The EMBO Journal vol.8 no.11 pp.3501 – 3510, 1989

Dual translational initiation sites control function of the λ *S* gene

Udo Bläsi, Kiebang Nam, Dieter Hartz[1], Larry Gold[1] and Ry Young

Department of Biochemistry and Biophysics, Biochemistry Building, Texas A and M University, College Station, TX 77843 and [1]Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309, USA

Communicated by V.Pirrotta

Lysis gene *S* of phage λ has a 107 codon reading frame beginning with the codons Met_1-Lys_2-Met_3. Genetic data have suggested that translational initiation occurs at both Met_1 and Met_3, generating two polypeptides, S107 and S105 respectively. We have proposed a model in which the proper scheduling of lysis depends on the partition of translational initiations between the two start codons. Here, using *in vitro* methods, we show that two stem–loop structures, one immediately upstream of the reading frame and a second ~10 codons within the gene, control the partitioning event. Utilizing primer-extension inhibition or 'toeprinting', we show that the two *S* start codons are served by two adjacent Shine–Dalgarno sequences. Moreover, the timing of lysis supported by the wild-type and a number of mutant alleles *in vivo* can be correlated with the ratio of ternary complex formation over Met_1 and Met_3 *in vitro*. Thus the regulation of the *S* gene is unique in that the products of two adjacent in-frame initiation events have opposing function.

Key words: lambda lysis/*sdi*/toeprinting/translation control

Introduction

The bacteriophage λ genome codes for three lysis genes: *S*, *R* and *Rz*, the first cistrons of the 28 kb transcript synthesized from the late promoter, $p_{R'}$ (Figure 1). The *Rz* gene, required for lysis in conditions that stabilize the bacterial envelope, has an unknown function (Young *et al.*, 1979). The *R* gene encodes a 17 500 M_r soluble monomeric transglycosylase (gp*R*) that degrades the peptidoglycan (Bienkowska-Szewczyk *et al.*, 1981). Lacking a signal sequence, gp*R* requires the function of gp*S*, a 9000 M_r inner membrane polypeptide (Altman *et al.*, 1985), to obtain access to the periplasm. Pleiotropic effects can be ascribed to *S*, including collapse of the membrane potential, disruption of active transport and an increase in permeability of the inner membrane (Adhya *et al.*, 1971; Reader and Siminovitch, 1971; Garrett and Young, 1982; Wilson, 1982). These physiological data and findings from a mutational analysis of *S* suggested that gp*S* oligomerizes to form a pore of sufficient size to permit transit of the transglycosylase activity across the inner membrane (Wilson, 1982; Raab *et al.*, 1986). More recently, direct molecular evidence for this model was obtained by the detection of SDS-resistant dimers and higher oligomeric forms of gp*S* (Zagotta, 1989).

The *S* gene and all λ late genes are separated from the strong constitutive late promoter $p_{R'}$ by the terminator $t_{R'}$, which prevents late gene expression during the early transcription phase (Herskowitz and Hagen, 1980). Expression of *S* commences at ~8–10 min, when the *Q* gene product begins anti-termination of the $p_{R'}$ transcript. No further regulatory events are known at the transcriptional level. Thus, *S* gene transcription is essentially constitutive over the last 30 min of the infective cycle. However, no effect of gp*S* is detectable until 40 min after infection, when the large excess of accumulated transglycosylase activity is released to the periplasm and lysis immediately ensues. These observations imply that *S* function is controlled post-transcriptionally. Genetic evidence has been presented which suggests a novel mode of translational regulation contributing to the 'lysis clock'. In a model devised to explain the dominant character of mutations immediately upstream from the *S* reading frame, it has been proposed that two translational starts are utilized (Raab *et al.*, 1988). The two starts, one codon apart, would lead to the synthesis of a 107 amino acid polypeptide (S107) and a 105 amino acid polypeptide (S105), differing only in the presence of a Met and a Lys residue at the amino terminus of the longer product (Figure 1). Moreover, the model specifies that the S107

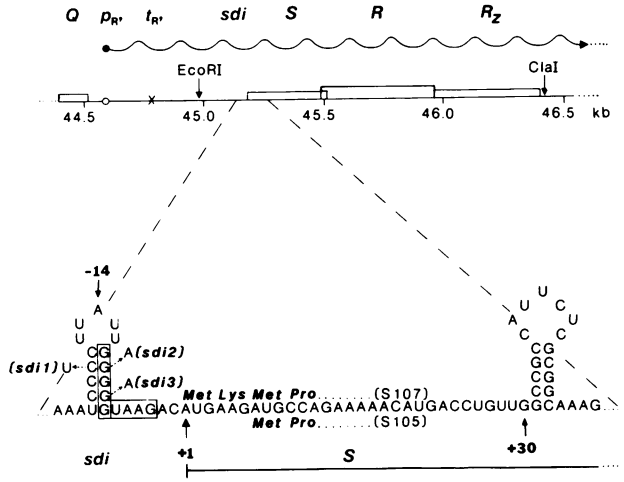

Fig. 1. Translational initiation region of the λ *S* gene. The promoter-proximal genes of the late operon of λ are shown. The bp numbers are taken from Daniels *et al.* (1983). The *Eco*RI–*Cla*I fragment containing the λ lysis genes *SRRz* has been cloned into the expression vector pBH20 (Itakura *et al.*, 1979), giving rise to plasmid pRG1 (Raab *et al.*, 1986). For this study the lysis cassette was reisolated on a *Eco*I–*Hin*dIII fragment from the latter plasmid. The translational initiation region of gene *S* is enlarged. +1, the first base of codon 1, corresponds to position 45 186 in the λ sequence (Daniels *et al.*, 1983). Sequences complementary to the 3′ end of the 16S rRNA are boxed. The positions and the base changes of the *sdi* mutations (Raab *et al.*, 1988) located in the *sdi* structure (*s*tructure *d*irected *i*nitiation) upstream of gene *S* are depicted. The N-terminal amino acid sequences of the two *S* gene products are given above (S107) and below (S105) the mRNA sequence.

