*Seminars in VIROLOGY* **8,** 176–185 (1997)
*Article No.* VI970120

# RNA Recognition by the MS2 Phage Coat Protein

Hans E. Johansson,*,[1] Lars Liljas,[†] and Olke C. Uhlenbeck*

**Department of Chemistry and Biochemistry, University of Colorado, Campus Box 215, Boulder, Colorado 80309-0215; and [†]Department of Molecular Biology, Uppsala University Biomedical Center, Box 590, S-75124 Uppsala, Sweden*

**The interaction of MS2 coat protein and its translational operator hairpin is a very well-characterized RNA–protein complex. The recent high-resolution cocrystal structure successfully explains many biochemical experiments measuring the affinity of the protein–RNA interaction for mutant proteins and chemically modified RNAs. However, an analysis of a tight binding variant of the RNA suggests that the conformation of the free RNA is also an important determinant of the affinity.** © 1997 Academic Press

***Key Words:*** **bacteriophage MS2 RNA; RNA–protein interaction; translational repression.**

## INTRODUCTION

About 10 min after infection of *Escherichia coli* by RNA bacteriophage MS2, the product of the phage coat protein gene reaches a high enough concentration to bind to an RNA hairpin at a unique site on the phage genome. This specific RNA–protein interaction has two important consequences. First, it represses translation of the phage-encoded subunit of the phage replicase, which is only expressed early in infection. Repression occurs by simple competitive inhibition, where the coat protein binds directly to the replicase initiation region and thereby blocks ribosome entry. Second, the protein–RNA interaction identifies the phage RNA for later assembly into virus particles after the coat protein reaches much higher concentrations later in infection. Since cellular RNAs are not packaged into virus unless the binding hairpin is introduced (1, 2), this can be considered a classic ''packaging'' site that acts by nucleating phage assembly.

Since the initial RNase protection experiment that defined this RNA–protein complex (3), this system has been studied as an example of a sequence-specific RNA–protein interaction. The relatively small size and high abundance of the protein, the comparative simplicity of the RNA target, and the ease of assaying the interaction *in vivo* or *in vitro* have led to numerous genetic (4–8), biochemical (9–18), and biophysical studies (15, 18–22) aimed at understanding how recognition is achieved. In addition, the high affinity and specificity of the interaction inside cells has made it useful for studying eukaryotic translation (23), transcription (24–26), and RNA transport (27). The MS2 system has also been used in the ''three-hybrid'' system to find new RNA–protein complexes (28). The interaction of phage coat proteins with RNA has been reviewed (9, 14). In this review, we will focus on the MS2 coat protein system and discuss how the available structural information can relate to biochemical assays measuring the affinity of the interaction.

The 129-amino-acid MS2 coat protein is present as a dimer at protein concentrations below 1 μM. The minimal RNA fragment containing all the determinants needed for binding is a hairpin which includes the Shine–Dalgarno sequence and initiation codon of the replicase gene (Fig. 1A). However, RNA mutagenesis (29) and *in vitro* selection (30) experiments show that neither of these signals is directly involved in coat protein recognition. Instead, tight binding requires the presence of four specific single-stranded residues held in place by the characteristic secondary structure (Fig. 1B). RNA binding by MS2 coat protein is very specific since little competition is observed when high concentrations of other RNAs are included in the presence of the RNA hairpin. In addition, many point mutations in the RNA-binding site reduce binding affinity 100-fold or more (29, 31–33).

[1]To whom correspondence should be addressed at Department of Medical Genetics, Uppsala University Biomedical Center, Box 589, S-75123 Uppsala, Sweden. current address: E-mail: Hans.Johansson@medgen.uu.se. Fax: +46-18-526849.

1044-5773/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.

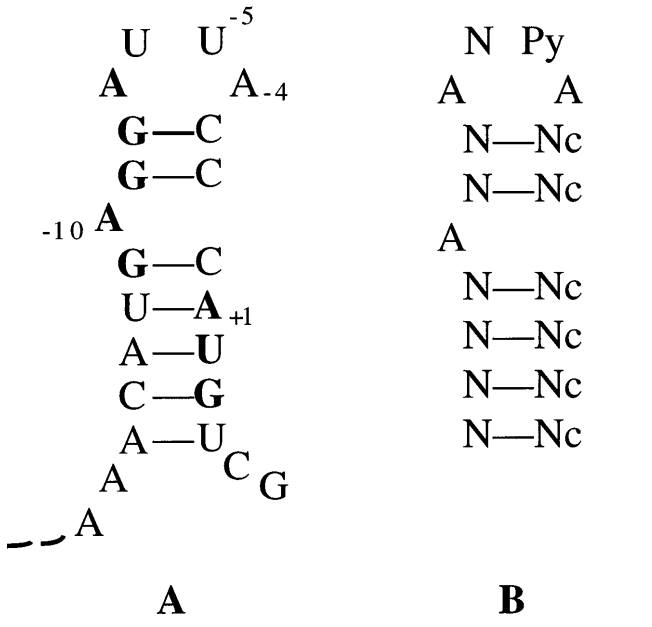



**FIG. 1.** (A) The replicase initiation region of MS2 where +1 is the first residue of the replicase gene. The Shine–Dalgarno and AUG are highlighted. (B) A minimal "consensus" RNA-binding site for MS2 coat protein where N is any nucleotide and $N_C$ is its complement.

At concentrations higher than 1 μM, MS2 coat proteins aggregate to form structures that closely resemble empty viral shells. Such coat protein capsids are also produced when the gene is overexpressed in *E. coli* (12, 34, 35). Coat protein capsids are extremely stable and can only be disaggregated under denaturing conditions. The formation of capsids *in vitro* is stimulated by the presence of RNA with a preference for RNAs containing the hairpin-binding site (1, 2).

## STRUCTURES

The structure of the complex between the MS2 coat protein dimer and the RNA hairpin was studied by soaking the hairpin into crystals of particles formed by recombinant coat protein expressed in *E. coli* (36, 37). These particles are very similar to native virus particles and contain 90 coat protein dimers. Due to the principles of quasi-equivalent packing of subunits in an icosahedral shell, 60 of the dimers are (slightly) asymmetric while the remaining 30 are symmetric. One RNA hairpin binds to each dimer. While the symmetric dimer binds RNA in two orientations making the interpretation more difficult, the asymmetric dimers bind RNA in a unique orientation allowing unambiguous modeling of the complex. A 10-stranded β-sheet in the dimer faces the interior of the particle, and the RNA hairpin binds to this sheet contacting both subunits of the dimer (Fig. 2). The RNA is bound to the sheet with the helical axis of the short stem roughly parallel to the strands of the sheet. The 5′ side of the stem and part of the loop is facing the protein. The lower part of the stem is disordered, prohibiting modeling of two base pairs. The total surface involved in the interaction is 1700 $Å^2$.

