In Vitro Selection of Bacteriophage Φ29 Prohead RNA Aptamers for Prohead Binding*

(Received for publication, August 26, 1997, and in revised form, November 19, 1997)

Feng Zhang‡ and Dwight Anderson¶

From the Departments of ‡Genetics and Cell Biology and ¶Microbiology and Oral Science, University of Minnesota, Minneapolis, Minnesota 55455

Prohead RNA (pRNA) of the Bacillus subtilis bacteriophage Φ29 is needed for in vitro packaging of DNA-gp3 product 3 (DNA-gp3). Residues 22–84 of the 174-base pRNA bind the portal vertex of the prohead, the site of DNA packaging. To define the nucleotides of pRNA needed for prohead binding and DNA-gp3 packaging and to seek biologically active variants of pRNA, segments of pRNA were randomized to obtain vast repertoires of RNA molecules. RNA aptamers, ligands best suited for prohead binding, were obtained by multiple rounds of in vitro selection. Evolution of pRNA aptamers was followed by a competition binding assay and nucleotide sequencing, and mutants were tested for DNA-gp3 packaging. Aptamers selected following randomization of the E stem and loop and a part of the C-E loop that were active in DNA-gp3 packaging were invariably wild-type. DNA-gp3 packaging activity also required nucleotides G82 and C98 that form base pairs intermolecularly with C97 and C48 to produce a novel hexameric oligomer of pRNA. The only mutant aptamers that retained full DNA-gp3 packaging activity showed changes of the U residues at positions 81, 84, and 85 of the D loop. Thus, the in vitro selections essentially recapitulated the natural evolution of pRNA.

The double-stranded DNA-gene product 3 (DNA-gp3) complex of the Bacillus subtilis bacteriophage Φ29 is packaged efficiently into a prohead with the aid of the ATPase gp16 and ATP hydrolysis in a completely defined in vitro system (1). A unique 174-base Φ29-encoded RNA, termed prohead RNA (pRNA), is present on the portal vertex (head-tail connector) of the prohead and is an essential constituent of the DNA packaging machine (2). pRNA is hypothesized to bind a supercoiled DNA-gp3-gp16 complex to link the DNA and prohead (3), recognize the left end of the DNA-gp3, which is packaged first (4), and unite with gp16 to form the DNA translocating ATPase (5).

The secondary structure of the pRNA was established by a phyllogenetic analysis (6). pRNA binding to proheads is specific, rapid, and irreversible in the presence of 10 mM Mg2+. Proheads protect nucleotides 22–84, 5′ to 3′, of pRNA from ribonuclease attack, and the use of site-directed mutants of pRNA has identified elements and sequences of pRNA that are required for prohead binding and DNA-gp3 packaging (7–9). The mutant studies also identified a pseudoknot involving a nine-membered bulge loop and a five-base hairpin loop in pRNA that is essential in DNA-gp3 packaging (9). Recently, the pseudoknot has been shown to be an intermolecular interaction, requiring just two base pairs, that links six identical molecules of pRNA into a structure that is positioned on the portal vertex of the prohead.2

The structure of pRNA as it interacts with the proteins of the packaging machine and DNA-gp3 is integral to understanding the mechanism of DNA packaging. To better define the sequence and structural elements of pRNA essential for prohead binding, a 62-base segment of the prohead binding domain (residues 30–91) was partially randomized, and pRNA aptamers, high affinity ligands that bind the prohead, were selected in vitro from a large pool of RNA molecules. Subsequently, 23 bases of this domain (residues 45–62 and 81–85) that include the intermolecular pseudoknot were completely randomized in a second in vitro selection analysis that was based on prohead binding and included assays of DNA-gp3 packaging. The results demonstrated more precisely the pRNA sequences and elements essential for prohead binding and DNA-gp3 packaging and identified the intermolecular base pairing required for pRNA oligomerization. In general, the selections recapitulated the natural evolution of pRNA.

EXPERIMENTAL PROCEDURES

Construction of RNA Libraries—The in vitro selections of pRNA aptamers that bind the prohead were patterned after the method termed systematic evolution of ligands by exponential enrichment (SELEX) (10), which utilizes variation, selection, and replication. A vast repertoire of RNA molecules produced from a randomized template are bound to a target, selected molecules are amplified as DNA that is competent for in vitro transcription, and the newly transcribed RNAs, enriched for better binding sequences, are subjected to selection to begin the next cycle.