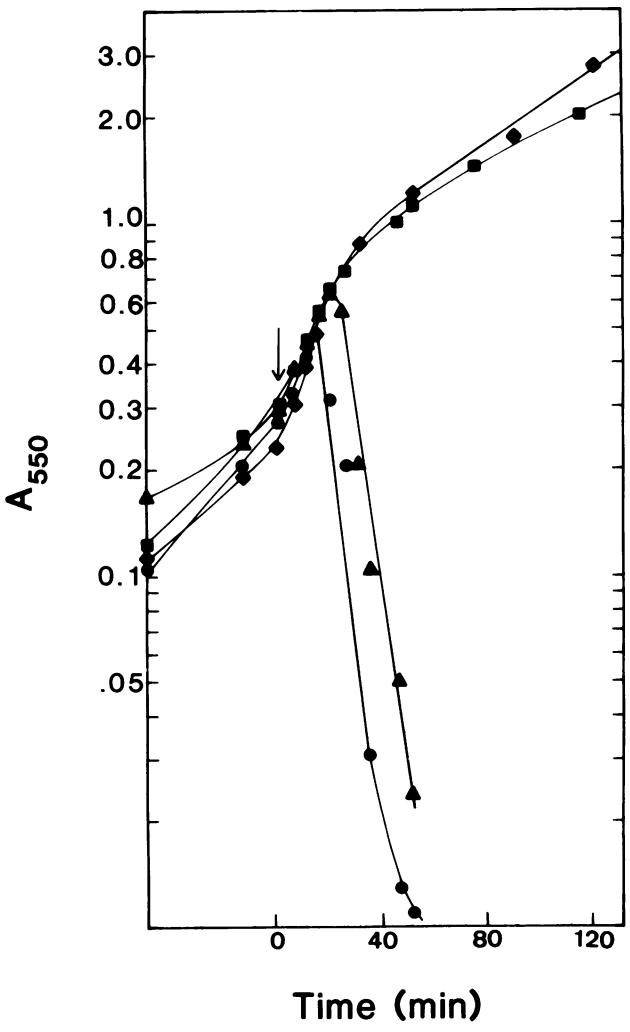

Fig. 2. Induction of S^+ and start codon mutant alleles. For description of the plasmids, see Table II. *E.coli* pop2135 carrying plasmids pLS105 (●; produces only S105), pLS107 (■; produces only S107), pLS157 (▲; wild-type allele, produces S105 and S107) and pLS130 (◆; produces neither S105 nor S107) respectively was grown as described in Materials and methods. At the times indicated by the arrow expression of the cloned lysis genes was induced by shifting the cultures to 42°C with continued aeration.

product would be defective in terms of lysis function and act as an antagonist for the S105 product, the real lysis effector. In addition, the model predicts that the ratio of S107 and S105 would significantly affect the scheduling of lysis (Raab *et al.*, 1988). A potential stem–loop structure in the mRNA upstream of the *S* reading frame (Figure 1) was implicated in the choice of initiation codons (Raab *et al.*, 1988). Single-base changes which destabilized the stem–loop structure (Figure 1) led to a significant delay of lysis time, implying an enhanced expression of the lysis defective S107 product. Therefore, the upstream stem–loop region, designed as the *sdi* (*s*tructure-*d*irected *i*nitiation) site, was postulated to sequester the putative Shine–Dalgarno sequence for the Met_1 start codon (Figure 1).

In the work presented here, the effects of altering the *sdi* structure and the *S* translational initiation region by site-directed mutagenesis are examined at the molecular level, using ribonuclease digestion, *lacZ* fusion analysis and extension inhibition analysis ('toeprinting').

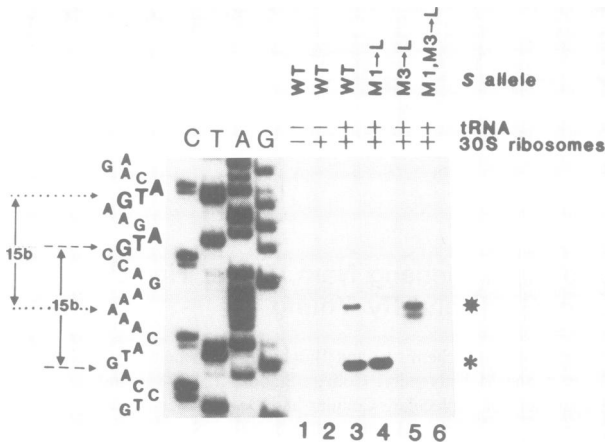

Fig. 3. Toeprint analysis of the S^+, $M_1 \rightarrow L$, $M_3 \rightarrow L$ and $M_{1,3} \rightarrow L$ alleles. Toeprinting was performed as outlined in Materials and methods. The DNA sequence of the 5'-flanking and initial coding sequences of the *S* gene are given at the left. Start codons 1 and 3 are in bold. **Lane 1**, primer extension reaction without addition of $tRNA^{fMet}$ and 30S ribosomes. **Lane 2**, primer extension reaction without addition of $tRNA^{fMet}$ but with addition of 30S ribosomes. Toeprint signals corresponding to initiation at start codon 1 (✱) and start codon 3 (*) obtained in the presence of $tRNA^{fMet}$ and 30S ribosomes are shown in **lanes 3–5**. **Lane 6**, toeprint performed on mRNA carrying the double null allele ($M_{1,3} \rightarrow L$).

