Specific Recognition of an rU$_2$N$_{15}$-rU Motif by VP55, the Vaccinia Virus Poly(A) Polymerase Catalytic Subunit*

Li Deng‡, Leonid Beigelman§, Jasenka Matulic-Adamic§, Alexander Karpeisky§, and Paul D. Gershon¶

From the Institute of Biosciences and Technology/Department of Biochemistry and Biophysics, Texas A&M University, Houston, Texas 77030-3303 and Ribozyme Pharmaceuticals Inc., Boulder, Colorado 80301

VP55, the vaccinia poly(A) polymerase catalytic subunit, interacts with oligonucleotide primers via two uridylate recognition sites (Deng, L., and Gershon, P. D. (1997) EMBO J. 16, 1103–1113). Here, we show that the cognate RNA sequence comprises a 5’-rU$_2$N$_{15}$-rU-3’ motif (where N = any deoxyribo or ribonucleotide), embedded within oligonucleotide primers 29–30 nucleotides (nt), or greater, in length. Nine residues separate the 3’-most ribouridylate of the optimally positioned motif from the primer 3’-OH. A ribose sugar at the extreme 3’-terminal nucleotide of the primer is absolutely required for VP55’s adenylyltransferase activity, but not for stable VP55-RNA interaction. A ribose at position −3 markedly stimulates both adenylyltransferase activity and stable binding. The use of uridine analogs indicated (i) those functional groups of the uracil base which contribute to stable VP55-primer interaction, and (ii) that VP55’s ability to discriminate uracil from cytosine stems largely from the requirement for a protonated N3 nitrogen within the pyrimidine ring. The rU$_2$N$_{15}$-rU motif was identified within the uridine-rich 3’ end of a naturally occurring vaccinia mRNA. However, oligonucleotides whose only internal uridylates comprised the motif supported only a 3–5 nt processive burst of oligo(A) tail addition, as opposed to the ∼30–35 nt burst observed with the naturally occurring 3’ end.

One unusual feature of the poly(A) polymerase (PAP)$^1$ encoded by vaccinia virus is its heterodimeric structure (2–4). However, roles for the individual subunits within the heterodimer have been established by examination of the in vitro properties of the individual subunits. Thus, the isolated larger (VP55) subunit possesses PAP catalytic activity, and is able to add ∼30–35 nt oligo(A) tails to RNA 3’ ends in a rapid and processive burst of polyadenylation, before switching, abruptly, to a very slow and non-processive mode of adenylyl addition (2, 5). The isolated smaller (VP39) subunit has no PAP catalytic activity, but its addition to VP55 permits tails that are greater than ∼35 nt in length to be processively elongated to an overall length of several hundred nucleotides, in vitro (6). In addition to its activity as a PAP processivity factor, VP39 has an entirely unrelated function, at the mRNA 5’ end. Thus, as an mRNA cap-specific 2’-O-methyltransferase, VP39 modifies the ribose sugar of the penultimate nucleotide of the mRNA type 0 cap structure to a 2’-O-methylated nucleotide (7).

Since VP55 extends oligo(A) primers greater than 30–35 nt in length in only a slow, non-processive mode of adenylyl addition, the ability of this subunit to catalyze the processive polyadenylation burst is apparently triggered by some signal other than oligo(A) or poly(A) (5). This signal was shown to comprise, in some manner, a high content of uridylate residues within the 3’-terminal 30–40 nt of the initial RNA primer, in no obvious arrangement or pattern (8). Consistent with this, VP55 interacts stably with uridylate-rich RNA segments that are 33–34 nt or greater in length (1, 8), in the electrophoretic mobility shift assay (EMSA). Comparisons of various oligonucleotide sequences have shown the abundance of the EMSA complex to correlate well with the processivity and salt resistance of the polyadenylation burst (6). Evidence has been presented that, during the processive burst, VP55 translocates with respect to its RNA substrate without dissociating (8), but can finally dissociate when the 3’ end has been extended sufficiently for the creation of a ∼30–35 nt uridylate-free tract.

In a recent study, the optimal positions of uridylates for stable interaction with VP55 and addition of the first nucleotide of the oligo(A) tail were determined (1). This study employed an in vitro “selection” approach in which VP55 preferentially either bound or transferred a labeled chain-terminating adenylyl to preferred substrates within oligonucleotide pools comprising an oligo(dC) tract spiked randomly with rU residues. The use of an oligo(dC) tract as the “background” sequence stemmed from observations that: (i) unlike oligo(U), oligo(C) is entirely refractory to interaction with, and oligoadenylylation by, VP55; and (ii) the presence of ribouridylates within a DNA primer can promote polyadenylation of the latter by VP55 (8). Using this approach, two essential patches of ribouridylates were identified within oligonucleotides, one positioned at −25 with respect to the oligonucleotide 3’-OH (denoted the “distal” patch), the other at position −10 (the “proximal A” patch). A 34-mer oligonucleotide possessing two appropriately positioned tetrauridylate patches (the “double-tetraUmer”) appeared indistinguishable in VP55-binding affinity from the “optimal” substrate, (rU)$_{34}$. However, the double-tetraUmer was distinguishable in that it received only a very short (∼5 nt) oligo(A) tail in the processive oligoadenylylation burst, in contrast to the ∼30–nt tail added to (rU)$_{34}$. This indicated that the minimal uridylic content for stable binding is incapable of supporting for the full translocatory activity of the polymerase. In addition to the two internal uridylate patches, other RNA characteristics were identified that might contribute to VP55-oligonucleotide interaction,

* This work was supported by National Institutes of Health Grant 1 R01 GM51953-01A1 (to P. D. G.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

‡ To whom correspondence should be addressed: Rm. 817, Albert B. Alkek Institute of Biosciences and Technology, Texas A&M University, 2121 W. Holcombe Blvd., Houston, TX 77030-3303. Tel.: 713-677-7864; Fax: 713-677-7870; E-mail: pgershon@ibt.tamu.edu.

