Nucleic Acids Research, 1998, Vol. 26, No. 5 **1337–1344**

Dissecting the key recognition features of the MS2 bacteriophage translational repression complex

Hugo Lago, Stephanie A. Fonseca, James B. Murray+, Nicola J. Stonehouse and Peter G. Stockley*

School of Biology, University of Leeds, Leeds LS2 9JT, UK

Received September 30, 1997; Revised and Accepted January 14, 1998

ABSTRACT

The MS2 RNA operator capsid offers an unparalleled opportunity to study sequence-specific protein–protein and RNA–protein interactions in molecular detail. RNA molecules encompassing the minimal translational operator recognition elements can be soaked into crystals of RNA-free coat protein shells, allowing the RNA to access the interior of the capsids and make contact with the operator binding sites. Correct interpretation of these structural studies depends critically on functional analysis in solution to confirm that the interactions seen in the crystal are not an artefact of the unusual approach used to generate the RNA–protein complexes. Here we present a series of *in vivo* and *in vitro* functional assays, using coat proteins carrying single amino acid substitutions at residues which either interact with the operator RNA or are involved in stabilizing the conformation of the FG loop, the site of the major quasi-equivalent conformational change. Variant operator RNAs have been assayed for coat protein affinity *in vitro*. The results reveal the robustness of the operator–coat protein interaction and the requirement for both halves of a protein dimer to contact RNA in order to achieve tight binding. They also suggest that there may be a direct link between the conformation of the FG loop and RNA binding.

INTRODUCTION

The translational repression complex which forms between bacteriophage MS2 coat protein (and its close relatives) and an RNA stem–loop of 19 nt has long been the paradigm for sequence-specific recognition of simple RNA target sites (1,2; Fig. 1). Our understanding of this interaction at the molecular level has recently made major advances due to the determination by X-ray crystallography of three-dimensional structures for the wild-type $T = 3$ phage particle (3,4), RNA-free capsids, produced by over-expression of a recombinant coat protein gene in the absence of genomic phage RNA (5), a non-assembling coat protein dimer (6) and, most importantly, the structures of several operator capsids in which the 19 nt operator fragment has been soaked into crystals of RNA-free capsids, where it makes sequence-specific contacts to coat protein dimers in the structure (7,8). Solution studies with chemically variant oligonucleotides, in which specific functional groups involved in the RNA–protein interaction have been substituted, leave little doubt that the complexes seen in the capsid crystals are representative of the repression complex which forms initially in solution between a coat protein dimer and the RNA operator (1,9,10). There is also a very good correlation between the intermolecular contacts seen in the operator capsid crystals and coat protein mutations which give rise to reduced RNA binding (2,11,12).

The canonical operator fragment consists of a 7 bp stem, interrupted by a bulged adenosine at position –10 (see Fig. 1 for numbering system), closed at one end by a 4 base single-stranded tetraloop. Extensive sequence variation experiments showed that the identity of only four nucleotide positions of the 19mer are important for coat protein recognition (1). These are adenosines at positions –4, –7 and –10 and a pyrimidine at –5. Interaction with the coat protein is preserved provided that the Watson–Crick base pairing of the stem is maintained and this recognition is insensitive to the identity of the base at –6. Solution studies suggested that all three important adenosines were recognized differentially (9,13) and this has been confirmed by the X-ray crystal structure of the operator complex (7,8). Within the capsid each dimer binds one operator fragment. In the quasi-equivalent protein shell there are two distinct types of dimer, A/B and C/C. Operators bind in either orientation to C/C dimers, however, the orientation at A/B dimers is unique and allows the molecular details of the interaction to be modelled in the resulting electron density maps (Figs 2 and 3). Adenosine residues at –4 and –10 interact with the protein in hydrophobic pockets in a roughly two-fold symmetrical arrangement across each coat protein dimer. The orientation of the base and the functional groups contacted at each site differ. Remarkably, the base at –7 is not contacted directly but is involved in an extended stacking interaction with the pyrimidine at –5, which in turn is stacked against Tyr85 of the protein. Solution structures of the RNA operator calculated from NMR data (14; Thompson *et al.*, unpublished results) suggest that the species bound by the protein is at best only a minor component of a set of conformers in equilibrium, confirming earlier predictions based on chemical reactivity of the various functional groups (13). It is clear that

*To whom correspondence should be addressed. Tel: +44 113 233 3092; Fax: +44 113 233 2835; Email: stockley@bmb.leeds.ac.uk

+Present address: Department of Chemistry, Indiana University, Bloomington, IN 47405, USA

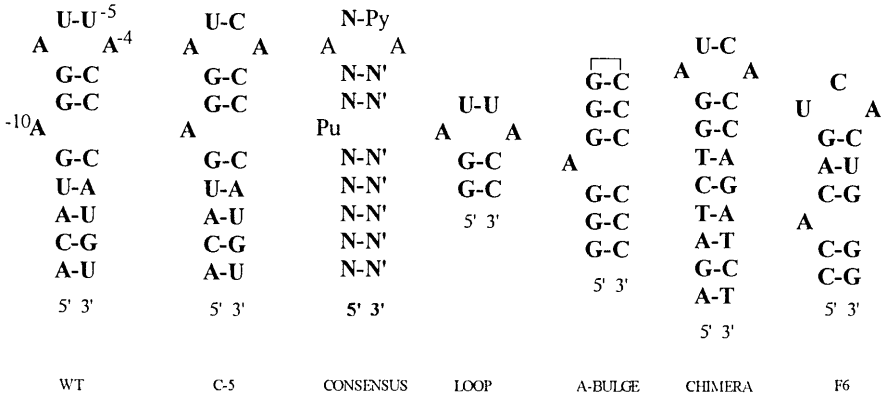

Figure 1. Structures of the RNA operator fragments used. Secondary structures of the MS2 operator sequences. The consensus is shown, together with sequences for the wild-type (WT), the C –5 variant, the loop, the A bulge (showing the position of the alkyl linker ∩), the RNA loop–DNA stem chimera and the F6 aptamer. Numbering is relative to the start of the replicase cistron for MS2. Py, pyrimidine; Pu, purine; N-N′, any complementary Watson–Crick base pair.