The DNA oligonucleotides used as templates in initial transcription for SELEX experiments I and II were purchased from Oligos Etc. Inc. In SELEX I, the randomized region of the template was based on positions 30–91 of the wild-type sequence (see Fig. 1, A and C) with a 16% mutation rate. The 110-nucleotide DNA template for the original RNA pool, with degenerate nucleotides in the 62-base prohead binding domain represented by N, was as follows: 5′-GGGGCCTTTG TCTAGATATAATACGACTCACTATAGGGCCTTTGTCATGT-3′ (the Xbal site is indicated with an asterisk, and the T7 promoter is underlined). Primer 1 (5′-GC*TCTAGATATAATACGACTCACTATAGGGCCTTTGTCATGT-3′ (the Xbal site is indicated with an asterisk, and the T7 promoter is underlined)) was used for making the double-stranded DNA template and for PCR amplification. Primer 2 (5′-AATGCA*GGGGCCTTTGTCTAGATATAATACGACTCACTATAGGGCCTTTGTCATGT-3′ (the Xbal site is indicated with an asterisk, and the T7 promoter is underlined)) was used for PCR amplification.

This paper is available on line at http://www.jbc.org
TCATG-3' (the PstI site is indicated with an asterisk) was used for reverse transcription and PCR amplification.

The 120-nucleotide template for the original RNA pool of SELEX II in the complete randomization of pRNA residues A45 through G62 and U81–U85 (see Fig. 1, A and D), with randomized nucleotide positions represented by N, was as follows: 5'-TTAGGAAAGT AGCGTGCACT TTTGCCATGA TTGACNNNNN ATCAACAAAG TATGTGGGNN NNNNNNNNN NNNNNNTAAT CCCCAACATA CACATGACAA TGGAAGTACC GTACCATTCC-3'. Primer 1 (5'-CCGG*AATTCTAA- TACGACTCACTATA*GGAATGGTACGGTACTT-3' (the EcoRI site and the transcription start site are indicated with an asterisk, and the T7 promoter is underlined)) was used for making the double-stranded DNA template and for PCR amplification. Primer 2 (5'-CGCG*GATCC*TTAGGAAAGTAGCGTGC-3' (the BamHI and DdeI sites are indicated with an asterisk)) was used for reverse transcription and PCR amplification.

Fig. 1. Secondary structure of pRNA, mutant pRNAs, and RNAs upon which templates for SELEX experiments were based. A, secondary structure of the 120-base form of wild-type pRNA (6). The helices are designated A, C, D, and E. Shading shows the RNase footprint of the prohead on pRNA (7). The boxes and line identify a pseudoknot, helix G (9), that was recently shown to be an intermolecular interaction that links identical molecules of pRNA into dimers and hexamers needed for DNA-gp3 packaging in vitro. B, compensatory changes to test the G helix interactions shown in A. C, RNA structure upon which the template for the original RNA pool of SELEX I was based. Shading shows the 62-nucleotide domain (nucleotides 30–91) that was partially randomized with a mutation rate of 16%. D, pRNA that served as the basis for the template for the original RNA pool of SELEX II. Shading marks residues 45–62 and 81–85, which were completely randomized.
dedicated with an asterisk) was used for reverse transcription and PCR amplification.

The 110-base single-stranded DNA template (SELEX I) and the 120-base single-stranded DNA template (SELEX II) were annealed to their respective primer-2 oligonucleotides and extended to form double-stranded DNA with five cycles of PCR. The double-stranded DNA templates in SELEX I and II were transcribed to produce original RNA pools with complexities of $2.5 \times 10^4$ and $7 \times 10^{13}$ molecules, respectively. T7 transcription was carried out as described (7) in a volume of 800 µl. T7 RNA polymerase was purchased from Ambion, Inc., and the reaction buffer was that of Life Technologies, Inc.

