The EMBO Journal vol.15 no.24 pp.6847–6853, 1996

Mutagenesis of a stacking contact in the MS2 coat protein–RNA complex

Karen A.LeCuyer[1], Linda S.Behlen and Olke C.Uhlenbeck[2]

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

[1]Present address: Department of Biochemistry, University of Connecticut Health Science Center, Farmington, CT 06032, USA

[2]Corresponding author

The thermodynamic contribution of a stacking interaction between Tyr85 in MS2 coat protein and a single-stranded pyrimidine in its RNA binding site has been examined. Mutation of Tyr85 to Phe, His, Cys, Ser and Ala decreased the RNA affinity by 1–3 kcal/mol under standard binding conditions. Since the Phe, His and Cys 85 proteins formed UV photocrosslinks with iodouracil-containing RNA at the same rate as the wild-type protein, the mutant proteins interact with RNA in a similar manner. The pH dependence of K_D for the Phe and His proteins differs substantially from the wild-type protein, suggesting that the titration of position 85 contributes substantially to the binding properties. Experiments with specifically substituted phosphorothioate RNAs confirm a hydrogen bond between the hydroxyl group of tyrosine and a phosphate predicted by the crystal structure.

Keywords: phosphorothioates/photocrosslinking/RNA–protein interactions

Introduction

Stacking interactions between nucleic acid bases and aromatic amino acids have been observed in several RNA–protein and single-stranded DNA–protein complexes (Mattaj and Nagai, 1995). The existence of such interactions was proposed 25 years ago, based on the observation that the intrinsic fluorescence of aromatic amino acids was quenched when they were mixed with nucleic acid bases (Montenay-Garestier and Helene, 1971). Although fluorescence quenching is often seen in nucleic acid–protein complexes, stronger evidence for the close proximity between a nucleic acid base and an aromatic amino acid comes from NMR spectroscopy and photochemical crosslinking experiments. The NMR chemical shifts of certain Tyr and Phe residues in the Ff bacteriophage gene V protein are substantially changed when single-stranded DNA is bound to the protein (King and Coleman, 1988). Irradiation of a complex of bacteriophage T4 gp32 protein and dT_8 results in a crosslink to Phe183 (Shamoo *et al.*, 1988). More recently, co-crystal structures of DNA–protein complexes have revealed several base–amino acid stacking interactions. For example, in the complex of dT_4 in the 3′,5′ exonuclease active site of Klenow polymerase, Phe473 stacks on the 3′ terminal T residue (Beese and Steitz, 1991). The more recent co-crystal structure of T4 gp32 protein and dT_6 confirmed the stacking of Phe183 on a thymidine residue (Shamoo *et al.*, 1995).

Many of the co-crystal structures of RNA–protein complexes also show examples of stacking between bases and aromatic amino acids. The single-stranded anticodon residue U35 in $tRNA^{Asp}$ stacks on Phe127 of the aspartyl-tRNA synthetase, a conserved residue in this group of synthetases (Cavarelli *et al.*, 1993). The terminal A76 of $tRNA^{Gln}$ stacks beneath Phe233 and Tyr211 of glutaminyl-tRNA synthetase (Perona *et al.*, 1993). Finally, the co-crystal structure of the U1A RNA binding domain with a small RNA hairpin reveals two such stacking interactions: Tyr13 with C-10 and Phe56 between A-11 and C-12 (Oubridge *et al.*, 1994). In the U1A complex, Tyr13 efficiently photocrosslinks to a 5-iodouracil-substituted RNA at position 10 (Stump and Hall, 1995).

The interaction of the MS2 (R17) bacteriophage coat protein with its operator RNA contains an example of a stacking interaction between a tyrosine and a cytidine residue. In the crystal structure of this RNA–protein complex within an assembled capsid, Tyr85 protrudes from the β sheet and stacks on one side of a single-stranded cytidine residue in a four-base RNA loop (Figure 1) (Valegård *et al.*, 1994). Mutation of this cytidine to other RNA bases strongly reduces the binding constant of the RNA–protein complex (Carey *et al.*, 1983b; Stockley *et al.*, 1995). This RNA–protein contact also shows efficient UV photocrosslinking when the RNA contains 5-iodouracil or 5-bromouracil in place of the cytidine residue (Willis *et al.*, 1993). In order to evaluate the contribution of this stacking interaction to the RNA–protein affinity, mutation of Tyr85 to different aromatic and aliphatic amino residues was performed, and the affinity of the mutant proteins for RNA was studied in detail. The mutant proteins were also studied for their ability to form UV photocrosslinks.

Results

Structural considerations

In order to interpret the thermodynamic effects of mutating an amino acid side chain, it is critical to know the structures of the protein, the RNA and the RNA–protein complex. In the MS2 system, a co-crystal structure of a 19-nucleotide synthetic RNA fragment bound to coat protein inside of an assembled capsid is available at 3 Å resolution (Valegård *et al.*, 1994). Tyr85 protrudes from the β sheet and is stacked onto the C-5. The other side of C-5 is stacked on A-7 and forms the end of the RNA structure (Figure 1). A hydrogen bond is predicted to form between the hydroxyl group of Tyr85 and the O1P 5′ to

© Oxford University Press

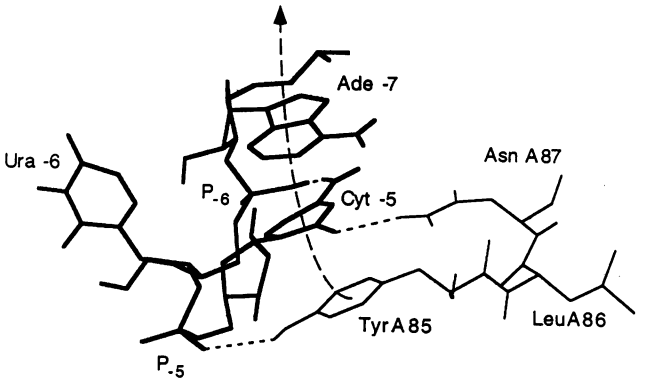

Fig. 1. Structure of the RNA–protein complex in the region around Tyr85 (Valegård *et al.*, 1994).

