The EMBO Journal vol.12 no.12 pp.4705–4714, 1993

Uridylate-containing RNA sequences determine specificity for binding and polyadenylation by the catalytic subunit of vaccinia virus poly(A) polymerase

Paul David Gershon and Bernard Moss

Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

Communicated by R.Perry

VP55, the catalytic subunit of vaccinia virus poly(A) polymerase, has the remarkable property of adding 30–35 adenylates to RNA 3′ ends in a rapid processive burst before an abrupt transition to slow, non-processive adenylate addition. Here, we demonstrate that this property results from the affinity of the enzyme for uridylate residues within the 3′ 31–40 nt of the RNA primer. At physiological salt concentrations, both polyadenylation and stable VP55 binding required the presence of multiple uridylates within a 31–40 nt length of RNA, though specific RNA sequences were not necessary. Even DNA in which the deoxythymidylate residues were replaced with ribouridylates, could be polyadenylated in a processive manner. Both the unmethylated pyrimidine ring and a 2′-OH on the associated sugar are features of ribouridylates that are important for priming. The abrupt termination of processive polyadenylation was attributed to translocation of VP55 along the nascent poly(A) tail, which lacks uridylates for stable binding. As evidence for translocation and interaction with newly synthesized RNA, other homopolymer tails were synthesized by VP55 in the presence of Mn^{2+}, which relaxes its donor nucleotide specificity. Only during poly(U) tail synthesis did processive nucleotide addition fail to terminate.

Key words: mRNA/poly(A) polymerase/poly(A) tail/RNA recognition/vaccinia

Introduction

Vaccinia virus is a large DNA virus that replicates in the cytoplasm of the infected cell. Although the transcription apparatus is virus-encoded (Moss *et al.*, 1991), viral mRNAs, like their cellular counterparts, possess capped 5′ ends (Wei and Moss, 1975) and poly(A) tails (Kates and Beeson, 1970; Nevins and Joklik, 1975). Poly(A) is added to the 3′ ends of vaccinia early transcripts following transcriptional termination rather than a specific RNA cleavage event (Rohrmann *et al.*, 1986). Factors and enzymes required for the synthesis and modification of early viral mRNAs are packaged within the infectious virus particle, facilitating their isolation and purification. The vaccinia virus poly(A) polymerase (PAP) has been purified as a heterodimer (Moss *et al.*, 1975; Nevins and Joklik, 1977) of the virus-encoded polypeptides VP55 and VP39 (Gershon *et al.*, 1991). Despite its apparent stability during purification, the heterodimer was readily dissociated by incubation with immobilized antibody to an N-terminal peptide of VP55 and subsequent washing (Gershon *et al.*, 1991). The resulting monomeric VP55 was found to possess polyadenylation catalytic activity. VP39 is present in molar excess over VP55 *in vivo*, and so exists in monomeric as well as heterodimeric form. No polyadenylation catalytic activity is associated with monomeric VP39. Instead, this subunit has an affinity for poly(A) and stimulates the formation of long poly(A) tails by VP55. VP39, either by itself or in association with VP55, is also a cap-specific mRNA (nucleoside-$O^{2'}$-)-methyltransferase (Schnierle *et al.*, 1992), indicating its participation in the modification of both ends of nascent mRNA. No role in 3′-polyadenylation for polypeptides other than VP55 and VP39 has been established, either *in vivo* or *in vitro*.

The properties of individually overexpressed and highly purified VP55 and VP39 have been determined (Gershon and Moss, 1992, 1993). VP55 polyadenylates a primer representing the 3′ end of a nascent vaccinia virus mRNA bimodally: 30–35 adenylates are added in a rapid, highly processive burst, after which polyadenylation decelerates abruptly and becomes non-processive. By using primers with preformed poly(A) tails it was shown that the transition to non-processive polyadenylation is regulated by the net length of the oligo(A) tail rather than by the number of adenylate additions catalyzed. VP55 adds adenylates to poly(A) primers greater than 34 nucleotides in length, but only in the slow non-processive mode. VP39 dramatically stimulates the rate of elongation of poly(A), or RNA primers possessing poly(A) tails greater than 30 nucleotides in length, by converting the slow, non-processive polyadenylation that occurs after the rapid burst in the presence of VP55 alone, to a more rapid, semi-processive and NaCl-resistant reaction. Nevertheless, VP39 does not compromise the selectivity of the polymerase for a non-polyadenylated mRNA 3′ end in the presence of a large molar excess of poly(A). The properties of the dissociable VP39 subunit are therefore consistent with a role in increasing the affinity of the polymerase for the growing poly(A) tail.

Here, we provide evidence that the processive addition of 30–35 adenylates by VP55 can be explained by its specific binding affinity for 31–40 nucleotide segments of RNA containing uridylates. According to this model, the translocation of VP55 to the nascent poly(A) tail, which lacks uridylates necessary for stable binding, causes an abrupt cessation of processivity.

Results

Stable, heparin-resistant, divalent cation-independent RNA binding by VP55

Since VP55 adds 30–35 adenylates to the VGF 60mer RNA [the 60 nt immediately preceding the poly(A) tail of a transcript of the vaccinia virus growth factor (VGF) gene (Yuen and Moss, 1986)] in a processive manner (Gershon and Moss, 1992), this RNA was used for initial VP55

binding studies. Fractions from the final column used in the purification of VP55 [poly(A)-Sepharose], over which recombinant VP55 had been purified to apparent homogeneity (Gershon and Moss, 1992), were incubated with ^{32}P end-labeled VGF 60mer RNA and analyzed by EMSA (Figure 1A). A novel band (labeled '1' in Figure 1A) migrating more slowly than the uncomplexed 60mer RNA was detected in the assays of fractions 35–55. These fractions were previously shown to contain VP55 (Gershon and Moss, 1992). The slowly migrating band disappeared on $CHCl_3$ extraction of samples prior to electrophoresis (data not shown), consistent with it being a non-covalent protein–RNA complex.

To determine the binding specificity of VP55 for nucleic acid, the VGF 60mer RNA was pre-incubated with excess amounts of the polyanionic compound, heparin, before the addition of VP55. EMSA complex formation was fully resistant to 40 μg/ml heparin and partially resistant to 200 μg/ml heparin (Figure 1B). At the latter concentration, heparin was present in a 150-fold molar excess over RNA. Next, the VP55–VGF 60mer RNA complex was shown to be divalent cation-independent. Thus, the $MgCl_2$ remaining from the RNA labeling reaction was removed by an additional round of electrophoresis in an EDTA-containing gel and this RNA was used in an EMSA experiment. In addition, EDTA was present in the non-denaturing gel used for resolution of the complex (Materials and methods). A stable complex was observed (Figure 1C).

VP55 binds stably to RNA segments 31–40 nt in length

To determine whether there is a minimum RNA size required for stable interaction with VP55, members of a nested set of RNAs representing the 3′ terminal 30, 40, 50, 60, 80, 100 and 110 nt of the model VGF gene transcript (VGF 30mer to 110mer RNAs) were individually incubated with VP55 and complexes were resolved in non-denaturing gels.