¶ The abbreviations used are: PAP, poly(A) polymerase; nt, nucleotide(s); EMSA, electrophoretic mobility shift assay; CoMP, cordycepin monophosphate.
namely features within the 3'-terminal ~5 nt (the “proximal B” site), along with base-independent contacts upstream of the distal patch (1).

The previous study (1) did not fully define RNA determinants recognized at the four VP55-oligonucleotide contact sites mentioned above. In the current study, we have characterized these determinants. In doing so, we have identified a pattern of ribouridylates recognized by VP55 for stable RNA interaction and addition of the first nucleotide of the poly(A) tail, and have defined the basis for ribouridylation as opposed to cytidylylation recognition. We have also identified the precise features recognized around the extreme 3’ terminus of the RNA and the number of nucleotides upstream of the distal site involved in base-independent contacts VP55-RNA.

**EXPERIMENTAL PROCEDURES**

*Materials—*3’-O-Phosphorimidylates of 5’-O-dimethoxytrityl-2’-O-tert-butylidemethylsil derivatives of ribothymidine (9), 6-methyluridine (10), 2- and 4-pyridinone ribosides (11), N3-methyluridine (12), 2-deazauridine, and 2-deazaadenosine (13) were prepared as described, with 4-O-diphenylcarbamoyl and N4-trifluoroacetyl groups being used to protect the 4-OH and 4-NH2 groups of the base, respectively. U-CE phosphorimidylase from Escherichia coli was purified using a PhosphorImager (Molecular Dynamics), by volume integration with TBE gels and quantitated using a PhosphorImager, as described (13). The mixture was incubated at 37 °C for 20 min and then transferred to 65 °C for 10 min to inactivate the enzyme. For binding reactions, 1 μl of the resulting mixture was combined with 7 μl of 50 mM Tris, pH 9.0, 5 mM dithiothreitol, 20 mM VP55, 10% glycerol, and incubated at 23 °C for 10 min before electrophoresing in an 8% polyacrylamide, 0.5 × TBE gel. Electrophoresis was at 18 V/cm for 2 h, after which the gel was vacuum-dried and both oligonucleotide and EMSA complex abundances were quantitated using a PhosphorImager (Molecular Dynamics), by volume integration with TBE gels and quantitated using a PhosphorImager, as described previously (12). All oligonucleotides were gel-purified prior to use (1). The highly purified, vaccinia-expressed VP55 employed in this study has been described (1, 5).

**Assays—**The EMSA was performed as described previously (1). Briefly, oligonucleotides were 5’-32P-labeled using T4 polynucleotide kinase. Kinase reactions contained 10 mM MgCl2, 5 mM dithiothreitol, 70 mM Tris, pH 7.6, 5 mM 2-mercaptoethanol, 50 mM NaCl, 50 mM Tris, pH 8.0, 5 mM dithiothreitol, 20 mM VP55, 10% glycerol, and incubated at 23 °C for 10 min before electrophoresing. The cordycepin monophosphate (CoMP) transferase time-course assay was performed as described (14). As shown in Fig. 1 of the previous study (1), the 3’-terminal one and two deoxyriboses sugars were switched to ribose sugars (Fig. 1C). Consistent with Fig. 1 (A and B), the effect of switching the extreme 3’-terminal sugar was dramatic, leading to a >50-fold increase in nucleotidyl transfer rate. Unexpectedly, switching the two 3’-terminal sugars to ribose led to a reaction rate lower than that observed upon switching the 3’-terminal sugar alone. The latter result indicated that, at position −2, a deoxyribose sugar might be slightly preferable to a ribose.

In a further experiment, we systematically examined the importance of a ribose sugar and uracil base at each of the three 3’-most nucleotides of a 34-mer oligonucleotide. To maximize the effects of changes around the oligonucleotide 3’ end, we eliminated all internal ribouridylates, basing oligonucleotides upon the template (dC)34-rU (denoted here (dC)34-dC-dC-rU). Both (dC)34-dC-dC-rU and a closely related oligonucleotide, (dC)31-dC-dC-rC, supported undetectable levels of CoMP transferase activity (Fig. 1D). While retaining the riboU residue at the extreme 3’-terminal position of (dC)34-dC-dC-rU, we examined the importance of a ribose sugar or uracil base at either the adjacent position (−2) or the adjacent two positions (−2 and −3) by substituting with riboU, deoxyriboU, or riboC nucleotides (Fig. 1D). Substitutions at the −2 position alone had little apparent effect upon CoMP transferase activity, as did deoxyriboU substitutions at both −2 and −3. However, substitution of positions −2 and −3 with either riboU or riboC stimulated activity significantly. These data indicated that a ribose sugar at position −3 has a significant positive effect upon CoMP transferase activity, irrespective of the associated base.

Although ribo or uracil substitutions at position −2 alone did not detectably improve the extremely low efficacy of (dC)34-rC-rC-rU as a CoMP transferase substrate (Fig. 1D, above), we wondered whether an effect of sugar species at position −2 might be apparent in the context of a more active oligonucleotide substrate, *i.e.* one with a ribose sugar at position −3. We therefore compared the activity of (dC)34-rC-rC-rU (which was active, Fig. 1D) with that of a variant, (dC)34-rC-dC-rU, differ-

---

2 A. Karpeisky and L. Beigelman, manuscript in preparation.
In the CoMP transferase time-course assay (Fig. 1E), the latter oligonucleotide appeared to be 2-fold more active than the former, indicating that a deoxyribose sugar may be preferred over a ribose at position 2. This would be consistent with the data shown in Fig. 1C, in which a ribose to deoxyribose switch at position 2 also led to a 2-fold increase in efficacy, despite being made in the context of a different oligonucleotide sequence. The effect of pyrimidine species at position 1, which had not been examined in Fig. 1 (A–D), was investigated by comparing the activity of (dC)31-rC-rC-rU (above) with that of a pyrimidine variant, (dC)31-rC-rC-rC-rU. Substitution of the uracil at position 1 with a cytosine led to a significant (2-fold) drop in activity of the oligonucleotide. We conclude that VP55's nucleotidyltransferase activity requires a ribose sugar at position 2 and is significantly enhanced by a ribose at position 3. In addition, a deoxyribose sugar at position 2 and a uracil base at position 1 each appear to impart a small positive influence. Additional experiments showed that neither a ribose sugar nor a uracil base at position 4 influence primer activity significantly (data not shown). The above data show that VP55's proximal site B interacts with the 3′-terminal three nucleotides of the primer and is primarily a ribose recognition site, with limited uracil recognition properties.