C-5. The C-5 is held in place by additional hydrogen bonds between the O2 of C-5 and the amide of Asn87, and the N4 of C-5 and the O1P of the residue at –6. There are two crystal structures available for the free MS2 coat protein. One is the structure of an empty capsid in which Tyr85 is not significantly displaced from its position in the RNA–protein complex (Golmohammadi *et al.*, 1993). In a structure of a coat protein dimer which does not assemble into capsids, Tyr85 is displaced 6Å from its position in the capsid (Ni *et al.*, 1995). It is unclear whether the difference between these two structures is a consequence of the different packing constraints in the two systems or a result of flexibility in Tyr85. The structure of a free RNA hairpin with a similar sequence determined by NMR was found to be quite different from the crystal structure of the RNA in the RNA–protein complex (Borer *et al.*, 1995). In the free RNA, the C-5 is stacked onto the A-4, while G-8, A-7 and U-6 form a separate stack. However, the NMR data suggest that the loop nucleotides are flexible and easily adopt the structure in the RNA–protein complex.

RNA binding properties of 85 position mutants

When the wild-type MS2 coat protein is overexpressed in *Escherichia coli*, phage-like capsid particles are formed. Treatment with 50% acetic acid is required to dissociate these capsids into dimers suitable for studying the RNA–protein interaction in solution (Beckett and Uhlenbeck, 1988). However, this acetic acid treatment frequently denatures mutant proteins. In a recent paper, we have shown that the Val75Glu;Ala81Gly double-mutant coat protein forms stable dimers that have identical affinity and specificity for RNA as wild-type coat protein, but cannot form capsids *in vivo* and *in vitro* (LeCuyer *et al.*, 1995). All of the mutants studied in this work also contain the Val75Glu and Ala81Gly mutations. These proteins were purified by ammonium sulfate precipitation and S-Sepharose chromatography as described in LeCuyer *et al.* (1995).

Plasmids where Tyr85 was changed to Phe, His, Cys, Ser, Trp, Leu and Ala were prepared by cassette mutagenesis. The Phe, His, Cys, Ser and Ala85 mutations all produced proteins which overexpressed to similar levels as the wild-type protein. The Trp and Leu mutations, however, produced no detectable protein upon overexpression.

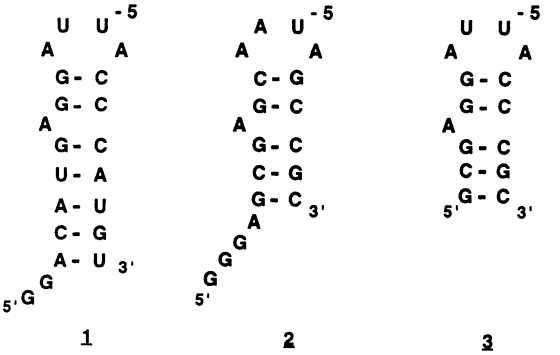

Fig. 2. Sequences and secondary structures of RNA hairpins. The phosphate 5′ to position –5 is substituted with a thiophosphate for experiments involving RNA 2 and RNA 3.

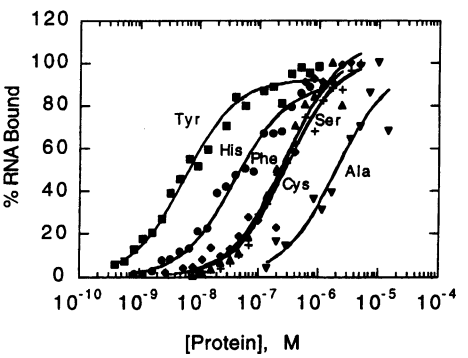

Fig. 3. Coat protein excess binding curves for the mutant proteins and RNA 1. The K_D values were determined by fitting a bimolecular equilibrium. Curves were normalized to 100% binding at the highest protein concentrations. The actual retention efficiencies varied from 30 to 60%. Symbols are: Tyr (■), His (●), Phe (▲), Cys (◆), Ser (+) and Ala (▼).

Dissociation constants for the various protein–RNA complexes were determined using a nitrocellulose filter binding assay (Carey *et al.*, 1983a). All filter binding experiments were performed with RNA 1 which has a U at the –5 position despite the fact that the structures of both the RNA–protein complex and the free RNA have a C at –5 (Figure 2). This was done because C-5 RNAs bind significantly tighter than U-5, making it difficult to obtain accurate equilibrium K_D measurements (Lowary and Uhlenbeck, 1987). The recently refined co-crystal structure containing a U-5 RNA demonstrates that U-5 and C-5 RNAs bind protein virtually identically (K.Valegård and L.Liljas, personal communication). In the standard binding conditions, the Phe, His, Cys, Ser and Ala proteins all bind less well than the wild-type Tyr (Figure 3, Table I). The Tyr85Ala substitution corresponds to the lowest-affinity protein with a ΔΔG relative to wild type of +3 kcal/mol. The Phe, Ser and Cys substitutions all result in a ΔΔG of about +2 kcal/mol. The His85 substitution is the least detrimental, with a ΔΔG of +1 kcal/mol.

