Mutants of the Bacteriophage MS2 Coat Protein That Alter Its Cooperative Binding to RNA†

Karen A. LeCuyer, Linda S. Behlen, and Olke C. Uhlenbeck*

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215

Received March 17, 1995; Revised Manuscript Received June 7, 1995®

ABSTRACT: An RNA binding assay measuring cooperative protein binding has been used to evaluate the effects of mutations in the MS2 phage coat protein expected to disrupt capsid assembly. By using the crystal structure of the virus as a guide, six different mutations in the FG loop structure were selected in which hydrophobic residues were replaced with charged residues. Most of these proteins form capsids in *Escherichia coli*, but not in an *in vitro* assembly assay, suggesting that interdimer interactions are weaker than wild type. These mutant proteins reduce the free energy of cooperative protein binding to a double-hairpin RNA from its wild-type value of −1.9 kcal/mol. Several of the variants that have large effects on cooperativity have no effect on RNA affinity, suggesting that protein−RNA interactions can be affected independently of dimer−dimer interactions. The V75E;A81G protein, which shows no measurable cooperativity, binds operator RNA equally well as the wild-type protein under a variety of buffer conditions. Because this protein also exhibits similar specificity for variant RNA sequences, it will be useful for studying RNA binding properties independent of capsid assembly.

Assembly of small *Escherichia coli* RNA phages such as MS2 and its close relative R17 occurs when 90 dimers of coat protein combine with one copy each of the genomic RNA and the A protein to form a $T = 3$ icosohedral virus. While the coat protein can assemble into empty capsids *in vitro*, the reaction is stimulated by the presence of RNA (Hohn, 1969; Matthews & Cole, 1972; Beckett & Uhlenbeck, 1988; Beckett et al., 1988; Mastico et al., 1993). A small RNA hairpin containing the initiation site of the replicase gene is particularly effective in stimulating assembly and is believed to be the packaging signal for the coat protein polymerization process (Hung et al., 1969; Ling et al., 1969; Beckett et al., 1988; Pickett & Peabody, 1993). The interaction of the coat protein dimer with the initiation site of the replicase gene is also involved in the regulation of replicase gene expression (Bernardi & Spahr, 1972). The affinity of the R17 phage coat protein dimer for variant hairpin RNAs has been studied extensively (Witherell et al., 1991).

The crystal structure of the MS2 bacteriophage reveals extensive protein−protein interactions between dimers (Valegård et al., 1990). The 15-residue loop between the F and G β-strands has received particular attention, as it is involved in numerous protein−protein contacts along both the 5- and 6-fold symmetry axes. This FG loop structure constitutes the main conformational difference in the 6-fold symmetry axis, with the two A and C subunits having extended loops and the B subunits having a more compact loop structure (Golmohammadi et al., 1993) (Figure 1). These different loop structures result in very different protein−protein contacts. The presence of a *cis*-proline in the FG loop of the B subunit has been proposed as a reason for the different B subunit packing (Valegård et al., 1990). The proximity of the 5′ end of the hairpin stem to the FG loop in the structure of the protein−RNA complex suggests that residues in the FG loop could influence RNA binding (Valegård et al., 1994). If RNA binding induces formation of the more compact B loop conformation, it is possible that the AB′ dimer−RNA complex constitutes the assembly initiation complex (Figure 1).

Mutations and deletions in the FG loops of both the MS2 and *fr* coat proteins have been shown to reduce the efficiency of capsid assembly (Peabody & Ely, 1992; Pushko et al., 1993). The extremely cooperative nature of the assembly reaction, however, makes it difficult to quantitatively evaluate the effect of such mutations on the dimer−dimer interaction. This difficulty is overcome by an assay that measures the cooperative binding of coat protein dimers to an RNA containing two protein binding sites (Witherell et al., 1990). The cooperative free energy measured in such an assay represents the contribution from a single interdimer interaction. In this work, we examine the effects of several mutations in the FG loops of the MS2 coat protein. These mutations are evaluated by using an *in vitro* encapsidation assay, as well as the quantitative cooperativity assay.

EXPERIMENTAL PROCEDURES

Coat Protein Expression and Purification. The overexpression system of Studier and Moffat (1986) was used to isolate large amounts of the MS2 bacteriophage coat protein. The expression plasmid for these experiments, pKCO, was constructed from the overexpression plasmid pTCT5 (Gott et al., 1991) and pJC4, a plasmid containing a coat protein gene with several new restriction sites (Giebler, 1992). pKCO contains the T7 promoter followed by 832 bp of sequence from the 3′ end of the phage maturation protein and the entire coat protein gene. In addition to the original sites for *Xba*I (−32), *Sal*I (+30), and *Eco*RI (+294), pKCO contains sites for *Nru*I (+113), *Bst*XI (+215), and *Bgl*II

† This work was supported by NIH Grant GM36944 to O.C.U. and by a postdoctoral fellowship from the American Cancer Society to K.A.L.

* Corresponding author.

® Abstract published in *Advance ACS Abstracts,* August 1, 1995.

