RNA binding characteristics and overall topology of the vaccinia poly(A) polymerase-processivity factor-primer complex

Lela Johnson and Paul D. Gershon*

Department of Medical Biochemistry and Genetics/Institute of Biosciences and Technology, Texas A&M University, Houston, TX 77030, USA

Received March 12, 1999; Revised and Accepted May 19, 1999

ABSTRACT

The vaccinia virus-encoded heterodimer responsible for poly(A) tail elongation comprises a polyadenylation catalytic subunit (VP55) and associated processivity factor (VP39). We show that monomeric VP39's affinity for RNA homopolymers follows the hierarchy poly(I) > poly(U) >> poly(G) > poly(A) > poly(C), that the heterodimer interacts stably with 40–45 nucleotide nucleic acid segments, and that its homopolymer preference for polyadenylation priming is comparable to the VP39 affinity hierarchy (above). For oligonucleotide ligands possessing the previously-identified (rU)2-(N)25-rU heterodimer-binding motif, the heterodimer's affinity and base-type preference are mediated via both the (rU)2 and rU portions, with the greater contribution coming from (rU)2. VP39's R107 sidechain contributes to specificity at the downstream rU. Substitution of each ribouridylate of the motif with either ribothymidine or 4-thiodeoxythymidine indicated that the downstream rU interacts with both heterodimer subunits, whereas the upstream (rU)2 interacts only with VP55. A 'crosslinking SELEX' approach indicated VP39-base proximity around position –10 of a 4-thioribouridine/deoxycytidine ligand pool. Upon incubating the heterodimer with a panel of identical-sequence oligonucleotides derivatized with azidophenacyl bromide at various phosphate positions, those derivatized at positions –11 to –21 photocrosslinked to both subunits in a coordinated manner. This region may therefore pass through a 'cleft' or enclosed 'channel' at the subunit interface.

INTRODUCTION

The vaccinia virus poly(A) polymerase (PAP) is a heterodimer of the virus-encoded proteins VP55 and VP39 (1). Although the VP55 subunit carries out the polyadenylation catalytic functions of the heterodimer (1), isolated VP55 can synthesize poly(A) tails no more than ~30 nt in length before abruptly dissociating from its primer (2). VP39 acts as a processivity factor for VP55 (3), enabling processivity to be maintained during the elongation of tails to several hundred nucleotides in length. VP39 has a second, entirely unrelated function at the mRNA 5' end, as a mRNA cap-specific 2'-O-methyltransferase (4).

Although the physical structure and structure–function relationships of VP55 remain undetermined, characteristics of its RNA substrate interaction have been studied in considerable detail. VP55 binds ~34 nt RNA segments (5,6), within which it recognizes an (rU)2-(N)15-rU motif (7). VP55 translocates as it elongates the tail, and its translocation to non-uridylate-containing RNA tracts apparently accounts for its dissociation from the RNA substrate after adding tails no more than ~30 nt in length (6). In contrast to VP55, the VP39 subunit has provided high-resolution crystal structures in combination with its methyltransferase cofactor (8), cofactor-product plus methyltransferase substrate RNA fragment (9) and cofactor-product plus intact methyltransferase substrate RNA (10). VP39 appears to possess multiple RNA binding sites (10–12). Consistent with this, VP39's two RNA end-modifying functions apparently utilize distinct regions of the protein surface (11,12). The mRNA 5' end binding site, which comprises a cleft juxtaposed to the protein's methyltransferase catalytic center, has been defined in atomic detail from the crystallized protein–RNA complex (10). Much less information is available regarding VP39's RNA binding site of for tail-elongation. Nonetheless, the VP55–VP39 heterodimer (the active species in tail elongation) has been shown to recognize an (rU)2-(N)25-rU motif, the downstream rU of which photocrosslinks strongly to VP39. The photocrosslinked VP39 sidechain (that of residue R107) is located away from the methyltransferase cleft, and appears to be intimately associated with the VP55–VP39 dimerization interface.

A better understanding of the topology of the heterodimer–primer ternary complex would contribute greatly to our understanding of VP39’s mechanism of processivity. Therefore, in this study, we have set out to more fully characterize this complex.

MATERIALS AND METHODS

Materials

VP55 (~30 µg/ml) was expressed from recombinant baculovirus-infected 'High Five' insect cells and purified as described (13).
The N-terminally modified VP39 variant VP39-ΔC26 is described in Shi et al. (13). The protein was expressed and purified as described previously (8) and obtained at a final concentration of 0.2 mg/ml. VP39 mutants were kindly provided by L. Deng (12). Oligonucleotides were synthesized using a Model 392 DNA/RNA synthesizer (ABI). The majority of oligonucleotides were synthesized using the reagents and methods of Cruachem, Inc. Beauchage reagent (Glen Research) was used for the synthesis of phosphorothioate-substituted oligonucleotides according to the manufacturer’s directions. For oligonucleotides containing ribothymidine, all RNA phosphoramidites were obtained from Chemgenes Inc. and used according to the manufacturer’s directions. All synthetic oligonucleotides were purified by urea–PAGE before use. Heterogeneous-length RNA homoribopolymer preparations were obtained from ICN and Amersham-Pharmacia, Inc. Azido-geneous-length RNA homopolymer and oligonucleotide labeling; purification of labeled RNA

Prior to 5'-end labeling, heterogeneous-length RNA homopolymers were treated with HK phosphatase (Epcentre Technologies) according to the manufacturer’s directions, followed by heat-inactivation of the enzyme. Reaction mixtures were then supplemented with [γ-32P]ATP (DuPont NEN), 5x polyribonucleotide kinase reaction buffer and T4 polyribonucleotide kinase (NEB), and incubated according to the manufacturer’s directions followed by heat inactivation of the enzyme. For the nitrocellulose-filter RNA binding assay, the labeled homopolymeric RNA was phenol/chloroform extracted, chloroform extracted, and applied to a 300 × 8 mm column containing Sephadex G-50 (DNA Grade Fine) that had been equilibrated with 20 mM triethylammonium bicarbonate (pH 8.5). After elution in the same buffer, fractions containing labeled RNA were dried under vacuum and redissolved in DEPC-treated water. For the ‘minimum size’ experiment (Fig. 3), 5'-end-labeled poly(U) (which had not been treated with phosphatase prior to labeling) was size fractionated by urea–PAGE, and the fractionated RNA isolated by ‘crush-and-soak’ of the resulting gel slice in 0.1% SDS/1 mM EDTA. Synthetic oligonucleotides were 5'-end-labeled using [γ-32P]ATP and T4 polyribonucleotide kinase followed by heat inactivation of the enzyme, as described previously (7). Oligonucleotides were 3'-end-labeled using [α-32P]cystidine triphosphate, VP55 and VP39, as described previously (7,14). After all oligonucleotide labeling procedures, oligonucleotides were NaOAc/ethanol precipitated and redissolved in water.

Oligonucleotide APB-derivatization

Phosphorothioate-containing oligonucleotides were derivatized with APB as described (15). Coupling efficiencies were determined by quantitation of gel shift bands as described (15).

