2808–2815 *Nucleic Acids Research, 1997, Vol. 25, No. 14*

MS2 coat protein mutants which bind Qβ RNA

Marc Spingola and David S. Peabody*

Department of Cell Biology, University of New Mexico School of Medicine and Cancer Research and Treatment Center, Albuquerque, NM 87131, USA

Received March 27, 1997; Revised and Accepted May 21, 1997

ABSTRACT

The coat proteins of the RNA phages MS2 and Qβ are structurally homologous, yet they specifically bind different RNA structures. In an effort to identify the basis of RNA binding specificity we sought to isolate mutants that convert MS2 coat protein to the RNA binding specificity of Qβ. A library of mutations was created which selectively substitutes amino acids within the RNA binding site. Genetic selection for the ability to repress translation from the Qβ translational operator led to the isolation of several MS2 mutants that acquired binding activity for Qβ RNA. Some of these also had reduced abilities to repress translation from the MS2 translational operator. These changes in RNA binding specificity were the results of substitutions of amino acid residues 87 and 89. Additional codon-directed mutagenesis experiments confirmed earlier results showing that the identity of Asn87 is important for specific binding of MS2 RNA. Glu89, on the other hand, is not required for recognition of MS2 RNA, but prevents binding of Qβ RNA.

INTRODUCTION

The coat protein of the RNA bacteriophage MS2 is a translational repressor which prevents expression of the replicase cistron by binding an RNA stem–loop containing the replicase translation initiation site. It has been a particularly useful model for the study of protein–RNA interactions. Its RNA target has been extensively characterized (1,2), the X-ray crystal structures of coat protein and coat protein–RNA complex have been solved (3–5) and its translational repressor (RNA binding) function is amenable to detailed genetic analysis (6). Moreover, MS2 is only one member of a large class of related RNA phages. Their coat proteins show clear sequence relatedness and possess similar tertiary structures, even though they bind different RNAs. Since each of the coat protein variants represents a slightly different solution to the problem of specific RNA recognition, their existence presents an opportunity to explore the structural basis of RNA recognition.

Recently we reported a comparison of the coat proteins of phages MS2 and GA, which possess ~62% amino acid sequence identity and bind translational operators that are closely related but differ in the nucleotide sequence of the RNA loop. By introducing GA-like amino acid substitutions into the RNA binding site of MS2 coat protein we were able to confer on it an RNA binding specificity similar to that of GA, thus defining the protein structural determinants of RNA binding specificity in that case (7).

In the present study we sought to confer on MS2 coat protein the RNA binding specificity of a more distant relative. The coat protein of phage Qβ shows only ~23% amino acid sequence identity to the coat protein of MS2. The essential structural properties of both the MS2 and Qβ operator RNAs have been determined previously (1,8) and are illustrated in Figure 1. Although there are some common features, notably the presence of a base paired stem with a bulged adenosine, there are also significant differences, especially above the bulge, where the Qβ operator has a longer stem and a loop of only 3 nt. Moreover, the bulged adenosine residue, which is a crucial feature of the MS2 operator, is largely dispensable in Qβ. Presumably the differences in RNA bindng specificity of the two proteins are the consequences of amino acid substitutions within their RNA binding sites. In this paper we describe our efforts to convert the MS2 coat protein to the RNA binding specificity of Qβ and report that as few as one or two amino acid substitutions are sufficient to do so.

MATERIALS AND METHODS

Plasmid construction

We have previously described a two-plasmid genetic system in which MS2 coat protein expressed from one plasmid (pCT119) represses the translation of a replicase–β-galactosidase fusion protein encoded on a second plasmid (pRZ5) (6). In the original system translational repression was mediated by binding of coat protein to the translational operator of the MS2 replicase sequence. This makes it possible to screen for repressor function *in vivo* by colony color on LB plates containing the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). In order to study the translational repression and RNA binding properties of Qβ coat protein we created an analogous two-plasmid system in which Qβ sequences replace their MS2 homologs. Thus Qβ coat protein produced from the plasmid pQCT1 represses β-galactosidase expressed from plasmid pRZQ5. The details of these plasmid constructions are described elsewhere (9).

*To whom correspondence should be addressed. Tel: +1 505 272 0071; Fax: +1 505 272 9494; Email: dpeabody@cobra.unm.edu

