JBC Papers in Press. Published on March 8, 2002 as Manuscript M112061200

1

Computer Modeling of 3D Structure of pRNA Monomer,

Dimer and Hexamer of Phi29 DNA Packaging Motor

Stephen Hoeprich, and Peixuan Guo*

Department of Pathobiology, Purdue University, West Lafayette, IN 47907. USA

*Corresponding author.

Please send correspondence to:

Peixuan Guo

Purdue Cancer Research Center, B-36 Hansen Life Science Research Building,

Purdue University, West Lafayette, IN 47907, USA

Phone: (765) 494-7561; FAX: (765) 496-1795; E-Mail: guop@purdue.edu

Summary

A striking common feature in the maturation of all linear dsDNA viruses is that their lengthy genome is translocated with remarkable velocity into the limited space within a preformed protein shell and packaged into near crystal density. A DNA-translocating motor, powered by ATP hydrolysis, accomplishes this task which would otherwise be energetically unfavorable. A virus-encoded 120-base RNA, pRNA, forms a hexameric complex to serve as a vital component of the DNA translocating motor of bacterial virus phi29. Sequential action of six pRNA ensures continual function in the DNA translocation process. Phi29 motor has been assembled with purified components synthesized by chemical or biotechnological approaches, and is able to pump the viral DNA into the protein shell in vitro. pRNA dimers are the building blocks of the hexamer. The computer models of the three-dimensional structure of the motor was constructed based on experimental data derived from photo-affinity cross-linking by psoralen, phenphi, and azidophenacyl; chemical modification and chemical modification interference with DMS, CMCT and Kethoxal; complementary modification; and nuclease probing by single-stranded and double-stranded specific RNases. The shape of these computer models is very similar to the published pRNA images of cryo-AFM (atomic force microscopy). pRNA hexamer docking with the connector crystal structure reveals a very impressive match with available biochemical, genetic, and physical data.

Introduction

The amazing diversity in RNA function is attributed to the astonishing variety of RNA species and the flexibility in RNA structure. To elucidate the question of how RNA molecules perform their versatile and novel functions, it is crucial to understand the principles and rules that regulate RNA structure. Due to its complexity and versatility, the criterion in RNA folding remains to be elucidated, and determination of RNA structure is an arduous task.

DsDNA viruses package their genomic DNA into a preformed protein shell, the procapsid [for reviews, see (1-3)]. In the case of phi29, a bacterial virus that infects *Bacillus subtilis*, translocation of dsDNA into the procapsid requires a virus-encoded RNA (4;5), called pRNA [for reviews, see (6-8)](Fig. 1). Mg^{++} induces appropriate folding of pRNA for dimerization (9;10). Dimers of pRNA bind to the connector (portal vertex, the unique site where DNA enters the procapsid (11;12)) and serve as the building block for hexamer assembly (9;13). The pRNA molecules interact intermolecularly via hand-in-hand interaction to form a hexameric complex that is a crucial part of the viral DNA translocation motor (14-17). The pRNA appears to be directly involved in the DNA translocation process, and leaves the procapsid after DNA packaging is completed (18). The sequential action of pRNA ensures the continuous function of the motor (18;19).

Dimerization of RNA has been shown to play a vital role in variety of biological functions. Dimerization of retrovirus RNAs via kissing loops governs essential steps in retroviral replication (20-24). We predict that dimerization of RNA might play other vital roles in cell cycles (25), for example, RNA-RNA interaction via alternating loops has also been reported for *bicoid* mRNA in *Drosophila* embryos (8;26). We believe that (25) the mechanism of *bicoid* mRNA interaction and translocation might be similar to that of phi29 pRNA. It is possible that a *bicoid* mRNA may also form hexameric rings to ride, track or rotate along *Staufen* protein during

its transportation. Indeed, there is evidence that bicoid mRNA can form dimers and multimers via RNA loop/loop interactions (26).

The use of computers to model the pRNA 3D structure has been attempted previously (16;27). At the time of these studies, very little experimental data regarding the three dimensional structure of pRNA monomer, dimer and hexamers was available. Considering the circumstances in building a pRNA 3D computer model based solely on data from partial secondary structure predictions and computer quantification, the aforementioned computer modeling work was laudable. Recently, the phi29 DNA packaging motor has become the subject of intense scrutiny (12;15-17;28;29). It has been reported (15;30) that the published partial hexamer model (16) does not match the size of the phi29 connector, the structure of which was recently solved by x-ray crystallography. Now that extensive data on pRNA three dimensional intra- and inter-molecular interaction has become available(9;10;31;32), it is time to construct computer models of the pRNA with the newly available data on structure and distance constraints. In addition, pRNA dimers are the building blocks in hexamer assembly (13), and a computer model of pRNA dimers has not been reported. This report describes the 3D structure of pRNA monomer, dimer and hexamer using computer modeling, based on experimental data derived from photo-affinity cross-linking, chemical modification and chemical modification interference, complementary modification, and nuclease probing by RNases. The models presented here integrate experimental data not previously used to make other models. These more realistic models can then be used to aid in understanding the role of pRNA in phi29 DNA packaging motor. Comparison of the computer models with the published pRNA images of cryo-AFM reveals high similarity in shape. Indeed, docking of pRNA hexamer model with the

connector crystal structure reveals a very impressive match with biochemical, genetic, and physical data currently available.

EXPERIMENTAL PROCEDURES

Crosslinking by psoralen and phenphi. The construction of pRNAs has been described previously (33). Crosslinking of pRNA with AMT (4'-aminomethyl-4,5',8-trimethylpsoralen) (10;34) and phenphi (*cis*-Rh(phen)(phi)Cl$_2^+$ (PHENPHI) (phen = 1,10-phenanthroline; and phi = 9,10-phenanthrenequinone diimine) was performed as described (35).

Crosslinking by azidophenacyl. GMPS-containing circularly permuted pRNA (cp-pRNAs) were prepared by *in vitro* transcription with T7 RNA polymerase in the presence of 40 mM Tris-HCl pH7.5, 12 mM MgCl$_2$, 2 mM spermidine, 10 mM NaCl, 1 mM ATP, 1 mM CTP, 1 mM UTP, 0.2 mM GTP, [a-^{32}P] GTP, 8 mM guanosine 5'-monophosphorothioate (GMPS) (Amersham Life Sciences) at 37^{0}C for 4 hrs (36). Transcripts containing the 5'-terminal phosphate of 5'-guanosine monophosphorothioate were coupled to an azidophenacyl group (36-38), then exposed to UV light (Phillips, UVB 20W-TL01, 311-nm) for 15-30 min. at 0°C to produce intramolecular crosslinks. Under these conditions, no photoagent independent crosslinks were detected (32).

Crosslinking of pRNA dimers were achieved by mixing equimolar amounts of photoagent-containing cp-pRNA I-a' with a transcomplementary pRNA A-i' in TMS (50 mM Tris-HCl pH 7.8, 10 mM MgCl$_2$, 100 mM NaCl). The pRNAs were incubated on ice for 15 min. and then exposed to UV light. Primer extension was performed to identify the crosslinking sites (36).

Chemical Modification. RNAs were modified with the chemicals dimethyl sulfate (DMS), 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate (CMCT), and β-

etoxy-α-ketobutyraldehyde (kethoxal) as reported previously (9;31). The concentration of chemicals were titrated empirically to produce, on the average, one base modification per pRNA molecule as described (9;27;39;40).

Chemical Modification Interference(41-43). Two pRNAs, 5'/3' B-a' and 23/97 A-b', were used to produce pRNA dimers (25;32). RNA 23/97 A-b' is a 75-base RNA that lacks bases 1-23 and 98-120. This RNA has been shown to be competent in dimer formation (13). pRNA 5'/3' B-a' was modified by chemicals and 23/97 A-b' was not. In addition, primers used in reverse transcriptase extension targeted pRNA 5'/3' B-a' specifically. This strategy was used to avoid ambiguous primer extension results (32).

The monomer RNA B-a' was treated with either DMS or CMCT as described (9;27;32), and then mixed with unmodified monomer A-b' to test their competency in dimer formation. After incubation, the reaction mixture was electrophoresed to separate the monomer from dimer. Both monomer and dimer were, after being isolated from the gel, subjected to primer extension as described (10). If a modified base is involved in dimer formation, pRNA B-a' carrying this modified base would not be able to form dimers with A-b', and thus will be present in the fast migrating band representing the monomer in the gel. The concentration of the modifying chemical was titrated so that on the average only one base of each pRNA was modified .

T1 ribonuclease probing. 10-20 ng [^{32}P]-labeled pRNA (approximately 3000 cpm) in 2 μl of H_2O was mixed with 6μl of carrier tRNA (1.75μg/μl), and dialyzed on a 0.025-μm type VS filter membrane (Millipore Corp.) against TBE (89mM Tris-borate, 2 mM EDTA, pH 8.0) for 15 minutes. Half of the sample was then transferred into an Eppendorf tube, and the other half was further dialyzed on VS filter membrane against TMS (50 mM Tris pH 7.8, 100 mM NaCl, 10 mM $MgCl_2$) for 30 minutes. One μl of T1 ribonuclease (1 unit/μl from the RNA sequencing kit

of USB) was added to both samples. After 15 minutes at ambient temperature, the reactions were stopped by adding an equal volume of stop solution (95% formamide, 0.025% xylene, 10 mM EDTA) and loaded onto an 8% sequencing gel. Both T1 ribonuclease and alkaline hydrolysis ladders were generated following the instructions of the RNA sequencing kit (USB) (10).

Creation of the 3D models.

Models of pRNA monomer, dimer and hexamer were produced on Silicon Graphics Indigo^2 and Octane computers running IRIX 6.2 or 6.5, using the programs NAHELIX, MANIP, PRENUC, NUCLIN, and NUCMULT (44;45). An SGI Dials Box attached to the computer was used to provide much of the user input to the MANIP program, including zooming in and out, linear movements and rotations. Using MANIP, one or more selected nucleotides can be moved, rotated, or torqued in bond angles. Nucleotides can be joined or separated.

The modeling was performed based on the following assumptions.

1). Fragments of the molecule were created using the program NAHELIX. All helices are created as a regular A-form double helix, and sequence dependent distortions of the helix are generally ignored.

2) Internal loops and mismatched bases are constructed by maintaining the integrity of the double helix while optimizing base pairing and stacking inside the loop, as suggested by other data from crystallography and NMR studies.

3) A general principle for the modeling of RNA hairpin loops has been proposed (46), and involves maximal stacking on the 3' side of the stem and enough nucleotides stacked on the 5' side to allow loop closure, as found in the loop of tRNA anticodon.

4) Bulges are constructed either protruding from stems, so that there is no helical distortion, or within the helical domain, forcing the helical axis to bend. The energies for stacking are considered to decide whether bulges should be protruding from or within the helical stems (47).

5) Helix untwisting or twisting, helix-helix interactions, triple base interactions (48), pseudoknots, or other higher order structures have been built into the model at constant geometrical distances but allowing certain torsion angle variation. The principle regarding RNA flexibility has been applied to the construction of the $U^{72}U^{73}U^{74}$ bulge at the three-helix junction of pRNA. This three-base bulge has been found to provide flexibility for the appropriate folding of the three-helix junction. Traditional computer algorithms involving the minimization of empirical energy functions have been considered.

6) Distances between atoms can be monitored as their positions are changed. 12 angstroms has been considered as a maximum distance constraint when bases are crosslinked by GMPS/Aryl Azide. Modified distance geometry and molecular mechanics algorithms have been considered to generate structures consistent with data from cross-linking, chemical modification and chemical modification interference. A constraint satisfaction algorithm provided by the program is used to refine the structure to take care of some poorly defined regions of pRNA in order to ensure a plausible representation of all atoms. PRENUC and NUCLIN create certain files needed for the refinement process to begin. NUCMULT refines the model by changing distances and angles to be closer to standard values. More refinements give the model more standard dimensions, if the user originally made the model with distances and angles that do not deviate too much from commonly accepted values.

RESULTS AND DISCUSSIONS

Previous work has identified the intermolecular interaction between the right hand loop (the loop closest to the 5' terminus of the pRNA) of one RNA molecule and the left hand loop (the loop closest to the 3' terminus of the pRNA) of another pRNA molecule (15-17). This intermolecular interaction between the loops for the formation of a hexamer is referred to as "hand-in-hand" interaction (25). In addition, pRNA dimer has been shown to be the building block for the formation of the hexameric complex (13). A model for the pathway of hexamer formation has been proposed (13).