A total number of 21-amino-acid residues in the dimer are involved in direct interactions with the hairpin (Fig. 3). Thirteen of these contacts are with the bases A −4, U −5, and A −10, known to be important for the RNA specificity (see below). The backbone contacts are mostly to phosphates; the single exception is a hydrogen bond between the 2′ OH of one ribose and a glutamic acid side chain. The phosphate interactions are mainly through lysine and arginine side chains, which form charged hydrogen bonds and ion pairs with the phosphate oxygens. There is thus a significant electrostatic contribution to the stability of the complex. The inner surface of the dimer has a net positive charge, with totally 10 positive charges and 6 negative charges well exposed. There are very few water-mediated contacts in the complex.

Two almost identical pockets, one on each protein subunit (denoted A and B), seem to partly determine the specificity by binding the bulged A −10 and A −4 in the loop. The bases are sandwiched between the aliphatic part of the Lys 61 side chain and to some extent the methyl group of Thr 59 on one side and the side chain of Val 29 on the other (Fig. 4). The pocket is limited by the Thr 45 and Ser 47 side chains which form direct hydrogen bonds to the bases. In spite of the similarity of the pockets, the bases are bound differently. A −4 is inserted such that N1 is bound to Ser 47 and N6 and N7 are bound to Thr 45. N6 is in addition forming a hydrogen bond with the carbonyl oxygen of Thr 59. A −10 is bound to Ser 47 by N3, while N1 and N6 form hydrogen bonds with Thr 45. This is due to the asymmetry of the RNA and the restricted possibilities of rearrangements without losing other important contacts. No interactions similar to these adenine-specific interactions have as yet been found in other protein–RNA complexes. The pocket is highly conserved in related phages; residues corresponding to Thr 45, Ser 47, and Lys 61 are found in all known small RNA phage sequences (9, 38).

The U −5 base in the loop is stacked below the Tyr 85 side chain (Fig. 5). This type of interaction is quite common in protein nucleic–acid complexes (39). In this

Copyright © 1997 by Academic Press

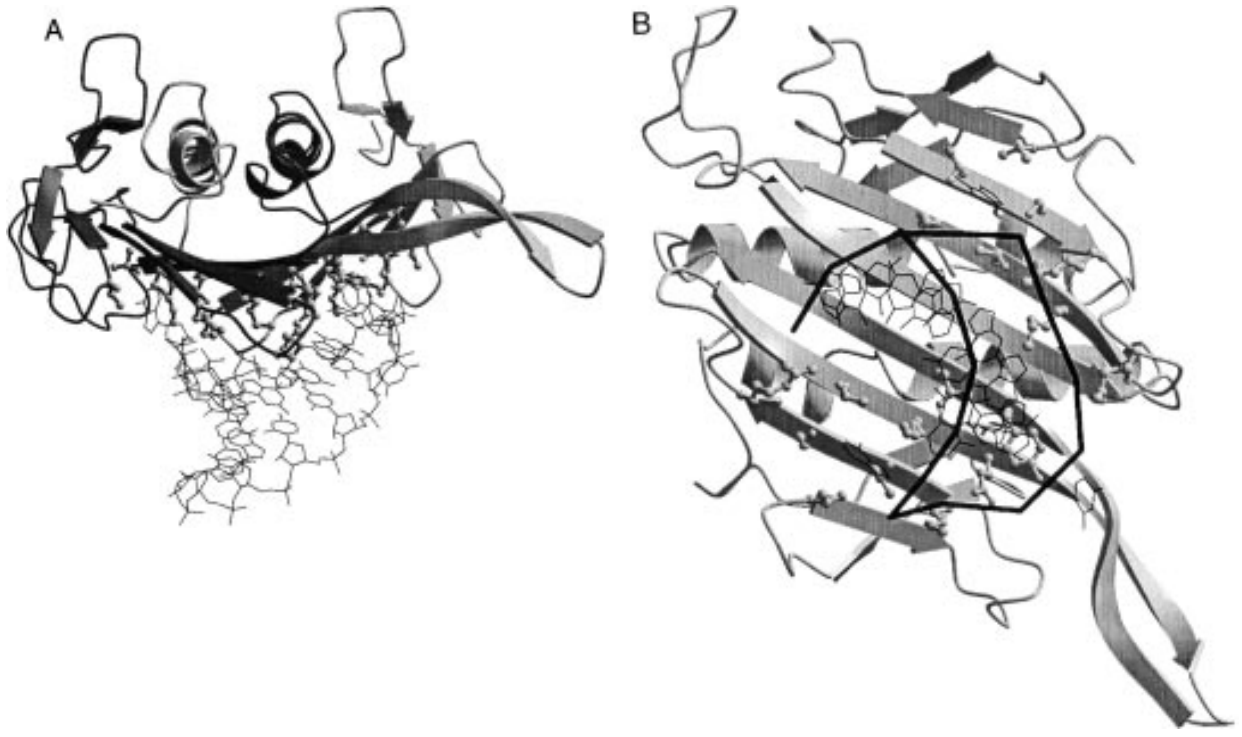



**FIG. 2.** Overview of RNA binding to an MS2 dimer. (A) Tangential view showing the RNA (thin lines) interacting with the central part of the 10-stranded sheet of the dimer. Side chains interacting with the RNA are drawn as ball-and-stick objects. (B) Radial view from the inside of the particle. The RNA is drawn schematically with thick lines connecting the phosphates and the bases drawn as thin lines. This figure and Figures 4–6 were drawn using the programs Raster3D (52, 53) and Molscript (54).

particular case, the hydroxyl group of the tyrosine forms a hydrogen bond to the phosphate of nucleotide −5, and a neighboring asparagine side chain makes a hydrogen bond with O2 of the base, an interaction which might be determining the specificity for pyrimidine at this position. The side chain of Thr 91 is also close to the base. The ribose 2′-hydroxyl of this nucleotide is forming a hydrogen bond with the side chain of Glu 63. The base of A −7 in the loop and the 5′ side of the stem (bases −8, −9, −11, −12, and −13) are stacked below the U −5 base.

