Degenerate in Vitro Genetic Selection Reveals Mutations That Diminish Alfalfa Mosaic Virus RNA Replication without Affecting Coat Protein Binding

JOURNAL OF VIROLOGY, Aug. 2004, p. 8036–8046 Vol. 78, No. 15
0022-538X/04/$08.00+0 DOI: 10.1128/JVI.78.15.8036–8046.2004

Degenerate In Vitro Genetic Selection Reveals Mutations That Diminish Alfalfa Mosaic Virus RNA Replication without Affecting Coat Protein Binding

Gail Rocheleau,† Jessica Petrillo,† Laura Guogas, and Lee Gehrke*

Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, and Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 11 December 2003/Accepted 3 April 2004

The alfalfa mosaic virus (AMV) RNAs are infectious only in the presence of the viral coat protein; however, the mechanisms describing coat protein's role during replication are disputed. We reasoned that mechanistic details might be revealed by identifying RNA mutations in the 3′-terminal coat protein binding domain that increased or decreased RNA replication without affecting coat protein binding. Degenerate (doped) in vitro genetic selection, based on a pool of randomized 39-mers, was used to select 30 variant RNAs that bound coat protein with high affinity. AUGC sequences that are conserved among AMV and ilarvirus RNAs were among the invariant nucleotides in the selected RNAs. Five representative clones were analyzed in functional assays, revealing diminished viral RNA expression resulting from apparent defects in replication and/or translation. These data identify a set of mutations, including G-U wobble pairs and nucleotide mismatches in the 5′ hairpin, which affect viral RNA functions without significant impact on coat protein binding. Because the mutations associated with diminished function were scattered over the 3′-terminal nucleotides, we considered the possibility that RNA conformational changes rather than disruption of a precise motif might limit activity. Native polyacrylamide gel electrophoresis experiments showed that the 3′ RNA conformation was indeed altered by nucleotide substitutions. One interpretation of the data is that coat protein binding to the AUGC sequences determines the orientation of the 3′ hairpins relative to one another, while local structural features within these hairpins are also critical determinants of functional activity.

Alfalfa mosaic virus (AMV) has a number of exceptional features that make it a useful system for studying RNA-protein interactions (reviewed in references 6 and 17). The three positive-sense genomic AMV RNAs (RNAs 1 to 3) and the subgenomic coat protein mRNA (RNA 4) are packaged individually into four bacilliform, aphid-transmitted particles. The viral genomic RNAs are infectious only in the presence of the viral coat protein (7), which binds specifically to the RNA 3′ termini. Unlike the arginine-rich RNA binding domains of the human immunodeficiency virus Tat and Rev proteins, the AMV coat protein RNA binding domain is lysine rich (3); however, it includes an arginine residue whose position and side chain identity are crucial for specific, high-affinity RNA binding (2). The AMV RNAs lack a poly(A) tail and are further distinguished from closely related plant viruses (e.g., brome mosaic virus) by the absence of a canonical CCA 3′ terminus and the inability to be charged by aminoacyl synthetases, both of which are defining features of tRNA-like ends (12).

The relationship between the formation of the AMV coat protein-RNA complex and functional activity in viral RNA translation, replication, and assembly is poorly understood. Previous biochemical characterization has defined specific binding determinants in both the RNA and coat protein (1, 3, 14, 15, 18, 26–28). An amino acid mutation that disrupts coat protein binding to the viral RNA (R17A) also blocks RNA replication (37), suggesting that coat protein binding is directly linked to replication functions. The coat protein has also been reported to enhance AMV RNA translation (20); however, the mechanism is not known.

The accumulated experimental evidence strongly suggests that coat protein binding is required for AMV RNA replication. We reasoned, therefore, that mechanistic details to explain coat protein's function(s) might be uncovered by identifying RNA mutations that maintained coat protein binding functions while either stimulating or inhibiting the accompanying replication process. To attempt to identify these mutations, we first used degenerate in vitro selection to isolate a pool of variant coat protein binding RNAs. Next, the nucleotide substitutions identified in the selected clones were transferred into viral genomic RNA 3, and viral RNA replication potential was assayed in transfected tobacco protoplasts. The degenerate in vitro selection revealed several distinct environments in the coat protein binding RNAs, including regions that were differentially invariant or highly variable, G-C rich, or found to contain G-U wobble pairs and/or covarying nucleotides. The five selected RNAs analyzed in biological activity assays all displayed severely limited viral RNA expression relating to defects in replication and/or viral RNA translation. Further experimentation revealed that the nucleotide substitutions identified in the selection caused conformational changes in the protein-free viral RNAs. The degenerate selec-

* Corresponding author. Mailing address: HST Division, MIT E25-545, 77 Massachusetts Ave., Cambridge, MA 02139. Phone: (617) 253-7608. Fax: (617) 253-3459. E-mail: Lee_Gehrke@hms.harvard.edu.

† These authors contributed equally to the work.

tion identified mutations that diminish viral RNA replication without effect on coat protein binding, generating new tools to study the role of the viral coat protein in genome activation. We interpret the data as evidence that activation of AMV replication is dependent on both the initial fold of the protein-free RNA and the final RNA fold established through viral coat protein binding.

MATERIALS AND METHODS

Pool construction. An oligonucleotide including the DNA sequence of the 39-nucleotide minimal RNA binding site for AMV coat protein (nucleotides 843 to 881 of viral RNA 4) (15), in addition to constant flanking sequences for reverse transcription and amplification, was chemically synthesized (MIT Biopolymers Lab). The DNA sequence was randomized in the coat protein binding region so that at each position there was an 80% probability of coupling the wild-type nucleotide and a 20% probability of coupling one of the other three non-wild-type nucleotides. The degenerate 39-nucleotide binding site was juxtaposed by a 42-nucleotide 5′ constant region containing the T7 RNA polymerase promoter sequence (underlined; 5′**ATAATACGACTCACTATAGGGAGAATTCT ACCATACG**GACCT3′), and a 25-nucleotide 3′ constant region (5′TCG**AGCA GCGAA*AGATCT*AGACGAT**3′). Primer binding sites used during PCR amplification are shown in bold, and a BglII restriction site is shown in italics. The 106-nucleotide DNA oligonucleotide was purified by electrophoresis into a 12% denaturing polyacrylamide gel, visualized by UV shadowing, and eluted at 4°C overnight in elution buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, and 0.1% sodium dodecyl sulfate) (8). The DNA was then extracted with phenol and chloroform and precipitated with ethanol. To generate the double-stranded pool zero DNA, the purified DNA oligonucleotide was subjected to five cycles of PCR amplification (5). A solution of DNA containing approximately 10^{15} sequences was transcribed with the MEGAshortscript T7 transcript kit (Ambion, Austin, Tex.) as described below. The RNA was gel purified as described for the DNA oligonucleotide, generating pool zero RNA.

Pool selection and amplification. Pool RNAs were selected on the basis of their ability to bind AMV coat protein with a nitrocellulose filter retention assay as described previously (14, 33), with minor modifications. For each round of selection, the gel-purified pool RNA was renatured by heating at 65°C for 3 min in REN buffer (10 mM Tris-HCl [pH 7.5], 50 mM NaCl, 3 mM $MgCl_2$, and 0.1 mM EDTA), followed by slow cooling. Next, the renatured RNA was incubated with AMV coat protein at room temperature for 1 h in 100 μl of binding buffer (10 mM Tris-HCl [pH 7.5], 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 5% glycerol). The binding reaction was filtered through HAWP nitrocellulose membranes (Millipore, Bedford, Mass.) that were prewetted with 2 ml of TMK buffer (100 mM Tris-HCl [pH 7.5], 10 mM magnesium acetate, 80 mM KCl) to retain protein-bound RNA. The filters were washed twice with 1 ml of TMK. The RNA was then eluted twice from the filter with 200 μl of a solution containing 7 M urea, 3 mM EDTA, and 100 mM sodium citrate at 65°C for 5 min.

The combined eluates were extracted with 500 μl of phenol, pH 5.2. The aqueous phase was next extracted with an equal volume of ether, and the RNA was precipitated by adding 2 volumes of water, 5 μg of linear polyacrylamide, and 3 volumes of ethanol. The eluted RNA was subjected to reverse transcription to synthesize the cDNA with Moloney murine leukemia virus reverse transcriptase (Gibco-BRL, Rockville, Md.) according to the manufacturer's instructions. The DNA was then amplified by thermal cycling over 15 cycles. The amplified DNA was purified by spin column chromatography (Clontech, Palo Alto, Calif.), transcribed into RNA, and gel purified for the next selection round. Four rounds of selection were carried out. Pool zero and pool 4 DNAs were cloned into the pGEM-T vector (Promega, Madison, Wis.), and the nucleotide sequence of the clones was determined with Sequenase 2.0 DNA polymerase (Amersham).

