Peptide display on live MS2 phage: restrictions at the RNA genome level

Meerten, D. van; Olsthoorn, R.R.C.L.; Duin, J. van; Verhaert, R.M.D.

Note: To cite this publication please use the final published version (if applicable).

Peptide display on live MS2 phage: restrictions at the RNA genome level

Dico van Meerten, Rene C. L. Olsthoorn, Jan van Duin and Raymond M. D. Verhaert†

Leiden Institute of Chemistry, Department of Biochemistry, Gorlaeus Laboratories, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands

The potential of the RNA phage MS2 to accommodate extra amino acids in its major coat protein has been examined. Accordingly, a pentapeptide was encoded in the genome as an N-terminal extension. In the MS2 crystal structure, this part of the coat protein forms a loop that extends from the outer surface of the icosahedral virion. At the RNA level, the insert forms a large loop at the top of an existing hairpin. This study shows that it is possible to maintain inserts in the coat protein of live phages. However, not all inserts were genetically stable. Some suffer deletions, while others underwent adaptation by base substitutions. Whether or not an insert is stable appears to be determined by the choice of the nucleic acid sequence used to encode the extra peptide. This effect was not caused by differential translation, because coat-protein synthesis was equal in wild-type and mutants. We conclude that the stability of the insert depends on the structure of the large RNA hairpin loop, as demonstrated by the fact that a single substitution can convert an unstable loop into a stable one.

Introduction

Several animal and plant viruses have been investigated for their potential use as carriers and presenters of epitopes and some recombinants have been effective in raising protective immunity in animals (Koo *et al.*, 1999; Burke *et al.*, 1988; Dedieu *et al.*, 1992; Kratz *et al.*, 1999; Lomonossoff & Johnson, 1996; Borisova *et al.*, 1996; Smith *et al.*, 1998).

Usually, epitopes are expressed as a fusion with a virus capsid protein, whose structure is known at high resolution, allowing rational design of foreign inserts on the surface of the virus particle. In contrast, filamentous DNA bacteriophages such as M13 are generally employed to construct large peptide and synthetic antibody libraries that can be used to identify known or unknown ligands (Zwick *et al.*, 1998).

RNA phages are attractive candidates to display epitopes because their capsid structures are known at atomic resolution (Valegård *et al.*, 1990). MS2 is a single-stranded RNA

bacteriophage of 3569 nucleotides that is able to infect male *E. coli* bacteria via adsorption to F-pili. Upon infection, each cell produces about 10^3–10^4 virions. These have icosahedral symmetry and contain 180 copies of the major coat protein and one copy of the maturation protein, which is required for infection (reviewed by van Duin, 1988).

Virions of the RNA phages are highly antigenic and some groups have examined the possibility of displaying peptides on the surface of empty capsids (Pushko *et al.*, 1993; Mastico *et al.*, 1993; Peabody, 1997; Heal *et al.*, 1999). For this purpose, the sequences of choice were inserted after amino acid 2 or 11 of the major coat protein. Both of these positions lie in loops exposed at the capsid surface and the insert would merely enlarge the size of this loop (Valegård *et al.*, 1990). The octapeptide inserted by Peabody (1997) at either position inhibited capsid formation, but half of the pentapeptides inserted by Pushko *et al.* (1993) at position 2 were compatible with the formation of capsids.

No reports have appeared on inserting peptides in viable RNA phage. Clearly, if one could display peptides on the surface of live phages, this would potentially expand the technique with selection from libraries or with natural *in vivo* selection. To test the feasibility of this approach, we took a mutant studied by Pushko *et al.* (1993) in which five amino

Author for correspondence: Jan van Duin.

Fax +31 71 527 4340. e-mail j.duin@chem.leidenuniv.nl

† Present address: Cargill R&D Europe, PO Box 34, 4600 AA Bergen op Zoom, The Netherlands.

acids (ASISI) were inserted at position 1 of the coat protein of fr, a close relative of MS2. We chose this mutant because its capacity to form (empty) capsids had remained intact.

Compared with these *ex vivo* studies on capsid formation, the *in vivo* approach faces several additional difficulties. For example, the peptide insertion must be compatible with the formation of an infectious phage. This requires the correct incorporation of the maturation protein in the now-modified virion. Furthermore, the new coat protein must still be able to act as a translational repressor by binding the replicase operator (Witherell *et al.*, 1991). Another problem facing peptide display on live phages is that the extra RNA encoding the peptide can come under selection pressure for various reasons. If it cannot adopt a proper secondary structure, it may fall prey to *E. coli* endonucleases (Arora *et al.*, 1996; Olsthoorn & van Duin, 1996; Klovins *et al.*, 1997; van Meerten *et al.*, 1999; Steege, 2000). On the other hand, the 15 extra nucleotides encoding the five amino acids are likely to change the local RNA secondary structure at the coat start and this might affect translational yield significantly (de Smit & van Duin, 1990 *a*).

The degeneracy of the genetic code allowed us to make various mutants, all encoding the same ASISI insert in the coat protein. Several of these mutants were unstable and suffered deletions that reduced the insert to two or three amino acids. Other mutants were fully stable. Thus, it is the sequence of the nucleic acid rather than that of the protein that determines the genetic stability of the mutants. These differences cannot be ascribed to differential translational yields, but they must be a result of different RNA structures.

Methods

■ **Bacterial strains and plasmids.** F$^-$ strain M5219 (M72 $trpA_{am}$, $lacZ_{am}$, Sm^r/$\lambda\Delta bio_{252}$, cI_{857}, ΔH_1), encoding the thermosensitive repressor (cI_{857}) and the transcriptional anti-termination factor N (Remaut *et al.*, 1981), was used to generate phages from their wild-type and mutant MS2 cDNA plasmids. *E. coli* F$^+$ KA797 (F′ $lacI^Q$, pro/ara, Δlac–pro, thi) was used as the host to allow the evolution of mutant MS2 phages. The strains were grown on LC broth containing (per litre): 10 g Bacto tryptone, 5 g yeast extract, 8 g NaCl, 2 g MgSO$_4$, 140 mg thymine and 1 ml 1 M Tris–HCl, pH 7·6.