If the mutant proteins are not fully active for RNA binding, the K_D values measured in protein excess may be in error. To determine the fraction of protein active in RNA binding, RNA excess titrations were performed (Witherell and Uhlenbeck, 1989). The amount of RNA bound increases with increasing amounts of RNA until saturation is reached. The ratio of RNA to protein at

Table I. RNA affinities of mutant proteins for RNA 1

Position 85	K_D (nM)	Relative K_D
Tyr	6.2	(1)
Phe	170	27
His	33	5.3
Cys	280	45
Ser	300	48
Ala	1000	160

The dissociation constants are determined in 100 mM Tris–Cl (pH 8.5), 10 mM Mg $(OAc)_2$ and 80 mM KCl at 4°C.

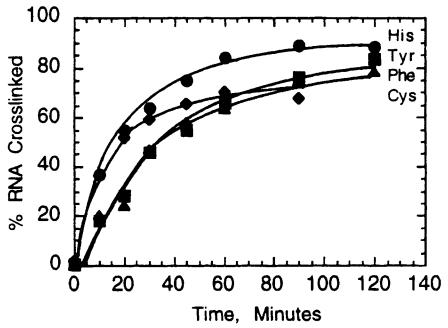

Fig. 4. Time courses of UV photocrosslinking to 5-iodouracil-substituted RNA 2 for the Tyr (■), His (●), Phe (▲) and Cys (◆) proteins.

saturation gives the fraction of active protein. All of the mutant proteins >95% were active for RNA binding when they were first purified (data not shown). Because the mutant proteins purify in an identical manner to the wild type and are fully active for RNA binding, we believe that the mutants are not trapped in a denatured form. Furthermore, the K_D values reported in the previous paragraph require no correction for the presence of inactive protein.

Photocrosslinking

Photocrosslinking of each mutant protein with 5-iodouracil-substituted RNA 2 was performed with 325 nm laser light in order selectively to excite the 5-iodouracil and prevent the degradation of RNA and protein observed at lower wavelengths (Figure 2) (Willis *et al.*, 1993). For each mutant, the protein concentration was saturating and time courses of photocrosslinking were determined. Proteins with Tyr, Phe, His and Cys at position 85 all crosslink to 5-iodouracil-substituted RNA 2 to similar extents with half-lives of 44, 47, 32 and 17 min, respectively (Figure 4). The Ser and Ala85 proteins do not crosslink, and are not expected to undergo this photocrosslinking reaction efficiently (data not shown) (Dietz and Koch, 1991). Because the Phe, His and Cys mutant proteins all crosslink similarly to the wild-type Tyr, they are presumed to interact with the RNA in a similar manner. Thus, the measured differences in RNA affinity are likely the result of small, local differences in the interaction and not large structural rearrangements.

pH dependence of K_D

Binding constants for the Tyr, Phe and His proteins were determined as a function of pH. As has been reported previously, the Tyr protein binds tightest at pH 7.5, with

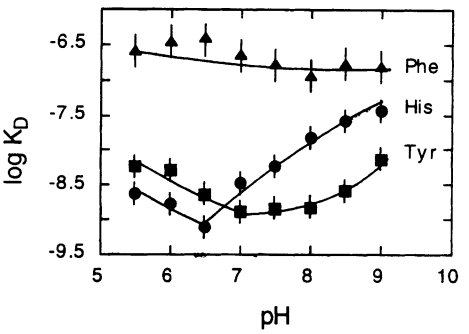

Fig. 5. Dependence of K_D on pH for the Tyr (■), His (●) and Phe (▲) proteins with RNA 1. The vertical lines indicate the error among duplicate measurements.

the K_D increasing modestly with either an increase or decrease in pH (Figure 5) (LeCuyer *et al.*, 1995). The Phe protein shows a modest pH dependence of K_D with a minimum at pH 8.0. In contrast, the His protein has a pronounced pH profile with a minimum around pH 6.5. It is, therefore, clear that the amino acid at the 85 position contributes significantly to the pH dependence of K_D. In the case of Tyr and His, this increase in K_D at higher pH may be due to deprotonation of Tyr or His. While the pK_a of free Tyr is 9.7 and that of His is 6.2, the pK_as of these residues in the protein are unknown. All three proteins have some decrease in RNA affinity below pH 6.0 or 7.0, possibly due to other titratable groups in the protein.

In order to determine whether His85 participates in an ion pair with the RNA when it is protonated, the ionic strength dependence of K_D was examined. When a protein and a nucleic acid form a complex, counter-ions are released from the surface of the macromolecules. This counter-ion release results in a dependence of the dissociation constant on salt concentration. The thermodynamic analysis of the salt dependence of K_D based on ion displacement has been developed for DNA homopolymers (Lohman *et al.*, 1980). Although the analysis was performed for DNA–protein complexes, the same analysis may be used to approximate the number of ionic contacts in an RNA–protein complex. Past experiments with MS2 coat protein conclude that, at pH 8.5, 4–5 ion pairs are formed upon RNA binding (Carey and Uhlenbeck, 1983), which correlates well with the five ionic contacts observed in the co-crystal structure (Valegård *et al.*, 1994). Although this analysis gives only an estimate of the number of ionic contacts in an RNA–protein complex, any differences between Tyr and His should be apparent.