A comparison of the protein dimer in capsids with and without the RNA hairpin generally shows very small changes in protein conformation, the exception being modest rearrangement of some side chains. This is what might be expected for a binding site on the floor of a large, rigid β-sheet. The RNA, in contrast, seems to undergo extensive conformational changes upon binding to the protein. In the NMR solution structure of a variant of the hairpin (40), the bulged A −10 is intercalated in the stem, and there is no evidence of stacking of bases −9 and −11 (Fig. 6). The conformation of the loop is uncertain due to flexibility, but in contrast to protein-bound RNA there is no evidence for stacking of U −5 and A −7.

## CORRELATING STRUCTURE WITH FUNCTION

The wealth of structural information and the comparative simplicity of this system have made it ideal for detailed structure–function studies. In such experiments the free energy of the protein–RNA interaction is measured with various protein mutations and/or chemically modified RNAs with three general objectives. The first is to confirm that the RNA–protein interaction in the cocrystal is the same as the one observed in solution. This issue is of greater concern than usual since the cocrystal structure is part of a virus-like capsid while the complex found in solution is a protein dimer bound to a hairpin. Furthermore, since there is substantial stimulation of capsid assembly by RNA, it is possible that some of the interactions are unique to the capsid. Nevertheless, all of the

Copyright © 1997 by Academic Press

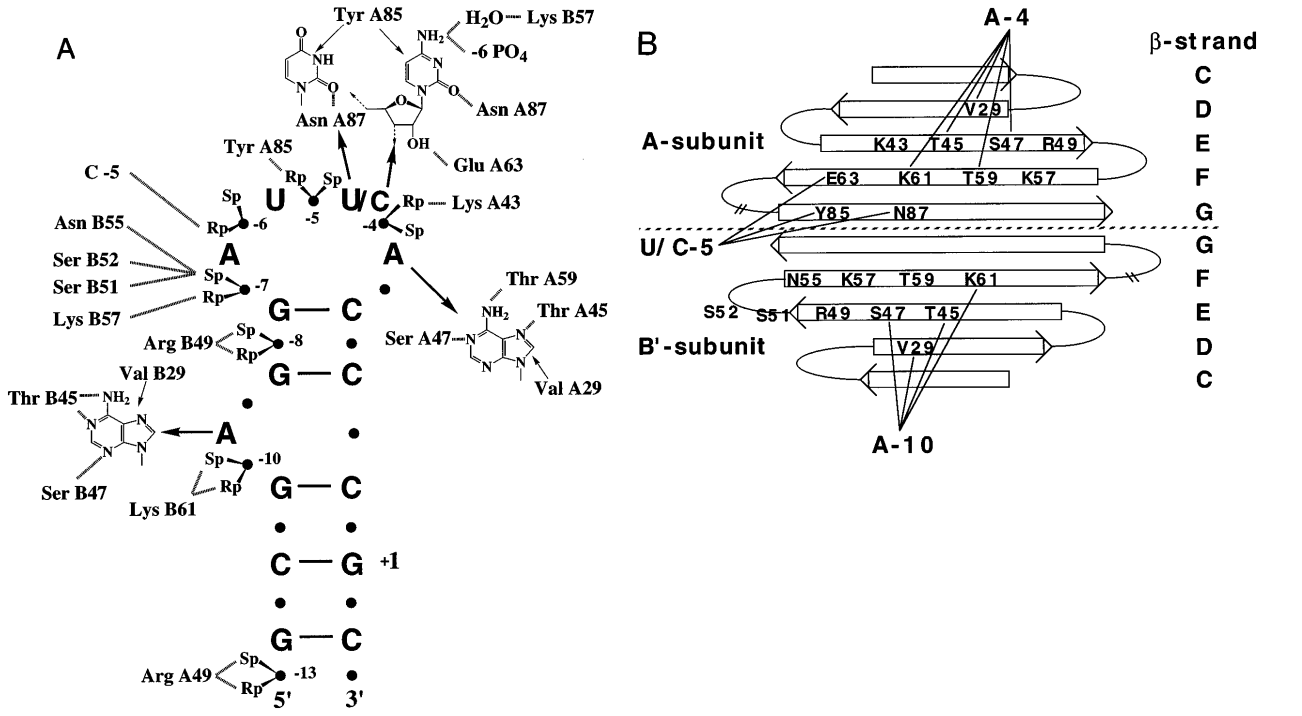



**FIG. 3.** Summary of the protein–RNA complex. (A) Side chains in A or B subunits which interact with hairpin residues. (B) RNA-binding sites on the β-sheets.

biochemical and mutagenic studies performed thus far are consistent with the two interactions being identical. However, this analysis is still incomplete and there are only relatively few experiments which directly compare the sequence specificity in both complexes.

A second important goal is understanding how the total free energy of binding is achieved. Thus, within the relatively large area of interaction between the two macromolecules, which intermolecular hydrogen bonds contribute most significantly to the binding energy and which are energetically neutral due to compensating interactions with the solvent in the unbound form? Are there ''hot spots'' in the binding free energy such as found in the human growth hormone–receptor complex (41)? The thermodynamic contribution of the ion pairs between phosphates and charged amino acids is also expected to be important. To estimate their contribution, an analysis must be performed of the counterions bound in the free state that must be released upon binding. Finally, evaluating the energetics of the large RNA conformational change observed upon protein binding is expected to be especially challenging since several residues are taken out of a favorable stacked environment in the free form and placed in another environment in the complex. Only the net free energy of this process will contribute to the overall $\Delta G$ of binding.

A final goal is to understand how the specificity of the RNA–protein interaction is achieved. Thus, why does the protein bind the target RNA tightly, but bind similar sequences with much lower affinity? While the cocrystal structure can help to rationalize the reduced affinity of other sequences, it is important to analyze other RNAs in detail in order to understand why they bind poorly.

