JOURNAL OF VIROLOGY, Mar. 1997, p. 2310–2319 Vol. 71, No. 3
0022-538X/97/$04.00+0

In Vitro Genetic Selection Analysis of Alfalfa Mosaic Virus Coat Protein Binding to 3′-Terminal AUGC Repeats in the Viral RNAs

FELICIA HOUSER-SCOTT, PATRICIA ANSEL-MCKINNEY, JUN-MING CAI, AND LEE GEHRKE*

Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02215, and Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 23 September 1996/Accepted 4 December 1996

The coat proteins of alfalfa mosaic virus (AMV) and the related ilarviruses bind specifically to the 3′ untranslated regions of the viral RNAs, which contain conserved repeats of the tetranucleotide sequence AUGC. The purpose of this study was to develop a more detailed understanding of RNA sequence and/or structural determinants required for coat protein binding by characterizing the role of the AUGC repeats. Starting with a complex pool of 39-nucleotide RNA molecules containing random substitutions in the AUGC repeats, in vitro genetic selection was used to identify RNAs that bound coat protein. After six iterative rounds of selection, amplification, and reselection, 25% of the RNAs selected from the randomized pool were wild type; that is, they contained all four AUGC sequences. Among the 31 clones analyzed, AUGC was clearly the preferred selected sequence at the four repeats, but some nucleotide sequence variability was observed at $AUGC_{865–868}$ if the other three AUGC repeats were present. Variant RNAs that bound coat protein with affinities equal to or greater than that of the wild-type molecule were not selected. To extend the in vitro selection results, RNAs containing specific nucleotide substitutions were transcribed in vitro and tested in coat protein and peptide binding assays. The data strongly suggest that the AUGC repeats provide sequence-specific determinants and contribute to a structural platform for specific coat protein binding. Coat protein may function in maintaining the 3′ ends of the genomic RNAs during replication by stabilizing an RNA structure that defines the 3′ terminus as the initiation site for minus-strand synthesis.

RNA-protein interactions in alfalfa mosaic virus (AMV) and the related ilarviruses are essential for virion assembly and stability (18, 19). In addition, a defining characteristic of AMV and the ilarviruses is that coat protein is required for viral infectivity; that is, the genomic RNAs are not infectious unless coat protein or coat protein mRNA is included in the inoculum (44). The viral genome is divided among three single-stranded positive-sense RNAs (Fig. 1). RNAs 1 and 2 encode the replicase subunits P1 and P2, respectively, whereas RNA 3 encodes both the movement protein P3 and coat protein. The coat protein is translated only from subgenomic RNA 4, which is coterminal with the 3′-terminal 881 nucleotides of genomic RNA 3. AMV and ilarvirus coat proteins are functionally interchangeable in genome activation (11, 12), despite the fact that in some cases there is little primary amino acid sequence homology. A consensus sequence with a crucial arginine residue in the RNA binding domain of the coat proteins may explain the cross-activation phenomenon (1).

During AMV and ilarvirus replication, the viral RNA-dependent RNA polymerase first copies the genomic RNAs into minus-strand RNAs, from which progeny positive strands are transcribed. In other bromoviruses, 3′-terminal tRNA-like or pseudoknot structures are thought to facilitate recognition of the genomic RNAs by viral RNA-dependent RNA polymerase (replicase) molecules (7, 13, 14, 30). In this context, the 3′ terminus of AMV and the related ilarvirus RNAs is structurally distinct because it is nonpolyadenylated, lacks pseudoknots, and cannot be aminoacylated (14). The 3′ termini of AMV and the ilarvirus RNAs do, however, contain high-affinity coat protein binding sites (15, 50, 51) (Fig. 1), and tetranucleotide AUGC repeats are a conspicuous feature of the 3′-terminal nucleotide sequences. That the AUGC sequences might be essential for coat protein binding was proposed earlier (20), and point mutations in the AUGC repeats diminish both coat protein binding and viral RNA replication (15, 33, 43). Houwing and Jaspars (16) proposed that coat protein binding to the 3′ termini of AMV and ilarvirus genomic RNAs induces an RNA conformational change, resulting in the formation of a replicase recognition signal. The results of Taschner et al., however, suggest that minus strands of AMV RNA 3 can be produced in the absence of coat protein (38), thereby arguing against the coat protein-replicase recognition model (16).

Recent evidence suggests that AMV coat protein may have multiple functions in the viral life cycle, and an unresolved issue is whether coat protein function is linked exclusively to viral RNA binding. In addition to genome activation and encapsidation, AMV coat protein has been implicated in several regulatory functions, including RNA stability (28), viral cell-to-cell movement (43), and modulating the ratio of positive- to negative-strand synthesis (6). Although functions in RNA stability and encapsidation imply RNA-protein interactions, a recent virus replication model suggests that coat protein may interact with replicase independently of RNA to promote plus-strand RNA accumulation (6). Other data indicate, however, that coat protein molecules defective in RNA binding are unable to activate the viral genome for replication. Zuidema et al. reported that trypsin treatment, which cleaves the amino-terminal RNA binding domain from coat protein, destroyed genome activation potential (49); moreover, Yusibov and Loesch-Fries (48) found that coat protein molecules containing a mutation in an arginine crucial for RNA binding (1) were also unable to activate viral RNA replication. Possible interpretations for these data are that the amino-terminal RNA binding

* Corresponding author. Mailing address: HST Division, Massachusetts Institute of Technology, 77 Massachusetts Ave., Building E25-545, Cambridge, MA 02139. Phone: (617) 253-7608. Fax: (617) 253-3459. E-mail: lgehrke@mit.edu.

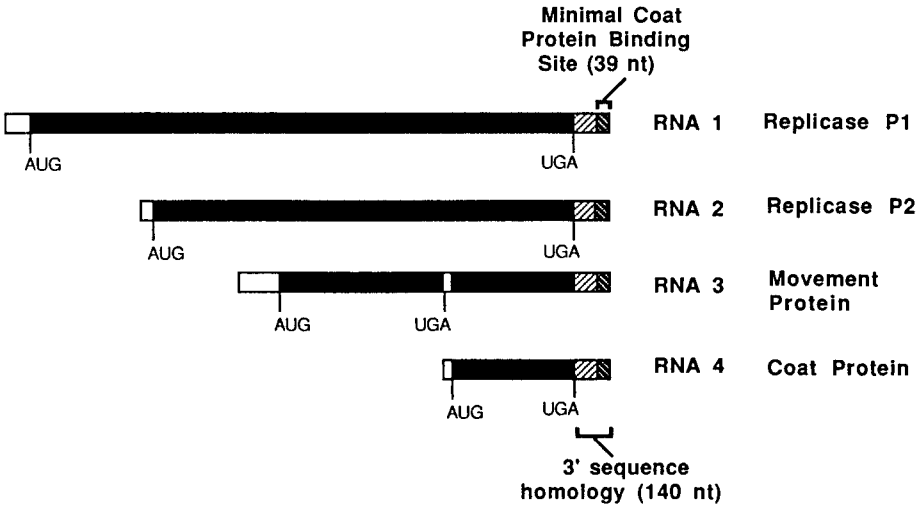


FIG. 1. Schematic representation of AMV RNAs. The 3′-terminal region (~140 nucleotides) of nucleotide sequence homology shared among the RNAs is shown, and the position of a minimal coat protein binding domain identified previously (15) is indicated. nt, nucleotides.

domain also interacts independently with replicase in the model of de Graaf et al. (6) or that a coat protein-RNA complex stimulates plus-strand accumulation.

To begin to understand the complexities of RNA-protein interactions in the life cycles of AMV and ilarviruses, we have focused on identifying determinants in both the RNA and the viral coat protein that are required for forming a specific ribonucleoprotein complex (1, 2, 15). We report here the use of in vitro genetic selection (8, 40) to analyze the functional significance of AUGC repeats for binding AMV coat protein. In contrast to biochemical mutagenesis techniques for analyzing small numbers of nucleotide substitutions, in vitro selection offered the potential for simultaneously analyzing over 4×10^9 nucleotide combinations in a pool of 39-nucleotide RNAs. After six iterative rounds of selection, amplification, and reselection, approximately one-fourth of the clones analyzed were wild type; that is, all four of the AUGC repeats were present. These data have been extended by introducing specific nucleotide changes into the viral RNA and then determining coat protein binding potential. The in vitro genetic selection data and nucleotide substitution and binding studies are evidence that the AUGC repeats are critical nucleotides that, along with arginine_{17} of AMV coat protein (1), represent key determinants of high-affinity binding of AMV coat protein and its amino-terminal peptides to the 3′ termini of the viral RNAs. The basis for the coat protein requirement in replication is not clear; however, coat protein binding to the 3′ termini of the genomic RNAs may stabilize an RNA structure favorable for correct initiation of minus-strand synthesis, thereby ensuring that the 3′ ends are maintained through replication.

