Proc. Natl. Acad. Sci. USA
Vol. 77, No. 5, pp. 2424–2428, May 1980
Biochemistry

Structure of genes for virus-associated RNA_I and RNA_{II} of adenovirus type 2

(DNA sequence/RNA sequence/secondary structures/RNA polymerase III recognition)

GÖRAN AKUSJÄRVI*, MICHAEL B. MATHEWS†, PER ANDERSSON*, BJÖRN VENNSTRÖM*‡, AND ULF PETTERSSON*

*Department of Microbiology, The Biomedical Center, University of Uppsala, Box 581, S-751 23 Uppsala, Sweden; and †Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, New York 11724

Communicated by J. D. Watson, January 11, 1980

ABSTRACT **A DNA sequence, 552 base pairs in length, encoding the two "virus-associated" (VA) RNAs of adenovirus type 2 is presented. Comparison of the oligonucleotide maps of VA RNA_I and VA RNA_{II} with the established sequence permits identification of the genes for these RNAs. VA RNA_I is 157–160 nucleotides long and VA RNA_{II} 158–163 nucleotides long, depending on the exact length of their heterogeneous 3′ ends. The genes are separated by a spacer of about 98 nucleotides. The RNAs exhibit scattered regions of primary sequence homology and can adopt secondary structures which resemble each other closely in their configuration and stability. VA RNA_{II} is also capable of assuming a different configuration that is energetically more favorable. The data suggest that the two RNA genes may have arisen by duplication of an ancestral gene and that the folding of the RNA chain may be of importance for the function of VA RNAs. Hypothetical RNA polymerase III recognition sequences and the coding potential of the region are discussed.**

The adenovirus genome contains genes for two low molecular weight RNAs, the "virus-associated" RNAs, VA RNA_I and VA RNA_{II} (1–4). The presence of a prominent class of low molecular weight RNA (VA RNA) in adenovirus-infected cells was reported in 1966 (5), but its function still remains obscure. The VA RNAs are of particular interest because they are unique among adenovirus transcripts in being synthesized by RNA polymerase III (2, 3, 6). In adenovirus-infected cells the two RNAs appear with different kinetics, synthesis of VA RNA_I becoming increasingly predominant at late times after infection (2). The two species of VA RNA are transcribed from the viral r-strand and their genes have been located by hybridization to restriction enzyme fragments. The genes map close to positions 29–30 on the adenovirus type 2 genome (Fig. 1*A*) (1, 2, 10) and are separated by a short DNA spacer (4).

The VA RNAs do not undergo posttranscriptional processing at their 5′ termini (2, 3, 6, 11), and therefore promoters for RNA polymerase III are presumed to be located in the vicinity of their structural genes. Recently Pan *et al.* (12) and Celma *et al.* (13) reported a nucleotide sequence of adenovirus-2 DNA that includes the gene for VA RNA_I. In the present study we have determined the nucleotide sequence of the region of the viral genome that includes the genes for VA RNA_I and VA RNA_{II} and the connecting spacer. Computer-aided model building shows that the two VA RNAs can form almost identical secondary structures although their primary sequences have only short regions of homology. An alternative and more stable secondary structure can be drawn for the VA RNA_{II} species.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

RESULTS

DNA Sequence Analysis of the VA RNA Genes and the Spacer Region. Hybridization experiments have mapped the VA RNA genes between positions 28.8 and 30.0 on the viral DNA (1, 2, 4), and the DNA sequence of the VA RNA_I gene and its immediate surroundings (map coordinates 28.6–29.4, approximately) has been reported (12, 13). We have confirmed most of the published sequence data and extended the DNA sequence in a rightwards direction through the VA RNA_{II} gene to map position 30.1, approximately, using the strategy summarized in Fig. 1*B*. The DNA sequence established and shown in Fig. 2*A* includes a stretch of 46 nucleotides at the extreme left of the region that was determined by Pan *et al.* (12) but has not been studied by us. For nucleotides 47–287, our data agree with the previously determined sequence (12, 13) at all positions except one: at nucleotide 200 our sequence reads C-C-G-C-C-C-G, compared with C-C-G-C-C-C-G in the sequence of Celma *et al.* (13). The presence of an extra C at this position is consistent with the ribonuclease T1 fingerprint of VA RNA_I, which clearly includes one copy of the tetranucleotide C-C-C-G (1, 14). This oligonucleotide cannot be accommodated in the published sequence (13), and we therefore assume that the missing C residue was overlooked because of "compressions" in the sequencing gels. Such compressions were experienced in several positions within the spacer, but these regions, which are presumably caused by secondary structure in the DNA, were resolved into multiple bands when the sequence of the complementary strand was determined. It should be noted that the entire sequence from nucleotide 168 to 551 established in this study has been determined from both of the complementary DNA strands (Fig. 1*B*). Furthermore, two-thirds of the sequence can be matched with the pancreatic and T1-ribonuclease digestion products from VA RNA_I and VA RNA_{II}, allowing a considerable measure of confidence in its accuracy.