A typical structure–function experiment involves making several point mutations in an amino acid that interacts with the RNA and measuring their affinities to the normal RNA as well as to RNAs containing changes in the interacting atoms or functional groups. In general, the easiest protein mutations to interpret are ones that reduce (or maintain) the size of the amino acid side chain so as to minimize the unpredictable rearrangements that would occur to accommodate steric bulk. A wide variety of modified RNAs is available to make the corresponding RNA ''mutagenesis'' comparatively easy. Although a decrease in affinity will often reflect the disruption of an interaction in the structure of the complex, it is important to note that the binding energy can also be affected by changes that

Copyright © 1997 by Academic Press

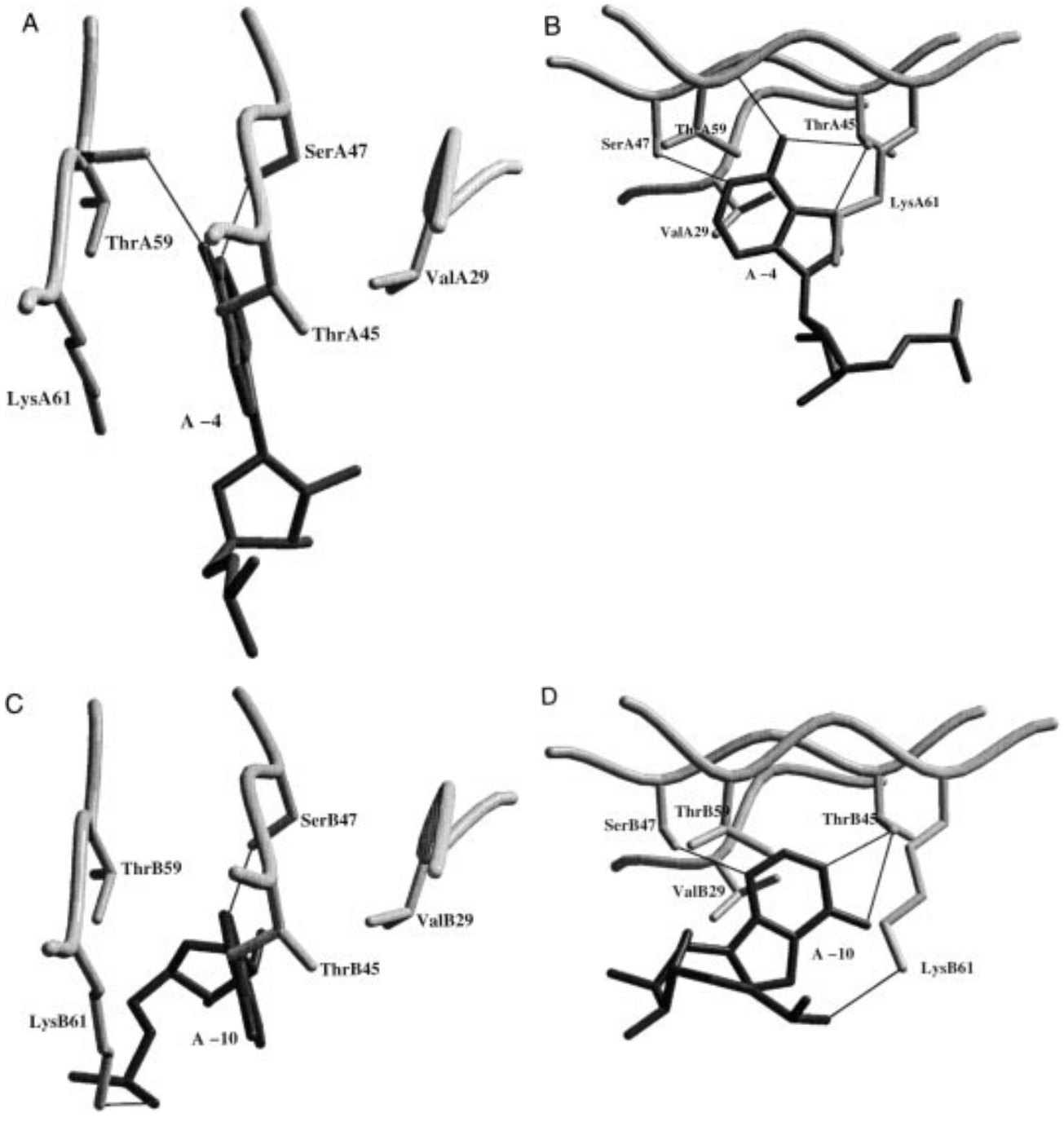



**FIG. 4.** The pockets for binding of A −4 and A −10. The protein is in light gray. The main chain of the protein is drawn as a smooth line with selected side chains attached. The RNA is in dark gray. Hydrogen bonds are shown as thin rods. Two roughly perpendicular views are shown. (A, B) Pocket for A −4 in the A subunit. (C, D) Pocket for A −10 in the B subunit.

Copyright © 1997 by Academic Press

alter the structure of the free protein or free RNA. In this regard, it is useful to consider the following thermodynamic cycle:

$$\begin{array}{ccc} P + R & \xleftrightarrow{\Delta G_{native}} & P \cdot R \\ \Delta G_{free} \updownarrow & & \updownarrow \Delta G_{complex} \\ P^* + R & \xleftrightarrow{\Delta G_{mutant}} & P^* \cdot R \end{array}$$

Thus, making a mutation in the protein (P → P*) can alter the energy of binding ($\Delta G_{native} - \Delta G_{mutant}$) because of a thermodynamic change in the complex ($\Delta G_{complex}$) and/or the free protein ($\Delta G_{free}$). For example, removing a charged amino acid could reduce $\Delta G_{complex}$ by deleting an ion pair with the RNA, but also affect $\Delta G_{free}$ by disrupting an ion pair that forms with another amino acid on the surface of the free protein.

In order to study the RNA-binding properties of coat protein mutants over a wide range of protein concentrations, it was necessary to prevent the ability of coat protein to aggregate into capsids. This was done by preparing mutations in the FG loop (12) or deleting it entirely (42). Since the FG loop participates in several interdimer contacts but is far away from the RNA-binding site, capsid formation was blocked without affecting the affinity or specificity for RNA. A more complicated problem is that the symmetric protein dimer can bind the asymmetric RNA in two orientations and, as shown in Fig. 3, several of the amino acids in each subunit interact with different parts of the RNA. Mutations of these amino acids will thus reflect an average of the effects in both contacts. In order to make mutations in one subunit selectively, a gene that fuses the two protein dimers has been prepared and the resulting protein has been shown to be active (5). Mutating this fusion protein in one of the amino acids that only contacts RNA in one of the subunits should give a protein that binds RNA in only one orientation. Such a protein can be used as background for analyzing mutations in the amino acids that bind in both subunits.