0006-2960/95/0434-10600$09.00/0

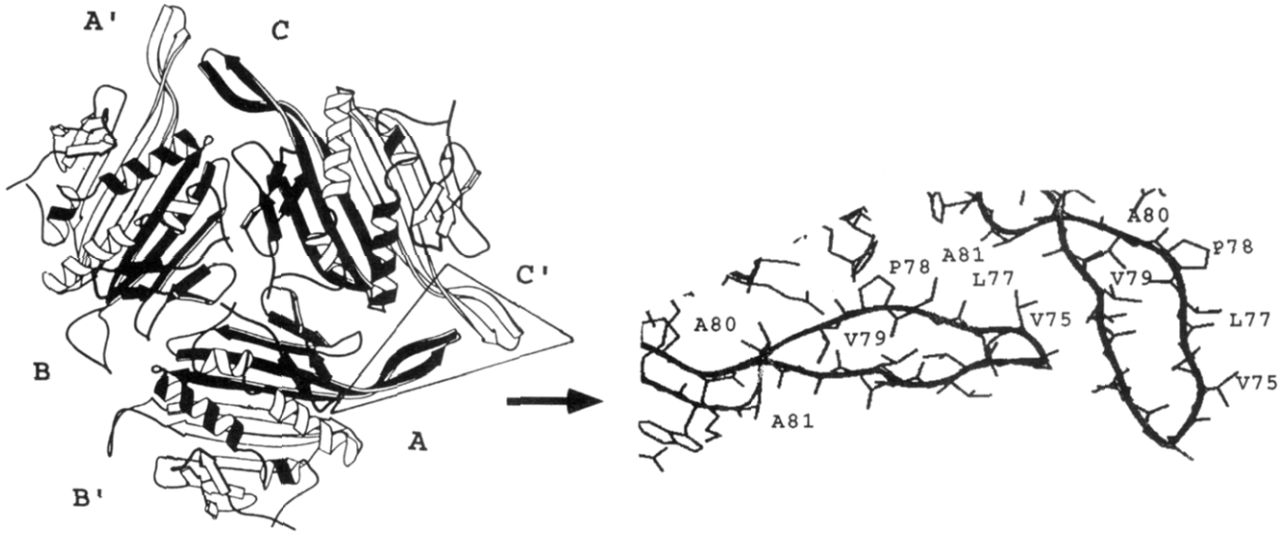

FIGURE 1: Schematic diagram of the AB′, A′B, and CC′ dimers in one hexamer of the R17 coat protein (Valegård et al., 1990). An enlarged view of one of the FG loop interactions of an AC pair is shown.

(+248) (numbering from the A of the AUG start codon). The FG loop variants selected for analysis were generated from random cassette mutagenesis of amino acid residues G73−W82 (Giebler, 1992). The *Xba*I/*Eco*RI fragments containing the variant coat protein genes were subcloned from the parent pJC4 into pKCO for overexpression. The coat protein plasmids were transformed into *E. coli* BL21-(DE3) (Studier & Moffatt, 1986). Proteins were prepared from capsids as described in Gott et al. (1991).

The V75E;A81G variant does not form capsids and was purified by a new method. *E. coli* Bl21(DE3)pLysS cells (Studier, 1991) containing the V75E;A81G plasmid were grown to mid-log phase in YTG medium supplemented with 150 μg/mL ampicillin and 30 μg/mL chloramphenicol. IPTG was added to a final concentration of 0.4 mM for T7 RNA polymerase production, and cells were incubated with shaking for an additional 3 h at 37 °C. Cells were harvested by centrifugation for 10 min at 8000*g* and then frozen. The frozen cells were thawed and resuspended in 50 mM NH_4-Cl, 20 mM Tris-Cl (pH 8.0), 3 mM DTT, 1 mM EDTA, and 5% glycerol with shaking for 1 h. The lysates were then sonicated with three 30-s pulses to assure complete resuspension and lysis. The lysate was centrifuged at 27000*g*, 4 °C for 50 min and then subjected to further centrifugation for 90 min at 186000*g*, 4 °C. The supernatant was ammonium sulfate-precipitated, with the fraction that precipitates between 35 and 50% ammonium sulfate used for further purification. The pellet was resuspended in buffer A [25 mM HEPES (pH 7.5), 25 mM NaCl, 0.1 mM EDTA, 2 mM DTT, and 0.02% NaN_3] and dialyzed overnight against buffer A to remove any residual ammonium sulfate. FPLC was performed on a HiLoad 16/10 SP Sepharose HP column (Pharmacia) with buffer A and buffer B (the same as buffer A but with 1 M NaCl). The coat protein was loaded at 1 mL/min in buffer A and equilibrated. Elution was performed with a gradient of buffer B from 0 to 50% B in 1 h at 3 mL/min. The coat protein eluted at approximately 20% B. Fractions were pooled and concentrated using a Centriprep-10 (Amicon). The protein is >95% pure as estimated from Coomassie Blue staining of SDS−PAGE gels. The protein concentration is calculated as a dimer using a molar extinction coefficient of 1.54×10^4 at 280 nm.