MATERIALS AND METHODS

Virion coat protein and peptides. AMV coat protein, a gift from Edward Halk, was purified as described by Kruseman et al. (21). The peptides corresponding to the 38- and 25-amino-terminal amino acids of the coat protein were synthesized by using *t*-butyloxcarbonyl chemistry with an Applied Biosystems model 430A solid-phase peptide synthesizer, and the N terminus was acetylated to correlate with the native form of the protein (41). The peptides were purified by high-pressure liquid chromatography. The peptide amino acid sequences are as follows:

CP25, *N*-acetyl-SSSQKKAGGKAGKPTKRSQNYAALR; and
CP38, *N*-acetyl-SSSQKKAGGKAGKPTKRSQNYAALRKAQLPKPPALKVP.

In vitro transcriptions for site-directed mutagenesis. RNAs were transcribed in vitro by using chemically synthesized DNA templates and bacteriophage T7 RNA polymerase. Transcription was initiated with 5′ GG to increase the transcriptional yield of these small RNAs (25, 26). RNAs were radiolabeled by including α-^{35}S-ATP or [α-^{32}P]UTP (Dupont/NEN) in the nucleotide mixtures used in the transcription reactions. Transcription products were separated by electrophoresis in denaturing polyacrylamide-urea gels followed by localization of the RNA band by UV light shadowing and overnight buffer elution at 4°C [0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, and 0.1% sodium dodecyl sulfate] (4). The eluted RNA was extracted twice with phenol and ethanol precipitated. RNA concentrations were determined spectrophotometrically. In vitro transcription of the fragment corresponding to the 3′-terminal 170 nucleotides of AMV RNA 4 ($\text{AMV}_{718-881}$ RNA) has been previously described (2).

To generate the randomized nucleotides at positions 866 to 868 within $\text{AMV}_{718-881}$ RNA, a *Bgl*II/*Sma*I fragment of $\text{AMV}_{718-881}$ DNA was PCR amplified with the following primers: 5′CACTATAGATCTGGGATGCT3′ and 5′ACGTCGCCCGGGCATCCCTTAGGGNNNTTCATGCA3′. The underlined sequences represent the restriction sites used in subcloning the individual fragments containing the randomized positions into the parent T7 transcription vector. The randomized nucleotides are indicated (boldface).

Electrophoretic mobility shift assay (EMSA). Bandshift assays were performed essentially as described by Weinberger et al. (46) as reported previously (2, 15).

Construction of the random sequence starting pool for AUGC selection. A 100-nucleotide DNA oligomer containing constant primer sequences, the T7 RNA polymerase promoter, and DNA sequences encoding $\text{AMV}_{843-881}$ RNA were chemically synthesized. The 39-nucleotide coat protein binding site was flanked by defined 5′ and 3′ sequences not present in AMV RNA 4. The 5′- and 3′-flanking sequences present in the transcribed RNAs used for selection were 5′GGGAGAAUUCCCGCGGCA3′ and 5′GAAAGAAUCUGCAGCCCGGG AAGCUU3′, respectively. The synthesizer was programmed to insert nucleotides randomly at cycles corresponding to $\text{AUGC}_{843-846}$, $\text{AUGC}_{849-852}$, $\text{AUGC}_{865-868}$, and $\text{AUGC}_{878-881}$ of $\text{AMV}_{843-881}$ RNA. Thus, 16 of 39 nucleotides of the 3′-terminal high-affinity coat protein binding site were random, yielding a starting pool with a theoretical complexity of 4^{16}, or 4.3×10^9 molecules. The synthetic single-stranded DNA template was purified by polyacrylamide gel electrophoresis under denaturing conditions.

For in vitro transcription, approximately 500 pmol of the gel-purified 100-mer DNA was annealed to an equimolar amount of oligonucleotide primer 1 containing the T7 RNA polymerase promoter (indicated in boldface). Primer 2 was used in subsequent amplification steps. The primer sequences were as follows: primer 1, 5′**TAATACGACTCACTATA**GGGAGAATTCCCGCGGCA3′; primer 2, 5′AAGCTTCCCGGGCTGCAGATTCTTTC3′.

The volume of the transcription reaction was 1 ml. Oligonucleotide primer 2 was also used as the 3′ primer for reverse transcriptase reactions (see below). After gel purification, the transcription yield was 2.6 nmol of starting pool RNA.

Selection and amplification procedure. Nitrocellulose filter binding assays were performed essentially as described by Tuerk et al. (39), with minor modifications. Gel-purified RNA was renatured by being heated to 65°C for 2 min in binding buffer (10 mM Tris-HCl [pH 7.5], 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 5% glycerol, 1 U of RNasin per μl, 80 μg of bovine serum albumin

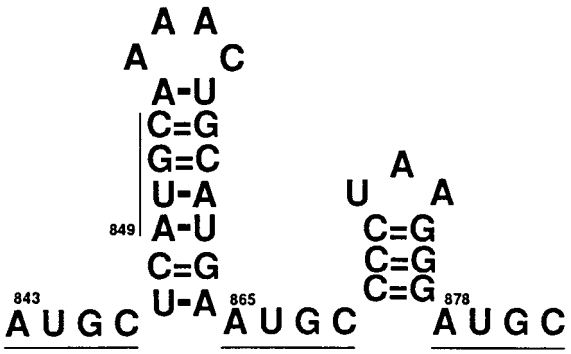


FIG. 2. Schematic representation of a 3′ minimal coat protein binding site showing nucleotide positions that were randomized for the in vitro selection analysis. Oligodeoxyribonucleotides were chemically synthesized as described in Materials and Methods, and the randomized RNA pool was generated by in vitro transcription using bacteriophage T7 RNA polymerase. RNAs in the pool had theoretically equal probabilities of adenosine, cytosine, uracil, or guanosine at each of the underlined positions.

per μl, and 0.15 μg per μl, of *Saccharomyces cerevisiae* phenylalanine tRNA per μl) and slowly cooled to room temperature over approximately 15 min. The RNA and protein were mixed and then incubated for 1 h at room temperature in a 100-μl volume of binding buffer. Following incubation, the mixture was vacuum filtered on HAWP nitrocellulose filters (Millipore, Bedford, Mass.) that had been prewetted with 2.0 ml of TMK buffer (100 mM Tris-HCl [pH 7.5], 80 mM KCl, 10 mM magnesium acetate, and 80 μg of bovine serum albumin per μl). The filters were washed twice with 1.0 ml of TMK buffer to reduce background. The selected RNAs were then eluted twice from the filters by repeated incubation in 200 μl of 7 M urea–3.0 mM EDTA–100 mM sodium citrate (pH 5.0) at 65 or 100°C for 5 min. The combined eluates were extracted with 500 μl of phenol that had been preequilibrated with 0.1 M sodium acetate, pH 5.2. The aqueous phase was subsequently extracted with an equal volume of ether, and the RNA was precipitated by adding 2 volumes of RNase-free water, 5 μg of linear polyacrylamide, and ethanol to a final concentration of 70%. Typically, 90 to 95% of the RNA was recovered from the filter by this method. Eluted RNA was copied into cDNA by using murine reverse transcriptase (Gibco BRL) according to the manufacturer's recommendations at 37°C in 50-μl volumes with a 3′-primer (oligonucleotide 1) concentration of 1 to 2 μM. The cDNA was then amplified by PCR (30 cycles) with oligonucleotides 1 and 2. The double-stranded amplified DNA was retranscribed into RNA and used in the next round of selection. All RNAs used for nitrocellulose filter selection were labeled internally during transcription by incorporating [α-^{32}P]UTP and also sized by electrophoresis in 8% polyacrylamide-urea gels. After the third and sixth selection rounds, PCR-amplified DNA products were cloned (TA cloning kit; In Vitrogen), and the nucleotide sequences of selected clones were determined by using Sequenase 2.0 DNA polymerase (United States Biochemical). For mobility bandshift assays, plasmid DNAs were linearized by digestion with *Sma*I and *Kpn*I and transcribed in vitro as described previously (1, 15).