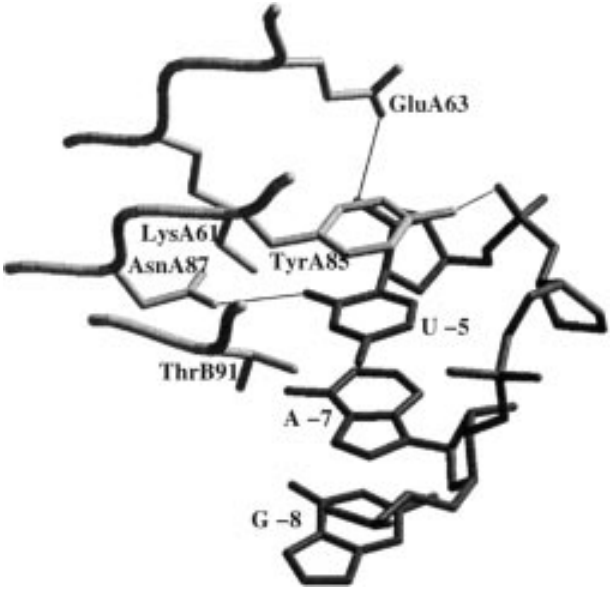



**FIG. 5.** Binding of U −5. The contact is mostly with the A subunit (hydrogen bonds to Glu 63, Tyr 85, and Asn 87). Three of the bases stacked to Tyr 85 are shown.

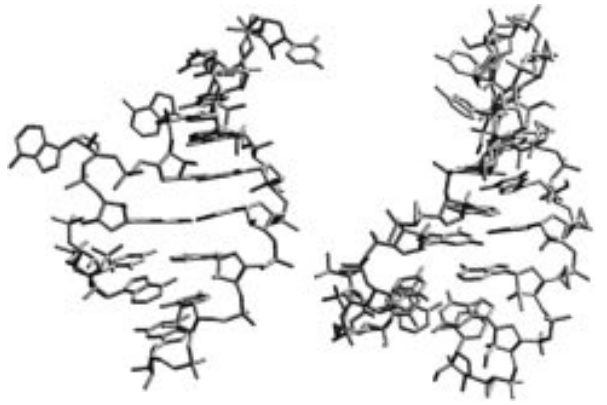


**FIG. 6.** The RNA conformation as (A) in complex with the protein and (B) in solution (NMR structure). The orientation is chosen to show the exposed adenines −10 (leftmost base) and −4 in the protein complex.

Peabody and co-workers have constructed an *in vivo* translational repression system that involves connecting the hairpin initiator to lacZ and preparing mutations in an overexpressed coat protein gene. Mutagenesis experiments in both the wild-type (4) and the dimer fusion (5) genes have confirmed that all the amino acids that participate in protein–RNA contacts observed in the cocrystal structure are sensitive to mutation. Since most of the mutations increase the size of the amino acid side chains that protrude from the β-sheets, steric clashes might explain the reduced affinity. Several interesting mutations which increase protein affinity have also been identified, some of which appear to introduce or alter ionic interactions (17). More recent experiments from this lab have focused on selections that alter the specificity of MS2 coat protein so that it binds the apparently quite

Copyright © 1997 by Academic Press

different Qβ operator. Remarkably, quite effective specificity swaps can be achieved with only a few amino acid changes that appear to introduce new contacts neither used by MS2 nor by Qβ (7, 42a).

The affinity of MS2 coat protein with RNA is generally measured by a nitrocellulose filter binding assay. Trace concentrations of labeled RNA are mixed with varying concentrations of coat protein, filtered, and counted. The apparent binding curve is fit to a binding constant and a filter retention efficiency. The fraction of active protein is determined in a separate titration experiment where RNA concentration is varied at a constant protein concentration above $K_d$. While the filter binding method involves separation of complex from free RNA, its validity has been confirmed by direct comparison with an equilibrium fluorescence titration assay (43). By using a microtiter well format and a phosphorimager, rapid acquisition of binding and rate constants of many different modified RNAs in a single experiment is feasible. In this system, gel mobility shift experiments only work for very tight binding RNAs ($K_d < 0.1$ nM) and thus are less convenient. Most structure–function experiments performed thus far have either focused on modified RNAs (13, 15, 33, 44) or those amino acids that interact with RNA in only one subunit. An example of such a structure–function experiment is presented in the next section.

## THE SPECIFICITY AT −5

Over the years the recognition of the −5 residue by coat protein has received a lot of attention. The wild type operator hairpin contains a U at this site (Figs. 1 and 5). Substitutions with A or G give hairpins with 15- to 100-fold weaker affinity for the coat protein (9, 45). The situation is reversed for hairpins with a C (C-RNAs) such that the affinity *in vitro,* measured by off-rate experiments, is 50- to 100-fold tighter (31, 46). Such operators would be predicted to confer a phenotype of superrepressors *in vivo* and might therefore be selected against by the virus. Early attempts to explain the dramatically higher affinity for C-RNAs proposed the formation of a Michael adduct between the C and a nearby cysteine in the coat protein (45), in analogy to those found between *E. coli* Ala- and Ile-tRNAs and their respective synthetases (47, 48). Biochemical experiments to directly test this model using $D_2O$ (46), halogenated base substitutions, and tritium exchange (44) all proved negative. In addition, the suggested pairing partner in the coat protein, the side chain of either cysteine 46 or 101, can be mutated without effect on RNA binding (49). The subsequent cocrystal strucdid not show cysteine or any other nucleophile near the C6 position of the pyrimidine (36). Clearly, the explanation had to be sought elsewhere.

A recent critical clue for explaining the tighter bonding of C-RNAs comes from a comparison of the cocrystal structures of C-RNA and U-RNA (37). The two RNAs are arranged the same way on the inner surface of the protein and the −5 bases are found in nearly identical positions. The side chains of Glu 63, Tyr 85, and Asn 87 interact with the −5 nucleotide the same way in both complexes (Figs. 7A and 7B), effectively locking it into place. Thr 91 (not shown) packs hydrophobically against both U and C −5. Remarkably, the only difference is found between the discriminatory groups and the neighboring −6 phosphate. O4 is located 3.7 Å away from O1P in the wt RNA. Thus, U −5 is most likely in its *keto*-form and does not participate in a direct hydrogen bond, which would require the base to be in an *enol*-tautomeric form. In the C-RNA the phosphate has moved ~0.5 Å closer and the 4-amine of C −5 directly contacts the backbone. Water molecules in the vicinity of the pyrimidines (not shown) could make additional contacts that may be involved in the specificity. Biochemical experiments with 4-thio U-RNA and 2-pyrimidone RNA showed that the 4 *keto*-oxygen of U contributed very little to the affinity (14, 18). Further experiments with RNAs containing modifications at −5 showed only slight differences in affinity, most likely contributable to inherent stacking differences (14, 33). In conclusion, both structural and biochemical data suggest that added binding energy for C-RNA is provided by a charged hydrogen bond between the 4-amine of C and phosphate −6.

