Translational Fusions with Fragments of the \textit{trpE} Gene Improve the Expression of a Poorly Expressed Heterologous Gene in \textit{Escherichia coli}

By \textsc{Andrew Makoff,* Nigel Parry and Lynn Dicken}

Wellcome Biotech Ltd, Langley Court, Beckenham, Kent BR3 3BS, UK

(Received 4 May 1988; revised 1 August 1988; accepted 14 September 1988)

A series of plasmids expressing fusions between the \textit{trpE} gene product, anthranilate synthase component I and the major immunogen (VP1) of foot and mouth disease virus were constructed such that increasing amounts of the 3' end of \textit{trpE} were deleted. Deletions removing up to 70\% of \textit{trpE} had little effect on the quantity of fusion protein expressed, while the number of molecules appeared to increase. Larger deletions led to a steady decrease in both the quantity of fusion protein produced and in the number of molecules. This is consistent with \textit{trpE} being responsible for the high levels of expression of VP1 by its gene product stabilizing VP1 against proteolytic degradation. Some out-of-frame deletion mutants were also produced. All deletion mutants in one wrong reading frame expressed low levels of two VP1-containing polypeptides. This observation is interpreted as being due to re-initiation of translation at a site inside the VP1 sequence which is activated by local termination of translation.

\textbf{INTRODUCTION}

High level expression of proteins in \textit{Escherichia coli} depends on a number of factors: (a) an efficient promoter; (b) absence of internal transcriptional termination signals; (c) a stable transcript; (d) efficient initiation of translation; (e) efficient elongation of translation and (f) a stable product. When the coding sequence of a well-expressed protein in an expression vector is replaced by the coding sequence of the foreign protein to be expressed, only factor (a) is unaltered. Consequently poor expression can be due to any combination of the other five factors. If the coding sequence of a poorly expressed foreign protein is placed downstream of the coding sequence of a well-expressed protein to generate a translational fusion to its C terminus, the RBS responsible for the efficient translation initiation of the well-expressed protein is preserved. In such a construction, since \textit{in vivo} degradation of transcripts appears to be in the 3' to 5' direction (von Gabain \textit{et al.}, 1983; Mott \textit{et al.}, 1985) the presence of the mRNA of the well-expressed protein at the 5' end of the mRNA of the foreign protein is unlikely to stabilize it against nuclease degradation. In contrast, the presence of the highly expressed stable polypeptide in a fusion with the foreign protein is likely to stabilize it against protease degradation as was seen with the expression of somatostatin (Itakura \textit{et al.}, 1977). Consequently if expression of the foreign polypeptide is improved by such a construction, as is the case with VP1, then it is likely to be due to changes in factors (d) and/or (f).

Repeated attempts to detect expression of VP1 in \textit{E. coli} as a mature protein from two different constructs have been unsuccessful (unpublished observations). While there are published examples of expression of VP1 as a fusion protein there are no examples of its expression as a mature protein (Kleid \textit{et al.}, 1981; Kupper \textit{et al.}, 1981). We have found that it is very well expressed as a fusion to the \textit{trpE} gene product, AS1. Fusing the VP1 sequence to \textit{trpE} ensures that the relatively efficient \textit{trpE} RBS is utilized and that VP1 will be efficiently translated.
translated. It also places VP1 on a protein complex which may fold in such a way as to protect VP1 from otherwise rapid degradation.

In order to determine whether more efficient translation initiation or reduced protein turnover were responsible for the improved expression of VP1, a series of deletion plasmids which varied in trpE content were constructed. If VP1 needed AS1 to protect against degradation one might expect that it would require a substantial portion of this protein to demonstrate increased stability. Furthermore, the exact fusion point might be critical and the addition or removal of even a single residue might dramatically affect the stability of the fusion protein. If, however, trpE was acting by providing an intact homologous RBS, then very little of the trpE coding sequence would be required, as only approximately 20 nucleotides either side of the initiation codon appear to participate in initiation of translation (Steitz, 1969; Scherer et al., 1980; Stormo et al., 1982). (For reviews see Steitz, 1979 and Gold et al., 1981.) The experience with expression of bovine growth hormone is consistent with this view (Scherer et al., 1984).