**Fig. 1. Characterization of proximal site B.** Low salt (50 mM NaCl) CoMP transferase time-course assays are shown, demonstrating ribose sugar and base requirements around the oligonucleotide 3′ end. Equivalent amounts of oligonucleotide were employed for each assay. The ordinate shows counts after electrophoresing time point samples and quantitating the bands corresponding to the labeled oligonucleotide. **A,** assay of the standard double-tetraUmer (●, 4 + 13; see text) versus an equivalent oligonucleotide differing only in the replacement of the 3′-terminal ribose with a deoxyribose sugar (○, double-tetraU-dC). **B,** Comparison of the activity of oligonucleotide tetraU-scanmer 1 ((rU)4-(dC)29-r, l) with sugar variants thereof, to compare the effect of switching the extreme 3′-terminal sugar from deoxyribose to ribose with that of a more extensive sugar switch throughout the oligonucleotide. ○, (rU)4-(rC)30; ●, (rU)4-(dC)30. Ordinate scale, counts normalized to values obtained for the optimal oligonucleotide, (rU)34. **C,** assays of various sugar-modified versions of the oligonucleotide tetraU-scanmer 14, made to elucidate the importance of the sugar species at the −2 position. In this experiment, the sugar species in the 3′-terminal two nucleotides were varied. d, deoxyriboC; other details are as for B. ●, (dC)26-(rU)4-(dC)2-rC; ○, (dC)26-(rU)4-(dC)2-dr; ▲, (dC)26-(rU)4-(dC)2-dd. **D,** assays of 3′ end variants of the oligonucleotide (dC)31-dU (referred to as (dC)31-dC-dC-rU), made to determine the effect of sugar and pyrimidine species at the 3′-terminal three nucleotides. Other details are as for A. □, (dC)31-dC-dC-dC-rU; ○, (dC)31-dC-dC-rU; ○, (dC)31-dC-dC-rU; ●, (dC)31-dC-dC-rU; ▲, (dC)31-dC-dC-rU; △, (dC)31-dC-dC-rU; ●, (dC)31-dC-dC-rU; ▲, (dC)31-dC-dC-rU. **E,** assays of 3′ end variants of the oligonucleotide (dC)31-dC-dC-rU, made to determine the effects of pyrimidine species at position −1 and sugar species at position −2. Other details are as for D. □, (dC)31-rC-rC-rU; ○, (dC)31-rC-rC-rU; ●, (dC)31-rC-rC-rU; ▲, (dC)31-rC-rC-rU; △, (dC)31-rC-rC-rU; ▲, (dC)31-rC-rC-rC.
portance of the 2′-OH of the extreme 3′-terminal nucleotide for stable VP55-oligonucleotide interaction, double-tetraU-dC was compared with the parental double-tetraUmer in the EMSA (Fig. 2C). Double-tetraU-dC exhibited a ~30% lower complex abundance, indicating that the 2′-OH of the extreme 3′-terminal sugar affects VP55-oligonucleotide interaction only slightly in the context of a 34-mer oligonucleotide containing internal uridylates. The experiments of Figs. 1 and 2C indicated that VP55’s catalytic functions can be uncoupled from stable VP55-oligonucleotide interaction through the identity of the 3′-terminal sugar. Whereas a ribose sugar within the 3′-terminal nucleotide appears to be absolutely necessary for catalytic function, the sugar species has only a marginal effect on the overall stability of VP55-oligonucleotide interaction.

What Are Minimum Numbers and Positions of Ribouridylates Required for Stable VP55 Interaction and Salt-resistant Polyadenylation?—Having characterized proximal site B, we set about characterizing VP55’s two uridylate recognition sites (proximal site A and the distal site), in terms of the minimum numbers and optimal positions of ribouridylates required for stable VP55-oligonucleotide interaction. We took a cautious approach, converting individual ribouridylates at either end of each of the two tetra-ribouridylate patches within the 4+13 double-tetraUmer oligonucleotide (1) to deoxyuridylates. This required a group of four new oligonucleotides, based upon the double-tetraUmer, in each of which the total number of ribouridylates was reduced from eight to seven. Fig. 2A shows the sequences of the 4+13 double-tetraUmer and the four new oligonucleotides, and their activities in the EMSA (in graphical form). For each oligonucleotide, reducing the total number of ribouridylates to seven did not seriously affect complex abundance with respect to the parental double-tetraUmer oligonucleotide. An oligonucleotide was therefore synthesized containing three ribouridylates in each patch, separated by 15 nt of oligo(dC) (“double-triUmer”), and individual ribouridylates at either end of each patch within the double-triUmer were converted to deoxyuridylates (Fig. 2B). Neither double-triUmer (Fig. 2C) nor the four variants thereof (Fig. 2B) exhibited significantly reduced complex abundance with respect to the double-tetraUmer. An additional oligonucleotide was therefore synthesized, containing only two ribouridylates in each patch, separated by 15 nt of oligo(dC) (“double-diUmer”), and individual ribouridylates were converted to deoxyuridylates at either end of each patch within this oligonucleotide (Fig. 2C). Complex abundances observed with the double-diUmer were comparable to those observed with both the double-triUmer and the double-tetraUmer (Fig. 2C), and reducing the number of ribouridylates in the proximal patch to only one (cf. mono/diUmer-3, mono/diUmer-4) did not seriously affect complex abundance with respect to the parental oligonucleotide (Fig. 2C). However, reducing the number of ribouridylates in the distal patch from two to one (mono/diUmer-1, mono/diUmer-2) had a seriously deleterious effect on complex abundance (Fig. 2C). These experiments indicated that the distal patch must possess at least two ribouridylates, and the proximal A patch at least one. To determine whether the two ribouridylates of the distal patch must be juxtaposed, an oligonucleotide based upon mono/diUmer-4 was synthesized in which the outside ribouridylate of the distal pair was moved one position upstream, without affecting the 15-nt inter-patch spacing (mono/diUmer-5). Since mono/diUmer-5 did not interact with VP55 to give an abundant EMSA complex (Fig. 2C), we suppose that stable interaction with VP55 requires the two ribouridylates of the distal patch to be juxtaposed.