Figure 2. Locations of residues in the operator binding site and the FG loop. Schematic drawing of an AB coat protein dimer, illustrating residues involved in RNA binding, the positions of mutations within the FG loop and the side chains they interact with and the positions of the tryptophan residue (Trp82) believed to contribute to the change in intrinsic fluorescence upon RNA binding. The A subunit is in yellow and B in green.

complex formation with coat protein must involve conformational changes from the dominant form of the RNA fragment in solution in which A –10 is intercalated in the base paired stem.

Despite this wealth of structural and functional data, a number of key questions remain to be answered. In the operator capsid both coat protein and RNA components appear to have undergone significant conformational changes from their solution forms. Precisely how these changes occur and what drives self-assembly of the $T = 3$ phage shell is still unknown. The contributions of each interacting partner to the overall stability of the complex is also unknown.

Here we correlate structural and functional studies and report the results of *in vivo* and *in vitro* analyses with RNA operator variants and coat protein mutants. The overall aim is to dissect the important interactions responsible for stability of the operator complex and for triggering self-assembly of the phage shell (15).

Two protein mutants carrying single amino acid substitutions, Thr45Ala and Thr59Ser (Fig. 2), have been studied because the side chains of these residues form part of the sequence-specific adenosine binding pockets. Thr45 is universally conserved amongst RNA bacteriophage coat protein sequences (1), consistent with its roles as a hydrogen bonding partner to the N1 and N6

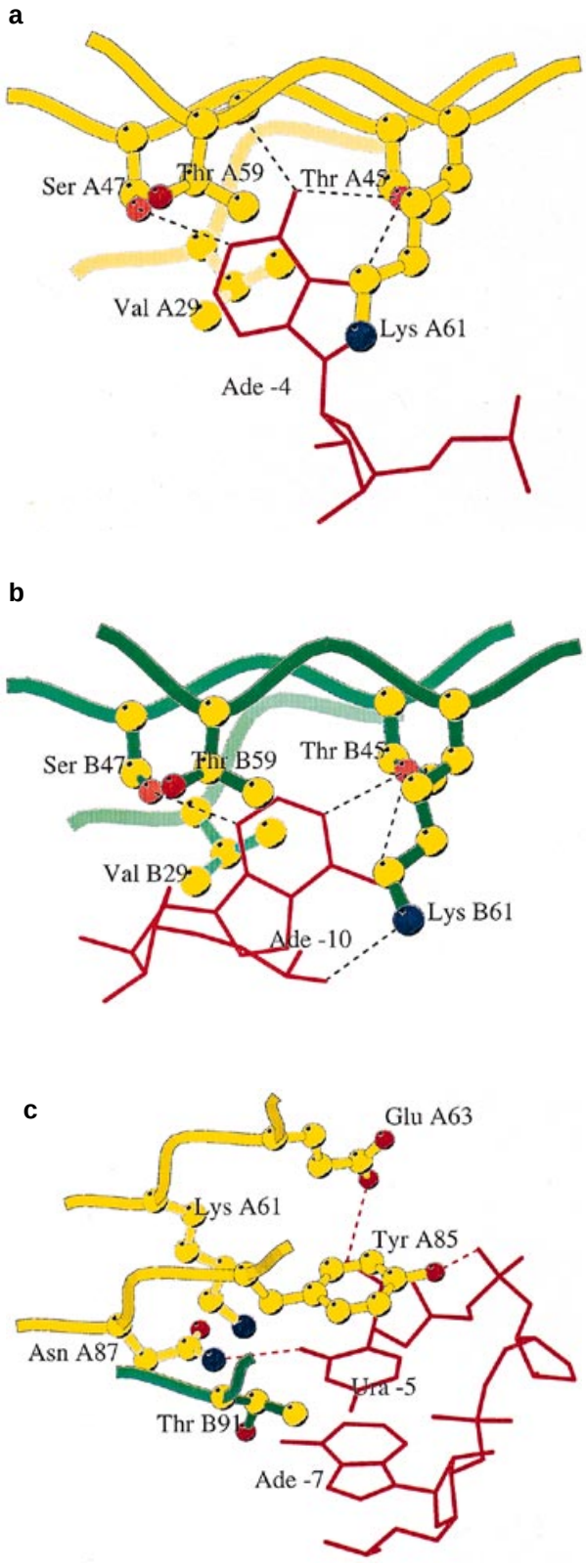

Figure 3. The base recognition pockets. Diagram showing details of the RNA–protein interactions at the two adenosine binding pockets, (**a**) –4 and (**b**) –10, and (**c**) the interaction at –5. The RNA is shown as a stick model (red); the protein chain is coloured as in Figure 2; hydrogen bonds are shown as dashed lines.

positions of A –10 and the N6 and N7 positions of A –4. Thr59 is less well conserved, its side chain, together with that of Lys61, forming one of the walls of the adenosine binding pockets. It is believed to make primarily hydrophobic contacts with operator RNA. The X-ray crystal structures of these two mutant capsids, with or without operator RNA, are reported by van den Worm *et al.* in this issue (16). In these structures the mutant proteins still bind operator RNA in a very similar fashion to wild-type, with only minor differences in the orientations of the adenines at –4 and –10, despite, at least for the Thr45Ala variant, a significant reduction in the number of intermolecular hydrogen bonds.