Plasmid pPLc2822-K33 (de Smit & van Duin, 1993) was derived from pPLc2833 (Remaut *et al.*, 1981) and carries the MS2 cDNA from position 1303 to 2057. It contains an engineered unique *Nsi*I site at position 1336. Plasmid pMS2000Δ, previously called pD (Licis *et al.*, 1998), was derived from the infectious clone pMS2000 (Olsthoorn *et al.*, 1994) by replacing the nt 1303–1901 MS2 cDNA fragment by a short linker. Plasmid pMS2000Δ confers kanamycin resistance. pMS2 contains the MS2 cDNA from position 103 to 2057 and has been described by Berkhout *et al.* (1987). pMS23 carries the MS2 cDNA from position 103 to the 3′ end and has been described previously (Berkhout & van Duin, 1985). Both plasmids pMS2 and pMS23 have unique *Xba*I (1303) and *Afl*II (1901) sites in the MS2 cDNA and carry the ampicillin-resistance gene. All MS2 cDNAs in these plasmids are under the control of the P$_L$ promoter.

■ **Construction of MS2 cDNA mutants.** The mutants were constructed by ligating complementary oligonucleotides in plasmid pPLc2822-K33 cleaved by *Nsi*I (1336) and *Sal*I (1365). The sequences of all oligonucleotides were checked after insertion in the plasmid.

Each set of complementary oligonucleotides was hybridized by mixing them in a 1:1 molar ratio, heating to 70 °C and slow cooling to room temperature. From the mutant plasmid, the *Xba*I (1303)–*Afl*II (1901) fragment was cloned in plasmid pMS2000Δ to produce the complete MS2 cDNA sequence. This procedure was followed for ASISI 1 and ASISI 3. Because of the high virulence of full-length MS2 cDNA, it is often difficult to transform cells stably with such a plasmid, as there is a strong selection against its presence. Thus, we sometimes used a two-plasmid system, where the genetic information for the phage is divided over two compatible replicons in an overlapping manner (Olsthoorn, 1996). In such a case, an RNA or DNA recombination event is required to obtain the contiguous sequence. Mutants ASISI 2, ANISI 1 and ANISI 2 were obtained in this way. The mutations were introduced in pMS23 and complemented with pMS2000 Δ. When this combination of plasmids contained wild-type sequences, the supernatant of an overnight culture contained 10⁹ p.f.u./ml. This is 100-fold less than from the contiguous sequence pMS2000. All titres mentioned in the Results are normalized to the one-plasmid value.

■ **Phage generation, titre estimation and phage evolution.** Bacteria carrying the wild-type or mutant MS2 cDNA plasmid(s) were grown overnight at 28 °C in LC in the presence of antibiotics. The supernatant (cycle 1) of the culture was plated on a lawn of KA797 (F$^+$) cells and incubated overnight at 37 °C. Plaques (cycle 2) were counted and individual plaques were taken and amplified overnight on KA797 cells in liquid culture (cycle 3). Phages were passaged for two more cycles of infection by growing them overnight at 37 °C in 2 ml liquid cultures containing KA797 cells (cycles 4 and 5). Routinely, at several evolutionary stages, the sequence between nt 1300 and 1365 was determined by RT–PCR. For some pseudorevertants, it was necessary to extend the sequenced region to nt 1551 (see below).

■ **RT–PCR and sequence analysis.** Phages from individual plaques or from 1 μl of an overnight-infected *E. coli* culture were suspended in 6 μl water and heated at 92 °C for 2 min. One μl was used for RT–PCR according to standard procedures recommended by the suppliers (Promega and Eurogentec). The primers used were biotin-labelled DUI715 (identical to nt 1182–1205) and unlabelled DUI114 (complementary to nt 1894–1878). PCR fragments were sequenced with DUI59 (complementary to nt 1422–1409) after separation and purification of the strands using Dynabeads (Dynal).

■ **Coat-protein expression measurements.** To measure the coat-protein yield of mutants and pseudorevertants, the phage RNA was amplified by RT–PCR as described above and the *Xba*I (1303)–*Bst*XI (1551) fragment was cloned into the expression plasmid pMS2 carrying the partial MS2 cDNA (nt 103–2057) behind the P$_L$ promoter. *E. coli* M5219 cells harbouring the expression plasmids were grown in LC at 28 °C to an OD$_{650}$ of 0·2, after which transcription was induced by heating to 42 °C. Samples for visualization of coat-protein synthesis by Western blot were taken 1 h after induction (de Smit & van Duin, 1990 *b*). An antiserum raised against SDS-denatured MS2 coat protein was used for detection.

■ **Reintroduction of pseudorevertant sequences in infectious MS2 cDNA plasmid combinations.** cDNA of ASISI 3.1, which is a pseudorevertant of ASISI 3, was prepared by ligating the appropriate complementary oligonucleotides covering the sequence 1336–1365 (*Nsi*I–*Sal*I) in pPLc2822-K33. The mutant sequence was then incorporated in pMS23 (position 103 to the 3′ end) as an *Xba*I–*Afl*II fragment (nt 1303–1901) and complemented with pMS2000Δ for the production of phages.

The *Xba*I–*Bst*XI fragments (nt 1303–1551) from pseudorevertants ASISI 1.1, ASISI 1.3, ASISI 2.3 and ASISI 2.5 were prepared from phage

RNA by RT–PCR and then used to replace the wild-type fragment in pMS2. The complete insert (1303–1551) was sequenced at the plasmid level after its introduction in pMS2. No mutations were found in addition to those reported in Results, except for a U → C transition in ASISI 2.3 at position 1440. For phage production, pMS2 containing the pseudo-revertant sequence was complemented with pMS2000Δ (lacking fragment 1303–1901). A double cross-over is required to obtain phages. Using wild-type sequences, the plasmid combination yielded 10^5 p.f.u./ml overnight culture.