Table I. Expression rates of SφlacZ fusions

Plasmid	*lacZ* fusion	β-Gal[a]
pSB87	SφlacZ	110
pSB88	S105φlacZ	70
pSB89	S107φlacZ	43
pSB90	S^{--}φlacZ	<3

Escherichia coli strain MC4100 bearing plasmids pSB87, pSB88, pSB89 and pSB90 respectively was grown in LB medium. Triplicate aliquots were taken of each culture at different densities during logarithmic growth. See Table III for details.
[a]β-Galactosidase values were determined as described by Miller (1972) and then averaged.

Results

Met_1 and Met_3 mutations have opposite effects on lysis timing

A key observation in the analysis of *S* regulation was the lysis-defective character of a mutation, $Met_3 \rightarrow$ Ile (Raab *et al.*, 1988). This finding and other data implicated codon 3 in translation initiation, but it was formally possible that the rare Ile codon involved, AUA, was affecting elongation (Königsberg and Godson, 1983). Moreover, since the mutational analysis was done using hydroxylamine mutagenesis, the possibility existed that cryptic lesions in the other lysis genes, *R* and *Rz*, contributed to the phenotype (Raab *et al.*, 1986). To eliminate these concerns, site-directed mutagenesis was employed to construct three new *S* alleles: $Met_1 \rightarrow$ Leu, $Met_3 \rightarrow$ Leu and $Met_{1,3} \rightarrow$ Leu, in each case with a CUG codon replacing the AUG. These alleles, and the wild-type cistron, were placed under p_L control along with the accessory lysis genes *R* and *Rz*, for visualization of the lytic event. Figure 2 shows that lysis is observed 20 min after induction of the wild-type *S* gene. Cells bearing plasmid pLS107, which carries the $Met_3 \rightarrow$ Leu change and

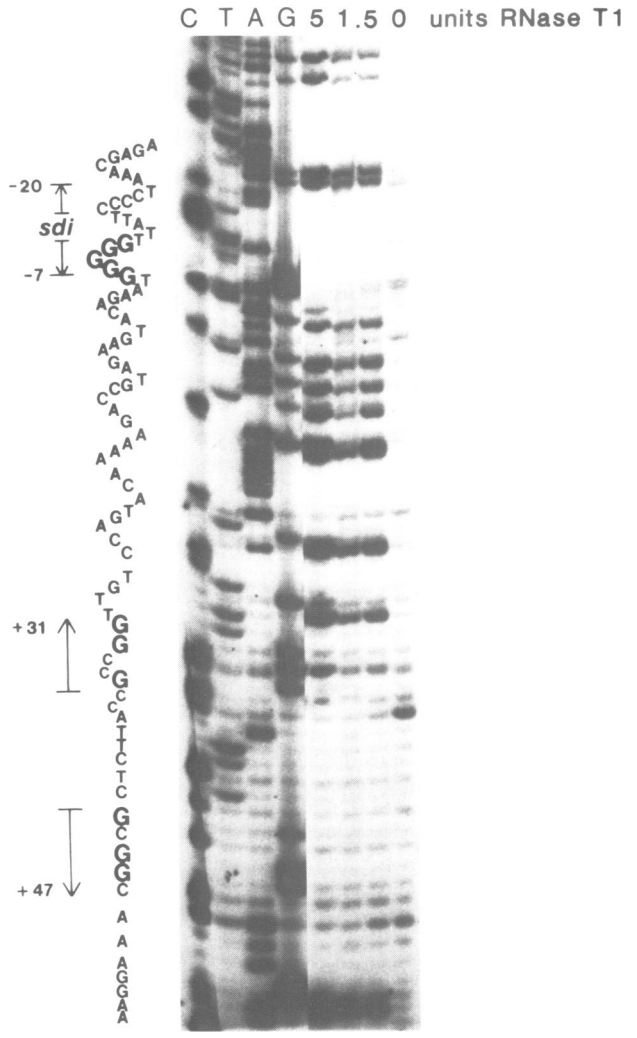

Fig. 4. Structural mapping of the gene *S* initiation region. Different concentrations of T1 ribonuclease were added to S^+ mRNA and the cleavage sites were then mapped using primer extension as described in Materials and methods. G residues inaccessible to RNase T1 are in bold. The positions of the *sdi* stem−loop (−20 to −7) and of the downstream stem−loop (+31 to +47) are shown by arrows.

thus produces only S107, do not lyse, nor do cells carrying plasmid pLS130, in which both start codons are eliminated. The Met_1 → Leu allele in plasmid pLS105, which produces only the S105 protein, caused lysis onset 15 min after induction, more rapidly than the wild-type (Figure 2). In addition, when the three *S* alleles were crossed back onto the λ genome, as described earlier (Raab *et al.*, 1988), similar results were obtained upon induction of the corresponding lysogens (not shown).