To monitor the progression of the pool RNAs, a competition binding assay was carried out (5). Both wild-type and pool RNAs were radiolabeled internally during transcription and renatured by heating to 65°C for 3 min, followed by slow cooling to room temperature in REN buffer. In a total volume of 100 μl of binding buffer, 0.5 μM radiolabeled wild-type and 2.5 μM radiolabeled pool RNA were mixed with 0.5 μM AMV coat protein and incubated at room temperature for 1 h. The reaction was filtered through a nitrocellulose membrane, and the RNA was eluted as above. The RNAs that bound to AMV coat protein and were retained on the membrane were eluted and separated by electrophoresis into an 8% polyacrylamide–urea gel. The pool and wild-type RNAs were electrophoretically distinguishable because the pool RNA includes constant regions needed for reverse transcription and PCR; therefore, it migrates with slower mobility. The quantity of individual RNAs was determined with a Phosphor Imager (Molecular Dynamics). The binding ratio is the ratio of pool RNA to wild-type RNA. In the early rounds, there is little enrichment for high-affinity coat protein binders; therefore, the binding ratio will be low. After successive selection rounds, the complexity of the pool decreases as the enrichment for high-affinity coat protein binders progresses, and the binding ratio increases. An increase in the binding ratio demonstrated that coat protein binding RNAs were being selected from the pool.

In vitro RNA transcription. RNAs used for the selection and gel shift experiments were transcribed in vitro with the Megashortscript T7 transcript kit (Ambion), essentially according to the manufacturer's instructions. RNAs transcribed for the selection experiments were radiolabeled by including 20 μCi of [α-^{32}P]UTP (PerkinElmer Life Sciences) in a 20-μl transcription reaction while decreasing the UTP concentration from 7.5 to 3.75 mM. Probe RNAs used in electrophoretic mobility shift assays were also radiolabeled during transcription by including 20 μCi of [α-^{32}P]UTP or [α-^{33}P]UTP in the presence of 5.6 mM each ATP, CTP, and GTP and 0.56 mM UTP. The RNAs were purified by electrophoresis into 12% denaturing polyacrylamide gels, visualized by either UV light-shadowing or autoradiography, and eluted overnight at 4°C in elution buffer (8). RNA transcripts were generated either from amplified DNA for the selection experiments or from annealed DNA oligonucleotides (19) for the nucleotide mutagenesis experiments.

RNAs used for the in vivo virus replication assays were transcribed with the mMessage T7 transcript kit (Ambion) according to the manufacturer's instructions with several modifications (18). Briefly, RNAs were 5′ capped during transcription with 6 mM $m^7G(5')ppp(5')G$ cap analog (Ambion) and 7.5 mM each ATP, CTP, and UTP and 1.5 mM GTP. The capped RNA was precipitated from a solution containing a final concentration of 3.75 M LiCl and 37.5 mM EDTA. The RNA pellet was washed with 70% ethanol, resuspended in water, quantified by spectrophotometry, and analyzed by electrophoresis into a 1% agarose gel containing ethidium bromide.

Electrophoretic mobility shift assay. Band shift assays were carried out as described previously (3, 15, 35). The band intensities were quantified with a PhosphorImager (Molecular Dynamics), and the relative K_d was defined as the protein or peptide concentration at which 50% of the RNA was bound (9). The amino acid sequence of the peptide AMV CP26 peptide, corresponding to the RNA binding domain of the viral coat protein, is SSSQKKAGGKAGKPT KRSQNYAALRK (3).

Site-directed mutagenesis of AMV RNA 3. The biological activities of genomic RNA 3 constructs containing the sequences of selected clones 14, 18, 21, 31, and 45 were tested in viral RNA replication assays in transfected tobacco protoplasts. To analyze invariant nucleotides, two additional RNA 3 constructs were generated, the G851A:C860U variant and the C869U/C870U variant. The nucleotide mutagenesis was carried out with the Quik-Change (Stratagene) kit, according to the manufacturer's recommendations. The nucleotide changes were confirmed by sequence analysis.

Viral RNA replication. Viral RNA replication assays in transfected tobacco protoplasts were carried out as described elsewhere (18). In some experiments, replication was activated with an RNA 4 clone that contained the complete coding region but lacked nucleotides 719 to 842 (18). This approach permitted the unequivocal identification, by Northern blot analysis, of newly transcribed RNA 4 generated as a result of viral RNA replication versus input activator RNA 4 molecules. To test the effects of RNA mutations on translation, wild-type RNAs 1 to 3 were transfected along with full-length RNA 4 RNA constructs containing 3′ sequences defined in the selection experiments. The activation of genomic RNA replication was used as an indirect assay of variant RNA 4 translation.

Protein analysis. Protoplasts were centrifuged at 5,000 × *g* for 5 min and resuspended in 200 μl of NT-1 extraction buffer (0.1 M $KHPO_4$, 1 mM EDTA, 10 mM dithiothreitol, and 5% glycerol). The cells were lysed by sonication, centrifuged for 10 min at 12,000 × *g* in 4°C, and 20 μl of supernatant was analyzed by electrophoresis into a 12% polyacrylamide-sodium dodecyl sulfate gel. AMV coat protein was detected by Western blotting with a polyclonal anti-AMV coat protein antibody and chemifluorescence detection (Amersham Biosciences).

AMV RNA analysis. Total protoplast RNA was isolated with Trizol reagent (Gibco-BRL) according to the manufacturer's instructions. Protoplasts from 3 ml of the suspension (approximately 500,000 cells) were sedimented and resuspended in 1 ml of Trizol reagent. Precipitated RNA was sedimented by centrifugation, and the resulting pellet was washed with 70% ethanol and resuspended in water at a concentration of 2 μg/μl.

Viral RNA replication was assayed by Northern blot analysis. The double-stranded probe used in these experiments is an NdeI-BstXI fragment of the coat

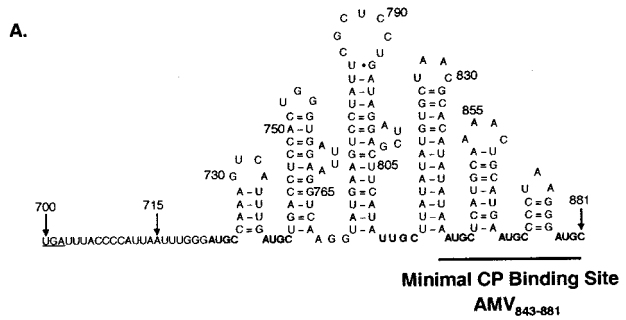

FIG. 1. (A) The secondary structure of the 3′ untranslated region of AMV RNA 3 and AMV RNA 4 is shown (25). The terminal 39-nucleotide minimal binding site for AMV coat protein is underlined. The nucleotide numbering shown corresponds to subgenomic AMV RNA 4. (B) The amino acid sequence of peptide CP26, representing the N-terminal 26 amino acids of AMV coat protein, is shown.

protein cDNA, which encompasses the majority of the protein coding region and therefore hybridizes to plus and minus strands of genomic RNA 3 and to the positive-strand subgenomic RNA 4. The DNA fragment was labeled with [α-^{32}P]dCTP (PerkinElmer) with the Prime-It II kit (Stratagene). Unincorporated nucleotide was removed with a Chromaspin gel filtration column (Clon-Tech). The heat-denatured probe was added directly to the prehybridization solution, and hybridization proceeded overnight at 65°C. The membrane was washed twice for 10 min at room temperature in 2× SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7)–0.1% sodium dodecyl sulfate and once for 10 min at room temperature in 0.2× SSC (0.03 M NaCl, 0.003 M sodium citrate, pH 7)–0.1% sodium dodecyl sulfate before exposure to X-ray film.

The conformations of the protein-free RNAs containing mutations identified in the degenerate selection were analyzed by native gel electrophoresis with the conditions described above for mobility shift assays. As a control, the RNAs were also analyzed by electrophoresis under partially denaturing conditions with the same percentage gel (10%) containing urea (Sequagel; National Diagnostics).

RESULTS

Selection of AMV coat protein binding RNAs. The AMV coat protein binds to the ≈170-nucleotide 3′ untranslated regions of the viral RNAs, wherein the 3′-terminal 39 nucleotides are highly conserved and represent a minimal coat protein binding site (Fig. 1) (15). For the selection, DNA oligomers containing this minimal coat protein binding site were randomized during synthesis so that, in theory, there was an 80% probability of coupling the wild-type nucleotide at each position and a 20% probability of coupling a non-wild-type nucleotide. Therefore, the theoretical probability of synthesizing a wild-type oligonucleotide was 0.8^{39}, which is 1 out of 6,250 molecules. In our experiments, RNA pool zero contained approximately 10^{15} sequences, providing complexity similar to that described in previous applications of the method (4). Four rounds of selection were performed. The progression of the selection was monitored by determining RNA binding ratios as described by Bartel et al. (5). The ratios showed that at each successive selection round, the pool was increasingly enriched for RNAs that bound coat protein.