RESULTS

In vitro genetic selection suggests that the AUGC motifs are necessary for high-affinity coat protein binding to $AMV_{843–881}$. The in vitro selection method is an iterative process of RNA-protein binding, selection, and amplification that can be used to identify protein-binding substrates from a pool of randomized RNAs (17, 40). The four AUGC sequences within a minimal coat protein binding site $AMV_{843–881}$ RNA (Fig. 2) (15) were randomized at positions 843 to 846, 849 to 852, 865 to 868, and 878 to 881 (Fig. 2) to probe the nucleotide sequence and structural specificity of the coat protein-RNA interaction. During the selection, the amount of AMV coat protein added to the reaction mixtures was limiting relative to the amount of pool RNA in order to promote competition and to identify the RNAs that bound with high affinity. The RNA concentrations in rounds 1 to 6 of the selection were 7.9, 1.6, 0.5, 0.5, 0.6, and 0.7 μM, respectively, while the coat protein concentration was 1/10 that of the RNA. After the third and sixth selection rounds, cDNA copies of the selected RNAs were cloned and the nucleotide sequences of a number of isolates were determined.

Of 31 clones analyzed from the sixth-cycle-selected RNA population, 23 unique sequences were identified (Table 1). The wild-type RNA sequence was represented with greatest frequency (8 of 31 clones), and 70% of the unique sequences (16 of 23) contained at least one AUGC. Only one sequence other than the wild type was represented more than once in the 31 clones; it contained three of four AUGC repeats (clone 9). A summary of the selection data showing the nucleotide preference at each position is shown in Fig. 3. Bar graphs in Fig. 3A and B show the percentages of clones analyzed that had the wild-type nucleotide sequence at each AUGC after three and six selection rounds, respectively. Little sequence preference was apparent after three selection rounds (Fig. 3A), while clear nucleotide preference was observed after the sixth selection round. The results presented in Fig. 3B reflect the sixth round selection data, and the numbers used to create the graph shown in Fig. 3B are presented in Fig. 3C. The data indicate that AUGC was preferred at each of the four randomized sites, although greater relative sequence variability was tolerated at $AUGC_{865–868}$ than at the other three repeats. Cytosine was not observed at position 843 or 849, adenosine was not found at position 879 or 880, and uridine was not identified at position 852.

The affinity of several of the cloned RNAs for coat protein was examined by electrophoretic mobility bandshift analysis (data not shown). Consistent with previous results (15), clones containing the wild-type RNA sequence bound coat protein with a relative binding affinity (i.e., the coat protein concentration at which 50% of the RNA is shifted into an RNA-protein complex) of 300 to 600 nM, while clones 9, 18, 33, 36, 37, 39, and 46 all had a relative binding affinity of >6 μM, which was the limit of sensitivity for the assay. Clone 21 (relative binding affinity, ~3.0 μM) was the only non-wild-type clone tested that had measurable relative binding affinity. This RNA differs from the wild type in only two nucleotide positions, $G_{867}U$ and $C_{868}G$ in the $AUGC_{865–868}$ repeat (see Fig. 5D). Overall, non-wild-type selected RNAs consistently bound coat protein with low affinity, and no "superbinding" RNAs (36) were identified. A single clone (clone 44) contained one of four AUGC sequences, and several clones without AUGC were also identified in the round six selected pool. With regard to clone 44, it was demonstrated previously (15) that coat protein binding to RNAs containing AAAA at the $AUGC_{865–868}$ position was not detectable by EMSA. Therefore, it is likely that a small percentage of the RNAs, including clone 44 and those lacking AUGC repeats, were carried nonspecifically throughout the selection, as has been noted by others using this technique (36). The high percentage of wild-type selected clones and the absence of identified covarying nucleotide pairs suggest that the tetranucleotide AUGC at the four repeat positions provides the optimal combination of sequence and structural determinants required for high-affinity coat protein binding.

$AUGC_{849–852}$ maintains structural integrity of a hairpin. In the four AUGC repeats found in the 3′-terminal 31 nucleotides, the $AUGC_{865–868}$ separating the two hairpins (Fig. 2) and the 3′-terminal $GC_{880–881}$ are the most highly conserved among AMV RNA 4 and ilarvirus RNAs (15). Conversely, $AUGC_{849–852}$ is distinct both because it is not conserved among other ilarvirus RNAs and because it is apparently base paired rather than being single stranded in solution (32). Although the in vitro selection data suggested the importance of $AUGC_{849–852}$ in coat protein binding, we considered the possibility that the selection process identified RNAs that retained the hairpin

TABLE 1. RNAs from round 6 of the in vitro genetic selection ranked according to the number of AUGC repeats present[a]

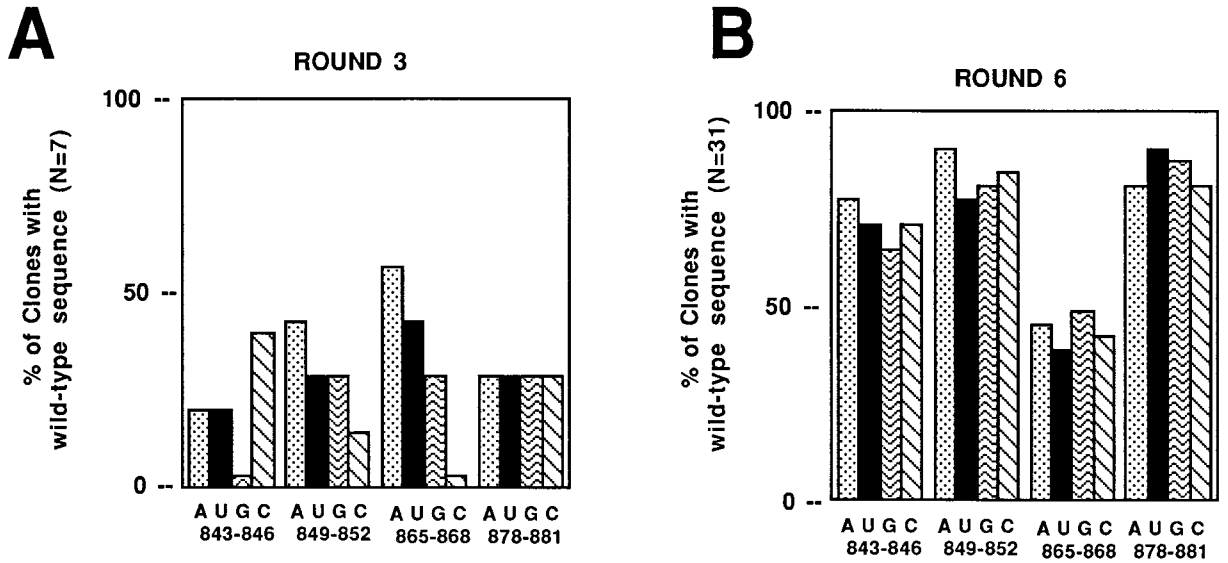
No. of AUGCs and clone[b]	Sequence at positions:				Frequency	Binding value[c]
	843–846	849–852	865–868	878–881		
4/4	AUGC	AUGC	AUGC	AUGC	8	300–600 nM
3/4						
CL 7	AUGC	AUGC	GAUU	AUGC	1	
CL 9	AUGC	AUGC	UGCG	AUGC	2	>6 μM
CL 15	AUGC	AUGC	GCGA	AUGC	1	
CL 17	AUGC	AUGC	CAUU	AUGC	1	
CL 18	AUGC	AUGC	GAAU	AUGC	1	>6 μM
CL 20	AUGC	AUGC	GAUC	AUGC	1	
CL 21	AUGC	AUGC	AUUG	AUGC	1	~3 μM
CL 22	AUGC	AUGC	ACGU	AUGC	1	
CL 46	AUGC	AUGC	CAUA	AUGC	1	>6 μM
CL 66	AUGC	AUGC	UUAA	AUGC	1	
CL 67	AUGC	AUGC	ACUG	AUGC	1	
2/4						
CL 24	GCAA	AUGC	GCAA	AUGC	1	
CL 27	ACUG	ACGC	AUGC	AUGC	1	
CL 65	AAUG	AUGC	GUGC	AUGC	1	
1/4						
CL 44	GUAG	GUUA	AAAA	AUGC	1	
0/4						
CL 33	UACU	AUGA	AAUG	CUCA	1	>6 μM
CL 36	UCAA	UAGC	CGGA	UCUU	1	>6 μM
CL 37	UCAG	AACC	CGGG	GUGC	1	>6 μM
CL 39	GUCC	GGCA	CCUA	GCGG	1	>6 μM
CL 42	AAAC	ACAG	CAGC	UGUU	1	
CL 68	AGUA	AAUC	GAAC	AUGA	1	
CL 69	GAAG	AAAA	CGCU	UUUA	1	