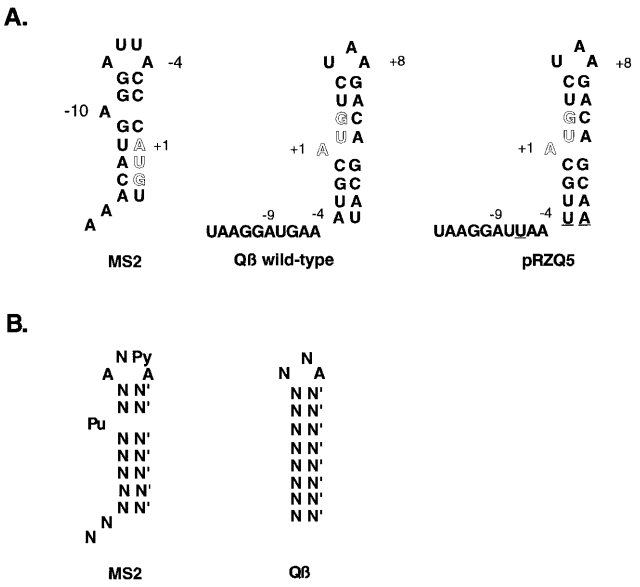

Figure 1. The translational operators of MS2 and Qβ. (**A**) The structures of the wild-type MS2 and Qβ operators. Also shown is the structure of the modified Qβ operator used in the construction of pRZQ5. Three base substitutions (underlined) were introduced to eliminate two AUG triplets while preserving the RNA secondary structure required for interaction with Qβ coat protein. (**B**) Summaries of the required structural features of the RNA targets of the MS2 and Qβ coat proteins. N refers to any of the four ribonucleotides, N′ to a complementary nucleotide and Pu and Py to purine and pyrimidine respectively.

Mutagenesis, selection and characterization of the mutants

To create a library of MS2 coat mutants containing a large number of amino acid substitutions in the RNA binding site three degenerate oligonucleotides were used as primers for mutagenesis of the coat gene on single-stranded pCT119-dlFG DNA (10) by the method of Kunkel *et al*. (11). The dlFG variant lacks a peptide loop essential for assembly of coat protein into capsids. The rationale behind its use is described in Results. The range of predicted amino acid substitutions generated in the mutagenesis is shown in Figure 2. After introduction of mutant DNAs into *Escherichia coli* ~5 000 000 independent transformants were obtained. These were divided into five pools and plasmid DNA was extracted by standard mini-preparation procedures. DNA from separate pools was then introduced by transformation into strain CSH41F$^-$ containing pRZQ5 and spread at a density of ~25 000 transformants/plate on M9 medium containing casamino acids and 0.2% lactose. This strain contains mutations rendering it both *lacZ*$^-$ and *galE*$^-$. Expression of β-galactosidase from pRZQ5 is lethal in *galE*$^-$ hosts, providing selection for translational repression (12). After 2 days growth at 42°C colonies were picked and streaked on LB plates containing X-gal. Plasmids were isolated from clones that gave rise to white colonies and these were subjected to DNA sequence analysis (13). Once the mutations were identified we took advantage of an appropriately located *Rsa*I site to transfer the mutations present in pCT119dlFG to the intact coat sequence in pCT119.

In some experiments mutagenesis was targeted to positions 87 and 89 using degenerate oligonucleotides capable of introducing the sequence NNG/C at the targeted codon. The resulting 32 possible triplets encode all 20 amino acids and one stop codon. These libraries of coat mutants were screened for translational repression in strain CSH41F$^-$ containing pRZ5 or pRZQ5 using colony color on LB plates containing X-gal. Plasmid DNA from selected colonies was subjected to DNA sequence analysis.

Mutant coat proteins were tested for their abilities to repress translation from the Qβ operator in pRZQ5 and from the MS2 operator in pRZ5 by comparing their abilities to inhibit β-galactosidase synthesis. Assay of β-galactosidase was by the method of Miller (14). Coat proteins were produced in *E.coli* and purified by a modification of methods we have described elsewhere (6). Cultures (500 ml) were grown in LB medium to saturation (overnight). Cells were pelleted by centrifugation and resuspended in 50 ml 100 mM NaCl, 50 mM Tris–HCl, pH 8.5, 10 mM EDTA, containing lysozyme at 2 mg/ml. After 60 min on ice, sodium deoxycholate was added to a concentration of 0.05%. The mixture was kept on ice for another 60 min and then sonicated to reduce viscosity. Polyethyleneimine was then added to a concentration of 0.2% and the lysate was held on ice for another 60 min. After centrifugation to remove the precipitate and cellular debris, ammonium sulfate was added to the supernatant to 50% saturation. The pellet was collected by centrifugation, dissolved in 100 mM NaCl, 10 mM Tris–HCl, pH 7.5, 0.1 mM $MgSO_4$, 0.01 mM EDTA and applied to Sepharose CL4B as described (6). Fractions containing coat protein were identified by polyacrylamide gel electrophoresis in SDS (15), pooled and concentrated in Centricon centrifugal concentrators. Capsids were disaggregated in 50% acetic acid and dialyzed against 10 mM acetic acid. This procedure yielded coat protein at purities estimated by gel electrophoresis to exceed 95%. Filter binding studies were conducted essentially as described by Carey *et al*. (16) using ^{32}P-labeled operator RNAs produced by run-off transcription *in vitro* from plasmids containing the appropriate operator sequences linked to a T7 promoter (17). The RNA binding curves we show were the results of experiments conducted in the TMK buffer of Carey *et al.* (16) under conditions favored by MS2 coat protein. Experiments were also performed in the MMK buffer preferred by Qβ coat protein (8), but under these conditions the MS2 coat protein and its variants bound poorly to both RNAs (not shown). Because different coat protein preparations sometimes differ in their fractional content of active protein, each preparation was also subjected to RNA-excess filter binding. RNA binding became saturated at values which indicated recoveries of active protein varying from 40 to 100% and these values were used to compute the concentrations of active protein in each sample (data not shown). The binding data shown in Figure 3 are the averages of two separate experiments. The curves were best fits to the data computed using Kaleidagraph (Abelbeck Software) and the equation $F = P/(K_d + P)$, where F is the fraction of RNA bound, K_d is the dissociation constant and P is the protein concentration.