METHODS

Strains. Two strains of Escherichia coli were used: MM294 [F- endA1 hsdR17(rk-, mk-) supE44 thi-1 lacI 17-1] (Meselson & Yuan, 1968); Y1083 [ΔlacIP07A] U169 proA+ Δlon araD139 strA thi hflA50 (chr: Tnl0)] (Young & Davis, 1983).

Plasmids. pFO, trp7 (Fig. 1) has trpE under trp promoter control, cloned between the EcoRI and HindIII sites of pAT153 (Twigg & Sherratt, 1980) with the 13 codons at the 3' end replaced by the VP1 coding sequence from FMDV strain O, BFS (Makoff et al., 1982). The plasmid excludes the five codons of VP1 5' to the HindIII site and includes 15 codons of p52, the adjacent gene in the FMDV genome. In addition, the BamHI site in the tetracycline resistance gene was removed by the Klenow fragment of DNA polymerase I (Biolabs) which generated a ClaI site in its place.

Deletion of part of trpE. This was achieved in three ways. In the first method, plasmid pFO, trp70 was constructed from pFO, trp7 by ligating the sequence 5' to the first BglII site in trpE to the DNA 3' to the BamHI site in the VP1 coding sequence, after filling in the cohesive ends with the Klenow fragment of DNA polymerase I. This construction would be predicted to delete a further three codons from the 5' end of the VP1-coding sequence and maintain it in the same reading frame as trpE.

The second procedure is outlined in Fig. 1. Plasmid pFO, trp7 was linearized with NdeI and then treated with the exonuclease Bal31 (Biolabs) over a tenfold range of incubation times. The digestion was done at 30°C in a high salt buffer (12 mM-CaCl2, 12 mM-MgCl2, 600 mM-NaCl, 20 mM-Tris/HCl, pH 8.0, 1 mM-EDTA) with 0.8 units enzyme and 10 μg DNA in a total volume of 100 μl. Every minute after the incubation had started, 10 μl of the reaction mixture was added to 12 ml of 0.8 mM-EDTA on ice in two pools (25 mM, final concentration), corresponding to 1–5 min (pool 1) and 6–10 min (pool 2). Electrophoresis of samples of the two pools on an agarose gel confirmed that the DNA had been shortened by the procedure and had become heterogeneous in length (data not shown). The two pools of DNA were then digested by BamHI, incubated with Klenow fragment in order to blunt the ends and self-ligated in order to recircularize the molecules. The two pools of DNA were used to transform E. coli cells (strain MM294) and the resulting colonies screened for expression of VP1-containing polypeptides by a solid phase immunoscreen.

This procedure would be expected to produce a heterogeneous population of plasmids which, provided that the nuclease did not digest beyond the trpE sequence, would have an intact VP1 coding sequence, one out of three of which would be in the same reading frame as trpE. The new junction would always be a Sau3A site (GATC) and one in four would also be a BamHI site (GGATCC). One-hundred colonies transformed with DNA from both pooled exposures to Bal31 were screened for production of VP1. As a positive control, cells transformed with pFO, trp7 were included in the screen. Cells transformed with pFO, trp80, made from pFO, trp7 as above but with no nuclease digestion (predicted to have VP1 translated in a wrong reading frame) were included in the screen as a negative control. Of the colonies derived from pool 1 DNA, 21 were positive as compared with only nine from pool 2 (data not shown). Restriction mapping of the plasmid DNA showed that of the plasmids derived from pool 1 DNA, 13 appeared to be identical to pFO, trp7, one appeared to have lost some of the VP1 coding sequence as well as trpE, and the remaining seven (pFO, trp71–77) appeared to be the expected trpE deletions (data not shown). Of the nine positive colonies from pool 2, plasmids from eight appeared to be pFO, trp7 and only one appeared to have a trpE deletion (pFO, trp78).