Fingerprint Analysis of VA RNA_{II}. To define the VA RNA_{II} sequence precisely and correlate it with the DNA sequence, we have carried out analyses of the RNA structure by digestion of purified ^{32}P-labeled VA RNA_{II} with ribonuclease T1 and pancreatic ribonuclease. The resulting oligonucleotides were separated in two dimensions and further characterized by redigestion with additional enzymes (15). The results obtained with pancreatic ribonuclease (Table 1), together with data on the ribonuclease T1 digest that is already published (4), position the VA RNA_{II} gene in the region of nucleotides 350–510, approximately. All of the pancreatic and T1-ribonuclease diges-

Abbreviation: VA, virus-associated.

‡ Present address: Department of Microbiology, University of California, San Francisco, CA 94143.

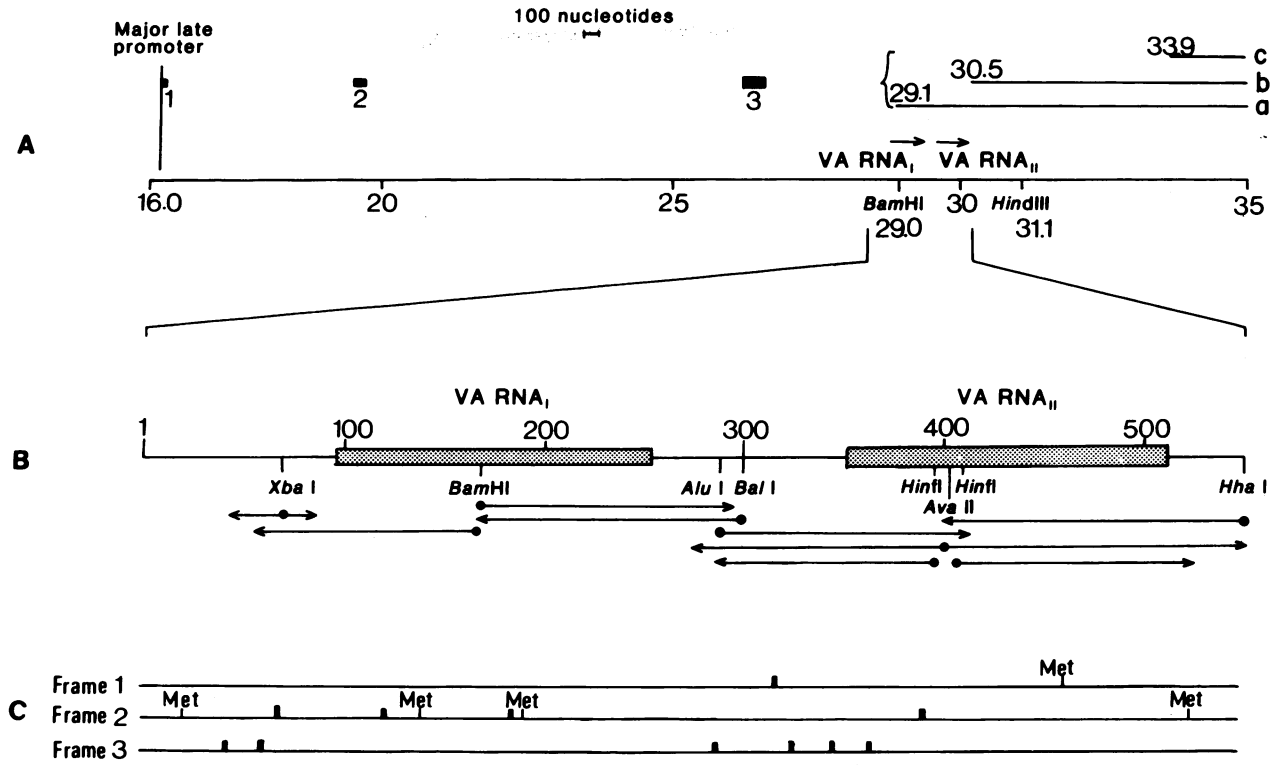

FIG. 1. Schematic representations of the VA RNA gene region. (*A*) Map of the VA RNA gene region on the adenovirus genome. The arrows indicate the location and direction of transcription of the VA RNAs. The location of the major late promoter at position 16.3 is indicated (7). The blocks numbered 1, 2, and 3 denote the three common leader segments of the late mRNAs, and the coordinates 29.1, 30.5, and 33.9 indicate the 5′ borders of the bodies of three mRNAs mapping in this region (8). (*B*) Enlargement of the VA RNA gene region around positions 29–30. The cleavage sites indicated were used for DNA sequence analysis by the method of Maxam and Gilbert (9). The closed circles represent the sites of 5′-terminal labeling by [γ-^{32}P]ATP and polynucleotide kinase; the arrows show the extent and direction of the sequences determined. (*C*) Sequence that covers the region in *B* shown in the three possible reading frames. The locations of termination codons (▮) and of AUG triplets (Met) are shown.

tion products are found within this span, and their molar yields are in excellent agreement with expectations from the DNA sequence.

The 5′ terminus of VA RNA$_{II}$ was identified by elution of the oligonucleotides from a pancreatic ribonuclease fingerprint, followed by digestion with ribonuclease T2 and one-dimensional separation of the products (16, 17). The structure of the oligonucleotide carrying a polyphosphorylated 5′ end was elucidated by determination of the radioactivity in the resulting mononucleotides. It yielded pGp, Gp, and Cp in a molar ratio of approximately 1:1:1. Because the oligonucleotide was produced by digestion with pancreatic ribonuclease, its structure must be pG-G-Cp, which is compatible with our previous findings (17) and matches the DNA sequence between nucleotides 353 and 355 (Fig. 2*A*).