RESULTS

Constructing a library of RNA binding site mutations

We have previously described a two-plasmid genetic system in which coat protein expressed from pCT119 represses translation of a replicase–β-galactosidase gene expressed from a second plasmid called pRZ5 (6). We wanted to create a large library of mutants with substitutions directed to binding site residues, since we intended eventually to screen such a library for a range of

A.

	βE				βF						βG				
	43	45	47	49	55	57	59	61	63	65	83	85	87	89	91
MS2	K	T	S	R	N	K	T	K	E	P	R	Y	N	E	T
GA	R	T	S	R	K	K	A	K	E	P	K	Y	S	D	T
Qß	R	T	S	S	R	N	K	Q	K	Q	Q	Y	D	T	S
SP	R	T	S	A	T	N	K	Q	K	Q	S	F	D	T	S
PP7	R	T	S	T	K	A	R	N	K	D	T	V	S	D	T

B.

43	**49**	**55**	**57**	**59**	**61**	**63**	**83**	**85**	**87**	**89**	**91**
K	R	N	K	T	K	E	R	Y	N	E	T
R	S	K	N	N	N	K	H	F	S	D	S
		S		K	Q		Q		D	A	
		R		S	H		S		G	K	
				R			K			N	
							N			T	

C.

	βE				βF						βG				
	43	45	47	49	55	57	59	61	63	65	83	85	87	89	91
MS2	K	T	S	R	N	K	T	K	E	P	R	Y	N	E	T
Qß (seq.)	R	T	S	S	**R**	**N**	**K**	**Q**	**K**	**Q**	Q	Y	D	T	S
Qß (str.)	R	T	S	S	N	K	Q	K	Q	P	Q	Y	D	T	S

Figure 2. (**A**) A comparison of the amino acid residues present on the putative RNA binding surfaces of five different RNA phage coat proteins based on the alignment presented in Liljas and Valegard (19). The numbering is that of MS2 coat protein and βE, βF and βG refer to the three β-strands that make up most of the RNA bindng site. (**B**) The diversity of RNA binding site amino acid sequences present in the mutational library used for isolation of MS2 coat mutants capable of binding the Qβ operator. (**C**) A comparison of the alignments of MS2 and Qβ RNA binding site residues based on sequence (seq.) and on structural considerations (str.). Incorrect alignments (in β-strand F) are shown in bold. See text for details.

novel RNA binding specificities. To do so it was desirable to create as large a library of sequence variants as possible without exceeding the capacities of conventional genetic and recombinant DNA methods to isolate and analyze them. Our previous mutational analysis of MS2 coat protein had implicated 10 amino acid residues on one surface of the coat protein β-sheet as constituents of the RNA binding site (18), but at least five more residues are present on the same β-sheet surface and these could become important for binding new RNAs. To completely randomize the amino acid composition of this β-sheet surface would require 20^{15} (or ~3.3×10^{19}) different amino acid sequences, clearly an unmanageable number. A comparison of the amino acids present at these 15 positions in coat proteins from five different RNA phages is shown in Figure 2A. Since these different coat proteins bind rather different RNA structures, it seemed that a fair range of specificities might be achieved by introducing into MS2 coat protein just the level of variation observed in Figure 2A. Accordingly, three degenerate oligonucleotides were used as primers in a site-directed mutagenesis reaction to create the amino acid sequence variation shown in Figure 2B. This should create ~3×10^6 different nucleotide sequences and ~730 000 different amino acid sequences, including all those shown in Figure 2B. These manipulations were carried out using plasmid pCT119-dlFG (10). This is a mutant version of pCT119 lacking sequences encoding the FG loop, a portion of the structure required for assembly of dimers into capsids (3). We had previously shown that mutations that confer certain assembly defects can cause increased repression without affecting RNA binding directly and wanted to avoid isolating more mutants of this type, even though they are probably very rare in our library since mutations were directed to the binding site. About 85% of the recombinants failed to repress the wild-type MS2 operator in pRZ5, indicating a high efficiency of mutagenesis.

It should be noted that the amino acid sequence comparisons shown in Figure 2A were based on a previously published alignment of the coat sequences (19). While this manuscript was in preparation the X-ray structure of Qβ coat protein became available (20). It shows that this alignment is in error in β-strand F. The correctness of the X-ray structure in this region was confirmed by our genetic studies of the RNA binding site of Qβ coat protein (9). The corrected alignment of amino acid sequences (see Fig. 2C) shows that the MS2 and Qβ RNA binding sites are somewhat more similar than Figure 2A indicates. Consequently, our mutant library contains more variation than is required to convert most of the MS2 binding site amino acids to their Qβ counterparts. However, this also means that two Qβ-like substitutions cannot be present in our library. We will describe later how we addressed this difficulty by the introduction of these specific substitutions.