Codons 1 and 3 are both used for initiation in vivo and in vitro

To obtain direct evidence for the utilization of both Met_1 and Met_3 as start codons, toeprint analysis (Hartz *et al.*, 1988) was performed on RNA prepared *in vitro* from templates carrying the wild-type, Met_1 → Leu, Met_3 → Leu and $Met_{1,3}$ → Leu alleles. Toeprinting has been used to reveal the site of 30S ribosomal subunits bound *in vitro* in ternary complexes (Winter *et al.*, 1987; Hartz *et al.*, 1988; McPheeters *et al.*, 1988). It has been shown that in the presence of $tRNA^{fMet}$, 30S particles bound to a translation initiation site on a mRNA block cDNA synthesis by reverse transcriptase primed downstream. Primer extension is inhibited at the 3′ edge of the 30S subunit, generating a stop signal usually 15 nucleotides downstream of the first base of the initiation codon AUG (Hartz *et al.*, 1988). The intensity of the 'toeprint' signal reflects the strength of the corresponding ribosome binding site and is a measure of translational initiation *in vivo* (D.Hartz, unpublished).

Using the wild-type mRNA, two toeprint signals were observed 15 bases downstream of the adenosine of the AUG initiation codons for S105 and S107, with relative intensities of ~2.5:1, as judged by densitometry (Figure 3, lane 3). Toeprint analysis performed on mRNAs carrying either the Met_1 → Leu or the Met_3 → Leu allele revealed signals corresponding to initiation at codon 3 and codon 1 respectively (Figure 3, lanes 4 and 5). The failure of the double null allele ($Met_{1,3}$ → Leu) to give ribosome-dependent signals (Figure 3, lane 6) confirmed the specificity of the signals obtained for Met_1 and Met_3.

To determine whether both start codons are used *in vivo*, each of the *S* alleles was fused after codon 10 to *lacZ*. A fusion joint early in the *S* reading frame was chosen to avoid the putative membrane insertion signals of *S* (Young and Young, 1982), which might confer toxicity to the hybrid protein and thus obscure the quantitation of the translation products. The expression rates of the fusion genes were monitored by measuring the β-galactosidase activity of the corresponding hybrid proteins. Table I shows that β-galactosidase activity obtained from the wild-type fusion was approximately the same as the sum of activities obtained from the fusions to the Met_1 → Leu (S105φlacZ) and Met_3 → Leu (S107φlacZ) alleles. Synthesis from codon 3 is observed in approximately a 2:1 ratio compared to codon 1. The ratio of β-galactosidase values obtained for the S105 and S107 fusions parallels the toeprint signals observed in Figure 3, indicating that ternary complex formation *in vitro* reflects translational initiations at the two codons *in vivo*.

The sdi structure is protected from ribonuclease T1 attack

The predicted *sdi* stem−loop (Figure 1) implicated in the control of the translation initiation sites was tested for stability *in vitro*. Ribonuclease T1 digests were performed on wild-type *S* mRNA and cleavage sites were then mapped using primer extension. Figure 4 shows that G residues in the *sdi* region are completely protected, whereas neighboring G bases are clearly sensitive, indicating that the *sdi* structure exists primarily in a double-stranded conformation. Another region of T1 resistance is found between the G at +31 and the G at +47 (Figure 4). The potential stem−loop structure for this sequence has a predicted free energy of formation of −7.6 kcal/mol, similar to that of the *sdi* structure (ΔG = −6.3 kcal/mol; Freier *et al.*, 1986). The experiment indicates that two RNA stem−loop structures are stable features of the 5′ region of the *S* mRNA.

Destabilization of the sdi structure increases S107 expression

The mutations *sdi1*, *sdi2* and *sdi3* (Figure 1) were originally isolated as hydroxylamine-induced single base changes which reduced the lethality of the *S* gene cloned under *lacPO* control (Raab *et al.*, 1986). λ lysogens carrying these mutations recombined onto the phage genome showed a

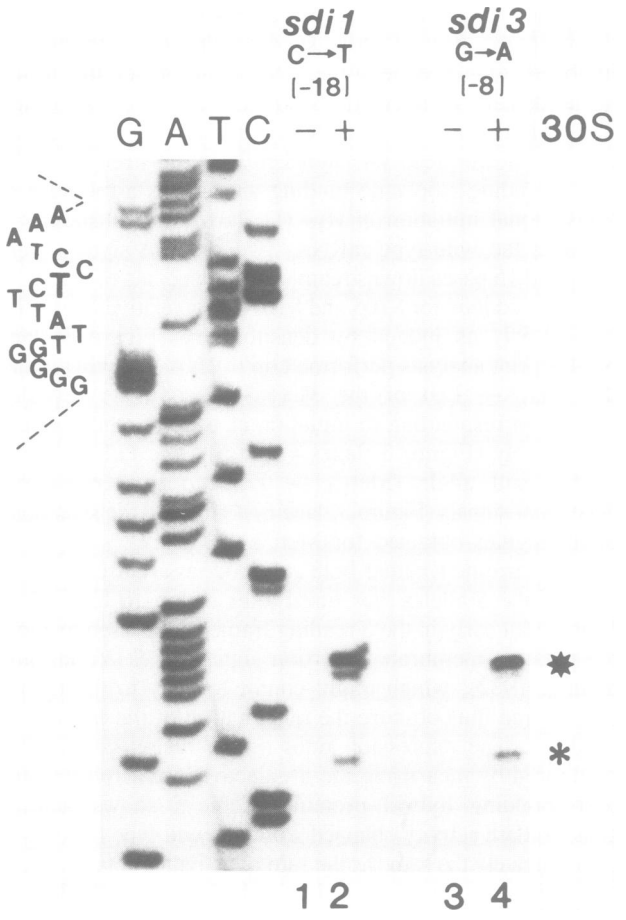

Fig. 5. Toeprint analysis of the *sdi* mutants. Primer extension reactions were performed in the absence (**lanes 1** and **3**) and in the presence of 30S ribosomes and $tRNA^{fMet}$ (**lanes 2** and **4**). The DNA sequence of the *sdi1* mutation is shown at the left. The C → T change at position −18 is indicated by a bold T. The toeprint signals corresponding to initiation at codon 1 (✱) and codon 3 (*) obtained with the *sdi1* and *sdi3* mutations are shown in **lanes 2** and **4** respectively.