The goal of modeling pRNA is to organize collections of structural data from cross-linking, chemical or ribonuclease probing, cryo-AFM and other genetic data into a three-dimensional form. Since a large number of structural constraints are available, computer programs can successfully construct three-dimensional structures.

Computer models of pRNA monomer[a], dimer[b], hexamer[c] and the hexamer/connector complex[d] are presented (Fig. 2). The pRNA molecule contains five primary regions (Fig. 1a). The first region is the 5'/3' helix that includes bases 1-28 and 92-117. The second region is the left hand stem loop incorporating bases 75-91. The third region is made of bases 40-44 and the right hand loop composed of bases 45-48. The fourth region is the head stem loop made of bases 49-62. Bases 30-39 and 71-63 comprise the fifth region, which is the helix between the three-helix junction point and the right hand loop. Hand-in-hand base pairing among 6 monomers forms a hexamer bound to a phi29 procapsid. The justification for the modeling is described below.

Data to justify for the construction of the monomer model[a]

**Complementary modification revealed that the 5' and 3' ends of the pRNA exist as a helix.** Complementary modification was used to confirm the presence of helical regions within the pRNA secondary structure (49-51) predicted by phylogenetics (52). An extensive series of helix disruptions by base substitutions virtually always resulted in the loss of DNA packaging activity. The inactive pRNAs in this category include pRNA FW/4, pRNA 14-16/10 and 7/101-103, 7/29 and 28/21; pRNA F3 (Fig. 3). Additional mutations that restored the predicted base pairing rescued pRNA activity, for example, pRNA FW/RV, pRNA 14-16/101-103, pRNA 28/29, pRNA F3/A5 with compensatory mutations are all active in phi29 DNA packaging. The secondary site suppression confirmed that these regions indeed are helical. The computer model of the pRNA monomer supports these data by showing that bases 1-2 are paired with bases 117-116; bases 7-9 are paired with bases 112-110; bases 14-16 are paired with bases 103-101; and bases 76-78 are paired with bases 90-88. The complementary modification data was incorporated into the 3D monomer model (Fig. 4).

Psoralen crosslinking shows that U^{69} is in close proximitye to U^{31}, U^{33}, and U^{36}. Psoralen

is a photoactive probe for pRNA structure (10) which intercalates into RNA helices. After

irradiation with 320-400 nm light, uridines that are in close proximity (helix or pseudoknot) are

covalently linked (34;53;54). The sites of crosslinks can be determined by primer extension (10)

and/or Mung bean nuclease treatment (55). The psoralen derivative, AMT (4'-aminomethyl-

4,5',8-trimethyl psoralen), was selected in this study due to its solubility (10). Crosslinking with

AMT revealed that in the absence of Mg^{++}, U^{69} is crosslinked to U^{31}, U^{33}, and U^{36}. Although our

model is created assuming Mg^{++} is present, our computer model of pRNA monomer still

provides useful information by showing that U^{69} is not distant from U^{31}, U^{33}, and U^{36} (Fig. 5).

Photoaffinity cross-linking with phenphi showed that base G75 is in close proximity to

G28 and G30. Phenphi [(cis-Rh(phen)(phi)Cl$_2{}^+$ (phen = 1,10-phenanthroline; and phi = 9,10-

phenanthrenequinone diimine)] was used to crosslink pRNA (35). UV light was used to activate

phenphi, which then formed covalent bonds between guanosine bases. Primer extension was

performed, and the reaction was electrophoresed to determine crosslinking sites. Stops in primer

extension reactions were observed at U^{29}, U^{31}, and U^{76}, corresponding to crosslinks to bases G^{28},

G^{30}, and G^{75}, respectively. The monomer model supports this data by showing that G^{75} is

proximate to G^{28} and G^{30} (Fig. 6).

Photoaffinity cross-linking with azidophenacyl (APA) to show that base G^{75} is in close

proximity to bases 26-30, while G^{78} is close to U^{31}, and G^{108} is close to bases 10-11. Circular

permutation allows the creation of new 5'/3' termini of pRNA while maintaining correct folding

(50;56;57), permitting labeling of any specific internal base by radioactive or photoaffinity

agents. Cross-linking was accomplished by attaching the photosensitive agent azidophenacyl

(APA) to the new 5' terminus of the circularly permuted pRNA (cp-pRNA) by the use of GMPS

as the first nucleotide incorporated in *in vitro* transcription and coupling with azidophenacyl bromide (32;36;50;57). Three nucleotides were selected as new 5' terminus for labeling with APA. One of the new 5' termini, G^{108}, is located within the helix necessary for DNA packaging, while two of the other sites, G^{75} and G^{78}, are located within interior sequences involved in procapsid binding. The particular nucleotides that cross-linked to the new termini of the cp-pRNAs labeled with APA were determined by primer extension after UV crosslinking. Stops in primer extension occur one base prior to crosslinked bases, thus a stop at base 32 would mean that base 31 was cross-linked. Extension products from cross-linked cp-pRNA were compared with that from non-cross-linked cp-pRNA to identify individual cross-linked nucleotides. It was found that G^{108} was cross-linked to C^{10} and G^{11}; base G^{75} was cross-linked to A^{26}, U^{27}, G^{28}, U^{29}, and G^{30}; and G^{78} was cross-linked to U^{31}. The azidophenacyl group is only 9 Å, but experimental data has demonstrated that the cross-linking group can reach distances of 12 Å (Norman Pace, personal communication). The data suggest that G^{108} is close to C^{10} and G^{11}, G^{75} is close to bases 26-30, and G^{78} is close to U^{31}, as seen in Fig. 7.

Chemical modification showed that the sequence $C^{18}C^{19}A^{20}$ forms a loop extended above the surface of the pRNA. Three different chemical probes were utilized to probe the structure of phi29 pRNA. The chemicals modify atoms in unpaired bases that, if paired instead, are involved in W-C base pairing. DMS methylates N1 of adenine and N3 of cytosine (39). CMCT reacts with guanines at N1 and uridines at N3 (39). Kethoxal reacts with guanines at N1 and N2 (39). Base modifications were detected by reverse transcriptase primer extension (39;40). The samples were subsequently electrophoresed on sequencing gels to determine stops in the extended primers. Stops occur one base prior to modified bases, thus a stop at base 21 would mean that base 20 was modified.

Chemical probing of pRNA revealed a large area of protection. However, the 3-base bulge $C^{18}C^{19}A^{20}$ (Fig. 8) was accessible to chemicals in monomer, dimer as well as procapsid-bound pRNA (9;27). A pRNA with three bases, 3'GGU5', inserted between A^{99} and A^{100} to pair with the bases $C^{18}C^{19}A^{20}$ in the bulge generates the pRNA 7/GGU (50). This pRNA was fully competent to form dimers and bind procapsids, however its activity in DNA packaging and virion assembly was completely lost (50). A pRNA with a deletion of all three bases of the CCA bulge (58) exhibited the same biological activity as pRNA 7/GGU concerning procapsid binding, DNA packaging and virion assembly. The results suggest that CCA, though not involved in procapsid binding, is present on the surface of the pRNA as a bulge to interact with other DNA packaging components (27).

Chemical modification reveals unpaired bases in loops and bulges. As already noted, chemical modification revealed that bases $C^{18}C^{19}A^{20}$ were modified by chemicals and confirmed that these three bases exist as a 3-base bulge. Additionally, bases 18-20, 42-48, 55-57, 82-86 are all strongly modified by chemicals. The monomer model supports these results by showing that all these bases are present in the model as single stranded loop or bulges (Fig. 9). Phylogenetic analysis of similar RNA from five different phages concurred with these results by showing that all of these bulges are present in similar regions of all five RNA molecules (Fig. 10). All five predicted single-base bulges (50;51) in phi29 pRNA are also modified fairly strongly, as are bases A^9, C^{10}, U^{36}, A^{93} and A^{100}. The monomer model concurs with these data by showing that all these bases are present in the model either as bulges, adjacent to a bulge or facing a bulge in the complementary strand.

The UUU presented as a bulge at the three-way junction provides flexibility in folding and serves as a hinge for the twisting of the lower stem-loop. Nucleotides $U^{72}U^{73}U^{74}$ were

presented in the model as a bulge located at the pRNA three helix junction (Fig. 11). The basis for such construction in the model is as follows: mutation studies have shown that deletion of these three nucleotides abolishes the activity of the mutant pRNA F5 with native 5'-/3'-ends (51). However, a circularly permuted cpRNA 75/71 that had a deletion of these three nucleotides but had new 5' and 3' ends located at bases 75 and 71, respectively, was fully active in *in vitro* phi29 assembly (50) (Fig. 11), suggesting that the UUU bulge in this area provided flexibility to the pRNA. In pRNA F5 with normal 5' and 3' ends, deletion of the UUU bulge eliminated the flexibility in folding of the three-way junction, therefore the mutant was misfolded. In cpRNA 71/75, this flexibility was compensated for by providing a new opening where the F5 pRNA mutant was missing the UUU bulge (51). It is our belief that the new termini in the area of deletion provided comparable flexibility through discontinuity of the phosphodiester bond as was provided by the hinge-like UUU bulge.

Comparison of computer mono model with published monomer images of *cryo-AFM*. Atomic force microscopy has been used by several investigators to detect images of RNA in a denatured conformation. Ph29 pRNA was utilized as the first attempt to examine the 3D structure of RNA in native conformation by cryo-AFM (9;13;32).

Cryo-AFM imaging revealed that the pRNA monomer folded into a "✓" (check mark)-shaped structure, resembling the computer monomer model. The color indicates the thickness and height of the image, but does not reflect the atom density observed end on. The brighter or whiter the color, the taller the surface is in the image. The darker the color, the lower the surface is in the image. The color and contrast of the image clearly indicate that the area around the head loop (the elbow of the "✓") is the thickest or tallest (Fig. 12), and strongly agrees with the computer model.

B. Data to justify the construction of dimer model[b]

Studies with pRNA mutants have revealed that two single stranded loops in the pRNA are involved in inter-RNA interactions to form a pRNA hexamer for phi29 DNA translocation (14;15;18) (for minireview see (19)). These two loops interact alternately to generate interlocking chains. Stable dimer pRNAs have been isolated and purified and are believed to be the intermediate for hexameric complex formation (9;13;16;25). Thus, it is logical to model the dimer and gain some insight about its structure.

Phylogenetics and mutation studies suggested that Bases 45-48 paired intermolecularly to base 85-82. Phylogenetic analysis of pRNAs from *Bacillus subtilis* phages SF5, B103 (59), phi29, PZA, M2, NF, and GA1 (60), shows very low sequence identity and few conserved bases, yet, the family of pRNAs appears to be very similar in predicted secondary structure (52) (25)(Fig. 10). All seven pRNAs of these phages contain both the right and left loops. Complementary sequences within the two loops were found in each of these pRNAs. The numbers of paired bases were from five (5'-GUUUU/CAAAA-5') for SF5 to four (5'-AACC/UUGG-5') for phi29/PZA and B103, to three (5'-AUC/UAG-5') for M2/NF, and two (5'-CC/GG-5') for GA1. A loop/loop interaction has been used as a parameter in modeling the pRNA dimer (Fig. 10F).

Genetics studies by in vitro mutagenesis. A series of mutant pRNAs carrying mutated right and/or left hand loop sequences were made. To simplify description, we used uppercase and lowercase letters to represent the right and left hand-in-hand loop sequences of the pRNA, respectively (Table 1). The same letter in upper and lower cases symbolizes a pair of complementary sequences. For example, in pRNA A-a', the right loop A (5'GGAC[48]) and left

loop a' (3'CCUG^{82}) are complementary, while in pRNA A-b', the four bases of the right loop A are not complementary to the sequence of the left loop b' (3'UGCG^{82}).

Determination of loop/loop interactions was accomplished by the mixing of inactive mutant pRNAs, each having interactive complementary loops, with each other to determine the loop/loop interaction (15). All mutant pRNAs that had unpaired right and left loops, such as pRNA J-i', were inactive in *in vitro* phi29 assembly when used alone. However, when two inactive pRNAs that were trans-complementary in their right and left loops, for example pRNA I-j' and J-i', were mixed in a 1:1 molar ratio, full activity was restored. The observed activity of a mixture of two inactive mutants (Table 1) suggests that the number of pRNAs in the DNA packaging complex was a multiple of two and confirmed that the right loop interacted with the left loop in dimer formation. Other combinations of pRNA mutants used in this manner suggested the number of pRNAs in the DNA packaging complex was also a multiple of three and a multiple of six.