Nitrocellulose filter RNA binding assay

Nitrocellulose filter RNA binding assays were conducted as described (1). Briefly, 5'-end-labeled and unlabeled (competitor) RNAs were combined with binding buffer (10% glycerol, 20 mM NaCl, 0.1 mM EDTA, 50 mM Tris, pH 8.0) in a total volume of 0.1 ml, then supplemented with highly purified VP39-ΔC26 [5 mg/ml (9)] to a final protein concentration of 50 µg/ml. After incubation at room temperature for 10–15 min, reactions were supplemented with 1 ml of ice-cold binding buffer and immediately deposited onto 24 mm nitrocellulose filters (BA85, Schleicher & Schuell) under suction. Filters were washed rapidly with 5 ml of ice-cold binding buffer under suction, then dried and the filter-bound radioactivity quantitated by Cerenkov counting.

Polyadenylation competition assay

Polyadenylation assays were performed as described (3), except that VP55 and VP39 were mixed with unlabeled competitor homopolymers prior to adding 5'-end-labeled oligonucleotide primer, buffer components and rATP. In all assays, primer was present in several-fold molar excess over VP55.

Photocrosslinking reactions

Photocrosslinking reactions with 4S rU- or 4S dT-substituted oligonucleotides contained 0.5 µM oligonucleotide, 30 nM VP55 and 40 nM VP39, in 50 mM Tris (pH 9.0), 5 mM dithiothreitol (DTT), 10% glycerol (TDG buffer). The total volume of the mixture was 15 µL. After 15 min incubation at room temperature, mixtures were irradiated for 10 min at 4°C using a 350 nm light source. For all photocrosslinking experiments (except those for the experiment shown in Fig. 8B), photoinactivated mixtures were applied directly to pre-electrophoresed (~1 h) electrophoretic mobility shift assay (EMSA) gels (below).

EMSA and competition EMSA

The EMSA was performed as described previously (7). Where necessary, bands were excised and the radiolabeled contents recovered by overnight soaking of the gel slices in 0.1% SDS/1 mM EDTA, followed by NaOAc/ethanol precipitation of the supernatant. For EMSA competition experiments, labeled probe oligonucleotides were mixed with various amounts of unlabeled competitor in TDG buffer before adding proteins. After electrophoresis, the 32P for competitor-containing gel lanes to the mean value of P for each crosslinking SELEX experiment, end-labeled oligonucleotides were NaOAc/ethanol precipitated and redissolved in water. Oligonucleotide sequences are given in Deng et al. (12). Within oligonucleotides, we refer to the Xth position from the 3' end as position ‘X’.

EMSA and competition EMSA

The EMSA was performed as described previously (7). Where necessary, bands were excised and the radiolabeled contents recovered by overnight soaking of the gel slices in 0.1% SDS/1 mM EDTA, followed by NaOAc/ethanol precipitation of the supernatant. For EMSA competition experiments, labeled probe oligonucleotides were mixed with various amounts of unlabeled competitor in TDG buffer before adding proteins. After electrophoresis, the 32P for competitor-containing gel lanes to the mean value of P for each crosslinking SELEX experiment, end-labeled oligonucleotides were NaOAc/ethanol precipitated and redissolved in water. Oligonucleotide sequences are given in Deng et al. (12). Within oligonucleotides, we refer to the Xth position from the 3' end as position ‘X’.

EMSA and competition EMSA

The EMSA was performed as described previously (7). Where necessary, bands were excised and the radiolabeled contents recovered by overnight soaking of the gel slices in 0.1% SDS/1 mM EDTA, followed by NaOAc/ethanol precipitation of the supernatant. For EMSA competition experiments, labeled probe oligonucleotides were mixed with various amounts of unlabeled competitor in TDG buffer before adding proteins. After electrophoresis, the 32P for competitor-containing gel lanes to the mean value of P for each crosslinking SELEX experiment, end-labeled oligonucleotides were NaOAc/ethanol precipitated and redissolved in water. Oligonucleotide sequences are given in Deng et al. (12). Within oligonucleotides, we refer to the Xth position from the 3' end as position ‘X’.

EMSA and competition EMSA

The EMSA was performed as described previously (7). Where necessary, bands were excised and the radiolabeled contents recovered by overnight soaking of the gel slices in 0.1% SDS/1 mM EDTA, followed by NaOAc/ethanol precipitation of the supernatant. For EMSA competition experiments, labeled probe oligonucleotides were mixed with various amounts of unlabeled competitor in TDG buffer before adding proteins. After electrophoresis, the 32P for competitor-containing gel lanes to the mean value of P for each crosslinking SELEX experiment, end-labeled oligonucleotides were NaOAc/ethanol precipitated and redissolved in water. Oligonucleotide sequences are given in Deng et al. (12). Within oligonucleotides, we refer to the Xth position from the 3' end as position ‘X’.

EMSA and competition EMSA

The EMSA was performed as described previously (7). Where necessary, bands were excised and the radiolabeled contents recovered by overnight soaking of the gel slices in 0.1% SDS/1 mM EDTA, followed by NaOAc/ethanol precipitation of the supernatant. For EMSA competition experiments, labeled probe oligonucleotides were mixed with various amounts of unlabeled competitor in TDG buffer before adding proteins. After electrophoresis, the 32P for competitor-containing gel lanes to the mean value of P for each crosslinking SELEX experiment, end-labeled oligonucleotides were NaOAc/ethanol precipitated and redissolved in water. Oligonucleotide sequences are given in Deng et al. (12). Within oligonucleotides, we refer to the Xth position from the 3' end as position ‘X’.

EMSA and competition EMSA

The EMSA was performed as described previously (7). Where necessary, bands were excised and the radiolabeled contents recovered by overnight soaking of the gel slices in 0.1% SDS/1 mM EDTA, followed by NaOAc/ethanol precipitation of the supernatant. For EMSA competition experiments, labeled probe oligonucleotides were mixed with various amounts of unlabeled competitor in TDG buffer before adding proteins. After electrophoresis, the 32P for competitor-containing gel lanes to the mean value of P for each crosslinking SELEX experiment, end-labeled oligonucleotides were NaOAc/ethanol precipitated and redissolved in water. Oligonucleotide sequences are given in Deng et al. (12). Within oligonucleotides, we refer to the Xth position from the 3' end as position ‘X’.

EMSA and competition EMSA

The EMSA was performed as described previously (7). Where necessary, bands were excised and the radiolabeled contents recovered by overnight soaking of the gel slices in 0.1% SDS/1 mM EDTA, followed by NaOAc/ethanol precipitation of the supernatant. For EMSA competition experiments, labeled probe oligonucleotides were mixed with various amounts of unlabeled competitor in TDG buffer before adding proteins. After electrophoresis, the 32P for competitor-containing gel lanes to the mean value of P for each crosslinking SELEX experiment, end-labeled oligonucleotides were NaOAc/ethanol precipitated and redissolved in water. Oligonucleotide sequences are given in Deng et al. (12). Within oligonucleotides, we refer to the Xth position from the 3' end as position ‘X’.
extracted (as above), NaOAc/ethanol precipitated and subjected to SDS–PAGE (10% polyacrylamide gel). Radiolabeled bands corresponding to photocrosslinked protein—oligonucleotide conjugates were again identified using the Phosphor-Imager, excised, and the gel slice incubated overnight in 1 mM EDTA. The resulting supernatant, containing ~5 to ~20 nM radiolabeled conjugate, was supplemented with 0.1 M Tris–HCl (pH 7.5), 10 mM EDTA, 0.5% SDS and 1 mg/ml pronase (Boehringer Mannheim; diluted from a 20 mg/ml stock solution in water) and incubated for 20 min at 55°C. The released oligonucleotide was recovered by NaOAc/ethanol precipitation and redissolved in 20 µl water. After supplementing with 13.5 M trifluoroacetic acid (TFA, Pierce) to a final TFA concentration of ~4.5 M and incubating for 5 min at 70°C, samples were neutralized by supplementing with 2.5 M NaOH (final) and 1 M Tris–HCl, pH 8.0 (final). After a final round of NaOAc/ethanol precipitation, samples were redissolved in 2 µl water, supplemented with an equal volume of formamide and subjected to urea–PAGE (38 cm, 13% polyacrylamide gel).