At the 3′ end, the most distal oligonucleotide that can be matched with the DNA sequence is G-A-G-C, generated by pancreatic ribonuclease digestion (nucleotides 503–516 in Fig. 2). We have not detected in significant yield any ribonuclease T1 product that lacks a G residue and that must, therefore, originate from the 3′ end of the RNA. The undecanucleotide C_4-U_6-G (nucleotides 506–516 in Fig. 2*A*) has not been detected in ribonuclease T1 fingerprints employing either DEAE-paper electrophoresis (4) or thin-layer homochromatography for the second dimension, implying that the 3′ end of VA RNA$_{II}$ is located to the left of nucleotide 516. Other studies of the termination sites for RNA polymerase III (reviewed in ref. 18) indicate that the enzyme usually terminates transcription within a cluster of T residues, and we conclude that the 3′ end of VA RNA$_{II}$ lies in the run of six T residues between nucleotides 510 and 515 (Fig. 2*A*). By analogy with VA RNA$_I$ (13), the 3′ end of VA RNA$_{II}$ is likely to be heterogeneous, and we have not attempted to establish its position more precisely.

Secondary Structures of the RNAs. Direct experiments (ref. 14; H. J. Monstein and L. Philipson, personal communication) and results from computer-assisted model building (19) have shown that VA RNA$_I$ has a high degree of secondary structure. This finding prompted a comparison between the possible secondary structures of the two RNAs, using a computer program based on the energy rules suggested by Tinoco *et al.* (20). The analysis showed that both VA RNAs can assume stable secondary structures. VA RNA$_I$ can form one stable structure, similar to that reported previously (19), with a free energy of formation of −108 kcal/mol (Fig. 3*A*) (1 cal = 4.184J), whereas VA RNA$_{II}$ can adopt two distinct conformations that have slightly different stabilities, with free energies of −125.8 and −108 kcal/mol (Fig. 3 *B* and *C*, respectively). One of the structures for VA RNA$_{II}$ has the same free energy as the structure formed by VA RNA$_I$, to which it is similar in overall shape, though the number of unpaired bases in the corresponding loops differs slightly. In the other, more stable, VA RNA$_{II}$ structure (Fig. 3*C*) the pairing in the center of the molecule is rearranged while the stem that contains the 5′ and 3′ ends of the RNA is conserved. *In vivo* an equilibrium between the two structures might be expected.