strong delay in lysis onset after thermal induction (Raab *et al.*, 1988). The *sdi* mutations should destabilize the stem−loop structure, since the prediced free energy of formation for the *sdi1* and *sdi3* mutant structures can be calculated as −3.3 and −1.4 kcal/mol respectively (Freier *et al.*, 1986). This should result in enhanced translational initiation at Met_1, since the putative Shine−Dalgarno sequence, GGGGG, at position −11 to −7, would become more accessible. Toeprint analysis of mRNA carrying the *sdi1* and *sdi3* mutations shows that with both *sdi1* and *sdi3*, ternary complex formation at the Met_1 start codon is much more frequent than at Met_3 (Figure 5, lanes 2 and 4). In addition, it should be noted that not only is the toeprint signal increased for Met_1 but also the signal for Met_3 is decreased, suggesting that the conformation of the *sdi* structure plays a role in controlling initiation at both start codons. When these *sdi* alleles were placed under p_L control, no lysis was obtained after thermal induction, regardless of the length of incubation period (Table II).

It should be noted that the mutation in the *sdi3* allele changes not only the predicted stability of the *sdi* structure but also alters the Shine−Dalgarno sequence to GGAG, which is closer to the Shine−Dalgarno consensus sequence (Gold and Stormo, 1987). However, *sdi1*, which has the wild-type Shine−Dalgarno sequence, GGGGG, and *sdi3* show the same Met_1:Met_3 toeprint signal ratio (Table II). This result shows that the increase in ternary complex formation over Met_1 in both alleles is primarily due to the destabilization of the *sdi* structure and not to the 'better' Shine−Dalgarno sequence in *sdi3*.

Table II. Toeprint analysis and lysis times of *S* alleles

Mutation[a]	Plasmid[b]	Lysis time[c]	Total binding[d]	Ratio 107/105[e]	Excess S105[f]
WT	pLS157	20	95 (1.00)	0.42	1
M_1 → L	pLS105	15	72 (0.75)	0	1.9
M_3 → L	pLS107	(−)	31 (0.32)	−	−
$M_{1,3}$ → L	pLS130	(−)	ND	−	−
sdi1	pLS101	(−)	32 (0.33)	2.70	−
sdi3	pLS99	(−)	74 (0.77)	2.70	−
sdi inv.	pLS111	18.5	83 (0.87)	0.27	1.1
G_{-3} → A	LS104	35	42 (0.44)	0.73	0.2
TAAG → GAGG	pLS115	10	81 (0.85)	0	2.1
C_{34},A_{48} → G	pLS118	30	37 (0.38)	0.28	1.4
C_{34},A_{48} → G					
G_{-3} → A	pLS123	50	23 (0.24)	0.80	0.1
T_{-21} → C	pLS124	15	79 (0.83)	0.24	1.2
C_{36} → G; A_{48} → C	pLS125	25	79 (0.83)	0.65	0.41

[a]Mutations in the translational initiation region of gene *S* (see text).
[b]Plasmids of the pLS series carry the λ lysis cassette, consisting of the corresponding *S* allele along with the accessory lysis genes *R* and *Rz* under p_L control.
[c]Cultures were grown at 30°C to an optical density of 0.2 and then shifted to 42°C. Lysis time is defined as the time (min) between shift to 42°C and onset of lysis. (−) signifies no lysis.
[d]The total binding parameter was determined as follows: after densitometry of each toeprint lane, the sum of S105 and S107 signals was divided by the sum of the read-through, S105 and S107 signals. This was taken as a measure of the total binding of the mRNA in 30S ternary complexes, expressed as percentage (%) of the total mRNA. The values given in parentheses are normalized to the value obtained for the wild-type. The value for the $M_{1,3}$ → L mutation was not determined (ND).
[e]The values were calculated from densitometric traces of S107 (Met_1) signal divided by S105 (Met_3) signal.
[f]Values were calculated by subtracting the S107 signal from the S105 signal and dividing the result by the total of the S105, S107 and read-through signal. The values were then normalized to the value obtained for the wild-type.

Distinct Shine−Dalgarno sequences direct initiation at Met_1 and Met_3

To verify that the GGGGG sequence at position −7 to −11 serves as the Shine−Dalgarno site for translational initiation at Met_1, the *sdi* stem−loop structure was inverted by site-directed mutagenesis. The inversion generates a spacing of 16 nt between the GGGGG sequence and Met_1, a spacing which is considered unfavorable for simultaneous annealing of the Shine−Dalgarno sequence to the 3′ end of the 16S rRNA and interaction of Met_1 codon with the anticodon of the initiator tRNA (Gold, 1988). Compared with the wild-type, the Met_1:Met_3 toeprint signal ratio in the *sdi* inversion is significantly decreased (Figure 6A), and the change is due to a reduction in the signal for Met_1, indicating the GGGGG sequence at position −11 to −7 functions as a Shine−Dalgarno site only for S107. The inversion mutation was placed under p_L control in plasmid pLS111. Onset of lysis was faster with the inversion mutation than with the wild-type (Table II).