Results

The system

In order to generate the desired MS2 phages, we transform F⁻ bacteria with a plasmid carrying the complete wild-type or mutant MS2 cDNA. This leads to the spontaneous formation of phages in the host and, after overnight growth of the transformed bacterial culture, the supernatant contains about 10^{11} p.f.u./ml for the wild-type sequence. For mutants, the titre is lower, depending on the gravity of the mutation. The MS2 cDNA need not be induced in order to obtain phages. There is apparently enough spurious transcription taking place, and a single transcript per cell may set the infection process in motion (Taniguchi *et al.*, 1978). As explained in Methods, we sometimes prefer to divide the MS2 cDNA over two compatible plasmids.

Choice of the peptide insert

For the present study, we used the pentapeptide ASISI, inserted after amino acid 1 of the MS2 coat protein, to yield the sequence (M)**ASISI**ASNF..., where the extra amino acids are in bold. The methionine that is normally split off from the mature protein is shown in parentheses. [Note that the insert can also be considered as an insertion after amino acid 3 (ISIAS) (underlined)]. Pushko *et al.* (1993) reported that this pentapeptide did not interfere with capsid formation.

To simplify insertion at the required position, we made use of an MS2 mutant that contained a unique *Nsi*I site just ahead of the coat start (Fig. 1). There is a large number of possibilities to encode these five new amino acids. Our choice was guided by several principles. Firstly, we avoided codons for which there is little tRNA, as this might affect translation efficiency (Spanjaard & van Duin, 1988). Secondly, the extra nucleotides should not add too many extra base-pairs to the initiator hairpin, as the resulting stabilization could inhibit coat expression (de Smit & van Duin, 1990*b*). On the other hand, the insert should not be present as a mere extension of the hairpin loop, since we have observed previously that very large hairpin loops are excised in RNA phages (Olsthoorn & van Duin, 1996). In Fig. 1, we show the three insertion mutants used in this study.

First nucleotide choice for the pentapeptide sequence

Our first choice of how to encode these extra amino acids is shown in Fig. 2 as ASISI 1. The start codon is boxed in grey
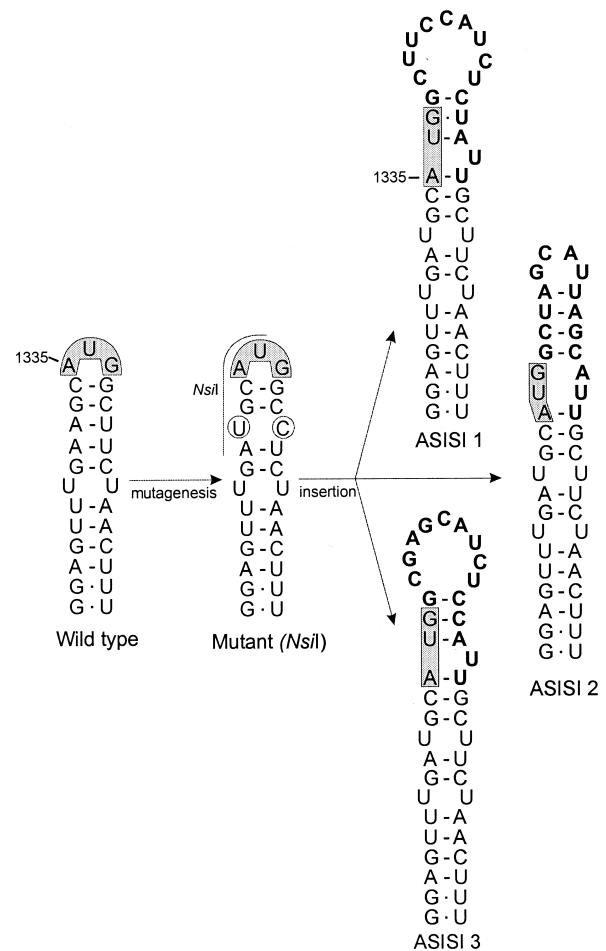

Fig. 1. Predicted structures of MS2 RNA at the start of the coat-protein gene for wild-type and mutant sequences. Mutant (*Nsi*I) is a helper construct. ASISI 1, 2 and 3 are the 15 nt insertion mutants. The start codon is shaded and inserted residues are in bold. The structure of the wild-type hairpin is strongly supported by probing and mutational analysis (Skripkin *et al.*, 1990; de Smit & van Duin, 1990*a*; Olsthoorn *et al.*, 1994). Mutant structures were deduced by eye and were confirmed by using MFold (Zuker *et al.*, 1999).