DISCUSSION

The nucleotide sequence of adenovirus-2 DNA that includes the genes for VA RNA$_I$ and VA RNA$_{II}$ has been determined. A comparison of the DNA sequence with the pancreatic and T1-ribonuclease oligonucleotides from the two RNAs establishes that VA RNA$_I$ has a chain length of 157–160 nucleotides and

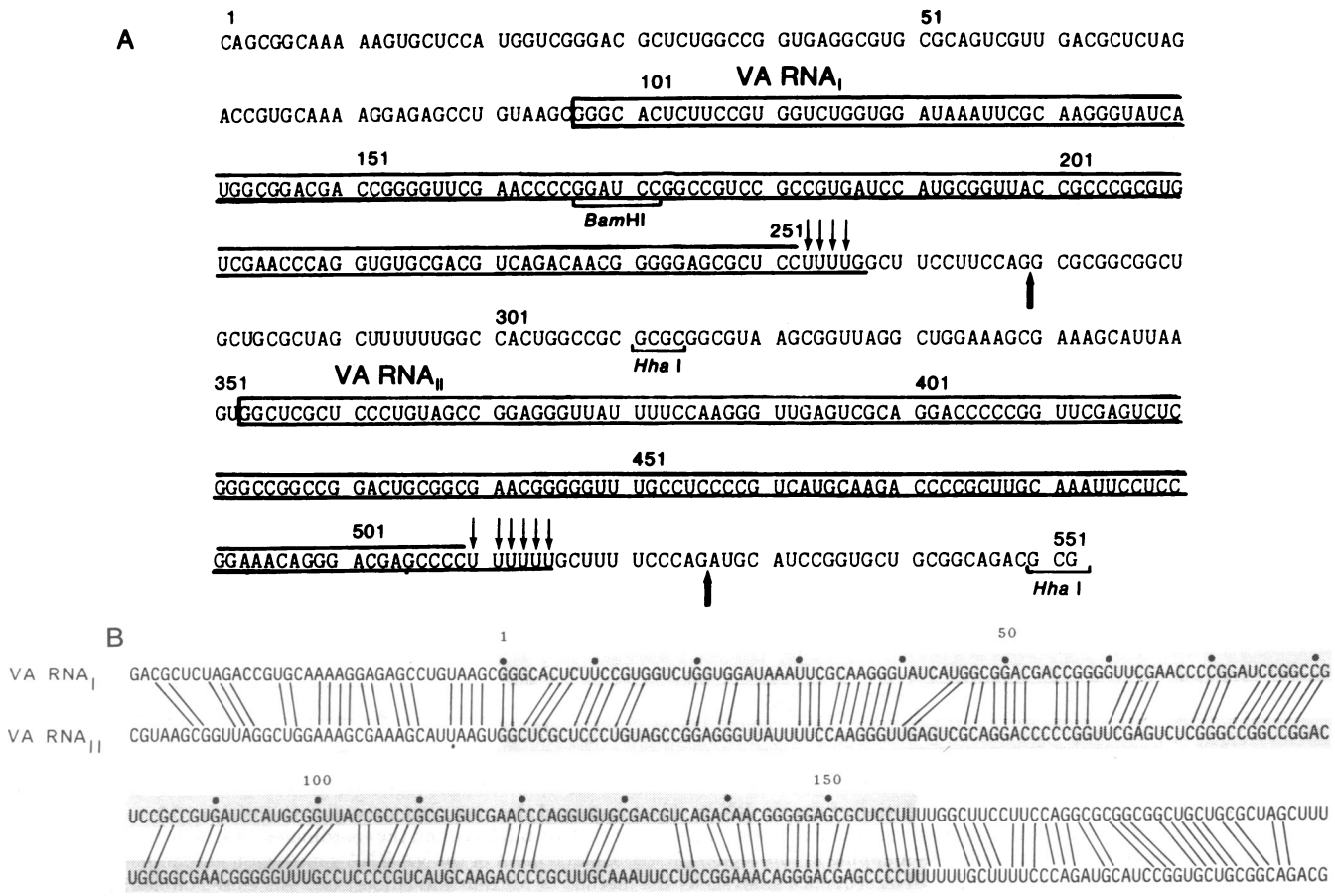

FIG. 2. Sequence of nucleotides in the VA RNA gene region of adenovirus-2. (*A*) Sequence of the VA RNA genes, given as the RNA sequence of the l-strand. Nucleotides 1–46 have not been determined in this study and are taken from Pan *et al.* (12). The genes for the VA RNAs are shown boxed. Thin arrows indicate the positions of the heterogenous 3′ ends of the VA RNAs. The locations of two hypothetical splice points (see *Discussion*) are shown with thick arrows. (*B*) Comparison of primary sequences of VA RNA$_I$ and VA RNA$_{II}$ (shaded) and of their flanking sequences, showing homologous regions where at least two consecutive bases can be matched.