**Line graph quantitation**

Line objects drawn through images of gel lanes were quantitated along their lengths (ImageQuant software, Molecular Dynamics), and the resulting values were exported as spreadsheets and plotted externally. To plot differences in radioactivity between two comparable gel lanes, numerical values derived from the corresponding line objects were subtracted within the spreadsheet after appropriate alignment of the relevant spreadsheet columns.

**RESULTS**

**VP39 is a base-type-specific RNA binding protein**

The previous observation that VP39 binds polypurines in preference to poly(C) (1) led us to examine VP39’s base-type specificity more closely. Homopolymer-competition experiments were performed in which heterogeneous-length poly(A), poly(C), poly(G), poly(U) and poly(I) homoribopolymer preparations were each radiolabeled and subjected to competition with the other, unlabeled, non-hybridizing members of the homopolymer set in the nitrocellulose filter RNA binding assay (Fig. 1). Taken together, the data showed that VP39 exhibits a clear base-type preference with the following hierarchy (high to low affinity): poly(I) > poly(U) > poly(A) > poly(G). This closely reflects the homopolymer binding-affinity hierarchy of the VP39 subunit (Fig. 1). However, the PAP catalytic subunit (VP55) also polyadenylates homoribopolymers in a base-type-specific manner (6), and also follows the above hierarchy [disregarding poly(I), which was not tested with VP55].

**The heterodimer interacts with oligonucleotide tracts 40–45 nt in length**

Since the polyadenylylation functions of VP39 require heterodimer formation, further investigations of VP39’s polyadenylation-specific RNA binding site were conducted in the context of the VP39–VP55 heterodimer. The minimum length of oligonucleotide commensurate with stable heterodimer binding was determined by end-labeling an aliquot of heterogeneous-length poly(U), isolating products in the size-range 40–45 nt or greater in length, and examining the resulting oligo(U) preparation with VP55–VP39 heterodimer, excising the labeled ternary complex from an EMSA gel and recovering the complexed oligo(U) fragments. As a control, the equivalent experiment was done with monomeric VP55. The isolated fragments were size-fractionated by urea–PAGE alongside a size marker comprising chemically synthesized (rU)_{34} (Fig. 3). Oligo(U) segments ≥29 nt in length were able to interact detectably with monomeric VP55 (Fig. 3), corresponding with the 30 nt minimum length previously identified for an in vitro-truncated naturally-occurring mRNA 3’ end segment (7). Oligo(U) segments ≥34 nt in length could apparently bind VP55 with full stability. For the heterodimer, oligo(U) segments ≥38 nt in length could bind detectably, while those ~45 nt or greater in length could apparently bind with full stability (Fig. 3).

**Contributions of the upstream (rU)\_2 and downstream rU of the heterodimer-binding motif to affinity and specificity**

The VP39–VP55 heterodimer [the active species in poly(A) tail elongation] interacts with an (rU)\_2-(N)\_12-rU motif, in which the three uridylates mediate protein recognition (12). To determine whether the heterodimer’s hierarchical base-type specificity in polyadenylylation (Fig. 2) might be mediated via the uridylates, competition EMSAs were performed in which interaction of the heterodimer with the standard motif-containing 50mer oligonucleotide [referred to as ‘UU(dC)\_5U’; sequence given in (12)] was assayed in the presence of various amounts of unlabeled heterogeneous-length homoribopolymeric competitor (Fig. 4A). EMSA complexes were far more susceptible to competition by unlabeled poly(U) than poly(C), consistent with the original identification of the motif as regions of U recognition in preference to C (5–7, 12). Unlabeled poly(A) competed to an intermediate extent (Fig. 4A). Since poly(U) was a significantly stronger competitor than “theoretical self”
(i.e. the profile that would be expected upon competition with unlabeled probe oligonucleotide), the standard probe apparently lacks heterodimer-binding determinants possessed by poly(U) (i.e. additional uridines and/or ribose sugars).

To determine whether the base-type preference shown in Figure 4A arises from the downstream rU of the heterodimer-binding motif, the EMSA competition experiment was repeated using a variant of the standard probe in which the downstream rU had been substituted with dC (Fig. 4B). Although the relative affinities of the poly(U), poly(A) and poly(C) competitors were unchanged with respect to the standard probe, the homopolymers all competed significantly more strongly (comparing Fig. 4B with A), indicating that the substituted probe had a reduced affinity for the heterodimer. Thus, the downstream rU of the

![Figure 1](image-url)
motif contributes both affinity and base-type specificity (for U over C) to the heterodimer–motif interaction. To address the role of the upstream (rU)₂ portion of the motif, this was substituted with (dC)₂, and the resulting oligonucleotide tested as a probe in an experiment equivalent to that of Figure 4B (Fig. 4C). Ready competition by the homopolymers showed that, like the downstream rU, the upstream (rU)₂ contributes both affinity and base-type specificity to the heterodimer–motif interaction. The stronger competition by the three homopolymers in Figure 4C than B indicated that the upstream (rU)₂ contributes significantly more affinity than the downstream rU. In a fourth experiment, a probe was used in which the downstream rU of the motif was substituted with the uridine analog 4-thiouridine (4S rU, Fig. 4D). Since Figure 4D resembled B more closely than A, results indicated a loss of recognition of the downstream uridine upon 4S rU substitution, suggesting that 4S rU is not significantly more advantageous than dC at the downstream position.

Base-type specificity at the heterodimer-binding motif was next examined using highly specific competitors as opposed to homopolymers (Fig. 5). The competitors used in Figure 5A were identical to the (standard) probe except that the downstream rU was substituted with either rA, rC, rG or 4S rU. The reproducibly unequal activities of the four competitors in the EMSA (Fig. 5A) indicated a clear base-type specificity at the downstream rU of the motif. Competitive efficacy at the downstream rU followed the following hierarchy (from strong to weak): rU (self) > rA > rG > 4S rU > rC. This is comparable to the hierarchies observed in the previous competition assays (Figs 1, 2 and 4). The experiment of Figure 5A was repeated using competitor oligonucleotides in which the upstream (rU)₂ of the motif was substituted with either (rA)₂, (rC)₂ or (rG)₂ (Fig. 5B). The resulting oligonucleotides were all much weaker competitors than those with substitutions in the downstream rU (comparing Fig. 5B with A) indicating that, as in Figure 4, the upstream (rU)₂ of the motif contributes significantly more affinity than the downstream rU. The hierarchy of competitive efficacy at the upstream (rU)₂ (from strong to weak) was rU (self) > rG > rA/rC.