and the insert is in bold. As can be seen, of the 15 extra nucleotides, nine form the pyrimidine-rich loop, while the remaining six extend the stem and add some stability to it. Rather unexpectedly, the presence of the extra 15 nucleotides caused a dramatic drop in fitness, since only three plaques could be recovered from a 1 ml overnight culture. For the wild-type sequence, 10^{11} plaques were recovered. Each of the plaques was allowed to evolve up to cycle 5 and the sequence was determined after each cycle (see Methods). For the sake of simplicity, we show only the final sequence after five cycles (Fig. 2). It was found that plaques 1 and 2 (ASISI 1.1 and ASISI 1.2) first had a hexanucleotide deleted from the insert, corresponding to Ile and Ser. Thereafter, two different suppressor mutations were selected in the loop (white on black lettering in Fig. 2). How these contribute to fitness is not clear. Plaque 3 suffered only a 9 nt deletion. No other suppressor mutations were found. The three pseudorevertants that were
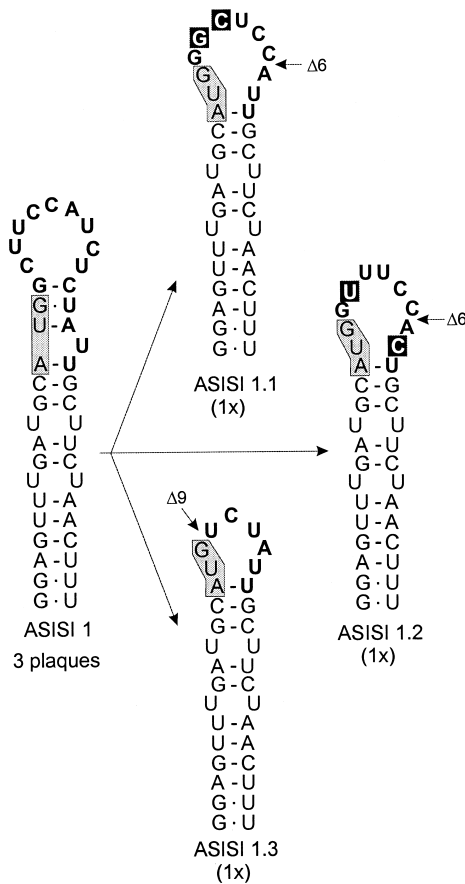
Fig. 2. Predicted RNA structures for ASISI 1 and its revertants. The start codon is boxed, the insertion is in bold and adaptive substitutions in the revertants are in white on black. Sites of deletions are shown by arrows (e.g. Δ6 represents a 6 nt deletion). The number in parentheses refers to the number of plaques found for each revertant.

Fig. 3. Predicted RNA structure for mutant ASISI 2 and its revertants. See legends to Figs 1 and 2 for further details.

obtained in this experiment produced about the same number of plaques as the wild-type phage. It is clear that the original insert caused great problems, and that drastic changes such as deletions sometimes followed by base changes were needed to recover the original viability.

Second nucleotide choice for the pentapeptide sequence

In the second experiment, we used a different set of codons to produce the same pentapeptide (ASISI 2). The two serine residues were now determined by AGC rather than UCU and UCC. The tentative hairpin resulting from these changes is shown in Fig. 3. The new choice avoids the presence of a large pyrimidine-rich loop at the top, which could be responsible for the deletions in the ASISI 1 mutant. The thermodynamic stability of the ASISI 2 hairpin is increased relative to the wild-type and the construct was thus predicted to yield less coat protein. If so, the right suppressor mutations might be selected to adjust the hairpin stability and protein yield.
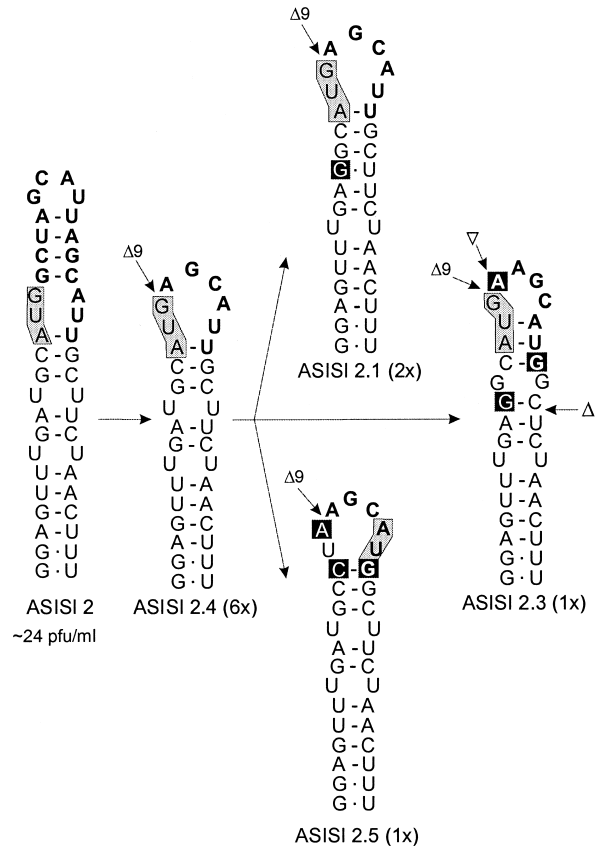
The titre of ASISI 2 was again very low. In fact, with our standard protocol, we did not get any plaques from the supernatant of an overnight culture of transformed cells. We therefore lysed the cells with lysozyme to release their contents. From two independent experiments, we recovered 21 and 27 p.f.u./ml culture in this way. From one of the cultures, six plaques were passaged up to cycle 5. All sequences were identical when they were first determined at cycle 3 (ASISI 2.4; Fig. 3). Nine contiguous nucleotides (encoding ASI) had been deleted. All six of these pseudorevertants were again sequenced after cycle 5. Two remained unaltered, while four had accumulated additional suppressor mutations in the stem and sometimes also in the loop. In ASISI 2.1, a U U mismatch had changed into a G·U pair. In ASISI 2.5, an A–U pair was replaced by C–G and, in ASISI 2.3, an extra base pair was added at the top of the stem. All these changes increase thermodynamic stability and this may be the pressure for their selection. In phage ASISI 2.5, one of the mutations recreated a start codon at the original position in the reading frame, thus producing a wild-type coat protein.