Intermolecular crosslinking data. The methods used for the dimer azidophenacyl crosslinking are essentially the same as those used for the monomer azidophenacyl crosslinking work described above. Circularly permutated pRNA B-a' was made with an azidophenacyl label on G^{82}. Labeled pRNAs were incubated with pRNA A-b' that has its left and right loop sequences complementary to the right and left loop, respectively, of pRNA B-a'. After UV-crosslinking, dimers were isolated from gels and the crosslinking site was identified by primer extension. G^{82} was found to cross-linked to G^{39}, G^{40}, A^{41}, C^{49}, G^{62}, C^{63}, and C^{64} (36). The dimer model support this finding by showing that G^{82}, G^{39}, G^{40}, A^{41}, C^{49}, G^{62}, C^{63}, and C^{64} are all in close proximity, and that the distance from G82 to these nucleotides is less than 12 angstroms (Fig. 13).

Chemical modification of dimer. Dimers consisting of A-b' and B-a' pRNAs were chemically modified using the same methods that were used for the modification of monomer. Bases C^{85}, C^{84}, U^{83}, G^{82}, A^{45}, C^{46}, G^{47} and C^{48} were not modified in dimers while they were modified in monomers (Fig. 8). Each of these bases is within the right/left hand loops, which are involved in inter pRNA interaction (15;16). Bases G^{57}, A^{56} and G^{55}, located in the head loop were also protected from chemical modification. Comparison of the modification patterns of monomers and dimers supported the computer model of dimers showing that all three major loops, the right, left and head loops, were involved in pRNA/pRNA contact to form dimers, since all these three loops were strongly modified in monomer but protected from modification in dimers (Fig.9).

Chemical modification interference distinguished bases involved from bases not involved in dimer formation. Chemical modification interference was performed to determine which bases were involved in dimer formation (Fig. 14). The monomer RNA B/a' was treated with either DMS or CMCT and then mixed with unmodified monomer A/b' to test the modified RNA's competency in dimer formation. After incubation, the reaction mixture was electrophoresed to separate monomers and dimers. Both monomers and dimers were, after isolation from gels, subjected to primer extension as other chemical modifications described above. If the modified base is involved in dimer formation, pRNA B-a' carrying this modified base would not be able to form dimers with A-b', and thus will be present in the fast migrating band representing the monomer in the gel. The concentration of the modifying chemical was titrated so that on the average only one base of each pRNA would be modified.

Bases 45-49, 52, 54-55, 59-62, 65-66, 68-71, 82-85, 88-90 showed a very strong involvement in dimer formation as revealed (32) by primer extension showing modification of

these bases in RNA isolated from the monomer band. The dimer model (Fig. 14) reveals that each of these bases is located at the interface between two pRNA monomers, coinciding with the data from chemical modification interference.

Comparison of computer dimer model with published dimer images of *cryo-AFM* We have used cryo-AFM to directly visualize purified pRNA dimers (9;13;32). The native dimers consisting of pRNAs A-b' plus B-a' had an elongated shape. Since the dimer is elongated, it appears that head to head contact was involved in dimer formation, resulting in a complex almost twice as long as a monomer. The computer dimer model has a very similar shape compared to the cryo-AFM images (Fig. 12).

C. *Data to justify the construction of hexamer model*[c]

Loop/loop interaction to form a hexamer. As already noted, dimers are the building blocks of the pRNA hexamer, and the pathway in assembling a hexamer is: dimer to tetramer to hexamer (13). It has also been shown that closed dimers, two molecules linked together by the holding of two pairs of hands (intermolecularly base paired sequences), were active in procapsid binding and DNA packaging, while open dimers, formed by the holding of only one pair of hands, are unstable in solution (13). Both tandem and fused pRNA dimers with complementary loops designed to form even-numbered rings were active in DNA packaging, while those without complementary loops were inactive (13;16). All of these findings imply that the true pRNA intermediate in hexamer assembly is the closed dimer with the holding of two pairs of hands, and that the two interacting loops played a key role in recruiting the incoming dimer (Fig. 15). Interestingly, hand-in-hand interaction has also been shown to be the mechanism in pRNA hexamer formation (15;25). In dimers, each pRNA monomer subunit only holds hands of ONE

additional pRNA. However, in hexamers each pRNA monomer subunit holds hands of TWO additional pRNAs. Thus the hand interaction in dimers and hexamers seems paradoxical, but can be explained by the finding that the pRNA has a strong tendency to form a circular ring by hand-in-hand contact regardless of whether the final product is a dimer, trimer or hexamer (to be published). Therefore, a conformational shift is expected during the transition from dimer to hexamer. We speculate that dimer formation is a prerequisite to generate an appropriate 3D interface for procapsid binding. One of the hands of the dimer would release after binding to the procapsid. The dimer with a released hand is similar to the open (linear) dimer that has been demonstrated to be unstable in solution but was still active in procapsid binding and DNA packaging (25). Such a conformation shift could be the intrinsic nature of such an intriguing RNA that could bear the task of DNA transportation. Indeed, pRNA conformational changes (for a review, see (19)) before and after binding to procapsid have been documented by nuclease probing, cross-linking and chemical modification (9;10;32;61). Recent studies have provided substantial information regarding the 3D structure of the pRNA (9;10;32;36). To comply with these new data, a new hexamer model was constructed. In this new model, the relative location of the stem loops has been manipulated to fulfill the aforesaid distance constraints (Fig. 2c), revealing that the distance between bases G78 and U31, and bases G75 and A26, U27, G28, U29 or G30 are shorter than 12 angstroms (see Cross-linking of monomer) (32). Also, within dimers the distance from bases G82 to G39, G40, A41, C49, G62, C63 or C64 (See Cross-linking of dimer) is less than 12 angstroms (36).

 Two functional domains of the pRNA. Extensive investigation reveals that the pRNA molecule contains two functional domains (Fig. 1a). One domain is for connector binding and the other is for DNA translocation (for a review, see (19)). This conclusion comes from the

results of: a) base deletion and mutation (33;49-51;60); b) ribonuclease probing (10;61); c) oligo targeting(62;63); d) competition assays to inhibit phage assembly (14;63;64); e) crosslinking to portal protein by UV (65); and f) psoralen crosslinking and primer extension (10). A truncated pRNA, comprised of bases 28-91, can still be specifically UV-crosslinked to the phi29 connector (65). A 75-base RNA segment, comprised of bases 23-97, was able to form dimers, interlock into hexamers, compete with full-length pRNA for procapsid binding, and thereby inhibit phi29 assembly *in vitro* (13). The connector binding domain is located at the central part of the molecule (13;61;65), bases 23-97 (Fig. 2C, E & F in green), and the DNA translocation domain is located in the 5'/3' paired ends (33)(Fig. 2C, E & F in red and cyan).

Protein/RNA crosslinking (65) and connector (portal vertex or gp10) binding assays (5) reveal that pRNA binds to the connector with its procapsid binding domain. Data from footprinting reveals that binding of pRNA to procapsid protects bases 26 to 83 of the pRNA from attack by nucleases (61). Chemical modification revealed that these same areas were inaccessible to chemicals after the pRNA bound procapsid (27)(Fig. 2e, 2f). Our hexamer model complies with the aforementioned data showing that the bases 23-97 (Fig. 2, E & F in green), which is the connector binding domain, interact with the predicted RNA-binding domain of connector (Fig. 2E and F in blue), while the 5'/3' paired region (Fig. 2E in red and cyan), which is the DNA translocation domain, extends away from the connector.

Docking of pRNA hexamer to connector[d]. The phi29 connector contains a wide end and a narrow end. The wide end is embedded in the capsid and the narrow end is exposed (12;66;67). By sequence homology comparison, it was predicted that the connector protein (gp10) contains a conserved RNA recognition motif (RRM), residues 148-214, located at the narrow end of the connector that protrudes from the procapsid (11;68) (for a review, see (19)). Our

connector/RNA docking model supports such a prediction by showing that the pRNA hexamer is attached to the RRM (Fig. 2E in blue), via its connector binding domain (Fig. 2C ,E & F in red and cyan). X-ray crystallography revealed that the connector contains three sections, a narrower section with a diameter of 6.6nm, a central section with a diameter of 9.4nm, and a wider section with a diameter of 13.8nm (12;67). The hexamer model presented here contains a central channel with a diameter of 7.6nm, that perhaps can sheath onto the narrow end of the connector and would be anchored by the connector central section, which is wider than the central channel of the RNA hexamer (Fig. 2E & F).

References

1. Black, L. W. (1988) DNA packaging in dsDNA bacteriophages. In Calendar, R., editor. *The Bacteriophages*, Plenum, New York

2. Bazinet, C. and King, J. (1985) *Ann.Rev.Microbiol.* **39,** 109-129

3. Guo, P. (1994) *Seminars in Virology (Editor's Introduction)* **5(1),** 1-3

4. Guo, P., Erickson, S., and Anderson, D. (1987) *Science* **236,** 690-694

5. Guo, P., Bailey, S., Bodley, J. W., and Anderson, D. (1987) *Nucleic Acids Res.* **15,** 7081-7090

6. Anderson, D. L. and Reilly, B. (1993) Morphogenesis of bacteriophage ϕ29. In Sonenshein, A. L., Hoch, J. A., and Losick, R., editors. *Bacillus subtilis and other gram-positive bacteria: Biochemistry, physiology, and molecular genetics*, American Society for Microbiology, Washington, D.C.

7. Guo, P. and Trottier, M. (1994) *Seminars in Virology* **5,** 27-37

8. Wagner, C., Palacios, I., Jaeger, L., St Johnston, D., Ehresmann, B., Ehresmann, C., and Brunel, C. (2001) *J Mol.Biol.* **313,** 511-524

9. Trottier, M., Mat-Arip, Y., Zhang, C., Chen, C., Sheng, S., Shao, Z., and Guo, P. (2000) *RNA* **6:1-10,** 1257-1266

10. Chen, C. and Guo, P. (1997) *J.Virol.* **71,** 495-500

11. Carazo, J. M., Donate, L. E., Herranz, L., Secilla, J. P., and Carrascosa, J. L. (1986) *J Mol Biol* **192,** 853-867

12. Simpson, A. A., Tao, Y., Leiman, P. G., Badasso, M. O., He, Y., Jardine, P. J., Olson, N. H., Morais, M. C., Grimes, S., Anderson, D. L., Baker, T. S., and Rossmann, M. G. (2000) *Nature* **408,** 745-750

13. Chen, C., Sheng, S., Shao, Z., and Guo, P. (2000) *J Biol Chem* **275(23),** 17510-17516

14. Trottier, M. and Guo, P. (1997) *J.Virol.* **71,** 487-494

15. Guo, P., Zhang, C., Chen, C., Trottier, M., and Garver, K. (1998) *Mol.Cell.* **2,** 149-155

16. Zhang, F., Lemieux, S., Wu, X., St.-Arnaud, S., McMurray, C. T., Major, F., and Anderson, D. (1998) *Mol.Cell.* **2,** 141-147

17. Hendrix, R. W. (1998) *Cell* **94,** 147-150

18. Chen, C. and Guo, P. (1997) *J.Virol.* **71,** 3864-3871

19. Guo, P. (2002) *Prog in Nucl Acid Res & Mole Biol.* **72,** 00

20. Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994) *Proc.Natl.Acad.Sci USA* **91,** 4945-4949

21. Laughrea, M. and Jette, L. (1996) *Biochemistry* **35 No. 5,** 1589-1598

22. Muriaux, D., Rocquigny, H. D., Roques, B. P., and Paoletti, J. (1996) *J Biol Chem* **271 No.52,** 33686-33692

23. Greatorex, J. S., Laisse, V., Dokhelar, M. C., and Lever, A. M. L. (1996) *Nucleic Acids Res.* **24 no .15,** 2919-2923

24. Clever, J. L., Wong, M. L., and Parslow, T. G. (1996) *J.Virol.* **70(9),** 5902-8

25. Chen, C., Zhang, C., and Guo, P. (1999) *RNA* **5,** 805-818

26. Ferrandon, D., Koch, I., Westhof, E., and Nusslein-Volhard, C. (1997) *EMBO J* **16,** 1751-1758

27. Zhang, C., Trottier, M., Chen, C., and Guo, P. (2001) *Virology* **281,** 281-293

28. Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L., and Bustamante, C. (2001) *Nature* **413,** 748-752

29. Davenport R.J. (2001) *Science* **291,** 2071-2072

30. Morais, M. C., Tao Y, Olsen, N. H., Grimes, S., Jardine, P. J., Anderson, D., Baker TS, and Rossmann, M. G. (2001) *J Struct Biol* **135,** 38-46