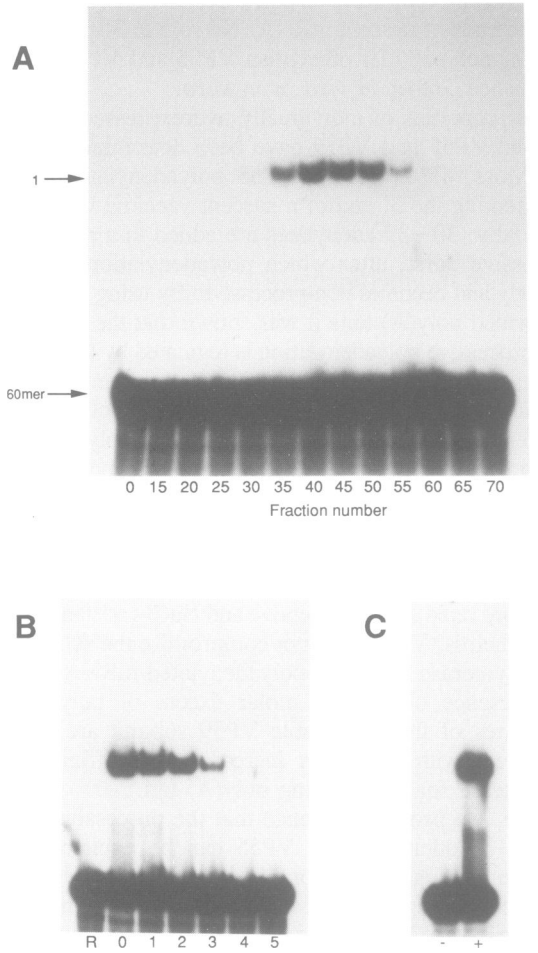

Fig. 1. RNA binding properties of VP55 examined by EMSA. (**A**) Fractions from the final VP55 purification step, poly(A)-Sepharose chromatography (Gershon and Moss, 1992), were assayed for stable interaction with the VGF 60mer RNA. Track 0, no added fraction; the numbers 15 to 70 under the remaining tracks are the fraction numbers; band labeled 1 (arrow), RNA–protein complex. (**B**) Heparin resistance of the VGF 60mer–VP55 complex. R, RNA incubated in absence of heparin or VP55; tracks 0, 1, 2, 3, 4 and 5, pre-incubation of the RNA with 0, 8, 40, 200, 1000 and 5000 μg/ml heparin, respectively, prior to VP55 addition. (**C**) EMSA of the VGF 60mer RNA after the removal of $MgCl_2$ remaining from the RNA labeling reaction by electrophoresis in an EDTA-containing polyacrylamide gel. – and + represent incubations in the absence and presence, respectively, of VP55.

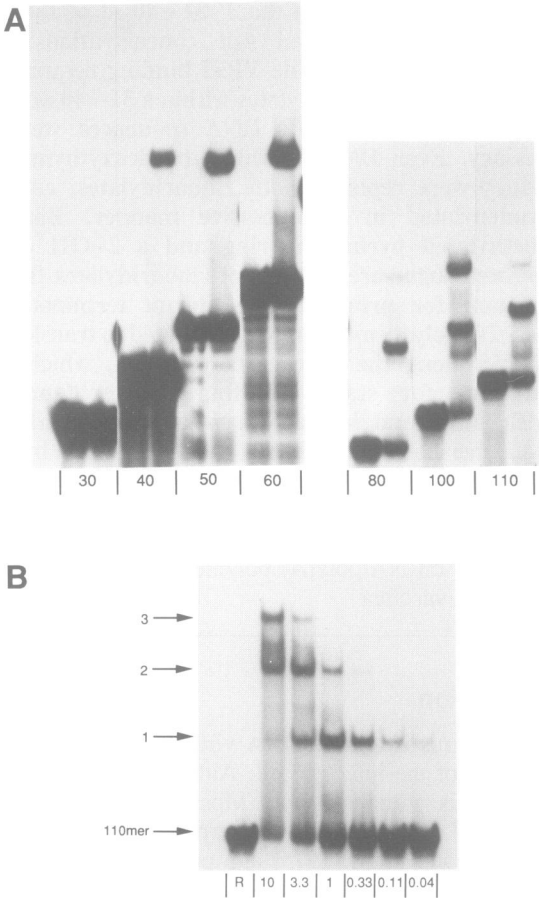

Fig. 2. Effect of RNA length on VP55 binding. (**A**) EMSAs of RNAs representing the 3′ terminal 30, 40, 50, 60, 80, 100 and 110 nt of the model VGF gene transcript. Numbers 30–110 represent lengths (in nt) of the 3′ co-terminal RNA segments. For each RNA, the first and second tracks represent incubations in the absence and presence, respectively, of VP55. The 80–110mer RNA complexes were resolved in a 6% polyacrylamide gel. (**B**) EMSAs in which VP55 was incubated with the VGF 110mer RNA at various VP55:RNA molar ratios. Three-fold dilutions of VP55 were incubated with fixed amounts of VGF 110mer RNA in the presence of 8 μg/ml heparin, and complexes were resolved in a 5% polyacrylamide gel. R, RNA alone; tracks 10 to 0.04, VP55:RNA molar ratio; arrows numbered 1, 2 and 3 indicate complexes corresponding to the binding of one, two and three molecules respectively of VP55 per RNA molecule.

With the exception of the VGF 30mer, VP55 interacted stably with each of the RNAs (Figure 2A). In similar experiments with various RNAs smaller than 30 nt in length, VP55–RNA complexes were not detected (data not shown). These results indicated that the minimum RNA size necessary for stable VP55 binding is 31–40 nt. With the 80, 100 and 110mer RNAs, two or more major complexes were observed (Figure 2A). The variation in abundance of the uppermost complex observed after incubation of VP55 with these RNAs correlated with differences in VP55:RNA molar ratio due to slightly different amounts of RNA in the incubations. To determine whether the additional complexes represented the binding of more than one molecule of VP55 per RNA molecule, EMSAs were performed in which VP55 was incubated with the 110mer at various VP55:RNA molar ratios between 10 and 0.04. Three major complexes were observed (numbered 1, 2 and 3; Figure 2B). The relative amounts of these complexes changed with VP55:RNA molar ratio in a manner consistent with their representing the binding of one, two and three molecules of VP55 per RNA molecule, respectively. The inability to obtain more than a single stable complex with the 60mer, two complexes with the 80mer (data not shown) and three with the 110mer (Figure 2B) was consistent with the binding of VP55 to 31–40 nt segments of RNA. The binding of multiple VP55 molecules suggested that the 3′ end of the RNA was not required for stable VP55–RNA interaction in the EMSA, but did not distinguish which of the bound VP55 molecules was potentially active in 3′ polyadenylation.

VP55 does not bind stably to poly(A)

Since VP55 extends poly(A) primers, albeit non-processively (Gershon and Moss, 1992), poly(A) was assayed for stable complex formation with VP55. VP55 was incubated with poly(A) molecules 25, 50, 75, 100, 150 and 200 nt in length, in the presence or absence of ATP, and the products were resolved in a non-denaturing gel (data not shown). Stable complexes were not observed, consistent with the non-processive polyadenylation of poly(A) primers by VP55.

Specific RNA sequences are not required for stable VP55–RNA interaction and NaCl-resistant polyadenylation

40mer RNAs were designed and synthesized to examine the sequence requirements for VP55 to conduct polyadenylation in a rapid, processive burst or in a slow non-processive fashion at 60 and 150 mM NaCl. Stable binding of the RNAs by VP55 was examined directly in the EMSA. The sequences of 40mer RNAs employed are given in Table I. During their design, they were each examined for regions with potential for Watson–Crick base-pairing, and these were avoided by minimal sequence modifications. The prototypic RNA was the VGF 40mer. RNAs of 40 nt were employed since this length is sufficient for stable VP55 binding (see above), and because the polyadenylation of RNAs 40 nt in length or greater is much more NaCl-resistant than that of shorter RNAs (data not shown). Five 40mers with identical base composition, but each possessing a random permutation of the VGF 40mer sequence ('scramblemers', Table I), were assayed for polyadenylation by VP55 in the presence of 150 mM NaCl, and for stable VP55 binding (Figure 3A and B respectively). Each of the five RNAs supported both stable VP55 binding and a burst of polyadenylation at 150 mM NaCl, indicating that specific RNA sequences are not required. The polyadenylation burst was shorter at 150 mM NaCl than at lower NaCl concentrations, as found with other

Table I. Sequences of 40mer RNAs used for determining the basis for RNA binding and processive polyadenylation by VP55