Two further singly substituted mutant proteins, Glu76Asp and Pro78Asn (Fig. 2), were studied to probe the mechanism of self-assembly and quasi-equivalent conformational switching. Glu76 and Pro78 are distal to the RNA binding site, lying in the FG loop, connecting the F and G β-strands. This is the site of the major conformational difference between quasi-equivalent subunits (4), resulting in three types of conformer in the capsid; A, B and C. In B-type conformers Glu76 forms a hydrogen bond to Gln40, locking the FG loop into a folded back position. This conformation is associated with a *cis* peptide linkage at Pro78, the equivalent peptides in A and C subunits being *trans*. *Cis–trans* isomerization and conformational switching at the FG loop have been proposed as the origin of quasi-equivalence in this system, consistent with the absolute conservation of proline at position 78 in available protein sequences (4) . Recent results with a deletion mutant of the related phage fr, however, show that it is possible to generate a $T = 3$ shell with coat protein subunits lacking an extended FG loop region (Axblom *et al.*, personal communication), although in the wild-type phage the FG loop is almost identical in conformation to the MS2 loop (17). The two MS2 FG loop variants, Glu76Asp and Pro78Asn, both still form $T = 3$ shells and the X-ray crystal structures of their RNA-free capsids have been reported (5). The Glu76Asp mutation is silent in terms of phage morphology, since the reduced length of the Asp side chain is compensated for by insertion of a bridging water molecule. The Pro78Asn mutation also does not affect the organization of the phage capsid, although all Asn linkages are *trans*. An infectious clone has been used previously to examine the phenotype of both mutants *in vivo*. The Glu76Asp substitution is viable, whereas the Pro78Asn is not. We have proposed that Pro78 may play a role in binding the essential maturation protein, the A protein, to the final capsid, which would be consistent with these data (18).

Finally, we have used the structural information to try and identify the key recognition features of the RNA operator. The interaction of the stem–loop with a coat protein dimer immediately suggests two operator 'half-sites'; one consisting of the residues of the tetraloop, which interacts with A-type subunits, and the other consisting of the bulged adenosine in the stem, which contacts B subunits. We have determined the *in vitro* coat protein affinities of a series of operator variants (Fig. 1) based on these ideas using a sensitive fluorescence titration assay. We have also compared coat protein binding to an RNA aptamer of the natural operator, which lacks the tetraloop and has a spacing of 3 bp between a 3 base loop and a bulged adenosine (19). In the context of a tetraloop this spacing is known to be extremely deleterious for coat protein recognition (1). The X-ray crystal structure of the aptamer–capsid complex has also recently been determined (20).

MATERIALS AND METHODS

Purification/synthesis of proteins and RNAs

Proteins were over-expressed and purified as RNA-free capsids as described previously (21,22). Protein concentrations were measured by UV spectrophotometry, using extinction coefficients at A_{280} estimated according to Gill and von Hippel (23). Disassembled coat protein was prepared by acetic acid extraction according to Sugiyama *et al.* (24) and kept on ice in 20 mM acetic acid, pH ≈ 3.4, until use. The RNA operator fragments were synthesized and purified as described previously (25,26). The triethyleneglycol phosphoramidite was purchased from Glen Research. RNA concentrations were determined from the absorbance at 260 nm using calculated extinction coefficients based on known values for mono- and dinucleotides as described (27).

Capsid thermal denaturation assays

Thermal denaturation assays were carried out as described previously (22). Briefly, 1 ml samples of capsids at ~1 mg/ml were dialysed at 4°C against 10 mM HEPES, 100 mM NaCl, 1 mM EDTA and 1 mM DTT, pH 7.2. Aliquots of 100 µl in Eppendorf vials were then incubated at defined temperatures in a heating block for 10 min prior to immediate processing for TEM. Samples were subjected to negative staining using 4% (w/v) uranyl acetate on formvar/carbon-coated copper grids, which were examined at 100 keV at a magnification of 50 000×.

In vivo superinfection assays

The mutant proteins were assessed for their ability to bind the wild-type translational operator *in vivo*. Mutant genes were cloned into the leaky expression plasmid ptachACP′ (28) and used to transform the piliated *Escherichia coli* strain EMG23 to ampicillin resistance (29). Using the soft agar overlay technique, 200 µl cells with an OD_{600} of 0.8–1.2 were plated with 100 µl wild-type phage (5.8×10^{11} p.f.u./ml at dilutions of from 10^{-3} to 10^{-9}) onto tryptone plates (containing 1 mM $CaCl_2$, 5 mM glucose and 100 µg/ml ampicillin). EMG23 and wild-type MS2 phage were purchased from ATCC (Rockville, MD). Proteins still able to bind the operator repress translation of the incoming replicase cistron and thus prevent replication of the phage and hence plaque formation. The numbers of plaques formed at different phage titres can therefore be used to assess operator affinity *in vivo* (28).

Fluorescence binding assays of RNA–protein complex formation

Fluorescence measurements were made on a Perkin-Elmer LS50B fluorimeter with spectral band widths of 3.5 (excitation) and 8.0 nm (emission), at excitation and emission wavelengths of 295 and 340 nm respectively. A water bath was used to keep the temperature constant at 25 ± 1°C in the cuvette during experiments, except for loop titrations, which were performed at 10°C (due to the low loop T_m). The temperature was measured directly using a thermocouple probe.