The salt dependence of K_D was determined to evaluate the number of ionic contacts in the Tyr and His complexes. A buffer containing 50 mM each of MES and HEPES was used so that both pH 6.0 and pH 8.2 could be tested in the same buffer. The cation concentration was changed by the addition of KCl. The slope of the log K_D versus log $[M^+]$ plot indicates that the wild-type coat protein makes five ionic contacts at pH 6.0 and four at pH 8.2 (Figure 6). Thus, the wild-type protein gains an additional ionic contact at low pH. In the case of the His85 protein, the data in Figure 6 indicate four ionic contacts at pH 8.2 and six at pH 6.0. The His protein thus gains one additional ionic contact at low pH, suggesting that His85 may form an ion pair with the RNA.

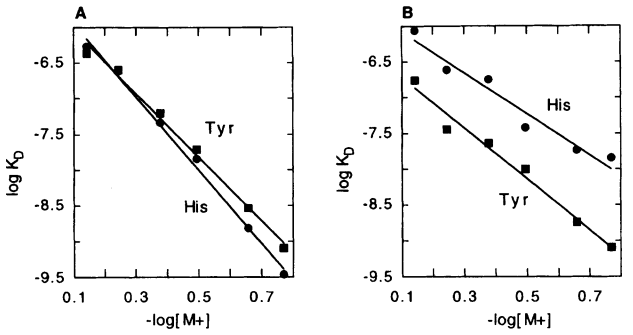

Fig. 6. Ionic strength dependence of K_D for the Tyr (■) and His (●) proteins with RNA 1 at pH 6.0 (**A**) and pH 8.2 (**B**). The number of ionic contacts is given by the slope of the line log K_D versus log [M+] divided by 0.85. The calculated slopes are 3.6 at pH 8.2 and 4.5 at pH 6.0 for Tyr, and 3.2 at pH 8.2 and 5.2 at pH 6.0 for His.

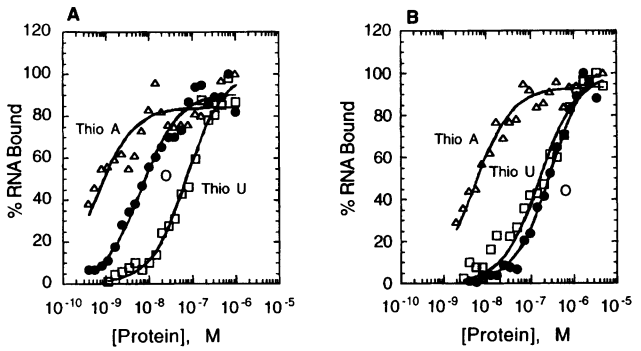

Fig. 7. Protein excess binding curves for the Tyr (**A**) and Phe (**B**) proteins with all-oxygen (●), (αS)-U-containing (□) and (αS)-A-containing (△) RNA 2.

Probing RNA backbone contacts with thiophosphate modification

The crystal structure of the MS2 RNA–protein complex shows that the hydroxyl group at Tyr85 hydrogen bonds to one of the phosphate oxygens of the C-5 (Valegård *et al.*, 1994). Thiophosphate substitution has been used previously to probe interactions of the MS2 coat protein with the RNA phosphate backbone (Milligan and Uhlenbeck, 1989). RNA 2 was used for these experiments because UTP (αS) can be inserted uniquely at the –5 position by transcription. When the affinity for UTP (αS)-substituted RNA 2 is examined, the wild-type protein shows a 15-fold increase in K_D as compared to all-oxygen RNA (Figure 7A). This increase in K_D observed upon thiophosphate substitution is presumably due to a disruption in the hydrogen bond between the hydroxyl of Tyr85 and the pro-R_P oxygen of the –5 phosphate. RNA 2 transcribed with ATP (αS) contains five thiophosphates and shows a 10-fold decrease in K_D compared to the all-oxygen RNA (Figure 7A). The hydrogen bonds of Lys61 to phosphate –10 and Lys57 to phosphate –7 seen in the co-crystal structure are presumably responsible for this increased RNA affinity. While the ATP (αS) experiment agrees with the previous results, the UTP (αS) experiment does not (Milligan and Uhlenbeck, 1989). Because no effect was previously seen with UTP (αS)-substituted RNA 2, it is possible that the RNA used in that experiment did not contain a thiophosphate.

The experiment with thiophosphate-substituted RNA 2 was repeated with the Phe85 protein (Figure 7B). Because

Table II. pH dependence of His85 protein binding to RNA 3 with phosphate –5 substitution

pH	K_D (nM)		
	Oxygen	pro-R_P sulfur	pro-S_P sulfur
6.0	1.4	1.5	3.6
7.5	2.8	3.1	5.4
8.5	9.0	9.0	26

Experiments were performed in 100 mM Tris–Cl, 10 mM $Mg(OAc)_2$ and 80 mM KCl at the indicated pH.