RNAs. RNA1 was prepared by *in vitro* transcription from synthetic DNA templates with T7 RNA polymerase (Gott et al., 1991). RNA2 was prepared by *in vitro* transcription from a plasmid template, treated with calf intestinal alkaline phosphatase, and end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP (Witherell et al., 1990). PacA+ RNA for assembly experiments was transcribed with [α-^{32}P]CTP as in Beckett et al. (1988), but was purified on a 5% denaturing polyacrylamide gel followed by crushing and soaking in 0.5 M ammonium acetate and 2 mM EDTA and then ethanol precipitation.

Nitrocellulose Filter Binding. Protein excess filter binding was performed in TMK buffer [100 mM Tris-Cl (pH 8.5 at 4 °C), 10 mM $MgCl_2$, 80 mM KCl] with 80 μg/μL BSA. A constant (<100 pM) concentration of ^{32}P-labeled RNA was mixed with varying concentrations of coat protein in a reaction volume of 150 μL. After incubation on ice for 90 min, the reactions were filtered through a 0.45-μm Nitro ME filter (MSI, Westborough, MA). Filtration was performed in a Schleicher and Schuell dot-blot apparatus modified as in Wong and Lohman (1993). The amount of radioactivity retained on the filter was quantitated by using a Molecular Dynamics phosphorimager. For each experiment with RNA1, the data points were fit to a retention efficiency and a K_d[1] value assuming a bimolecular equilibrium. For each experiment with RNA2, the K_d value determined from binding to RNA1 was used in the cooperative binding equation to calculate a value for K_c, the cooperativity constant (Witherell et al., 1990). Each binding curve was repeated at least three times, and the reported values for K_d and K_c represent averages. The K_d values are reproducible within a factor of 2 from the reported value.

To determine the fraction of protein that is active for RNA binding, 50 nM protein dimer was titrated with a series of RNA concentrations from 0.5 to 200 nM (Witherell & Uhlenbeck, 1989). All of the proteins reported in this work were between 80 and 95% active. The K_d values determined in the protein excess experiments were subsequently corrected for the fraction of active protein.

[1] Abbreviations: K_d, dissociation constant; K_c, cooperativity constant.

Capsid Assembly. Capsids were reconstituted as in Beckett et al. (1988). Briefly, PacA+ RNA and coat protein dimer were incubated in TMK buffer (pH 8.5) for 12 h at 4 °C. Coat protein concentrations were 2 μM for wild type, 40 μM for L77R, 70 μM for V79E, and 100 μM for the others. PacA+ RNA was added in a 1:90 molar ratio of RNA to protein, with the RNA concentration determined from the optical density at 260 nm. Capsids were separated from free RNA on a 1% agarose gel (18 × 18 cm) run in 50 mM phosphate buffer (pH 7.0) with 1 mM $MgCl_2$. Gels were run at 100 mA for 10−12 h, dried, and visualized on a Molecular Dynamics phosphorimager. As a standard for capsid migration, a capsid capture experiment was performed in which the wild-type capsid was incubated with ^{32}P-labeled RNA1. It has previously been observed that assembled capsids will take up small operator RNAs (Beckett et al., 1988).

RESULTS

The crystal structure of the MS2 bacteriophage at 2.8 Å resolution has recently been solved (Valegård et al., 1990; Golmohammadi et al., 1993). In this structure, each of the 180 coat protein monomers forms extensive contacts with another monomer to form a dimer. The coat protein dimer is the basic structural unit and is the form of the coat protein that is favored in solution (Shafranski et al., 1975; Beckett & Uhlenbeck, 1988). The coat protein monomer consists of a 5-stranded antiparallel β-sheet facing the inside of the virus and two α-helices on the outer surface. Most of the connections between the β-sheets are short turns or loops, except the 15-residue loop connecting the F and G β-strands. This FG loop is found at the 5-fold and quasi-6-fold axes of symmetry in the assembled virus. While the FG loop clearly is involved in interdimer interactions, it is external to the central fold of the protein and thus is not likely to be critical for the formation of the dimer in solution. The conformation of the FG loops differs in the three different dimer interactions. The CC′, AB′, and A′B dimers make contact at the quasi-6-fold axis of symmetry to give two AC and one B′B loop−loop interactions. Figure 1 shows the three dimers at the quasi-6-fold axis of symmetry with an enlarged view of an AC loop interaction.

To evaluate the effect of mutation of the FG loops on interdimer interactions, six single or double FG loop mutations were selected from a library generated in a genetic screen of the coat protein binding site (Giebler, 1992). The mutations that were chosen introduce a charged amino acid in place of a hydrophobic residue at different positions in the FG loop protein−protein interaction face. The replacement of at least one hydrophobic residue with a charged residue is expected to alter the protein packing interactions (Dill, 1990). A more detailed description of the environment of each mutated amino acid is found in Table 1.

Most of the variant proteins listed in Table 2 formed capsids when overexpressed in *E. coli,* indicating that interdimer contacts still form. It has been shown that a mutant deficient for dimer formation does not form capsids and does not bind RNA (Peabody & Ely, 1992). These proteins were purified in a manner analogous to that of the overexpressed wild-type protein (Gott et al., 1991). This protocol involves centrifugation, a DEAE column, and subsequent denaturation with acetic acid followed by renaturation by dialysis. The V75E;A81G variant did not form capsid particles when overexpressed, suggesting the presence of much weaker protein−protein interactions. Because this protein does not form capsids, it was purified by a protocol involving ammonium sulfate precipitation and S-Sepharose chromatography.