The RNA/coat protein titrations were performed as follows. A sample of coat protein in 20 mM acetic acid was diluted to the desired concentration in TMK buffer (0.1 M Tris–HCl, pH 7.5, 10 mM $MgCl_2$ and 80 mM KCl; 15), placed in a stirred cuvette and allowed to equilibrate for 30 min prior to addition of RNA, also in TMK. RNA was added in small aliquots, each addition being followed by a 5 min incubation to ensure equilibration prior to making the measurement. Twelve single measurements with an integration time of 10 s were averaged for each aliquot. The total volume added to the cuvette was ≤5% of the initial volume. Measurements were corrected for buffer effects and when necessary for inner filter effects according to Carpenter and Kneale (30). Quenching of tryptophan by photobleaching was controlled by shutting off the excitation beam between measurements. Addition of RNA to saturating levels resulted in a decrease of ~30% in fluorescence at 340 nm.

In vitro operator affinities

Binding data from the fluorescence titrations were fitted to a simple binding model (see equation below; 31) using the program SigmaPlot (Jandel Scientific).

$$F = \{(L_t + E_t + K_d) - [(L_t + E_t + K_d)^2 - 4 \times L_t \times E_t]^{0.5}/(2 \times E_t)\} \times (F_e - F_s) + F_s$$

where F is the measured fluorescence intensity at a given concentration of RNA, F_s is the fluorescence intensity of the coat protein in the absence of RNA, F_e is the fluorescence intensity at the saturation point, L_t is the total concentration of RNA present in the cuvette for each measurement, E_t is the total concentration of binding sites present in the cuvette and K_d is the apparent equilibrium dissociation constant.

The binding data fitted well to this model with correlation coefficients in all cases >98%, suggesting that the dominant complex being formed was of a coat protein dimer bound to a RNA operator fragment.

RESULTS AND DISCUSSION

Functional assays of complex formation

Three assays have been used here to examine the effects of the mutations. Effects on capsid stability have been assessed by observing thermal disruption of the $T = 3$ shells directly by TEM of negatively stained samples, as described previously (22). Effects on wild-type operator RNA binding have been studied *in vivo* using a superinfection assay (2,28), in which wild-type phage are used to infect cells carrying a leaky expression plasmid (ptachACP′) for the mutant protein of interest. If the mutant protein is stable and can still bind the operator fragment it represses translation of the incoming phage replicase gene, preventing formation of plaques, the appearance of which can be used as an indication of a defect in operator binding. Binding to operator fragments *in vitro* has also been monitored, using the intrinsic fluorescence of the protein as a probe of RNA binding, as described previously (15). This is an equilibrium assay and is therefore more accurate than the nitrocelluose filter binding assays traditionally used in this system (9,32,33). Interactions between molecules lacking some of the essential recognition features of the wild-type species are likely to result in complexes with shorter half-lives, which are not suited to techniques involving separation of equilibrium mixtures. The details of these assays and their results are listed in Tables 1 and 2 and Figure 4.

Table 1. *In vivo* superinfection and thermal denaturation assays

Coat protein	Denaturation temperature (°C)	*In vivo* phenotype	
		10^{-7}	10^{-9}
Wild-type	69	–	–
Thr45Ala	69	+	+
Thr59Ser	71	+	+
Glu76Asp	66	–	–
Pro78Asn	66	–	–

+, plaques/confluent lysis on agar plates; –, no plaques, binds operator RNA like wild-type.

Mutant coat proteins were assessed for their ability to bind the operator *in vivo* and for their thermal stability *in vitro*. Plaque formation by wild-type MS2 phage on cells expressing the mutant proteins was used to assess operator affinity *in vivo*. The presence of ~250 Å diameter $T = 3$ shells in TEM images was used to estimate the denaturation temperatures of RNA-free capsids. Denatured samples appeared as amorphous aggregates from which it was not possible to reform capsids by cooling below the denaturation temperature.

Mutants in the RNA binding site

The Thr45Ala mutant capsid has a wild-type thermal stability, as expected, since mutations in the RNA binding site would not be expected to affect capsid stability. *In vivo* Thr45Ala cannot repress replicase translation and *in vitro* its affinity for operator RNA is reduced by a factor of 5. These results correlate well with the lost hydrogen bonds from the operator capsid structure reported in this issue (16), a total of five hydrogen bonds per operator–coat protein dimer complex having been lost and a less favourable salt link between the phosphate oxygen O2P of G –11 and LysB61 being generated (Fig. 3).

Interpretation of the Thr59Ser data is not so clear cut. As for Thr45Ala, this mutant fails to repress *in vivo*. *In vitro*, however, its affinity for operator RNA is only very slightly reduced, consistent with a reduced hydrophobic interaction due to the loss of 25 Å^2, out of a total interaction surface of 700 Å^2 on the coat protein, contributed by the Thr59 methyl groups (16). The magnitude of this effect *in vivo* is perhaps surprising, but is probably due to the fact that the assay measures competition between coat protein and ribosome for operator binding, rather than simply the intrinsic RNA–protein affinity. Surprisingly, the Thr59Ser mutant capsid has a slightly increased denaturation temperature. The crystal structure of the RNA-free capsid suggests a possible explanation for this. The smaller serine side chain at residue 59 allows the side chain of GluA89 to re-orientate and move away from the main chain of Lys57, facilitating interactions with ArgB49 and SerB59. Such an increase in electrostatic interaction between residues on different strands of the extended β-sheet of the subunit would be consistent with the increased thermal stability observed.