Since an oligonucleotide possessing a 4S rU substitution at the downstream rU of the heterodimer-binding motif can be
Whether the R107 sidechain of VP39 might play a role in the heterodimer’s specificity for the downstream rU, in addition to being proximal to it, an EMSA competition experiment identical to that in Figure 5A was performed, except that wild-type VP39 was substituted with a non-photocrosslinkable VP39 mutant (R107A) lacking the R107 sidechain (12; Fig. 5C). The mutant heterodimer was much less effective than the wild-type one in discriminating among the four competitors (comparing Fig. 5C with A). A similar loss of specificity was observed in experiments using the R107K mutant of VP39 (12; data not shown). This would indicate that the R107 sidechain of VP39 does indeed contribute to the heterodimer’s base-type specificity at the downstream rU. Finally, the R107A–VP55 heterodimer was combined with the upstream (rU)₂-specific competitors (Fig. 5D). Upon comparing mutant and wild-type heterodimers (Fig. 5D versus B), the most noticeable effects were a decreased efficacy of all three competitors with the mutant and an increased discrimination among them. This is entirely consistent with the role of R107 in specific recognition of the downstream rU since, with loss of specificity at the downstream rU (by substitution of R107), a greater proportion of the overall protein–oligonucleotide affinity is mediated via the upstream (rU)₂. Therefore, affinity differences due to substitutions in the latter portion of the motif are magnified.

Which heterodimer subunit interacts with each of the three uridines of the heterodimer-binding motif?

We next wondered which of the two heterodimer subunits interacts with each of the three uridines of the heterodimer-binding motif. Photocrosslinking approaches with ⁴S rU-substituted oligonucleotides have proven uninformative, since oligonucleotides with ⁴S rU substitutions at any of the motif’s three rUs cannot be photocrosslinked to the VP55 subunit, and VP39 photocrosslinking is only observed for oligonucleotides substituted at the downstream rU (12; data not shown). We therefore took an alternative approach, based upon the distinctive protein-binding characteristics of oligonucleotides possessing ribothymidine substitutions. Thus, the interaction of monomeric VP55 with its cognate (rU)₂-(N)₁₅-U motif is hyper-stabilized upon substitution of the extreme upstream and downstream rU positions of the motif with ribothymidine, and destabilized upon ribothymidine substitution at the central rU (7). This effect can also be seen in the control experiment of Figure 6A (upper panel) in which three oligonucleotides, each possessing a ribothymidine substitution at one of the three uridylates of the motif were tested as EMSA competitors for VP55 binding by the parental [(rU)₂-(N)₁₅-U motif-containing] oligonucleotide. As an indication of whether the three uridylates of the (rU)₂-(N)₁₅-U heterodimer-binding motif might interact with the VP55 subunit of the heterodimer, the three uridylates of this motif were individually substituted with ribothymidine, and the resulting oligonucleotides tested as EMSA competitors for heterodimer binding by the parental [(rU)₂-(N)₁₅-U motif-containing] oligonucleotide (Fig. 6A, lower panel). As in the control experiment, substitution at either the upstream or the downstream rU led to an apparent hyper-stabilization of binding, and substitution at the central rU was destabilizing (7). The similarity between the profiles shown in the two panels of Figure 6A would indicate that, within the heterodimer–oligonucleotide ternary complex, all three uridylates of the heterodimer-binding motif are recognized by the VP55 subunit.
Since the heterodimer’s VP55 subunit appears to interact more strongly with thymine than uracil at the two outside positions of the motif (Fig. 6A), we wondered whether the photosensitive thymidine analog 4-thiodeoxythymidine (4S dT) might photocrosslink to VP55 more strongly than the corresponding uridine analog, 4S rU (12). Oligonucleotides individually substituted with 4S dT at the uridylates of the heterodimer-binding motif were therefore end-labeled, incubated with the heterodimer and UV-irradiated. The resulting mixtures were purified from an EMSA gel and subjected to SDS–PAGE (Fig. 6B). At both the EMSA and SDS gel steps (data not shown and Fig. 6B, respectively), substitution at the two outside positions (−37 and −10) led to more abundant complexes than substitution at the central position (−36) consistent with the effect of ribothymidine substitution upon binding stability (Fig. 6A). Moreover, each of the 4S dT-substituted oligonucleotides could be photocrosslinked to the VP55 subunit, providing further evidence that all three uridylates of the motif interact with this subunit within the heterodimer–oligonucleotide ternary complex. The VP39 subunit could be photocrosslinked only to the oligonucleotide substituted at the downstream position (−10, Fig. 6B), consistent with VP39’s very strong photocrosslinking of the equivalent 4S rU-substituted oligonucleotide (7). Overall, the data of Figure 6 indicate that, whereas the downstream uridylate of the heterodimer-binding motif interacts with both
heterodimer subunits, the upstream (rU)₂ portion may interact only with VP55.

Which heterodimer subunit is proximal to each base of the primer within the ternary complex?

Having characterized subunit interactions with the three uridylates of the heterodimer-binding motif, we next wondered which of the two subunits might be juxtaposed to each base throughout the oligonucleotide primer. We aimed to do this in a single experiment using a variation of the 'crosslinking SELEX' (16) approach with 4S rU as the photocrosslinking 'tag'. Although 4S rU is not recognized as rU (above), it was nonetheless considered to be potentially photocrosslinkable. A (dC)₄₀-based oligonucleotide pool was synthesized, spiked randomly throughout with 4S rU. We aimed for an ~1.25% spiking level (i.e. approximately one 4S rU per two 40mer oligonucleotides) to ensure that, for the vast majority of molecules possessing a pH-hydrolyzable ribose sugar, its position always corresponded with that of the photocrosslinkable base. Before synthesizing the pool, dC/4S rU mixtures were tested to identify a monomer-doping level (during synthesis) commensurate with a spiking level (in the product) of ~1.25%, using a previously devised method (17). Whereas doping and spiking levels correlated in a linear fashion for mixed dC/rU positions (Fig. 7A) such as those present in the oligonucleotide pools used previously (5), the relationship for mixed dC/4S rU positions was distinctly non-linear (Fig. 7A), requiring an ~9% doping level to achieve ~1.25% spiking.

The (dC/4S rU)₄₀ pool was 5'-end-labeled and mixed with the heterodimer. The resulting mixture was irradiated at 350 nm, electrophoresed in an EMSA gel (Fig. 7B), and material corresponding to the ternary complex [labeled 'Dimer-'] recovered. SDS–PAGE of this material (Fig. 7C) indicated that ~10-fold more of the pool oligonucleotide had become photocrosslinked.
to the VP39 subunit than to the VP55 subunit. The oligonucleotide–VP39 and oligonucleotide–VP55 covalent conjugates were individually isolated from the SDS gel and thoroughly digested with pronase, leaving a selected population of oligonucleotide molecules whose relative abundances within the population were taken to reflect the efficiency of photocrosslinking at each position within the oligonucleotide. After repeating the experiment using a 3'-end-labeled pool, the selected populations from both experiments were hydrolyzed with 33% TFA and subjected to urea–PAGE (Fig. 7D). For the 5'-end-labeled pool, hydrolysis ladders for the free (uncross-linked) and VP55-selected material showed no discernible nodes. However, two nodes were observed in the hydrolysis ladder from the VP39-selected population (as plotted in Fig. 7E); a strong and a weak one corresponding to hydrolysis ~25 to ~32 nt and ~4 to ~11 nt, respectively, from the 3'OH. For the 3'-end-labeled pool, two distinct nodes were observed in the hydrolysis ladders prior to crosslinking SELEX (Fig. 7D), presumably because the VP55–VP39 heterodimer had been used for the initial oligonucleotide 3'-end-labeling step, leading to the selective labeling of good substrates. Although a similar nodal pattern was also observed in the VP39-crosslinked population after the crosslinking SELEX procedure, the nodes had become further emphasized as indicated in the right-most 'difference' plot of Figure 7E. As with the 5'-end-labeled pool, nodes were clearly apparent ~25 to ~32 nt and ~4 to ~11 nt, from the 3'OH (Fig. 7E). Overall, the experiment shown in Figure 7 indicated that 4S rU is within photocrosslinkable distance of VP39 within the heterodimer–oligonucleotide ternary complex at two positions within the oligonucleotide.