The third series of deletion plasmids was constructed in a similar way to pFO, trp70 in that a restriction enzyme site in trpE was used in conjunction with the BamHI site of the VP1 sequence to delete a defined length of the trpE gene. In these constructions a BamHI linker (Biolabs) was used to join the two sequences. In order to keep VP1 in the same reading frame as trpE the linker CGGATCCG was used with the sites AhaII, BglII and AccI to
Expression of fusion proteins in E. coli

give plasmids pF0, trp750, 752 and 754 respectively (see Figs 1 and 3). For plasmid pF0, trp753 the linker CGCGGATCCGCG was used with the Trhi111 site of trpE.

Screen for plasmids expressing VP1. The method used is a modification of that first described by Broome & Gilbert (1978). Colonies were grown in triplicate: two sets on nitrocellulose discs (Schleicher and Schüll) on plates containing 1-5% (w/v) agar in M9 medium (Maniatis et al., 1982) containing 40 µg tryptophan ml⁻¹ in order to repress the trp promoter and one set on reference plates containing L broth, agar and ampicillin. After overnight incubation at 37 °C, the discs were transferred to fresh plates containing M9 medium but with 10 µg IAA ml⁻¹ in place of tryptophan in order to induce the trp promoter. After further incubation for 3 h, the nitrocellulose discs were removed and the cells were lysed by chloroform vapour for approximately 30 min. The discs were then blotted with Whatman 3MM filter paper to remove most of the cell debris and then treated as in the Western blot procedure described below.

Induction and analysis of products. Cultures from fresh colonies were grown overnight at 37 °C in M9 medium in shake flasks; 2 ml of this overnight culture was then diluted to 10 ml with pre-warmed M9 medium containing IAA at a final concentration of 10 µg ml⁻¹. After incubation for 3–4 h, the equivalent of 1 ml of cells at an OD₆₅₀ of 1-0 were resuspended in 100 µl Laemmli loading buffer and 10 µl of each sample (approximately 20 µg total cell protein) were run on an 11.5% (w/v) polyacrylamide gel (Laemmli, 1970). The gel was analysed either by staining of the proteins by Coomassie Blue (Fluka) or by Western blotting (Towbin et al., 1979). The VP1-containing proteins were detected using rabbit antisera raised against four successive inoculations of 20 µg gel-eluted VP1 (FMDV strain O, BFS 1860/67) and ¹²⁵I-labelled protein A (Pharmacia).

DNA sequencing. The region of the deletions in most of the plasmids described in this paper was sequenced. The appropriate EcoRI–NarI fragment from each plasmid (see Fig. 4) was subcloned between the EcoRI and AecI sites of M13mp18 (BRL) and sequenced by the chain termination method (Sanger et al., 1977).

Assay of VP1. The VP1 content of induced extracts was estimated by an indirect ELISA. The extracts in Laemmli loading buffer were serially diluted in carbonate coating buffer (15 mM-Na₂CO₃/35 mM-NaHCO₃) and adsorbed to flexible PVC microplates overnight at 4 °C. The plates were washed in PBS (Dulbecco's phosphate-buffered saline ‘A’; Oxoid) containing 0-05% Tween 20, blotted dry and 50 µl antisera diluted in this buffer plus 1% (w/v) bovine serum albumin was added to each well. After 1 h at 37 °C, the plates were washed and dried and 50 µl diluted goat anti-rabbit peroxidase conjugate (Miles Laboratories) was added. After further 1 h the plates were again washed and dried and 50 µl o-phenylenediamine peroxide substrate was added. After colour development the reaction was stopped by the addition of 12-5% (v/v) sulphuric acid and the A₄₉₀ read. The VP1 content was estimated by comparison with an extract of a known concentration of purified viral particles (FMDV strain O, BFS 1860/67), disrupted by boiling in Laemmli loading buffer and treated in the same way. Pre-immune rabbit serum and E. coli extracts from a culture which did not express VP1 were also included as controls for non-specific binding.