The FG loop mutants

From the structural data (5) it appears that the reduced thermal stability of both Pro78Asn and Glu76Asp could be due to the absence of the same hydrogen bond, namely that between the Oε1 position of the side chain of GluB76 and the Nε2 group on GlnB40 (Fig. 2). This is lost completely in the Pro78Asn mutant, where minor adjustments to the polypeptide conformation adjacent to the site of mutation lead to a large (5 Å) shift in the position of the carboxyl group of Glu76. In the Glu76Asp mutant the contact to Gln40 is, however, maintained via a bridging water molecule. Glu76 also stacks on TrpB82, an interaction which is lost completely in the Pro78Asn mutant, but does occur with the shorter aspartate side chain in Glu76Asp molecules. The water-mediated contact is probably easily lost at elevated temperatures, resulting in similar overall reductions in thermal stability for both mutant capsids.

Table 2. *In vitro* operator binding data

Coat protein	K_d (nM)	*n*	Relative affinity (% wild-type)	E_t (μM)
Wild-type	39 (± 5)	4	100	0.34 (± 0.03)
Protein mutants				
Thr45Ala	190 (± 29)	3	20	–
Thr59Ser	49 (± 7)	2	79	0.28 (± 0.03)
Pro78Asn	23 (± 2)	2	167	0.26 (± 0.03)
Operator variants				
C variant	17 (± 3)	2	231	0.21 (± 0.03)
Loop	310 (± 22)	1	12	–
A bulge	330 (± 30)	2	11	–
Chimera	360 (± 8)	1	10	–
F6	82 (± 6)	1	47	0.28 (± 0.03)

The apparent equilibrium dissociation constants (K_d) for the binding of wild-type MS2 coat protein or its mutants to the wild-type operator fragment, together with values for several operator-related sequences binding to wild-type coat protein derived from Figure 4, are listed. Data were derived from the curve fitting algorithm described in Materials and Methods. The protein concentration in all the titrations was 1 μM coat protein (monomer). Experimental errors, in parentheses, are standard deviations of the mean value. *n* is the number of repetitions performed. In the case of single experiments the errors are those of the fitting procedure. E_t is the calculated concentration of binding sites.

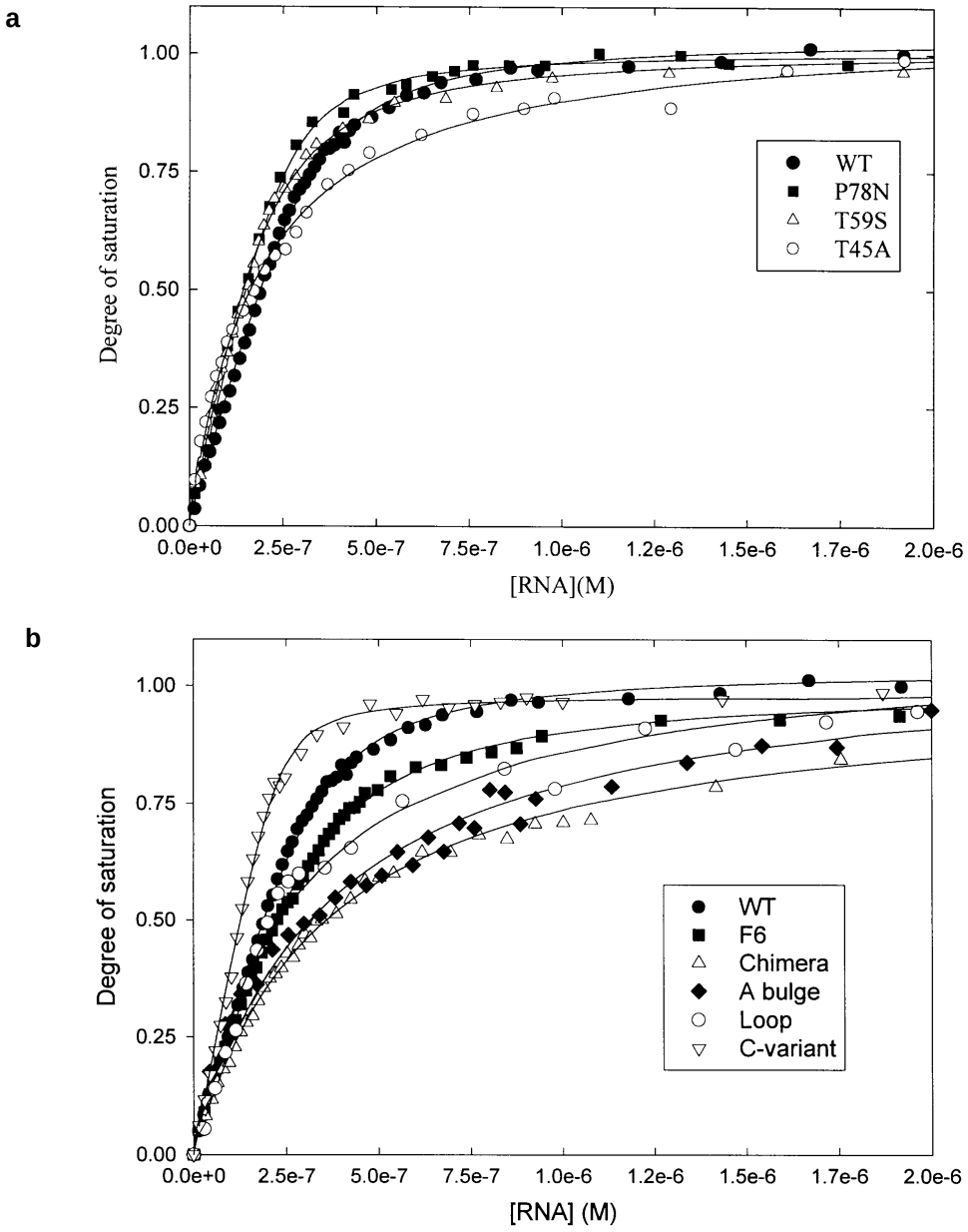

Figure 4. Fluorescence binding assays of RNA–coat protein complex formation. (**a**) Wild-type (WT) and mutant proteins with wild-type operator RNA and (**b**) wild-type (WT) protein with variant RNAs. Assay details are as described in Materials and Methods.