In general, the properties and fate of the ASISI 2 insert were similar to those of the previous series. The titre was extremely low and the return to high titre could apparently only be
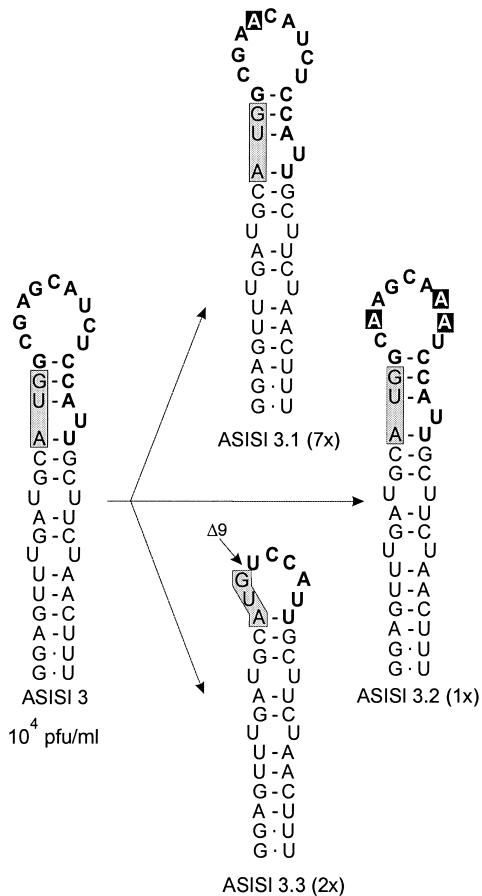
Fig. 4. Predicted RNA structure for mutant ASISI 3 and its revertants. See legends to Figs 1 and 2 for further details. MFold draws a pair between the G and the C in the loops of ASISI 3 and ASISI 3.1. However, this putative pair is lost in ASISI 3.2 and therefore it is insignificant or non-existent.

Fig. 5. Western blot showing the translational efficiency of the coat-protein gene (COAT) in several mutants and revertants. EF-Tu was stained briefly with anti-EF-Tu to provide an internal standard for the amount of cell extract applied in each lane. In the negative control, the empty expression vector is shown. Cultures with the plasmid-encoded coat protein were grown at 28 °C until the OD_{650} was 0·2. The P_L promoter, controlling transcription of the coat gene, was then induced for 60 min by heating the culture to 42 °C.

achieved by a deletion, followed by several adaptive mutations, the advantages of which are only partly understood.

Third nucleotide choice for the pentapeptide sequence

We next constructed ASISI 3, in which we returned to the 9 nt loop but included more purines in it than in ASISI 1 (Fig. 4). This choice was based on the loop composition found in the pseudorevertants of ASISI 1. These revertants, at least ASISI 1.1 and 1.2, show also that a large loop of 10 nt does not have to be a problem. In two independent transformations, the ASISI 3 construct yielded about 10^4 p.f.u./ml overnight culture. This construct is therefore about 1000 times better than ASISI 1.

Two colonies from independent transformations were selected. Five plaques from each were sequenced. As shown in Fig. 4, two had a 9 nt deletion (ASISI 3.3). Surprisingly, eight had not suffered any deletion but had maintained all 15 nt. There were changes in the sequence, however; one in ASISI 3.1 and three in ASISI 3.2. Interestingly, all of these changes were in the hairpin loop, but they did not seem to add
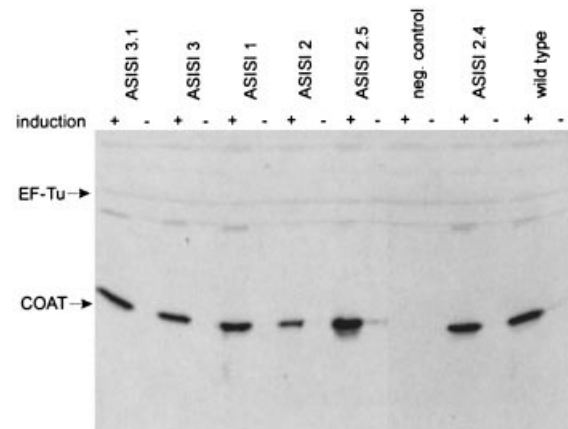
conventional base pairs. However, they resulted in different pentapeptide sequences.

In order to understand better the deletions and substitutions in all of our pseudorevertants, it was necessary to identify the evolutionary pressure that led to their selection. Is it the sequence of the peptide insert, the RNA hairpin-loop structure or the efficiency of coat gene translation that adversely affects viability in our starting mutants?

The inserts have only a marginal effect on coat-protein gene translation

The effect of the insertion on coat-protein yield can be readily measured by using expression vectors carrying the coat-protein gene sequences of the wild-type, the initial mutants and their pseudorevertants. Western blots such as the one shown in Fig. 5 revealed that the differences in coat-protein yield between wild-type, mutants and pseudo-revertants were rather small, at most a 3- to 4-fold decrease for ASISI 2 relative to wild-type. Such differences, however, have never been seen to cause changes in titre of 10 orders of magnitude. For instance, in a previous study, initiation mutants with a 30- to 1000-fold decrease in coat-protein yield still showed a titre that was only decreased by three to four logs (Olsthoorn *et al.*, 1994). Moreover, ASISI 1 produced about the same amount of coat protein as wild-type but still suffered the same 10 logs drop in titre as ASISI 2.

The low titre of our mutants therefore might be due either to the extra nucleotides in the genome or to the extra amino acids in the coat protein. The extended protein may be incompatible with some control or virion function. Alternatively, the protein may be fully functional but the extra RNA may become a handicap of some sort. Pseudorevertants ASISI
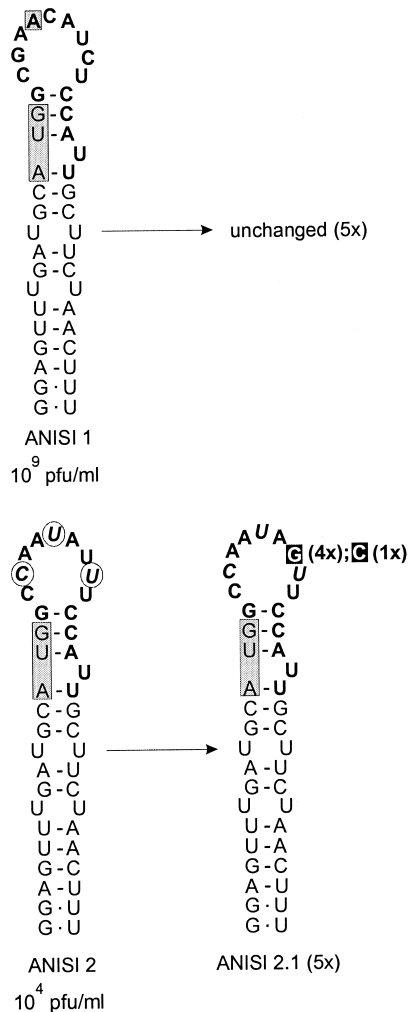
Fig. 6. Predicted RNA structure for ANISI 1 and ANISI 2 and their revertants. See legends to Figs 1 and 2 for further details.