The importance of the −5 interaction to the overall binding energy has been confirmed by studies with mutated coat proteins (15 and H. E. Johansson *et al.*, unpublished results). Replacement of the tyrosine ring of the conserved residue 85 by alanine reduced binding affinity >200-fold. Changing Tyr 85 to Phe reduced binding affinity less, but suggested that the hydrogen bond to the −5 phosphate was important, a fact confirmed by modification of the neighboring −5 phosphate to phosphorothioate (15). Somewhat surprisingly, changes of Asn 87, which directly contacts the −5 base, to Ala or Ser only marginally perturb the interaction (15 and H. E. Johansson *et al.*, unpublished results). Finally, changing either partner in the contact between −5 2′-OH and Glu A63 indicates that it also contributes substantially to protein affinity (13 and

Copyright © 1997 by Academic Press

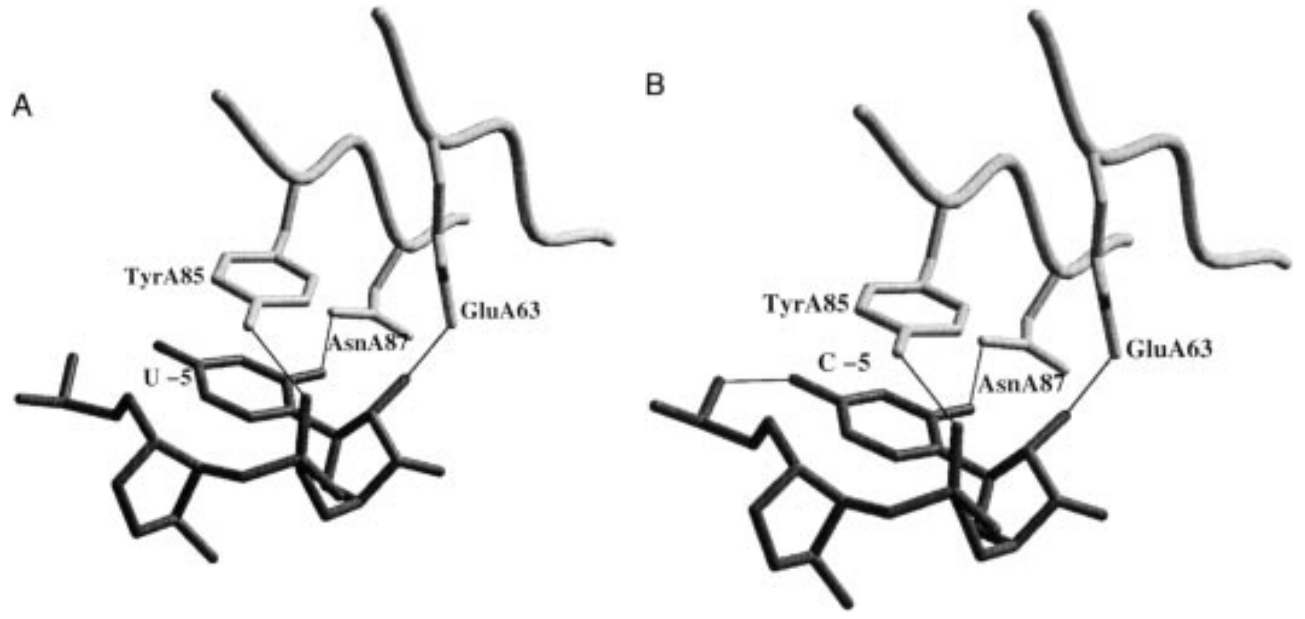



**FIG. 7.** Details of the interactions at position −5. (A) In the wild-type complex. (B) In the complex of the C variant showing an extra hydrogen bond between N2 of the base and the phosphate group of −6. Other interactions are the same.

H. E. Johansson, K. A. LeCuyer, and O. C. Uhlenbeck, unpublished results).

How do these protein modifications affect the specificity of the interaction? Steric exclusion seems most important in the discrimination against purines at −5. The size of the binding pocket may be modulated by the side chain of residue 87 and Ala or Ser mutant proteins accommodate purines more easily (16 and H. E. Johansson *et al.*, unpublished). From inspection of the structures with a −5 U or C, it is clear that other protein side chains close to the pocket would have to adopt different conformations in order for purines to fit. The related phage GA, which carries an A −5 in the loop, alters four nearby amino acids (Lys 43 Ser, Asn 55 Lys, Asn 87 Ser, and Ser 51 Ala). Apart from Asn 87 Ser, none of the other amino acids was found to alter the specificity for U over A at −5 alone (16). To extend the current knowledge on purine discrimination, cocrystal structures with A- and G-RNAs would clearly be beneficial. The pyrimidine specificity for coat proteins with mutations of Asn 87 and Tyr 85 has also been assessed (H. E. Johansson *et al.*, unpublished results). Removal of the hydrogen bond with O2 of the −5 base by mutation of Asn 87 to Ala did not alter binding to either U- or C-RNA. Mutations of Tyr 85 to Phe, His, or Ser only slightly reduced the preference of C over U even though the absolute affinity varied substantially. This small effect might be attributable to the differences in stacking properties for both the amino acid side chains and the −5 base.

Both structural analysis and binding experiments with chemically modified RNAs indicate that the tighter binding of the C-RNA must involve one or more phosphates. The ionic strength dependence of $K_d$ suggested that four to five ion pairs contributed to the overall binding energy (50). This number correlates well the five salt bridges found in the cocrystal structures (36, 37). Binding experiments with thiophosphate-modified RNAs made by transcription located four phosphates that affect the affinity (51). Recently, a more thorough analysis of RNAs with singly substituted phosphorothioates has identified eight phosphates sensitive to this modification (D. Dertinger, L. S. Behlen, and O. C. Uhlenbeck, unpublished results). One of these is phosphate −6. U- and C-RNAs, which contain a single sulfur atom in place of either of the nonbridging oxygens at −6, bound coat protein with affinities only twofold tighter than all-oxygen U-RNA. U- and C-RNA with a dithiophosphate at −6 had equal affinities that were both approximately 10 times greater than the all-oxygen U-RNA. Finally, singly methylphosphonate substituted U- and C-RNAs bound with the same affinity as the all-oxygen U-RNA (H. E. Johansen *et al.*, unpublished). In other words, although the absolute affinity of the different phosphate −6 modifications varied, the C-RNA no longer bound more

Copyright © 1997 by Academic Press

tightly. Thus the selective binding of C-RNA over U-RNA by the MS2 coat protein is primarily due to the internal charged RNA hydrogen bond between the amino group of C and the phosphate at −6. Although this hydrogen bond is not present in the structure of the free RNA, the higher affinity must be a consequence of the increased propensity of this conformation to form in the free state.