RESULTS

Analysis of colonies positive on the immuno-screen

Induced cultures of cells carrying plasmids pF0, trp7, 70 and the eight plasmids which were identified by the antibody screen as expressing VP1 (pF0, trp71–78; see Methods and Fig. 1) were analysed by gel electrophoresis (Fig. 2a). The levels of expression of fusion proteins by cells carrying pF0, trp70 and 71–75 were all very high and compare well with the expression of fusion proteins by pF0, trp7. By contrast, induced cultures of pF0, trp76–78 expressed fusion protein at much lower levels (Fig. 2a). These cultures seemed to be much more variable in their levels of expression. A Western blot of a similar gel is shown in Fig. 2(b) and confirms that the expressed fusion proteins contain VP1. In all cases the major band was accompanied by many fainter bands. The total protein in these fainter bands was almost certainly very much less than in the corresponding major band, and the relative proportions of protein in the minor bands and in the major band was about the same for each plasmid. This is clear when examining different exposures of the autoradiograph of the Western blot of Fig. 2(b) (data not shown). There are several possible ways in which these minor bands could have arisen. They could have arisen as a result of initiation or termination of transcription or re-initiation of translation at alternative sites within the coding sequence for the fusion protein. Some are probably breakdown products due to proteolysis, but any fragments lacking VP1 would not have been detected.

Plasmids pF0, trp71–78 were partially sequenced as described in Methods. Fig. 3 shows the sequence of trpE (Yanofsky et al., 1981) up to the NdeI site (see Fig. 1), on which is indicated where the VP1 coding sequence joins the remaining trpE sequence for each plasmid. From the
sequence the predicted size of each expressed fusion protein was calculated and presented diagrammatically in Fig. 4, along with the fusion proteins from pFO, trp7 and pFO, trp70 for comparison. The sizes of the fusion proteins predicted from the nucleotide sequence are entirely consistent with the observed sizes of the stained bands shown in Fig. 2(a). In pFO, trp78, Bal31 had evidently removed all of trpE and some of the upstream sequence as well, leaving the VP1 sequence fused to the third codon of the trpL sequence. Because of this, it was not analysed further.

**Construction of further trpE deletions**

As can be seen in Fig. 2(a), the level of expression of fusion protein decreases sharply between pFO, trp75 and pFO, trp76. Since, as shown in Fig. 3, the deletion in pFO, trp76 is 273 nucleotides larger than in pFO, trp75, it was decided to construct plasmids with deletions of intermediate size. This was done as described in Methods to give plasmids pFO, trp750, 752, 753 and 754 (see Figs 1, 3 and 4).

 Cultures of *E. coli* strain MM294 transformed with these plasmids were induced and analysed as before. Fig. 5 shows a stained polyacrylamide gel which compares the levels of expression from the four new plasmids with that from pFO, trp74–77. Cells containing pFO, trp750 express a fusion protein at levels comparable to cells containing pFO, trp70–75. Cells containing the other three new plasmids express the fusion proteins at levels intermediate between this and cells containing pFO, trp76 and 77.
Fig. 2. Induction of fusion proteins by plasmids with different amounts of the trpE gene deleted. Samples were run on an 11.5% (w/v) SDS-polyacrylamide gel. Molecular masses of protein markers are given in kDa. (a) Gel stained with Coomassie Blue. Where the fusion protein band is partly obscured by a host band, its position is indicated by an arrow. (b) Gel was blotted onto nitrocellulose and incubated with antiserum raised against VPI followed by 125I-labelled protein A.
Expression of fusion proteins in E. coli

The results with the plasmids which were positive on the antibody screen indicate a steady decrease in expression levels as a function of increasing size of deletion. In order to investigate the possibility that some poorly expressing plasmids might have been missed by the screen, some of those which were negative were also examined. Induced extracts from 34 of these were