The photocrosslinking experiments of Figure 7B–E employed a 40mer pool because this was close to the minimum oligonucleotide length able to interact detectably with the heterodimer (Fig. 3). Use of a minimum-length oligonucleotide was expected to minimize the possibility of multiple isoforms of the protein–RNA complex resulting from the absence of a motif to 'anchor' the heterodimer at a single position. Since oligonucleotides ≥45 nt in length bind the heterodimer significantly more stably than 40mers (Fig. 3), we repeated the crosslinking SELEX experiment using a longer (50mer) dC/4S rU pool (Fig. 7F). In comparison with the 40mer pool, the position of the downstream node was unchanged. However, the location of the upstream node was 10 nt more distant from the oligonucleotide 3' end.

4S rU photocrosslinking position on the VP39 surface

To confirm the conclusions of the 'crosslinking SELEX' approach (above), we synthesized and tested a set of 20 individually 4S rU-substituted (dC)₃₋₅C-based oligonucleotides, each differing in the position of the single 4S rU substitution. These were 5'-end-labeled, mixed with the heterodimer and irradiated. Ternary complexes were isolated from EMSA gels and the resulting material resolved by SDS–PAGE. As in the 'crosslinking SELEX' approach (Fig. 7), photocrosslinking to the VP39 subunit was much more efficient than to VP55. After repeating the experiment, the mean proportion of each EMSA complex-derived oligonucleotide that had become photocrosslinked to VP39 was quantitated and plotted against 4S rU position (Fig. 8A). Two nodes were observed, whose positions (~2 to ~38 region and ~10 to ~12 region) were comparable to those observed by crosslinking SELEX (Fig. 7E).

As in the 5'-end-labeled crosslinking SELEX experiment (Fig. 7E), the node located closer to the oligonucleotide 5' end was the stronger one (in which up to ~50% of the oligonucleotide from the excised ternary complex had become photocrosslinked to VP39).

The downstream node observed in both the crosslinking SELEX (Fig. 7) and the ordered oligo series (Fig. 8A) experiments, corresponds to the position of 4S rU substitution (~10) that was previously shown to photocrosslink to the sidechain of VP39 residue R107 (12). To identify the region of VP39 photocrosslinking to the upstream node, a variant of the standard motif-containing 50mer was synthesized in which the upstream (rU)₂₅ portion was replaced with a single 4S rU. The position of the single 4S rU residue (~36) was therefore equivalent to the upstream node in the crosslinking SELEX experiment with the 50mer pool (Fig. 7F). VP39-NAC26 (a variant possessing an N-terminal protein kinase A recognition site) was N-terminally 3P-labeled, mixed with the oligonucleotide (which is denoted 4S (dC)₃₋₅C); see Materials and Methods for sequence), photoradiated and the conjugate recovered from an SDS gel and subjected to partial proteolysis and electrophoresis as described previously (12). A control experiment employed an oligonucleotide in which the single 4S rU substitution replaced the downstream rU of the motif [denoted 'UU(dC)₃₋₅C' (12). Identi-cal peptide mapping results were obtained with both oligonucleotides (Fig. 8B), indicating the interaction of similar or identical regions of VP39 with both upstream and downstream portions of the motif. To determine whether the upstream position had become photocrosslinked to VP39 residue R107 (VP39’s photocrosslinking site for oligonucleotides substituted at the downstream position; 12), heterodimers with alanine substitutions at residues N104, L106, R107 and D108 of the VP39 subunit were each mixed with four oligonucleotides containing single 4S rU substitutions at positions equivalent to either the upstream or the downstream node, and the mixtures photoradiated. Ternary complexes were isolated from EMSA gels and the resulting material resolved by SDS–PAGE (Fig. 8C). Whereas mutants N104A, L106A and D108A retained the photocrosslinking efficacy of wild-type VP39, R107 mutations eliminated photocrosslinking to each of the four oligonucleotides (Fig. 8C), indicating a role for R107 in photocrosslinking to both the upstream and the downstream crosslinking SELEX nodes. This is discussed further below (see Discussion).