31. Zhang, C., Trottier, M., Chen, C., and Guo, P. (2001) *Virology* **281,** 281-293

32. Mat-Arip, Y., Garver, K., Chen, C., Sheng, S., Shao, Z., and and Guo, P. (2001) *J Biol Chem* **276,** 32575-32584

33. Zhang, C. L., Lee, C.-S., and Guo, P. (1994) *Virology* **201,** 77-85

34. Wassarman, D. A. (1993) *Molecular Biology Reports* **17,** 143-151

35. Mohammad, T., Chen, C., Guo, P., and Morrison, H. (1999) *Bioorg.Med.Chem Lett.* **9,** 1703-1708

36. Garver, K. and Guo, P. (2000) *J Biol Chem* **275(4),** 2817-2824

37. Nolan, J. M., Burke, D. H., and Pace, N. R. (1993) *Science* **261,** 762-765

38. Harris, M., Nolan, J., Malhotra, A., Brown, J., Harvey, S., and Pace, N. (1994) *The EMBO Journal* **13,** 3953-3963

39. Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P., and Ehresmann, B. (1987) *Nucleic Acids Res.* **15,** 9109-9128

40. Moazed, U., Stern, S., and Noller, H. F. (1986) *J Mol Biol* **187,** 399-416

41. Siebenlist, U. and Gilbert, W. (1980) *Proc.Natl.Acad.Sci.USA* **77,** 122-126

42. Conway, L. and Wickens, M. (1987) *EMBO J.* **6,** 4177-4184

43. Rymond, B. C. and Rosbash, M. (1988) *Genes Dev.* **2,** 428-439

44. Massire, C. and Westhof, E. (1998) *J Mol.Graph.Model.* **16,** 197

45. Westhof, E., Dumas, P., and Moras, D. (1985) *J.Mol.Biol.* **184,** 119

46. Haasnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Broom, J. H., Singh, U. C., Pattabiraman, N., and Kollman, P. A. (1986) *J.Biomol.Struct.& Dyn* **3,** 843-857

47. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) *Annu.Rev.Biophys.Chem.* **17,** 167-192

48. Michel, F., Ellington, A. D., Couture, S., and Szostak, J. W. (1990) *Nature* **34,** 578-580

49. Zhang, C. L., Tellinghuisen, T., and Guo, P. (1995) *RNA* **1,** 1041-1050

50. Zhang, C. L., Tellinghuisen, T., and Guo, P. (1997) *RNA* **3,** 315-322

51. Reid, R. J. D., Zhang, F., Benson, S., and Anderson, D. (1994) *J Biol Chem* **269,** 18656-18661

52. Zuker, M. (1989) *Science* **244,** 48-52

53. Tyc, K. and Steitz, J. A. (1992) *Nucleic Acids Res.* **20,** 5375-5382

54. Cimino, G. D., Gamper, H. B., Isaacs, S. T., and Hearst, J. E. (1985) *Ann.Rev.Biochem.* **54,** 1151-1193

55. Hui, C. F. and Cantor, C. R. (1985) *Proc.Natl.Acad.Sci USA* **82,** 1381-1385

56. Pan, T., Gutell, R. R., and Uhlenbeck, O. C. (1991) *Science* **254,** 1361-1364

57. Zhang, C. L., Trottier, M., and Guo, P. X. (1995) *Virology* **207,** 442-451

58. Reid, R. J. D., Bodley, J. W., and Anderson, D. (1994) *J.Biol.Chem.* **269,** 9084-9089

59. Pecenkova, T., Benes, V., Paces, J., Vlcek, C., and Paces, V. (1997) *Gene* **199,** 157-163

60. Bailey, S., Wichitwechkarn, J., Johnson, D., Reilly, B., Anderson, D., and Bodley, J. W. (1990) *J Biol Chem* **265,** 22365-22370

61. Reid, R. J. D., Bodley, J. W., and Anderson, D. (1994) *J Biol Chem* **269,** 5157-5162

62. Zhang, C. L., Garver, K., and Guo, P. (1995) *Virology* **211,** 568-576

63. Trottier, M., Garver, K., Zhang, C., and Guo, P. (1997) *Nucleic Acids Symposium Series* **36,** 187-189

64. Trottier, M., Zhang, C. L., and Guo, P. (1996) *J.Virol.* **70,** 55-61

65. Garver, K. and Guo, P. (1997) *RNA.* **3,** 1068-1079

66. Jimenez, J., Santisteban, A., Carazo, J. M., and Carrascosa, J. L. (1986) *Science* **232,** 1113-1115

67. Simpson, A. A., Leiman, P. G., Tao, Y., He, Y., Badasso, M. O., Jardine, P. J., Anderson, D. L., and

    Rossmann, M. G. (2001) *Acta Crystallogr.D.Biol.Crystallogr.* **57,** 1260-1269

68. Grimes, S. and Anderson, D. (1990) *Mol.Biol.* **215,** 559-566

---

[a] The atomic coordinates for the structure of this pRNA monomer are available in The Protein Data Bank (http://www.rcsb.org/pdb/index.html) under PDB # 1L4R.

[b] The atomic coordinates for the structure of this pRNA dimer are available in The Protein Data Bank (http://www.rcsb.org/pdb/index.html) under PDB # 1L4Q.

[c] The atomic coordinates for the structure of this pRNA hexamer are available in The Protein Data Bank (http://www.rcsb.org/pdb/index.html) under PDB # 1L4O.

[d] The atomic coordinates for the structure of this hexamer/connector (nucleic acid/protein) complex are available in The Protein Data Bank (http://www.rcsb.org/pdb/index.html) under PDB # 1L4P.

Table Legend

Table 1. Interlocking pRNAs form a hexamer. Different combinations of two, three, or six mutated pRNAs are used to demonstrate that six pRNAs form a hexamer.

Figure Legends

Fig. 1**. Secondary structure of pRNA and procapsid/hexamer complex.** A. Diagram showing the predicted pRNA secondary structure. The right and left-hand loops, the head loop the $U^{72}U^{73}U^{74}$ bulge and the $C^{18}C^{19}A^{20}$ bulge are in boxes. The DNA packaging domain (5'/3' ends) and the procapsid binding domain (the larger area) are shaded. The curved line points to the two interacting loops. B. Diagram showing hand-in-hand-interaction between six pRNA monomers to form a hexamer. The hexamer is shown to bind to the connector (the hashed hexagon) on the procapsid.

Fig. 2. **Computer models showing the monomer, dimer, hexamer, and connector**. A. The model of monomer in spacefill format showing the $U^{72}U^{73}U^{74}$ bulge (in white), the right (in red) and left- (in green) hand loop. B. The model of dimer in spacefill format with one unit in blue and the other unit in yellow. The right and left-hand loops are highlighted in red and green, respectively. C. The model of hexamer in spacefill format showing the procapsid binding domain in green, and the DNA translocating domain in red (the 5'-end) and cyan (the 3'-end). The DNA translocating domain of the 5'/3'-paired region points down and to the left. D. The crystal structure of the connector (12) in wire-frame format. The RNA Recognition Motif (RRM)(68) (19)is colored blue. E. Docking of the pRNA hexamer to the RNA binding domain (RRM) of the connector. The connector-binding domain is in green and the DNA translocating

domain is in red and cyan. F. Illustration of the pRNA hexamer/connector complex as part of phi29.

Fig. 3. **Activity (pfu/ml) of pRNAs by complementary modification**(49) (50) (51). Complementary modification verifies predicted secondary structure by comparing the activity, or lack of activity, to the wild-type sequence.
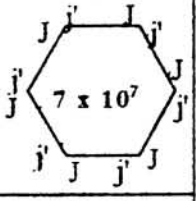
Fig. 4**. Computer model of pRNA monomer to illustrate the results of complementary modification.** Bases mutated in the complementary modification studies are shown in spacefill mode in the 3D model. If the secondary site complementary mutation could restore the pRNA activity, these bases are presented as a helical stretch. The model shows that bases 1-2 are paired with bases 117-116; bases 7-9 are paired with bases 112-110; bases 14-16 are paired with bases 103-101; and bases 76-78 are paired with 90-88.

Fig. 5. **Computer model of pRNA monomer to illustrate the results of intramolecular psoralen photo affinity crosslinking.** The model reflects the experimental data that U^{69} (black) crosslinks to U^{31} and U^{33} U^{36} (gray).

Fig. 6. **Computer model of pRNA monomer to illustrate the results of intramolecular phenphi photo affinity crosslinking**. The model reflects the experimental data that G^{75} (black) crosslinks to G^{28} and G^{30} (gray).

Fig. 7. **Computer model of pRNA monomer to illustrate the results of intramolecular azidophenacyl photo affinity crosslinking.**   A. G^{75} (black) crosslinks to A^{26}, U^{27}, G^{28}, U^{29}, and G^{30} (gray). B. G 78 (black) crosslinks to U^{31} (gray), and G^{108} (black) crosslinks to C^{10} and G^{11} (gray).

Fig. 8. **Comparison of chemical modification pattern of monomer (A) and dimer (B).** The black arrow, gray square, and double-lined arrow indicate a strong, moderate, and weak modification, respectively. C is a model to portray the formation of dimer. The four base-paired (45-48/85-82) were modified in monomer, but were protected from chemical modification in dimer.

Fig. 9. **Computer model of pRNA monomer to illustrate the results of chemical modification in the presence of Mg^{2+}.** Heavily modified bases, moderately modified bases, and lightly modified bases are shown in black sticks, gray sticks, and light gray sticks, respectively. It is notable that the single-stranded right-hand loop, head loop, left-hand loop, and the CCA bulge are in black sticks, indicating a strong modification.

Fig. 10. **Phylogenetics analysis of pRNA.**   Phylogenetic analysis of pRNAs from *Bacillus subtilis* phages SF5, B103 (59), phi29, PZA, M2, NF, and GA1 (60), shows very low sequence identity and few conserved bases, but their predicted secondary structures resemble each other (25;52).   All seven pRNAs of these phages contain both the right and left loops with complementary sequences. The dimer model of phi29 pRNA (F) is in concordance with the data of phylogenetic analysis concerning intermolecular loop/loop interaction in dimer formation.

The bases shown in black spacefill format (bases 45-48) and gray black spacefill format (base 85-82) represent the right and left-hand loop, respectively.

Fig. 11. **Complementary Modification of pRNAs in the $U^{72}U^{73}U^{74}$ region suggests that the UUU sequence provides flexibility to pRNA**. In pRNA F5, the native 5'/3' ends have been kept and U^{72}, U^{73}, and U^{74} have been deleted, resulting in an inactive pRNA(51). In cpRNA (circularly permutated pRNA), the native 5'/3' ends have been joined by an AAA sequence and U^{72}, U^{73}, and U^{74} have been deleted to make new 5'/3' ends, resulting in an active pRNA molecule(50).

Fig. 12. **Comparison of cryo-AFM images (A and B) with computer models (C and D).** The direct observations of the monomer (A) and dimer (B) by Cryo-AFM are compared to the 3D structure of monomer (C) and dimer (D) observed from different viewpoint. The color in cry-AFM images indicates the thickness and tallness of the image, but does not reflect the atom density observed end on. The brighter or whiter the color, the taller the surface is in the image. The darker the color, the lower the surface is in the image. Dimers were about twice the length of monomers. The models of monomer in C and dimer in D have been tuned from different angle and aligned with the AFM images.
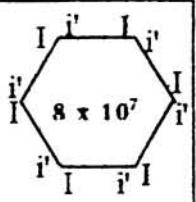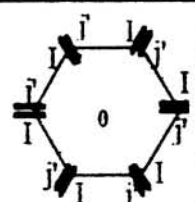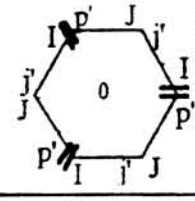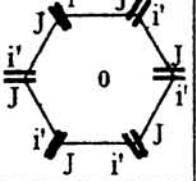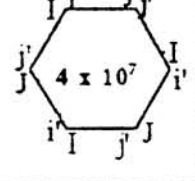
Fig. 13. **Computer model of pRNA dimer to illustrate the results of intermolecular azidophenacyl photo affinity crosslinking.** The dimer model is in agreement with the empirical data showing that G^{82} (in black spacefill) in one pRNA unit is proximate G^{39}, G^{40}, A^{41}, C^{49}, G^{62}, C^{63}, and C^{64} (in gray wireframe) of the other pRNA unit.