	−40	−30	−20	−10
Prototype	·	·	·	·
VGF 40mer	UUAUGUAUAUUUAUAAAAAUGCUAAGUAUGCGAUGUAUCU			
Scramblemers	·	·	·	·
1	GGAAUUUUGCUAAUUUGAUAUUAAAAUCGUAAACUAUAUGU			
2	GACGAAUGAUUGUUAAUUUAUAUACAGUAUGAUCUUAAUU			
3	GGACAAUGAUUGUUAAUUUAUAUACAGUAUGAUCUUAAUU			
4	GUUAUAUUGAGUAUUUUUAAAAGACAGUACUAGUUUAACU			
5	GGUUAUAUUAGUAUUUUUAAAAGACAGUACUAGUUUAACU			
Omitmers	·	·	·	·
ACGmer	GGGCAGAGACAAGCCACGCAACCAGGGACGCCGCAAGAGC			
ACUmer	GUCAUCCUAAUCUCCAACAUCUUACUAUUCCAAUAACUCA			
AGUmer	GAGUUGGUAUUAGAGAGAAGUGAUUGGAAGAUAGUUAUGU			
CGUmer	GGCUGCUCGCCGCCCUGGUCUGCCGGCUUGGUUUGUUCUG			
Umers	·	·	·	·
U1mer-1	GCUGGCAGCGAGCACACGACCGAACAGCACGCAGAGAGAC			
U1mer-2	GAAGAGCCACGACCGAGCCACCCAGGCGAGAGCUACGGACA			
U5mer-1	GCCUCACAGACUGGACAGUGACCGAGGACUGACGAUGCCA			
U5mer-2	GUGGCAACCUCGCACGAGGACUCGACACAAGUGAUCGCAG			
U8mer-1	GGCUACAGCAUGUGACCACUGCUCUCAGCGAGGAAAGCUU			
U8mer-2	GUCACGCGUACACUAGGAUGACGCUACUCUAGGCACGUGA			
Tetra-Umers	·	·	·	·
1	GGGCAGAUUUUAGCCACGCAACCAGGGACGCCGCAAGAGC			
2	GGGCAGAGACAAGCCACUUUUCCAGGGACGCCGCAAGAGC			
3	GGGCAGAGACAAGCCACGCAACCAGGGUUUUCGCAAGAGC			
4	GGGCAGAGACAAGCCACGCAACCAGGGACGCCGCAAUUUU			
Alternating copolymers	·	·	·	·
$(AG)_{20}$	ACACACACACACACACACACACACACACACACACACACAC			
$(AC)_{20}$	AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG			

Homoribopolymers are not shown. The positions of uridylates in the 'Umers' and 'tetra-Umers' are underscored.

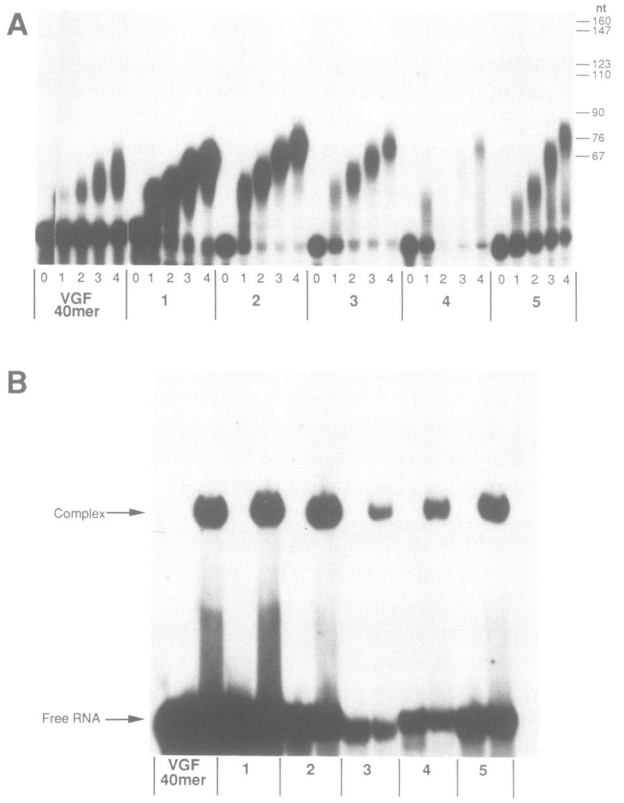

Fig. 3. Effect of RNA sequence on polyadenylation and stable binding by VP55. Five 40mer RNAs with identical base composition, but each possessing a random permutation of the VGF 40mer sequence ('scramblemers', Table I) are designated 1–5 (bold numerals). (**A**) Polyadenylation of the VGF 40mer RNA and the five scramblemers in the presence of 150 mM NaCl. Tracks 0–4 (small numerals), products sampled at 0, 15, 40, 120 and 300 s. The smaller amount of product RNA with scramblemer 4 is due to experimental variation, and is not considered significant. (**B**) EMSAs of the six RNAs. The first and second tracks for each RNA represent incubations in the absence and presence, respectively, of VP55.

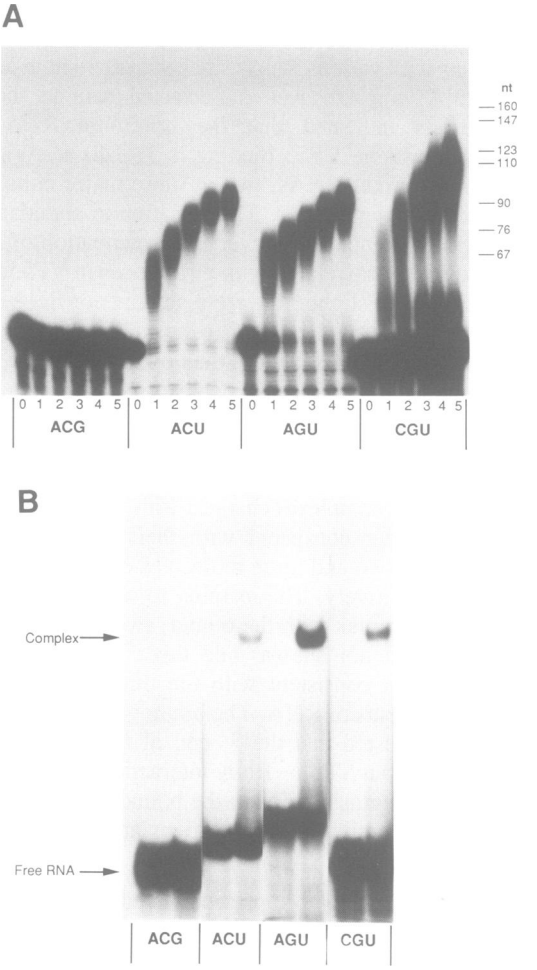

Fig. 4. Effect of base composition on polyadenylation and stable binding by VP55. Four 40mer RNAs were synthesized in each of which one of the four standard mRNA base types was omitted but the remaining three were included in equal proportion and random sequence ('omitmers', Table I). ACG, ACU, AGU and CGU designate the ACGmer, ACUmer, AGUmer and CGUmer RNAs respectively. (**A**) Polyadenylation of the four RNAs in the presence of 60 mM NaCl. Tracks 0 to 5 show products sampled at 0, 15, 40, 120, 300 and 600 s. (**B**) EMSAs of the four RNAs. The first and second tracks for each RNA represent incubations in the absence and presence, respectively, of VP55.

RNA primers such as the VGF 60mer (Gershon and Moss, 1993).

Uridylate-containing RNA is necessary and sufficient for stable VP55–RNA interaction and NaCl-resistant polyadenylation

Since rapid processive polyadenylation was not RNA sequence specific, the effects of the four standard types of mRNA base were considered. Four 40mer RNAs ('omitmers', Table I) were synthesized, in each of which one of the four base-types was omitted but the remaining three were included in equal proportion and random sequence. When assayed for polyadenylation by VP55 in the presence of 60 mM NaCl, three of the RNAs supported a burst of polyadenylation; however, the RNA lacking uridylates (the ACGmer), did not (Figure 4A). Furthermore, in contrast to the other three RNAs, the ACGmer did not stably bind VP55 (Figure 4B). Thus, uridylates within the RNA appear to be necessary for both a burst of polyadenylation and stable RNA binding by VP55. The ability of the RNA lacking adenylates (the CGUmer) to bind VP55 and to prime polyadenylation at 60 mM NaCl indicated that the inability of the ACGmer to so function could not be attributed solely to high G+C content. Nevertheless, high G+C content may be slightly inhibitory, since polyadenylation of the CGUmer was inhibited by 150 mM NaCl (data not shown), and appears to be partially inhibited by 60 mM NaCl (Figure 4A).