Phosphate-anchored longer-range photocrosslinkers

The majority of photocrosslinking experiments performed thus far employed 4S rU as the photoactive species. However, disadvantages of 4S rU include its photocrosslinking over only a short range and its negative influence upon uridine recognition by the heterodimer. To complement these deficiencies, we enlisted azidophenacyl bromide (APB)—a relatively long-range photocrosslinker whose attachment to nucleic acid phospho- moieties leaves the bases unmodified (15,18,19). Briefly, a family of 10 heterodimer-binding motif-containing 50mers was synthesized, differing only in the position of substitution with a single phosphorothioate. These were coupled with APB, mixed with the VP55–VP39 heterodimer and irradiated. Heterodimer–oligonucleotide ternary complexes were isolated from an EMSA gel (an equivalent EMSA complex abundance being observed for each oligonucleotide, data not shown) and
Figure 7. Identifying positions of base-protein proximity using a ‘crosslinking SELEX’ approach. (A) Relationship between proportional doping of dC phosphoramidite bottle with either rU or 4S rU phosphoramidite for solid-phase synthesis of 9mer test oligonucleotides (abscissa), and ribonucleotide incorporation (ordinate). Various dC/rU and dC/4S rU mixtures were tested, and ribonucleotide incorporation was monitored by subjecting the resulting oligonucleotides to acid hydrolysis at the internal ribose linkage [as outlined in (17)]. (B) EMSA gel showing complexes obtained after mixing VP55 and/or VP39 with a synthetic 40mer dC/4S rU pool oligonucleotide and photolabelling (see text for experimental details). Symbols are as described in Figure 8B legend. (C) SDS–PAGE of material recovered from the major complex present in each of the three right-most lanes of the EMSA gel shown in (B). Free oligonucleotide from the left-most EMSA gel lane was also recovered and re-electrophoresed. Samples were loaded on the EMSA and SDS gels in corresponding order. Insignificant amounts of photocrosslinked conjugate were obtained after mixing the heterodimer with a control oligonucleotide lacking a photosensitive 4S rU residue [(dC) 39 rC] and irradiating equivalently (data not shown). (D) Acid hydrolysis of free oligonucleotide and pronase-treated conjugates recovered from slices of the SDS–PAGE gel shown in (C). 5', 3', experiment initiated with 5'- or 3'-end labeled pool oligonucleotide, respectively. Distance (nt) from the oligonucleotide 3' end is indicated down the sides of the gel. (E) Line graph quantitation of gel lanes (indicated) in (D). (F) Hydrolysis products from two equivalent crosslinking SELEX experiments employing 3'-end-labeled pool oligonucleotides of identical composition but different length. Lanes 40 and 50, (dC/4S rU) 40 pool [as used in (D)] and (dC/4S rU) 50 pool, respectively. As with the 40mer (E), nodes present after the initial labeling step were stronger after selection also. For additional details, see legends to other parts, this figure.
the resulting material resolved by SDS–PAGE (Fig. 9A). The proportion of oligonucleotide that had become photocrosslinked to either subunit was quantitated (Fig. 9B). Although each oligonucleotide, this value (ordinate) is plotted against the distance of the single 4S rU substitution from the oligonucleotide 3'OH (abscissa). Although EMSA complexes were of low abundance due to the absence of a heterodimer-binding motif and the slightly suboptimal oligonucleotide length (data not shown), they were equal in abundance from one oligonucleotide to another. (B) Protease mapping. This was done as described previously (12). Briefly, 4S rU-substituted oligonucleotides were incubated with VP35 and an N-terminally 32P-labeled, terminally-modified VP39 derivative (VP39-NAC26), then irradiated. The resulting VP39–oligonucleotide conjugates were isolated by SDS–PAGE, treated with various amounts of trypsin and subjected to high-resolution (peptide) gel electrophoresis (shown). VP39-S(dC)25 U, UU(dC)25 S–VP39 and VP39 (above lanes), analyses of VP39-NAC26 conjugates with oligonucleotides S(dC)25 U and UU(dC)25 S, and of free VP39–NAC26, respectively. For each conjugate, ‘0’ denotes an aliquot with no added trypsin, filled wedge denotes increasing amounts of trypsin (0.1, 0.5 and 1 µg, or 0.1 and 0.5 µg per reaction for conjugates and unconjugated protein, respectively). The extreme right-hand lane (VP39/CNBr) shows a cyanogen bromide (CNBr) partial digest of VP39-NAC26. The methionine cleavage sites corresponding to individual labeled bands in the digest are indicated to the right of the gel. Bands corresponding to free and conjugated VP39-NAC26 are denoted ‘VP39’ and ‘VP39-’, respectively. As previously (12), the largest proteolytic fragment shared by the unconjugated and conjugated protein samples migrated slightly ahead of the CNBr-cleavage fragment corresponding to residue M92, indicating that the photocrosslinking site for both oligonucleotides is close to the C-terminal side of residue K90. In combination with data obtained using C-terminally-labeled protein (data not shown), this indicated the photocrosslinking site to lie between residues K90 and R114. (C) Four oligonucleotides were each photocrosslinked to heterodimers containing wild-type VP39 and muteins thereof bearing single amino acid substitutions. Ternary complexes were recovered from an EMSA gel, and their contents subjected to SDS–PAGE. Oligonucleotides and muteins are denoted as in the text and previously (12), other symbols are as given in Figure 6B legend. As previously, the VP39 subunit photocrosslinked more readily to oligonucleotide UU(dC)25 S than to oligonucleotides S(dC)25 U or S(dC)25 U (12), the image shown for the latter two oligonucleotides being the product of a much longer exposure. S(dC)25 U photocrosslinked with intermediate efficiency.

**DISCUSSION**

In this study, we have identified the overall length of oligonucleotide that interacts with the heterodimer, the heterodimer subunit(s) interacting with each of the three ribouridylates of the oligonucleotides (Fig. 9A), which increased in abundance closer to the oligonucleotide 3' end. Additional experiments indicated that these might represent the crosslinking of two oligonucleotide molecules per VP55 molecule (data not shown).
the (rU)₂–(N)₂–rU heterodimer-binding motif, and have addressed specificity at each of these uridylates. We have also identified non-base-specific interactions between the heterodimer and all parts of the priming oligonucleotide and addressed the overall topology of the VP55–VP39-primer ternary complex. We conclude (Fig. 10) that: (i) base-type-specific RNA-binding properties of the heterodimer are mediated via protein–oligonucleotide contacts characterized by APB photocrosslinking. These indicate the location of a cleft or enclosed channel in the heterodimer surface formed at the interface of the two subunits, through which we suggest that the 3' end portion of the ~45 nt primer may pass. Evidence for this comes from (i) the coordinated photocrosslinking of both subunits to primer molecules derivatized with APB 11–21 nt from the oligonucleotide 3'OH and (ii) the apparent interaction of both subunits with the downstream rU of the motif.

![Figure 10. Summary of proposed interactions within the VP55–VP39-primer ternary complex](image)

The internal loop in the primer represents the ~10 nt greater spacing between the two uridylate-containing portions of the motif for the VP55–VP39 heterodimer than for monomeric VP55. Whereas both subunits are shown contacting the downstream rU of the primer’s heterodimer-binding motif, only the VP55 subunit apparently contacts the upstream (rU)₂. The portion of VP39 facing upstream is drawn with a broken line because although we have found no evidence to suggest that VP39 interacts directly with the 5' region of the primer, we cannot formally discount it. The two symmetrical sets of four arrows located between the primer ‘loop’ and the downstream rU denote the uridylate-independent protein–oligonucleotide contacts characterized by APB photocrosslinking. These indicate the location of a cleft or enclosed channel in the heterodimer surface formed at the interface of the two subunits, through which we suggest that the 3' end portion of the ~45 nt primer may pass. Evidence for this comes from (i) the coordinated photocrosslinking of both subunits to primer molecules derivatized with APB 11–21 nt from the oligonucleotide 3'OH and (ii) the apparent interaction of both subunits with the downstream rU of the motif.

The (rU)₂–(N)₂–rU heterodimer-binding motif, and have addressed specificity at each of these uridylates. We have also identified non-base-specific interactions between the heterodimer and all parts of the priming oligonucleotide and addressed the overall topology of the VP55–VP39-primer ternary complex. We conclude (Fig. 10) that: (i) base-type-specific RNA-binding properties of the heterodimer are mediated via all three uridylates of the heterodimer-binding motif; (ii) each of the three ribouridylates of the motif interacts with the VP55 subunit while the downstream rU contacts VP39 also; (iii) the upstream ~half of the complete ~45 nt priming site [including the upstream (rU)₂ of the heterodimer-binding motif] interacts with the VP55 subunit only, whereas the downstream ~half sits within a ‘channel’ (i.e. a ‘tubular enclosed passage’) (20) or ‘cleft’ (i.e. ‘indented formation’) (20) at the interface of both subunits, proximal to which (or within which) the downstream rU of the motif is situated prior to the initiation of polyadenylation. We suggest that, within this topological arrangement, VP55 (the adenylyltransferase catalytic subunit) may be responsible for actively translocating the PAP heterodimer with respect to the RNA (consistent with the ability of the isolated VP55 subunit to translocate; 6), and that the putative channel/cleft serves to restrain the polymerase from dissociating from the downstream portion of the RNA during translocation. The channel/cleft might act by topologically enclosing the RNA or by allowing both subunits to contribute binding determinants to the same RNA region.