Selection and characterization of mutant MS2 coat proteins able to bind the Qβ operator

The plasmids pQCT119, which contains the Qβ coat gene, and pRZQ5, which contains a fusion of the Qβ operator to the *lac*Z gene, were constructed by analogy to pCT119 and pRZ5 and are the components of a two-plasmid system for the genetic characterization of the RNA binding site of Qβ coat protein which we have described in detail elsewhere (9). Note that the pRZQ5 operator differs from the wild-type Qβ sequence in three positions (underlined in Fig. 1). These mutations were introduced to eliminate two extra AUG triplets. Their removal was necessary for efficient translational repression *in vivo*, but did not affect binding to Qβ coat protein *in vitro*. In Table 1 are shown the translational repressor activities of the various coat proteins described in this study. Note that each wild-type protein is a good repressor of translation from its cognate operator, but poorly represses translation from the non-cognate operator. In an attempt to isolate mutant MS2 coat proteins with the ability to bind the Qβ operator, the mutant library described above was introduced by transformation into strain CSH41F$^-$ containing pRZQ5 and plated on medium containing lactose. Strain CSH41F$^-$ is *galE*$^-$ and expression of β-galactosidase is lethal to such strains in the presence of lactose (12), providing a selection for the acquisition of translational repressor function. Clones which survived on lactose were streaked on LB medium containing X-gal and those yielding white colonies were picked for further analysis. This last step screens out reversions of the *galE* defect and any mutations in lactose permease. Nucleotide sequence analysis revealed the nature of the nucleotide and amino acid substitutions that characterize each of the mutants (Table 2). All the mutants possessed amino acid substitutions at residue 89. Many also had substitutions at position 87. Although some mutants also showed changes at positions 85 and/or 43, their translational repression behaviors seemed not to be affected significantly by these

additional changes (results not shown). For this reason, subsequent analyses focused on the three mutants called E89T, N87S-E89T and N87S-E89K.

Table 1. The translational repressor activities of the various mutants described in the text

Repressor	Fold repression	
	pRZ5	pRZQ5
MS2 (wild-type)	41	2
N87S	4	5
E89K	2	4
E89T	19	6
N87S-E89T	6	12
N87S-E89K	4	19
T59Q	29	2
E63Q	5	3
E89K-T59Q	3	4
E89K-E63Q	2	3
N87S-E89K-T59Q	1	4
N87S-E89K-E63Q	1	3
N87S-E89T-T59Q	2	25
N87S-E89T-E63Q	2	13
E89T-T59Q	8	4
E89T-E63Q	5	5
E89D	28	3
E89A	16	6
E89H	17	8
E89C	11	7
E89S	13	8
E89V	12	6
N87H	10	6
Qβ (wild-type)	1	28

Translational repression is expressed as the fold reduction in β-galactosidase activity produced from pRZ5 in strains containing the indicated repressors compared with the amount produced in the absence of repressor.

Table 2. The amino acid substitutions found in the MS2 coat mutants isolated for their abilities to repress translation from the Qβ translational operator of pRZQ5

Substitutions	No. of isolates
E89T	24
N87S-E89T	8
N87S-E89K	2
Y85F-E89T	1
Y85F-N87S-E89T	1
K43R-Y85F-N87S-E89T	1

Table 3. Dissociation constants (nM) for the interaction of MS2 and Qβ coat proteins and the various MS2 coat mutants with the MS2 and Qβ operators determined by filter binding

Repressor	Operator	
	MS2	Qβ
MS2	1.8 (0.2)	148
Qβ	>1000	170
E89K	2.3 (0.5)	4.1 (0.5)
E89T	0.7 0.04)	18.6 (7.9)
N87S-E89K	1.6 (0.7)	3.0 (0.6)
N87S-E89T	5.4 (3.4)	28.5 (0.7)

The numbers in parentheses are standard deviations.

Each of the mutations was transferred into pCT119 (i.e. non-dlFG) and the resulting plasmids were introduced into strain CSH41F$^-$ containing either pRZQ5 (Qβ operator) or pRZ5 (MS2 operator). Restoring the FG loop confers the capacity for virus assembly. This makes the proteins easy to purify (6) and eliminates the super-repressor activity that results from the assembly defect of the dlFG constructs. In previous work we had already isolated the N87S, N87L and E89K single substitutions; they were members of a collection of mutants defective for binding of the MS2 operator (18). We had also studied N87S and produced another residue 89 substitution (E89D) during a comparison of the RNA binding specificities of MS2 and GA coat proteins (7). We tested all of these mutants for their abilities to repress the MS2 and Qβ operators using the two-plasmid system by comparing the relative amounts of β-galactosidase activity produced. One mutant, E89D, retains the translational repression specificity of wild-type MS2 coat protein. The N87S mutant loses the ability to repress MS2, with some enhancement of repression of the Qβ operator. In other words, it is repressor-defective for both operators. E89K also loses the ability to tightly repress the MS2 operator and is somewhat improved in its repression of Qβ. Another variant, E89T, improves the ability to repress Qβ while retaining most of its activity for the MS2 operator. Two double mutants, N87S-E89K and N87S-E89T, switch specficities, so that they now show a preference for the Qβ operator.