Fig. 14. **Computer model of pRNA dimer to illustrate the results of chemical modification interference.** Bases that are demonstrated to interfere with dimer formation are shown as gray spacefilled bases in the pRNA subunits. The dimer model is in agreement with the empirical data by showing that these bases are located at the interface of two pRNA.
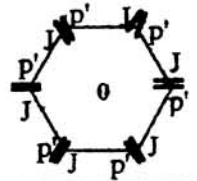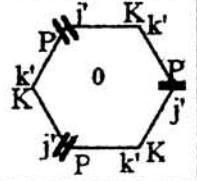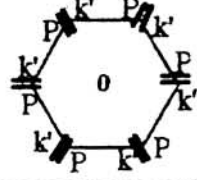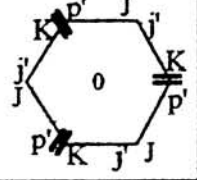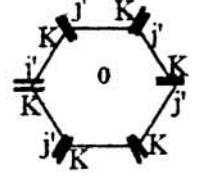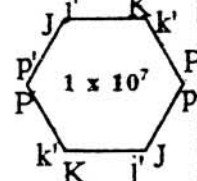
Fig. 15. **A model to depict the assembly of pRNA hexamer from three dimers.** pRNA dimer is produced in solution via the hand-in-hand and head-to-head contact. Binding of pRNA dimer to connector, which is composed of 12 subunits of protein gp10, results in conformational change of pRNA dimer that release one pair of interacting hands (II). The free hand is used to recruit the on coming dimer via hand-in-hand interaction (III). Sequential addition of three dimers resulted in the formation of pRNA hexamer.

Table 1.

Two Interlocking pRNAs

pRNAs	Predicted Hexamer	pRNAs	Predicted Hexamer
I-i' (wild type pRNA)	8×10^7	I-j' (unpaired loop)	0
J-j' (Compensatory modification)	7×10^7	(I-j') + (J-p') (miss one link)	0
J-i' (Unpaired loop)	0	(I-j') + (J-i') (Compensatory pair)	4×10^7

Three Interlocking pRNAs

pRNAs	Predicted Hexamer	pRNAs	Predicted Hexamer
J-p' (Unpaired loop)	0	(P-k') + (K-j') (miss one link)	0
P-k' (Unpaired loop)	0	(K-j') + (J-p') (miss one link)	0
K-j' (Unpaired loop)	0	(J-p') + (P-k') + (K-j') (compensatory trimer)	1×10^7

Six Interlocking pRNAs
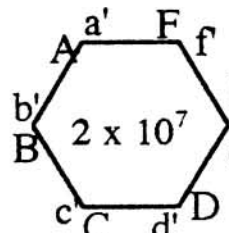
pRNAs	Predicted Hexamer
(A-b')+(B-c')+(C-d')+ (D-e')+(E-f')+(F-a')	2×10^7

Fig. 1

Fig. 2

A. Monomer

B. Dimer

C. pRNA hexamer

7.6 nm

D. Connector

13.8 nm

6.8 nm

9.4 nm

E.

F. Procapsid
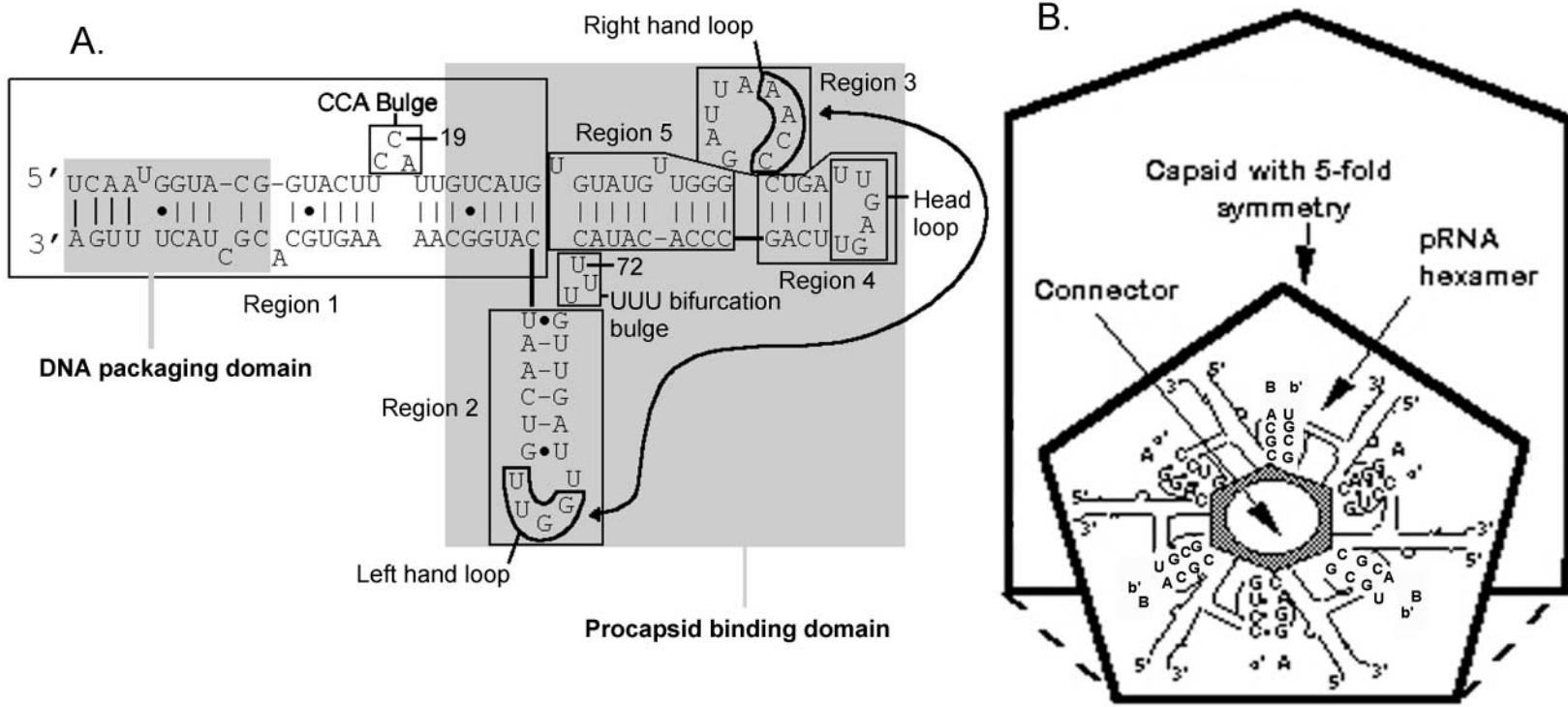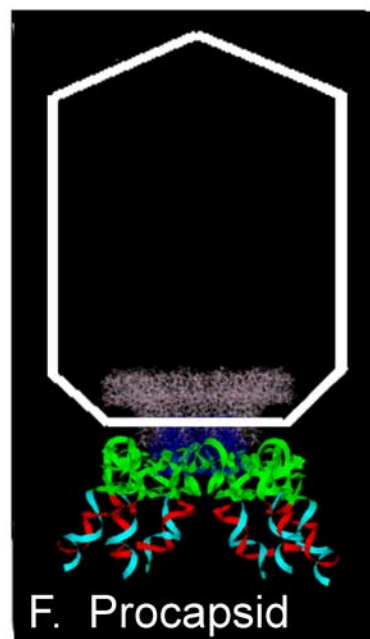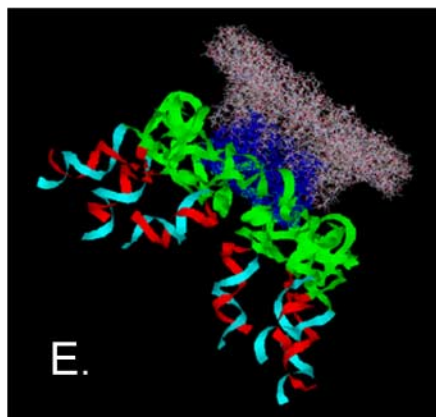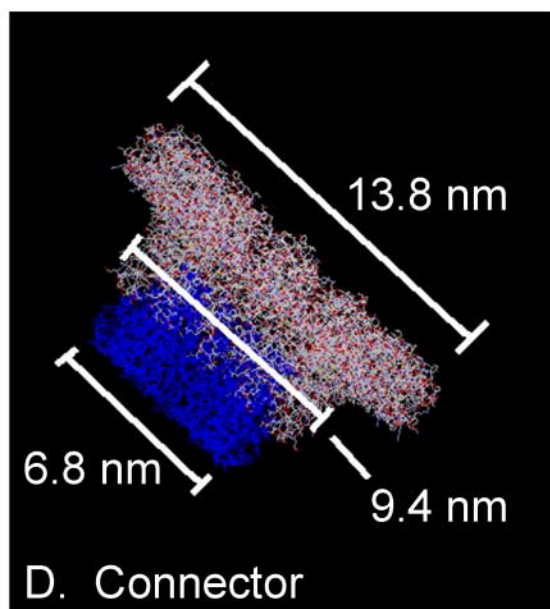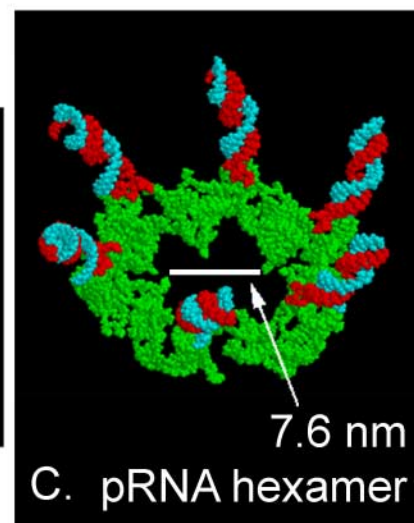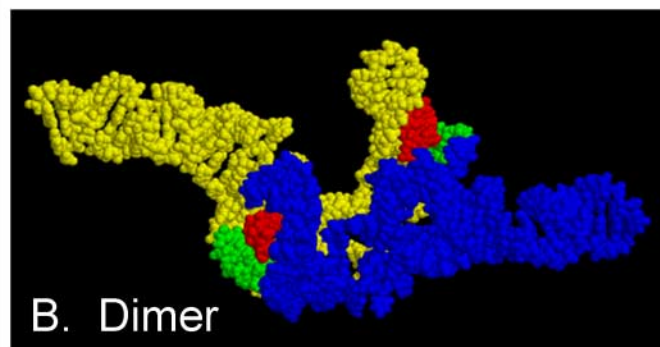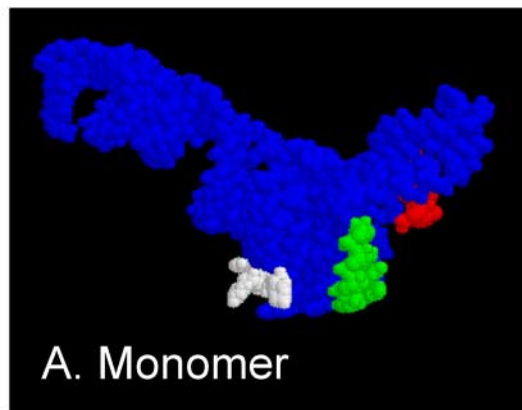
Fig. 3