The wild-type behaviour of the Glu76Asp mutant in the superinfection assay (Table 1) was as expected, due to the separation of residue 76 and the RNA binding site. It was not possible to perform a fluorescence binding assay on this mutant protein due to the unusual behaviour of the mutant capsids under dissociation conditions. Wild-type capsids are fully dissociated using the method described previously (24), whereas Glu76Asp capsids are not, as judged by TEM (data not shown). A possible explanation of this behaviour comes from the crystal structures of wild-type and Glu76Asp capsids, in which the side chains of the residue at position 76 stabilize the B-type conformer by making a contact to Gln40. Protonation of the acidic side chain at residue 76 would disrupt this contact and might well be the cause of the acid-induced dissociation of bacteriophage capsids. This would be more difficult in the mutant protein capsid because of the lowered pK_a of aspartate relative to glutamate. Alternatively, reassembly with this mutant protein might occur in the dilute acetic acid solutions used to stabilize the dimeric form of the wild-type protein subunit.

In vivo the Pro78Asn mutant coat protein also behaves like the wild-type as judged by the superinfection assay. The increased apparent *in vitro* affinity of this mutant for operator RNA is intriguing, since this change was not expected to affect RNA binding, for the reasons discussed above. The initial hypothesis regarding control of quasi-equivalence in this system, however, invoked *cis–trans* isomerization at this proline residue, implying that the conformation of the coat protein dimer forming an operator complex might be linked in some way with RNA

binding and self-assembly (4,15). This would also be consistent with the unique orientation of the operator fragment bound at A/B coat protein dimers within the phage capsid. The recent results with a fr mutant lacking most of the FG loop make this seem less plausible (Axblom *et al*., personal communication). One possibility to account for the effect of this mutation might be that capsid assembly is occurring in the fluorimeter cuvette, making the apparent here affinity dependent on the *cis–trans* isomerization reaction in the wild-type. The MS2 Asn78 mutant might be able to assemble more readily because of absence of this isomerization, consistent with the increased flexibility of the FG loop in capsids of this mutant (5). The increased affinity, however, is clearly visible throughout the binding curves (Fig. 4a), even at stoichiometries of coat protein:RNA of between 10 and 100:1, where it has been reported previously that a maximum of only 30% of the protein would be assembled into capsid and even then only after much longer incubation times than used here (15).

The magnitude of the increased affinity seen here (~2-fold), therefore, appears too large for such an explanation and suggests that there *is* a direct linkage between operator RNA binding and coat protein conformation, consistent with earlier proposals on the role of the operator in assembly initiation (15), although charge neutralization is clearly also important (34). Further support for a direct linkage comes from quenching of the intrinsic fluorescence upon RNA binding. The fluorescence spectrum of the coat protein is dominated by the two tryptophan residues, Trp32 and Trp82. Of these, Trp32 is in the D β-strand, does not change its environment between quasi-equivalent conformers and is distal to the bound RNA. Trp82, on the other hand, is adjacent to the FG loop, changes its orientation markedly between A, B and C conformers but is still no closer than 7.2 Å (to the Cα carbon) to the RNA when bound. CD spectra of the coat protein as free dimers or as capsids suggest that there is very little, if any, rearrangement of the secondary structures in these two states (35). These data, therefore, suggest that quenching of the fluorescence signal is a direct effect of RNA binding, probably correlated with conformational changes at the FG loop and not due to effects of dimerization and/or refolding. If such changes do occur, the likely candidate for a switch linking loop conformation and RNA binding is the side chain at AlaB80, which would clash sterically if the orientation of the operator RNA on an A/B coat protein dimer were reversed (7).

Binding of operator variants

As expected (7,25,32,33), the C –5 operator variant had a higher *in vitro* affinity for coat protein than the wild-type in the fluorescence titration assay, although here the effect was less marked (~2.5 times) than reported previously for filter binding assays (6–30 times; 9,32). This is consistent with the significantly increased half-life of the C –5 operator complex, which would have a proportionally larger effect in such assays. We have proposed that the higher affinity is due to the additional internal hydrogen bond, compared with the wild-type, made between the exocyclic amino group of C –5 and the adjacent 5′ phosphodiester (Fig. 3c). This probably stabilizes the RNA in solution in a conformer closer to that bound by the protein, hence reducing the free energy required to distort the RNA during complex formation (7,14).