Phe is missing the hydroxyl group of Tyr, it cannot hydrogen bond with the phosphate oxygen at –5. Unlike the case of the Tyr protein, the affinity of the Phe protein for the UTP (αS)-substituted RNA 2 is identical to that for the all-oxygen RNA. Like the wild-type Tyr protein, the Phe85 protein shows a 10-fold decrease in K_D with the ATP (αS)-substituted RNA, indicating that it contacts the –10 and –7 phosphates in a similar manner to the wild-type protein. These thiophosphate substitution experiments provide biochemical evidence for the hydrogen bond between the hydroxyl group of Tyr85 and the –5 pro-R_p oxygen.

In order to test the hypothesis that the protonated His85 protein forms a charged hydrogen bond with the –5 phosphate oxygen, the same thiophosphate substitution experiments were performed with His85 at pH 6.0 and pH 8.2. When the affinity of His for the UTP (αS)-substituted RNA 2 is determined, the K_D is identical to that of the all-oxygen RNA at both pHs (data not shown). This result does not support the idea that the protonated His protein is stabilized by interaction with the pro-R_P oxygen at –5. In order to investigate this result more thoroughly, experiments were performed with RNA 3 which can be chemically synthesized with a phosphorothioate at position –5 as either a pro-R_P or pro-S_P isomer (Figure 2). The pro-R_P isomer of 3 has the same affinity as the all-oxygen RNA, in agreement with the results obtained with RNA 2 (Table II). However, the pro-S_P isomer shows a 2- to 3-fold decrease in affinity, suggesting that the His protein may interact with one of the phosphate oxygens 5′ to U-5. Unexpectedly, the K_D values for all three RNAs vary with pH in an identical manner (Table II), indicating that the weaker binding of the pro-S_P isomer does not depend on the protonation state of His85.

Discussion

Mutagenesis of Tyr85 of the MS2 bacteriophage coat protein clearly indicates that this amino acid side chain contributes substantially to the free energy of binding to the RNA hairpin. The total binding energy of 12 kcal/mol is reduced by 1–3 kcal/mol when the Tyr is replaced by Phe, His, Cys, Ser or Ala. This result is consistent with the observation that RNA mutations of C-5, the stacking partner of Tyr85, also substantially alter complex stability (Carey *et al.*, 1983b). Stacking contacts are thus not only common in complexes between proteins and single-stranded nucleic acids, but they make a significant contribution to the binding energy. A similar conclusion was

reached when Tyr13 of the U1A RNP domain was substituted with a Phe or a Thr (Jessen *et al.*, 1991; Stump and Hall, 1995).

Attempts to interpret the decrease in binding energy of the different position 85 mutations in terms of molecular interactions are complicated by several issues. First of all, the interactions of the tyrosine with the RNA involve not only stacking, but also a hydrogen bond between the hydroxyl group of tyrosine and a phosphate oxygen. Since this hydrogen bond involves a charged residue, its intrinsic energy may be quite large (Fersht *et al.*, 1985). In addition, the hydrogen bond will fix the dipole of the Tyr with respect to the dipole of the uridine ring, thus influencing energy derived from stacking. The second major complexity in interpreting the mutagenesis data in a detailed way is shared with many mutagenesis experiments (Jessen *et al.*, 1991; Brown *et al.*, 1994). Since the structure of the mutant protein in either its free form or when complexed with RNA is unknown, it is difficult to attribute a change in binding energy to a particular molecular interaction. For example, the K_D could be increased because of an alternative structure of the mutant protein which must be disrupted in order to bind RNA. Alternatively, the structure of the mutant protein in the complex may have altered side chain packing or extensive changes in solvent hydration that could have substantial energy associated with them. Although the free energy differences measured for the mutant proteins are more complex than the simple removal of an interaction, conclusions can be drawn about the contribution of the mutated amino acid to the formation of a stable complex.

The only available information about the structure of the mutant protein complexes is that efficient UV photocrosslinking is observed when the adjacent uridine is replaced with 5-iodouridine. Proteins with Tyr, Phe, His and Cys all crosslink with similar efficiencies and rates. The details of the photocrosslinking reaction are likely to be different in each case and oriented model compound studies are not yet available (Dietz and Koch, 1987). Thus, the only conclusion which can be made is that the side chain must remain close to the U-5 in the complex. Crosslinks to Tyr, His and Phe have been previously reported (Willis *et al.*, 1993, 1994; Stump and Hall, 1995). A crosslink between 5-iodouracil and Cys, although predicted from model compound studies, has not been previously observed in an RNA–protein complex (Dietz and Koch, 1991). As expected, the Ser and Ala proteins do not crosslink to the 5-iodouracil-substituted RNA (Dietz and Koch, 1991). Our results are in agreement with those for the U1A protein in which a 5-iodouracil-substituted RNA crosslinks with Tyr13 and Phe13, but not with Thr13 (Stump and Hall, 1995). It therefore appears that this UV crosslinking reaction is an efficient means of detecting stacking interactions in RNA–protein complexes.

Of all the mutations of position 85, the 27-fold weaker binding of the Phe protein corresponding to a $\Delta\Delta G$ of 2 kcal/mol is the easiest to understand. A significant fraction of the reduced binding energy is due to the inability of the Phe protein to form a hydrogen bond with the –5 phosphate oxygen. The importance of this hydrogen bond in the wild-type protein is confirmed by the observation that replacement of the phosphate with a pro-R_P thiophosphate reduces affinity by 10-fold. Although the sulfur in a thiophosphate contains a greater portion of the electronegativity, it is a poor hydrogen bond acceptor compared to oxygen (Fersht, 1987). As would be expected, the Phe protein does not show an altered affinity for the thiophosphate-substituted RNA. It is unlikely, however, that the entire reduction in binding energy can be attributed to the absence of this hydrogen bond. The dipole moment of Tyr is substantially greater than that of Phe, resulting in improved stacking with a nucleotide ring (Montenay-Garestier and Helene, 1971). In addition to its weaker dipole moment, Phe lacks the hydroxyl group of Tyr which may act favorably to stack the Tyr on the U-5. The weaker binding of the Phe protein is thus due to a combination of its altered hydrogen bonding and stacking properties as compared to Tyr.