RNA binding properties of the mutant proteins

The wild-type and four MS2 coat mutants, N87S-E89K, N87S-E89T, E89K and E89T, were subjected to *in vitro* analyses of their RNA binding activities using the protein-excess nitrocellulose filter binding assay described by Carey *et al.* (16). Dissociation constants for the various proteins were determined using both the MS2 and Qβ operator RNAs. Binding curves are shown in Figure 3 and the K_d values are summarized in Table 3. We previously described the RNA binding properties of N87S, which binds MS2 RNA with a K_d of 60 nM (7). It should be noted that the solution conditions employed in these experiments were those previously determined as optimal for the MS2 coat protein–RNA interaction. Qβ coat protein prefers somewhat different conditions (8) and, for that reason, binds even its own RNA relatively poorly under these conditions. Under its preferred conditions it exhibits a K_d of ~2 nM for the Qβ operator (see Fig. 3). Using the solution conditions favored by Qβ, all the MS2 coat protein variants bind both operators poorly (results not

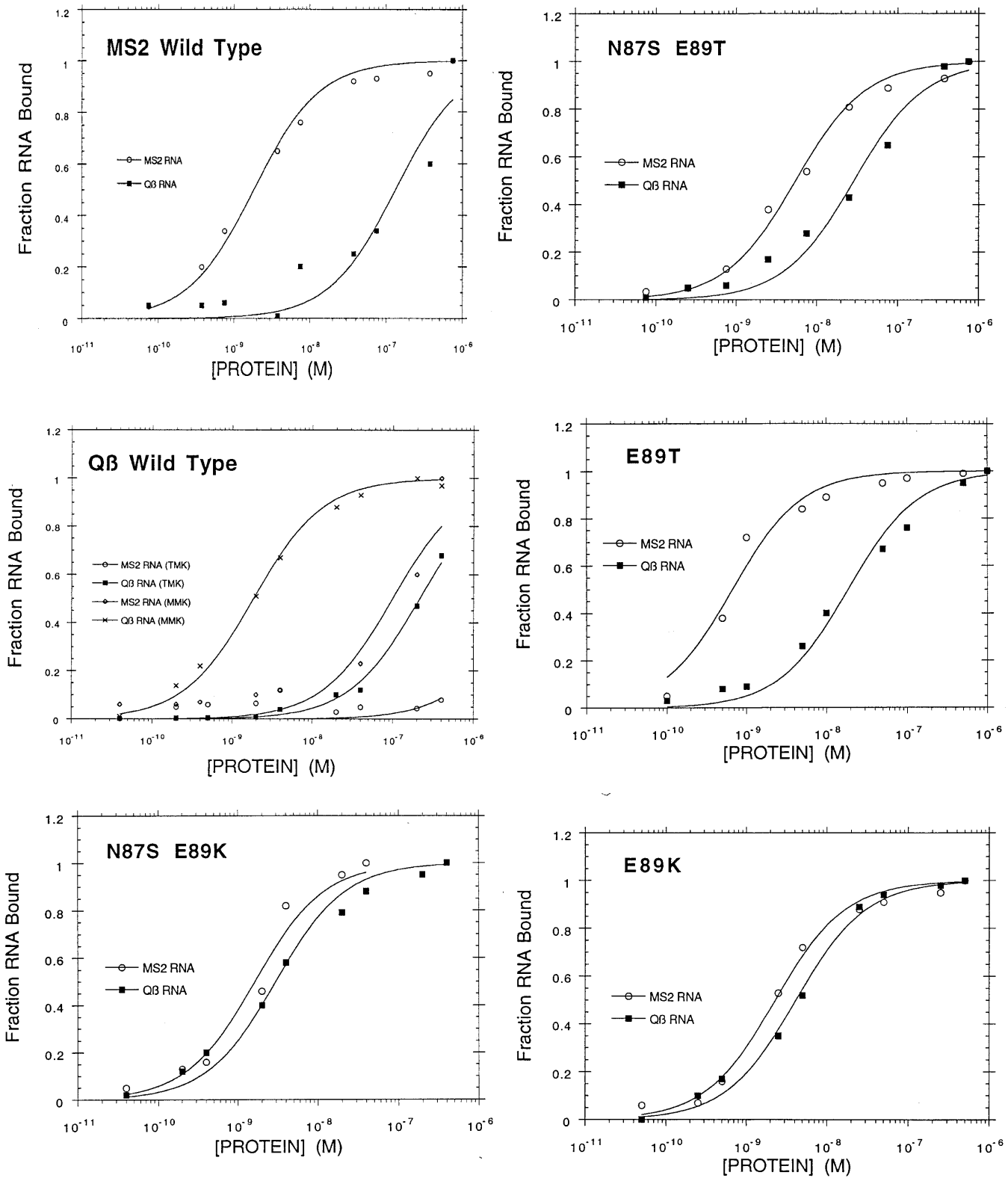

Figure 3. Binding of MS2 or Qβ operators by the wild-type and various mutant proteins as determined in a protein-excess nitrocellulose filter binding assay. The curves shown here and the dissociation constants reported in the text and in Table 3 are the averages of two experiments. Each panel shows the binding behavior of the indicated protein for the translational operators of MS2 and Qβ. All the proteins were bound to RNA in TMK buffer under conditions previously defined as optimal for RNA binding by MS2 coat protein (16). However, for comparison we also show the results of RNA binding by Qβ coat protein conducted in its preferred buffer, MMK (8), as indicated in the figure.