[a] The frequency of each sequence in the selection is indicated. Several clones were tested for coat protein binding activity by EMSA at 22°C.
[b] CL, clone.
[c] Apparent binding value, defined as the protein concentration at which 50% of RNA shifted into an RNA-protein complex.

independent of nucleotide sequence (note that the nucleotides paired with $AUGC_{849–852}$ in the wild-type RNA were not randomized during the selection). Therefore, variant RNAs with nucleotide changes in the 847 to 864 hairpin were synthesized and tested by EMSA for coat protein binding. The variant RNA shown in Fig. 4A maintained the $AUGC_{849–852}$ sequence, but nucleotide substitutions prevented formation of base pairs that maintained the hairpin structure. Alternatively, the variant RNA shown in Fig. 4B altered the $AUGC_{849–852aaaa}$ sequence but maintained the potential for base pairing. The accompanying EMSA data shown in Fig. 4 indicate that the RNA lacking potential to form the hairpin (Fig. 4A) did not bind coat protein or peptides at concentrations up to 5 μM, as evidenced by the absence of mobility shifts. However, an RNA with potential for forming A · U base pairs in place of $AUGC_{849–852}$ (Fig. 4B) coat protein retained binding potential (relative binding affinity, ~2.5 μM). These data are evidence that structural features of the 847 to 864 hairpin are important for coat protein binding, independent of the AUGC nucleotide sequence. A related approach was recently reported by Reusken and Bol (34).

Minor nucleotide substitutions are tolerated for coat protein binding in the $AUGC_{865–868}$ repeat. The in vitro selection data (Table 1 and Fig. 3) suggested that although AUGC is clearly the preferred nucleotide sequence at $AUGC_{865–868}$, other sequences at this site would support low-affinity coat protein binding. Indeed, 20 of 31 selected RNAs had at least three perfect AUGC repeats out of four possible repeats (i.e., wild type or three of four AUGC; Table 1), and the imperfect repeat was always $AUGC_{865–868}$ (Table 1). To test nucleotide sequence flexibility at the $AUGC_{865–868}$ repeat, RNAs randomized at positions 866 to 868 within $AMV_{718–881}$ RNA (comprising the majority of the 3′ untransplated region (15)) were analyzed. These RNAs were not tested by in vitro genetic selection; rather, oligonucleotides with random nucleotides inserted at three positions (U_{866}, G_{867}, and C_{868}) were synthesized and used to prime PCR amplification of the ~160-nucleotide 3′ untranslated region of AMV RNA 4. The resultant DNA fragments were subcloned, and plasmid DNAs were used to transform competent bacteria. Isolated colonies were picked, and the nucleotide sequence of the insert DNAs was determined. Of the 64 possible clones (4^3), our analysis was limited to 18 of 30 unique sequences that were identified after the sequences of about 150 clones were analyzed. RNA binding to coat protein peptide CP25 (2) was determined both by EMSA using radiolabeled mutant RNAs and by analysis of the capacity of the mutant RNAs to compete against wild-type $AMV_{718–881}$ RNA for peptide binding. Peptide CP25 was used in the binding experiments because prior data had indicated that $AUGC_{865–868}$ was a major determinant for peptide binding (15).

The results are summarized in Table 2. As predicted, several RNAs with nucleotide substitutions at $AUGC_{865–868}$ (AAGC, ACGG, ACGU, ACUG, ACUU, AGGC, and AUGA) retained low-affinity peptide binding capacity in the context of wild-type AUGC at the three other repeat positions (843 to 846, 849 to 852, and 878 to 881). Significantly, two of these sequences (ACGU and ACUG) were also identified by in vitro selection (Table 1, clones 22 and 67), while none of the variant RNA sequences that lack binding capacity (Table 2) were identified by selection (Table 1). Reusken et al. (33) also reported that a comparable 208-nucleotide AMV RNA fragment



C

	A843	U844	G845	C846	A849	U850	G851	C852	A865	U866	G867	C868	A878	U879	G880	C881
A	**24**	4	6	3	**28**	4	2	4	**14**	9	5	7	**25**	0	0	3
U	3	**22**	3	1	1	**24**	2	0	3	**12**	8	5	3	**28**	3	2
G	4	1	**20**	5	2	1	**25**	1	7	5	**15**	6	2	1	**27**	1
C	0	4	2	**22**	0	2	2	**26**	7	5	3	**13**	1	2	1	**25**

FIG. 3. Nucleotide sequence representation in the clones selected at the third and sixth iterative rounds. (A) Nucleotide preference observed after three selection rounds. After three rounds, cDNA copies of the selected RNAs were generated, cloned, and sequenced. The graph, based on the analysis of seven sequences, shows the percentages of clones analyzed that had the wild-type nucleotide sequence at each of the AUGC positions. (B) Nucleotide preference observed after six selection rounds. The graph, based on the analysis of 31 sequences, shows the percentages of clones analyzed that had the wild-type nucleotide sequence at each of the AUGC positions. (C) The numbers used to generate the graph shown in panel B. In this table, representation of each AUGC nucleotide in the 31 clones analyzed is shown.

containing AGGC at nucleotides 865 to 868 bound coat protein. Therefore, the in vitro genetic selection data (Table 1) and the mutagenesis data (Table 2) correlate well. Considered together, the in vitro genetic selection results (Table 1) and the $AUGC_{865–868}$ variant RNA data (Table 2) suggest that the optimal RNA substrate for high-affinity coat protein binding contains four AUGC sequences, with $AUGC_{849–852}$ fulfilling a structural role. Furthermore, limited sequence variability at $AUGC_{865–868}$ is permitted (although the binding affinity is lowered) in an environment consisting of hairpins 847 to 864 and 869 to 877 and AUGC repeats 843 to 846 and 878 to 881. The striking selection and preference for the nucleotide sequence AUGC at positions 843 to 846 and 878 to 881 strongly suggest sequence-specific interactions with coat protein or peptides. As we discuss in the next section, it is likely that these nucleotides contribute to RNA structural features required for coat protein binding as well.

DISCUSSION

This report focuses on the significance of repeated tetranucleotide AUGC motifs for specific binding of AMV coat protein and N-terminal peptides to AMV and ilarvirus RNAs. In previous work (2, 15), evidence that the 3′-terminal 39-nucleotide fragment of AMV RNAs is a minimal coat protein or peptide binding site was presented. The AUGC sequences are a defining feature of AMV and ilarvirus RNAs, the 3′ termini of which can be folded into similar secondary structures (15, 18). The AUGC separating the two hairpins shown in Fig. 2 (i.e., nucleotides 865 to 868) is completely conserved among AMV and ilarvirus RNAs, as is the 3′-terminal GC (i.e., nucleotides 880 to 881) (15). Conversely, $AUGC_{849–852}$ is not conserved, while $AUGC_{843–846}$, found at the 5′ end of the minimal coat protein binding site shown in Fig. 2, and $AU_{878–879}$ are partially conserved. Although enzymatic structure mapping data of deproteinized RNA suggest that hairpin 847 to 864 is stable in solution, the extreme 3′ end of the RNA (including G-C hairpin 869 to 877) can assume multiple conformations (1a, 32). To evaluate the role of the repeated AUGC sequences for coat protein binding and to better define the structural details of the AMV RNAs, we have used a combination of in vitro genetic selection and mutagenesis approaches.