**In Vitro Selection—**pRNA-free proheads were produced by isopropyl-1-thio-β-D-galactopyranoside induction of *Escherichia coli* strain HMS174(DE3) (pAR7-8.65-10) as described (7, 8). The binding affinity of proheads to pRNA was titrated by the use of $^{32}$P-labeled wild-type pRNA (8, 9). The ratio of molecules of RNA to proheads was 100. For each selection, 200 µg of proheads (approximately $10^{12}$) were used. Competitive binding was performed in 400 µl of TM (0.05 M Tris-HCl (pH 7.5), 0.01 M MgCl$_2$) buffer with 180 µg of yeast tRNA at room temperature for 30 min. RNA bound to proheads was separated from free RNA by centrifugation in a 5–20% sucrose gradient in the Beckman SW55 rotor at 35,000 rpm for 30 min at 4 °C. The prohead band was isolated, and 1 volume of H$_2$O was added to dilute the sucrose. RNA was recovered by phenol extraction and ethanol precipitation in the presence of 0.5 M NaAC. cDNA was produced by Superscript RNase H reverse transcriptase (Life Technologies, Inc.) and amplified by 30 cycles of PCR in four 100-µl reactions, and the DNA was transcribed to start another round of selection. For reverse transcription, the selected RNA was annealed to 1 µg of primer 2 by heating in water for 3 min at 100 °C and then cooling the sample in ice water. The reverse transcriptase was from Life Technologies, Inc., and reverse transcription was carried out in 40 µl as recommended by the manufacturer. PCR cycle conditions, adapted from Longel et al. (11), were as follows: 4 min at 94 °C for initial denaturation, 1 min at 94 °C, 2 min at 45 °C, 3 min at 72 °C, and a final 7 min at 72 °C. Wild-type pRNA was transcribed from plasmid PRT71 digested by DdeI as described (7, 8). DNAs from different cycles were cloned into pUC19 at XbaI and PstI sites (SELEX I) or EcoRI and BamHI sites (SELEX II).

**Prohead Binding Assay—**The ability of RNA pools from each selection to compete with wild-type pRNA for prohead binding was used as a measure of relative binding affinity. The competition binding assay was carried out as described (8, 9). Briefly, 1.4 pmol of 120-residue wild-type $^{32}$PpRNA was mixed with 14 pmol of an unlabeled RNA from each selection pool and 60 pmol of tRNA. 0.35 pmol of proheads were added to the RNA, and the fraction of $^{32}$PpRNA bound was measured by filtration. A competitor activity for each RNA in the competition assay was derived by fitting values for the fraction of wild-type $^{32}$PpRNA bound to the standard curve described by

$$B_C = \frac{1}{1 + C}$$  \hspace{1cm} (Eq. 1)

and solving for $C$, where $B_C$ is the fraction bound and $C$ is the moles of competitor RNA per mole of $^{32}$PpRNA. Values were normalized to the no competitor control ($C = 0$) and the wild-type pRNA competitor control ($C = 100$) to give

$$C_C = \frac{B_C - B_{10}}{B_C - B_{no}} \times 100$$  \hspace{1cm} (Eq. 2)

where $C_C$ compares the competitor activity of the test RNA to the wild-type pRNA, $B_{10}$ is the fraction of $^{32}$PpRNA bound when a 10-fold excess of wild-type pRNA is used as competitor, $B_C$ is the fraction of $^{32}$PpRNA bound with no competitor RNA, and $B_{no}$ is the fraction of $^{32}$PpRNA bound when a 10-fold excess of test competitor is added.

**Cloning and Sequencing of Variants—**DNA amplified from different cycles of selection was cloned into the PstI and XhoI sites (SELEX I) of pUC19 and the EcoRI and BamHI sites (SELEX II). Clones with inserts were identified by negative α complementation. The plasmids were sequenced with Sequenase 2.0 (United States Biochemicals) using a protocol adapted from that provided by the manufacturer.

**DNA Packaging Assays—**In vitro DNA-gp3 packaging was carried out as described (8, 12). Briefly, wild-type or mutant pRNAs were bound to RNA-free proheads, and the reconstituted proheads were then added to DNA-gp3, gp16, and ATP. After 30 min at 25 °C, unpackaged DNA was digested with DNase I, EDTA was added, and packaged DNA was extracted from filled heads and quantified by agarose gel electrophoresis.