A subset of the oligonucleotides of Fig. 2 (A–C) was polyadenylated in the presence of 150 mM NaCl. Those which did not form an abundant EMSA complex (mono/diUmer-1, mono/diUmer-2) were refractory to polyadenylation at 150 mM NaCl (data not shown). Although the others could be polyadenylated, they received only short tails in the initial processive burst (i.e., in the first two time points of the assay: Fig. 2D). Thus, the double-tetraUmer and double-triUmer received ~7–8 nt tails, the double-diUmer received 5–6-nt tails, and mono/diUmers-3 and -4 received 4–5-nt tails (Fig. 2D). This contrasted with the ~25-nt tails rapidly added to the control oligonucleotide, (rU)34.

The data of Fig. 2 (A and B) suggested that 15 nt might be the preferred spacing of the two ribouridylate patches. Thus, in Fig. 2A, a slightly greater complex abundance was observed with the oligonucleotides possessing 15 nt of oligo(dC) between the two patches (tri/tetraUmer-1 and tri/tetraUmer-3) than with those possessing 14 nt of oligo(dC) (4+13 double-tetraUmer, tri/tetraUmer-2, and tri/tetraUmer-4). Similarly, in Fig. 2B, a significantly greater complex abundance was observed with the oligonucleotides possessing 15 nt of oligo(dC) between the two patches (di/triUmer-2 and di/triUmer-4) than with those possessing 16 nt of oligo(dC) (di/triUmer-1 and di/triUmer-3). The tolerated spacing between the two patches was further examined using ligand selection methodology (1).

Thus, a 34-mer oligonucleotide was synthesized whose downstream half corresponded with mono/diUmer-3, but whose upstream half was replaced by a dC/rU pool (Fig. 2E). The oligonucleotide was subjected to the double-selection procedure described previously (1). After partial hydrolysis (1), a discrete node corresponding to the proximal A patch ribouridylate was visible. In addition, a hydrolysis node corresponding to the distal patch was observed. This latter node was somewhat broader than the presumed minimum of two nucleotides, indicating that the anchoring of VP55’s proximal site did not rigidly fix the position of the distal site. This would be expected if the two uridylate recognition sites of VP55 were able to flex slightly with respect to one another. This experiment provided further evidence that the optimal spacing between the two sites is ~14 to ~16 nt. The data of Fig. 2 can be summarized as showing that, within the context of a 34-mer oligonucleotide, an rU21–dC15–rU motif can support stable VP55-oligonucleotide interaction.

Ribose Sugars at Internal Positions within the Oligonucleotide Contribute to VP55-Oligonucleotide Interaction—We next investigated the possible contributions of internal sugars to VP55-oligonucleotide interaction. Using the EMSA, we initially compared the optimal substrate, (rU)34, with a (dU/rU)34 pool. The latter oligonucleotide is an oligouridylate whose backbone sugars are predominantly deoxyribose, but which is spiked at the 10% level with ribose (apart from the extreme 3′-terminal sugar which is 100% ribose (1)). EMSA complexes with the latter oligonucleotide were only ~0.5 times as abundant as with the former (Fig. 3A), indicating a requirement for ribose sugar(s) at internal positions for EMSA complex stability. To identify the internal positions requiring a ribose sugar, a set of eight sugar variants of the minimal oligonucleotide mono/diUmer-4 (above) was synthesized. In this set, the three essential ribouridylates (at positions 27, 26, and 10) were substituted with deoxyriobytidylates in each of the eight possible rU/dU combinations (Fig. 3B). EMSA results with the oligonucleotides (Fig. 3C) indicated that position 26 is the most sensitive to dU substitution, ~10 is almost as sensitive, ~27 is the least sensitive. Substitution at all three positions together was highly deleterious. Substitution of the resulting triple-dU variant with single rU residues reversed the deleterious effect to the greatest extent when the substitution was made at position ~26, and to a lesser extent when substitution was at
FIG. 2. Experiments to elucidate the minimum numbers and optimal positions of internal uridylates required for stable VP55-oligonucleotide interaction, and the effect of sugar species at the extreme 3'-terminal nucleotide. With the exception of double-tetraU-dC, the extreme 3'-terminal nucleotide sugar of each of the oligonucleotides tested was ribose. All sequences are shown with 5' to 3' polarity. A, upper panel, names and sequences of the oligonucleotides tested, in which single ribouridylates at the outer edges of each of the two tetraU patches of the 4+13 double-tetraUmer oligonucleotide were individually converted to deoxycytidylates. Lower panel, plot showing EMSA complex abundances for the oligonucleotides. B, upper panel, names and sequences of the oligonucleotides tested, in which single ribouridylates at the outer edges of each of the two triU patches of the double-triUmer oligonucleotide were individually converted to deoxycytidylates. Lower panel, plot showing EMSA complex abundances for the oligonucleotides. Other details are as in A. C, upper panel, names and sequences of the oligonucleotides tested, in which single ribouridylates at outer edges of each of the two diU patches of the double-diUmer oligonucleotide were individually converted to deoxycytidylates. Center panel, one of the two EMSA gels contributing to the duplicate data plotted in the lower panel. The gel shows complexes (C) obtained for the three progenitor oligonucleotides (double-tetraUmer, double-triUmer, double-diUmer, along with double-tetraU-dC, the (rU)₃₄ positive control, and the oligonucleotides whose sequences are shown in upper panel. F, free oligonucleotide. Lower panel, plot showing EMSA complex abundances averaged from the data shown in center panel and an additional, identical experiment. Values are shown as a
position -10 (Fig. 3C). Thus, the data were consistent between the two mirror-image parts of the experiment, i.e. invasion of the triple-rU variant with single dU residues, and invasion of the triple-dU variant with single rU residues. The eight oligonucleotides were also tested as CoMP transferase substrates, in the presence of both 50 and 150 mM NaCl (Fig. 3D). The depressive effect of deoxy sugar substitutions at the essential uridylates was more pronounced at elevated salt indicating that, in the CoMP transferase assay, the stringency of ribose sugar recognition increases with salt concentration. The binding assay (Fig. 3C) and the enzymatic assay (Fig. 3D) were consistent with regard to the relative activities of the eight oligonucleotides, with ribose sugars apparently being more critical at positions 2 and 10 than at position 27.