The above experiment suggested the importance of uridylates within a 40mer RNA primer. To determine whether uridylates are sufficient, the homoribopolymer U_{40} was assayed for polyadenylation by VP55 in the presence of 60 and 150 mM NaCl (Figure 5A), and for stable VP55 binding in the EMSA (Figure 5B). The homoribopolymers C_{40} and G_{40} were used as controls. U_{40} was polyadenylated with a burst at both 60 and 150 mM NaCl, and the RNA interacted stably with VP55, indicating that the presence of uridylates is indeed sufficient. C_{40} was refractory to polyadenylation at both 60 and 150 mM NaCl, and did not form a stable complex with VP55. G_{40} also did not form a stable complex with VP55, and was refractory to polyadenylation at 150 mM NaCl. However, somewhat

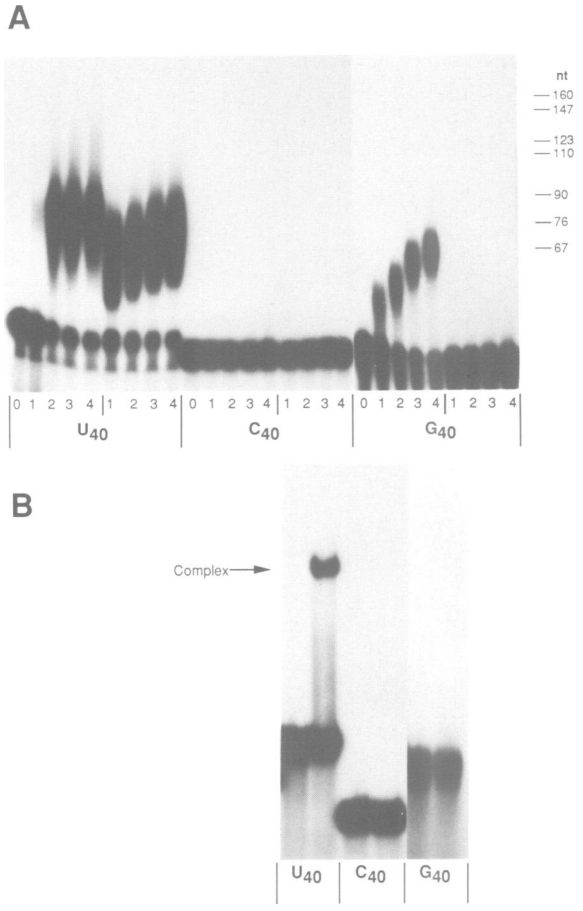

Fig. 5. (**A**) Polyadenylation of homoribopolymers U_{40}, C_{40} and G_{40} in the presence of 60 and 150 mM NaCl. Tracks 0 to 4 are labeled as for Figure 3. For each RNA, the left- and right-most assays were performed in the presence of 60 and 150 mM NaCl, respectively. (**B**) EMSAs of the three homoribopolymers. The first and second tracks for each RNA represent incubations in the absence and presence, respectively, of VP55.

surprisingly, G_{40} was polyadenylated with a short burst at 60 mM NaCl. Addition of adenylates to G_{40} was semi-processive, as only a portion of the primer was initially extended, but the rate of addition of the first 15–20 adenylates was dependent on the concentration of VP55 (data not shown).

Effects of uridylate content and position on binding and polyadenylation by VP55

U_{40}, the 'scramblemers' and the uridylate-containing 'omitmer' RNAs, which could be polyadenylated in the presence of 150 mM NaCl and could bind stably to VP55 (above), each possessed a relatively high uridylate content (>33%), and also possessed groups of juxtaposed uridylates. To assess the importance of juxtaposed uridylates in RNAs when the uridylate content is high, the alternating copolymer $(UC)_{20}$ was tested (Figure 6A). This RNA could be polyadenylated in the presence of 60 mM or 150 mM NaCl and could form a stable complex with VP55, indicating that uridylate juxtaposition is not necessary within a uridylate-rich RNA.

To investigate the importance of uridylate content, random-sequence 40mer RNAs possessing relatively low uridylate contents, of between 2.5 and 20% were synthesized ('Umers', Table I). Within each RNA, A, C and G residues were present in similar proportions, and for each uridylate content, two RNAs were synthesized. Polyadenylation of each of the six resulting RNAs was assayed in the presence of 60 and 150 mM NaCl. None could be polyadenylated in the presence of 150 mM NaCl (data not shown). In the presence of 60 mM NaCl, four of the RNAs could be polyadenylated (U1mer-2, U5mer-2, U8mer-1 and U8mer-2; Figure 6B), though only two or three adenylates were added to the latter RNA. Thus, polyadenylation of the 'Umer' RNAs did not correlate simply with uridylate content. The processivity of polyadenylation of the four permissive 'Umer' RNAs in Figure 6B could be determined, since each was assayed at three RNA:VP55 molar ratios (primer challenge cannot be used to determine processivity due to the rapidity of the burst (Gershon and Moss, 1992). Thus, the single uridylate of U1mer-2 was unable to confer processivity but did support non-processive polyadenylation. Furthermore, the rate of non-processive polyadenylation of this RNA was greater than that of the prototypic non-processive primer, poly(A), at equivalent VP55:RNA molar ratios and NaCl concentrations (Gershon and Moss, 1992, 1993; data not shown). Polyadenylation of U5mer-2 and U8mer-1 proceeded in short, saltatory, processive bursts interspersed with non-processive polyadenylation. In EMSAs of the six Umer RNAs (Figure 6C), a stable complex with VP55 was detected only with U5mer-2 and U8mer-1, the two RNAs that became polyadenylated in short processive bursts. The very low abundances of the observed VP55–RNA complexes, compared with that of the prototypical VGF 40mer–VP55 complex, indicated relatively low stabilities.

In the 'Umer' RNAs, the reduced uridylate content resulted in reduced numbers of juxtaposed uridylates. To distinguish between these two parameters, four 40mer RNAs were synthesized, each possessing a low (10%) uridylate content, but in which the uridylates were juxtaposed ('tetra-Umers', Table I). In these RNAs, a four-uridylate cluster replaced various four base segments of the ACGmer RNA, which was refractory to polyadenylation and stable VP55 binding (above). As with the 'Umers', polyadenylation of each of the 'tetra-Umers' was inhibited by NaCl at a concentration of 150 mM; however, at 60 mM NaCl, each of the four RNAs could be polyadenylated (Figure 6D). Since each RNA was assayed at three RNA:VP55 molar ratios, the processivity of the burst of polyadenylation could be determined. Processivity appeared high for tetra-Umer-2, while the polyadenylation of the other three RNAs appeared more saltatory in character, and may include bursts of processive polyadenylation interspersed with non-processive addition. When assayed for stable interaction with VP55 in the EMSA (Figure 6E), three of the RNAs formed a detectable complex with VP55. The high abundances of the complexes compared with those with the 'Umer' RNAs (Figure 6C), indicated a much greater stability. It is not known why a complex between VP55 and tetra-Umer-4 was not observed in repeated experiments, but this may correlate with the low processivity of the first two or three adenylate additions, apparent on careful inspection of Figure 6D.