The selection protocol applied involved randomizing only the 16 AUGC nucleotides in $AMV_{843–881}$, with the remaining nucleotides unchanged. Because the complexity of the randomized pool was relatively low (4^{16}), each of the predicted random RNA sequences was represented multiple times in the initial RNA pool. The most striking result was that after six iterative rounds, approximately one-fourth of the selected RNAs were wild type and most of the remaining selected RNAs contained two or three perfect AUGC repeats. No

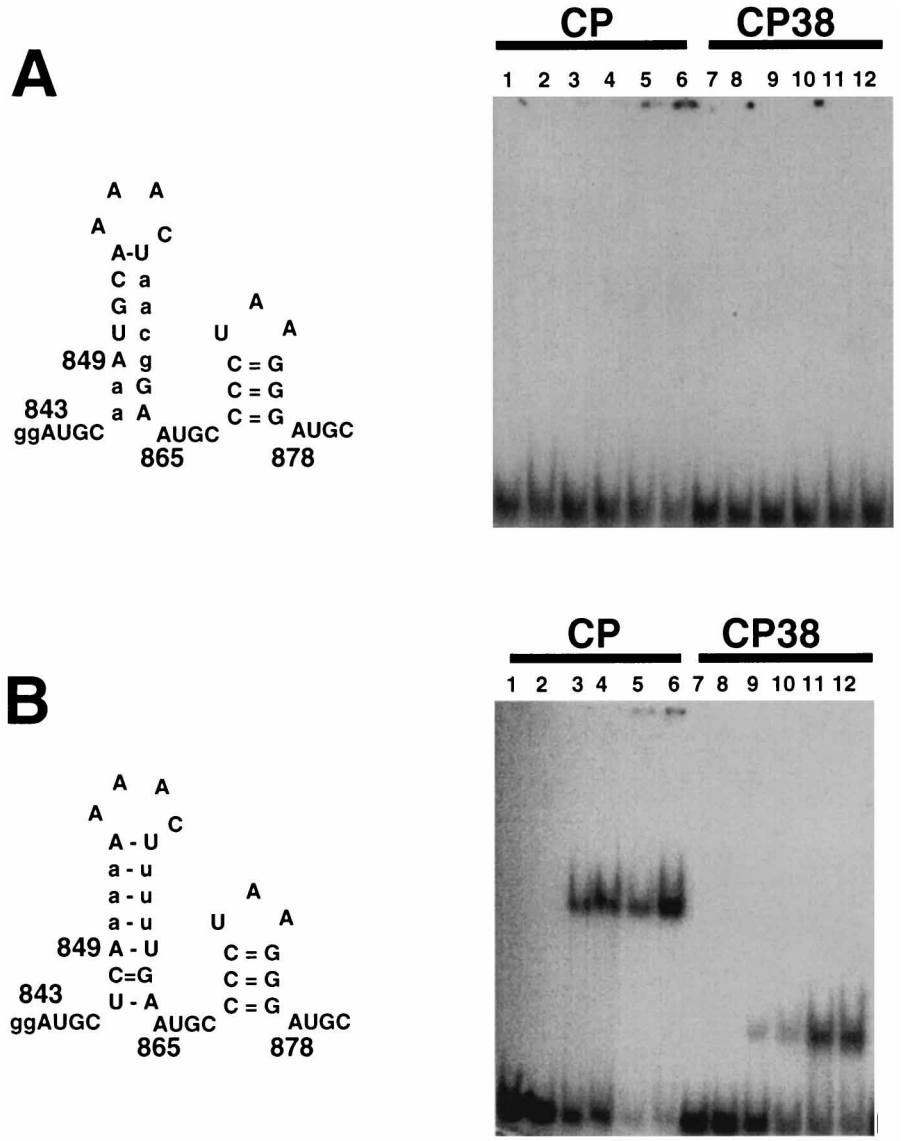


FIG. 4. Analysis of sequence and structural determinants for coat protein binding in hairpin $AUGC_{849–852}$. (A) Effect of maintaining the AUGC nucleotide sequence while preventing the formation of the hairpin at positions 847 to 864. (Left) Schematic representation of the RNA. Substitutions that prevent formation of hairpin 847 to 864 are indicated (lowercase letters). (Right) EMSA of AMV coat protein and peptide binding. The concentration of radiolabeled RNA in each reaction mixture was 100 nM. Lanes 1 and 7, labeled RNA only; lanes 2 to 6, labeled RNA plus coat protein at concentrations of 100 nM, 500 nM, 1 μM, 2.5 μM, and 5 μM, respectively; lanes 8 to 12, labeled RNA plus peptide CP38, corresponding to the N-terminal 38 amino acids of AMV coat protein. The CP38 concentrations were 100 nM, 500 nM, 1 μM, 2.5 μM, and 5 μM, respectively. (B) Effect of maintaining potential for hairpin formation while altering the AUGC nucleotide sequence. (Left) Schematic representation of the substituted RNA. Nucleotides 850 to 852 were changed from UGC to AAA, and compensatory changes at positions 860 to 862 (UUU) were also introduced to retain basepairing potential. (Right) EMSA of AMV coat protein and peptide binding. The concentrations of labeled RNA, peptide, and coat protein are the same as for panel A.

superbinding RNAs were identified; in fact, the coat protein binding affinities of all non-wild-type selected RNAs were diminished at least 10- to 20-fold. These data suggest that within the constraints of the experimental design, the number of RNA sequence solutions for high-affinity coat protein binding is very limited; in fact, AUGC repeats seem to represent the single optimal solution.

Why is the sequence AUGC, in the context of other AMV RNA determinants, optimal for high-affinity coat protein and peptide binding? Possibilities include sequence-specific interactions with coat protein, presentation of structural features that define a platform for binding, or both. Although in vitro selection can provide insight into RNA secondary structure by covariation analysis (3), nucleotide covariation was not apparent in our selections. This result suggests either that coat protein interacts specifically with every unpaired AUGC nucleotide or that peptide binds specifically to some AUGC nucleotides and also stabilizes structural features based on interactions between others. At least two lines of evidence support the latter hypothesis, and a secondary structure model that incorporates the selection and mutagenesis data is presented in Fig. 5A. First, hydroxyl radical footprinting (1) and chemical modification interference studies (1a) strongly suggest that coat protein and coat protein peptides interact specifically with some, but not all, of the AUGC nucleotides. Second, circular dichroism data and unusual electrophoretic

TABLE 2. Summary of the CP25 binding activity of $AMV_{817–881}$ RNA containing random nucleotides at positions 866 to 868[a]

$AUGC_{865–868}$ nucleotides	Binding[b]
AAAA	−
AAAC	−
AACC	−
AACA	−
AAGC	+
AAUA	−
ACGG	+
ACGU	+
ACUG	+
ACUU	+
AGAC	−
AGCG	−
AGGC	+
AGUC	+/−
AUAC	+/−
AUAG	−
AUGA	+
AUGC	+

[a] After chemically synthesized DNA fragments containing the randomized sequences were subcloned, competent bacteria were transformed with the plasmid DNA. Colonies were selected and the nucleotide sequence of the clone was determined. RNA was generated by in vitro transcription and tested by EMSA for CP25 binding.

[b] +, the mutant RNA bound CP25 as determined by direct EMSA and also competed with wild-type $AMV_{718–881}$ for peptide binding; +/−, the RNA bound CP25 in direct EMSA but failed to compete effectively against $AMV_{817–881}$ for peptide binding; −, the mutant RNA lacked detectable CP25 binding in both the EMSA and the competition assay.

mobility of AMV RNA-peptide complexes in nondenaturing polyacrylamide gels suggest an AMV RNA conformational change upon peptide binding (2). Together, the strong selection for the specific AUGC sequences, the sequence conservation in AMV and ilarvirus RNAs, the evidence for conformational change upon peptide binding (2), and chemical modification interference data (1a) suggest that the AUGC sequences have inherent sequence as well as structural significance for coat protein binding.