Table 1: Interactions Made by FG Loop Residues[a]

residue	AC interactions	BB′ interactions	comments
Val75	intra, Glu76 inter, Ala68, Gln70	intra, Leu77 inter, Gln54	tip of loop
Leu77	inter, Ala81	intra, Thr71 inter, Thr69	middle of loop
Pro78	inter, Phe82	intra, Trp82, Thr69	*cis/trans* isomerization
Val79	intra, Ala68, Thr69	intra, Gln70, Val72 inter, Gln40	middle of loop
Ala80	intra, Ala68, Val67	intra, Thr97, Leu103	twist at base of loop
Ala81	intra, Ala68, Val67 inter, Leu77	intra, Thr97, Leu103	twist at base of loop

[a] Intra refers to intradimer interactions, and inter refers to interactions between two dimers.

Table 2: Capsid Assembly with FG Loop Variants

variant	capsid formation in *E. coli*	concentration required for assembly *in vitro* (μM)
wild type	+	2
V79E	+	>70
L77R	+	40
V75E;V79E	+	>100
P78S;V79R	+	>100
A80E	+	>100
V75E;A81G	−	>1000

Selective encapsidation of the viral RNA is required for the assembly of infectious viral particles. An *in vitro* model for assembly uses an 1100-nucleotide RNA (PacA+) containing the operator hairpin to mimic the viral RNA (Beckett et al., 1988). In a reaction containing a 1:90 molar ratio of PacA+ RNA to coat protein, the wild-type coat protein efficiently assembles into capsids at protein concentrations above 2 μM. These *in vitro*-assembled particles have the same agarose gel mobility as those isolated from *E. coli. In vitro* capsid assembly experiments performed with the mutant proteins in pH 8.5 TMK buffer are shown in Figure 2. While the L77R mutant assembles at protein concentrations above 40 μM, none of the other FG loop mutants form capsids at concentrations of 100 μM (or 70 μM for V79E) (Table 2). Similar results were observed at pH 7.5, where viral assembly is more favorable (Matthews & Cole, 1972) (data not shown). This inability to assemble is not because the mutant proteins are denatured since they are all active for RNA binding (Table 3). Thus, the *in vitro* assembly data suggest that although most of the FG loop variants form capsids when overexpressed in *E. coli*, they are impaired in their ability to form productive dimer−dimer contacts.

To evaluate the effect of the FG loop mutations on a single dimer−dimer interaction, an assay was used that measures binding to an RNA containing two protein binding sites (Witherell et al., 1990). This assay allows quantitation of the free energy contributed to RNA binding by the interaction of two protein dimers (ΔG_{COOP}). Both assembly and ΔG_{COOP} are favored by conditions of high ionic strength, high temperature, and low pH, lending support to the view that the value of ΔG_{COOP} correlates to the cooperative interactions seen in the assembly reaction (Witherell et al., 1990). Thus, one would expect that the mutations in the FG loops that

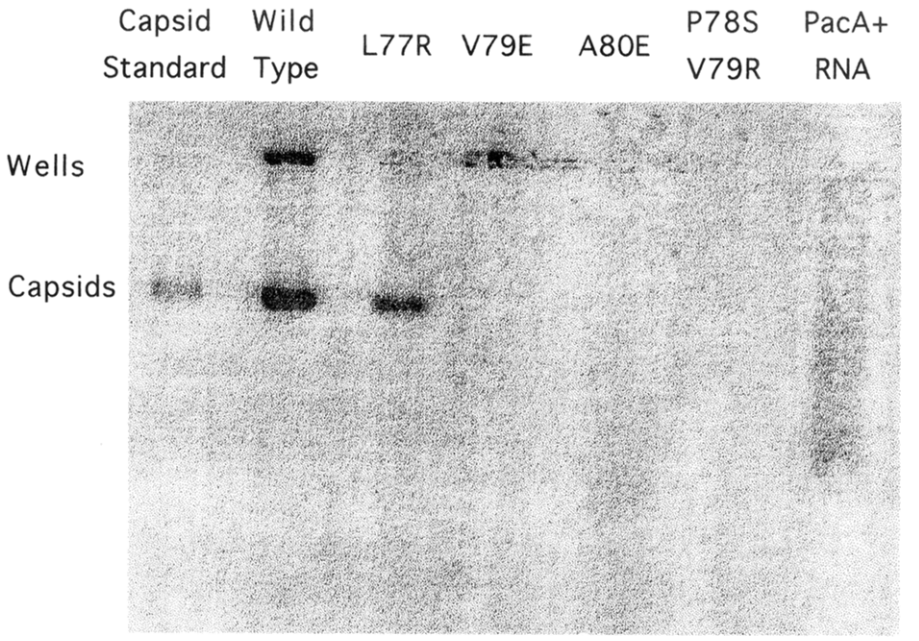

FIGURE 2: *In vitro* capsid assembly experiment for wild-type protein and the FG loop variants. Assembly was performed in TMK buffer at 4 °C with PacA+ RNA and protein concentrations of 2 μM for wild type, 70 μM for V79E, 40 μM for L77R, and 100 μM for A80E and P78S;V79R. The PacA+ RNA lane consists of RNA alone incubated in TMK buffer. The capsid standard consists of an empty wild-type capsid into which RNA1 has been soaked.