**RESULTS AND DISCUSSION**

**Structure of pRNA and Rationale for in Vitro Selection of pRNA Aptamers for Prohead Binding—**Fig. 1A shows the proposed secondary structure of the 120-base form of the 174-residue pRNA (6) that is necessary and sufficient for DNA-gp3 packaging in the defined in vitro system (2, 12). The shaded region that marks nucleotides 22–84 is the footprint of proheads on pRNA generated with the ribonucleases A, T1, and V1 (7). Analyses utilizing oligonucleotide-directed mutant pRNAs show that both the sequence and the secondary structure of residues 40–91 are important for prohead binding and that elements of the A helix formed from residues 1–28 and 92–117 are needed for DNA packaging functions other than prohead binding (8, 13). The secondary and tertiary structures of the prohead binding region of pRNA have been probed by measuring binding of mutant pRNAs to RNA-free proheads and in vitro of packaging of DNA-gp3. A truncated pRNA of 62 residues (residues 30–91) retains prohead binding activity but is insufficient for DNA-gp3 packaging (9).

A pseudoknot in pRNA inferred from the phylogenetic studies was confirmed by constructing mutants to change positions $^{45}$AACC$^{48}$ to GCGA (mutant F6), $^{52}$GGUGU$^{55}$ to UCGR (mutant F7), or $^{55}$AACC$^{62}$ to GCGA and $^{52}$GGUGU$^{65}$ to UCGR (mutant F6/F7) to restore the hypothetical base pairing of the G helix (9) (Fig. 1B). Neither F6 nor F7 pRNAs could reconstitute proheads for *in vitro* DNA-gp3 packaging, whereas the F6/F7 double mutant pRNA was as effective as wild-type pRNA in reconstituting proheads for DNA packaging. Moreover, proheads reconstituted with either F6 or F7 pRNA could not promote in vitro phage assembly in a prohead-defective extract, whereas the F6/F7 pRNA was as effective as wild-type pRNA in reconstituting proheads for phage assembly. Recently, a mixture of F6 and F7 pRNAs was found to be as active in DNA-gp3 packaging as the double mutant F6/F7 and wild-type pRNA, showing that the pseudoknot is intermolecular rather than intramolecular. This intermolecular base pairing results in the formation of an oligomer from six identical molecules of pRNA that is essential for efficient DNA-gp3 packaging. To define
the specific nucleotides of pRNA needed for prohead binding and formation of the intermolecular pseudoknot, segments of pRNA were partially or completely randomized, and RNA aptamers for prohead binding were obtained by multiple rounds of in vitro selection.

Partial Randomization of the Prohead Binding Domain of pRNA—The 62-base prohead binding domain (residues 30–91) shown as the shaded region in the template of Fig. 1C was partially randomized with a mutation rate of 16% as described under “Experimental Procedures.” The DNA template utilized the T7 promoter, and an initial pool of RNA molecules with a complexity of $5 \times 10^{14}$ molecules was produced. RNA pools were obtained from seven rounds of in vitro selection by prohead binding and separated from unbound RNA by isolation of the prohead-pRNA complexes in sucrose density gradients. cDNA was produced, amplified by PCR, cloned, and sequenced.

The ability of pools of RNA from the various rounds of in vitro selection to compete with the binding of $[32P]$wild-type pRNA to proheads is shown in Fig. 2. From the fifth round, RNA pools were equivalent to wild-type pRNA as competitors.

**Fig. 3.** Sequences of 93 pRNA clones from selection cycle five of SELEX I after partial randomization of the 62-nucleotide prohead binding domain.

The 62-base wild-type sequence is shown above the mutant sequences. Numbers under the sequence are the positions in pRNA. Capital letters above the wild-type sequence designate the helices and loops of pRNA. Deviations from the wild-type nucleotide at each position are given, and conservation of the wild-type base is indicated with a dash. Circled bases designate covariation within helices to maintain base pairing.
cloned, and sequenced. From the fifth round, RNA pools were bound to proheads. cDNA was produced, amplified by PCR, (residues 45–62 and 81–85) (Fig. 1 needed for DNA-gp3 packaging assays. Twenty-three bases focused on the sequences and elements of pRNA crucial for packaging, and therefore, this study was limited to prohead binding. In the mutant pRNAs 4-1 (Table II) and 5-1, which pair with 62G, 61A, and 60C, respectively) and in the residues 53U, 54U, 55G, and 56A of the E loop.

The 93 sequences are given in Fig. 3. Bases in boldface type are needed for DNA-gp3 packaging assays. Fifty-nine of these sequences showed changes at practically every position of the prohead binding domain (not shown). The contrary, in the fifth round, changes in residues 46–48 of the C-E loop and 81–84 of the D loop, which are involved in intermolecular base pairings that produce dimers and hexamers of pRNA, were infrequent in 93 sequences, and most of the RNA sequences are wild-type at these positions and at 46A and 84U which may contribute to the intermolecular interaction. Also, changes were rare in three of four base pairs of the E helix (residues 49C, 50U, and 51G, which pair with 62G, 61A, and 60C, respectively) and in the residues 53U, 54U, 55G, and 56A of the E loop.