The experiments shown in Fig. 3 (C and D) indicated that internal ribose sugars contribute to stable VP55-oligonucleotide-proportion of that obtained with (rU)34. Error bars indicate range; other details are as in A. D, polyadenylation of a subset of the oligonucleotides examined by EMSA in C. Reactions were carried out in the presence of 150 mM NaCl. For each assay, samples were taken at (left to right) 0, 10, 40, and 120 s. E, flexibility of spacing between the two essential ribouridylate patches, analyzed by double-round ligand selection (see text). A pool oligonucleotide was used whose 3'9 half was identical to the 3'9 half of mono/diUmer-3, and whose 5'9 half (labeled P and underlined in the sequence) was synthesized as a (dC/rU) pool (at a ratio of 9:1). All cytidylates possessed a deoxyribose sugar except for the extreme 3'-terminal nucleotide, which contained a ribose sugar. 0, 10, 20, and 30 denote duration (min) of hot alkaline hydrolysis. Prox. A denotes the node resulting from hydrolysis at the single ribouridylate of the minimal proximal A patch. The double band after 30 min of hydrolysis presumably results from the partial loss of the adenosine moiety of CoMP from the extreme 3' end of the oligonucleotide. The artifactual band observed previously (1), denoted Art, and the calibration ladder (left) were identified by comparison with hydrolysis patterns of other oligonucleotides on the same gel (data not shown). Vertical arrows on the sequence denote positions of hydrolysis.

FIG. 3. Importance of ribose sugars at the essential ribouridylate positions. A, EMSA of two oligonucleotides, (rU)34 and (dU/rU)34, which differ only in their sugar composition (see text). C, complex; F, free oligonucleotide. Data from the gel shown, and from a duplicate experiment, were averaged and plotted. Error bars show range. B, sequences showing all eight possible combinations of ribose (r) and deoxyribose (d) sugars at the three essential uridylates of the minimal oligonucleotide, mono/diUmer-4 (representing each of the eight possible states of a 3-bit binary code). The oligonucleotides are grouped into two sets, showing substitutions of a triple-rU-containing oligonucleotide with single dU residues, and vice versa. The three-letter names of the oligonucleotides are given to the right. C, duplicate EMSAs with the oligonucleotides shown in B. Complex abundances were quantitated and mean values plotted as a proportion of the values obtained with mono/diUmer-4. Error bars show range. D, 10-min CoMP transferase assays of the eight oligonucleotides shown in B. Mean values from duplicate assays are shown, with error bars representing range. Filled and open bars represent assays in the presence of 50 and 150 mM NaCl, respectively. E, hydrolysis ladder showing the positions of important ribose sugars within the (rU/dU)34 oligonucleotide pool. This pool, which had been analyzed previously (1), was now subjected to a higher stringency double-selection procedure comprising high salt CoMP transfer followed by EMSA. D denotes the node corresponding to the distal patch; A denotes the artifactual band observed in all experiments. 0, 10, and 25 denote duration (min) of hot alkaline hydrolysis.
The experiments of Figs. 1–3 indicated that VP55-oligonucleotide interaction is promoted by uracil recognition properties of Poly(A) Polymerase.

The experiments of Figs. 1–3 indicated that VP55-oligonucleotide interaction is promoted by uracil recognition properties of Poly(A) Polymerase.

The experiments of Figs. 1–3 indicated that VP55-oligonucleotide interaction is promoted by uracil recognition properties of Poly(A) Polymerase.