VA RNA$_{II}$ has a length of 158–163 nucleotides. The range derives from uncertainty in the precise location of the 3′ termini, which lie in runs of T residues. Stable hairpin-like structures can be drawn for both species. Taking the 3′ terminus of RNA$_I$ to lie at nucleotide 254, the separation between the two VA RNA genes is 98 base pairs. The disparity between this value and the estimate of 75 base pairs for the spacer made on the basis of less direct results (4) is attributable to the existence of a previously disregarded 21-base-pair DNA fragment (nucleotides 394–414) released from the VA RNA$_{II}$ gene by digestion with the endonuclease *Hin*fI.

Two lines of evidence suggest that the VA RNAs are transcribed from independent promoters: (*i*) polyphosphorylated 5′ termini have been detected on both species synthesized *in vivo* and *in vitro* (2,6); (*ii*) transcription of fragmented adenovirus-2 DNA *in vitro* can give rise to VA RNA$_{II}$ without VA RNA$_I$ (21). Specifically, viral DNA cleaved with the endonucleases *Bam*HI or *Hha* I, which cleave in the center of the VA RNA$_I$ gene and close to its 3′ end, respectively, fail to direct synthesis of VA RNA$_I$ but still generate full-length VA RNA$_{II}$. The DNA sequence (Fig. 2*A*) shows that the VA RNA$_{II}$ gene is included in an endonuclease *Hha* I-derived fragment that extends 40 nucleotides upstream from the 5′ end of the RNA and about 37 nucleotides downstream from its 3′ end. Thus it is possible to conclude that no more than 40 nucleotides upstream from the initiation site are needed for proper polymerase binding and initiation of VA RNA$_{II}$ transcription. In the case of a *Xenopus* tRNA gene, no more than 22 upstream nucleotides are required for transcription by RNA polymerase III (22).

Korn and Brown (18) have identified features of the sequences adjacent to the initiation point that are common to genes transcribed by RNA polymerase III. For their comparison they used the genes for ribosomal 5S RNA of several species, and adenovirus-2 VA RNA$_I$. Examination of the VA RNA$_{II}$ gene shows that it shares several of these common features. (*i*) The region immediately adjacent to the initiation site is rich in purine residues. Similar purine-rich sequences are also found upstream from RNA polymerase II promoters on the adenovirus-2 genome (7, 23), and this appears to be a feature shared by promoter regions for the two classes of eukaryotic RNA polymerases. (*ii*) A number of sequences occur as direct repeats, including the octamer G-C-G-C-G-G-C-G found between nucleotides 270–277 and 311–318. (*iii*) The oligonucleotides G-A-A-A-G, A-G-C-G-G, and G-C-C, corresponding to the canonical sequences A-A-A-A-G, A-G-A-A-G, and G-A-C (18), are present about 15, 25, and 40 bases before the 5′ end of the transcript. On the other hand, differences between the two VA RNA promoters seem to be biologically significant, being reflected in the kinetics of synthesis of the two species in infected HeLa cells (2) and in the ability of cell-free transcriptional systems from various sources to discriminate between them (21).

The conventional assumption that polymerase III promoters necessarily lie upstream from their structural sequences is undermined by the recent demonstration that the main control signals governing transcription of the *Xenopus* gene for 5S ribosomal RNA appear to be buried within the structural gene itself, upstream sequences playing a subsidiary role (24, 25). The