The nine sequences of residues 45–62 that were not wild-type were of particular importance in prohead binding and DNA-gp3 packaging (8). The 62-base pRNA is not active in DNA-gp3 pseudoknot formation, i.e. pseudoknot formation dominated by the wild-type is dramatic. From cycle three, only 2 of 18 clones were wild-type. From cycle four, 74 of 90 clones (82%) were wild-type, and from cycle five, 190 of 226 clones (84%) were wild-type. Representative mutant sequences of the randomized regions from cycles 4 and 5 are given in Tables II and III, respectively.

Tables II and III illustrate that residues 45–62 are wild-type in 22 of 31 sequences. On the contrary, changes occurred in residues 81–85 in all but three of the mutant pRNAs. The base 45–62 interval includes the C-E bulge loop and the E stem and loop. Bases 81–85 make up the D loop, residues of which form base pairs intermolecularly with the C-E loop in pRNA oligomerization (addressed below). Substantial change in the D loop shows that prohead binding per se does not require pseudoknot formation, i.e. pseudoknot formation may follow prohead binding. Moreover, mutant pRNA F6, which cannot form the pseudoknot, still has 20% prohead binding competitor relative to wild-type pRNA (9). The results suggest that the intermolecular pseudoknot is not important for the initial interaction of pRNA with the prohead but is important for subsequent interactions critical to DNA-gp3 packaging activity.

The nine sequences of residues 45–62 that were not wild-type showed abundant change (Tables II and III). Eight of these sequences showed covariation in base pairs of the E helix, suggesting that maintenance of this element is important in prohead binding. In the mutant pRNAs 4-1 (Table II) and 5-1, 5-14, and 5-15 (Table III), covariation occurred in two of four base pairs of the E stem, and in the mutant 4-10 (Table II), covariation occurred in all four bases of the E stem.

Folding of pRNA mutant sequences was predicted by the computer algorithm MFOLD (14). Minor, even single, base

### TABLE I

| 5S | 5U | 32A | 33I | 34G | 35U | 36D | 37O | 38G | 39O | 40G | 41A | 42U |
|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| A  | 1  | 1  | 1  | 4  | 6  | 7  | 5  | 8  | 3  | 8  | 7  | 0  |
| U  | -  | -  | 1  | 3  | 4  | 2  | 6  | 2  | 0  | -  | -  | -  |
| C  | 6  | 2  | 1  | 2  | 1  | 5  | 5  | 7  | 3  | 2  | 0  | 4  |
| G  | -  | 0  | 0  | 1  | -  | 16 | 3  | -  | -  | -  | 6  | 7  |

**Summary of base changes at each position of the partially randomized 62-base prohead binding domain of pRNA in SELEX I**

The 93 sequences are given in Fig. 3. Bases in boldface type are needed for DNA-gp3 packaging assays. Fifty-nine of these sequences showed changes at practically every position of the prohead binding domain (not shown). Amounts of [32P]pRNA bound were as follows: input, 14964 cpm; no proheads, 142 cpm. The competition activities, which normalize the [32P]pRNA fraction bound with wild-type pRNA competitor (100%) and no competitor (0%) (8, 9), were as follows: 5S rRNA, 2%; pRNA, 100%; cycle 0, 4%; cycle 1, 12%; cycle 2, 16%; cycle 3, 46%; cycle 4, 83%; cycle 5, 97%; and cycle 6, 91%.