Table 3: Binding and Cooperativity Data for FG Loop Variants[a]

variant	K_d (nM)	K_c	ΔG_{COOP} (kcal/mol)
wild type	7	30	−1.9
V79E	8	10	−1.3
L77R	10	6	−1.0
V75E;V79E	40	3.4	−0.7
P78S;V79R	6	2	−0.4
A80E	15	1.2	−0.1
V75E;A81G	5	1	0

[a] K_d is the equilibrium dissociation constant, K_c is the cooperativity constant, and ΔG_{COOP} is the same cooperative free energy calculated using K_c.

weaken capsid assembly would also show reduced values of ΔG_{COOP}.

The effects of the FG loop mutations on the cooperative free energy of binding were measured by using two different RNA molecules (Witherell et al., 1990). RNA1, which contains only one coat protein binding site, was used to measure the K_a for each of the variant proteins (Figure 3). RNA2, which contains two coat protein binding sites separated by a single-stranded spacer region, binds the coat protein in a cooperative manner. Because the two hairpins in RNA2 have the same sequence, binding to RNA2 can be fit to a model of cooperative binding to two identical sites:

$$Y = (K_a[P] + K_a^2[P]^2K_c)/(1 + 2K_a[P] + K_a^2[P]^2K_c)$$

where Y is the fraction of RNA bound, K_a is the equilibrium binding association constant for RNA1, [P] is the concentration of coat protein, and K_c is the cooperativity constant (Ackers et al., 1983). This model assumes that the K_a measured for RNA1 equals the K_a for each protein binding site on RNA2 and that complexes of one or two dimers bound to the RNA are retained on nitrocellulose filters with the same efficiency. By using the K_a value determined from binding to RNA1, the value of K_c is determined by fitting the binding data for RNA2 to the cooperative binding

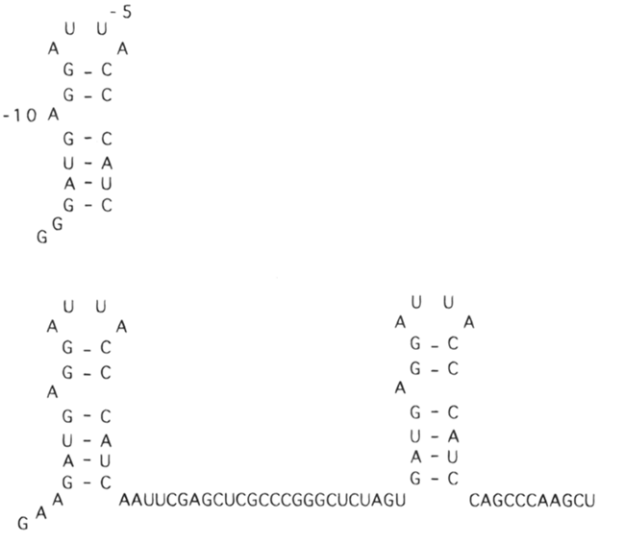

FIGURE 3: Sequences and secondary structures of RNA1 and RNA2. The RNA1 hairpin contains the sequence and structural elements of the consensus binding site (Witherell et al., 1991).

equation. In order to minimize error, both K_a and K_c are determined for the same set of protein dilutions. In a case where there is no cooperative free energy, binding to RNA1 and RNA2 will be identical.

In general, the RNA binding data clearly demonstrate that most of the FG loop mutations reduce the strength of interaction between coat protein dimers. As the amount of cooperative free energy decreases for the mutants, the RNA1 and RNA2 binding curves indeed become increasingly similar (Figure 4). The values for cooperative free energy vary from values close to that for the wild-type interaction to those that are barely measurable by this assay (Table 3). The replacement of hydrophobic amino acids with charged residues resulted in a decrease in ΔG_{COOP} in all cases.

Since many of these variant proteins show small RNA binding effects, but large cooperative free energy effects, it is clear that interactions between coat protein dimers are