I. Bases 1-4 paired to 117-114

pRNA FW/4

1.6×10^{2} pfu/mL

pRNA 7/27

0 pfu/mL

pRNA FW/RV

1.6×10^{7} pfu/mL
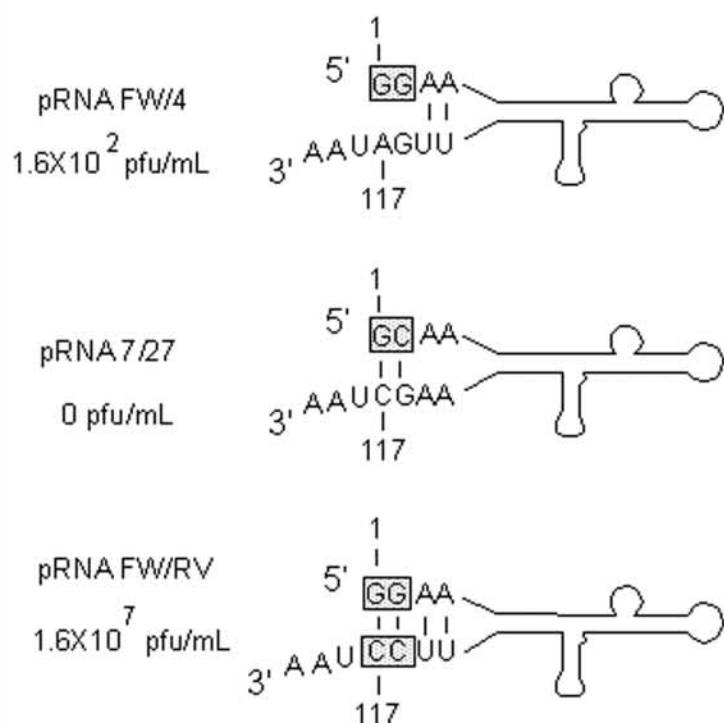
II. Bases 14-16 paired to 103-101

pRNA 14-16/10

80 pfu/mL

pRNA 7/101-103

2×10^{4} pfu/mL

pRNA 14-16/101-103

9×10^{7} pfu/mL
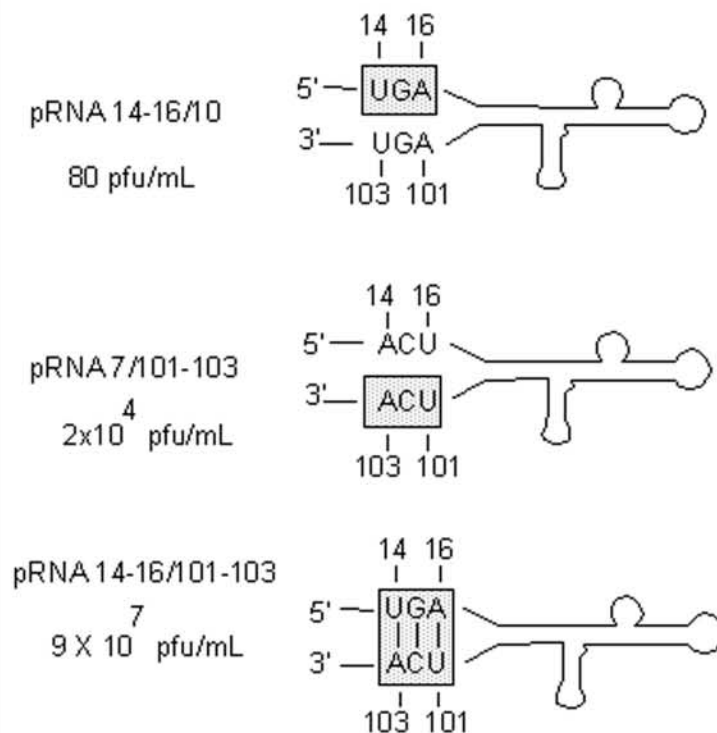
III. Bases 7-9 paired to 112-110

pRNA 7/29

0 pfu/mL

pRNA 28/11

0 pfu/mL

pRNA 28/29

1.6×10^{7} pfu/mL
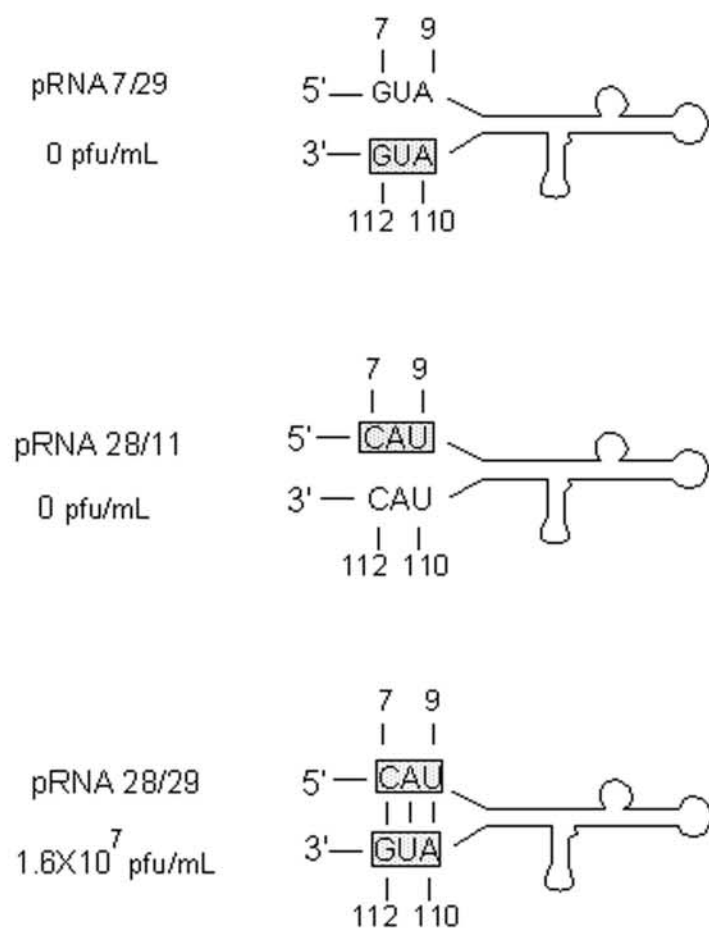
IV. Bases 76-78 paired to 90-88

pRNA F3
inactive

pRNA F3/A5
active

Figure 4

Figure 5

U 69

U 31
U 33
U 36

Fig. 6

Figure 7

A.

B.
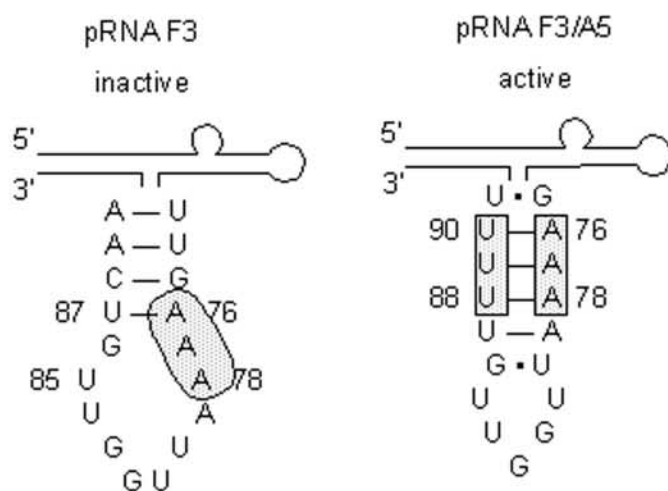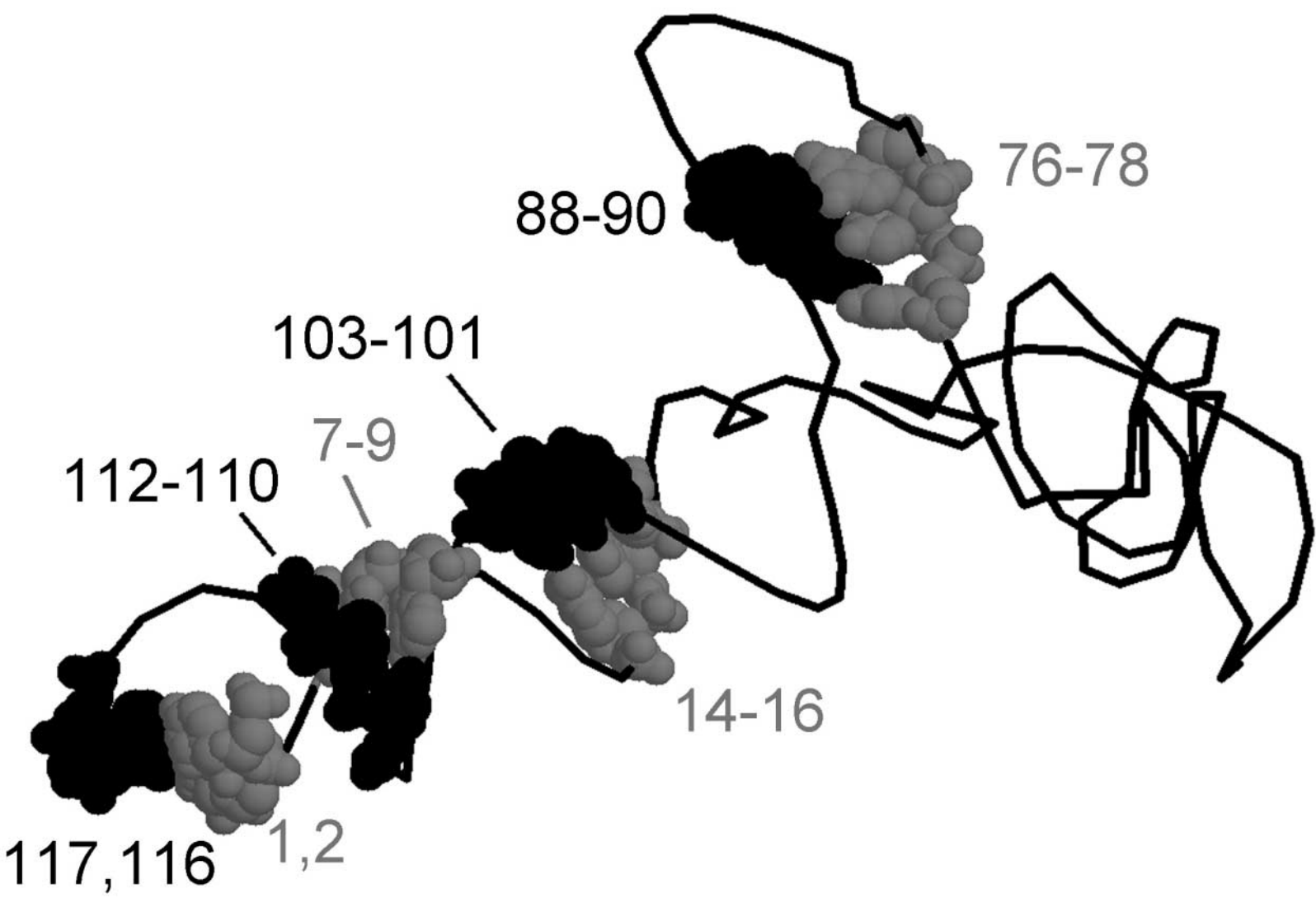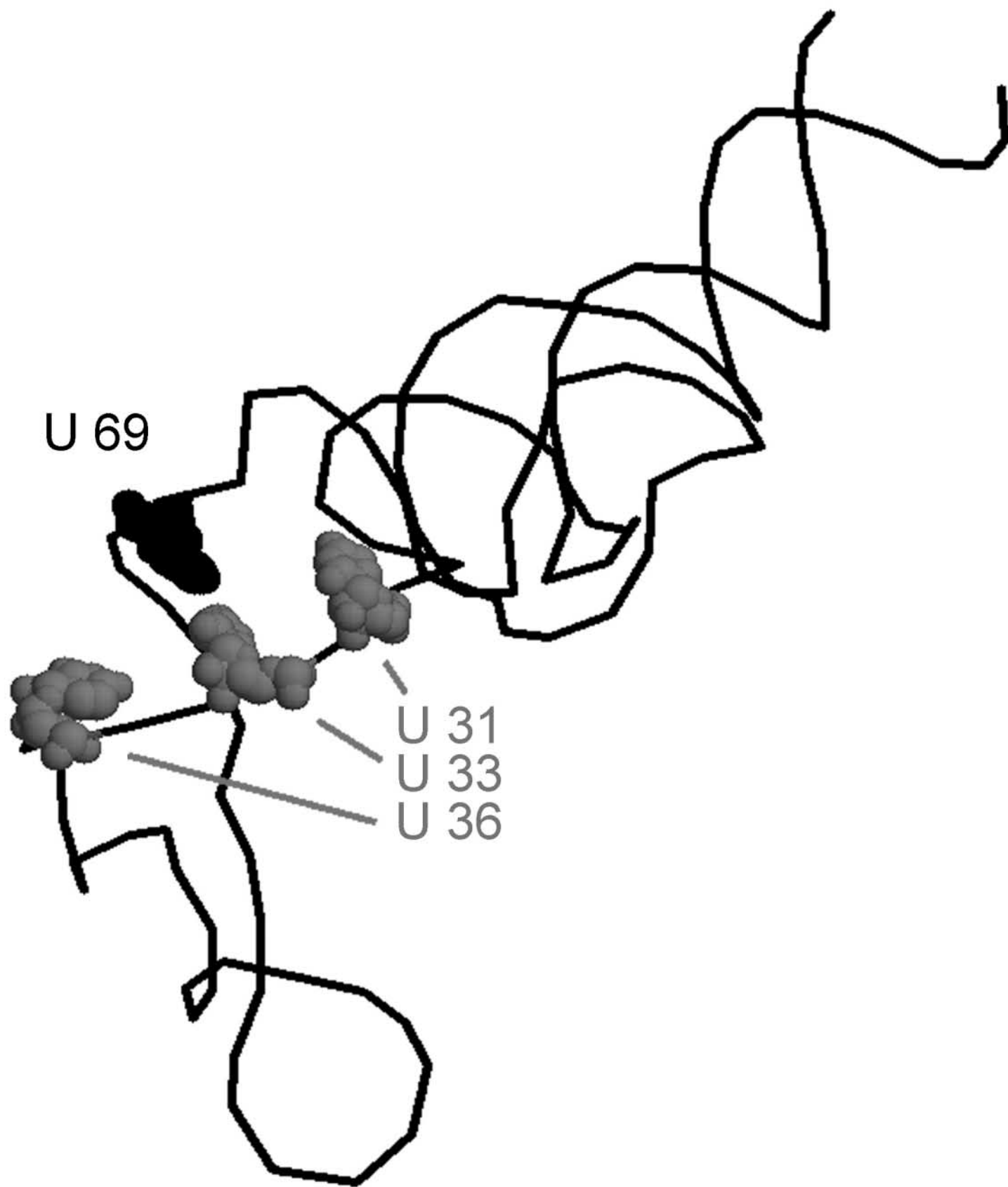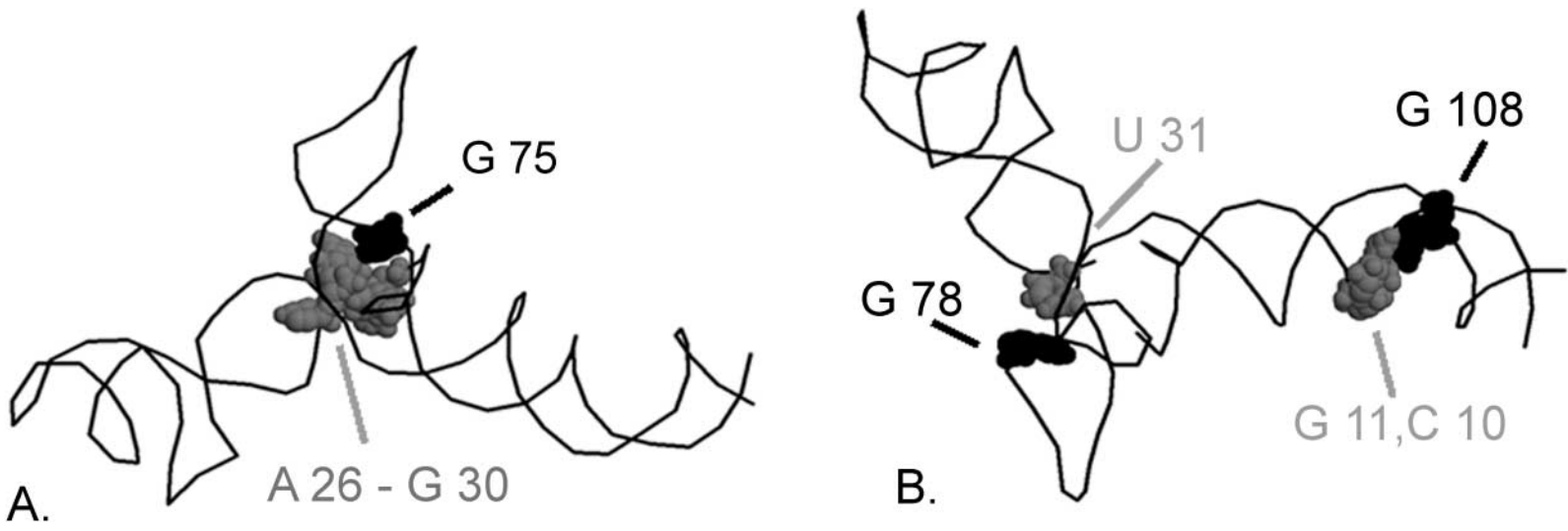
Fig. 8
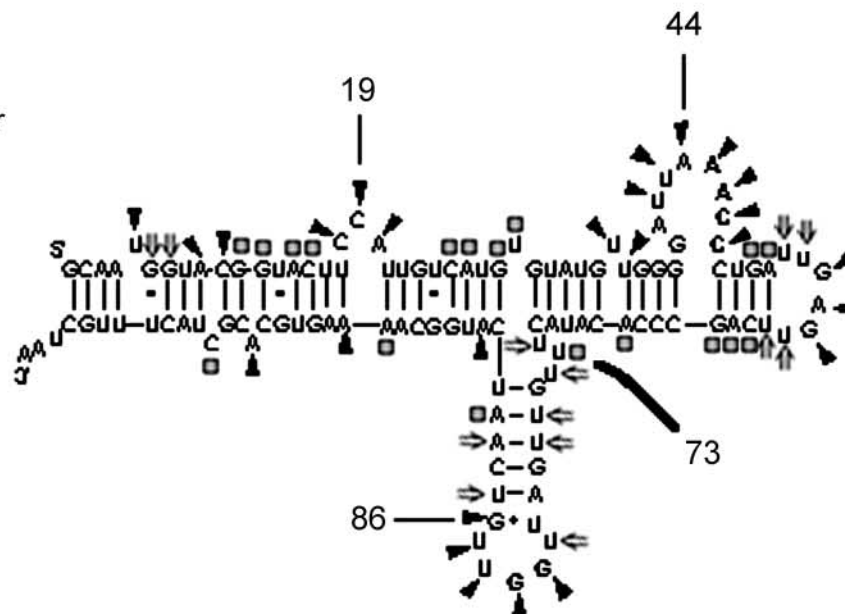
A. Monomer