Sequence analysis revealed covarying and invariant nucleotides. Following the fourth selection round, cDNAs corresponding to the selected RNAs were cloned, and the nucleotide sequences of 38 independent isolates were determined. The sequences were aligned and are presented in Fig. 2 in order of decreasing coat protein binding affinity. The relative dissociation constant for each cloned RNA, determined by band shift analysis, is shown at the right of Fig. 2. Binding was analyzed in a coat protein titration at five different concentrations, and the apparent dissociation constant was defined visually as the peptide concentration at which half of the RNA was shifted into a complex. The nucleotides in purple letters (Fig. 2) are invariant among the selected RNAs. Based on the experimental design of the degenerate selection, the random probability of finding 38 clones that share seven identical nucleotides is 0.0002. To ensure that the nucleotide conservation was not an artifact of DNA oligonucleotide synthesis, the nucleotide sequences of 10 clones from the unselected pool zero were determined. The data (not shown) confirmed the predicted nucleotide variability in the starting pool.

The red letters in Fig. 2 show the positions of nucleotide covariations, while G-U wobble pairs are indicated by underlined letters. Nucleotide covariations were identified only in hairpin 847 to 864 (Fig. 2 and 3C). It has been reported that the primer binding regions added for reverse transcription and PCR amplification during in vitro selection may influence the outcome (11); therefore, all selected RNAs were tested for binding activity in the absence of the 3′ constant region. To remove the 3′ primer binding region, the isolated cloned DNAs were linearized by digestion with the BglII restriction enzyme, which cleaved all clones immediately downstream of the randomized sequences. In none of the 38 cases did removing the 3′ primer sequence have a significant effect on coat protein binding affinity (data not shown).

We did additional controls to rule out the possible contributions of 5′ sequences representing the primer binding regions. These experiments required synthesizing exact oligonucleotides for in vitro transcription (19), and we sampled nine such constructs (clones 6, 11, 12, 14, 21, 25, 31, 45, and 46). With a single exception (clone 25), removing the 5′ and 3′ primers did not have a significant effect on coat protein binding. Coat protein binding to clone 25 RNA was dependent on the 5′ primer region; therefore, we excluded clone 25 from further analyses. Subsequent experiments focused on selected RNAs that retained the functional capacity to bind coat protein with binding constants less than 0.55 μM (Fig. 2). All of the raw sequence data are shown in Fig. 2, but further analyses described below excluded clone 25 (whose coat protein binding is dependent on 5′ primer sequences) and clones 7, 8, 13, 27, 28, 39, and 40 because of their low coat protein binding affinities. Therefore, conclusions about the significance of the selection data were based primarily on the 30 remaining clones.

Figure 3A summarizes what was known about coat protein binding to this RNA prior to the work described in this paper. The stoichiometry of peptide binding to the RNA is 2:1 (1), and the boxes enclose nucleotides that are protected from hydroxyl radical cleavage by coat protein or the coat protein peptides (2). The red and blue letters represent nucleotides that were identified as important for coat protein binding by chemical modification interference analysis (1). The localization of nucleotides in the chemical modification interference study suggested that there are two repeated, partially overlapping peptide binding sites on the RNA. The repeated se-

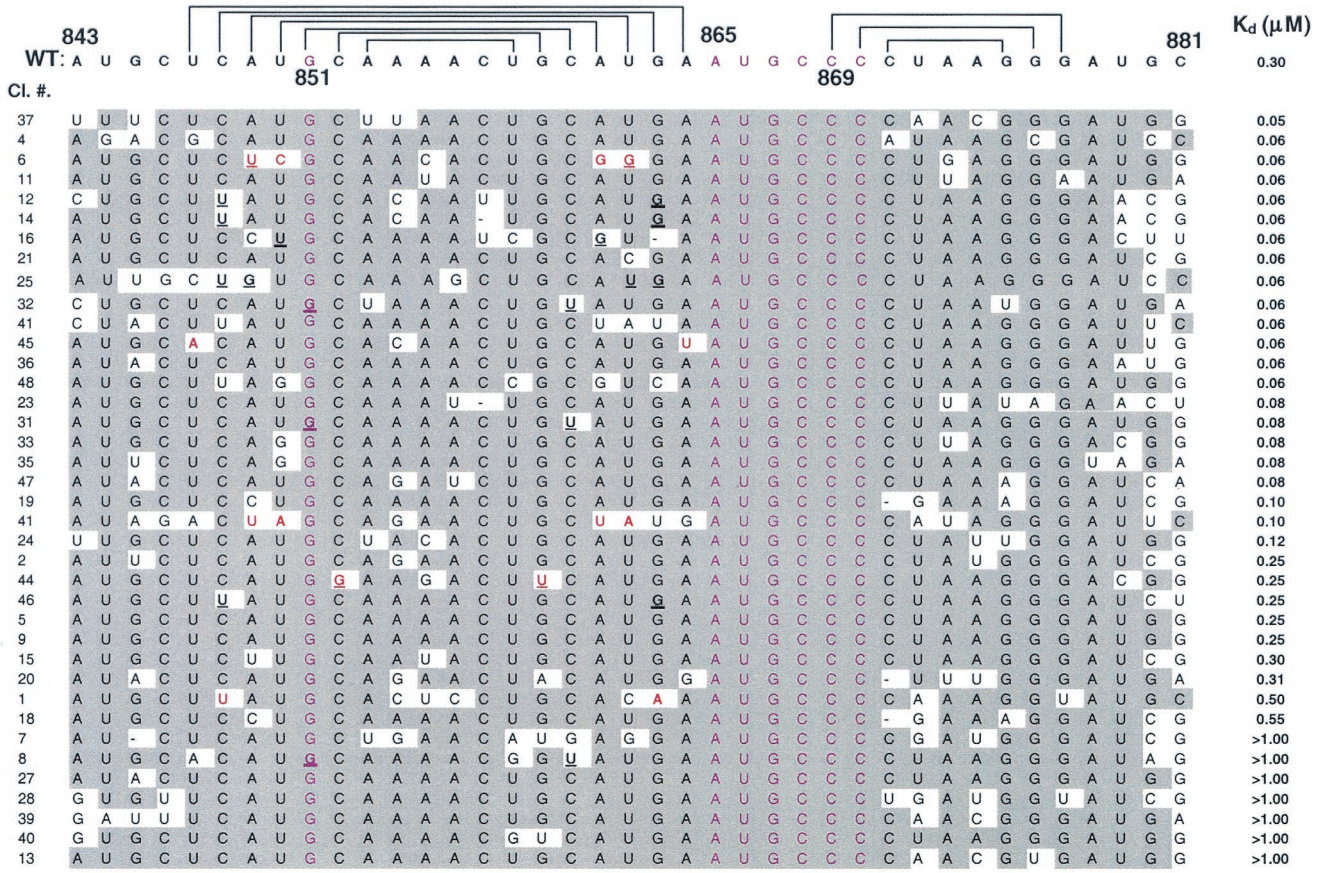

FIG. 2. Nucleotide sequences of round 4 selected clones. The selected clones are ordered, from the top down, by decreasing coat protein binding affinities. The AMV wild-type sequence at the top of the figure is noted by WT, and the sequences that are predicted to pair to form the 5′ and 3′ hairpins are joined by lines above the linear sequence. The shaded selected sequences match the AMV wild-type RNA. Nucleotides shown in purple are invariant in the selected clones. Covarying base pairs are shown in red, while the selected guanosine or uridine nucleotides that are involved in G-U pairs are underlined. The approximate K_d values for coat protein binding to the AMV wild-type and selected clone RNAs are shown to the right of the sequences. The K_d values were calculated as described in Materials and Methods.

quences form two domains with the motif U-C-hairpin-R-A-G (where R is a purine), forming binding sites 1 and 2 (Fig. 3A) (1). Recent evidence supports the significance of these two binding sites (18).

The selected invariant and highly represented nucleotides are presented on the RNA secondary map in Fig. 3B. The cluster of invariant nucleotides at nucleotides 865 to 870 (P = 0.001) includes $AUGC_{865\text{-}868}$, which was identified previously for its role in coat protein binding (14). The selection of $AUGC_{865\text{-}868}$ in these experiments serves as an important positive control for the selection method. The significance of nucleotides C869 and C870 in coat protein binding was not revealed previously; neither was the specific significance of invariant nucleotide G851. In addition to these invariant nucleotides, several other positions were very highly conserved. The invariant and highly conserved nucleotides are clustered into local environments, suggesting commonality of function. The red letters show positions that were common in 29 of 30 clones (P = 0.009), while the blue letters in Fig. 3B indicate positions that were conserved in 28 of 30 clones (P = 0.03). The apex of hairpin 847 to 864, including paired nucleotides 851 and 860 and 852 and 859, represented a G-C rich environment in the selection, while the base of the proposed 3′ hairpin 869 to 877 was similarly G-C rich. Highly represented nucleotides U844 (29 of 30 clones), C846 (28 of 30 clones), and A878 (29 of 30 clones) were also identified as being important for coat protein binding in previous chemical modification interference and hydroxyl radical footprinting experiments (1, 2), providing additional evidence in support of a two-site binding model (1, 18). Conversely, two 3′-terminal nucleotides were found to be highly variable in the selection (Fig. 3B, green letters).