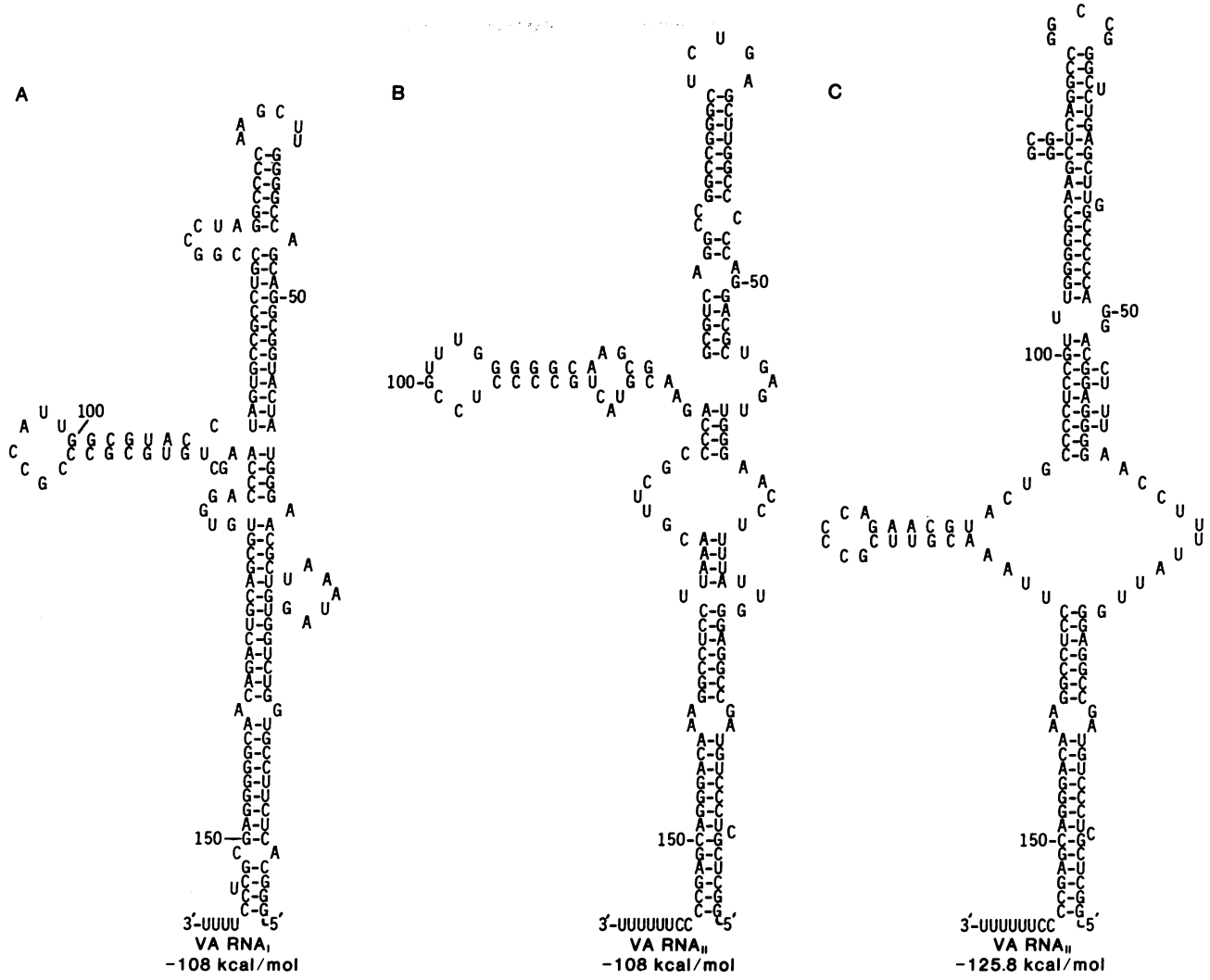

FIG. 3. Secondary structure of VA RNAs. The structure of VA RNA$_I$ (*A*) and two possible structures for VA RNA$_{II}$ (*B* and *C*) were derived by computer analysis.

importance of downstream sequences for the transcription of a tRNA gene has also been established (26). Whether the same is true of the VA RNA genes is not yet known, but few homologies are apparent between the VA RNAs and the central promoter-like region of 5S RNA. The effect of a small deletion (of nucleotides 71 and 72, Fig. 2*A*) upstream from the initiation site for VA RNA$_I$ in adenovirus-5 (27) clearly demonstrates that external sequences exert some influence. This deletion eliminates the synthesis of a minor form of VA RNA$_I$ that is initiated at the A residue (nucleotide 94) three residues before the normal start (28, 29) but does not affect production of the major species.

The exact position of the 3′ end of VA RNA$_{II}$ has not been established, but almost certainly it lies in the T cluster between positions 510 and 515. It remains to be proven that the 3′ ends of the VA RNAs are generated by termination of transcription rather than by processing of a large nuclear transcript as is the case for the 3′ ends of most late adenovirus mRNAs (30). An extended version of VA RNA$_I$, V200, has been observed both *in vivo* and *in vitro* (3, 21). The 3′ end of this RNA is also located within a cluster of T residues (nucleotides 292–297) and may result from transcriptional readthrough beyond the normal termination site for VA RNA$_I$. Whether V200 represents a precursor form of VA RNA$_I$ or merely a product of transcriptional error is presently unknown.