3.1 and ASISI 3.2 partly resolved this issue, however, by showing that neither the protein nor the RNA insert was unacceptable per se.

Selection is against the inserted RNA, not the inserted peptide

Although the two revertants ASISI 3.1 and 3.2 tolerated the insert, it is clear that the starting sequence of ASISI 3 is in need of improvement. The suppressor mutations selected in ASISI 3.1 and 3.2 prescribe a different protein insert sequence, thus raising the question whether the evolutionary pressure for the changes originated from a non-functional protein or from an unfavourable RNA structure. To distinguish between the two possibilities, we started from the evolved phage ASISI 3.1, carrying a single substitution G8A that turned serine into asparagine (A of the AUG is residue 1). Either serine is a bad amino acid at this position in the protein, or G8 is somehow deleterious at the RNA level. To answer this question, we made two new MS2 cDNA constructs. In one, ASISI 3.1, the

single G8A suppressor mutation was cloned back in the infectious clone. It is now called ANISI 1 because of the Ser-to-Asn change (Fig. 6). This sequence should stay unaltered in the emerging phage and the titre of the construct should be high. In the other construct, ANISI 2, we left the new amino acid sequence unchanged but made three substitutions in wobble positions (circled italics in Fig. 6).

As expected, the infectious cDNA construct with the revertant sequence (ANISI 1) had a titre of 10^9 p.f.u./ml, while the original ASISI 3 with a G at position 8 had a titre of only 10^4 p.f.u./ml. This shows that the G8A exchange was not an accidental replication error but was selected through improved phage fitness, which increased by five orders of magnitude. The G8A mutation was the only difference between ASISI 3 and ANISI 1 (see below).

With construct ANISI 2, we tested whether the amino acid or the nucleotide sequence was the target of selection. ANISI 2, encoding the same pentapeptide as ANISI 1, had a titre that was 10^5 lower than that of ANISI 1, and it formed very small plaques. It is thus clear that it is the RNA sequence and not the resulting amino acid sequence that hampered phage growth. Five plaques derived from ANISI 2 were sequenced. We found the U11G substitution in four of them and U11C in one. Again, these substitutions are in the loop and suggest that, from the point of view of fitness of the phage, there was something wrong with the loop in ANISI 2. The suppressor mutations corrected this defect. That the mutations changed the peptide sequence at the same time was probably a coincidence.

The phenotypic reversions are the result of the recorded suppressor mutations

A priori, we cannot exclude that the phenotypic reversion of our pseudorevertants was due to mutations outside the sequenced region. Therefore, it was necessary to clone back the sequenced section in an otherwise wild-type MS2 cDNA background and to show maintenance of the revertant phenotype. ANISI 1 is an example. The cloned-back sequence extended only from 1336 (*Nsi*I) to 1365 (*Sal*I). Therefore, the four logs difference in titre between ASISI 3 cDNA and ANISI 1 cDNA could be fully ascribed to the single G8A suppressor mutation, which constitutes the only difference between the constructs.

A similar analysis was carried out for pseudorevertants ASISI 1.1, ASISI 1.3, ASISI 2.3 and ASISI 2.5. Here, the revertant region 1303–1551 (*Xba*I–*Bst*XI) was cloned back in pMS2 (103–2057) and, after complementation with pMS-2000Δ (Δ1303–1901), the titres of overnight cultures were compared to those of the wild-type equivalent. The original mutant sequences ASISI 1 and ASISI 2 produced no phages in this plasmid combination, but all pseudorevertant sequences showed titres that were the same as that found for wild-type (see Methods).

The cloned-back cDNA of the four pseudorevertants was sequenced (nt 1303–1551). No mutations were found outside

the initiator hairpin except for a U → C transition at position 1440 in one of the four sequenced clones (ASISI 2.3). We do not know whether this mutation is part of the phenotype or was introduced by the PCR. At any rate, for the three other pseudorevertants, the sequence changes in the initiator hairpin accounted fully for the phenotype.

Discussion

In this study, we have explored the possibility of inserting pentapeptides in the coat protein of live virions of the RNA phage MS2. We chose the N terminus as the site of the insertion because the inserted peptide would be in a loop that was directed outwards (Valegård et al., 1990). The sequence of the peptide ASISI was taken from a study by Pushko et al. (1993), who reported that this insert did not interfere with the formation of (empty) capsids.

There are about 1000 ways to encode a pentapeptide and we restricted our choice somewhat by excluding rare codons. In our first mutant (ASISI 1), we made the 9 nt loop of the initiator stem highly pyrimidine rich because there is a preference for these in MS2 RNA hairpin loops (Olsthoorn & van Duin, 1996). Coat-protein expression from this mutant was not much different from wild-type, but its fitness was reduced dramatically, as reflected in a drop in titre of about 10 orders of magnitude. Reversion to high titre was achieved by a deletion of 6 or 9 nt (two or three amino acids) followed by some adaptive mutations in the loop of the initiator hairpin.

In our next construct (ASISI 2), two other serine codons were tested. Their presence predicts the formation of an initiator hairpin with a more-common four-membered loop. The properties of this mutant were, broadly speaking, the same as those of ASISI 1. The titre was down by about 10 logs and the cure to high titre was the deletion of 9 nt followed by several adaptive mutations, some of which were in the loop of the initiator hairpin and others in its stem. The third try (ASISI 3) was successful. Here, the 9 nt loop of the initiator hairpin contained more purines than in ASISI 1. This choice reflected the loop composition in the revertants from ASISI 1 and, to some extent, from ASISI 2. The full insert was still present in the majority of plaques. Only some adaptive mutations in the loop had occurred and these changed the amino acids of the insert. Were these adaptations selected to improve the protein or the RNA structure? The answer to this question came from further study of revertant ASISI 3.1, which had a G8A (Ser → Asn) substitution in the hairpin loop.