Another interesting characteristic of the Phe mutant is the modest pH dependence of its K_D. While the wild-type Tyr has a dramatic decrease in RNA affinity at higher pH, the K_D for Phe changes very little in the 7.5–9 pH range. This result suggests that the sharp pH dependence of K_D for the wild-type coat protein at higher pH is due to Tyr85. It is possible that deprotonation of the Tyr (pK_a 9.7) is responsible for this decrease in RNA affinity. Both Phe and Tyr show an increase in K_D at lower pH. Because this feature is present in both proteins, it must be due either to other titratable groups on the protein or pH-dependent changes in the RNA structure.

The His mutation is interesting because it introduces a significant pH dependence to the RNA affinity. At high pH, the His protein binds less well than the wild-type Tyr, while at low pH, it binds somewhat tighter. Since the apparent pK_a of this 18-fold change in binding affinity is close to the pK_a of free histidine, the protonated form of His must be able to form an additional productive interaction with the RNA. Because the ionic strength dependence of the K_D for this His protein is greater at low pH rather than at high pH, it appears that the extra binding energy has substantial ionic character. A somewhat similar situation has been observed for the interaction of the Mnt transcriptional repressor with its DNA operator (Vershon *et al.*, 1985). In the case of Mnt, a His present in the wild-type protein is responsible for a 13-fold decrease in K_D between pH 8.0 and 6.5. This pH dependence of K_D is attributed to protonation of the His residue which can then donate a hydrogen bond at low pH.

Although the structure of the His protein complexed to RNA is unknown, the efficient crosslinking of the His protein to 5-iodouracil-substituted RNA suggests that its structure must be quite similar to Tyr. The additional binding energy at low pH may thus be the result of interaction with the –5 phosphate group. Protonation of His results in a charged hydrogen bond donor. In tyrosyl-tRNA synthetase, deletion of a side chain that forms a hydrogen bond with the substrate decreases the free energy of binding by 0.5–1.5 kcal/mol (Fersht *et al.*, 1985). Deletion of a side chain which hydrogen bonds to a charged group on the substrate decreases the binding energy by 3.5–4.5 kcal/mol. Thus, hydrogen bonds, where one of the partners is charged, are higher energy than normal hydrogen bonds. If, however, the His hydrogen bonds to a negatively charged phosphate oxygen, an ion pair will be formed.

If the His residue forms an ionic contact with the same

phosphate group which contacts the Tyr hydroxyl group, thiophosphate interference with the phosphate group 5′ to U-5 should be observed. Thiophosphate-substituted RNA 3 was used so that both the pro-R_P and pro-S_P isomer could be examined. While the pro-R_P isomer has the same affinity as the all-oxygen RNA, the pro-S_P isomer has decreased affinity. The His thus interacts with one of the phosphate oxygens 5′ of U-5. Interpreting the interference results, however, is complicated because changing O to S changes the identity of the hydrogen bonding group as well as the charge distribution. In the all-oxygen RNA, each of the phosphate oxygens will have a partial negative charge. In a thiophosphate, the sulfur takes the bulk of the negative charge (Frey and Sammons, 1985). Thus, in a pro-R_P thiophosphate, the oxygen in the pro-S_P position will effectively have no charge. A further complication is that oxygen is a better hydrogen bond acceptor than sulfur. A study of thiophosphate interference at many positions in the MS2 hairpin gives results which are difficult to interpret in detail (D.Dertinger, personal communication). The simplest conclusion one can draw is that if there is a contact, one or both of the isomers will show a change in K_D. In the case of the His mutant, there is an interaction between the protonated His and one of the phosphate oxygens 5′ to U-5.

The current study uses a variety of amino acid changes to assess the importance of Tyr85 to the formation of a stable RNA–protein complex. The Ala protein is of interest because it removes the possibility of any stacking or hydrogen bonding and results in a $\Delta\Delta G$ of 3 kcal/mol. Both Ser and Cys may restore the ability to hydrogen bond and bind 1 kcal/mol better than Ala. Phe, which can stack but not hydrogen bond, also binds 1 kcal/mol better than Ala. Thus, an amino acid which either stacks or hydrogen bonds restores part of the interaction energy. This conclusion is supported by His which can both stack and hydrogen bond, and binds RNA tighter than Ser, Cys or Phe. The combination of amino acid mutagenesis and RNA modification thus allows one to probe in detail an interaction observed in the crystal structure. In the MS2 complex, Tyr85 is clearly important for RNA recognition. All of the mutants, except His at low pH, show a significant loss in binding energy. Many of the other *E.coli* bacteriophage coat proteins, including fr, GA and Qβ, have a Tyr at the equivalent position (Witherell *et al.*, 1991). Interestingly, all of these proteins recognize different hairpin RNAs. It may be that stacking of Tyr is a key feature of all of these complexes.