The sequence UAAG, complementary to the extreme 3′ end of the 16S rRNA, is located in the upstream sequence

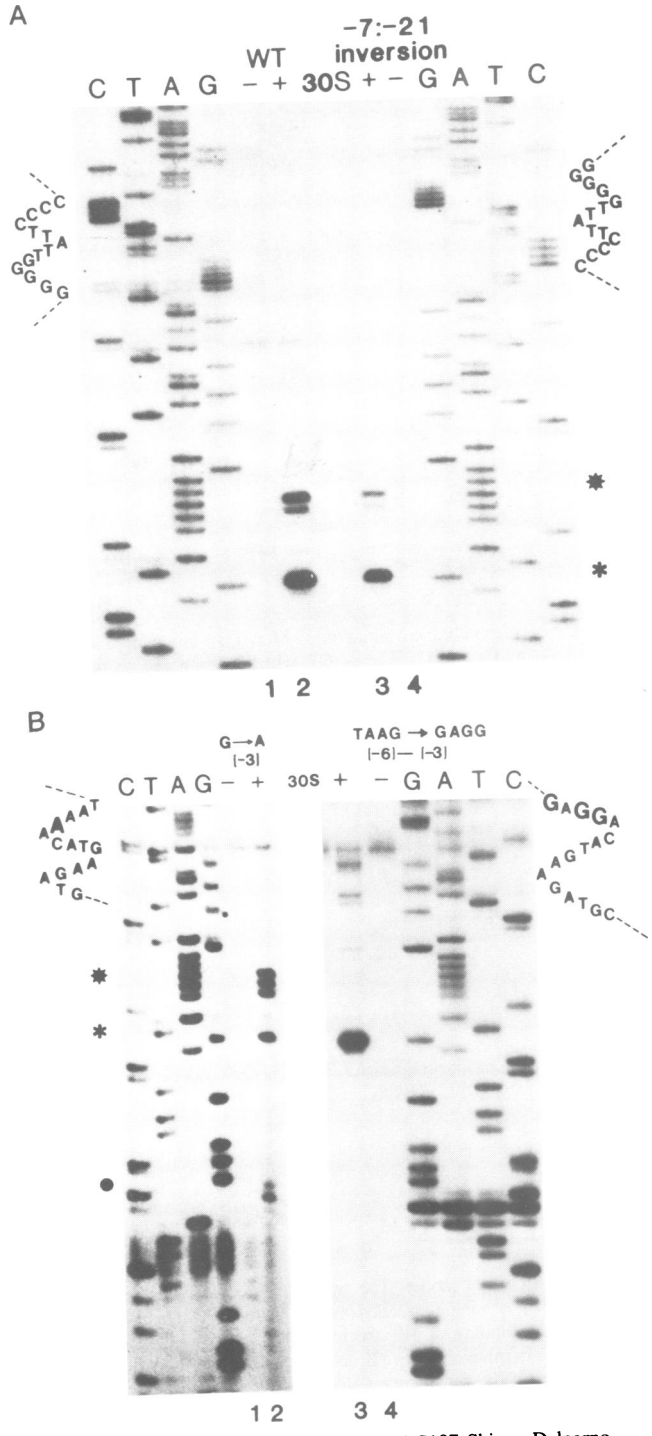

Fig. 6. Toeprint analysis of the S105 and S107 Shine−Dalgarno mutants. (**A**) The toeprints corresponding to initiation at codon 1 (✱) and codon 3 (*) obtained with mRNA carrying the inverted *sdi* structure are shown in **lane 3**. **Lane 4**, no 30S ribosomes or $tRNA^{fMet}$ added to the primer extension reaction. The DNA sequence of the *sdi* inversion is shown at the right. The toeprint signals obtained for codon 1 and codon 3 (**lane 2**) of the wild-type mRNA and the wild-type DNA sequence are shown at the left for comparison. (**B**) Toeprinting performed on mRNA carrying the G → A mutation at nucleotide −3 (**lanes 1** and **2**) and the TAAG → GAGG change in the −6 to −3 region (**lanes 3** and **4**). Signals corresponding to initiation at codon 1 (✱) and codon 3 (*) are shown in **lanes 2** and **3** respectively. **Lanes 1** and **4**, no 30S ribosomes added. (●) toeprint signal corresponding to an out-of-frame AUG at position +19 to +21. The DNA sequence of the G → A mutation at −3 is given at the left. The sequence of th TAAG to GAGG change is shown at the right. Base changes are in bold.

at position −6 to −3 nt relative to Met_1 (Figures 1 and 8). The spacing between this sequence and the third codon is 8 nt, within the boundaries of an optimal spacing (Gold, 1988). Changing the G at −3 to an A by site-directed mutagenesis eliminated the strongest base-pairing possible between the mRNA and the 3′ end of the 16S rRNA. Induction of this allele resulted in a significantly delayed lysis time (Table II), as would be expected if Met_3 starts are reduced. The toeprint pattern (Figure 6B, lane 2) supports this interpretation, with both the total ribosome binding at Met_1 and Met_3 and the fraction bound to Met_3 greatly reduced (Table II). Additional toeprint signals are also observed at +16 and +17 nt from the Met_1 codon, and a weaker signal corresponding to a downstream, out-of-frame AUG codon at nt 19−21. Assuming the extra signals are generated by the same ribosome-binding geometry, it seems unlikely that any of these extra out-of-frame starts could result in the synthesis of a functional *S* polypeptide *in vivo*. In support of this rationale, it should be noted that although lysis is delayed with this allele, no lysis would be expected at all if the +16 and +17 signal reflected additional Met_1 starts *in vivo*, since this would make the Met_1:Met_3 ratio >1 (see Discussion). The opposite effect was observed when we replaced the 'weak' Shine−Dalgarno domain UAAG with the more canonical sequence GAGG. This alternative Shine−Dalgarno sequence differs in two respects from the UAAG sequence: firstly, it should interact more strongly with the 3′ end of the 16S rRNA, and, secondly, the predicted complementarity will be to nt 1536−1539 (CUCC) of the 16S rRNA, as is usually the case (Gold and Stormo, 1987) rather than to nt 1539−1542 (CUUA), the extreme 3′ terminus of the mature rRNA (Figure 8). A strong toeprint signal was obtained corresponding to ternary complex formation over Met_3 (Figure 6B, lane 3). No toeprint signal was observed for Met_1. Onset of lysis directed by a plasmid (pLS115) bearing this mutation started 10 min earlier than with wild-type (Table II). We conclude that the UAAG sequence serves as the Shine−Dalgarno sequence for Met_3. Also, its base-pairing with the extreme 3′ end of the 16S rRNA seems to be required for the usage of Met_1 as a start codon.