Our primary aim, namely to determine which subunit contacts each of the three uridylates within the (rU)₂–(N)₂–rU heterodimer-binding motif, was initially addressed taking photocrosslinking approaches. Proximity of the VP39 subunit to the downstream rU of the motif had already been demonstrated from its ready photocrosslinking to an oligonucleotide in which the downstream rU was substituted with δrU (12). Although it seemed likely that the VP55 subunit was primarily responsible for interaction with the upstream (rU)₂ portion, since two adjacent rUs are also a characteristic of the (rU)₂–(N)₁–rU motif that interacts with monomeric VP55 (7), this was difficult to prove by δrU photocrosslinking because oligonucleotides with δrU substitutions at the upstream (rU)₂ could not be photocrosslinked to either subunit in significant yield (12). Attempts to use alternative photosensitive uridine analogs (5-bromo- and 5-iodouridine), even in combination with tuned lasers (L. Johnson, A. Oraevsky, P. Gershon, unpublished data) were fruitless. Two possible explanations for the absence of photocrosslinking of either subunit to the upstream (rU)₂ were
proposed previously (12), namely, the absence of a proximal photocrosslinkable protein moiety within the ternary complex and/or elimination of uridine recognition upon analog substitution. A combination of the two is probably true. Thus, in none of our experiments could we demonstrate proximity between VP39 subunit moieties and the 5' portion of the primer containing the upstream (rU)2 (Fig. 10). These experiments included photocrosslinking assays employing two highly effective photocrosslinkers, namely 4SdT (Fig. 6B) and APB (Fig. 9). By contrast, strong zero-length photocrosslinking of 4SdT to the VP55 subunit (Fig. 6B) indicated the proximity of this subunit to the upstream (rU)2. The absence of significant VP55 crosslinking to oligonucleotides containing 4SdT substitutions in the upstream region (12) therefore presumably results from disrupted uridine recognition. Indeed, the sensitivity of monomeric VP55 to substitutions in the upstream (rU)2 with uracil analogs (including the photosensitive analog 5-iodouridine), has been previously documented (7). Although current experiments show a loss of recognition at the downstream rU upon 4SdT substitution (Figs 4D and 5A), sufficient proximity is apparently retained in this case for efficient photocrosslinking to the VP39 subunit. What structural features of 4SdT render it a more proficient photocrosslinker than 4SdT? Since ribothymidine substitution at the upstream and downstream uridines of the VP55- and heterodimer-binding motifs stabilizes the binding of VP55 and the heterodimer, respectively (7 and Fig. 6A), the positive influence of the 5-CH2 group apparently supercedes any negative effects arising from the 4-thio moiety. Positive influences of the 5-CH2 might include some intrinsic property such as its hydrophobicity, or a secondary effect—perhaps upon the pKa of position N3 in the pyrimidine ring.

A secondary aim of this work was to determine which of the two heterodimer subunits contacts every part of a 50 nt primer. This was of particular interest because the spacing between the upstream (rU)2 and downstream rU portions of the (rU)2(N)2-rU heterodimer-binding motif is 10 nt greater than that for the corresponding (rU)2(N)2-rU VP55-binding motif. Thus, we wondered (for example) whether the additional 10 nt might participate in an excursion of the primer away from the VP55 subunit, perhaps towards VP39. The 'crosslinking SELEX' procedure (16) represented an attempt to address this by examining which bases throughout the heterodimer-bound oligonucleotide are proximal to which of the two subunits. Although the approach was hampered by very low levels of VP55 photocrosslinking at any position within the oligonucleotide pool (as also found in crosslinking experiments with the discrete 4SdU-substituted oligonucleotides of Fig. 8A, data not shown), data were obtained for the VP39 subunit. Elements in common with the published crosslinking SELEX procedure (16) included our use of an oligonucleotide pool contaminated with a photocrosslinkable nucleotide, protein photocrosslinking and isolation of the photocrosslinked oligonucleotide–protein conjugate by PAGE, and recovery of the oligonucleotide therefrom by protolysis. In contrast to the published procedure, we did not attempt to amplify or reselect the selected ligands (16). Partial hydrolysis of crosslinking SELEX pools that had been 3'-end-labeled using the VP55–VP39 heterodimer as the labeling enzyme yielded two nodes in the hydrolysis ladder prior to any crosslinking-selection (Fig. 7D). Apparently, pool molecules possessing a 4SdT at either of the two nodal positions had become preferentially labeled. Due to the poor or non-existent recognition of the 4-thiouracil base by the heterodimer (Figs 4D and 5A), it was presumably the ribose sugar that had been selected for during labeling, due to the deoxyribose background of the pool. The uneven 3'-end labeling of the pool molecules did not hamper photocrosslinking selection experiments, since this effect could be subtracted during the quantitation of gel lanes (Fig. 7E).

Crosslinking SELEX showed two apparent nodes of proximity between the VP39 subunit and the oligonucleotide ligand pool (Fig. 7D and E). These two nodes were also observed with a series of oligonucleotides representing discrete molecules from the crosslinking SELEX pool (Fig. 8A). It was initially assumed that the two nodes indicated the interaction of VP39 with both upstream and downstream portions of the primer. However, unexpectedly, two of the discrete oligonucleotides used in Figure 8A, whose photocrosslinker positions were equivalent to those of the two nodes (oligos 12 and 28) photocrosslinked to the same VP39 sidechain (that of R107, data not shown). Increasing the length of the oligonucleotide pool from 40 to 50 nt led to an increased distance between the upstream node and the oligonucleotide 3' end by 10 nt (Fig. 7F). Again, discrete oligonucleotides with 4SdT substitutions at the positions of the two nodes photocrosslinked only to the R107 sidechain of VP39 (Fig. 8B and C). Photocrosslinking of the upstream and downstream node-regions to the same VP39 sidechain contradicted the simplest interpretation of the crosslinking SELEX data, namely that, within the ternary complex, the VP39 subunit contacts the upstream and downstream regions of the primer via distinct RNA binding sites. To address further the possibility that individual VP39 molecules can interact simultaneously with two distinct sites in the primer, we tried mixing the heterodimer with an end-labeled 50mer oligonucleotide in which a single centrally-located rG residue was flanked by two photocrosslinkable 4SdU residues (at positions corresponding to the two crosslinking SELEX nodes), followed by photoradiation, exhaustive digestion with RNase T1 and SDS–PAGE. Although the gel clearly showed a labeled band corresponding to a covalent complex between VP39 and half-size oligonucleotide fragments, no band was observed whose mobility corresponded to VP39 plus the full-size oligonucleotide (i.e. two half-size fragments). This provided evidence that, within the vast majority of ternary complexes, both halves of the oligo cannot simultaneously interact with VP39, consistent with the absence of two protein–RNA photocrosslinking sites corresponding to the two nodes. These data might also argue against a model in which the oligonucleotide loops around VP39 within the heterodimer–oligonucleotide ternary complex such that two positions within the oligonucleotide can simultaneously contact the R107 sidechain. The latter model is disfavored anyway, since it would be difficult to reconcile with other topological data.