Overall, the data of Figure 6B–E suggest that VP55 can interact simultaneously with multiple uridylates. These need not be juxtaposed, though in uridylate-poor RNAs,

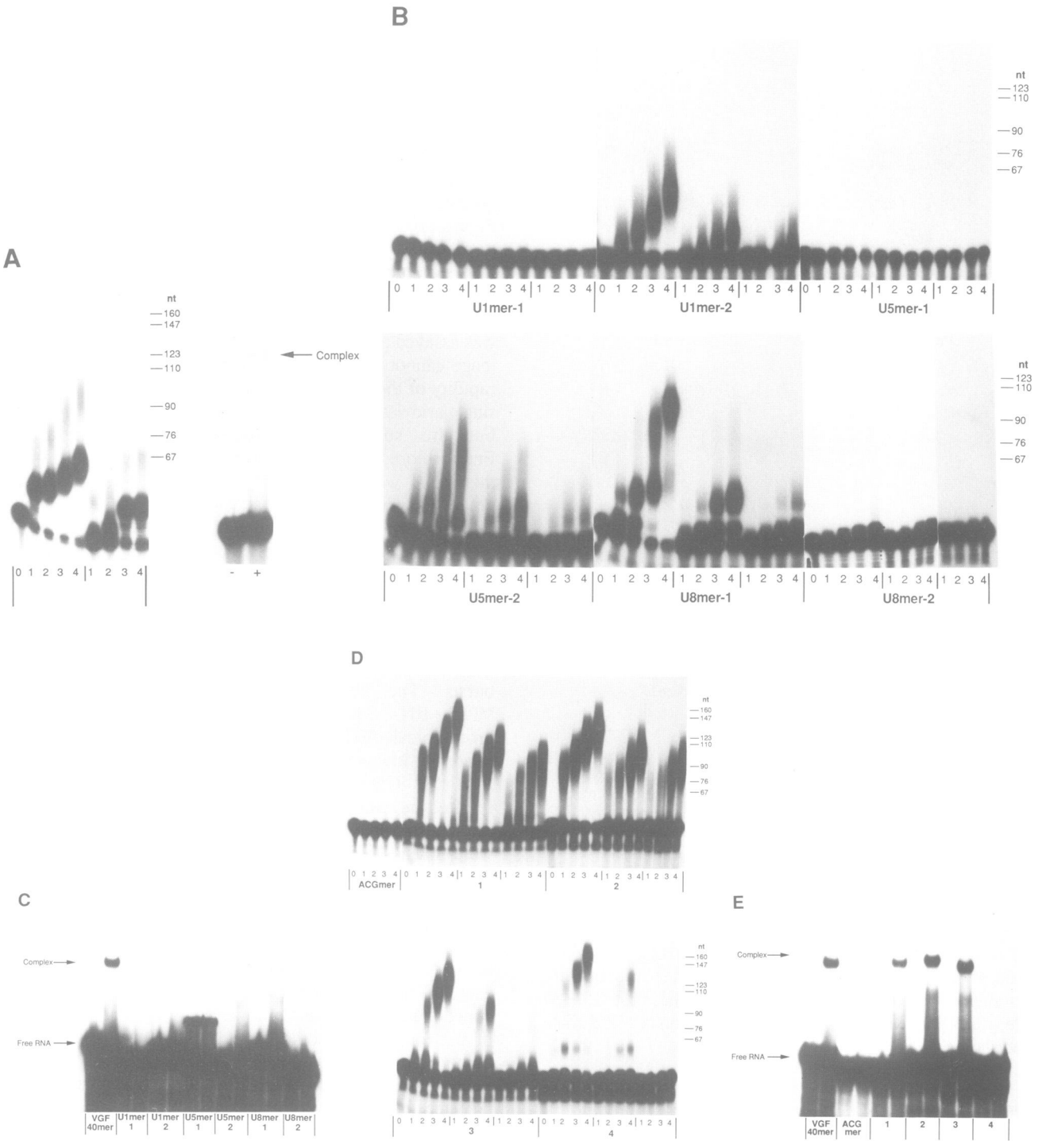

Fig. 6. Effects of low uridylate content and juxtaposition on polyadenylation and stable binding of RNAs by VP55. (**A**) Left panel: two polyadenylation assays of the RNA $(UC)_{20}$. The left- and right-most assays were performed in the presence of 60 and 150 mM NaCl, respectively. Tracks 0 to 4 are labeled as for Figure 3. Right panel: EMSA of the RNA $(UC)_{20}$. − and + indicate incubation in the absence and presence, respectively, of VP55. (**B**) Polyadenylation of six 40mer RNAs with low uridylate content ('Umers', Table I), in the presence of 60 mM NaCl. The three assays for each RNA represent (left to right) RNA:VP55 molar ratios of 1, 3 and 9, varied by dilution of VP55. Tracks 0 to 4 are labeled as for Figure 3. (**C**) EMSAs of the six Umer RNAs. The first and second tracks for each RNA represent incubations in the absence and presence, respectively, of VP55. The slowly migrating band observed with U5mer-1 was not observed in denaturing gels and is presumed to result from limited secondary structure formation. (**D**) Polyadenylation of the ACGmer RNA and four related 40mer RNAs, each possessing a differently positioned cluster of four uridylates ('tetra-Umers', Table I), in the presence of 60 mM NaCl. 1, 2, 3 and 4 (bold numerals) designate the RNAs. The three assays for each RNA represent different RNA:VP55 molar ratios as in (B). Tracks 0 to 4 (small numerals) are labeled as for Figure 3. (**E**) EMSAs of the VGF 40mer and ACGmer control RNAs and the four 'tetra-Umer' RNAs. The first and second tracks for each RNA represent incubations in the absence and presence, respectively, of VP55.

juxtaposition may increase the likelihood of multiple, simultaneous interactions occurring.

Both an unmethylated pyrimidine ring and the presence of a 2′-OH on the associated sugar are features of ribouridylates required to mediate polynucleotide recognition and polyadenylation by VP55

To investigate the chemical determinants of ribouridylate recognition by VP55, sequences equivalent to that of the VGF 40mer RNA were synthesized as DNA, two RNA−DNA chimeras, and as DNA with a modified base-type. One of the two chimeras, RNA-dT, was composed entirely of ribonucleotides except for the ribouridylates, which were replaced with deoxythymidylates. The other chimera, DNA-U, was composed entirely of deoxyribonucleotides except for the deoxythymidylates, which were replaced with ribouridylates. In the DNA with a modified base-type, DNA-dU, deoxythymidylates were replaced with deoxyuridylates. Along with the VGF 40mer RNA, these four polynucleotides were assayed for polyadenylation in the presence of 60 and 150 mM NaCl (Figure 7A), and for stable VP55 binding (Figure 7B). The DNA sequence equivalent to that of the VGF 40mer RNA was not recognized by VP55. Like the DNA, the RNA-dT chimera was almost completely refractory to polyadenylation by, or stable interaction with, VP55. Conversely, the DNA-U chimera, like the RNA, could be polyadenylated in the presence of 60 and 150 mM NaCl, and could form a stable complex with VP55. These results show that the recognition of RNA by VP55 is mediated solely via the chemical differences between ribouridylates and deoxythymidylates, and that VP55 does not discriminate between 2′-OH and 2′-H groups on the sugar moieties of neighboring adenylate, cytidylate and guanylate bases. The DNA-dU polynucleotide was then tested to determine the importance of the 2′-OH on the sugars associated with ribouridylates. Since DNA-dU was polyadenylated less efficiently than the DNA-U chimera, it is concluded that the 2′-OH contributes to ribouridylate recognition. Since it was polyadenylated more efficiently than the DNA, it would appear that the unmethylated pyrimidine ring contributes also.