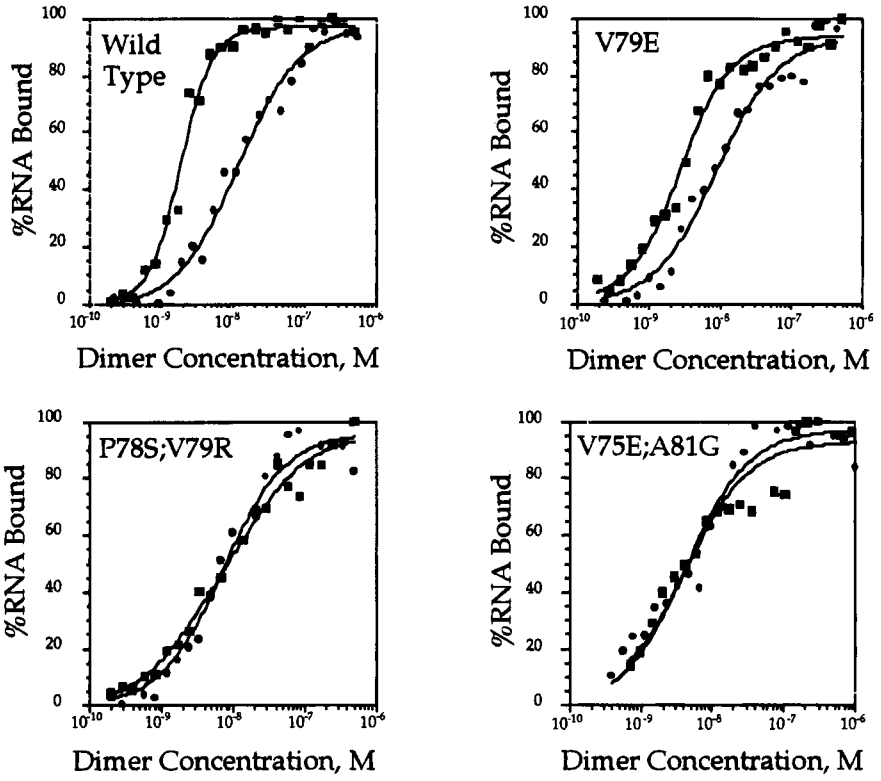

FIGURE 4: Filter binding experiments for ^{32}P-labeled RNA1 (●) and RNA2 (■) with the indicated proteins in TMK buffer (pH 8.5) at 4 °C. The binding curves for RNA1 fit a bimolecular equilibrium, while the binding curves for RNA2 fit a cooperative binding curve (see text).

Table 4: Specificity of RNA Binding

RNA	K_d (nM)	
	wild type	V75E;A81G
U_{-5} (RNA1)	2.6	6.2
C_{-5}	0.013	0.023
A_{-5}	26	66
G_{-5}	250	290
U_{-10}	>500	>500
G_{-10}	8	12

independent from dimer−RNA interactions. For example, the P78S;V79R variant has the same K_d as the wild-type protein, but has a K_c of 2, which is 15-fold less than the wild-type value of 30. In other cases, however, both RNA binding and cooperative free energy are affected. For example, the V75E;V79E variant also has a K_c of 3.4, yet shows a 6-fold increase in K_d.

A variant of the MS2 bacteriophage coat protein that does not form capsids under any circumstances, yet binds RNA normally would be desirable for many applications. On the basis of the data in Table 3, the V75E;A81G protein could be an excellent candidate for such a nonaggregating coat protein background. Capsid assembly experiments performed at coat protein concentrations up to 1 mM show no evidence of assembled particles. As shown in Table 4, both the wild-type coat protein and the V75E;A81G protein have the same specificity for the nucleotides at the −5 and −10 positions, both of which are important determinants for the specific interaction (Figure 3). The RNAs used in this experiment are variants of RNA1 with the specified changes. The affinity of RNA1 for the V75E;A81G protein was determined as a function of pH and ionic strength (data not shown). In both cases, the data were nearly identical to those for the wild-type protein, further confirming that both proteins bind RNA in the same manner.

DISCUSSION

In the current study, we used the crystal structure of the MS2 bacteriophage to select several coat protein mutations that were expected to disrupt cooperative binding to RNA. The FG loop region of the coat protein was targeted because it is not an essential part of the folded protein structure and mutations in this region have been previously shown to affect assembly. Using a two-plasmid system to study translational repression by coat protein, Peabody and Ely (1992) identified several mutations in the FG loops that repress translation better than the wild-type protein. This phenotype results from the facts that dimers, but not capsids, are able to repress translation and that these mutants have a higher intracellular concentration of free dimer than of wild type. A deletion of residues 67−79 from the FG loop was also deficient for capsid formation, yet was still active in repression (Peabody & Ely, 1992). These results are consistent with those for the related *fr* phage in which successive deletions from the FG loop region cause an increase in the number of dimers relative to capsids (Pushko et al., 1993).

When overexpressed in *E. coli*, five of six FG loop mutants formed capsids. However, *in vitro* experiments indicated that all of the mutant proteins were at least partially defective in the assembly of capsids. While the wild-type coat protein assembles into capsid particles *in vitro* at concentrations above 2 μM protein (Beckett et al., 1988), only the L77R mutant protein forms capsids at the much higher protein concentrations used for the FG loop mutants. There are several possible reasons why capsids form inside *E. coli* cells but not *in vitro*. While the *in vitro* experiment involves the

addition of coat protein in a single step, the expression of protein in the overexpression system results in a slow increase in concentration over time (Studier & Moffatt, 1986). The addition of coat protein in a single step may allow nonspecific RNA binding to occupy a large fraction of the coat protein added. The final concentration of coat protein *in vivo* is also substantially higher than that used *in vitro*. The coat protein concentration in the protein overexpression system is estimated to be a least 1 mM, which is 10-fold higher than the protein concentrations obtained *in vitro*. Finally, the *in vitro* buffer conditions that were optimized for RNA binding are different from the conditions inside the cell.