Both RNA variants encompassing operator half-sites had significantly reduced affinities for the coat protein, suggesting that tight binding requires that both sets of interactions are made simultaneously. (The A bulge site was synthesized with a triethyleneglycol linker between the two oligonucleotide strands in order to stabilize the duplex produced.) The low affinity of the loop oligonucleotide was surprising, since it is in this region that the majority of intermolecular contacts are made with the intact wild-type operator. The T_m of this fragment was determined to be 24.1 ± 2.0°C and, since the binding assays were carried out at ~20°C, it is possible that the loop RNA was not fully folded under these conditions. Repeating the assay at 10°C, where we would expect the loop to be mostly fully folded, did not result in an increased affinity. We then produced and tested an RNA–DNA chimera in which the MS2 loop sequence was stabilized on an extended base paired deoxynucleotide stem. The deoxynucleotide 'support' would not be expected to make the additional canonical contacts between the operator backbone and the protein in the wild-type operator complex (Fig. 1). Once again the affinity of the chimera was similar to the loop alone.

Preliminary X-ray data have been obtained for capsids soaked with the loop, A bulge and chimera oligonucleotides (Grahn *et al.*, personal communication; Stonehouse *et al.*, personal communication). The results confirm the importance of the loop interactions, since there is good density for the loop fragments bound to A- and C-type coat protein subunits, as expected. The capsids soaked with bulge RNA and the chimeric molecule, however, showed only poorly ordered density for the RNA.

Finally, we assayed binding of the wild-type protein to an RNA aptamer produced via *in vitro* selection experiments (19). Schneider *et al.* (36) have shown previously that RNAs based on the wild-type operator are the highest affinity ligands for the coat protein. However, it has now been possible to isolate a series of aptamer families with reasonable affinity for protein having differences in the canonical recognition sites. The oligonucleotide studied here is the consensus of Family 6 (F6) of Hirao *et al.* (19) and has a 3 base single-stranded loop, compared with the wild-type tetraloop. There are also 3 bp instead of two between the loop and the bulged adenosine (Fig. 1). The X-ray crystal structure of the aptamer operator complex, however, shows that this RNA is able to fulfil most of the key recognition contacts by adopting a flattened loop conformation, allowing G –7 to participate both in the stacking interaction seen with wild-type operators and to make a Watson–Crick base pair (20; Fig. 5). Thus the adenine contacts at –4 and –10 are still made, as is the pyrimidine stacking interaction at –5. Backbone and sugar contacts are similar to the wild-type operator structures. There are, however, some significant differences, including rearrangement of non-sequence-specific phosphate contacts and elongation of others (LysA43 to the phosphate at –4 and the contact between the 2′-OH of –5 and GluA63). In addition, one water-mediated contact has been elongated and another shortened. The sequence-specific contact between N4 of –5 and the phosphate backbone at –6 is lengthened, as is the contact between O2 of –5 and AsnA87. The decreased apparent affinity *in vitro* is consistent with the loss or modification of these important contacts (37).

Conclusions

The Thr45Ala mutant gave the expected phenotype *in vitro* and *in vivo*, given the degree of sequence conservation. The X-ray crystal structure of this operator capsid, however, shows that this

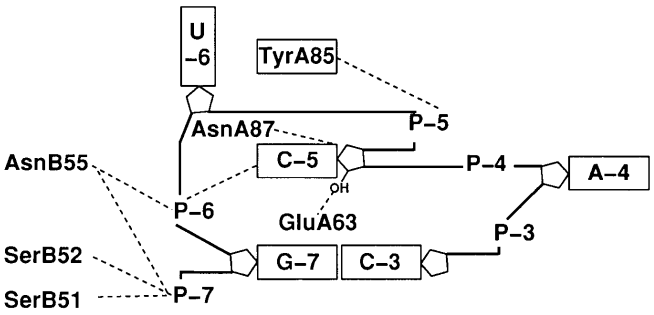

Figure 5. The loop contacts of the F6 aptamer. Diagram of the loop contacts made in the F6 RNA aptamer complex. RNA numbering is relative to the MS2 operator based on A –4.

residue is not critical for formation or stabilization of the operator complex (see figure 3 in 16). Similar conclusions can be reached for the Thr59Ser mutant. The superinfection assay shows the finely balanced nature of the molecular interactions *in vivo*, since a fall of ~20% in intrinsic RNA–protein affinity yields a variant unable to repress translation. This mutant also illustrates the complexity of the system and the subtle ways mutation can effect more than one feature of the bacteriophage, in this case operator binding and capsid stability.

The results with the Glu76Asp and Pro78Asn mutants have revealed the importance of the folded FG loop in B-type subunits for stabilizing the $T = 3$ capsid. The *in vitro* data with Pro78Asn also suggest some direct linkage between RNA binding and conformation of the FG loop, a conclusion supported by the change in intrinsic protein fluorescence. This result may be critically important in defining the pathway of self-assembly.

The data for the operator variants show that both sets of RNA–protein interactions in each coat protein dimer contribute similar interaction free energies, as judged by their similar K_d values. Tight binding requires that both interactions are made simultaneously. The results with the RNA aptamer show that similar interaction surfaces (20) and interaction free energies ($\Delta\Delta G \approx 1.5$ kJ/mol) can form from distinct secondary structures, illustrating the plasticity of RNA molecules.

ACKNOWLEDGEMENTS

We thank Lars Liljas for preparation of Figures 2 and 3 and he and his colleagues for their assistance in interpretation of the structural data. We also thank Siân Rowsell and Máire A. Convery for preparation of Figure 5, Andrew Ellington (Indiana University) and David Peabody (University of New Mexico) for access to their unpublished data on the RNA aptamer and Rachael Hill for helpful comments. This work was supported by grants from the UK BBSRC, the MRC, the Wellcome Trust and the Leverhulme Trust.