Materials and methods

Cloning and purification

The parent plasmid for cloning (pKCONA) contains substitutions Val75-Glu;Ala81Gly and produces a coat protein that does not form capsids when overexpressed (LeCuyer *et al.*, 1995). Self-complementary oligonucleotides containing the 85 position mutations were cloned into the *Bgl*II–*Eco*RI fragment of this plasmid. The sequence of the mutagenic oligonucleotides was: 5′-TGC AGA TCT NNN TTA AAT ATG GAA CTA ACC ATT CCA ATT TTC GCT ACG AAT TCG T-3′, where NNN = TTC for Phe, CAC for His, TGC for Cys, TCC for Ser, GCC for Ala, TGG for Trp and CTC for Leu. The insert was annealed, extended by adding dNTPs and Klenow polymerase, and then digested with *Bgl*II and *Eco*RI. Vector and insert were ligated overnight with T4 DNA ligase (NEB). Ligation mixtures were transformed into JM83 cells. Sequencing was performed with the USB sequenase protocol. Clones with the correct sequence were transformed into BL21(DE3)pLysS cells (Studier *et al.*, 1990). All of the mutant proteins were purified as described previously (LeCuyer *et al.*, 1995). The proteins were stored at high concentrations and at 4°C in the FPLC buffer [25 mM HEPES (pH 7.0), ~40 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT)].

RNAs

[α-^{32}P]CTP-labeled RNAs were prepared by *in vitro* transcription from synthetic DNA templates as described in Gott *et al.* (1991). RNA 2 containing either 5-iodouracil, UTP(αS) or ATP(αS) was prepared using 2 mM modified nucleotide in place of the unmodified nucleotide in the transcription reaction. Thiophosphate-substituted RNA 3 was chemically synthesized and the isomers separated by HPLC (D.Dertinger, unpublished experiments).

UV photocrosslinking

Mixtures (250 µl) containing 50 nM ^{32}P-labeled RNA 2, either 0.5 µM (Tyr, His) or 10 µM (Phe, Cys, Ser) protein, 80 µg/ml bovine serum albumin (BSA), 5 mM DTT in TMK buffer [100 mM Tris (pH 8.5); 10 mM $Mg(OAc)_2$; 80 mM KCl] were incubated on ice for 30 min prior to crosslinking. An Omnichrome HeCd laser at 35 mW power was used to obtain 325 nm light. Samples were irradiated in methacrylate cuvettes. Samples of 20 µl were removed over the course of 2 h. Crosslinked and uncrosslinked RNA were separated on 20% denaturing polyacrylamide gels, and the yields were quantitated using a Molecular Dynamics PhosphorImager. Rates of crosslinking were determined from the slope of the line log (% crosslinked) versus time.

Nitrocellulose filter binding assays

Protein excess filter binding was performed in TMK buffer. A constant (<100 pM) concentration of ^{32}P-labeled RNA 1 was mixed with varying 0.7-fold dilutions of coat protein in a reaction volume of 100 µl. After incubation for 45 min, the reactions were filtered through a 0.45 µm Nitro ME filter (MSI, Westborough, MA) using the Schleicher and Schuell dot-blot apparatus described in Wong and Lohman (1993). The amount of radioactivity retained on the filter was quantitated using a Molecular Dynamics PhosphorImager. The data points were fit to a retention efficiency and K_D assuming a bimolecular equilibrium using the Kaleidagraph program.

To determine the active fraction of protein, a fixed concentration of protein above the measured K_D value was mixed with increasing amounts of RNA 1 in ratios ranging from 1:10 to 4:1 RNA to protein. All proteins were >95% active when first purified, with activity decreasing to 50% over several months. The K_D values determined in the protein excess experiments were corrected for the fraction of active protein.

The pH dependence of K_D was determined in 100 mM of several different buffers to give the appropriate pH range: MES (5.5, 6.0, 6.5), HEPES (7.0, 7.5, 8.0), Tris (8.5) and Ches (9.0).

The salt dependence of K_D was determined in a buffer of 50 mM MES + 50 mM HEPES so that both pH 6.0 and pH 8.2 could be obtained in the same buffer. The $Mg(OAc)_2$ concentration is 10 mM in all buffers and the KCl concentration is varied from 50 to 600 mM. The total ion concentration, [M+], is defined as the sum of [MES] + [HEPES] + 2[$Mg(OAc)_2$] + [KCl].

Acknowledgements

We wish to thank Kristen Meisenheimer and Tad Koch for help with photocrosslinking experiments, and Dagmar Dertinger for phosphorothioate RNA. This work was supported by National Institutes of Health Grant GM36944 to O.C.U. and an American Cancer Society postdoctoral fellowship to K.A.L.

References

Beckett,D. and Uhlenbeck,O.C. (1988) Ribonucleoprotein complexes of R17 coat protein and a translational operator analog. *J. Mol. Biol.*, **204**, 927–938.

Beese,L.S. and Steitz,T.A. (1991) Structural basis for the 3′-5′ exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.*, **10**, 25–33.

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.

Brown,B.M., Milla,M.E., Smith,T.L. and Sauer,R.T. (1994) Scanning mutagenesis of the Arc repressor as a functional probe of operator recognition. *Nature Struct. Biol.*, **1**, 164–168.

Carey,J.C. and Uhlenbeck,O.C. (1983) Kinetic and thermodynamic characterization of the R17 coat protein-RNA interaction. *Biochemistry*, **22**, 2610–2615.