The point of transition to non-processive polyadenylation by VP55 is determined by the number of nucleotides between the growing 3′ end of the oligo(A) tail and a stable VP55 binding site within the RNA

The transition from processive to non-processive extension by VP55 is regulated by the net length of the growing oligo(A) tail (Gershon and Moss, 1992). To monitor oligo(A) tail length, VP55 could use, as a reference point, either the 5′ end of the tail or a stable (uridylate-containing) VP55 binding site within the RNA. If the reference point is the latter, then VP55 should include the length of a 3′ non-binding tract of RNA when 'measuring' the oligo(A) tail. Such a tract could comprise poly(C) since, like poly(A), poly(C) is refractory to stable VP55 binding (above). Two RNAs were therefore synthesized, comprising the VGF 40mer sequence with additional preformed 3′ oligo(C) tails of 5 or 15 nt (40mer-C_5 and 40mer-C_{15}). The two RNAs were assayed for polyadenylation, along with the parental RNA, the VGF 40mer, in the presence of 60 mM NaCl (Figure 8A). Each of the three RNAs was extended to the same overall length of ~20−25 nt beyond the 3′ end of the VGF 40mer sequence, in a rapid burst of polyadenylation, before the transition to slow, non-processive extension. This result favors the alternative that VP55 uses, as a reference point, a stable binding site within the RNA, rather than the 5′ end of the growing oligo(A) tail.

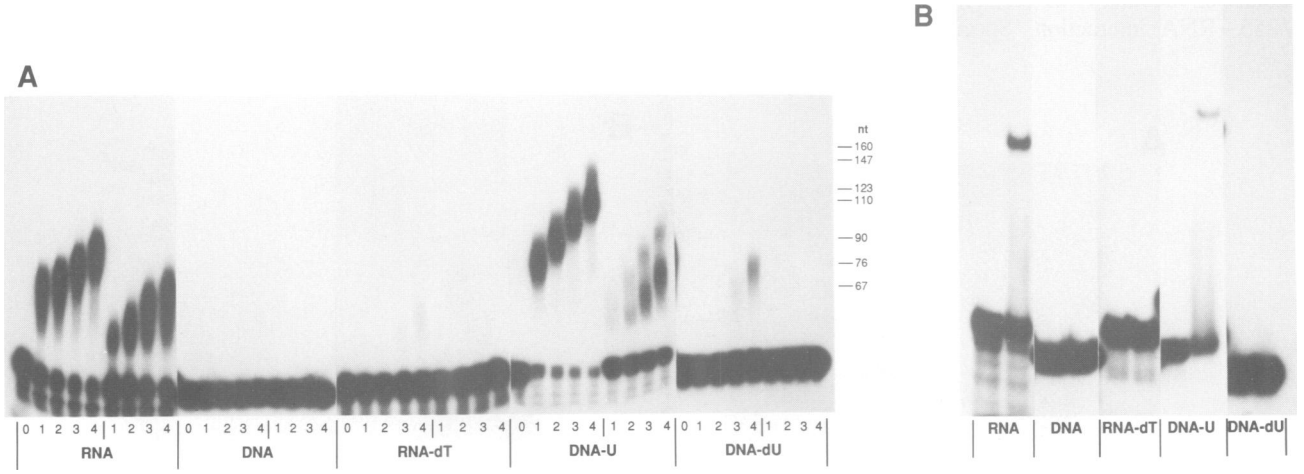

Fig. 7. Investigation of features of ribouridylates that mediate polynucleotide recognition and polyadenylation by VP55. The following polynucleotides, which were equivalent in sequence, were tested: RNA, the VGF 40mer; DNA, the DNA version of the VGF 40mer; RNA-dT, the VGF 40mer in which uridylates were replaced by deoxythymidylates; DNA-U and DNA-dU, the DNA version of the VGF 40mer in which deoxythymidylates were replaced by ribouridylates or deoxyuridylates respectively. (**A**) Polyadenylation of the five polynucleotides. Tracks 0 to 4 are labeled as for Figure 3. For each polynucleotide, the left- and right-most assays were performed in the presence of 60 and 150 mM NaCl, respectively. (**B**) EMSAs of the five polynucleotides. The first and second tracks for each polynucleotide represent incubations in the absence and presence, respectively, of VP55. The mobilities of the VP55−RNA and VP55−DNA-U complexes appear slightly different since they were resolved on separate gels.

VP55 translocates with respect to the RNA primer whilst adding nucleotides to the growing 3′ end

Whether VP55 translocates with respect to the RNA primer whilst adding nucleotides to the growing 3′ end was addressed here. Translocation would predict that, if VP55 continuously synthesizes a stable binding site, viz. oligo(U), it would be unable to terminate processive synthesis after the addition of 30–35 nucleotides. It is known that VP55 can add uridylates, as well as cytidylates and guanylates, to an RNA primer in the presence of $MnCl_2$ (Gershon and Moss, 1992). Thus, if the model is correct and translocation occurs during addition, the processive synthesis of oligo(A), oligo(C) and oligo(G) tails should terminate, but the processive synthesis of poly(U) tails should not. The VGF 40mer RNA was therefore polyadenylated, uridylated, cytidylated and guanylated, using a longer assay time than employed previously (Figure 8B). It was found that uridylation can indeed occur rapidly and processively, and poly(U) tails greater than 400 nt in length were synthesized, in a manner reminiscent of poly(A) tail synthesis in the presence of the processivity factor VP39. By contrast, the rapid, processive synthesis of poly(A) and poly(G) by VP55 terminated during the assays, and poly(C) tail formation proceeded only very slowly. The termination of rapid, processive poly(A) synthesis is more clearly visible in experiments in which more timepoints are taken during the first minute of the assay (Gershon and Moss, 1992). The experiment of Figure 8B strongly indicates that VP55 indeed translocates with respect to the RNA during nucleotide addition, and can detect the composition of the newly synthesized RNA 3′ end.

Discussion

We have analyzed the RNA characteristics required for a processive burst of polyadenylation by VP55, the catalytic subunit of vaccinia virus poly(A) polymerase, and for stable VP55–RNA interaction. Specific RNA sequences were found not to be required for the burst of polyadenylation by VP55. However, at physiological salt concentrations, the presence of uridylates within the RNA was both necessary and sufficient. Uridylate density appeared to be critical for a processive burst, with a uridylate content of 33% or greater ensuring processive polyadenylation at 150 mM NaCl (the physiological salt concentration) and stable VP55 binding. At lower salt concentrations fewer uridylates were sufficient, and clustering of small numbers of uridylates increased the stability of the RNA–VP55 interaction. For nearly all the 40mers tested, stable VP55 binding in the EMSA, as determined by the abundance of the observed complexes, corresponded with the efficacy of the RNA as a primer for a processive burst of polyadenylation in the presence of 60 mM NaCl. VP55 shows no significant regions of homology with RNA binding proteins known to interact with stretches of uridylates, such as the hnRNP C proteins (Swanson and Dreyfuss, 1988; Wilusz and Shenk, 1990), the 64 kDa subunit of CstF (Takagaki *et al.*, 1992), *Sxl* (Bell *et al.*, 1988; Horabin and Schedl, 1993), U2AF (Zamore *et al.*, 1992) and PTB (Gil *et al.*, 1991; Patton *et al.*, 1991). Notably, the RNP consensus sequence/recognition motif (Bandzulius *et al.*, 1989; Query *et al.*, 1989), which is present in each of these proteins except PTB, could not be found in the sequence of VP55. The absence of sequence-specific RNA binding by VP55 would account for our inability to define the VP55 footprint size by RNase protection (data not shown).