Analysis of colonies negative on the immuno-screen

The results with the plasmids which were positive on the antibody screen indicate a steady decrease in expression levels as a function of increasing size of deletion. In order to investigate the possibility that some poorly expressing plasmids might have been missed by the screen, some of those which were negative were also examined. Induced extracts from 34 of these were
analysed as before. The majority of these exhibited no obvious stained band on a polyacrylamide gel (data not shown). Gel analyses of four of the five which did (pFO, trp81–84) and also pFO, trp80, referred to earlier, are shown in Fig. 2(a). All five plasmids expressed high levels of a protein ranging in size from 14 to 26 kDa, consistent with translation of a fusion between part of trpE and the VP1 coding sequence in the wrong reading frame. This was confirmed by sequencing (see Figs 3 and 4). The Western blot showed, as expected, that none of these fusion proteins recognized antisera against VP1 (Fig. 2b). Surprisingly, three of these plasmids (pFO, trp81, 83 and 84) also expressed a doublet of approximately 21 and 20 kDa which did recognize the antisera. The same doublet could be seen faintly in all the other tracks if the Western blot autoradiograph was exposed for a much longer time (data not shown).
Expression of fusion proteins in E. coli

Fig. 6. (a) Expression of VP1 as a percentage of total cell protein plotted against the number of amino acids of AS1 remaining for each fusion protein. Each point was derived from the ELISA data by a linear regression of the linear part of the mean of two independent titrations curves (see Methods). The standard errors for each point are indicated. (b) Expression of fusion protein as a percentage of total cell protein. Each point was calculated from the corresponding point in (a) by dividing each yield of VP1 by the appropriate proportion of VP1 in the fusion protein.

Interestingly, the three plasmids which expressed the doublet have the VP1 coding sequence in one wrong reading frame while the other two plasmids have the other. The significance of this finding is discussed later. Of the remaining 30 negative colonies, only extracts of induced cultures from one (pFOl trp830) gave an obvious band on a polyacrylamide gel (see below) and none gave any band on a Western blot.

A further eight plasmids obtained from the 34 colonies which were negative on the antibody screen were analysed by restriction mapping. One of these (pFOl trp830) mapped between pFOl trp74 and 75. Since it expresses an induced protein of approximately 16 kDa which does not appear on a Western blot and has a deletion of approximately 1120 nucleotides deduced by restriction mapping (data not shown), its VP1 sequence must be translated in the same wrong reading frame as pFOl trp80 (see Fig. 4). Another (pFO1 trp85) mapped between pFOl trp752 and 753 and its exact position could be deduced by the fortuitous creation of a BspMII site (TCCGGA) (data not shown) and is indicated in Fig. 3. This plasmid would be predicted to have its VP1 sequence also translated in the same reading frame as pFOl trp80 and to express a fusion protein of 9.1 kDa, which would be difficult to visualize on a stained polyacrylamide gel. The remaining six plasmids have deletions into the VP1 sequence, into the trp promoter region, or both.

Estimates of expression levels

In order to obtain a quantitative comparison of the levels of expression of fusion protein by the deletion mutants, extracts of cultures were assayed for VP1 using an ELISA. Using these results and assuming a value of 200 μg total cell protein ml⁻¹ for cells at an OD₆₅₀ of 1, the levels of both VP1 and fusion protein as a percentage of total cell protein were calculated and are displayed graphically in Fig. 6. Fig. 6(a) shows initially an increase in levels of VP1 followed by
a decrease as increasing amounts of trpE are deleted. The levels of fusion protein (Fig. 6b) show initially a gradual reduction followed by a more rapid reduction as more than 70% of the trpE gene is deleted. A similar pattern was observed when the fusion protein was estimated by gel scans of each lane of a stained polyacrylamide gel (data not shown).

The ELISA estimates all fragments containing VP1 and therefore will include all those smaller bands seen in Fig. 2(b). This will overestimate the amount of VP1 present in each intact fusion protein. However, we judge the overestimate to be slight, since the minor bands represent much less protein than the major band; in any event they will barely affect the relative estimates since the minor bands together appear to be in approximately the same ratio to the major band for each extract in Fig. 2(b).

**Expression of the fusion proteins in a lon strain**

*E. coli* rapidly turns over abnormal proteins, including many heterologous proteins (Lin & Zabin, 1972). One degradative system involved in this has been characterized and involves the protease La which is encoded by the lon gene (Chung & Goldberg, 1981). Strains lacking this gene have been shown to stabilize certain unstable polypeptides (Bukhari & Zipser, 1973; Gottesman & Zipser, 1978). In order to test the possibility that proteolysis involving La was responsible for the lower levels of expression of the smaller fusion proteins, a number of the plasmids described above were transformed into the lon strain Y1083 and the cells were induced and extracts analysed as before. All of the plasmids including the poorly expressing pFO, trp76–78 directed expression at comparable levels and with a similar pattern of anti-VP1-binding minor bands to those seen in strain MM294 (data not shown). Clearly if proteolysis is involved it appears to be independent of the lon system.