![Fig. 4. Competition filter binding with RNA pools of SELEX II.](https://www.jbc.org/)

**TABLE II**

| 56A | 57U | 59U | 59P | 61O | 62A | 62U | 62G | 63C | 64C | 65A | 65C | 66A | 67C |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| A  | 6  | 2  | 0  | 0  | 0  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| U  | 0  | 4  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| C  | 6  | 2  | 1  | 1  | 1  | 6  | 3  | 0  | 0  | 0  | 0  | 0  | 0  |
| G  | 3  | 1  | 1  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |

**TABLE III**

| 69U | 70A | 71C | 72I | 73D | 75O | 75U | 77U | 78G | 79A | 80U | 81D |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| A  | 6  | 2  | 0  | 0  | 0  | 2  | 0  | 0  | 0  | 0  | 0  |
| U  | -  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| C  | 2  | 1  | 5  | 2  | 3  | 2  | 5  | 3  | 3  | 9  | 1  |
| G  | 2  | 3  | 1  | 3  | 7  | 1  | 5  | 2  | 1  | 7  | 1  |

**Summary of base changes at each position of the partially randomized 62-base prohead binding domain of pRNA in SELEX II**

Competition binding assays were performed as described under “Experimental Procedures.” 5S rRNA was used as a control. Amounts of [32P]pRNA bound were as follows: input, 14964 cpm; no proheads, 142 cpm. The competition activities, which normalize the [32P]pRNA fraction bound with wild-type pRNA competitor (100%) and no competitor (0%) (8, 9), were as follows: 5S rRNA, 2%; pRNA, 100%; cycle 0, 4%; cycle 1, 12%; cycle 2, 16%; cycle 3, 46%; cycle 4, 83%; cycle 5, 97%; and cycle 6, 91%.

![Fig. 4. Competition filter binding with RNA pools of SELEX II.](https://www.jbc.org/)
changes in the pRNA sequence can produce drastic alterations in the predicted foldings (15). Studies on pRNA mutants with limited numbers of G to A or C to U changes generated by bisulfite mutagenesis showed that the foldings of pRNA have prognostic value. In each case in which the overall predicted folding of a mutant differed substantially from the wild-type model, production of RNA was not detected \textit{in vivo} from the cloned genes, presumably because of degradation of misfolded molecules by endogenous ribonucleases. All 22 of the pRNA sequences of Tables II and III that are wild-type for residues 45–62 were predicted to fold as the model did (Fig. 1A), regardless of the D loop sequence. This suggests that all major determinants of secondary structure are in place in these mutants, and even changing all five bases of the D loop did not perturb the predicted folding. However, only 6 of these 22 mutant pRNAs were active in DNA-gp3 packaging, in all cases because the residues G82GG were retained in the D loop, providing the capability to form the intermolecular base pairs needed for oligomerization (see below). None of the pRNAs with changes in the segment that included residues 45–62 could reconstitute RNA-free proheads for DNA-gp3 packaging \textit{in vitro}.

The SELEX experiments defined the minimal requirements for intermolecular base pairing leading to formation of pRNA dimers and hexamers required for DNA-gp3 packaging.2 Previous work suggested that the pseudoknot was a four-base pair interaction involving nucleotides 45–48 and 83–82 (9) (see Fig. 1, A and B). Full DNA-gp3 packaging activity was obtained with the mutants 4-2, 4-5, 4-6, 5-6, 5-8, and 5-12, in which bases were changed from U85 to C; U85 to G; U85 to C; U81 to A; and U84UU85 to GA, respectively (Tables II and III); thus, potential base pairing involving U residues at positions 84 and 85 was not needed for biological activity. The DNA-gp3 packaging results from SELEX II are shown in Fig. 6. All of the biologically active pRNAs maintained G at positions 82 and 83. A change of G82 to C in mutants 4-3, 5-9 and 5-11; to A in mutants 4-8 and 4-11; and to U in mutants 5-10 and 5-18 while maintaining G83 to U resulted in loss of DNA-gp3 packaging activity. Similarly, a change of G83 to U in mutants 4-9 and 5-5 while G82 was maintained resulted in loss of DNA-gp3 packaging activity. As expected, when both G82 and G83 were changed to CC in mutants 4-4, 5-3, and 5-13; AA in mutants 4-7 and 5-15; and AC in mutants 4-1 and 5-20, DNA-gp3 packaging was not observed. These

### Table II

| pRNA | C-G Loop | G sizes | E Loop | E sizes | D Loop | Predicted DNA-gp3 Packaging |
|------|----------|---------|--------|---------|--------|-----------------------------|
| 4-1  | C-U      | 85-81   | C-A    | 45-62   | Y      | Y                           |
| 4-2  | -        | -       | -      | -       | -      | -                           |
| 4-3  | G-U      | -       | -      | -       | -      | -                           |
| 4-4  | U-G      | -       | -      | -       | -      | -                           |
| 4-5  | -        | -       | -      | -       | -      | -                           |
| 4-6  | -        | -       | -      | -       | -      | -                           |
| 4-7  | -        | -       | -      | -       | -      | -                           |
| 4-8  | -        | -       | -      | -       | -      | -                           |
| 4-9  | -        | -       | -      | -       | -      | -                           |
| 4-10 | -        | -       | -      | -       | -      | -                           |
| 4-11 | -        | -       | -      | -       | -      | -                           |