Ribouridylate recognition by VP55 was found to be mediated solely via the chemical differences between ribouridylates and deoxythymidylates, and VP55 was found not to discriminate between 2′-OH and 2′-H groups on the sugar moieties of neighboring adenylate, cytidylate and guanylate bases. By substituting ribouridylates with deoxyuridylates, it was found that ribouridylate recognition includes that of its 2′-OH group. Interaction with this group in addition to the uracil moiety may help VP55 orientate the bound RNA. The absence of a 5-methyl group on the

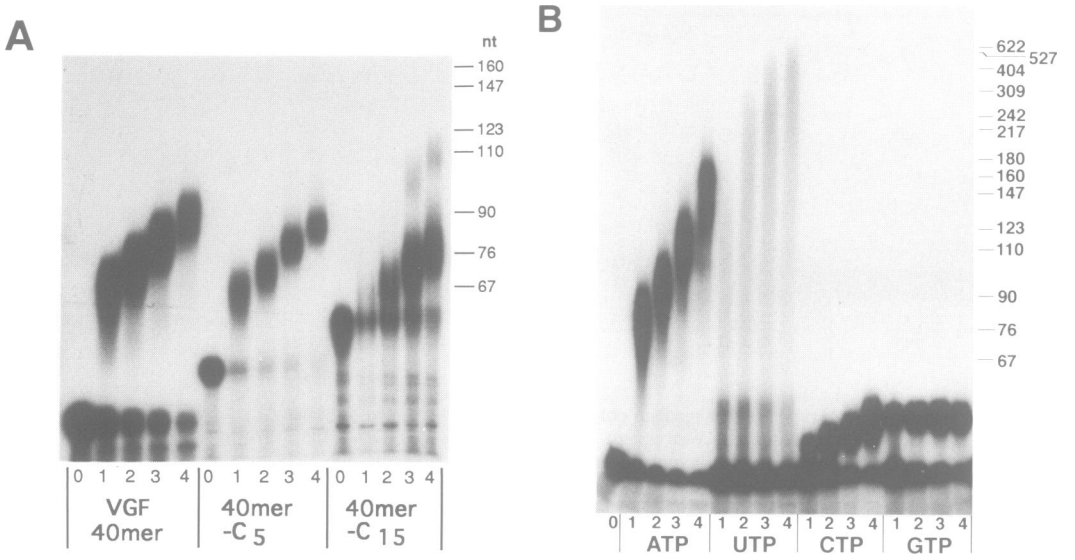

Fig. 8. Effects of homopolymer tracts at RNA 3′ ends. **(A)** Polyadenylation of VGF 40mer and two related RNAs, 40mer-C_5 and 40mer-C_{15}, which contain pre-formed oligo(C) tails of 5 and 15 nt, respectively. Tracks 0–4 are labeled as for Figure 3. **(B)** Time-course assays of polyadenylation, uridylation, cytidylation and guanylation of the VGF 40mer RNA, designated ATP, UTP, CTP and GTP, respectively. Tracks 0 to 4, products sampled at 0, 1, 5, 20 and 80 min. The NaCl concentration was 60 mM.

pyrimidine ring was also detected by VP55 and, since poly(C) was not recognized, a 4-amino substitution of the ring appears to be inhibitory. In addition to its unique chemical structure and pattern of potential hydrogen bonds, some other feature of the unsubstituted uracil moiety may contribute to its specific interaction with VP55. Thus, amongst the commonly occurring bases, uracil is unique in its possession of a rigid, planar structure. The substitution of methyl and amino groups would reduce this planarity which may, in turn, inhibit its insertion into a 'pocket' in the VP55 structure, or its stacking with aromatic amino acids. The ability of poly(G) to be polyadenylated semi-processively at low salt concentrations may indicate a specific role for a carbonyl group attached to a 6-membered ring. Such a carbonyl is positioned similarly in guanine and uracil.

Although we show that more than one molecule of VP55 can bind to a 110 nt RNA representing a non-polyadenylated vaccinia mRNA 3′ end *in vitro*, there is strong evidence that VP55 molecules bound at positions away from the RNA 3′ end are not active in polyadenylation. This comes from our unpublished studies with VGF 40mer to 110mer RNAs possessing preformed oligo(A) tails greater than 35 nt in length. Although these RNAs were polyadenylated only non-processively by VP55, as would be expected, stable complexes, which normally correlate with processive addition, were observed in the EMSA. Since VP55 cannot bind stably to poly(A), these stable but inactive complexes would have resulted from VP55 binding to internal RNA segments. Other evidence that non-productive VP55−RNA complexes can form *in vitro* has been presented (Gershon and Moss, 1993). If VP55 must interact with RNA at, or very close to, its 3′ end to effect nucleotidyl transfer, as these data suggest, then the catalytically active species of VP55 would need to translocate during nucleotide addition. The processive synthesis of long poly(U) tails supports the prediction of VP55 translocation. That experiment additionally indicated that translocation does not require VP39 or ATP hydrolysis, though a role for NTP hydrolysis cannot be discounted.

A model is presented (Figure 9) to account for the termination of rapid, processive polyadenylation after oligo(A) tails of 30−35 adenylates have been synthesized. In this model, the polyadenylating species of VP55 interacts with RNA 30−35 nt distal to the 3′ end via a site which interacts stably with uridylates, but only unstably with oligo(A). Processive polyadenylation would be accompanied by the translocation by VP55 until the 'uridylate-recognition site' has come into contact with newly synthesized oligo(A). At this point processive polyadenylation would cease. As well as accounting for the termination of processive polyadenylation and the translocation of VP55 during the processive polyadenylation burst, the model accounts for the inability of VP55 to polyadenylate poly(A) RNAs less than 34 nt in length (Gershon and Moss, 1992) or to stably bind VGF 3′ end RNAs less than 31−40 nt in length. Interactions required for the semi-processive polyadenylation of poly(G) may also be mediated through the 'uridylate-recognition site', since poly(G) primers shorter than 40 nt in length are refractory to polyadenylation (data not shown). Since the uridylate-recognition site cannot interact with poly(C), the model would account for the early termination of processive polyadenylation when short pre-formed oligo(C) tracts are placed 3′ to a uridylate-containing region (Figure 8A), and the inability of VP55 to polyadenylate RNAs containing longer (40 nt) 3′ oligo(C) tracts (data not shown). The polyadenylation of RNAs with a low uridylate content often proceeds in intermediate-sized processive jumps interspersed with non-processive addition (Figure 6B and D). This may result from the translocation of the small numbers of uridylates into and out of the uridylate-recognition site.

Although the model emphasizes the presence of a 'uridylate-recognition site', which interacts with a local region of RNA 30−35 nt distal to the extreme 3′ end, we do not exclude the possibility that VP55 also contacts the extreme 3′ end itself, to provide additional binding stability and to effect nucleotide transfer, along with regions between −30 and the extreme 3′ end. Consistent with this would be the following observations. (i) Short oligo(U) RNAs (between 5 and 20 nt in length) cannot bind stably to VP55 (data not shown). Furthermore, a 110-fold molar excess of the trinucleotide UUU is unable to compete with the VGF 40mer RNA for stable VP55 binding and polyadenylation (data not shown). (ii) VP55 cannot polyadenylate RNAs possessing a 14 or 15 nt 3′ oligo(G) tract (Figure 8B; data not shown), indicating that VP55 can make some type of discriminatory contact with regions 14 nt distal to the RNA 3′ end. The detailed mechanics of VP55−RNA interaction await further studies.