A second explanation for the photocrosslinking of both upstream and downstream oligonucleotide regions to the same VP39 sidechain might be that VP55 and VP39 exist within EMSA 'ternary' complexes as a 1:2 heterotrimer as opposed to a 1:1 heterodimer, with the two VP39 molecules positioned over the oligonucleotide's two photocrosslinking nodes. However, 1:1 subunit stoichiometry was demonstrated by an experiment in which end-labeled VP55 and VP39 of known specific activity were mixed with unlabeled oligonucleotide, and the resulting complex recovered from an EMSA gel and
subJECTED TO SDS–PAGE AND QUANTITATION. WE THEREFORE FAVOR A THIRD EXPLANATION, NAMELY THAT A SINGLE BAND IN THE EMSA GEL CONTAINS A HETEROGENEOUS MIXTURE OF 1:1:1 VP55–VP39–OLIGONUCLEOTIDE TERNARY COMPLEXES, PERHAPS DUE TO THE ‘SLIDING’ OF OLIGONUCLEOTIDE MOLECULES BETWEEN MULTIPLE (TWO) STATIC POSITIONS WITH RESPECT TO THE HETERODIMER. POOL OLIGO MOLECULES MIGHT BE PARTICULARLY PRONE TO SLIDING DUE TO THEIR LACK OF A HETERO-DIMER-BOUNDING MOTIF FOR ‘ANCHORING’, AND THE 50MERS OF FIG. 8B AND C MIGHT ALSO SLIDE BECAUSE OF THE DISSOLVED MOTIF RESULTING FROM $^{32}$P SUBSTITUTION. PERHAPS ONLY IN THOSE COMPLEXES POISED FOR PHOTOCROSS-LINKING AT THE DOWNSTREAM NODE IS THE OLIGONUCLEOTIDE OPTIMALLY POSITIONED FOR POLY-ADENYLATION OF ITS 3' END, CONSISTENT WITH THE KNOWN INTERACTION OF VP39 WITH THE DOWNSTREAM-NODE REGION IN THE CONTEXT OF MOTIF-CONTAINING OLIGONUCLEOTIDES (12; FIGS 6B, 9 AND 10). THE DATA SHOWN IN FIGURE 7F WOULD BE CONSISTENT WITH THE SECOND (i.e. UPSTREAM) NODE REPRESENTING OLIGONUCLEOTIDE ‘SLIDING’ TO WITHIN A FIXED DISTANCE OF ITS 5' END.

A TERTIARY AIM OF THIS STUDY WAS TO FURTHER INVESTIGATE THE BASE-TYPE-SPECIFIC PROPERTIES OF VP39 AND THE HETERO DIMER. CONSISTENT WITH AN EARLIER FINDING THAT NONOMERIC VP39 IS A BASE-TYPE-SPECIFIC RNA BINDING PROTEIN THAT INTERACTS WITH POLY(G) AND POLY(A) IN PREFERENCE TO POLY(C) (1), THE BINDING HIERARCHY OF VP39 USING AN ENLARGED SET OF HOMORIBOPOLYMERS WAS (STRONG TO WEAK): POLY(I) > POLY(U) >> POLY(G)/POLY(A) > POLY(C) (FIG. 1). POLY(DA) DID NOT COMPETE AT ALL FOR POLY(A) BINDING (DATA NOT SHOWN), INDICATING SOME BACKBONE SPECIFICITY ASSOCIATED WITH THE BASE-TYPE-SPECIFIC RNA BINDING SITE.

IT IS UNCLEAR WHY VP39, BEING A PROCESSIVITY FACTOR FOR POLY(A) TAIL ELONGATION, WOULD EXHIBIT GREATER AFFINITY FOR URIDINE AND INOSINE THAN ADENINE. ALTHOUGH WE DID NOT CHARACTERIZE FUNCTIONALITIES OF THE BASES RESPONSIBLE FOR BINDING SELECTIVITY, STRUCTURAL CHARACTERISTICS SHARED BY THE STRONGER BINDERS (inosine, uracil and guanine) INCLUDE A PROTONATED N3 (PYRIMIDINE)/N1 (PURINE)—A FEATURE WHICH APPEARS TO BE IMPORTANT FOR URIDINE RECOGNITION BY NONOMERIC VP55 (7)—AND/OR A 6-KETO (PURINE)/4-KETO (PYRIMIDINE) FUNCTION.

THE BASE-TYPE-SPECIFIC PROPERTIES OF THE HETERO DIMER SEEM TO BE MEDIATED VIA BOTH SUBUNITs. THUS, THE UPSTREAM (RU)2 PORTION OF THE HETERO-DIMER-BINDING MOTIF, WHERE BASE-TYPE SPECIFICITY IS OBSERVED (FIG. 5B), APPARENTLY CONTACTS THE VP55 SUBUNIT ONLY (FIG. 10). CONSISTENT WITH THIS, THE ISOLATED VP55 SUBUNIT EXHIBITS A HIERARCHICAL BASE-TYPE PREFERENCE (6). BOTH SUBUNITs INTERACT WITH THE DOWNSTREAM RU OF THE MOTIF (FIG. 10) WHERE BASE-TYPE SPECIFICITY IS ALSO OBSERVED. THE R107 SIDECHAIN OF VP39 NOT ONLY DIRECTLY CONTACTS THE DOWNSTREAM RU (12), BUT ALSO CONTRIBUTES TO URIDINE SPECIFICITY (FIG. 5C). WHETHER OR NOT R107 IS VP39'S SOLE BASE-TYPE SPECIFICITY DETERMINANT IS UNKNOWN. THE LOCATION OF R107 ON THE PROTEIN SURFACE IS REMOTE FROM VP39'S MOST WELL CHARACTERIZED RNA BINDING SITE, NAMELY THE CLEFT RESPONSIBLE FOR BINDING THE 5' END OF THE METHYLTRANSFERASE SUBSTRATE RNA OF VP39 (12). CAPPED RNA INTERACTION AT THE LATTER SITE HAS BEEN CLEARLY SHOWN, BY CO-CRYSTALLOGRAPHIC ANALYSIS, TO BE SEQUENCE-NON-SPECIFIC IN CHARACTER (10). CONSISTENT WITH THIS, THE FIVE HOMOPOLYMERS EMPLOYED HERE WERE EQUIVALENT IN THEIR ABILITY TO COMPETE WITH A SHORT METHYLTRANSFERASE SUBSTRATE [m1G(5')ppp(5')G(A)] FOR VP39 BINDING IN AN EMSA ASSAY DESCRIBED PREVIOUSLY (11) (DATA NOT SHOWN). HOWEVER, CAPPED HOMORIBOPOLYMERS HAVE PREVIOUSLY BEEN SHOWN TO BEUNEQUAL IN THEIR EFFICACIES AS METHYLTRANSFERASE SUBSTRATES FOR VP39; Whereas capped poly(A) and poly(I) COULD BE STRONGLY METHYLATED DURING THE REACTION, CAPPED poly(G), poly(U) AND poly(C) COULD NOT (21). OTHER DATA HAVE INDICATED THAT DOWNSTREAM REGIONS OF THE METHYLTRANSFERASE SUBSTRATE OF VP39 MAY INTERACT WITH VP39 AT A SITE THAT IS DISTINCT FROM THE METHYLTRANSFERASE 'CLEFT' (22). IT REMAINS TO BE SEEN WHETHER THIS "DOWNSTREAM" SITE INCLUDES THE SIDECHAIN OF R107 OR REPRESENTS YET ANOTHER REGION OF THE PROTEIN SURFACE.

ACKNOWLEDGEMENTS

WE THANK DR XUENONG SHI FOR KINDLY PROVIDING PURIFIED VP39-NAc26 PROTEIN, DR ALEC HODEL FOR PROVIDING HIGHLY PURIFIED VP39-AC26, DR LI DENG FOR MUTENES AND FOR ADVICE ON THE PARTIAL PROTEOLYSIS OF END-LABELED PROTEIN AND SHI-MEI WANG FOR PURIFIED BACULOVIRUS-EXPRESSED VP59. THIS WORK WAS SUPPORTED BY NIH GRANT R01 GM51953.

REFERENCES