REFERENCES

1 Witherell,G.W., Gott,J.M. and Uhlenbeck,O.C. (1991) *Prog. Nucleic Acid Res. Mol. Biol.*, **40**, 185–220.
2 Stockley,P.G., Stonehouse,N.J., Walton,C., Walters,D.A., Medina,G., Macedo,M.B., Hill,H.R., Goodman,S.T.S, Talbot,S.J., Tewary,H.K., Golmohammadi,R., Liljas,L. and Valegård,K. (1993) *Biochem. Soc. Trans.*, **21**, 627–633.
3 Valegård,K., Liljas,L., Fridborg,K. and Unge,T. (1990) *Nature*, **345**, 36–41.
4 Golmohammadi,R., Valegård,K., Fridborg,K., and Liljas,L. (1993) *J. Mol. Biol.*, **234**, 620–639.
5 Stonehouse,N.J, Valegård,K., Golmohammadi,R., van den Worm,S., Walton,C., Stockley,P.G. and Liljas,L. (1996) *J. Mol. Biol.*, **256**, 330–339.
6 Ni,C.-Z., Syed,R., Kodandapani,R., Wickersham,R., Peabody,D.S. and Ely,K.R. (1995) *Structure*, **3**, 255–263.
7 Valegård,K., Murray,J.B., Stonehouse,N.J., Van den Worm,S.H.E., Stockley,P.G. and Liljas,L. (1997) *J. Mol. Biol.*, **270**, 724–738.
8 Valegård,K., Murray,J.B., Stockley,P.G., Stonehouse,N.J. and Liljas,L. (1994) *Nature*, **371**, 623–626.
9 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) *Nucleic Acids Res.*, **23**, 2512–2518.
10 Baidya,N. and Uhlenbeck,O.C. (1995) *Biochemistry*, **34**, 12363–12368.
11 Peabody,D.S. (1993) *EMBO J.*, **12**, 595–600.
12 Peabody,D.S. and Lim.F. (1996) *Nucleic Acids Res.*, **24**, 2352–2359.
13 Talbot,S.J., Medina,G., Fishwick,C.W.G., Haneef,I. and Stockley,P.G. (1991) *FEBS Lett.*, **283**, 159–164.
14 Borer,P.N., Lin,Y., Wang,S., Roggenbuck,M.W, Gott,J.M., Uhlenbeck,O.C. and Pelezer,I. (1995) *Biochemistry*, **34**, 6488–6503.
15 Beckett,D. and Uhlenbeck,O.C. (1988) *J. Mol. Biol.*, **204**, 927–938.
16 Van den Worm,S.H.E., Stonehouse,N.J., Valegård,K., Murray,J.B., Walton,C., Fridborg,K., Stockley,P.G. and Liljas,L. (1998) *Nucleic Acids Res.*, **26**, 1345–1351.
17 Liljas,L., Fridborg,K., Valegård,K., Bundule,M., Pumpens,P. (1994) *J. Mol. Biol.*, **244**, 279–290.
18 Hill,H.R., Stonehouse,N.J., Fonseca,S.A. and Stockley,P.G. (1997) *J. Mol. Biol.*, **266**, 1–7.
19 Hirao,I., Peabody,D.S. and Ellington,A.D. (1998) *Mol. Diversity*, submitted for publication.
20 Convery,M.A., Rowsell,S., Stonehouse,N.J., Ellington,A.D., Hirao,I., Murray,J.B., Peabody,D.S., Phillips,S.E.V. and Stockley,P.G. (1998) *Nature Struct. Biol.*, in press.
21 Mastico,R.A., Talbot,S.J. and Stockley,P.G. (1993) *J. Gen. Virol.*, **74**, 541–548.
22 Stonehouse,N.J. and Stockley,P.G. (1993) *FEBS Lett.*, **334**, 355–359.
23 Gill,C. and von Hippel,P.H. (1989) *Anal. Biochem.*, **182**, 319–326.
24 Sugiyama,T., Hebert,R.R. and Hartman,K.A. (1967) *J. Mol. Biol.*, **25**, 455–463.
25 Talbot,S.J., Goodman,S.T.S., Bates,S.R.E., Fishwick,C.W.G. and Stockley,P.G. (1990) *Nucleic Acids Res.*, **18**, 3521–3528.
26 Murray,J.B., Collier,A.K. and Arnold,J.R.P. (1994) *Anal. Biochem.* **218**, 177–184.
27 Puglisi,J.D. and Tinoco,I. (1989) *Methods Enzymol.*, **80**, 304–325.
28 Walton,C. (1993) PhD thesis, University of Leeds, UK.
29 Clowes,R.C. and Hayes,W. (1968) *Experiments in Microbial Genetics*, Blackwell Scientific Publications, Oxford, UK.
30 Carpenter,M.L. and Kneale,G.G. (1994) *Methods Mol. Biol.*, **30**, 313–325.
31 Clarke,A.R. (1996) In Engel,P.C. (ed.), *Enzymology Labfax*. Bios Scientific Publishers and Academic Press, San Diego, USA and London, UK.
32 Lowary,P.T. and Uhlenbeck,O.C. (1987) *Nucleic Acids Res.*, **15**, 10483–10493.
33 Carey,J., Lowary,P. and Uhlenbeck,O.C. (1983) *Biochemistry*, **22**, 4723–4730.
34 Hohn,T. (1969) *J. Mol. Biol.*, **43**, 191–200.
35 Scott,D.J. (1996) PhD Thesis, University of Leeds.
36 Schneider,D., Tuerk,C. and Gold,L. (1992) *J. Mol. Biol.*, **228**, 862–869.
37 Fersht,A.R., Leatherbarrow,R.J. and Wells,T.N.C. (1986) *Phil. Trans. R. Soc. Lond. A*, **317**, 13–28.