The sequences of the two VA RNA genes exhibit scattered homologies amounting to about 60% of their nucleotides in the coding regions and exceeding 70% in the flanking regions within 20 bases of the 5′ and 3′ termini (Fig. 2*B*). In view of the RNAs' similarity in length and possible secondary structure and their genes' close proximity to one another, it is tempting to speculate that the two VA RNA genes have arisen through duplication of a single ancestral gene. It should be pointed out however, that no homologous regions longer than seven nucleotides exist, corroborating earlier results indicating that the VA RNAs do not cross-hybridize (2, 10). The role of the two VA RNAs in the infectious cycle remains a matter of speculation (31), but their similar secondary structures suggest that the folding of the RNA chain is essential for their function.

Electron microscopic (8) and S1-nuclease mapping (S. Berget, personal communication) studies have demonstrated the presence of three late cytoplasmic mRNA species that have the tripartite leader spliced to an mRNA body in the vicinity of the VA RNA genes. Cell-free protein synthesis experiments (32) indicate that the shortest RNA (species c in Fig. 1*A*) encodes virion component IIIa. One of the longer RNAs, probably the a species of Fig. 1*A*, codes for a pair of polypeptides of 52,000 and 55,000 daltons, whereas the translational activity of the b species is uncertain. Examination of the DNA sequence reveals that nonsense codons block protein synthesis in all three reading

Table 1. Oligonucleotide analysis of VA RNA_{II}

Pancreatic ribonuclease products			
Spot no.	Structure deduced	Copies per molecule	
		Observed	Predicted
0	G_5-U	0–1	1
1	(G_4,A-G)U	0–1	1
2	(A-A-G,G_2)U	1	1
3	(A-G,G_2)A-C	1	1
4	(G,A-G)U	2	2
5	G-G-U	1	1
6	G-G-G-C	1	1
7	G-G-A-A-A-C	1	1
8	(A-G,G)A-C	1	1
9a	G-G-A-C	1	1
9b	(G,A-G)C	1	1
10	G-U	2–3	2
11	G-G-C	2	2
12	A-A-A-U	1	1
13	A-A-G-A-C	1	1
14	G-A-A-C	1	1
15	A-G-C	1	1
16	A-U	2	2
17	G-C	7–8	7
18	C	≈22	33
19	U	≈18	18–24

The molar yields of the oligonucleotides obtained by digestion of VA RNA_{II} with pancreatic ribonuclease are listed: observed yields were experimentally determined by estimation of the radioactivity in each spot (numbered according to M. B. Mathews and T. Grodzicker, unpublished); predicted yields are from the DNA sequence (Fig. 2*A*). Satisfactory resolution of both larger and smaller oligonucleotides required the preparation of fingerprints employing both thin-layer homochromatography and DEAE-paper electrophoresis as alternative second dimensions (15).

frames near map position 29.5 (Fig. 1*C*). This finding, and the existence of viable mutants with a deletion of nucleotides 71 and 72 (26) or a four-base insertion between nucleotides 167 and 168 (unpublished data), makes it unlikely that the sequence to the left of position 29.5 is translated. Bearing in mind the observations that translation of eukaryotic mRNAs generally starts at the initiation codon closest to the 5′ end of the RNA (33) and that the tripartite leader is free of these signals (34, 35), it is probable that translation of the a messenger species commences with the A-U-G triplet located at nucleotides 463–465 (coordinate 29.8). The rules governing the location of splice points have not yet been clearly established, but one useful generalization is that acceptor sites often have the structure Y-Y-N-Y-A-G (Y, pyrimidine nucleoside; N, any nucleoside) and follow a stretch of pyrimidines, especially U residues (36). These conditions are met at the 3′ end of the VA RNA_I gene, and the splice site between the tripartite leader and the body of the a message mapped at coordinate 29.1 could well be located at nucleotide 269 (Fig. 2*A*). It is interesting to note that a similar sequence arrangement occurs at the 3′ end of the VA RNA_{II} gene, allowing a hypothetical splice point at nucleotide 526 that might correspond to the leader–body junction for the b messenger species. If these hypothetical sites functioned *in vivo* during the maturation of adenovirus mRNAs, an economical organization of processing sites on the viral genome would emerge: the U-rich regions that terminate RNA transcription would also serve as recognition signals for the enzyme(s) that splice the tripartite leader to the bodies of two late mRNAs.

We thank E. Johnsen, M. Pettersson, S. Weirich, I. Wendel, and C. Thummel for skillful technical help and L. Richardson, M. Gustafson, and C. Sjoholm for excellent secretarial assistance. We also thank H. J. Monstein and L. Philipson for valuable discussions about VA RNA structures and V. Tate for comments on the manuscript. This investigation was supported by grants from the Swedish Cancer Society, the Swedish Medical Research Council, and the U.S. National Cancer Institute.