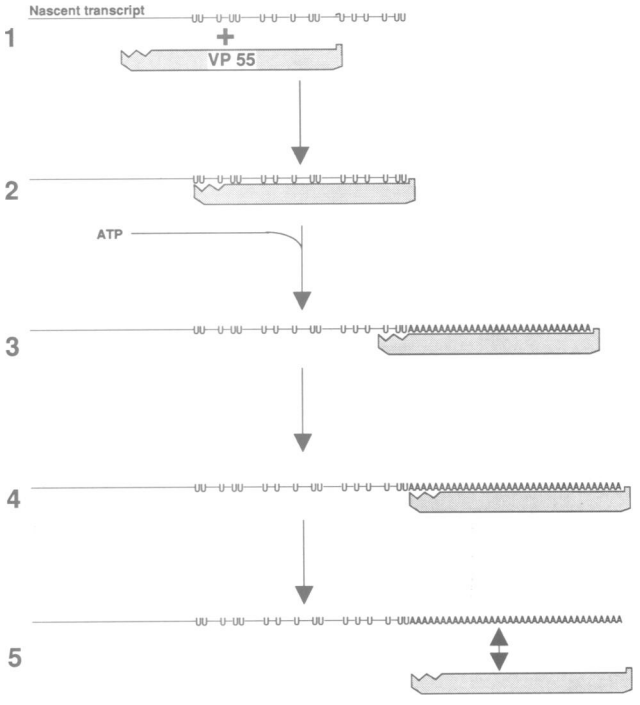

Fig. 9. Proposed role of the uridylate-dependent VP55−RNA interaction in the termination of processive polyadenylation after extension of the oligo(A) tail to a net length of 30−35 nt. **Steps 1** and **2**: uridylate-dependent interaction of VP55 with the 3′ end of the nascent transcript. Stable interaction with uridylates occurs primarily via a site on VP55 that interacts with RNA 30−35 nt from its 3′ end (wavy line). This site interacts only unstably with adenylates. Although VP55 may possess no intrinsic ability to recognize the RNA 3′ end *in vitro*, VP55 molecules bound at other positions would be unable to polyadenylate the RNA (see text). **Step 3**: in the presence of ATP, VP55 processively adds 30 adenylates to the RNA 3′ end whilst translocating and interacting with uridylates in the RNA. **Steps 4** and **5**: on extension of the oligo(A) tail to a net length of 35 nt, newly synthesized oligo(A) has translocated to the 'uridylate-recognition site' of VP55, causing dissociation from the RNA and a cessation of processive polyadenylation.

As discussed above, the most likely role for uridylates would be in the binding of VP55 to RNA. Other roles, such as providing a template for adenylate addition, seem much less likely, particularly since the polyuridylation reaction also requires a uridylate-containing primer. Indeed, poly(A) is refractory to polyuridylation (data not shown). The vaccinia early gene transcriptional termination signal, UUUUUNU, is usually found 20–50 nt from the 3′ ends of vaccinia early transcripts (Yuen and Moss, 1987). Although this uridylate cluster may contribute to VP55–RNA interaction, its specific sequence is not required for polyadenylation. The UUUUUNU signal does, however, specify the site of poly(A) tail addition through its recognition by the transcriptional termination apparatus, just as the site of poly(A) tail addition to mammalian mRNAs is specified by the position of the RNA cleavage signal, AAUAAA. The newly formed 3′ ends of vaccinia transcripts could then be targeted for polyadenylation by spatial considerations, such as the physical juxtaposition of VP55 with the vaccinia transcription apparatus. Despite the possible coupling of polyadenylation to the formation of the RNA 3′ end, the vaccinia polyadenylation system appears profoundly different from that of the mammalian nucleus. Notably the nuclear system is RNA sequence specific and coupled to RNA cleavage *in vivo* (Bienroth *et al.*, 1993), and utilizes a PAP which is not homologous to VP55 and is not, by itself, processive.

The low primer sequence specificity of VP55, especially at low salt, the processive addition of relatively short poly(A) tracts and the ability to catalyze addition of uridylates, cytidylates and guanylates indicate its potential usefulness for the *in vitro* tailing of RNA.

Materials and methods

Materials

Unless otherwise stated, the previously described preparation of highly purified recombinant VP55 [poly(A)-Sepharose column fraction 45 (Gershon and Moss, 1992)] was used here. In this preparation, VP55 had been purified to apparent homogeneity.

RNAs

The VGF 60mer RNA (Gershon and Moss, 1993), and the VGF 30mer and 110mer RNAs (Gershon and Moss, 1992), have been described. Phosphoramidite chemistries were used to synthesize those RNAs referred to as the VGF 40mer, ACGmer, CGUmer, the 'Umers', the 'tetra-Umers' and the alternating copolymers (Table I), as well as the RNAs U_{40}, C_{40}, G_{40}, and the DNA, the DNA containing deoxyuridylates and the two DNA–RNA chimeras based on the VGF 40mer sequence. The RNAs referred to as 'scramblemers', ACUmer, AGUmer (Table I), the VGF 50mer, 80mer and 100mer, 40mer-C_5 and 40mer-C_{15}, were synthesized enzymatically using methods described previously (Gershon and Moss, 1992, 1993). Each of the RNAs was purified by polyacrylamide gel electrophoresis and 5′ end-labeled as described by Gershon and Moss (1992). For the purification of G_{40}, gels contained 40% formamide.

Assays

Unless otherwise stated, RNA was present in 2- to 4-fold molar excess over VP55 in all assays. Polyadenylation assays were performed as described previously (Gershon and Moss, 1992). Incubation conditions for the electrophoretic mobility shift assay (EMSA) were similar to those used in the polyadenylation assay except for the presence of glycerol and the absence of ATP. Thus, 5′ end-labeled RNA probes were diluted 20-fold into 50 mM Tris–HCl pH 9.0, 5 mM DTT, 10% glycerol. After addition of VP55 followed by 5–6 min incubation at 22°C, complexes were resolved in 8% (unless otherwise noted) polyacrylamide gels buffered with 68 mM Tris, 23 mM borate, 1 mM EDTA, pH 8.0. Samples were electrophoresed at 18 V/cm for 1–2 h.

Acknowledgements

We thank Jerry Sisler for the synthesis of DNA oligonucleotides.

References

Bandzulius,R.J., Swanson,M.S. and Dreyfuss,G. (1989) *Genes Dev.*, **3**, 431–437.
Bell,L.R., Maine,E.M., Schedl,P. and Cline,T.W. (1988) *Cell*, **55**, 1037–1046.
Bienroth,S., Keller,W. and Wahle,E. (1993) *EMBO J.*, **12**, 585–594.
Gershon,P.D. and Moss,B. (1992) *Genes Dev.*, **6**, 1575–1586.
Gershon,P.D. and Moss,B. (1993) *J. Biol. Chem.*, **268**, 2203–2210.
Gershon,P.D., Ahn,B.-Y., Garfield,M. and Moss,B. (1991) *Cell*, **66**, 1269–1278.
Gil,A., Sharp,P.A., Jamison,S.F. and Garcia-Blanco,M.A. (1991) *Genes Dev.*, **5**, 1224–1236.
Horabin,J.I. and Schedl,P. (1993) *Mol. Cell. Biol.*, **13**, 1408–1414.
Kates,J. and Beeson,J. (1970) *J. Mol. Biol.*, **50**, 19–33.
Moss,B., Rosenblum,E.N. and Gershowitz,A. (1975) *J. Biol. Chem.*, **250**, 4722–4729.
Moss,B., Ahn,B.-Y., Amegadzie,B., Gershon,P.D. and Keck,J.G. (1991) *J. Biol. Chem.*, **266**, 1355–1358.
Nevins,J.R. and Joklik,W.K. (1975) *Virology*, **63**, 1–14.
Nevins,J.R. and Joklik,W.K. (1977) *J. Biol. Chem.*, **252**, 6939–6947.
Patton,J.G., Mayer,S.A., Tempst,P. and Nadal-Ginard,B. (1991) *Genes Dev.*, **5**, 1237–1251.
Query,C.C., Bentley,R.C. and Keene,J.D. (1989) *Cell*, **57**, 89–101.
Rohrmann,G., Yuen,L. and Moss,B. (1986) *Cell*, **46**, 1029–1035.
Schnierle,B.S., Gershon,P.D. and Moss,B. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 2897–2901.
Swanson,M.S. and Dreyfuss,G. (1988) *Mol. Cell. Biol.*, **8**, 2237–2241.
Takagaki,Y., MacDonald,C.C., Shenk,T. and Manley,J.L. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 1403–1407.
Wei,C.M. and Moss,B. (1975) *Proc. Natl Acad. Sci. USA*, **72**, 318–322.
Wilusz,J. and Shenk,T. (1990) *Mol. Cell. Biol.*, **10**, 6397–6407.
Yuen,L. and Moss,B. (1986) *J. Virol.*, **60**, 320–323.
Yuen,L. and Moss,B. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 6417–6421.
Zamore,P.D., Patton,J.G. and Green,M.R. (1992) *Nature*, **355**, 609–614.

Received on June 24, 1993; revised on August 9, 1993