Atomic Determinants of Uracil Base Recognition at the Distal and Proximal A Sites—Oligo(C) is the least active of the four common homoribopolymers (oligo(A), oligo(C), oligo(G), and oligo(U)) as a VP55 primer (8), whereas oligo(U) is the most active. Despite this, uracil (structure 1, Fig. 5A) shares more structural features in common with cytosine (structure 2, Fig. 5A) than with the other bases. Cytosine differs from uracil in (i) the transplacement of a keto with an amino group at position 4 of the pyrimidine ring, and (ii) the pKa of the nitrogen at position 3 (~9.2 for uracil, ~4.2 for cytosine (15)). Whereas the cytosine 4-amino group hydrogen atoms are good hydrogen bond donors due to their partial positive charge, the N-3 nitrogen is a good acceptor. By contrast, the 4-keto oxygen of uracil is a good acceptor and the protonated N-3 (at physiological pH) a good donor. We first set out to determine whether the dramatic difference in activity for uracil and cytosine resulted from a requirement for the 4-keto group. We addressed this by individually substituting each of the three essential uracils of the minimal oligonucleotide mono/diUmer-4 (at positions ~27, ~26, and ~10), with an analog lacking the 4-keto group. Since an analog possessing only this change (2-pyridinone) is apparently base-sensitive during RNA deprotection (16), a comparable analog, 2-pyridinone, was employed (structure 3, Fig. 5A). This analog differs from 2-pyridinone through the replacement of the N-3 nitrogen with a CH group (a “3-deaza” substitution); however, it retains aromaticity. We used 3-deazauridine (structure 4, Fig. 5A) to control for the absence of nitrogen N-3 in 2-pyridinone. Fig. 5B shows the results of duplicate EMSA analyses for the analog-substituted oligonucleotides. 2-Pyridinone does not appear to have incurred any additional binding lesion with respect to 3-deazauridine at any of the three oligonucleotide positions, indicating that the 4-keto function does not contribute to uracil recognition. Indeed, loss of this function even appeared slightly efficacious at each po-
We next wondered whether the 4-amino function of cytosine exerts a negative effect upon VP55 binding. This was analyzed by comparing a 4-keto function with a 4-amino function in both the absence and presence of nitrogen N-3, i.e. determining whether 3-deazacytidine (structure 5, Fig. 5A) was as deleterious with respect to 3-deazauridine as cytidine is with respect to uridine. Individual substitutions with ribocytidylate (Fig. 5B) clearly demonstrated cytosine’s negative effect with respect to uracil, particularly at the distal site. At position −27, 3-deazacytidine was also detrimental with respect to 3-deazauridine, indicating that cytosine’s 4-amino function may have a deleterious effect at this position in the oligonucleotide. However, the equally deleterious effects of the two analogs at positions −26 and −10 argued against a negative effect of the 4-amino function at the latter two positions. Finally, the possible importance of protonation of the N-3 nitrogen for uracil recognition was investigated using N-3-methyluridine (structure 6, Fig. 5A). The N-3-methyl modification was deleterious at all three essential ribouridylate positions (though more so at oligonucleotide position −26 than at either −27 or −10, Fig. 5B). Though we cannot rule out steric hindrance due to the methylation of N-3, the data would indicate a significant role for a protonated N-3 nitrogen in uracil recognition, particularly at position −26. We conclude that VP55 discriminates uracil from cytosine at each of the three oligonucleotide positions where a uridylate is required (−27, −26, and −10) largely through the presence of a protonated nitrogen (i.e. a hydrogen bond donor) at position 3 of the uracil ring, and to a lesser

**FIG. 5.** Effects of functional group substitutions in the uracil ring, at each of the three essential uridylic positions (−27, −26, and −10) of the “minimal” oligonucleotide mono/diUmer-4. A, structures of the pyrimidines and pyrimidine analogs employed. B, effects of the analogs on EMSA complex abundance. Data are grouped according to the position of substitution within the oligonucleotide (abscissa). Complex abundances (ordinate) were quantitated from dried gels using a PhosphorImager. Data are represented as a proportion of values obtained with the positive control oligonucleotide (mono/diUmer-4, which contains three unmodified uridines) in the same gel. 1.0 (horizontal line) represents the complex abundance of the positive control. All data represent mean values from duplicate experiments; error bars show range.
extent through the negative effect of a 4-amino function at position −27.

We expanded the experiment, taking advantage of the availability of additional uracil analogs to determine whether other functional groups around the uracil ring contribute to stable interaction with VP55, at each of the three essential ribouridylate positions of mono/diUmer-4. Thus, the importance of the 2-keto function was probed using the analog 4-pyridinone (structure 7, Fig. 5A), which lacks this function but retains aromaticity. Since 4-pyridinone was deleterious with respect to the parental analog, 3-deazauridine, at both distal positions of the oligonucleotide (i.e., −27 and −26; Fig. 5B), the 2-keto function appears to be important for VP55 interaction with the distal patch. Surprisingly, the absence of a 2-keto function appeared to be advantageous (with respect to 3-deazauridine) at position −10. The space around position 5 of the uracil ring was probed by the substitution of either a methyl (in ribothymidine, structure 8, Fig. 5A) or an iodide function (in 5-iodouridine, structure 9, Fig. 5A). Although the methyl and iodide groups are isosteric, the substituted analogs differ from each other, and from uridine, in the pKα of nitrogen N-3 (−9.7 for ribothymidine and −7.7 for 5-iodouridine (15). Ribothymidine and 5-iodouridine substitutions were both deleterious at position −26, possibly due to steric hindrance and/or water displacement. At oligonucleotide positions −27 and −10, the 5-methyl substitution appeared to be advantageous, and the 5-iodo deleterious. This difference may result from the opposing effects of the two substitutions on the pKα of N-3 (which increases from −9.2 to −9.7 with ribothymidine, but decreases to −7.7 with iodouridine), strengthening and weakening protonation at physiological pH, respectively. Finally, 6-methyluridine (structure 10, Fig. 5A) was tested. This analog probes not only steric constraints around position 6 of the ring, but also the sensitivity of the VP55-oligonucleotide interaction to local changes in syn-anti equilibrium, since 6-methyluridine is constrained in the syn conformation about the glycosidic bond (10, 17) as distinct from the conformation of uridine, which is mostly anti. The negative effect of 6-methyluridine, particularly at position −26 (Fig. 5B), indicates the possible occurrence of a prohibitive syn-constraint. Overall, each of the three essential uracils showed a different spectrum of important functional groups.

We note that structural requirements at position −26 appeared significantly more strict than at either −27 or −10. Thus, position −26 might be considered the most conservative “gate-keeper” of uracil recognition by VP55.