B. Monomer subunit of the dimer

C. Dimer

Figure 9

head
loop

left-hand loop

CCA bulge

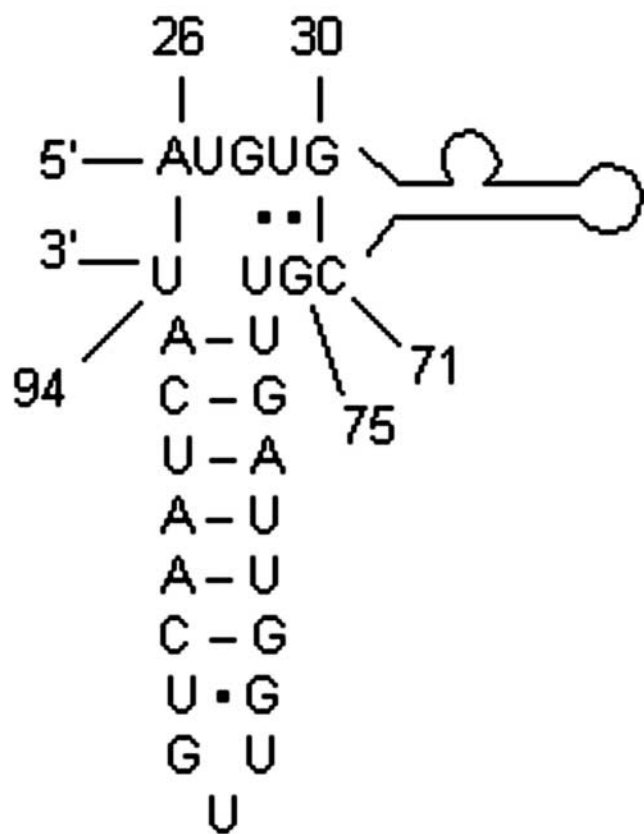right-
hand
loop
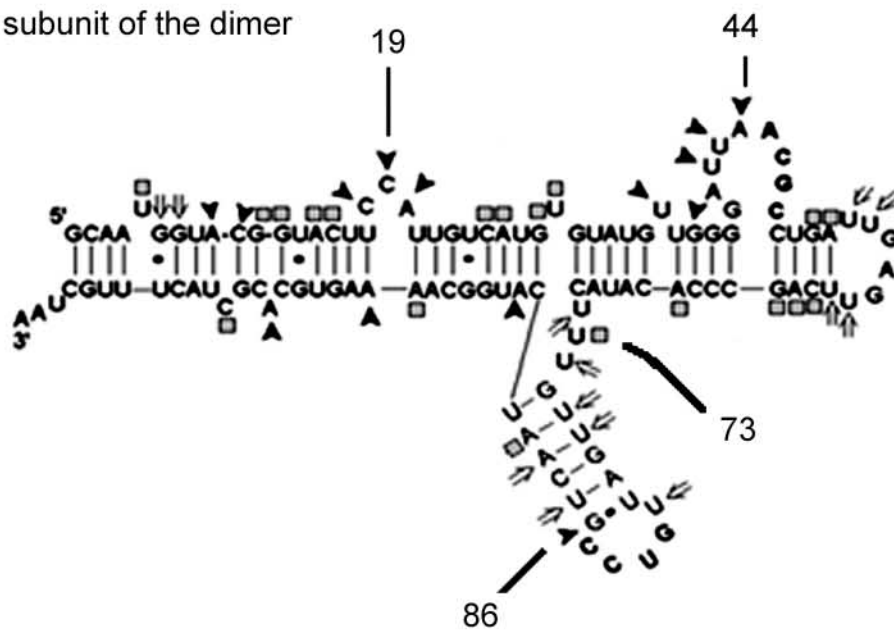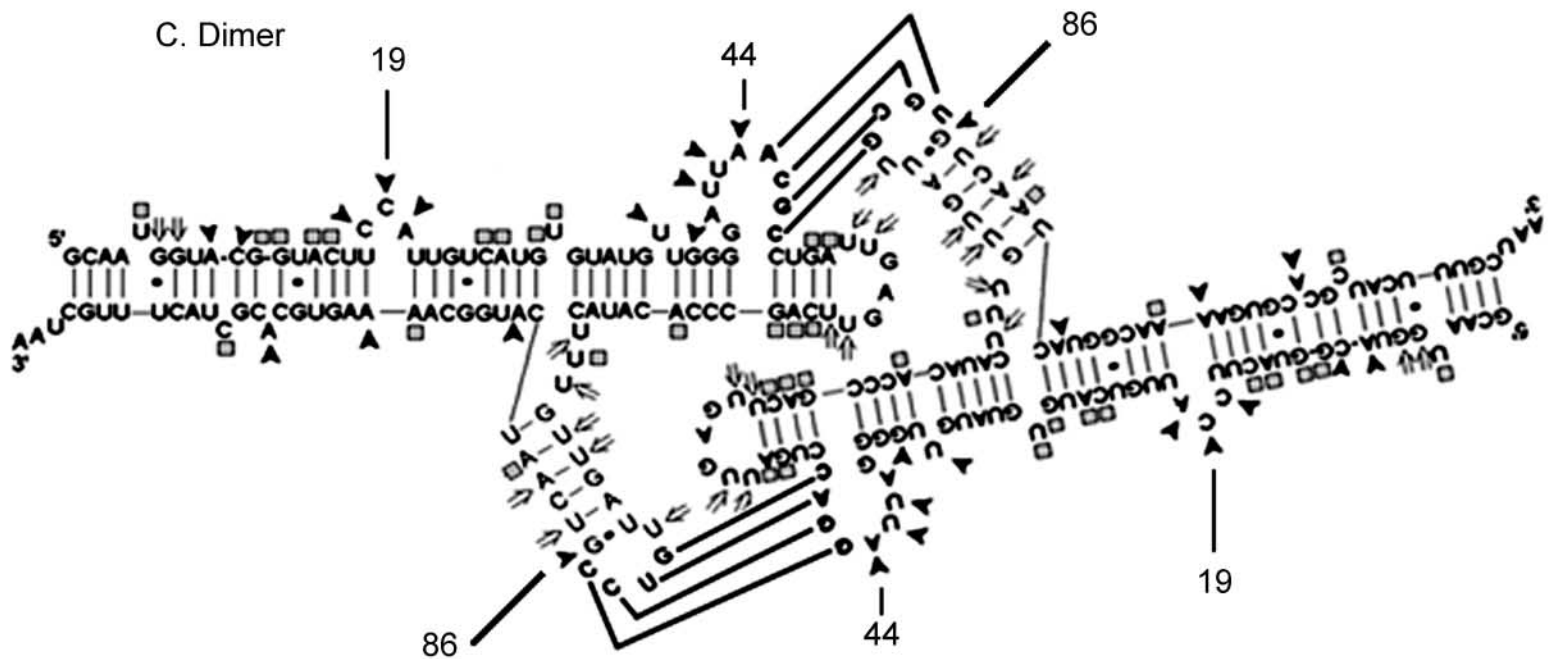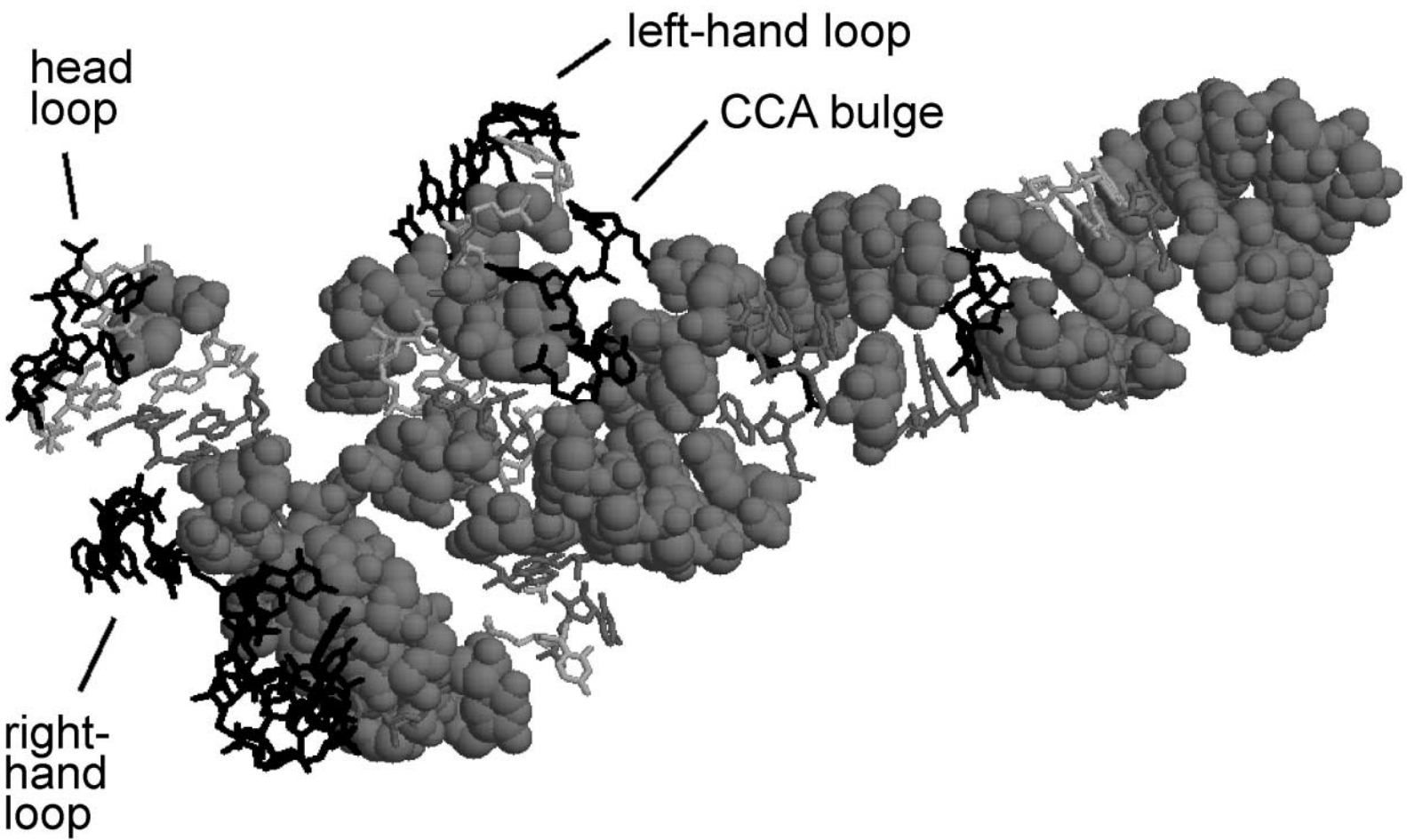
Figure 10