While Fig. 3B illustrates the selected invariant and highly conserved nucleotides, Fig. 3C localizes the selected nucleotide covariations and G-U wobble base pairs. The covariation and wobble data are also summarized in Table 1, which identifies the individual cloned RNAs. The data present five examples of nucleotide covariation in helix 847 to 864. Nucleotide covariations provide compelling evidence in support of base pairing (21, 22), and the selection data presented here strongly suggest that helix 847 to 864 is indeed a critical determinant of AMV coat protein binding. At the same time, however, nucleotide covariation was not observed in the 3′ helix. Nucleotides C869 and C870 of the 3′ helix were invariant, while their

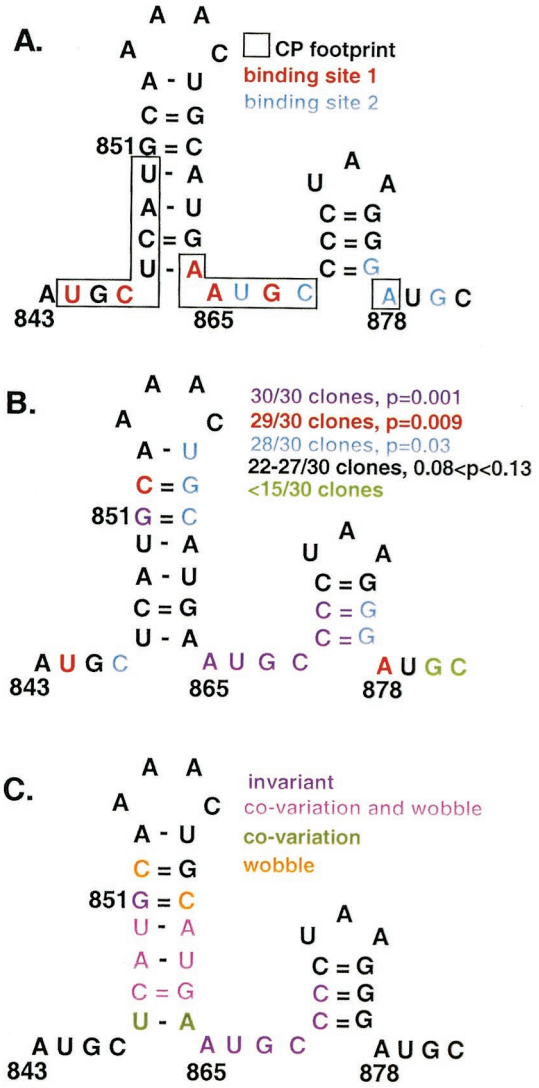

FIG. 3. (A) Summary of coat protein binding determinants in the minimal binding site RNA. Nucleotides protected from hydroxyl radical cleavage by bound coat protein or bound peptide are enclosed in boxes (2). The nucleotides shown in color were identified as important for coat protein binding by chemical modification interference analysis (1). These nucleotides form two partially overlapping binding sites with the motif U-C-hairpin-R-A-G, where R is a purine (1, 18). (B) Representation of wild-type nucleotide sequences in the selected RNA clones. The nucleotide sequence and proposed secondary structure of the randomized 39-nucleotide coat protein binding domain are shown. The numbers and color coding indicate how many of the 30 selected clones had the wild-type nucleotide at each position. The *P* values indicate the probability of identifying each distribution of nucleotide sequence as a random event. Purple, 30 of 30 clones; red, 29 of 30; blue, 28 of 30; black, nucleotides that were wild type in 22 to 27 of the 30 clones. The green letters at the 3′ end represent RNAs found to have the wild-type nucleotide sequence in fewer than 15 of 30 clones. (C) The covarying, invariant, and wobble nucleotides are shown in the secondary-structure map of the 39-nucleotide RNA. Pink, positions where both covarying and wobble substitutions were observed; green, covariation at one site where wobble changes were not found; orange, two positions where wobble substitutions were seen but covariation was not observed. Nucleotides shown in purple were invariant. The nucleotide numbering shown corresponds to AMV RNA 4.

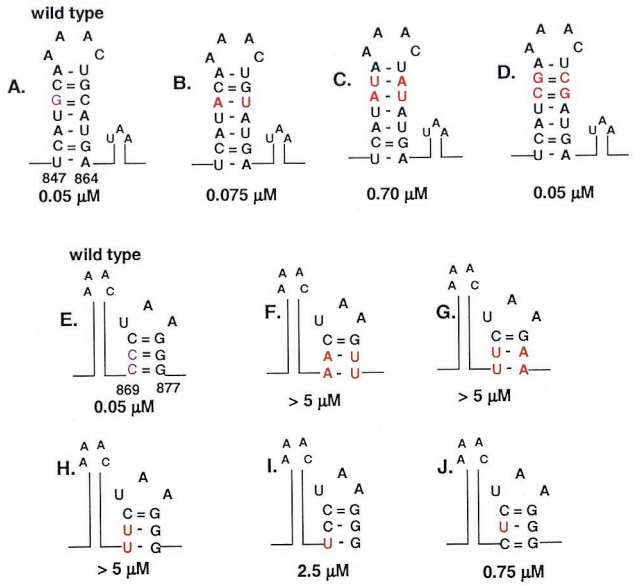

FIG. 4. Schematic representations showing mutations tested in the two hairpins of the 39-nucleotide RNA along with the affinity constants reflecting binding to AMV coat protein peptide CP26. The wild-type sequences of the 5′ and 3′ hairpins are presented in panels A and E, respectively. Panels A to D show nucleotide changes in helix 847 to 864; panels E to J show changes in helix 869 to 877.

potential base pairing partners G877 and G876 were present in 28 of 30 clones (Fig. 3B). Although there was strong nucleotide conservation in the hairpin 869 to 877 selection, there were no covariation data that could be used to define base pair features. We also identified eight examples of G-U wobble pairs occurring in six out of the seven base pairs of helix 847 to 864 (Fig. 3C and Table 1). The localization of the G-U wobble pairs correlated strongly with the region of hairpin 847 to 864 that is protected by coat protein from hydroxyl radical cleavage (compare Fig. 3A and C). These results suggest that the wobble base pairs are not detrimental to AMV coat protein binding.

Biochemical characterization of invariant nucleotides in coat protein binding. Site-directed mutagenesis and electrophoretic mobility shift assays were used to test the selection results experimentally. Peptide CP26, corresponding to the RNA binding domain of the viral coat protein (3) (Fig. 1B), was used in the mobility shift experiments. The peptide was

TABLE 1. Covariation and G:U wobble pairs in helix 847 to 864

Wild-type base pair[a]	Clone no.[b]		
	C = G/G = C	A-U/U-A	(U-G/G-U)
C852:859G			44
G851:860C			31, 32
U850:861A	6	41	16
A849:862U		41	6
C848:863G		1	12, 14, 46
U847:864A		45	

[a] This column shows the Watson-Crick base pairs in helix 847 to 864 of the 39-nucleotide RNA.
[b] The numbers correspond to the clones shown in Fig. 2.

used in place of the full-length coat protein in these experiments for several reasons. First, the behavior of the CP26 peptide mirrors that of the full-length protein in hydroxyl radical footprinting experiments (2); moreover, coat protein peptides are functional in activating viral RNA replication in mesophyll tobacco protoplasts (3). Peptide binding is also a sensitive indicator of AMV RNA conformational changes that accompany coat protein binding (3). One difference between peptide and the full-length coat protein is that the peptide binds the minimal 39-nucleotide RNA binding site with a higher affinity ($K_d \approx 0.05$ μM) than the full-length coat protein ($K_d \approx 0.3$ μM).

Invariant nucleotide G851 is located near the top of helix 847 to 864, a local environment that showed a high level of overall nucleotide conservation in the selected RNAs (Fig. 3B). To assay the importance of the absolute identity of invariant G851, we changed the G851:C860 base pair to A851:U860 (Fig. 4B). These substitutions diminished binding affinity slightly, from 0.05 μM to 0.075 μM. We also noted that the adjacent nucleotide C852 varied in only 1 of 38 clones, suggesting that it is also important for coat protein binding. To test this possibility, base pairs at G851:C860 and C852:G859 were changed to either A851:U860 and U852:A859 (Fig. 4C) or C851:G860 and G852:C859 (Fig. 4D). We found that the A851:U860/U852:A859 substitutions (Fig. 4C) had a significant effect on peptide binding, diminishing the affinity by about 10-fold. However, by maintaining stacked C=G/G=C pairs while inverting their positions, binding affinity was maintained (Fig. 4D) (1). Overall, the mutagenesis data suggest that G851 need not be absolutely conserved for high-affinity binding; rather, the results suggest that a thermodynamically stable GC-rich environment is important for high-affinity binding.