shown) and it is for this reason that MS2 conditions were employed. E89T and N87S-E89T each show significant improvements in binding to Qβ RNA. The biggest improvements in Qβ RNA binding, however, are seen with E89K and N87S-E89K, which bind Qβ RNA about as well as either MS2 or Qβ coat proteins bind their normal, cognate operators under their respective optimal conditions.

The *in vitro* binding behaviors generally agree with the *in vivo* translational repression experiments in that mutations which improve repression of the Qβ operator also improve binding of Qβ RNA *in vitro*. E89K, however, binds strongly to both operators *in vitro*, even though it scarcely represses the MS2 operator *in vivo*. We also note that although N87S improves the *in vivo* translational repressor activities of both E89T and E89K for the Qβ operator in the double mutants, it improves only the activity of E89K *in vitro*. We have observed small disparities between *in vivo* and *in vitro* behaviors of various mutants on other occasions (for example see 7). They must reflect differences in the conditions of the two binding reactions, possibly including pH, ionic strength and the presence *in vivo* of potential nucleic acid competitors.

Codon-directed mutagenesis of positions 87 and 89

What are the roles of the wild-type residues at positions 87 and 89 for translational repression of the MS2 translational operator? We previously showed that some substitutions at these sites render the coat protein defective for repression of the MS2 operator in pRZ5 (7,18), but we made no attempt to determine the range of acceptable substitutions. To this end we used codon-directed mutagenesis with degenerate oligonucleotide primers designed to introduce all possible amino acid substitutions at position 87 or 89. In each case the template for mutagenesis already contained the N87S or E89K repressor-defective mutations. The resulting mutant mixtures were introduced into strain CSH41F$^-$ (pRZ5) and plated on X-gal plates. White colonies represent reversions of the original repressor defects and reveal the alternative amino acid substitutions that permit binding of MS2 RNA. The results of DNA sequence analysis of 12 such position 87 revertants suggest that only asparagine (i.e. the wild-type residue) may be fully functional at this site for binding of the MS2 translational operator (Table 1). However, one partially functional isolate contained histidine at this position. In contrast, a wide range of amino acid substitutions of residue 89 yield repressor activities for the MS2 translational operator within a few-fold of wild-type. The conservative substitution of aspartic acid (E89D) was nearest to wild-type behavior. Among the 22 revertants we analyzed, seven different amino acid substitutions were found, including re-isolation of the already described E89T. Each substitution of position 89 led to improved repression of the Qβ operator in pRZQ5. The small effect of the aspartic acid substitution, however, suggests that the presence of an acidic residue at position 89 prevents tight binding of the Qβ operator.

Site-directed substitutions of residues 63 and 59

The design of our mutant library was based on amino acid sequence comparisons (19) which, at least in the case of Qβ coat protein, were partially in error. When the structure of Qβ coat protein became available we recognized that certain Qβ-like amino acid substitutions were not present in our library. In particular, residues T59 and E63 reside in positions where they could play a role in RNA binding, but, because of the design of the mutagenic oligonucleotides used in its construction, our library cannot contain the relevant T59Q and E63Q substitutions (Fig. 2B). For this reason, we introduced these mutations into the MS2 sequence, either alone or in various combinations with the N87S, E89K and E89T mutations, and measured their effects on translational repression *in vivo*. The results, summarized in Table 1, show that the T59Q substitution by itself has no effect on translational repression of the Qβ operator. Decreases in Qβ repression were generally observed when T59Q was combined with the other mutations. However, in combination with N87S-E89T, T59Q boosts repression to a level nearly as high as that achieved by Qβ coat protein itself. Although certain previously characterized substitutions of T59 (i.e. T59A and T59S) were strongly repressor defective (18), the T59Q substitution results in only a small reduction in MS2 repression. It usually also reduces repression in combination with other amino acid substitutions. The E63Q mutation, on the other hand, significantly reduces translational repression of the MS2 operator alone or in combination with the other amino acid substitutions. Meanwhile, E63Q has little or no effect on repression of the Qβ operator, except in combination with N87S-E89K, when repression is significantly reduced.

Unfortunately, we were unable to determine the affinities of these triple mutants for the Qβ and MS2 RNAs *in vitro*. We sometimes find that mutants containing multiple amino acid substitutions fail to refold properly after the acid denaturation step used in their purification (6). This can be the case even when the mutant proteins apparently fold properly *in vivo*, as judged by their abilities to assemble into capsids. This was apparently the case for N87S-E89T-T59Q and N87S-E89T-E63Q. RNA-excess filter binding experiments showed that neither protein exhibited much RNA binding activity after purification.