### Table III

| pRNA | C-G Loop | G sizes | E Loop | E sizes | D Loop | Predicted DNA-gp3 Packaging |
|------|----------|---------|--------|---------|--------|-----------------------------|
| 5-1  | G-U      | 85-81   | C-A    | 45-62   | Y      | Y                           |
| 5-2  | -        | -       | -      | -       | -      | -                           |
| 5-3  | -        | -       | -      | -       | -      | -                           |
| 5-4  | -        | -       | -      | -       | -      | -                           |
| 5-5  | -        | -       | -      | -       | -      | -                           |
| 5-6  | -        | -       | -      | -       | -      | -                           |
| 5-7  | -        | -       | -      | -       | -      | -                           |
| 5-8  | -        | -       | -      | -       | -      | -                           |
| 5-9  | -        | -       | -      | -       | -      | -                           |
| 5-10 | -        | -       | -      | -       | -      | -                           |
| 5-11 | G-U      | 85-81   | C-A    | 45-62   | Y      | Y                           |
| 5-12 | -        | -       | -      | -       | -      | -                           |
| 5-13 | -        | -       | -      | -       | -      | -                           |
| 5-14 | G-U      | 85-81   | C-A    | 45-62   | Y      | Y                           |
| 5-15 | -        | -       | -      | -       | -      | -                           |
| 5-16 | -        | -       | -      | -       | -      | -                           |
| 5-17 | -        | -       | -      | -       | -      | -                           |
| 5-18 | -        | -       | -      | -       | -      | -                           |
| 5-19 | -        | -       | -      | -       | -      | -                           |
| 5-20 | -        | -       | -      | -       | -      | -                           |

### Fig. 5

Sequencing gel of pools of RNA from the initial selection through the fifth cycle of SELEX II. Numbers above the lanes refer to the selection cycles, and numbers at the left refer to positions in pRNA. Emergence of the wild-type is clear in selection cycle 3.
results showed that base pairing between $47CC48$ and $85GG82$ is necessary and sufficient for DNA-gp3 packaging. Data showing that this intermolecular interaction produces a hexameric oligomer of pRNA will be presented. Site-directed mutagenesis of these residues has confirmed the results presented here and demonstrated that pRNAs with a change of the wild-type bases $47CC48$–$85GG82$ to GC–CG, CG–GC, CU–GA, UC–AG, or GA–CU retain efficient DNA-gp3 packaging activity. Mg$^{2+}$ is needed for formation of dimers and hexamers of pRNA in solution. Dimers of pRNA have been detected by native polyacrylamide gel electrophoresis at room temperature, and dimers and hexamers were detected at 4°C; proof of the dimeric and hexameric forms of pRNA in solution was obtained by analytical ultracentrifugation.

Conclusions and Perspective—The SELEX experiments recapitulated the natural evolution of the C-E bulge loop and the E stem and loop of pRNA. No alternatives to wild-type pRNA in terms of the sequence and secondary structure of these elements were identified by the SELEX experiments. The only analogues of wild-type pRNA with apparently identical structure and function in DNA-gp3 packaging were the mutants 4-2 and 2-5, which had D loop changes of U$65$ to C, U$65$ to G, $44UU45$ to CA, $U79$ to A, and $44UU45$ to GA, respectively (Tables II and III); all of these RNAs maintain $43GG42$, which forms intermolecular base pairs with $47CC48$ of the C-E bulge loop that are needed for DNA-gp3 packaging. Thus, the D loop functions primarily in intermolecular base pairing of pRNA rather than prohead binding; however, deletion of both the D stem and loop results in loss of prohead binding competitor activity (8).