The results of substituting invariant nucleotides C869 and C870 in the 3′ helix of the binding site were tested next (Fig. 4E to J). Both base pairs C869:G877 and C870:G876 were mutated simultaneously to either purine-pyrimidine pairs (A-U) (Fig. 4F) or pyrimidine-purine pairs (U-A) (Fig. 4G). These substitutions significantly reduced coat protein binding affinities. At 5 μM AMV coat protein (the highest concentration tested), less than 15% of the RNA was bound by coat protein (data not shown). To further dissect the specific role of C869 and C870, the two nucleotides were changed to U, permitting the formation of U-G wobble base pairs (Fig. 4H). The binding experiments revealed that the K_d was greater than 5 μM, representing more than a 100-fold decrease in binding affinity compared to wild-type RNA. Finally, C869 and C870 were changed individually to U (Fig. 4I and J), leading to 50- and 15-fold decreases in coat protein binding affinities, respectively. Reusken and Bol reported previously that converting the C=G pairs at 869/877 and 870/876 to G=C pairs retained coat protein binding potential (26). Coupled with the strong nucleotide conservation in the selection, these results suggest that the stability of C-G base pairs in hairpin 869 to 877 is an important determinant of AMV coat protein binding.

Identification of RNAs that bind coat protein but are functionally inactive. To complete the second part of the screen aimed at gaining further insight into coat protein's function(s), we tested viral RNA replication with a subset of the selected RNAs. To do so, the nucleotide changes identified in the selection were transferred into infectious clones of viral RNA 3. The variant RNA 3 molecules were cotransfected with genomic RNAs 1 and 2 (to provide helicase/methyl transferase and replicase activities, respectively) and RNA 4 (to provide translated coat protein that activates replication). Viral RNA replication was assayed by Western blot to detect coat protein production (Fig. 5A) and by Northern blot (Fig. 5B) to examine subgenomic RNA 4 expression. A number of the clones identified in the selection were analyzed. Clone 14 was identified as a representative example among the selected RNAs because it includes a G-U wobble pair and binds coat protein with high affinity (Fig. 2). Clone 18 was selected because it binds coat protein with an affinity slightly less than that of wild-type RNA (500 nM versus 300 nM). Clone 21 RNA binds

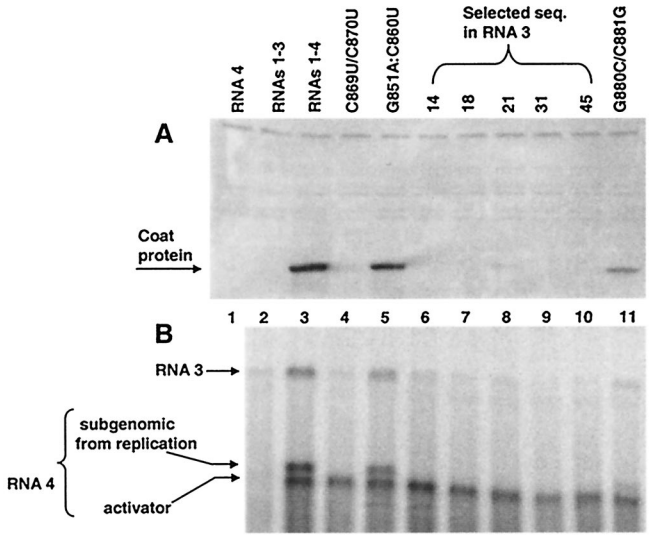

FIG. 5. Virus replication in tobacco protoplasts. Protoplasts were transfected with AMV genomic RNAs, including mutant RNA 3 molecules carrying the selected RNA sequences. (A) Western blot analysis showing coat protein expression by the infected cells. Cells were harvested 48 h after electroporation. Western blot membranes were probed with a polyclonal anti-coat protein antibody and developed with chemifluorescence technology (Amersham Biosciences). Lane 1, cells transfected with coat protein mRNA only (RNA 4); lane 2, cells transfected with genomic RNAs 1 to 3 only (noninfectious in the absence of coat protein); lane 3, wild-type viral RNAs 1 to 4 were transfected into the protoplasts. In lanes 4 to 11, all cells received wild-type genomic RNAs 1 and 2 in addition to an activator RNA 4 construct with a wild-type coat protein coding sequence and a truncated 3′ UTR wherein nucleotides 719 to 842 were deleted. The following additions were made to the reactions shown in lanes 4 to 11: lane 4, genomic RNA 3 containing the CC to UU double mutation (C869U/C870U) in the downstream hairpin (see Fig. 4H); lane 5, genomic RNA 3 containing the G851A:C860U mutation (Fig. 4B). Lanes 6 to 10, genomic RNA 3 containing the nucleotide sequences of selected clones 14, 18, 21, 31, and 45, respectively. Lane 11, genomic RNA 3 containing mutations in the 3′-terminal two nucleotides, G880C and C881G (RNA 4 numbering). We note that the faint band seen in the coat protein region of lane 4 was not seen in other experiments; it most likely represents spillover from lane 3. (B) Northern blot analysis of RNAs extracted from transfected protoplasts. In lanes 2 to 11, replication was activated with a truncated AMV RNA 4 molecule (activator) lacking nucleotides 719 to 842 (RNA 4 numbering). Subgenomic RNA 4 transcribed during viral RNA replication is labeled subgenomic. The probe used was a double-stranded DNA probe corresponding to the coat protein coating region common to RNAs 3 and 4.

coat protein with high affinity, lacks a G-U wobble substitution, and has a nucleotide mismatch in hairpin 847 to 864. Clone 31 was of interest because it has a G-U wobble pair at invariant G851 and also because there are only two nucleotide substitutions in the clone (Fig. 2). Finally, clone 45 has a covarying A-U base pair change in hairpin 847 to 864, without a G-U substitution.

Control experiments for the replication assay are presented in Fig. 5A and B, lanes 1 to 3. Both viral coat protein and subgenomic RNA 4 were detected when the cells received a complete inoculum including wild-type RNAs 1 to 4 (Fig. 5, lane 3). Coat protein generated by translation of the transfected RNA 4 alone (i.e., without replication) was not detected (lane 1); moreover, genomic RNAs 1 to 3 alone were not infectious in the absence of coat protein or coat protein mRNA (lane 2). In these experiments, replication was activated with a modified RNA 4 construct, wherein nucleotides 719 to 842 of the untranslated region were deleted, leaving the 3′ minimal coat protein binding site intact (18). The migration of this RNA (activator RNA 4) is faster than that of the full-length subgenomic RNA 4. Therefore, the detection of subgenomic RNA (Fig. 5B, lane 3) is unequivocal evidence of viral RNA replication; this RNA cannot represent transfected input activator RNA 4.

Evidence of viral RNA replication was not observed upon inserting the nucleotide sequences of selected RNA clones 14, 18, 31, and 45 into the 3′ terminus of genomic RNA 3 constructs (Fig. 5, lanes 6, 7, 9, and 10). Similarly, the RNA 3-(C869U, C870) construct was found to lack replication potential (Fig. 5, lanes 4). All of the selected RNAs bound coat protein with high affinity (clone 18 had the lowest affinity, as shown in Fig. 2); however, they did not show correspondingly robust activity in the viral RNA replication assay (Fig. 5). We noted previously that many of the selected clones included nucleotide substitutions at the extreme 3′ terminus (nucleotides 880 and 881 by RNA 4 numbering) that are likely a result of technical issues associated with the selection technique (11).

It has been reported that the 3′-terminal nucleotide is important for viral RNA function (10); therefore, we generated an additional construct to determine if RNAs containing G880C and C881G substitutions were functional. As shown in Fig. 5, lanes 11, genomic RNA 3 containing these substitutions supported viral RNA replication and the production of viral coat protein, although the levels were lower than those with the wild-type genomic RNA. The activity of the RNA 3-G880C/C881G construct was reproducibly significantly greater than that of selected clones 14,18, 21, 31, and 45, suggesting that the functional effects of the substitutions were not limited to the 3′ GC nucleotides. Clone 21 showed a low but reproducible level of viral RNA viral coat protein expression (Fig. 5A, lane 8) that reflected viral RNA replication; however, the levels of replicated or transcribed RNAs (Fig. 5B) were below the level of detection.