## ACKNOWLEDGMENTS

We thank Linda Behlen, Dagmar Dertinger, Kerstin Fridborg, and Drs. Narayan Baidya, Karen LeCuyer, and Karin Valegård for help with figures and stimulating discussions. Work in the Uhlenbeck laboratory is supported by a grant from the National Institutes of Health (GM36944) and in the Liljas laboratory by the Swedish Natural Science Research Council. HEJ was supported by a postdoctoral fellowship from Swedish Natural Science Research Council.

## REFERENCES

1. Beckett, D., Wu, H., and Uhlenbeck, O. C. (1988) Roles of operator and non-operator RNA sequences in bacteriophage R17 capsid assembly. *J. Mol. Biol.* **204,** 939–947.
2. Pickett, G. G., and Peabody, P. S. (1993) Encapsidation of heterologous RNAs by bacteriophage MS2 coat protein. *Nucleic Acids Res.* **21,** 4621–4626.
3. Bernardi, A., and Spahr, P.-F. (1972) Nucleotide sequence at the binding site for coat protein on RNA of bacteriophage R17. *Proc. Natl. Acad. Sci. USA* **69,** 3033–3037.
4. Peabody, D. S. (1993) The RNA binding site of MS2 coat protein. *EMBO J.* **12,** 595–600.
5. Peabody, D. S., and Lim, F. (1996) Complementation of RNA binding site mutations in MS2 coat protein heterodimers. *Nucleic Acids Res.* **24,** 2352–2359.
6. Fouts, D. E., and Celander, D. W. (1996) Improved method for selecting RNA-binding activities in vivo. *Nucleic Acids Res.* **24,** 1582–1584.
7. Lim, F., Spignola, M., and Peabody, D. S. (1996) The RNA-binding site of bacteriophage Qβ coat protein. *J. Biol. Chem.* **271,** 31839–31845.
8. Wang, S., True, H. L., Seitz, E. M., Bennet, K. A., Fouts, D. E., Gardner, J. F., and Celander, D. W. (1997) Direct genetic selection of two classes of R17/MS2 coat proteins with altered capsid assembly properties and expanded RNA-binding activities. *Nucleic Acids Res.* **25,** 1649–1657.
9. Witherell, G. W., Gott, J. M., and Uhlenbeck, O. C. (1991) Specific interaction between RNA phage coat proteins and RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **40,** 185–220.
10. Gott, J. M., Pan, T., LeCuyer, K. A., and Uhlenbeck, O. C. (1993) Using circular permutation analysis to redefine the R17 coat protein binding site. *Biochemistry* **32,** 13399–13404.
11. Stockley, P. G., Stonehouse, N. J., Walton, C., Walters, D. A., Medina, G., Macedo, J. M., Hill, H. R., Goodman, S. T., Talbot, S. J., Tewary, H. K., Golmohammadi, R., Liljas, L., and Valegård, K. (1993) Molecular mechanism of RNA-phage morphogenesis. *Biochem. Soc. Transact.* **21,** 627–633.
12. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1995) Mutants of the bacteriophage MS2 coat protein which alter its cooperative binding to RNA. *Biochemistry* **34,** 10600–10606.
13. Baidya, N., and Uhlenbeck, O. C. (1995) The role of 2′-hydroxyl groups in an RNA–protein interaction. *Biochemistry* **34,** 12363–12368.
14. Stockley, P. G., Stonehouse, N. J., Murray, J. B., Goodman, S. T. S., Talbot, S. J., Adams, C. J., Liljas, L., and Valegård, K. (1995) Probing sequence-specific RNA recognition by the bacteriophage MS2 coat protein. *Nucleic Acids Res.* **23,** 2512–2518.
15. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1996) Mutagenesis of a stacking contact in the MS2 coat protein–RNA complex. *EMBO J.* **15,** 6847–6853.
16. Lim, F., Spingola, M., and Peabody, D. S. (1994) Altering the RNA binding specificity of a translational repressor. *J. Biol. Chem.* **269,** 9006–9010.
17. Lim, F., and Peabody, D. S. (1994) Mutations that increase the affinity of a translational repressor for RNA. *Nucleic Acids Res.* **22,** 3748–3752.
18. McGregor, A., Vaman Rao, M., Duckworth, G., Stockley, P. G., and Connolly, B. A. (1996) Preparation of oligonucleotides containing 4-thiouridine using Fpmp chemistry: Photo-crosslinking to RNA binding proteins using 350 nm irradiation. *Nucleic Acids Res.* **24,** 3173–3180.
19. Gott, J. M., Willis, M. C., Koch, T. H., and Uhlenbeck, O. C. (1991) A specific, UV-induced RNA–protein crosslink using 5-bromouridine substituted RNA. *Biochemistry* **30,** 6290–6295.
20. Willis, M. C., Hicke, B. J., Uhlenbeck, O. C., Cech, T. R., and Koch, T. H. (1993) Photocrosslinking of 5-iodouracil-substituted RNA and DNA to proteins. *Science* **262,** 1255–1257.
21. Willis, M. C., LeCuyer, K. A., Meisenheimer, K. M., Uhlenbeck, O. C., and Koch, T. H. (1994) An RNA–protein contact determined by 5-bromouridine substitution, photocrosslinking and sequencing. *Nucleic Acids Res.* **22,** 4947–4952.
22. Meisenheimer, K. M., Meisenheimer, P. L., Willis, M. C., and Koch, T. H. (1996) High yield photocrosslinking of a 5-iodocytidine (IC) substituted RNA to its associated protein. *Nucleic Acids Res.* **24,** 981–982.
23. Stripecke, R., and Hentze, M. W. (1992) Bacteriophage and spliceosomal proteins function as position-dependent *cis/trans* repressors of mRNA translation in vitro. *Nucleic Acids Res.* **20,** 5555–5564.
24. Carroll, R., Peterlin, B. M., and Derse, D. (1992) Inhibition of human immunodeficiency virus type 1 Tat activation by coexpression of heterologous trans activators. *J. Virol.* **66,** 2000–2007.
25. Mathews, S. A., and Sriprakash, K. S. (1994) R17 coat protein binding site: a convenient reporter for in vitro transcription. *Nucleic Acids Res.* **22,** 534–535.
26. Selby, M. J., and Peterlin, B. M. (1990) Trans-activation of HIV-1 Tat via a heterologous RNA binding protein. *Cell* **62,** 769–776.
27. Venkatesan, S., Gerstberger, S. M., Park, H., Holland, S. M., and Nam, Y. (1992) Human immunodeficiency virus type 1 Rev activation can be achieved without Rev-responsive element RNA if Rev is directed to the target as a Rev/MS2 fusion protein which tethers the MS2 operator RNA. *J. Virol.* **66,** 7469–7480.
28. SenGupta, D. J., Zhang, B., Kraemer, B., Pochart, P., Fields, S., and Wickens, M. (1996) A three-hybrid system to detect RNA-protein interactions in vivo. *Proc. Natl. Acad. Sci. USA* **93,** 8496–8501.
29. Romaniuk, P. J., Lowary, P., Wu, H. N., Stormo, G., and Uhlenbeck, O. C. (1987) RNA binding site of R17 coat protein. *Biochemistry* **26,** 1563–1568.

