Data were processed with the HKL2000 software (Otwinowski and Minor, 1997). The orthorhombic crystals contained one heterodimer in the asymmetric unit. Eight selenium sites were found with Shake-and-Bake (Howell et al., 2000). Phases were obtained with SHARP (de La Fortelle and Bricogne, 1997) and solvent flattened with SOLOMON (Abrahams and Leslie, 1996) at 3.0 Å. The resulting electron density map was interpretable. VP39 was placed in the map by molecular replacement using CNS (Brunger et al., 1998). VP55 was traced in the experimental electron density map as a polyalanine model in O (Jones et al., 1991). VP55 side chains were fitted iteratively in maps calculated with combined experimental and calculated phases. To confirm the validity of VP55 sequence, we sequenced the DNA of the plasmid used to overexpress VP55. We identified a mutation, on residue Leu 36 → Ser with respect to the WR VP55 sequence (accession number YP_232939). This residue is located on the surface of VP55, away from the active site and the heterodimeric interface. The mutation reflects differences in virus phylogeny. The model was refined in CNS by conjugate gradient minimization and isotropic B factor refinement by using the peak structure factors. The refined coordinates were used as a molecular replacement model to solve the P1 structure (data not shown) and for refinement of the ATP-γ-S bound data set. Water molecules, divalent ions, and ATP-γ-S were located in $F_o - F_c$ maps. Accessible surface area calculations were performed in CNS. Although previous crystallographic studies on VP39 alone showed an endogenous AdoMet in the methyltransferase active site (Hodel et al., 1996), no electron density for the cofactor was found at the V39 active site in the structures reported here. Figure 1A was prepared with MOLSCRIPT (Kraulis, 1991) and rendered with RASTER-3D (Merrit and Murphy, 1994). Figures 2–4 were made with PyMOL (DeLano, 2002).

### Acknowledgments

We thank the personnel at Advanced Photon Source for their assistance and Dr. K Bataille for the data collected at beamline 17-ID and Drs. N. Duke and S. Ginell for the data collected at beamline 19-ID that is supported by the U. S. Department of Energy, Office of Energy Research, under Contract No. W-31-109-ENG-38. We thank C. Li and H. Li for purifying VP55. The bulk of the work was supported by a Howard Hughes Medical Institute grant to F.A.Q. This work was further supported by National Institutes of Health grant GM51953 and other funds available to P.D.G and the Welch Foundation to F.A.Q (Q-0581). This manuscript is dedicated to the memory of Alec Hodel.

Received: January 25, 2006
Revised: March 3, 2006
Accepted: March 10, 2006
Published: May 4, 2006

### References

Abrahams, J.P., and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. D Biol. Crystallogr. *52*, 30–42.

Aravind, L., and Koonin, E.V. (1999). DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res. *27*, 1609–1618.

Bard, J., Zhelkovsky, A.M., Helmling, S., Earnest, T.N., Moore, C.L., and Bohm, A. (2000). Structure of yeast poly(A) polymerase alone and in complex with 3'-dATP. Science *289*, 1346–1349.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta. Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Castano, I.B., Heath-Pagliuso, S., Sadoff, B.U., Fitzhugh, D.J., and Christman, M.F. (1996). A novel family of TRF (DNA topoisomerase I-related function) genes required for proper nuclear segregation. Nucleic Acids Res. *24*, 2404–2410.

Davies, J.F., 2nd, Almassy, R.J., Hostomska, Z., Ferre, R.A., and Hostomsky, Z. (1994). 2.3 A crystal structure of the catalytic domain of DNA polymerase beta. Cell *76*, 1123–1133.

de La Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472–494.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System on World Wide Web (http://www.pymol.org).

Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990). An attempt to unify the structure of polymerases. Protein Eng. *3*, 461–467.

Delarue, M., Boule, J.B., Lescar, J., Expert-Bezancon, N., Jourdan, N., Sukumar, N., Rougeon, F., and Papanicolaou, C. (2002). Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. EMBO J. *21*, 427–439.

Deng, L., and Gershon, P.D. (1997). Interplay of two uridylate-specific RNA binding sites in the translocation of poly(A) polymerase from vaccinia virus. EMBO J. *16*, 1103–1113.