In addition to the clone 21 results, viral coat protein expression was observed when the G851A:C860U construct (illustrated in Fig. 4B) was tested (Fig. 5A and B, lanes 5). The G851A:C860U RNA is not itself a selected RNA; rather, it is a variant RNA generated to determine if G851 is required for coat protein binding and functional activity. The G851A:C860U substitutions diminished peptide binding affinity to the 39-mer RNA by less than twofold (Fig. 4B); the same was true for coat protein binding to the 180-nucleotide 3′ untranslated region (UTR) RNA containing the same substitutions (J. Petrillo and L. Gehrke, unpublished data). The G851A:C860U substitutions were active in the context of genomic RNA 3 (Fig. 5, lane 5), suggesting that a paired guanosine at G851 is not absolutely required for coat protein binding and viral RNA replication.

The majority of selected clones (28 of 30) retained the G851:C860 pair, suggesting that a GC-rich environment at the apex of helix 843 to 864 may be favored for coat protein binding. The remaining two clones (2 of 30) (clones 31 and 32) showed a C860U substitution, leading to the potential formation of a G-U pair. This G-U is one of a number of G-U substitutions identified in the selection (Fig. 2, Table 1). These RNAs bind coat protein with high affinity; however, RNA 3 clone 31 was inactive in viral RNA replication (see Discussion). The principal conclusion from the experiments presented in Fig. 5 is that variant RNAs containing nucleotide substitutions in the coat protein binding domain show diminished biological function in viral RNA replication while retaining coat protein binding activity.

Effects of the nucleotide substitutions on RNA 4 translation. Neeleman et al. (20) reported that coat protein binding may autoregulate AMV RNA translation. This phenomenon would manifest in our experiments if the mutations inserted into genomic RNA 3 or copied into newly transcribed subgenomic coat protein mRNA (i.e., RNA 4) decreased viral RNA replication by reducing viral protein expression as a result of limited viral mRNA translation. To distinguish between replication and transcription-translation effects, we introduced the nucleotide sequences of several selected RNAs into the 3′ UTR of the coat protein mRNAs (RNA 4) used to activate replication. If the mutations in the 3′ UTR of the RNA 4 diminished translation, the expected result would be a corresponding decrease in the activation of replication. Viral RNA replication experiments were done by transfecting wild-type genomic RNAs 1 to 3, along with RNA 4 constructs containing the selected RNA sequences.

The results (Fig. 6) revealed that viral RNA replication was activated in all experiments, as evidenced by the detection of coat protein in the Western blot. However, we observed a range of activation levels that was reproducible over several experiments. RNA 4 clone 18 yielded the lowest level of activation (Fig. 6, lane 8), while RNA 4 clone 14 was also relatively weak (Fig. 6, lane 7). This assay, which is sensitive but also indirect, provides evidence suggesting that the translation of RNA 4 constructs containing the selected RNA mutations may be diminished.

Nonetheless, the combined results shown in Fig. 5 and 6 strongly suggest the levels of translated coat protein reflected in Fig. 6 were, at least in several cases, sufficient to activate viral RNA replication. We note that the amount of coat protein translated from RNA 4 clone G851A:C860U (Fig. 6, lane 6) correlates with strong replication when the mutations were in RNA 3 (Fig. 5, lane 5). Because the amounts of coat protein translated from RNA 4 clones 21, 31, and 45 were equal to or greater than that of RNA 4 clone G851A:C860U (Fig. 6, compare lanes 13 to 15 with lane 6), it follows that the failure to detect significant viral RNA and viral coat protein in Fig. 5

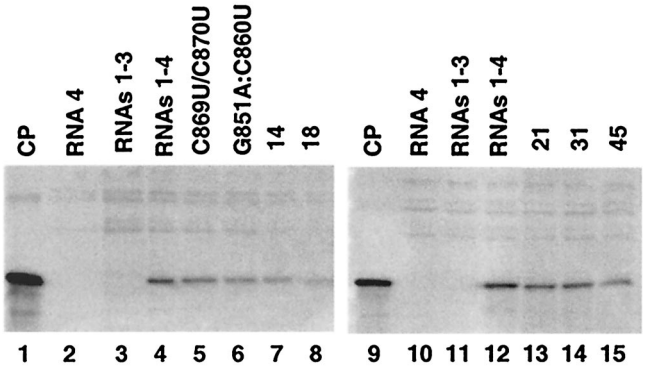

FIG. 6. Replication activated by subgenomic RNA 4 variants carrying nucleotide substitutions identified in the selection. (A) The figure shows a Western blot representing coat protein expression in protoplasts that were incubated for 48 h following transfection. Excepting the control reactions shown in lanes 1, 2, 9, and 10, all reactions contained wild-type genomic RNAs 1 to 3. The specific RNA 4 added to each transfection is identified for each lane. Lanes 1 and 9, coat protein standard; lanes 2 and 10, wild-type RNA 4; lanes 3 and 11, genomic RNAs 1 to 3 only, without added RNA 4; lanes 4 and 12, wild-type RNA control with genomic RNAs 1 to 3 plus RNA 4; lane 5, RNA 4 containing substitutions in the invariant C869 and C879 (see Fig. 4H); lane 6, RNA 4 containing substitutions G851A:C860U (see Fig. 4B); lane 7, RNA 4 containing the clone 14 substitutions; lane 8, RNA containing the clone 18 substitutions; lane 13, RNA containing the clone 21 substitutions; lane 14, RNA containing the clone 31 substitutions; lane 15, RNA containing the clone 45 substitutions.

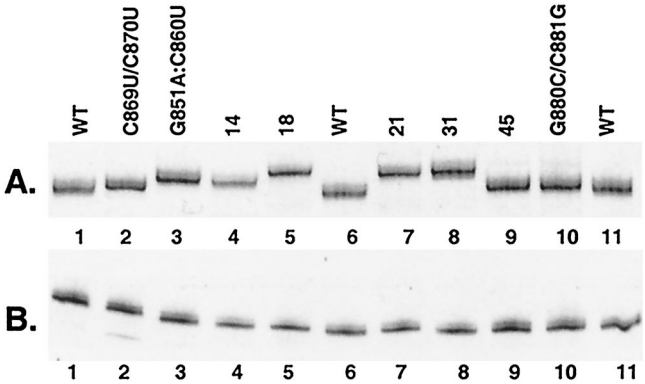

FIG. 7. Polyacrylamide gel electrophoresis of protein-free 3′-terminal 180-nucleotide fragments of AMV RNA 4 containing nucleotide substitutions from selected RNA clones. (A) Native polyacrylamide gel electrophoresis. The migration of the wild-type 180-nucleotide RNA (Fig. 1) is shown in lanes 1, 6, and 11. Lane 2, 3′ UTR containing the substitutions at the invariant C869 and C870 positions (see Fig. 4H); lane 3, 3′ UTR containing the G851A:C860U mutations at the invariant C851 base pair positions (see Fig. 4B); lane 4, 3′ UTR containing the clone 14 nucleotide substitutions; lane 5, 3′ UTR containing the clone 18 substitutions; lane 7, 3′ UTR containing the clone 21 nucleotide substitutions; lane 8, 3′ UTR containing the clone 31 substitutions; lane 9, 3′ UTR containing the clone 45 substitutions; lane 10, 3′ UTR containing the G880C, C881G substitutions. (B) RNAs were analyzed by electrophoresis into a polyacrylamide-urea gel under denaturing conditions. The RNAs loaded and the lane labeling are the same as in panel A.

(lanes 8 to 10) resulted at least in part from replication defects. On the basis of these data, we conclude that the mutations identified in the selection seem to affect both viral RNA translation and replication (or subgenomic RNA transcription) without significant impact on coat protein binding.

Nucleotide changes in the 3′ UTR alter RNA conformation. The selection and subsequent replication-translation experiments identified AMV RNAs that bind coat protein but exhibit diminished biological activity. We next returned to the question of why coat protein is required to activate AMV and ilarvirus replication. We reasoned that insight into the mechanism might be gained if it were possible to correlate the loss of a biological activity with a feature that is common among the selected RNAs but distinct from the wild-type RNA. The selected RNA mutations were scattered throughout the 3′-terminal 39 nucleotides and did not localize to a potential discrete interaction domain. Our hypothesis was, therefore, that RNA conformational changes caused by the nucleotide substitutions might be related to the changes in replication and/or translation observed in Fig. 5 and 6.