Because the *in vitro* results show that dimer−dimer interactions have been altered in the mutants, an assay was used that measures the ΔG_{COOP} for a single dimer−dimer interaction (Witherell et al., 1990). Since all of the FG loop mutants studied show a decrease in their ΔG_{COOP}, it is clear that binding to the double-hairpin RNA reflects the dimer−dimer interaction. The mutant proteins have a wide range of values for ΔG_{COOP}, yet most still form capsids when overexpressed in *E. coli*. This result suggests that the protein−protein interaction face is very robust and can accommodate single or double amino acid substitutions.

A detailed structural analysis of the cooperativity mutants is complicated by the lack of a solution structure for the dimer. In the case of the free dimer in solution, the FG loops are likely symmetric and flexible. The dimer−RNA1 complex, however, is asymmetric (Valegård et al., 1994). It is possible that the two FG loops adopt different structures in the complex. In the dimer−RNA2 complex, two FG loops remain free in solution, while two interact at the dimer interface. While the structure of the loops in solution is unknown, the loops that interact likely adopt one of the conformations observed in the intact bacteriophage. There are two different FG loop conformations observed in the crystal structure of the bacteriophage: the open A or C conformation and the tightly folded B conformation (Valegård et al., 1990). We expect that the FG loops at the dimer interface of the RNA2 complex are in a folded conformation resembling either the AC or BB′ conformation.

Because the BB′ interaction differs from that of the AC, it has been proposed that the conversion of the FG loop to the B conformation is an early step in capsid assembly (Valegård et al., 1994). This conversion of the FG loops to the B conformation may facilitate binding of additional dimers to form an assembled capsid. It is thus possible that the mutations studied have perturbed the ability of the FG loops to form the B loop conformation seen in the structure of the virus. Formation of the B loop conformation has been proposed to occur through a structural perturbation due either to P78 *cis/trans* isomerization and/or to RNA binding (Valegård et al., 1994). The P78S;V79R mutant allows us to address this first hypothesis. While Pro78 is in the *trans* conformation in the A and C loop conformations, it is in the *cis* conformation in the more compact B subunits. The presence of a *cis*-proline in the B loop structure may account for its tightly folded structure. Because the P78S mutant still has measurable cooperativity and assembles into capsids in *E. coli*, it is likely that Pro78 isomerization is not required for capsid formation. It is, of course, possible that S78 is in a conformation similar to that of *cis*-P78 in the capsid form of this mutant protein. If S78 is not in the *cis* conformation, new protein−protein interactions must be able to compensate for the loss of the *cis*-proline in the structure of this mutant protein.

The second hypothesis for formation of the B loop structure is that it is caused by binding to RNA (Valegård et al., 1994). Without a knowledge of the structure of the free dimer with and without RNA, it is not possible to directly address this hypothesis. Because several of the FG loop variants have decreased cooperative free energies and wild-type RNA affinity, it is clear that protein−protein interactions can be affected independently of protein−RNA interactions. There are, however, two examples where FG loop mutations have an effect on RNA binding. In the first case, the V79E;V75E mutant shows a 6-fold decrease in RNA binding. These same mutations taken singly in either the V79E or the V75E;A81G protein have no effect on RNA binding. The significantly increased K_d in the V75E;V79E variant thus is not likely a result of disruption of a specific protein−RNA contact, but rather an electrostatic effect. The second example is the A80E variant, which has a 2-fold decrease in RNA affinity. This small decrease may be explained by the fact that A80 is near the 5′ end of the RNA in the crystal structure (Valegård et al., 1994). In general, however, a decrease in cooperative free energy does not necessarily result in a corresponding perturbation of the RNA−protein interaction.

The experiments presented here clearly demonstate that there is no direct thermodynamic coupling between protein−protein interactions and RNA binding. RNA binding is not required for capsid assembly because capsids can be formed in the absence of RNA (Beckett et al., 1988; Mastico et al., 1993). FG loop variants with significant decreases in ΔG_{COOP} can bind RNA with wild-type affinity. The observed improvement in assembly upon RNA binding thus is likely due to a general electrostatic effect. Once the protein binds to RNA, it is likely to migrate to the growing capsid without dissociating from the RNA. This conclusion is supported by the assembly of phage particles in the presence of poly(vinyl sulfate), as well as by the observation of improved assembly without RNA at high ionic strength (Hohn, 1969; Matthews & Cole, 1972). The critical role of the operator RNA sequence may be during an infection cycle when the protein concentration is initially low and gradually increases with time (Nathans et al., 1969). The high affinity of the coat protein for the operator hairpin will ensure selective encapsidation of the viral RNA over nonviral sequences. Specific encapsidation of the viral genome does not require direct coupling between protein−RNA and protein−protein interactions.

In this study, a combination of RNA binding assays and *in vitro* assembly was used to evaluate the influence of FG loop mutations on protein−protein and protein−RNA interactions. The mutants examined in the current study provide a starting point for a more detailed study of the dimer−dimer interaction face in the MS2 bacteriophage coat protein. The approach of isolating an individual protein−protein interaction may be useful for the study of assembly in other bacteriophage or viral coat proteins, such as alfalfa mosaic virus, satellite tobacco mosaic virus, or hepatitis B virus (Larson et al., 1993; Pollack & Ganem, 1993; Houser-Scott et al., 1994). In addition, engineering such proteins so that they no longer assemble will facilitate the determination of their consensus RNA binding sites.