DISCUSSION

Mutations that converted MS2 coat protein to the RNA binding specificity of Qβ were readily isolated. All the original mutations substituted either a threonine (E89T) or lysine (E89K) for Glu89, indicating a key role for this amino acid residue in discriminating between the two operators. Note that the Qβ coat protein contains threonine at the position equivalent to residue E89 of MS2. Some mutants contained an additional substitution of serine for the asparagine normally present at position 87 (N87S). Extensive substitution of these same two sites by codon-directed mutagenesis indicated that a variety of amino acids could replace E89 of MS2 coat protein without much loss of binding activity for the MS2 operator. This makes sense, since in the structure of the coat protein–RNA complex E89 does not directly contact RNA. It should be noted, however, that the identity of residue 89 is not entirely irrelevant to the binding of MS2 RNA. Although their effects were modest, most substitutions reduced translational repression by a few-fold. Of the substitutions we characterized, the conservative replacement of E89 with aspartic acid (E89D) led to the smallest decrease in repressor activity for the MS2 operator. It repressed nearly as well as the wild-type, indicating that an acidic residue at this site favors MS2 RNA binding. At the same time, substitutions of residue 89 generally increased binding activity for Qβ RNA (Table 1). E89 seems to allow discrimination against the Qβ operator. This function apparently requires an acidic amino acid, since E89D shows a barely detectable increase

in repressor activity for Qβ RNA. We do not mean to imply that discrimination against Qβ RNA is a normal function of MS2 coat protein. Presumably, mixed infections in nature are rare. However, E89 may play a role in preventing the binding of non-specific hairpins in cellular or viral RNAs. Note that each of the mutants with E89 substitutions binds MS2 operator RNA *in vitro* with wild-type or higher affinity, yet many E89 substitutions are somewhat less effective repressors *in vivo*. This is consistent with the idea that other RNAs may compete with the operator for the mutant coat proteins.

The identity of residue 87 is strongly important for recognition of MS2 RNA. Our results suggest that probably only asparagine is fully functional at this site, although histidine (E89H) is able to partially fulfil this role. This is consistent with the structure of the coat protein–RNA complex, where a hydrogen bond is observed between the amide nitrogen of the asparagine side chain and an oxygen atom on U –5 in the translational operator (4). The presence in histidine of a similarly positioned hydrogen bond donor may explain its ability to partially replace asparagine. In experiments reported elsewhere we have conferred to Qβ coat protein the ability to bind MS2 RNA by changing Asp91 (the homolog of Asn87 in MS2) to asparagine (9), thus confirming the important role of this amino acid in recognition of the MS2 operator.

Two additional sites that might have played a role in acquisition of Qβ RNA binding specificity, T59 and E63, were not mutated to their Qβ counterparts in our mutational library. Therefore, we created these substitutions, T59Q and E63Q, by site-directed mutagenesis. The effects of the substitutions on MS2 RNA binding were surprising. Previously we found that T59S and T59A substitutions result in repressor-defective phenotypes (7,18), yet T59Q is only slightly reduced in its MS2 repressor activity. It is not obvious to us why this substitution is tolerated. In combination with the E89K, E89T, E89K-N87S and E89T-N87S mutations T59Q generally also led to small effects (usually reductions) of MS2 repression. The T59Q substitution usually also had negligible effects on repression of the Qβ operator, whether alone or in combination with the other mutations, except in the case of T59Q-N87S-E89T. This triple mutant was the best MS2 mutant repressor of Qβ we found, repressing nearly as well as Qβ coat protein itself.

The E63Q substitution results in a repressor defect for the MS2 operator, whether alone or in combination with the E89K, E89T, E89T-N87S and E89K-N87S mutations. Extensive random mutagenesis (18) and direct structural analyses (4) initially failed to identify residue 63 as a component of the RNA binding site, but a recent report suggests a possible role of the E63 side chain in forming a hydrogen bond with the 2′-hydroxyl of the uridine at position –5 in the operator loop (21). This assertion is consistent with our results. On the other hand, the E63Q substitution had little effect on repression of the Qβ operator, except in the case of its combination with N87S-E89K, where it led to a significant loss of repressor activity compared with the double mutant itself. These results emphasize that the effects of a given substitution can be highly dependent on their context.

We previously reported a comparison of the RNA binding sites of the coat proteins of MS2 and GA (7). They bind highly similar RNAs, the important difference being the substitution of a single nucleotide (U –5 in the loop; see Fig. 1). Thus it is not surprising that it was a relatively simple matter to confer the RNA binding specificity of GA on the MS2 coat protein. In contrast, Qβ coat protein binds an RNA possessing less obvious similarity to the operators of MS2 and GA. Although all three operator RNAs can be described as stem–loops with bulged adenosines, Qβ differs from the other two in the length of the stem, the size of the loop and in the relative unimportance of its bulged nucleotide (8). It is striking, therefore, that MS2 coat protein is so easily modified to bind Qβ RNA. Moreover, we have shown that single amino acid substitutions can endow Qβ coat protein with the ability to bind MS2 operator RNA (9). The ready interconvertability of the RNA binding specificities of Qβ and MS2 coat proteins must reflect similarities in the structures of their RNA binding sites. We previously reported the genetic identification of MS2 RNA binding site amino acids (18) and recently completed a similar analysis of amino acids required for operator recognition by Qβ coat protein (9). In Figure 2C we have compared the amino acids present at 15 structurally equivalent positions on the β-sheets of the MS2 and Qβ coat proteins. Although eight of the 15 are not conserved, many of the residues actually required by each coat protein for recognition of its respective operator are, in fact, identical or conservatively substituted. Thus each of these proteins is only one or two amino acid substitutions removed from acquisition of the alternative RNA binding specificity.