The effects of the selected nucleotide substitutions on RNA conformation were evaluated with native polyacrylamide gel electrophoresis. In addition to the C869U/C870U, G851A:C870U, and G880C/C881G substitutions, five 3′ UTR constructs containing the nucleotide sequences of selected clones 14, 18, 21, 31, and 45 were generated and analyzed in the context of the complete 180-nucleotide 3′ UTR RNA (illustrated in Fig. 1A). The results (Fig. 7A, lanes 2 to 5 and 7 to 10) show that all of the variant RNAs migrated with diminished mobility compared to the wild-type RNA (Fig. 7A, lanes 1, 6, and 11). The magnitude of the shift correlated with the position of the nucleotide changes in the hairpins. The greatest mobility shifts were observed with clones that included nucleotide changes in the upper region of hairpin 847 to 864; that is, G851A:C860U, 14, 18, 21, and 31 (Fig. 7A, lanes 3, 4, 5, 7, and 8). The smallest mobility shifts were observed with clone 45 or correlated with nucleotide changes in downstream hairpin 869 to 877: clones C869U/C870U and G880C/C881G (Fig. 7A, lanes 2 and 9 to 10). The slower migration of all of the RNAs containing nucleotide substitutions suggests that their conformations are distinct from and less compact than the wild-type RNA. As a control, the electrophoretic migration of the same RNAs was analyzed under denaturing conditions (Fig. 7B). In contrast to the results of the native gel electrophoresis, the RNAs comigrated in the polyacrylamide-urea gel, indicating that the mobility differences observed in Fig. 7A reflect altered RNA conformations.

These results suggest that all of the nucleotide substitutions tested caused RNA conformational changes; however, the magnitude of the mobility shift did not correlate uniformly with biological activity. The G851A:C860U substitution caused a significant mobility shift (Fig. 7A, lane 3), with only a modest decrease in replication activity (Fig. 5, lane 5). Conversely, clone 45 showed a small mobility shift in the native gel (Fig. 7A, lane 9) while correlating with a significant decrease in replication potential (Fig. 5, lane 10). It is conceivable that some conformational changes could be neutral with respect to functional activity while others are inhibitory; however, defining RNA conformational changes as the mechanism underlying biological activity in the clones will require further experimentation. We conclude that the nucleotide substitutions identified in the selection share the common effect of altering

the RNA conformation, which may have an impact on RNA functional activity in replication or translation.

DISCUSSION

Coat protein-mediated genome activation is unique to AMV and ilarviruses; however, despite many years of study, the mechanism is poorly understood. A novel contribution described in this report is the identification of mutations in the coat protein binding domain of the viral RNA that affect viral RNA functions without diminishing coat protein binding. The study was accomplished with a dual screen based on degenerate in vitro genetic selection followed by in vivo virus replication. The screen identified variant AMV RNAs that bind coat protein with high affinity but fail to replicate efficiently in transfected cells. The outcome is the identification of new tools to study the mechanism of coat protein-mediated genome activation in AMV and ilarviruses. The broader significance of this issue lies in understanding how positive-strand RNA viruses accomplish precise initiation of negative-strand RNA synthesis in the apparent absence of primer sequences. Another reason for probing AMV RNA-coat protein interactions is to understand how nonpolyadenylated RNAs are translated compared to cellular polyadenylated RNAs (13, 30). It is currently not known if AMV RNA-coat protein binding shares structural or functional similarities to other complexes involving nonadenylated RNAs, including the histone RNA stem-loop binding protein (36) or the rotavirus RNA-NSP3 (23) ribonucleoproteins.

Structural features of the coat protein binding domain revealed by the degenerate selection. The degenerate in vitro genetic selection process revealed significant new information about the structural features of the 39-nucleotide minimal coat protein binding site (15). In contrast to a random selection, the degenerate selection limits the complexity of the starting RNA pool, thereby maintaining the overall sequence and structural features of the RNA while introducing a level of variation sufficient to identify structural details (4, 5). The nucleotide sequences of the selected RNAs (Fig. 2) show that, on average, approximately six nucleotide substitutions were identified per clone, consistent with the expected distribution in a 20% randomization. $AUGC_{865\text{-}868}$ was invariant in the selection (Fig. 2 and 3A), a finding that is consistent with the results of a higher complexity selection (14). A second internal control relates to the finding that nucleotides U844, C846, and A877 were identified as important for coat protein binding both by chemical modification interference analysis (1) and by the selection in this study. The convergence of the selection and chemical modification interference data is consistent with a two-site binding model (Fig. 3A) (2, 18).

Several conclusions about hairpin 847 to 864 and its role as a determinant of coat protein binding can be drawn from the selection data. First, the large number of nucleotide covariations (Fig. 3C, Table 1) strongly suggests that the hairpin is essential for forming the RNA-coat protein complex. This result is consistent with mutagenesis data showing that disrupting the hairpin was accompanied by a loss of coat protein binding, while restoring the hairpin with different sequences was permissive (15, 26). An unexpected finding was that G-U wobble pairs in hairpin 847 to 864 support high-affinity coat protein binding (Fig. 3C). These wobble pairs occupied positions included the coat protein contact area suggested by hydroxyl radical footprint analysis (Fig. 3A and C) (2).

We considered the possibility that the wobble pairs surfaced in the screen simply because the probability of selecting a G-U pair is enhanced because only a single base change is needed to convert a G-C or C-G to G-U or U-G. We note, however, that G-U wobble pairs appeared in two clones as a result of nucleotide covariation rather than a single nucleotide substitution. In clones 6 and 44, both nucleotides of Watson-Crick base pairs are changed to form a G-U wobble pair. Furthermore, the single conversion of a C-G pair to an A-U contrasts with eight conversions of G-C/C-G pairs to A-U/U-A or G-U/U-G wobble pairs (Table 1). Together, these data suggest that both canonical Watson-Crick base pairs and noncanonical G-U pairs in hairpin 847 to 864 were favorable substrates for viral coat protein binding. Although coat protein bound with high affinity to variant AMV RNAs containing noncanonical G-U pairs, the native RNA sequence has little potential for forming G-U pairs in this region by short-range folding interactions. This result might suggest that the viral RNA sequence has evolved to minimize the potential for forming G-U pairs that, because of their conformational properties (34), might disrupt functional activity.

None of the loop nucleotides in either hairpin was highly conserved in the selection, suggesting that their absolute identity does not play a significant role in coat protein binding. These data are consistent with prior evidence showing that mutations in the loop nucleotide sequences did not have a significant effect on viral RNA-coat protein interactions (1, 26). The lack of involvement by the loop nucleotides distinguishes the AMV RNA-coat protein interaction from a number of other RNA-protein complexes, wherein specific binding is dependent on interactions with bulge or loop nucleotides (16, 24, 29). Although the loops do not contact the AMV peptide, some loop nucleotides are highly conserved across AMV and the ilarvirus RNAs, suggesting roles in viral life cycle functions other than coat protein binding (15).

Nucleotides C869 and C870 were invariant in the selection, and their proposed base-paired partners were very highly conserved (Fig. 3B). The importance of these nucleotides in coat protein binding was underscored by mutagenesis experiments (Fig. 4E to J). In contrast to hairpin 847 to 864, hairpin 869 to 877 showed no evidence of nucleotide covariation (Fig. 3C) in the selection. Houser-Scott et al. demonstrated that the 5′ hairpin 843 to 868 with flanking AUGC sequences is not sufficient for coat protein binding (15), and Reusken et al. reported that both hairpins shown in Fig. 3 are required for coat protein binding (26). Mutagenesis experiments showed that replacing the C-G pairs with A-U pairs resulted in a loss of coat protein binding (Fig. 4F and G). Two interpretations of the degenerate selection data can be stated. The first interpretation, taking into account the absence of covariations, is that the nucleotides are highly conserved for a role in coat protein binding that does not require base pairing. The second interpretation is that both base pairing and the stability and nucleotide identity of a CG-rich environment are essential for binding. The latter interpretation is consistent with the mutagenesis data (Fig. 4E to J) (26).

Finally, the selection data showed that the 3′-terminal GC

nucleotides were poorly represented in the degenerate selection (nucleotides 880 and 881, Fig. 3B). In fact, the occurrence of wild-type nucleotides at positions 880 and 881 of the selected clones fell below the value predicted from the degenerate synthesis parameters, possibly suggesting that there was a selection against the wild-type nucleotides. At the same time, comparative sequence analysis and data from other biochemical analyses indicate that the 3′-terminal nucleotides (especially G880) are indeed important for coat protein binding. Comparative sequence analysis demonstrates that the 3′-terminal AUGC sequence is highly conserved among AMV and ilarvirus RNAs (15), although 3′-terminal A*A*GC (31, 32) and AUG*A* (PubMed nucleotide entry AF434922) have been reported.

In a prior in vitro selection analysis, the four AUGC repeats found in the 3′-terminal 39 nucleotides were randomized (14). The results showed that the wild-type $AUGC_{878\text{-}881}$ sequence was selected in the majority of clones, suggesting importance for coat protein binding. Nucleotide deletion experiments demonstrated that the 3′-terminal cytosine can be removed without loss of coat protein binding; however, removing both the terminal G and C resulted in loss of coat protein binding in the context of the 39-nucleotide RNA (1). In summary, published data indicate that the 3′-terminal AUGC is indeed important for AMV coat protein binding. In interpreting the combined data, we considered again that the flanking constant regions required for reverse transcription and amplification might have affected the selection data in this region (11). Overall, the selection data for other regions of the 39-nucleotide RNA are consistent with the biochemical results; however, the combined findings for the 3′ positions 880 and 881 were somewhat ambiguous. As a result, we have not formed conclusions about RNA-coat protein interactions at nucleotides 880 and 881 based on the selection data.