Copyright © 1997 by Academic Press

30. Schneider, D., Tuerk, C., and Gold, L. (1992) Selection of high affinity RNA ligands to the bacteriophage R17 coat protein. *J. Mol. Biol.* **288,** 862–869.
31. Carey, J. C., Lowary, P. T., and Uhlenbeck, O. C. (1983) Interaction of R17 coat protein with synthetic variants of its RNA binding site. *Biochemistry* **22,** 4723–4730.
32. Wu, H.-N., and Uhlenbeck, O. C. (1987) Role of a bulged A residue in a specific RNA–protein interaction. *Biochemistry* **26,** 8221–8227.
33. Talbot, S. J., Goodman, S., Bates, S. R., Fishwick, C. W., and Stockley, P. G. (1990) Use of synthetic oligoribonucleotides to probe RNA–protein interactions in the MS2 translational operator complex. *Nucleic Acids Res.* **18,** 3521–3528.
34. Mastico, R. A., Talbot, S. J., and Stockley, P. G. (1993) Multiple presentation of foreign peptides on the surface of an RNA-free spherical bacteriophage capsid. *J. Gen. Virol.* **74,** 541–548.
35. Peabody, D. S. (1990) Translational repression by bacteriophage MS2 coat protein expressed from a plasmid: A system for genetic analysis of a protein–RNA interaction. *J. Biol. Chem.* **265,** 5684–5689.
36. Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., and Liljas, L. (1994) Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature* **371,** 623–626.
37. Valegård, K., Murray, J. B., Stonehouse, N. J., van den Worm, S., Stockley, P. G., and Liljas, L. (1997) The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions. *J. Mol. Biol.* **270,** 727–728.
38. Golmohammadi, R., Fridborg, K., Bundule, M., Valegård, K., and Liljas, L. (1996) The structure of bacteriophage Qβ at 3.5 Å resolution. *Structure* **4,** 543–554.
39. Mattaj, I. W., and Nagai, K. (1995) Recruiting proteins to the RNA world. *Nature Struct. Biol.* **2,** 518–522.
40. Borer, P. N., Lin, Y., Wang, S., Roggenbuck, M. W., Gott, J. M., Uhlenbeck, O. C., and Pelczer, I. (1995) Proton NMR and structural features of a 24-nucleotide RNA hairpin. *Biochemistry* **34,** 6488–6503.
41. Clackson, T., and Wells, J. A. (1995) A hot spot of binding energy in a hormone–receptor interface. *Science* **267,** 383–386.
42. Peabody, D. S., and Ely, K. R. (1992) Control of translational repression by protein–protein interactions. *Nucleic Acids Res.* **20,** 1649–1655.
42a. Spingola, M., and Peabody, D. S. (1997) MS2 coat protein mutants which bind Qβ RNA. *Nucleic Acids Res.* **25,** 2808–2815.
43. Beckett, D., and Uhlenbeck, O. C. (1988) Ribonucleoprotein complexes of R17 coat protein and a translational operator analog. *J. Mol. Biol.* **204,** 927–938.
44. Milligan, J. F. (1988) Identification of protein–phosphate contacts in the R17 coat protein binding site. Ph.D. thesis. University of Illinois at Urbana–Champaign.
45. Uhlenbeck, O. C., Carey, J., Romaniuk, P., Lowary, P. T., and Beckett, D. (1983) Interaction of R17 coat protein with its RNA binding site for translational repression. *J. Biomol. Struct. Dynam.* **1,** 539–552.
46. Lowary, P. T., and Uhlenbeck, O. C. (1987) An RNA mutation that increases the affinity of an RNA–protein interaction. *Nucleic Acids Res.* **15,** 10483–10493.
47. Schoemaker, H. J. P., and Schimmel, P. R. (1977) Effect of transfer RNA synthetases on H-5 exchange of specific pyrimidines in transfer RNAs. *Biochemistry* **16,** 5454–5464.
48. Starzyk, R. M., Koontz, S. W., and Schimmel, P. (1982) A covalent adduct between the uracil ring and the active site of an aminoacyl tRNA synthetase. *Nature* **298,** 136–140.
49. Peabody, D. S. (1989) Translational repression by bacteriophage MS2 coat protein does not require cysteine residues. *Nucleic Acids Res.* **17,** 6017–6027.
50. Carey, J. C., and Uhlenbeck, O. C. (1983) Kinetic and thermodynamic characterization of the R17 coat protein-RNA interaction. *Biochemistry* **22,** 2610–2615.
51. Milligan, J. F., and Uhlenbeck, O. C. (1989) Determination of RNA–protein contacts using thiophosphate substitutions. *Biochemistry* **28,** 2849–2855.
52. Bacon, D. J., and Anderson, W. F. (1988) A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6,** 219–220.
53. Merritt, E. A., and Murphy, M. E. P. (1994) Raster3D Version 2.0: A program for photorealistic molecular graphics. *Acta Crystallogr.* **D50,** 869–873.
54. Kraulis, P. J. (1991) Molscript: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24,** 946–950.

Copyright © 1997 by Academic Press