When the E stem and loop were wild-type in SELEX II sequences, the C-E loop was invariably wild-type, whereas the D loop sequences showed many changes (Tables II and III). Thus, the C-E loop must have prohead binding function unrelated to pseudoknot formation, and the pseudoknot is not a determinant of prohead binding. However, the potential for base pairing of the D loop and the C-E loop that produces pRNA oligomers was present in all SELEX II sequences active in DNA-gp3 packaging (Tables II and III), confirming studies with site-directed mutants. Surprisingly, covariation in bases of the base 45–48 and 81–85 segments that form pseudoknot pairs was common in sequences of cycle 3 of SELEX II that showed many changes in the base 45–62 and 81–85 intervals (data not shown); thus, pseudoknot formation, although not required for prohead binding, may aid prohead binding prior to evolution of a wild-type E stem and loop.

The nine-membered C-E bulge loop, the four-base pair E stem, and the six-residue E loop constitute the crux of the prohead binding domain. Deletion of the C-E loop results in loss of prohead binding competitor activity (8). Inversion of the E stem residues $49CUGA52$ and the complementary residues $62GACU59$ to AGUC and UCAG, respectively, results in prohead binding competitor activity of only 11% and no detectable DNA-gp3 packaging activity, even though the pRNA was predicted to fold as the wild-type did (8). Very few changes were observed in these residues in SELEX I (Table I), and no changes were observed in the G$51$–C$56$ pair. Covariation occurred in the E stem in eight of nine mutants with multiple changes in the residue 45–62 segment in SELEX II (Tables II and III). The composite results show the importance of the E stem in both prohead binding and DNA-gp3 packaging. Considering the E loop, no changes were observed in residues U$53$, U$54$, and A$56$ in SELEX I, and G$55$ was replaced only with A (Table I). Previously, the highly conserved residues U$54$ and A$56$ were changed to C and respectively. Modest decreases were observed in prohead binding competitor activity and a 4-fold and 2-fold drop, respectively, in capacity to reconstitute proheads and support phage assembly in prohead-defective extracts (15). The SELEX I results suggested that the least important residues of the E loop in prohead binding were $G57$ and $U58$, in which change was quite frequent (Table I). Change of the four residues $54UGAG57$ to CUUU resulted in drastic loss of prohead binding competitor activity (8), underscoring the importance of the E loop sequence.

The A helix of pRNA (see Fig. 1A) may be the principal site of interaction with gp16 that results in formation of the unique RNA-dependent DNA translocating ATPase (5, 13). The A helix has been partially randomized with a mutation rate of 40%, and the selection and sequencing of pRNA aptamers that bind gp16 is in progress. Other goals include investigation of the role(s) of oligomerization of pRNA in DNA-gp3 packaging, visualization of pRNA hexamers on the prohead portal vertex, and determination of the higher order structure of pRNA as it interacts with proteins of the DNA packaging machine and the packaging substrate DNA-gp3.

Acknowledgment—We thank Charlene Peterson for preparation of the manuscript.

REFERENCES

1. Guo, P., Grimes, S., and Anderson, D. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 3505–3509
2. Guo, P., Erickson, S., and Anderson, D. (1987) Science 236, 690–694
3. Grimes, S., and Anderson, D. (1997) J. Mol. Biol. 266, 901–914
4. Grimes, S., and Anderson, D. (1989) J. Mol. Biol. 209, 101–108
5. Grimes, S., and Anderson, D. (1989) J. Mol. Biol. 215, 559–566
6. Bailey, S., Wichitwechkarn, J., Johnson, D., Reilly, B. E., Anderson, D., and Boldey, J. W. (1990) J. Biol. Chem. 265, 22365–22370
7. Reid, R., Boldey, J. W., and Anderson, D. (1994) J. Biol. Chem. 269, 5157–5162
8. Reid, R., Boldey, J. W., and Anderson, D. (1994) J. Biol. Chem. 269, 9084–9089
9. Reid, R., Zhang, F., Benson, S., and Anderson, D. (1994) J. Biol. Chem. 269, 16856–16861
10. Tuerk, C., and Gold, L. (1990) Science 249, 505–510
11. Lampel, J., Dill, D., and Leonard, S. (1992) Biotechniques 13, 604–610
12. Wichitwechkarn, J., Bailey, S., Boldey, J. W., and Anderson, D. (1989) Nucleic Acids Res. 17, 3459–3468
13. Zhang, C., Lee, C.-S., and Guo, P. (1994) Virology 201, 77–85
14. Zunker, M. (1989) Science 244, 48–52
15. Wichitwechkarn, J., Johnson, D., and Anderson, D. (1992) J. Mol. Biol. 223, 991–996