Loss of biological activity correlates with RNA conformational changes in the protein-free RNA. The conclusions stated in this paper differ from those of previous models in raising the possibility that the functional activity of the viral RNAs is dependent on RNA structure both before (Fig. 7) and after (3) coat protein binding. Although we observed conformational differences among the viral RNAs with native polyacrylamide gels (Fig. 7), this technique does not provide information at the level of local conformation within the hairpins. However, several observations, taken together, are consistent with the possible significance of conformational changes in explaining the differential activities of the clones. The wild-type RNA, clone G851A:C860U RNA, and clone 31 RNA have G=C, A-U, and G-U sequences at the 851:860 base pair, respectively. Both the wild-type and clone G851A:C860U RNAs are active in the replication assay (Fig. 5, lanes 3 and 5); however, clone 31 (G-U) is inactive (Fig. 5, lane 9).

G-U pairs have nearly the same thermodynamic stability as A-U pairs; however, they have unique chemical and structural properties (34). Because clone 31 has, in addition to the G-U pair, a 3′-terminal C881G substitution, we generated and tested an RNA construct wherein C860U was the only substitution. The results (J. Petrillo and L. Gehrke, unpublished data) demonstrated that the single nucleotide change from C860 to U860 was sufficient to cause the RNA conformational change and to diminish replication levels to undetectable levels. Because G851 and U860 are both active in the context of G=C and A-U pairs (Fig. 5, lanes 3 and 5), the loss of activity with the G-U substitution is likely to be an effect of RNA conformation rather than being related to the nucleotide identity.

We interpret the results presented here as evidence that the mechanism of AMV genome activation centers on coat protein-mediated RNA conformational changes that organize the 3′ terminus of the RNA for replication functions. Although the selected RNA mutations did not affect coat protein binding, their impact may have been to perturb local structure or to interfere with RNA conformational changes that accompany coat protein binding. We propose that coat protein binding to the AMV RNA AUGC sequences may determine the orientation of the 3′ hairpins, whose local folding is optimized, by the nature of the primary sequence determinants, for functional activity in replication or translation.

ACKNOWLEDGMENTS

We thank David Bartel for helpful discussions about the degenerate selection method. We also thank John McQuade of Pine Cliff Consulting (Framingham, Mass.) for performing the statistical analyses.

This work was supported by NIH award GM42504.

REFERENCES

1. **Ansel-McKinney, P., and L. Gehrke.** 1998. RNA determinants of a specific RNA-coat protein peptide interaction in alfalfa mosaic virus: conservation of homologous features in ilarvirus RNAs. J. Mol. Biol. **278:**767–785.
2. **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.
3. **Baer, M., F. Houser, L. S. Loesch-Fries, and L. Gehrke.** 1994. Specific RNA binding by amino-terminal peptides of alfalfa mosaic virus coat protein. EMBO J. **13:**727–735.
4. **Bartel, D. P., and J. W. Szostak (ed.).** 1994. Study of RNA-protein interactions by in vitro selection. IRL Press, New York, N.Y.
5. **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.
6. **Bol, J. F.** 1999. Alfalfa mosaic virus and ilarviruses: involvement of coat protein in multiple steps of the replication cycle. J. Gen. Virol. **82:**947–951.
7. **Bol, J. F., L. Van Vloten-Doting, and E. M. J. Jaspars.** 1971. A functional equivalence of top component a RNA and coat protein in the initiation of infection by alfalfa mosaic virus. Virology **46:**73–85.
8. **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.
9. **Calnan, B. J., B. Tidor, S. Biancalana, D. Hudson, and A. D. Frankel.** 1991. Arginine-mediated RNA recognition: the arginine fork. Science **252:**1167–1171.
10. **Degraaff, M., C. Thorburn, and E. M. J. Jaspars.** 1995. Interaction between RNA-dependent RNA polymerase of alfalfa mosaic virus and its template: Oxidation of vicinal hydroxyl groups blocks in vitro RNA synthesis. Virology **213:**650–654.
11. **Ellington, A. D., and J. W. Szostak.** 1990. In vitro selection of RNA molecules that bind specific ligands. Nature **346:**818–822.
12. **Fechter, P., J. Rudinger-Thirion, C. Florentz, and R. Giege.** 2001. Novel features in the tRNA-like world of plant viral RNAs. Cell. Mol. Life Sci. **58:**1547–1561.
13. **Hentze, M.** 1997. eIF4G: a muiltipurpose ribosome adapter? Science **275:** 500–501.
14. **Houser-Scott, F., P. A. Ansel-McKinney, J. M. Cai, and L. Gehrke.** 1997. In vitro genetic selection analysis of alfalfa mosaic virus coat protein binding to 3′-terminal AUGC repeats. J. Virol. **71:**2310–2319.
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.
16. **Jaffrey, S. R., D. J. Haile, R. D. Klausner, and J. B. Harford.** 1993. The interaction between the iron-responsive element binding protein and its cognate RNA is highly dependent upon both RNA sequence and structure. Nucleic Acids Res. **21:**4627–4631.
17. **Jaspars, E. M.** 1999. Genome activation in alfamo- and ilarviruses. Arch. Virol. **144:**843–863.

18. **Laforest, S. M., and L. Gehrke.** 2004. Spatial determinants of the alfalfa mosaic virus coat protein binding site. RNA **10:**48–58.
19. **Milligan, J. F., D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck.** 1987. Oligoribonucleotide synthesis using T7 polymerase and synthetic DNA templates. Nucleic Acids Res. **15:**8783–8798.
20. **Neeleman, L., R. C. Olsthoorn, H. J. Linthorst, and J. F. Bol.** 2001. Translation of a nonpolyadenylated viral RNA is enhanced by binding of viral coat protein or polyadenylation of the RNA. Proc. Natl. Acad. Sci. USA **98:** 14286–14291.
21. **Noller, H. F., and C. R. Woese.** 1981. Secondary structure of 16S ribosomal RNA. Science **212:**403–411.
22. **Pace, N. R., B. C. Thomas, and C. R. Woese.** 1999. Probing RNA structure, function, and history by comparative analysis, p. 113–141. *In* R. F. Gesteland, T. R. Cech, and J. F. Atkins (ed.), The RNA world, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
23. **Piron, M., P. Vende, J. Cohen, and D. Poncet.** 1998. Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly(A) binding protein from eIF4F. EMBO J. **17:**5811–5821.
24. **Puglisi, J. D., L. Chen, S. Blanchard, and A. D. Frankel.** 1995. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. Science **270:**1200–1203.
25. **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.
26. **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.
27. **Reusken, C. B. E. M., 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.
28. **Reusken, C. B. E. M., L. Neeleman, F. T. Brederode, and J. F. Bol.** 1997. Mutations in coat protein binding sites of alfalfa mosaic virus RNA 3 affect subgenomic RNA 4 accumulation and encapsidation of viral RNAs. J. Virol. **71:**8385–8391.
29. **Romaniuk, P. J., P. Lowary, H.-N. Wu, G. Stormo, and O. C. Uhlenbeck.** 1987. RNA binding site of R17 coat protein. Biochemistry **26:**1563–1568.
30. **Sachs, A. B., P. Sarnow, and M. W. Hentze.** 1997. Starting at the beginning, middle, and end: translation initiation in eukaryotes. Cell **89:**831–838.
31. **Scott, S. W., M. T. Zimmerman, and X. Ge.** 2003. Viruses in subgroup 2 of the genus Ilarvirus share both serological relationships and characteristics at the molecular level. Arch. Virol. **148:**2063–2075.
32. **Scott, S. W., M. T. Zimmermann, X. Ge, and D. J. Mackenzie.** 1998. The coat proteins and putative movements proteins of isolates of Prunus necrotic ringspot virus from different host species and geographic origins are extensively conserved. Eur. J. Plant Pathol. **104:**155–161.
33. **Tuerk, C., and L. Gold.** 1990. Syst. evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science **249:**505–510.
34. **Varani, G., and W. H. McClain.** 2000. The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. EMBO Rep. **1:**18–23.
35. **Weinberger, J., D. Baltimore, and P. Sharp.** 1986. Distinct factors bind to apparently homologous sequences in the immunoglobulin heavy-chain enhancer. Nature **322:**846–848.
36. **Williams, A. S., and W. F. Marzluff.** 1995. The sequence of the stem and flanking sequences at the 3′ end of histone mRNA are critical determinants for the binding of the stem-loop binding protein. Nucleic Acids Res. **23:**654–662.
37. **Yusibov, V., and L. S. Loesch-Fries.** 1998. Functional significance of three basic N-terminal amino acids of alfalfa mosaic virus coat protein. Virology **242:**1–5.