The FPLC-purified V75E;A81G protein is an appropriate nonaggregating background for studying the coat protein−RNA interaction. The RNA binding affinity and specificity of the V75E;A81G variant were identical to those of the wild type under all conditions tested. The FG loop mutant background will be useful for making mutations in the RNA binding face of the coat protein. Previous work in our laboratory has shown that although many mutants in the RNA binding face form capsids, it is not possible to obtain active dimers in solution (J. M. Gott, unpublished observations). The harsh acetic acid denaturation protocol used to disassemble the capsid also partially denatures the monomer. Although the wild-type protein renatures to an active form, mutant proteins usually do not. A mutant protein that does not form capsids and can be purified by a more gentle means makes it possible to obtain active mutants in the RNA binding face. Several RNA binding face variants that were inactive for RNA binding when purified from capsids were active when purified in the nonaggregating background (K. A. LeCuyer and O. C. Uhlenbeck, unpublished results). In addition, to study proteins with weakened RNA binding affinities, it would be advantageous to perform RNA binding studies at protein concentrations above 1 μM, where the wild-type protein assembles into capsids. Since the V75E;A81G protein does not form capsids at 1 mM protein concentration, it is suitable for detecting weak RNA affinities and could be used to study protein−RNA interactions by NMR.

ACKNOWLEDGMENT

We thank Steve Schultz and Jodi Ryter for help in protein purification and Hans Johannson, Maria Stubenrauch, and Tracy Stage for critical reading of the manuscript.

REFERENCES

Ackers, G. K., Shea, M. A., & Smith, F. R. (1983) *J. Mol. Biol. 170,* 223−242.

Beckett, D., & Uhlenbeck, O. C. (1988) *J. Mol. Biol. 204,* 927−938.

Beckett, D., Wu, H. N., & Uhlenbeck, O. C. (1988) *J. Mol. Biol. 204,* 939−947.

Bernardi, A., & Spahr, P.-F. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69,* 3033−3037.

Dill, K. A. (1990) *Biochemistry 29,* 7133−7155.

Giebler, H. (1992) Master's Thesis, University of Colorado, Boulder, CO.

Golmohammadi, R., Valegård, K., Fridborg, K., & Liljas, L. (1993) *J. Mol. Biol. 234,* 620−639.

Gott, J. M., Willis, M. C., Koch, T. H., & Uhlenbeck, O. C. (1991) *Biochemistry 30,* 6290−6295.

Hohn, T. (1969) *J. Mol. Biol. 43,* 191−200.

Houser-Scott, F., Baer, M. L., Liem, K. F., Jr., Cai, J.-M., & Gehrke, L. (1994) *J. Virol. 68,* 2194−2205.

Hung, P. P., Ling, C. M., & Overby, L. R. (1969) *Science 166,* 1638−1640.

Larson, S. B., Koszelak, S., Day, J., Greenwold, A., Dodds, J. A., & McPherson, A. (1993) *J. Mol. Biol. 231,* 375−391.

Ling, C. M., Hung, P. P., & Overby, L. R. (1969) *Biochemistry 8,* 4464−4469.

Mastico, R. A., Talbot, S. J., & Stockley, P. G. (1993) *J. Gen. Virol. 74,* 541−548.

Matthews, K. S., & Cole, R. D. (1972) *J. Mol. Biol. 65,* 1−15.

Nathans, D., Oeschger, M. P., Polmar, S. K., & Eggen, K. (1969) *J. Mol. Biol. 39,* 279−292.

Peabody, D. S., & Ely, K. R. (1992) *Nucleic Acids Res. 20,* 1649−1655.

Pickett, G. G., & Peabody, P. S. (1993) *Nucleic Acids Res. 21,* 4621−4626.

Pollack, J. R., & Ganem, D. (1993) *J. Virol. 67,* 3254−3263.

Pushko, P., Kozlovskaya, T., Sominskaya, I., Brede, A., Stankevica, E., Ose, V., Pumpens, P., & Grens, E. (1993) *Protein Eng. 6,* 883−891.

Shafranski, P., Zagorski, V., Khrobochek, Y.-a., & Zagorska, L. (1975) *Mol. Biol. (Moscow) 9,* 78−85.

Studier, F. W. (1991) *J. Mol. Biol. 219,* 37−44.

Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189,* 113−130.

Valegård, K., Liljas, L., Fridborg, K., & Unge, T. (1990) *Nature 345,* 36−41.

Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., & Liljas, L. (1994) *Nature 371,* 623−626.

Witherell, G. W., & Uhlenbeck, O. C. (1989) *Biochemistry 28,* 71−76.

Witherell, G. W., Wu, H.-N., & Uhlenbeck, O. C. (1990) *Biochemistry 29,* 11051−11057.

Witherell, G. W., Gott, J. M., & Uhlenbeck, O. C. (1991) *Prog. Nucleic Acid Res. Mol. Biol. 40,* 185−220.

Wong, I., & Lohman, T. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 5428−5432.

BI950608Y