Within a Naturally Occurring Sequence, the Most Stable VP55 Binding Site May Not Juxtapose VP55’s Catalytic Site with the Extreme 3′-Terminal Nucleotide—The 3′-co-terminal VGFmer RNAs (5, 8) represent the portion of a natural vaccinia early mRNA immediately preceding its poly(A) tail (18). Previously, by using 5′-terminal truncations of the VGFmer sequence, we showed that 34 nt was the minimum length of VGFmer RNA that could support stable VP55 binding in the EMSA (1). Upon inspection of the VGF 34-mer sequence, the rU2-N15-rU motif (which interacts with VP55, above), was found to occur once within the oligonucleotide (Fig. 6A) with its rU2 portion falling within the only multi-uridylate patch present in the RNA. A single rU → rC change in the sequence, made to eliminate the rU2 portion, led to a drop in EMSA complex abundance by −80% (VGF34-C, Fig. 6B). This indicated that the rU2-N15-rU motif plays a major role in stable VGFmer interaction with VP55. In the VGF 34-mer, the motif occurs 4 nt closer to the RNA 5′ end than in mono/diUmer-4, the 34-mer oligonucleotide in which the motif was initially characterized (Fig. 2C). Consequently, the length of oligonucleotide downstream of the motif is extended from 9 to 13 nt. Since previous experiments showed that a −9–10 nt spacing between the proximal A patch and the 3′ end is optimal for nucleotide addition (1), we synthesized a version of the VGF 34-mer in which the 3′ end was truncated by 4 nt, to 9 nt total (VGF34D3, Fig. 6A). As would be expected, this RNA was more active than the parental VGF 34-mer in the CoMP transferase time-course assay (Fig. 6C). In the EMSA, complex abundance with VGF34D3′ was equivalent to that with the VGF 34-mer (Fig. 6B). This indicated that the truncation did not affect VP55-binding stability, and therefore that VP55 can interact stably with RNA tracts only 30 nt in length, a reduction from the previously characterized minimum of 33–34 nt (1). Moreover, since only 3 nt is present upstream of the distal patch in the 30-mer RNA, this must be the maximum size for the nonspecific (“N”) RNA binding site proposed previously (1).

**DISCUSSION**

In this study, we have characterized four RNA binding sites of VP55 that were identified previously (1), namely the distal site, proximal sites A and B, and the N site, occurring −25 nt, −10 nt, 0–5 nt, and 28–34 nt, respectively, from the 3′ terminus.

Previously, contacts at proximal site B were not clearly characterized (1) because (i) all oligonucleotides were synthesized with a 3′-terminal riboC nucleotide, and (ii) during the
[23P]CoMP-labeling reaction an artificial reaction involving the 3’-terminal ~8 nucleotides led to hypo-hydrolysis of these positions during hydrolytic analysis of selected oligonucleotide species. Here, we examined proximal site B by determining the necessity for a uracil base, a ribose sugar, or both, at each of the three 3’-terminal nucleotides (~1 to ~3). Proximal site B appeared to be predominantly a ribose recognition site, rather than a uridylate recognition site as previously thought. Although a uracil at position ~2 (the extreme 3’-terminal position) enhanced CoMP transferase activity somewhat, it was not essential. A ribose sugar (as opposed to a deoxyribose) at position ~1 was, however, essential, demonstrating the disproportionately large influence of a single oxygen atom at the 2’-OH of the extreme 3’-terminal nucleotide. This being so, an RNA molecule 3’-terminated with a single 3’-dAMP nucleotide would be refractory to further elongation by VP55. Thus, with 2’-dATP as the sole nucleoside triphosphate in an in vitro polyadenylation reaction, extension of an RNA primer could be expected to proceed rapidly for the addition of only a single nucleotide, with further elongation being extremely slow. These exact extension kinetics have indeed been observed under the above reaction conditions (see Fig. 4 in Ref. 19). Although the result was interpreted as a nucleotide-binding site phenomenon in that earlier study, it would now seem likely to result from the nucleic acid recognition properties of proximal site B. The importance of a 2’-OH at the extreme 3’-terminal nucleotide for catalysis but not for stable binding demonstrates, for the first time, an uncoupling of VP55’s RNA binding and catalytic functions, thus accounting for the ease with which VP55 interacts with internal RNA segments (8). The stringent recognition of ribose by VP55’s catalytic center provides a mechanism for ensuring RNA as opposed to DNA polyadenylation. It also predicts that an RNA substrate terminated with a single deoxynucleotide might be an efficient inhibitor of polyadenylation by VP55, since such a molecule could be neither extended nor released from the enzyme in a rapid manner. This being so, the addition of a deoxynucleotide or chain terminating nucleotide to in vitro reactions in which activated vaccinia virions are undergoing mRNA synthesis, processing, and extrusion, might be predicted to inhibit full extrusion.

In addition to the requirement for a ribose sugar at position ~1, a ribose at position ~3 appeared to significantly enhance nucleotidyl transfer and to stabilize VP55-oligonucleotide interaction (Fig. 1D), consistent with the apparent selection of substrates possessing a ~3 ribose in earlier experiments (see Figs. 4B and 5 of Ref. 1). Since the stimulatory effect upon VP55’s CoMP transferase activity of a ~3 ribose was only ~2% that of correctly positioned internal ribouridylates (Fig. 1, compare A and D), the latter apparently play a much greater role in bringing the 3’-OH to VP55’s catalytic center for nucleotidyl transfer. Nonetheless, an oligonucleotide possessing a ~3 ribose in combination with correctly positioned internal uridylates is demonstrably superior to its equivalent possessing the latter but not the former (Fig. 4).