Predicted structures for several other RNAs that have been tested in binding assays (Table 1) are shown in Fig. 5B to D. Both the selected clone 9 RNA (Fig. 5B) and $AUGC_{866–868aaa}$ (Fig. 5C) bind coat protein poorly, if at all (relative binding values of >6 μM). Clone 9, which contains $AUGC_{843–846}$, $AUGC_{849–852}$, and $AUGC_{878–881}$, had an unexpected binding phenotype. Here, $AUGC_{865–868}$ is replaced by $UGCG_{865–868}$; however, an AUGC is reconstituted if the A at position 864 is included to present $\underline{AUGC}G_{864–868}$ (Fig. 5B). The poor binding phenotype of this RNA suggests that $AUGC_{865–868}$ must be single stranded and/or maintained in the correct spatial orientation for high-affinity coat protein binding. Although RNA $AUGC_{866–868aaa}$ (15) (Fig. 5C) may fold into a structure similar to that proposed in Fig. 5A, the putative $U_{866} \cdot A_{878}$ base pair is absent and the substitution of adenosine at positions 866 to 868 may prevent base-specific contacts with coat protein and peptides. Excepting the wild-type RNAs, none of the other selected RNAs bound coat protein with affinities comparable to that of the wild type; however, clone 21 (relative binding affinity, 3 μM) had a measurable affinity. A predicted structure for this RNA (Fig. 5D) shows that all four of the 3′-terminal AUGC nucleotides could base pair. The diminished binding affinity may result from nucleotide substitutions in the bulged nucleotides at positions 867 and 868.

Secondary structure predictions suggest that although their 3′-terminal nucleotide sequences vary, AMV and ilarvirus RNAs have the potential to fold into similar conformations. In Fig. 6, theoretical secondary structure folding patterns for the extreme 3′ termini of tobacco streak virus RNA 4 and citrus variegation virus (CVV) RNA 4 are shown. The length of the 5′ hairpins varies, and the CVV RNA may contain a bulge loop; however, the AUGC repeats can be positioned at the base of the hairpins and can form a junction structure similar to that shown for AMV RNAs 3 and 4 in Fig. 5A. Although several lines of evidence (cited above) are consistent with these RNA structures, direct experimental evidence in support of these folding patterns is limited at this time. The secondary structure presented in Fig. 5A resembles folding patterns that have been proposed for the 3′ terminus of bacteriophage ϕ6 RNAs, a region that has been implicated in polymerase recognition or function (27).

Potential roles for hairpins and AUGC sequences in viral RNA replication have been described not only for AMV and ilarviruses, but also for yeast viruses and human alphaviruses. Using a deletion mutant of L-A virus, Esteban et al. identified an essential hairpin and a critical 3′-terminal AUGC required for in vitro replication (9), and a model for interaction of the 3′ terminus of yeast L-A virus RNA with replicase has been proposed by Fujimura and Wickner (10). Comparative sequence analysis also reveals conserved AUGC repeats near the 3′ terminus of alphavirus RNAs (Fig. 7). The alphavirus AUGC repeats are found in a region where 11 of 12 alphavirus RNA nucleotides align perfectly with nucleotides 858 to 869 of AMV RNA 4 (Fig. 7). The region of homology shown in Fig.

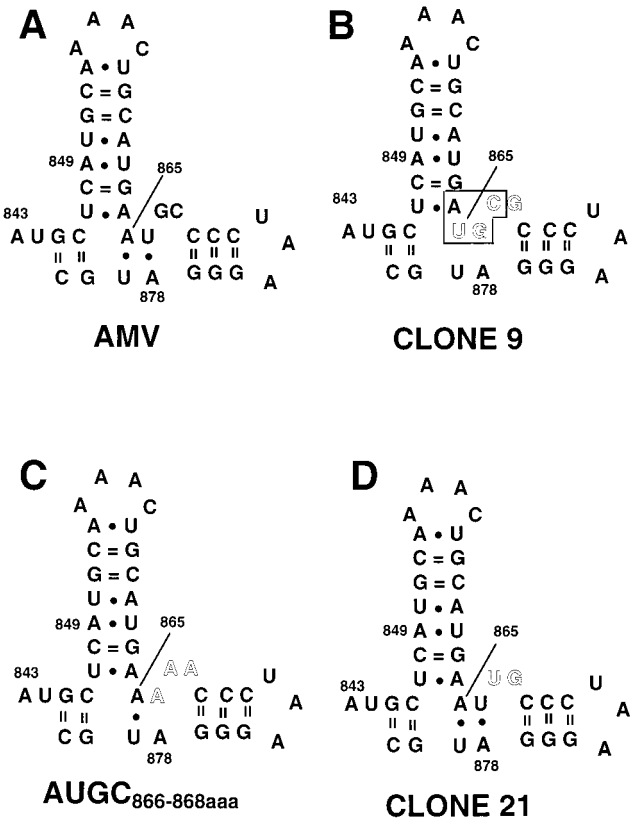


FIG. 5. Secondary structure folding models. (A) Wild-type $AUGC_{843–881}$ RNA; (B) clone 9 from the in vitro genetic selection. Boxed region, an alternate AUGC formed by including A_{864} with $UGC_{865–867}$ of selected clone 9 RNA. (C) RNA $AUGC_{866–868aaa}$ (15); (D) clone 21 from the in vitro genetic selection. Outlined letters indicate non-wild-type sequences selected in $AUGC_{865–868}$.

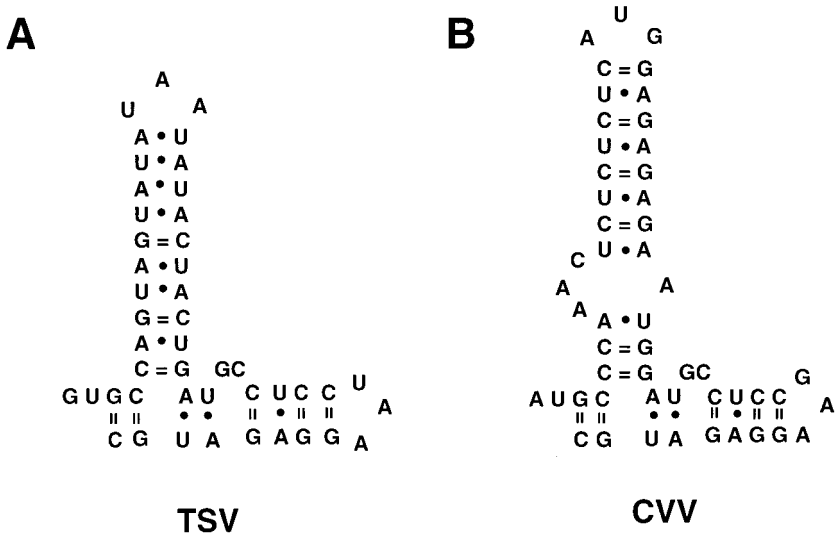


FIG. 6. Predicted secondary structure folding for the 3′ termini of ilarvirus RNAs from tobacco streak virus (TSV) RNA 3/4 (5) (A) and CVV RNA 3/4 (37) (B).

7 is repeated several times in alphavirus RNA 3′ untranslated regions (29) and has been implicated in replicase binding and host factor interactions (22). Although the AUGC sequences are found in common among alphaviruses, AMV, and ilarviruses, the extended homology shown in Fig. 7B is shared only by alphaviruses and AMV RNAs 1 to 4. AMV and alphaviruses are arthropod transmitted (AMV by aphids and alphaviruses by mosquitoes), a feature that distinguishes AMV from ilarviruses and has resulted in the classification of AMV as the sole member of the genus *Alfalmovirus*.