Deng, L., Beigelman, L., Matulic-Adamic, J., Karpeisky, A., and Gershon, P.D. (1997). Specific recognition of an rU2–N15-rU motif by VP55, the vaccinia virus poly(A) polymerase catalytic subunit. J. Biol. Chem. *272*, 31542–31552.

Deng, L., Johnson, L., Neveu, J.M., Hardin, S., Wang, S.M., Lane, W.S., and Gershon, P.D. (1999). A polyadenylylation-specific RNA-contact site on the surface of the bifunctional vaccinia virus RNA modifying protein VP39 that is distinct from the mRNA 5′ end-binding "cleft". J. Mol. Biol. *285*, 1417–1427.

Doublie, S. (1997). Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. *276*, 523–530.

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. Nature *391*, 251–258.

Gershon, P.D. (2004). Studying vaccinia virus RNA processing in vitro. Methods Mol. Biol. *269*, 151–168.

Gershon, P.D., and Moss, B. (1992). Transition from rapid processive to slow nonprocessive polyadenylation by vaccinia virus poly(A) polymerase catalytic subunit is regulated by the net length of the poly(A) tail. Genes Dev. *6*, 1575–1586.

Gershon, P.D., and Moss, B. (1993a). Stimulation of poly(A) tail elongation by the VP39 subunit of the vaccinia virus-encoded poly(A) polymerase. J. Biol. Chem. *268*, 2203–2210.

Gershon, P.D., and Moss, B. (1993b). Uridylate-containing RNA sequences determine specificity for binding and polyadenylation by the catalytic subunit of vaccinia virus poly(A) polymerase. EMBO J. *12*, 4705–4714.

Gershon, P.D., Ahn, B.Y., Garfield, M., and Moss, B. (1991). Poly(A) polymerase and a dissociable polyadenylation stimulatory factor encoded by vaccinia virus. Cell *66*, 1269–1278.

Gershon, P.D., Shi, X., and Hodel, A.E. (1998). Evidence that the RNA methylation and poly(A) polymerase stimulatory activities of vaccinia virus protein VP39 do not impinge upon one another. Virology *246*, 253–265.

Hodel, A.E., Gershon, P.D., Shi, X., and Quiocho, F.A. (1996). The 1.85 A structure of vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. Cell *85*, 247–256.

Hodel, A.E., Gershon, P.D., Shi, X., Wang, S.M., and Quiocho, F.A. (1997). Specific protein recognition of an mRNA cap through its alkylated base. Nat. Struct. Biol. *4*, 350–354.

Hodel, A.E., Gershon, P.D., and Quiocho, F.A. (1998). Structural basis for sequence-nonspecific recognition of 5′-capped mRNA by a cap-modifying enzyme. Mol. Cell *1*, 443–447.

Holm, L., and Sander, C. (1996). Mapping the protein universe. Science *273*, 595–603.

Howell, P.L., Blessing, R.H., Smith, G.D., and Weeks, C.M. (2000). Optimizing DREAR and SnB parameters for determining Se-atom substructures. Acta Crystallogr. D Biol. Crystallogr. *56*, 604–617.

Hu, G., Gershon, P.D., Hodel, A.E., and Quiocho, F.A. (1999). mRNA cap recognition: dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains. Proc. Natl. Acad. Sci. USA *96*, 7149–7154.

Hu, G., Oguro, A., Li, C., Gershon, P.D., and Quiocho, F.A. (2002). The "cap-binding slot" of an mRNA cap-binding protein: quantitative effects of aromatic side chain choice in the double-stacking sandwich with cap. Biochemistry *41*, 7677–7687.

Johnson, A., and O'Donnell, M. (2005). Cellular DNA replicases: components and dynamics at the replication fork. Annu. Rev. Biochem. *74*, 283–315.

Johnson, L., and Gershon, P.D. (1999). RNA binding characteristics and overall topology of the vaccinia poly(A) polymerase-processivity factor-primer complex. Nucleic Acids Res. *27*, 2708–2721.

Johnson, L., Liu, S., and Gershon, P.D. (2004). Molecular flexibility and discontinuous translocation of a non-templated polymerase. J. Mol. Biol. *337*, 843–856.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. *24*, 946–950.

Martin, G., Keller, W., and Doublie, S. (2000). Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. EMBO J. *19*, 4193–4203.