**DISCUSSION**

Proteins are frequently expressed as fusion proteins when reliably high expression of a gene is more important than the exact nature of the resulting polypeptide. While it is widely recognized that fusing a heterologous gene to a homologous gene such as trpE makes high expression more likely, it is not entirely clear how the homologous sequence aids expression or how much of it is required.

In the experiments described in this paper we have studied the effects on expression of large deletions in plasmids which express fusions with trpE. We have shown that deletions of up to 70% of the trpE gene lead to an increase in the expression of VP1 (and therefore in the number of molecules of fusion protein) and a slight decrease in the quantity of fusion protein. (This discrepancy arises because the increased expression of VP1 is in smaller molecules.) As the trpE gene is deleted beyond the 150 codons at the 5' end, both the level of VP1 and fusion protein decrease sharply.

This result can be viewed as the sum of two conflicting effects. The smaller effect, which is for the number of molecules of fusion protein to increase as the trpE content decreases could be due to several possible causes. It might be brought about if trpE contained weak transcriptional termination signals, discrete endonucleolytic cleavage sites in its mRNA or a high proportion of rare codons. None of these possibilities are very likely as trpE is the homologous gene. It is much more probable that the small increase in the number of fusion protein molecules as a function of decreasing trpE content is due to a saturation effect.

The more significant effect is for the level of expression to decrease as trpE is progressively deleted, and this decrease is most marked as the last 30% is removed. As was pointed out in the Introduction the most likely cause of improved expression of a fusion protein over a mature heterologous protein is either the presence of an intact homologous RBS or the greater resistance of the product to proteolysis. The largest deletions (excluding pFO, trp78) are in pFO, trp76 and 77 where respectively 120 and 105 nucleotides of the 5' end of trpE are retained. Even in these cases the trpE RBS is clearly intact. However, it is possible that the RBS is constrained by secondary structure in the more heavily deleted plasmids (Iserentant & Fiers, 1980; Hall et al., 1982). Using the rules of Tinoco et al. (1973), a computer-aided search for potential secondary
Expression of fusion proteins in *E. coli*

structure showed no such structure whose presence correlated with poor expression. Because of these considerations it is hard to see how the different *trpE* content of the plasmids described here could affect the RBS.

By contrast, the observed effect on expression of deleting increasing amounts of *trpE* is consistent with AS1 protecting VP1 from proteolysis. This conclusion could be confirmed by half-life studies of pulse-labelled fusion proteins produced by some of the plasmids described here. It is known that *E. coli* rapidly turns over abnormal proteins, although it is not known how they are distinguished from normal proteins (Lin & Zabin, 1972; Gottesman & Zipser, 1978). One protease involved in this is the protease Lα; this is known to be regulated by the *hptR* gene product, which also regulates the heat-shock genes. Production of abnormal proteins as well as elevated temperature induces this protease via the *hptR* gene. (For a review, see Goldberg & Goff, 1986.) Although we have shown that protease Lα is probably not involved in the suspected high turnover of some AS1–VP1 fusion proteins, it is likely that other proteases are also induced by the same stimuli. The use of a *hptR* strain with some of these plasmids might resolve this question.

Our results suggest that the N-terminal 100 to 150 amino acids of AS1 are crucial in protecting the molecule from proteolysis. Because of the steady decrease in expression with increasing size of deletion, it would appear that the rest of the molecule is unimportant in influencing this protection. However, it must be considered that the steady decrease in expression might result from the use of an expression screen to identify in-frame gene fusions. This is unlikely for two reasons. Firstly, the mutants pFO, trp750, 752, 753 and 754 were constructed according to the location of convenient restriction enzyme sites; yet all four express VP1 at levels which fit the same trend as seen with the other mutants. Secondly, as already indicated, 34 out of 79 colonies which were negative on the antibody screen were further investigated and none produced fusion proteins which would recognize antisera on a Western blot. Restriction mapping of plasmids from eight of these showed that six have deletions extending into the trp promoter and/or VP1 sequence and the other two are clearly out-of-frame fusions. (This high proportion of very large deletions would explain why the number of VP1-expressing mutants represented much less than the predicted one-third of the total.)