1. Mathews, M. B. (1975) *Cell* **6**, 223–229.
2. Söderlund, H., Pettersson, U., Vennström, B., Philipson, L. & Mathews, M. B. (1976) *Cell* **7**, 585–593.
3. Weinmann, R., Brendler, T. G., Raskas, H. J. & Roeder, R. G. (1976) *Cell* **7**, 557–566.
4. Mathews, M. B. & Pettersson, U. (1978) *J. Mol. Biol.* **119**, 293–328.
5. Reich, P. R., Baum, S. G., Rose, J. A., Rowe, W. P. & Weissman, S. M. (1966) *Proc. Natl. Acad. Sci. USA* **55**, 336–341.
6. Price, R. & Penman, S. (1972) *J. Virol.* **9**, 621–626.
7. Ziff, E. B. & Evans, R. M. (1978) *Cell* **15**, 1463–1475.
8. Chow, L. T. & Broker, T. (1978) *Cell* **15**, 497–510.
9. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
10. Pettersson, U. & Philipson, L. (1975) *Cell* **6**, 1–4.
11. Wallace, R. & Kates, J. (1972) *J. Virol.* **9**, 627–635.
12. Pan, J., Celma, M. L. & Weissman, S. M. (1977) *J. Biol. Chem.* **252**, 9047–9054.
13. Celma, M. L., Pan, J. & Weissman, S. H. (1977) *J. Biol. Chem.* **252**, 9032–9042.
14. Ohe, K. & Weissman, S. M. (1971) *J. Biol. Chem.* **246**, 6991–7009.
15. Barrell, B. G. (1971) in *Procedures in Nucleic Acid Research*, eds. Cantoni, G. L. & Davies, D. R. (Harper & Row, New York), Vol. 2, pp. 751–779.
16. Volckaert, G., Min Jou, W. & Fiers, W. (1976) *Anal. Biochem.* **72**, 433–446.
17. Vennström, B., Pettersson, U. & Philipson, L. (1978) *Nucleic Acids Res.* **5**, 205–220.
18. Korn, L. J. & Brown, D. D. (1978) *Cell* **15**, 1145–1156.
19. Zain, S., Gingeras, T. R., Bullock, P., Wong, G. & Gelinas, R. E. (1979) *J. Mol. Biol.* **135**, 413–433.
20. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41.
21. Weil, P. A., Segall, J., Harris, B., Ng, S. Y. & Roeder, F. G. (1979) *J. Biol. Chem.* **254**, 6163–6173.
22. Telford, J. L., Kressman, A., Koski, R. A., Grosschedl, R., Muller, F., Clarkson, S. G. & Birnstiel, M. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2590–2594.
23. Alestrom, P., Akusjärvi, G., Perricaudet, M., Mathews, M. B., Klessig, D. & Pettersson, U. (1980) *Cell* **19**, 671–681.
24. Sakonju, S., Bogenhagen, D. F. & Brown, D. D. (1980) *Cell* **19**, 13–26.
25. Bogenhagen, D. F., Sakonju, S. & Brown, D. D. (1980) *Cell* **19**, 27–36.
26. Kressmann, A., Hofstetter, H., DiCapua, E., Grosschedl, R. & Birnstiel, M. L. (1979) *Nucleic Acids Res.* **7**, 1749–1763.
27. Thimmappaya, B., Jones, N. & Shenk, T. (1979) *Cell* **18**, 947–954.
28. Celma, M. L., Pan, J. & Weissman, S. M. (1977) *J. Biol. Chem.* **252**, 9043–9046.
29. Vennström, B., Pettersson, U. & Philipson, L. (1978) *Nucleic Acids Res.* **5**, 195–204.
30. Nevins, J. R. & Darnell, J. E. (1978) *Cell* **15**, 1477–1493.
31. Mathews, M. B. (1980) *Nature (London)*, in press.
32. Miller, J. S., Ricciardi, R. P., Roberts, B. E., Paterson, B. M. & Mathews, M. B. (1980) *J. Mol. Biol.*, in press.
33. Kozak, M. (1978) *Cell* **15**, 1109–1123.
34. Akusjärvi, G. & Pettersson, U. (1979) *Cell* **16**, 841–850.
35. Zain, S., Sambrook, J., Roberts, R. J., Keller, W., Fried, M. & Dunn, A. R. (1979) *Cell* **16**, 851–861.
36. Seif, I., Khoury, G. & Dhar, R. (1979) *Nucleic Acids Res.* **6**, 3387–3398.