Martin, G., Moglich, A., Keller, W., and Doublie, S. (2004). Biochemical and structural insights into substrate binding and catalytic mechanism of mammalian poly(A) polymerase. J. Mol. Biol. *341*, 911–925.

Merrit, E., and Murphy, M. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. Acta. Crystallogr. D Biol. Crystallogr. *50*, 869–873.

Moss, B., Rosenblum, E.N., and Gershowitz, A. (1975). Characterization of a polyriboadenylate polymerase from vaccinia virions. J. Biol. Chem. *250*, 4722–4729.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Niessing, D., Huttelmaier, S., Zenklusen, D., Singer, R.H., and Burley, S.K. (2004). She2p is a novel RNA binding protein with a basic helical hairpin motif. Cell *119*, 491–502.

Oguro, A., Johnson, L., and Gershon, P.D. (2002). Path of an RNA ligand around the surface of the vaccinia VP39 subunit of its cognate VP39–VP55 protein heterodimer. Chem. Biol. *9*, 679–690.

Ollis, D.L., Kline, C., and Steitz, T.A. (1985). Domain of E. coli DNA polymerase I showing sequence homology to T7 DNA polymerase. Nature *313*, 818–819.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H., and Kraut, J. (1994). Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. Science *264*, 1891–1903.

Pelletier, H., Sawaya, M.R., Wolfle, W., Wilson, S.H., and Kraut, J. (1996). Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. Biochemistry *35*, 12742–12761.

Quiocho, F.A., Hu, G., and Gershon, P.D. (2000). Structural basis of mRNA cap recognition by proteins. Curr. Opin. Struct. Biol. *10*, 78–86.

Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H., and Kraut, J. (1994). Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. Science *264*, 1930–1935.

Sawaya, M.R., Prasad, R., Wilson, S.H., Kraut, J., and Pelletier, H. (1997). Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: Evidence for an induced fit mechanism. Biochemistry *36*, 11205–11215.

Schnierle, B.S., Gershon, P.D., and Moss, B. (1992). Cap-specific mRNA (nucleoside-O2'-)-methyltransferase and poly(A) polymerase stimulatory activities of vaccinia virus are mediated by a single protein. Proc. Natl. Acad. Sci. USA *89*, 2897–2901.

Shamoo, Y., and Steitz, T.A. (1999). Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. Cell *99*, 155–166.

Shi, X., Bernhardt, T.G., Wang, S.M., and Gershon, P.D. (1997). The surface region of the bifunctional vaccinia RNA modifying protein VP39 that interfaces with Poly(A) polymerase is remote from the RNA binding cleft used for its mRNA 5′ cap methylation function. J. Biol. Chem. *272*, 23292–23302.

Shimizu, K., Santocanale, C., Ropp, P.A., Longhese, M.P., Plevani, P., Lucchini, G., and Sugino, A. (1993). Purification and characterization of a new DNA polymerase from budding yeast Saccharomyces cerevisiae. A probable homolog of mammalian DNA polymerase beta. J. Biol. Chem. *268*, 27148–27153.

Shuman, S., and Moss, B. (1988). Vaccinia virus poly(A) polymerase. Specificity for nucleotides and nucleotide analogs. J. Biol. Chem. *263*, 8405–8412.

Smith, C.A., Calabro, V., and Frankel, A.D. (2000). An RNA-binding chameleon. Mol. Cell *6*, 1067–1076.

Steitz, T.A. (1997). DNA and RNA polymerases: structural diversity and common mechanisms. Harvey Lect. *93*, 75–93.

Tomita, K., Fukai, S., Ishitani, R., Ueda, T., Takeuchi, N., Vassylyev, D.G., and Nureki, O. (2004). Structural basis for template-independent RNA polymerization. Nature *430*, 700–704.

Varani, G., and Nagai, K. (1998). RNA recognition by RNP proteins during RNA processing. Annu. Rev. Biophys. Biomol. Struct. *27*, 407–445.

Zuccola, H.J., Filman, D.J., Coen, D.M., and Hogle, J.M. (2000). The crystal structure of an unusual processivity factor, herpes simplex virus UL42, bound to the C terminus of its cognate polymerase. Mol. Cell *5*, 267–278.

## Accession Numbers

Two sets of coordinates with accession numbers 2GA9 (with bound ATP-γ-S) and 2GAF (without bound nucleotide) have been deposited in the Protein Data Bank.