Three internal ribouridylates were found to be required for stable oligonucleotide-VP55 interaction, and their occurrence at positions ~10, ~26, and ~27 in the oligonucleotide constituted an rU2-N17-rU motif. This motif was identified once within the VGF 34-mer RNA (whose sequence represents the segment of a characterized vaccinia mRNA immediately preceding its poly(A) tail; see Refs. 1 and 13) and was found to be necessary for stable VP55-RNA interaction. However, the position of the motif within the VGF 34-mer, 13 nt from its 3’ end, is significantly greater than the ~9 nt optimum distance for CoMP transfer characterized in selection experiments (1). Although deletion of the 3’-most 4 nt of the VGF 34-mer significantly improved its efficiency as a CoMP transferase substrate, the parental VGF 34-mer was nonetheless active (Fig. 6C), indicating that VP55 can tolerate some deviation from the optimal spacing during adenylate addition. Such flexibility could permit VP55 to remain static during short bursts of processive polyadenylation, such as the 3–8-nt bursts observed with the minimal U-containing oligonucleotides (Fig. 2D). The position of the motif in the VGF 34-mer, only 3 nt from the 5’ end of the RNA, indicated that the nonspecific contacts upstream of the distal site previously named N site (nonspecific) contacts (1), are less extensive than previously thought. Nonetheless, the importance of this short upstream region is emphasized by the loss of EMSA complex abundance upon truncation of two additional residues from the 5’ end of the VGF 34-mer (1).

The three essential internal ribouridylates were found to be recognized primarily via their uracil moieties. The recognized functional groups of each of the three essential uracils were explored using uridine analogs. Each of the three uracils exhibited a distinctive fingerprint of recognized functional groups, with the ~26 position being the most sensitive to alterations in the pyrimidine ring. The residue at position ~26 therefore appears to act as the most stringent gatekeeper of uracil recognition. Despite the structural similarity between uracil and cytosine, VP55 discriminates strongly against ribocytidine at the three essential ribouridylate positions. Although it is not clear what role this could play in vivo, the mechanism was readily elucidated. Cytosine and uracil differ in having opposing polarities of hydrogen bonding donors and acceptors along the N-3–C-4 edge of the pyrimidine ring (15). For each of the three uridylates, discrimination appeared to stem predominantly from the necessity for a hydrogen bond donor at nitrogen N-3, although, at the ~27 uridylate, an inhibitory effect of cytosine’s 4-amino function was also apparent. The six-membered ring of guanine possesses a nitrogen (N-1) at a position equivalent to that of uracil’s N-3 nitrogen, with an equivalent pKa (~9.4), along with a keto function equivalent to uracil’s 4-keto. Interestingly, an oligo(G) primer is able to support limited semi-processive polyadenylation by VP55 at 60 mM NaCl (8), although unlike oligo(U), oligo(G) was inactive at 150 mM NaCl and did not form a detectable EMSA complex with VP55. The lower affinity of VP55 for guanine than for uracil presumably stems from the distinctive structures within the guanine ring at positions equivalent to those in uracil where substitutions adversely affect interaction with VP55. Ribose sugars appeared to stabilize VP55-oligonucleotide interaction at the positions of the three essential ribouridylates. Whether VP55 somehow co-ordinates its recognition of the uracil base and the ribose sugar at the internal ribouridylate recognition sites, perhaps to orientate the RNA chain, is a matter for conjecture, since inadequate VP55 binding affinity might preclude an examination of the contribution of ribose in the context of bases other than uracil.

The above data provide a static view of the VP55-oligonucleotide contacts necessary for stable complex formation and addition of the first nucleotide of the poly(A) tail. Additional studies would be necessary to obtain a dynamic picture of the contacts occurring during translocation by VP55 and 30-nt oligo(A) tail formation.

Acknowledgment—We thank Steve W. Lockless for reading the manuscript.

REFERENCES
1. Deng, L., and Gereshon, P. D. (1997) EMBO J. 16, 1103–1113
2. Gershon, P. D., Ahn, B.-Y., Garfield, M., and Moss, B. (1991) Cell 66, 1269–1278
3. Moss, B., Rosenblum, E. N., and Gereshowitz, A. (1975) J. Biol. Chem. 250, 31551
4722–4729
4. Nevins, J. R., and Joklik, W. K. (1977) J. Biol. Chem. 252, 6939–6947
5. Gershon, P. D., and Moss, B. (1992) Genes Dev. 6, 1575–1586
6. Gershon, P. D., and Moss, B. (1993) J. Biol. Chem. 268, 2203–2210
7. Schnierle, B. S., Gershon, P. D., and Moss, B. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 2897–2901
8. Gershon, P. D., and Moss, B. (1993) EMBO J. 12, 4705–4714
9. Strobel, S. A., Cech, T. R., Usman, N., and Beigelman, L. (1994) Biochemistry 33, 13824–13835
10. Beigelman, L., Karpeisky, A., and Usman, N. (1995) Nucleosides & Nucleotides 14, 895–899
11. Matulic-Adamic, J., Gonzalez, C., Usman, N., and Beigelman, L. (1996) Bioorg. Med. Chem. Lett. 6, 373–378
12. Burgin, A. B., Gonzalez, C., Matulic-Adamic, J., Karpeisky, A., Usman, N., McSwiggen, J., and Beigelman, L. (1995) Biochemistry 35, 14090–14097
13. Yuen, L., and Moss, B. (1986) J. Virol. 60, 320–323
14. Thomson, J. G., and Gershon, P. D. (1995) Biotechniques 19, 416–425
15. Saenger, W. (1984) Principles of Nucleic Acid Structure, Springer-Verlag, New York
16. Murray, J. B., Adams, C. J., Arnold, J. P., and Stockley, P. G. (1995) Biochem. J. 311, 487–494
17. Suck, D., and Saenger, W. (1972) J. Am. Chem. Soc. 94, 6520–6526
18. Yuen, L., and Moss, B. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 6417–6421
19. Shuman, S., and Moss, B. (1988) J. Biol. Chem. 263, 8405–8412