We recloned the relevant part of the revertant sequence (ASISI 3.1) in the MS2 infectious cDNA (ANISI 1) and made one derivative in which the new 'phage-approved' amino acid sequence was encoded differently (ANISI 2). As expected, the recloned pseudorevertant sequence (ANISI 1) had a high titre and the sequence stayed unchanged in phages derived from it. In contrast, ANISI 2 gave a low titre and needed a suppressor mutation in the loop to boost its fitness.

The conclusion must be that it is the nucleic acid sequence of the loop that determines the fitness and genetic stability of the insert. This is particularly clear for ASISI 3.1 and 3.2 and for the progeny of ANISI 2, where a single mutation in the loop was apparently sufficient to make a stable phage (ANISI 2.1).

The analysis carried out here is reminiscent of another study from this lab, in which a non-coding hairpin loop was enlarged by 26 nt (Olsthoorn & van Duin, 1996). As seen here, there was a drop in titre and there were two kinds of adaptations. Sometimes, the big loop was resected to an acceptable size. In other pseudorevertants, a few substitutions in the loop sufficed to remove most of the genetic burden and raise the fitness.

We tentatively suggest that these big hairpin loops create a target for RNase E, one of the major endonucleases of *E. coli*. Suppressor mutations in the large loop such as those found in ASISI 3.1 and 3.2 possibly induce resistance. This may result from interactions between loop nucleotides, reminiscent of what has been found for GNRA and UNCG loops (Cheong et al., 1990; Heus & Pardi, 1991).

Why is the drop in viability of the insertion mutants so large?

An unsolved question is why the insertions ASISI 1 and ASISI 2 have such a dramatic effect on the titre. In our previous study, where we introduced 26 single-stranded nucleotides in a non-coding loop (Olsthoorn & van Duin, 1996), the titre dropped by only three orders of magnitude. Even considering that, in an evolutionary sense, it is simpler to remove nucleotides from a non-coding sequence than from a reading frame, it is difficult to account for the dramatic drop in titre. We cannot rigorously exclude that the mutations affect replication, although it is hard to envisage why certain loop sequences would support replication while others would not. Unfortunately, unlike results obtained for Q_β phage, the replicase of MS2 cannot be isolated in a straightforward, reproducible way (Federoff, 1975) and the influence of the mutations on replication therefore cannot be tested easily.

Deletions versus substitutions

Another question remaining is why ASISI 1, for instance, with a 9 nt loop, is not rescued by substitutions and suffers a deletion, whereas one or more substitutions suffice for the ASISI 3 and ANISI 2 mutants.

The likely answer is that ASISI 1, with its all-pyrimidine loop, would need more simultaneous substitutions than can be found in the quasispecies pool. For instance, the chances of finding a revertant with three transversions can be estimated to be about $(10^{-5})^3$ or 10^{-15}. A deletion can then be the more likely event to save the virus. Here, we would have to assume that the three substitutions in ASISI 3.2 have accumulated sequentially. This idea, that deletions are more likely than triple or double substitutions, is supported by a recent study of Licis

et al. (2000), where it was shown that a single stop codon in the MS2 lysis frame was repaired by a substitution to sense, but a double stop codon could not be cured in this way. Instead, the region containing the stop codons was excised, allowing rescue of the phage (at the expense of a shorter lysis protein).

References

Arora, R., Priano, C., Jacobson, A. B. & Mills, D. R. (1996). *cis*-acting elements within an RNA coliphage genome: fold as you please, but fold you must! *Journal of Molecular Biology* **258**, 433–446.

Berkhout, B. & van Duin, J. (1985). Mechanism of translational coupling between coat protein and replicase genes of RNA bacteriophage MS2. *Nucleic Acids Research* **13**, 6955–6967.

Berkhout, B., Schmidt, B. F., van Strien, A., van Boom, J., van Westrenen, J. & van Duin, J. (1987). Lysis gene of bacteriophage MS2 is activated by translation termination at the overlapping coat gene. *Journal of Molecular Biology* **195**, 517–524.

Borisova, G., Borschukova Wanst, O., Mezule, G., Skrastina, D., Petrovskis, I., Dislers, A., Pumpens, P. & Grens, E. (1996). Spatial structure and insertion capacity of immunodominant region of hepatitis B core antigen. *Intervirology* **39**, 16–22.

Burke, K. L., Dunn, G., Ferguson, M., Minor, P. D. & Almond, J. W. (1988). Antigen chimaeras of poliovirus as potential new vaccines. *Nature* **332**, 81–82.

Cheong, C., Varani, G. & Tinoco, I. (1990). Solution structure of an unusually stable RNA hairpin, 5′GGAC(UUCG)GUCC. *Nature* **346**, 680–682.

Dedieu, J. F., Ronco, J., van der Werf, S., Hogle, J. M., Henin, Y. & Girard, M. (1992). Poliovirus chimeras expressing sequences from the principal neutralization domain of human immunodeficiency virus type 1. *Journal of Virology* **66**, 3161–3167.

de Smit, M. H. & van Duin, J. (1990a). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proceedings of the National Academy of Sciences, USA* **87**, 7668–7672.

de Smit, M. H. & van Duin, J. (1990b). Control of prokaryotic translational initiation by mRNA secondary structure. *Progress in Nucleic Acid Research and Molecular Biology* **38**, 1–35.

de Smit, M. H. & van Duin, J. (1993). Translational initiation at the coat-protein gene of phage MS2: native upstream RNA relieves inhibition by local secondary structure. *Molecular Microbiology* **9**, 1079–1088.