AMV and ilarviruses belong to the family *Bromoviridae*; however, the RNAs lack the characteristic 3′-terminal tRNA-like structure found in other bromoviruses that is believed to be important for RNA replication (24, 35). Houwing and Jaspars proposed that coat protein binding may be important for specific RNA-replicase interactions at the early stages of replication (16). However, Taschner et al. reported that minus-strand RNA 3 synthesis can take place in the absence of coat protein in transgenic protoplasts expressing replicase subunits P1 and P2 (38). This latter finding argues that coat protein is

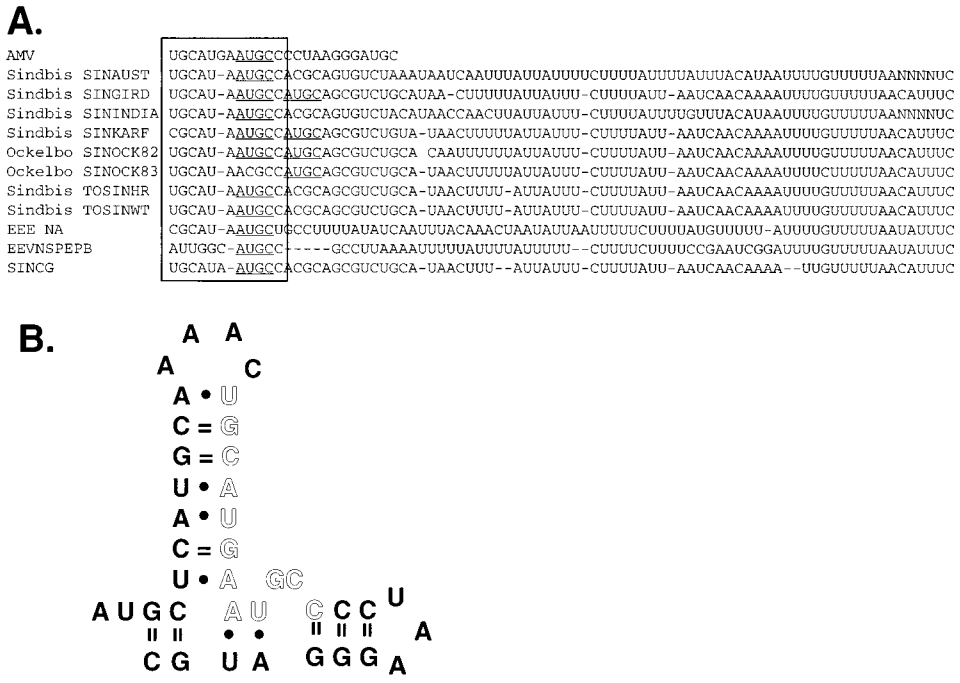


FIG. 7. AUGC sequences at the 3′ termini of alphavirus RNAs. (A) Alphavirus sequences aligned at their 3′ termini, with AUGC sequences (underlined), an extended region of homology between AMV RNAs and alphavirus RNAs (boxed), and single gaps in the alignment (dashes) indicated. The database descriptor is adjacent to each sequence. (B) Location of the AMV-alphavirus homology in AMV RNAs 3 and 4 at the extreme 3′ terminus. The nucleotides that are boxed in panel A are shown outlined. The nucleotide numbering is the same as for Fig. 5A.

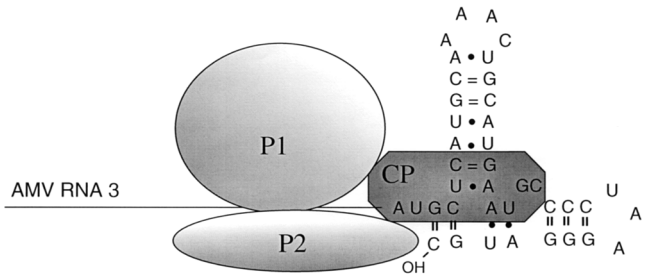


FIG. 8. Model for AMV coat protein (CP) binding to viral RNAs and potential interactions with viral replicase. The line represents the 3′ end of genomic AMV RNA 3, and the proposed terminal structure stabilized by coat protein binding is indicated. Viral replicase subunits P1 (putative helicase activity) and P2 (putative RNA-dependent RNA polymerase) (23) are suggested to bind to a promoter region in the 3′ untranslated region (42), and the putative coat protein-replicase interaction in the figure is based on coisolation experiments reported by Quadt et al. (31).

not required for replicase recognition of plus-strand RNA. At the same time, however, mutations in the AUGC sequences that are important for coat protein binding also interfere with viral RNA replication in plant protoplasts (43). Similarly, addition of the 39-nucleotide minimal binding site RNA (Fig. 2) to an inoculum containing genomic RNAs 1 to 3 plus coat protein severely diminishes AMV replication in protoplasts, while a mutant RNA unable to bind coat protein ($AUGC_{865–868}$ changed to AAAA) has no effect (15a). Overall, the data suggest that mutations affecting coat protein-RNA interactions correlate directly with diminished replication efficiency.

AMV and ilarviruses are distinct among bromoviruses because of both absence of a 3′-terminal tRNA-like structure and the fact that coat protein is required for viral infectivity (a phenomenon called genome activation). The tRNA-like structures found in both coliphage Qβ RNA and plant viral RNAs have been termed "molecular fossils" that tag genomic RNAs for replication and have the added function of a telomere, preventing loss of 3′-terminal nucleotides during successive replication rounds (47). In the absence of a tRNA-like 3′ end, how do AMV and ilarviruses deal with "the telomere problem" (45) and maintain their 3′ termini? We hypothesize that the coat protein-RNA interaction may establish the functional equivalent of a tRNA-like structure present in other plant viral RNAs and that this ribonucleoprotein complex may serve to prevent loss of 3′-terminal nucleotides during successive rounds of replication. A model suggested by the composite data is presented in Fig. 8. The 3′ terminus of the genomic RNAs is presented with replicase subunits P1 and P2 bound to a promoter region in the 3′ untranslated region (42). Coat protein, which coisolates with viral replicase (31), is shown interacting both with the RNA and with the replicase. A prediction of this model is that coat protein binding to sequence and structural determinants involving the AUGC repeats stabilizes a conformation that positions the 3′-terminal nucleotide in the correct orientation for accurate initiation of replication. This mechanism is analogous to the proposed role of elongation factors Tu and Ts in immobilizing the extreme 3′ end of the Qβ genome for accurate replication (47). It is possible that AMV RNA 3 minus strands can be replicated without coat protein (38) but with diminished efficiency and heterogeneous 3′ ends. Since the AUGC repeats are conserved in the context of similar predicted secondary structures, the mechanism proposed in Fig. 8 could apply generally to AMV and the ilarviruses. Further insight into this question will require quantitation of replication efficiencies and nucleotide sequence analysis of the 5′ termini of minus-strand RNAs generated in the presence and absence of coat protein.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM-42504 from the National Institute of General Medical Sciences.

We thank Philippe Leboulch for insightful comments and for many helpful discussions and L. Sue Loesch-Fries for comments on the manuscript.