The similarity of the two sites implies that certain aspects of the interactions of the two coat proteins with their respective operators may be similar, despite the apparent structural dissimilarities of their operator RNAs. The X-ray structure of the MS2 coat protein–operator RNA complex reveals that the adenosines at –4 and –10 (Fig. 1) interact in a pseudo-symmetrical manner with the symmetrical coat protein dimer. A binding site for each adenosine is formed on different halves of the dimer by residues V29, T45, S47 and K61 (4). Identical amino acid residues are found in the homologous positions of Qβ coat protein, indicating conservation of the adenosine binding sites and raising the possibility that the essential adenosine in the loop of the Qβ operator may participate in a similar interaction with its coat protein. However, given the relative indifference of Qβ coat protein to the bulged adenosine, the pseudo-symmetry, which is an essential feature of the MS2 coat protein–RNA interaction, must be broken in Qβ. Additional interactions, probably with the stem, must compensate for the loss of this contact. Work currently in progress will determine whether this particular mode of interaction with the adenosine residue in the operator loop is a conserved feature of both coat protein–RNA complexes.

One wonders what would be the effects of these specificity mutations on virus viability if they were introduced into the viral genome. It is commonly assumed that the specific interaction of coat protein with the translational operator is important for virus viability, because of its apparent role in genome encapsidation. On this basis it would be predicted that the specificity changes we observe *in vivo* should require compensatory mutations in the translational operator. However, we have recently reported experiments showing that the interaction of coat protein with the operator RNA is not required for virus viability; mutations that inactivate the translational operator do not result in non-viability of the virus (22). On the other hand, genomes containing repressor-defective coat mutations cannot form plaques. This suggests the possibility that coat protein interacts with other sites to accomplish genome encapsidation. Since nothing is known of these other interactions, the effects on the virus life cycle of specificity changes in coat protein cannot be predicted with confidence.

ACKNOWLEDGEMENT

This work was supported by a grant from the National Institutes of Health.

REFERENCES

1 Romaniuk,P.J., Lowary.P., Wu,H., Stormo,G. and Uhlenbeck,O.C. (1987) *Biochemistry*, **26**, 1563–1568.
2 Wu,H.-N. and Uhlenbeck,O.C. (1987) *Biochemistry*, **26**, 8221–8227.
3 Golmohammadi,R., Valegard,K., Fridborg,K. and Liljas,L. (1993) *J. Mol. Biol.*, **234**, 620–639.
4 Valegard,K., Murray,J.B., Stockley,P.G., Stonehouse,N. and Liljas,L. (1994) *Nature*, **371**, 623–625.
5 Ni,C.-Z, Syed,R., Kodandapani,R., Wickersham,J., Peabody,D.S. and Ely,K.R. (1995) *Structure*, **3**, 255–263.
6 Peabody,D.S. (1990) *J. Biol. Chem.*, **265**, 5684–5689.
7 Lim,F., Spingola,M. and Peabody,D.S. (1994) *J. Biol. Chem.*, **269**, 9006–9010.
8 Witherell,G.W. and Uhlenbeck,O.C. (1989) *Biochemistry*, **28**, 71–76.
9 Lim,F., Spingola,M. and Peabody,D.S. (1996) *J. Biol. Chem.*, **271**, 31839–31845.
10 Peabody,D.S. and Ely,K.R. (1992) *Nucleic Acids Res.*, **20**, 1649–1655.
11 Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) *Methods Enzymol.*, **154**, 367–382.
12 Malamy,M.H. (1966) *Cold Spring Harbor Symp. Quant. Biol.*, **31**, 189–201.
13 Sanger,F., Nicklen,S. and Coulson,A. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5468.
14 Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15 Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
16 Carey,J., Cameron,V., de Haseth,P.L. and Uhlenbeck,O.C. (1983) *Biochemistry*, **22**, 2601–2610.
17 DraperD.E., White,S.A. and Kean,J.M. (1988) *Methods Enzymol.*, **164**, 221–237.
18 Peabody,D.S. (1993) *EMBO J.*, **12**, 595–600.
19 Liljas,L and Valegard,K (1990) *Semin. Virol.*, **1**, 467–475.
20 Golmohammadi,R., Fridborg,K., Bundule,M., Valegard,K. and Liljas,L. (1996) *Structure*, **4**, 543–554.
21 Baidya,N. and Uhlenbeck,O.C. (1995) *Biochemistry*, **34**, 12363–12368.
22 Peabody,D.S. (1997) *Mol. Gen. Genet.*, in press.