Carey,J.C., Cameron,V., deHaseth,P.L. and Uhlenbeck,O.C. (1983a) Sequence specific interaction of R17 coat protein with its RNA binding site. *Biochemistry*, **22**, 2601–2610.

Carey,J.C., Lowary,P.T. and Uhlenbeck,O.C. (1983b) The interaction of R17 coat protein with synthetic variants of its RNA binding site. *Biochemistry*, **22**, 4723–4730.

Cavarelli,J., Rees,B., Ruff,M., Thierry,J.-C. and Moras,D. (1993) Yeast $tRNA^{Asp}$ recognition by its cognate class II aminoacyl-tRNA synthetase. *Nature*, **362**, 181–184.

Dietz,T.M. and Koch,T.H. (1987) Photochemical coupling of 5-bromouracil to tryptophan, tyrosine and histidine, peptide-like derivatives in aqueous fluid solution. *Photochem. Photobiol.*, **46**, 971–978.

Dietz,T.M. and Koch,T.H. (1991) Mechanistic studies toward an understanding of photocrosslinking of 5-bromouracil-containing nucleic acids to associated proteins. *Trends Photochem. Photobiol.*, **2**, 121–136.

Fersht,A.R. (1987) The hydrogen bond in molecular recognition. *Trends Biochem. Sci.*, **12**, 301–304.

Fersht,A.R., Shi,J.-P., Knill-Jones,J., Lowe,D.M., Wilkinson,A.J., Blow,D.M., Brick,P., Carter,P., Wayne,M.M.Y. and Winter,G. (1985) Hydrogen bonding and biological specificity analysed by protein engineering. *Nature*, **314**, 235–238.

Frey,P.A. and Sammons,R.D. (1985) Bond order and charge localization in nucleoside phosphorothioates. *Science*, **228**, 541–545.

Golmohammadi,R., Valegård,K., Fridborg,K. and Liljas,L. (1993) The refined structure of bacteriophage MS2 at 2.8 Å resolution. *J. Mol. Biol.*, **234**, 620–639.

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.

Jessen,T.H., Oubridge,C., Teo,C.H., Pritchard,C. and Nagai,K. (1991) Identification of molecular contacts between the U1 A small nuclear ribonucleoprotein and U1 RNA. *EMBO J.*, **10**, 3447–3456.

King,G.C. and Coleman,J.E. (1988) The Ff gene 5 protein-d(pA) 40–60 complex: 1H NMR supports a localized base-binding model. *Biochemistry*, **27**, 6947–6953.

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.

Lohman,T.M., deHaseth,P.L. and Record,M.T.,Jr (1980) Pentalysine-deoxyribonucleic acid interactions: a model for the general effects of ion concentrations on the interactions of proteins with nucleic acids. *Biochemistry*, **19**, 3522–3530.

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.

Mattaj,I.W. and Nagai,K. (1995) Recruiting proteins to the RNA world. *Nature Struct. Biol.*, **2**, 518–522.

Milligan,J.F. and Uhlenbeck,O.C. (1989) Determination of RNA–protein contacts using thiophosphate substitutions. *Biochemistry*, **28**, 2849–2855.

Montenay-Garestier,T. and Helene,C. (1971) Reflectance and luminescence studies of molecular complex formation between tryptophan and nucleic acid components in frozen aqueous solutions. *Biochemistry*, **10**, 300–306.

Ni,C.Z., Syed,R., Kodandapani,R., Wickersham,J., Peabody,D.S. and Ely,K.R. (1995) Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly. *Structure*, **3**, 255–263.

Oubridge,C., Ito,N., Evans,P.R., Teo,C.H. and Nagai,K. (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432–438.

Perona,J.J., Rould,M.A. and Steitz,T.A. (1993) Structural basis for transfer RNA aminoacylation by *Escherichia coli* glutaminyl-tRNA synthetase. *Biochemistry*, **32**, 8758–8771.

Shamoo,Y., Williams,K.R. and Konigsberg,W.H. (1988) Photochemical crosslinking of bacteriophage T4 single-stranded DNA binding protein (gp32) to oligo (dT)8: identification of phenylalanine-183 as the site of crosslinking. *Proteins: Struct. Funct. Genet.*, **4**, 1–6.

Shamoo,Y., Friedman,A.M., Parson,M.R., Konigsberg,W.H. and Steitz,T.A. (1995) Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature*, **376**, 362–366.

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.

Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, **185**, 60–89.

Stump,W.T. and Hall,K.B. (1995) Crosslinking of an iodo-uridine RNA hairpin to a single site on the human U1A N-terminal RNA binding domain. *RNA*, **1**, 55–63.

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.

Vershon,A.K., Youdeian,P., Weiss,M.A., Susskind,M.M. and Sauer,R.T. (1985) Mnt repressor-operator interactions: Altered specificity requires N-6 methylation of operator DNA. In Calendar,R. and Gold,L. (eds), *Sequence Specificity in Transcription and Translation.* A.R.Liss, New York, pp. 209–218.

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.

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.

Witherell,G.W. and Uhlenbeck,O.C. (1989) Specific RNA binding by Q beta coat protein. *Biochemistry*, **28**, 71–76.

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.

Wong,I. and Lohman,T. (1993) A double-filter method for nitrocellulose-filter binding: Application to protein-nucleic acid interactions. *Proc. Natl Acad. Sci. USA*, **90**, 5428–5432.

Received on July 19, 1996