REFERENCES

1. **Ansel-McKinney, P., S. W. Scott, M. Swanson, X. Ge, and L. Gehrke.** 1996. A plant viral coat protein RNA-binding consensus sequence contains a crucial arginine. EMBO J. **15:**5077–5084.
1a. **Ansel-McKinney, P., and L. Gehrke.** Unpublished data.
2. **Baer, M., F. Houser, S. L. Loesch-Fries, and L. Gehrke.** 1994. Specific RNA binding by amino-terminal peptides of alfalfa mosaic virus coat protein. EMBO J. **13:**727–735.
3. **Bartel, D. P., M. L. Zapp, M. R. Green, and J. W. Szostak.** 1991. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. Cell **67:**529–536.
4. **Calnan, B. J., S. Biancalana, D. Hudson, and A. D. Frankel.** 1991. Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. Genes Dev. **5:**201–210.
5. **Cornelissen, B. J. C., H. Janssen, D. Zuidema, and J. F. Bol.** 1984. Complete nucleotide sequence of tobacco streak virus RNA 3. Nucleic Acids Res. **12:**2427–2437.
6. **de Graaf, M., M. R. Man in't Veld, and E. M. Jaspars.** 1995. In vitro evidence that the coat protein of alfalfa mosaic virus plays a direct role in the regulation of plus and minus RNA synthesis: implications for the life cycle of alfalfa mosaic virus. Virology **208:**583–589.
7. **Duggal, R., F. C. Lahser, and T. C. Hall.** 1994. *cis*-acting sequences in the replication of plant viruses with plus-sense RNA genomes. Annu. Rev. Phytopathol. **32:**287–309.
8. **Ellington, A. D., and J. W. Szostak.** 1990. In vitro selection of RNA molecules that bind specific ligands. Nature **346:**818–822.
9. **Esteban, R., T. Fujimura, and R. B. Wickner.** 1989. Internal and terminal *cis*-acting sites are necessary for in vitro replication of the L-A double-stranded RNA virus of yeast. EMBO J. **8:**947–954.
10. **Fujimura, T., and R. B. Wickner.** 1992. Interaction of two *cis* sites with the RNA replicase of the yeast L-A virus. J. Biol. Chem. **267:**2708–2713.
11. **Gonsalves, D., and S. M. Garnsey.** 1974. Infectivity of multiple nucleoprotein and RNA components of citrus leaf rugose virus. Virology **64:**343–353.
12. **Gonsalves, D., and S. M. Garnsey.** 1975. Infectivity of heterologous RNA-protein mixtures from alfalfa mosaic, citrus leaf rugose, citrus variegation, and tobacco streak viruses. Virology **67:**319–326.
13. **Haenni, A.-L., S. Joshi, and F. Chapeville.** 1982. tRNA-like structures in the genomes of RNA viruses. Prog. Nucleic Acid Res. Mol. Biol. **27:**85–104.
14. **Hall, T. C.** 1979. Transfer-like structures in viral genomes. Int. Rev. Cytol. **60:**1–26.
15. **Houser-Scott, F., M. L. Baer, K. F. Liem, Jr., J.-M. Cai, and L. Gehrke.** 1994. Nucleotide sequence and structural determinants of specific binding of coat protein or coat protein peptides to the 3′ untranslated region of alfalfa mosaic virus RNA 4. J. Virol. **68:**2194–2205.
15a. **Houser-Scott, F., L. Gehrke, and L. S. Loesch-Fries.**, Unpublished data.
16. **Houwing, C. J., and E. M. Jaspars.** 1978. Coat protein binds to the 3′-terminal part of RNA 4 of alfalfa mosaic virus. Biochemistry **17:**2927–2933.
17. **Irvine, D., C. Tuerk, and L. Gold.** 1991. SELEXION. Systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. J. Mol. Biol. **222:**739–761.
18. **Jaspars, E. M. J.** 1985. Interaction of alfalfa mosaic virus nucleic acid and protein, p. 155–221. *In* J. W. Davies (ed.), Molecular plant virology. CRC Press, New York, N.Y.
19. **Kaper, J. M.** 1973. Arrangement and identification of simple isometric viruses according to their dominating stabilizing interactions. Virology **55:**299–304.
20. **Koper-Zwarthoff, E. C., and J. F. Bol.** 1980. Nucleotide sequence of the putative recognition site for coat protein in the RNAs of alfalfa mosaic virus and tobacco streak virus. Nucleic Acids Res. **8:**3307–3318.
21. **Kruseman, J., B. Kraal, E. Jaspars, J. Bol, F. Brederode, and H. Veldstra.** 1971. Molecular weight of the coat protein of alfalfa mosaic virus. Biochemistry **10:**447–455.
22. **Kuhn, R. J., H. Zhang, and J. H. Strauss.** 1990. Mutagenesis of the 3′ nontranslated region of Sindbis virus RNA. J. Virol. **64:**1465–1476.
23. **Matthews, R. E. F.** 1991. Plant virology. Harcourt Brace Jovanovich, New York, N.Y.
24. **Miller, W. A., J. J. Bujarski, T. W. Dreher, and T. C. Hall.** 1986. Minus-strand initiation by brome mosaic virus replicase within the 3′ tRNA-like structure of native and modified templates. J. Mol. Biol. **187:**537–546.

25. **Milligan, J. F., D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck.** 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. **15:**8783–8798.
26. **Milligan, J. F., and O. C. Uhlenbeck.** 1989. Synthesis of small RNAs using T7 RNA polymerase. Methods Enzymol. **180:**51–62.
27. **Mindich, L., X. Qiao, S. Onodera, P. Gottlieb, and J. Strassman.** 1992. Heterologous recombination in the double-stranded RNA bacteriophage Phi6. J. Virol. **66:**2605–2610.
28. **Neeleman, L., E. A. G. van der Vossen, and J. F. Bol.** 1993. Infection of tobacco with alfalfa mosaic virus cDNAs sheds light on the early function of coat protein. Virology **196:**883–887.
29. **Ou, H.-J., D. W. Trent, and J. H. Strauss.** 1982. The 3′-non-coding regions of alphavirus RNAs contain repeating sequences. J. Mol. Biol. **156:**719–730.
30. **Pleij, C. W. A., K. Rietveld, and L. Bosch.** 1985. A new principle of RNA folding based on pseudoknotting. Nucleic Acids Res. **13:**1717–1731.
31. **Quadt, R., H. J. Rosdorff, T. W. Hunt, and E. M. Jaspars.** 1991. Analysis of the protein composition of alfalfa mosaic virus RNA-dependent RNA polymerase. Virology **182:**309–315.
32. **Quigley, G. J., L. Gehrke, D. A. Roth, and P. E. Auron.** 1984. Computer-aided nucleic acid secondary structure modeling incorporating enzymatic digestion data. Nucleic Acids Res. **12:**347–366.
33. **Reusken, C. B., L. Neeleman, and J. F. Bol.** 1994. The 3′-untranslated region of alfalfa mosaic virus RNA 3 contains at least two independent binding sites for viral coat protein. Nucleic Acids Res. **22:**1346–1353.
34. **Reusken, C. B. E. M., and J. F. Bol.** 1996. Structural elements of 3′-terminal coat protein binding site in alfalfa mosaic virus RNAs. Nucleic Acids Res. **24:**2660–2665.
35. **Schimmel, P.** 1989. RNA pseudoknots that interact with components of the translation apparatus. Cell **58:**9–12.
36. **Schneider, D., C. Tuerk, and L. Gold.** 1992. Selection of high affinity RNA ligands to the bacteriophage R17 coat protein. J. Mol. Biol. **228:**862–869.
37. **Scott, S. W., and X. Ge.** 1995. The complete nucleotide sequence of RNA 3 of citrus leaf rugose and citrus variegation ilarviruses. J. Gen. Virol. **76:**957–963.
38. **Taschner, P. E. M., A. C. van der Kuyl, L. Neeleman, and J. F. Bol.** 1991. Replication of an incomplete alfalfa mosaic virus genome in plants transformed with viral replicase genes. Virology **181:**445–450.
39. **Tuerk, C., S. Eddy, D. Parma, and L. Gold.** 1990. Autogenous translational operator recognized by bacteriophage T4 DNA polymerase. J. Mol. Biol. **213:**749–761.
40. **Tuerk, C., and L. Gold.** 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science **249:**505–510.
41. **van Beynum, G. M., J. M. de Graaf, A. Castel, B. Kraal, and L. Bosch.** 1977. Structural studies on the coat protein of alfalfa mosaic virus. The complete primary structure. Eur. J. Biochem. **72:**63–78.
42. **van der Kuyl, A. C., K. Langereis, C. J. Houwing, E. M. Jaspars, and J. F. Bol.** 1990. *cis*-acting elements involved in replication of alfalfa mosaic virus RNAs in vitro. Virology **176:**346–354.
43. **van der Vossen, E. A., L. Neeleman, and J. F. Bol.** 1994. Early and late functions of alfalfa mosaic virus coat protein can be mutated separately. Virology **202:**891–903.
44. **van Vloten-Doting, L.** 1975. Coat protein is required for infectivity of tobacco streak virus: biological equivalence of the coat proteins of tobacco streak and alfalfa mosaic viruses. Virology **65:**215–225.
45. **Watson, J. D.** 1972. Origin of concatemeric T7 DNA. Nature New Biol. **239:**197–201.
46. **Weinberger, J., D. Baltimore, and P. Sharp.** 1986. Distinct factors bind to apparently homologous sequences in the immunoglobin heavy-chain enhancer. Nature (London) **322:**846–848.
47. **Weiner, A. M., and N. Maizels.** 1987. tRNA-like structures tag the 3′ ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. Proc. Natl. Acad. Sci. USA **84:**7383–7387.
48. **Yusibov, V. M., and L. S. Loesch-Fries.** 1995. N-terminal basic amino acids of alfalfa mosaic virus coat protein involved in the initiation of infection. Virology **208:**405–407.
49. **Zuidema, D., M. A. Bierhuizen, and E. M. J. Jaspars.** 1983. Removal of the N-terminal part of alfalfa mosaic virus coat protein interferes with the specific binding to RNA 1 and genome activation. Virology **129:**255–260.
50. **Zuidema, D., M. F. Bierhuizen, B. J. Cornelissen, J. F. Bol, and E. M. Jaspars.** 1983. Coat protein binding sites on RNA 1 of alfalfa mosaic virus. Virology **125:**361–369.
51. **Zuidema, D., and E. M. J. Jaspars.** 1984. Comparative investigations on the coat protein binding sites of the genomic RNAs of alfalfa mosaic and tobacco streak viruses. Virology **135:**43–52.