Both the sdi and the downstream stem−loop are involved in positioning of the initiation complex

It has been suggested that secondary structures upstream or downstream of a gene start might serve to facilitate or to prevent translation initiation at a target ribosome binding site (Stanssens *et al.*, 1985; Gold, 1988). To test the notion that the downstream stem−loop structure (Figure 1, nt 31−47) is involved in the ribosome-binding pathway, silent mutations were created at positions +33 (C → G) and +48 (A → G) to destabilize the secondary structure without changing the amino acid sequence. Also, the new cognate codons were chosen such that the tRNA availability, and thus the decoding frequency, was approximately the same as for the original codons (Grantham *et al.*, 1980). The toeprint signals for this mutant allele showed a strong decrease in overall ribosome binding (Figure 7, lane 1; Table II), indicating that the downstream stem−loop is indeed a feature involved in the binding pathway. Interestingly, the Met_1:Met_3 ratio is reduced ~2-fold with the alteration of the downstream stem−loop. This indicates that the downstream stem−loop is required not only for efficient ribosome binding to the

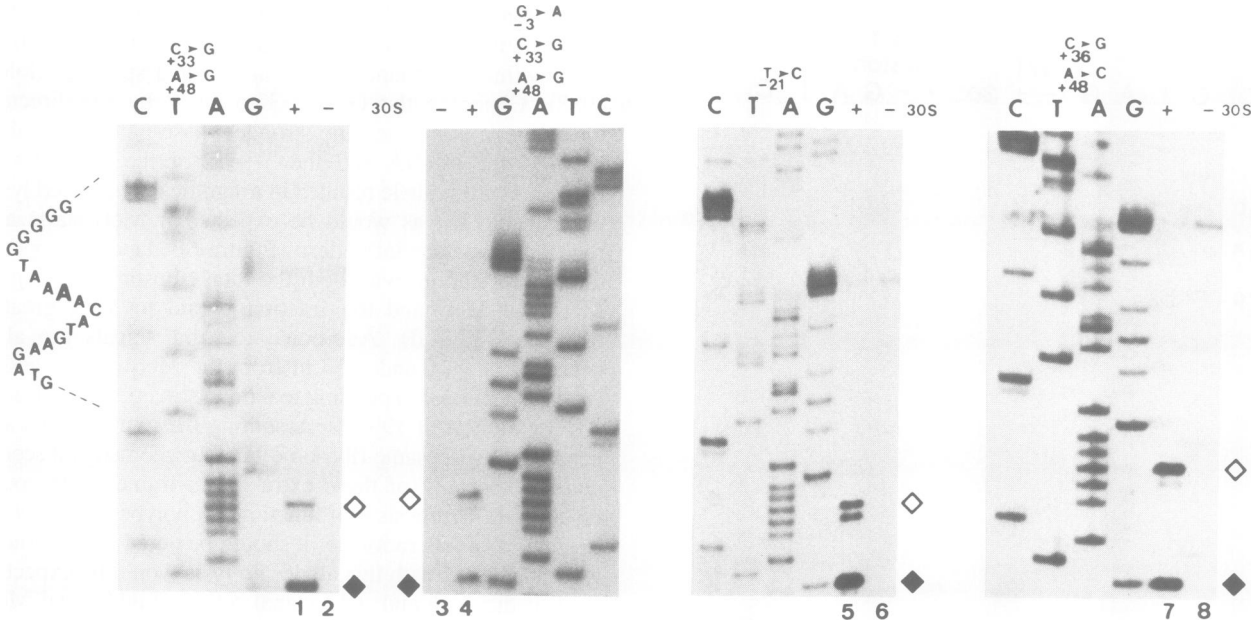

Fig. 7. *Sdi* and downstream stem−loop control start codon partition. The toeprint signals corresponding to initiation at Met_1 (◇) and Met_3 (◆) obtained with the double mutation (C_{33} → G; A_{48} → G), with the triple mutation (G_{-3} → A; C_{33} → G; A_{48} → G), with mutation (T_{-21} → C; strengthening of *sdi*) and with the double mutation (C_{36} → G; A_{48} → C; strengthening of downstream stem−loop) are shown in **lanes 1, 4, 5** and **7** respectively. **Lanes 2, 3, 6** and **8**, primer extension carried out in the absence of 30S ribosomes. The translational initiation site of gene *S* is shown at the left. For example, The G_{-3} → A change of the triple mutation is given by a bold A.

5′ region of the *S* mRNA but also for proper partitioning of starts between Met_1 and Met_3.