The results, therefore, suggest that the N-terminal part of AS1 is able to protect the whole fusion protein from degradation and this is unaffected by the amount of the C-terminal part of AS1 present in the fusion protein. It is most unlikely that the N-terminal part of the molecule is able to fold around the rest of the molecule if it can vary so much in length. It is more likely that the N-terminal part is a folding domain. This structure by itself may act as a signal, preventing recognition of the molecule by one or more proteins involved in an *E. coli* degradative pathway. It may act as a nucleation centre, increasing the probability of the rest of the molecule folding into a definite structure, thereby avoiding rapid degradation. Alternatively, it may interact with VP1 and mask potential cleavage sites or recognition sites for a factor like the eukaryotic ubiquitin (Hershko, 1983), the binding of which would target the molecule for proteolysis. The relevance of the VP1 part of the fusion proteins could be investigated by replacing it by other poorly expressed heterologous genes.

Of the seven plasmids where the VP1 sequence is in a different reading frame from the remaining *trpE* sequence, plasmids pFO, trp81, 83 and 84 express two VP1-containing polypeptides of approximately 21 and 20 kDa (Fig. 2b) in addition to the out-of-frame fusion proteins (Fig. 2a). These three plasmids all have their fusion proteins translated in reading frame II which terminates at UGA-95, while the other four use reading frame III which terminates at UGA-21 (see Fig. 7a). Because the three plasmids express the same two VP1-containing polypeptides, re-initiation of translation within the VP1-coding sequence is likely to be taking place. Fig. 7(a) shows the positions of all possible start codons and Shine–Dalgarno sequences (Shine & Dalgarno, 1974) in the VP1-coding sequence that could generate such polypeptides. In order to see if secondary structure could affect the function of the above potential re-initiation sites, the predicted mRNA sequence of the VP1 coding sequence in the region of these sites was searched for possible base-pairing in the same way as was described earlier. The most stable structure appears to be one which has the two hairpin loops shown in
Fig. 7. (a) Predicted RNA sequence encoding VPI for pFO, trp7 and all trpE deletion plasmids. All three translational reading frames are given and their stop codons are boxed. Potential re-initiation codons are underlined and their predicted amino acid given in capitals. The numbers underneath refer to the molecular mass in kDa of the predicted polypeptide synthesized from each re-initiation codon. The possible Shine–Dalgarno sequences upstream of these codons are indicated by.

(b) Predicted secondary structure for the RNA sequence in (a). The stop codons, potential re-initiation codons and Shine–Dalgarno sequences are indicated as in (a). The free energy for each hairpin loop is shown.
Expression of fusion proteins in E. coli

Fig. 7(b). If it could form, the predicted upstream hairpin loop would prevent re-initiation of translation from AUG-82 and GUG-94. Only in reading frame II would this secondary structure be disrupted by ribosomes terminating translation of the out-of-frame fusion protein at UGA-95. Re-initiation at these two sites would generate two polypeptides of approximately 22.4 and 21.8 kDa, consistent with the two bands in the Western blot of Fig. 2(b). A similar mechanism for re-initiation of translation has been described for the lacI gene by Cone & Steege (1985a, b).

We would like to thank Mr H. Spence for writing the computer program to predict RNA secondary structure, and Dr N. Fairweather for useful comments.

REFERENCES


Scherer, G. F., Walkinshaw, M. D., Arnott, S. & Morre, D. J. (1980). The ribosome binding sites recognized by E. coli ribosomes have regions with signal character in both the leader and protein coding segments. Nucleic Acids Research 8, 3895–3907.


Shine, J. & Dalgarno, L. (1974). The 3' terminal sequence of E. coli 16S ribosomal RNA: complem-