Federoff, N. (1975). Replicase of the phage f2. In *RNA Phages*, pp. 238–258. Edited by N. D. Zinder. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Heal, K. G., Hill, H. R., Stockley, P. G., Hollingdale, M. R. & Taylor-Robinson, A. W. (1999). Expression and immunogenicity of a liver stage malaria epitope presented as a foreign peptide on the surface of RNA-free MS2 bacteriophage capsids. *Vaccine* **18**, 251–258.

Heus, H. A. & Pardi, A. (1991). Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* **253**, 191–194.

Klovins, J., van Duin, J. & Olsthoorn, R. C. L. (1997). Rescue of the RNA phage genome from RNase III cleavage. *Nucleic Acids Research* **25**, 4201–4208.

Koo, M., Bendahmane, M., Lettieri, G. A., Paoletti, A. D., Lane, T. E., Fitchen, J. H., Buchmeier, M. J. & Beachy, R. N. (1999). Protective immunity against murine hepatitis virus (MHV) induced by intranasal or subcutaneous administration of hybrids of tobacco mosaic virus that carries an MHV epitope. *Proceedings of the National Academy of Sciences, USA* **96**, 7774–7779.

Kratz, P. A., Böttcher, B. & Nassal, M. (1999). Native display of complete foreign protein domains on the surface of hepatitis B virus capsids. *Proceedings of the National Academy of Sciences, USA* **96**, 1915–1920.

Licis, N., van Duin, J., Balklava, Z. & Berzins, V. (1998). Long-range translational coupling in single-stranded RNA bacteriophages: an evolutionary analysis. *Nucleic Acids Research* **26**, 3242–3246.

Licis, N., Balklava, Z. & van Duin, J. (2000). Forced retroevolution of an RNA bacteriophage. *Virology* **271**, 298–306.

Lomonossoff, G. P. & Johnson, J. E. (1996). Use of macromolecular assemblies as expression systems for peptides and synthetic vaccines. *Current Opinion in Structural Biology* **6**, 176–182.

Mastico, R. A., Talbot, S. J. & Stockley, P. G. (1993). Multiple presentation of foreign peptides on the surface of an RNA-free spherical bacteriophage capsid. *Journal of General Virology* **74**, 541–548.

Olsthoorn, R. C. L. (1996). *Structure and evolution of RNA phages*. PhD Thesis, Leiden University, The Netherlands.

Olsthoorn, R. C. L. & van Duin, J. (1996). Random removal of inserts from an RNA genome: selection against single-stranded RNA. *Journal of Virology* **70**, 729–736.

Olsthoorn, R. C. L., Licis, N. & van Duin, J. (1994). Leeway and constraints in the forced evolution of a regulatory RNA helix. *EMBO Journal* **13**, 2660–2668.

Peabody, D. S. (1997). Subunit fusion confers tolerance to peptide insertions in a virus coat protein. *Archives of Biochemistry and Biophysics* **347**, 85–92.

Pushko, P., Kozlovskaya, T., Sominskaya, I., Brede, A., Stankevica, E., Ose, V., Pumpens, P. & Grens, E. (1993). Analysis of RNA phage fr coat protein assembly by insertion, deletion and substitution mutagenesis. *Protein Engineering* **6**, 883–891.

Remaut, E., Stanssens, P. & Fiers, W. (1981). Plasmid vectors for high-efficiency expression controlled by the P_L promoter of coliphage lambda. *Gene* **15**, 81–93.

Skripkin, E. A., Adhin, M. R., de Smit, M. H. & van Duin, J. (1990). Secondary structure of the central region of bacteriophage MS2 RNA. Conservation and biological significance. *Journal of Molecular Biology* **211**, 447–463.

Smith, A. D., Geisler, S. C., Chen, A. A., Resnick, D. A., Roy, B. M., Lewi, P. J., Arnold, E. & Arnold, G. F. (1998). Human rhinovirus type 14:human immunodeficiency virus type 1 (HIV-1) V3 loop chimeras from a combinatorial library induce potent neutralizing antibody responses against HIV-1. *Journal of Virology* **72**, 651–659.

Spanjaard, R. A. & van Duin, J. (1988). Translation of the sequence AGG–AGG yields 50% ribosomal frameshift. *Proceedings of the National Academy of Sciences, USA* **85**, 7967–7971.

Steege, D. A. (2000). Emerging features of mRNA decay in bacteria. *RNA* **6**, 1079–1090.

Taniguchi, T., Palmieri, M. & Weissmann, C. (1978). Q_β DNA-containing hybrid plasmids giving rise to Q_β phage formation in the bacterial host. *Nature* **274**, 223–228.

Valegård, K., Liljas, L., Fridborg, K. & Unge, T. (1990). The three-dimensional structure of the bacterial virus MS2. *Nature* **345**, 36–41.

van Duin, J. (1988). Single-stranded RNA bacteriophages. In *The Viruses*, pp. 117–167. Edited by H. Fraenkel-Conrat & R. Wagner. New York: Plenum Press.

van Meerten, D., Zelwer, M., Régnier, P. & van Duin, J. (1999). *In vivo* oligo(A) insertions in phage MS2: role of *Escherichia coli* poly(A) polymerase. *Nucleic Acids Research* **27**, 3891–3898.

Witherell, G. W., Gott, J. M. & Uhlenbeck, O. C. (1991). Specific interaction between RNA phage coat proteins and RNA. *Progress in Nucleic Acid Research and Molecular Biology* **40**, 185–220.

Zuker, M., Mathews, D. H. & Turner, D. H. (1999). Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In *RNA Biochemistry and Biotechnology*, pp. 11–43. NATO ASI Series. Edited by J. Barciszewski & B. F. C. Clark. Amsterdam: Kluwer Academic.

Zwick, M. B., Shen, J. & Scott, J. K. (1998). Phage-displayed peptide libraries. *Current Opinion in Biotechnology* **9**, 427–436.