A. B103
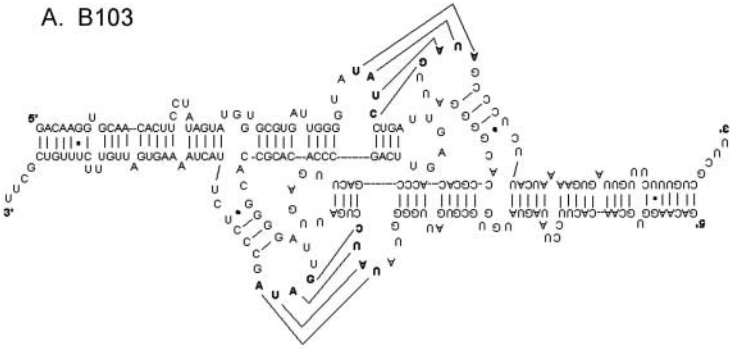
D. GA1
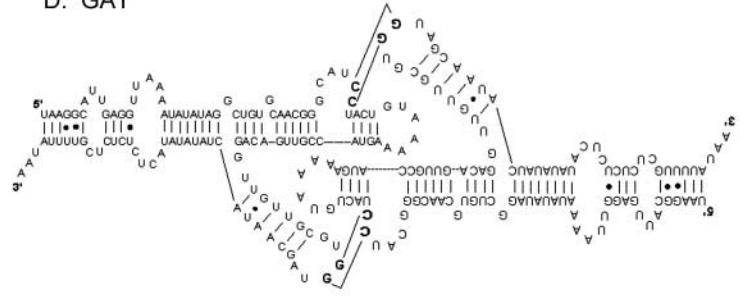
B. M2-NF
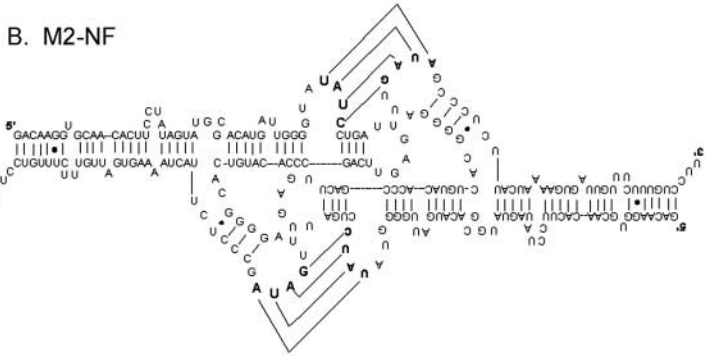
E. Phi29-PZA
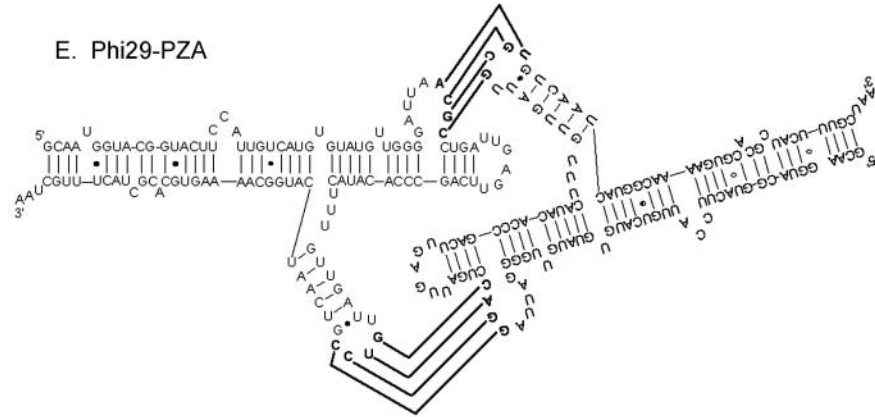
C. SF5
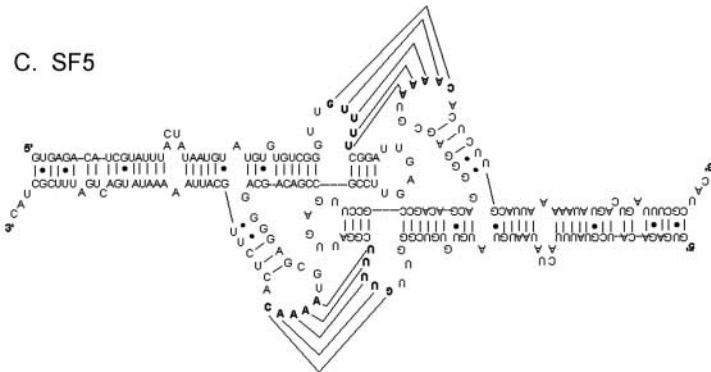
F. Phi29 pRNA computer model

Fig. 11

pRNA F5
inactive

cpRNA 75/71
active
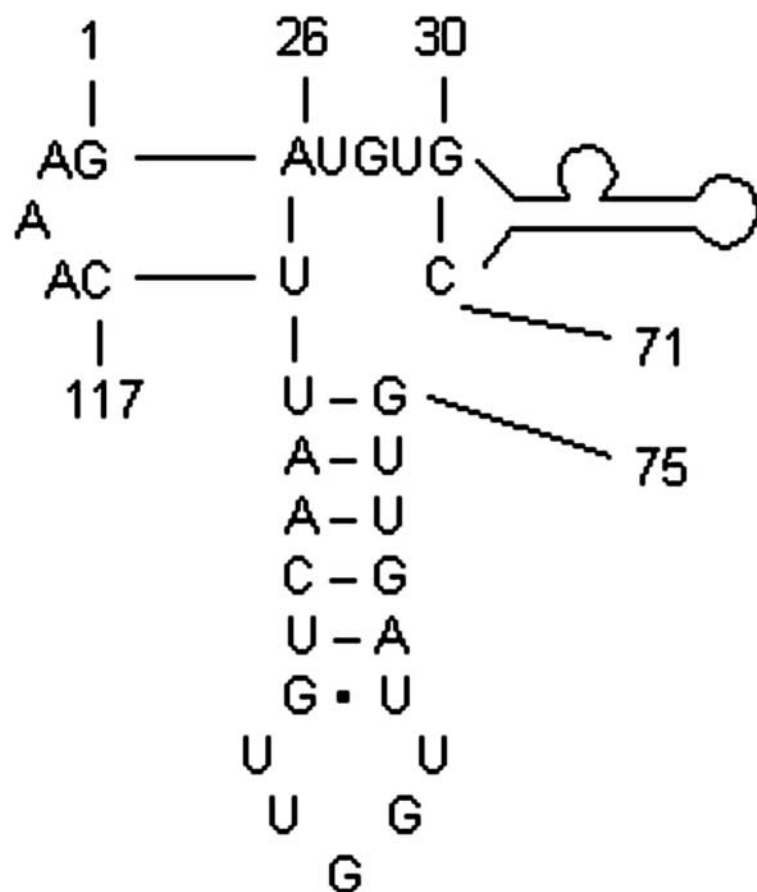
Figure 12

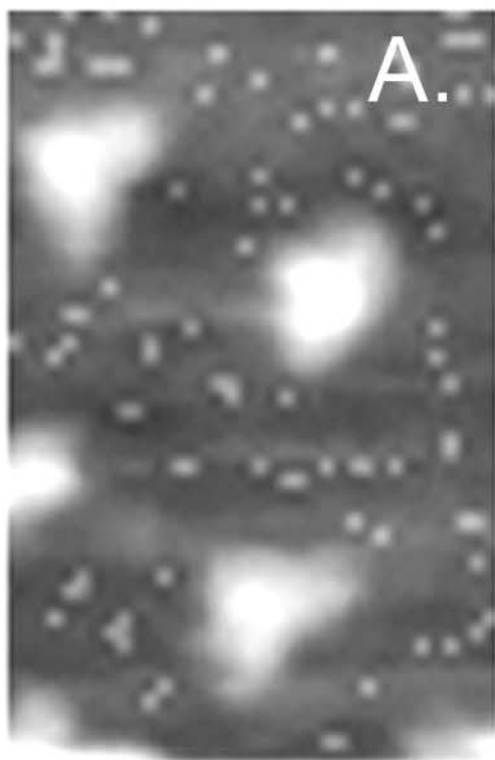**Monomer**

Dimer

Cryo-AFM

Computer Model

Figure 13

G 82

G39, G40,
A41, C49,
G62, C63,
C64

G39, G40,
A41, C49,
G62, C63,
C64

G 82

Figure 14

Fig. 15

pRNA dimer in solution

Procapsid shell

Connector

I   II   III   IV