Since both the destabilization of the downstream stem−loop structure and the alteration of the Met_3 Shine−Dalgarno sequence (UAAG → UAAA) resulted in a significant reduction in total ribosome binding, an allele containing both changes (three single base changes: G_{-3} → A; G_{+33} → G; A_{+48} → G) was made. Toeprinting performed on mRNA carrying this construct revealed that the overall ribosome binding was ~4-fold lower than with wild-type (Table II). The few binding events that do occur are partitioned almost equally between Met_1 and Met_3 (Figure 7, lane 4) and induction of the plasmid pLS123 carrying this construct resulted in lysis being delayed 30 min when compared to wild-type (Table II). Comparison of the toeprints of two alleles with the downstream stem−loop defect (Figure 7, lanes 1 and 4) again shows that the UAAG Shine−Dalgarno sequence is important for starting at Met_3.

The fact that the destabilization of the *sdi* and downstream stem−loop structures have opposite effects on the choice of initiation codons suggests that the 30S ribosome is initially positioned by binding between these two structures. To test this idea, both the *sdi* and downstream structures were altered towards increased stability by adding additional G:C base pairs. Increasing the *sdi* stem structure by a single G:C base pair (T_{-21} → C; ΔG changes from −6.3 to −9.3 kcal/mol) results in a decrease in the Met_1:Met_3 ratio *in vitro* and faster lysis *in vivo* (Figure 7, lane 5; Table II). In contrast, strengthening the downstream stem loop (C_{+36} → G; A_{+48} → C; ΔG changes from −7.6 to −12.2 kcal/mol) results in opposite changes: an increase in the Met_1:Met_3 ratio and delayed lysis (Figure 7, lane 7; Table II). Since the changes in the downstream structure do not affect the amino acid sequence, the delay in lysis must be due to an alteration in the ratio of the gene products. Taken with the destabilization results cited above, these data suggest that

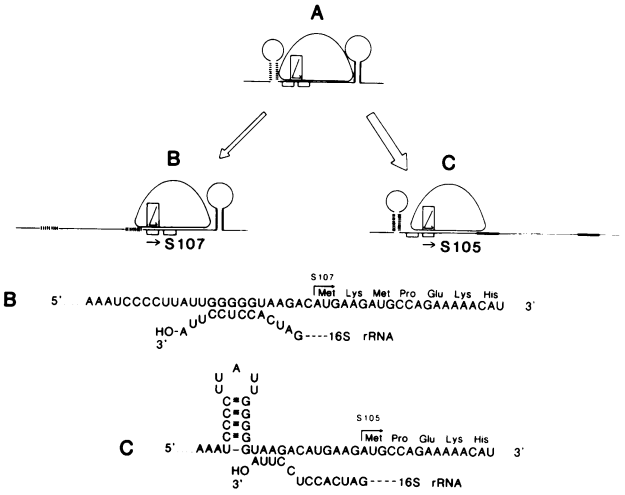

Fig. 8. The *sdi* control model. The two stem−loop structures and the intervening 37 nt of single-stranded mRNA define a domain which serves for initial binding of 30S particles (**A**). Lateral diffusion then results in two alternative routes (**B** and **C**). In two-thirds of binding events, melting of the downstream stem−loop is followed by ternary complex formation over codon Met_3, leading to expression of the lysis effector S105 (**C**). In one-third of the binding events, melting of the *sdi* structure allows ternary complex formation over codon Met_1, resulting in expression of the lysis inhibitor S107 (**B**).

the two stem−loop structures both participate in positioning the 30S ribosome in the *S* translational start region. Moreover, the relative stability of the two structures is set to effect a particular ratio of initiation events at the two start codons.

Discussion

Prokaryotic genes, with rare exceptions, have a single translational start site, usually identifiable by an AUG 6−10 nt downstream of a Shine−Dalgarno sequence (Gold *et al.*,

Table III. *Escherichia coli* strains and plasmids used in this study

Strain	Relevant features	Source or reference
Bacterial strains		
MC4100	F⁻Δ *lac araD* U169 sm^R *thi*⁻ *relA*⁻	Raab *et al.* (1988)
JM103	Δ (*lac pro A,B*) *supE, thi,* (r_k^+, m_k^+)/*F' tra D36, pro AB lac* ΔzM15 I^q	Messing *et al.* (1981)
CJ236	*dut ung thi relA1* F'::Tn*9*	Kunkel *et al.* (1987)
pop2135	*endA thi hsdR malT* (*cI857* p_R)::*malPQ*	O.Raibaud (Institute Pasteur, Paris)
Plasmids		
pLc236	p_L promoter, amp^r	Remaut *et al.* (1981)
pSKS107	*lacZ* fusion vector, amp^r	Shapira *et al.* (1983)
Sϕ*lacZ* fusion plasmids		
pSB87	wild-type *S* fused at codon 10 to *lacZ* (S107−β-Gal and S105−β-Gal products)	This study
pSB88	M_1 → L allele fused at codon 10 to *lacZ* (S105−β-Gal product)	This study
pSB89	M_3 → L allele fused at codon 10 to *lacZ* (S107−β-Gal product)	This study
pSB90	$M_{1,3}$ → L allele fused at codon 10 to *lacZ* (no hybrid protein product)	This study
In vitro transcription vectors		
pSP65	SP6 promoter, amp^r	Promega Biotec
pGEM™-3Z	SP6, T7 promoters, amp^r	Promega Biotec
pGEM™-4Z	SP6, T7 promoters, amp^r	Promega Biotec
Bluescript M13−	T3, T7 promoters, amp^r	Stratagene Cloning Systems

1981). This is necessary since prokaryotic ribosomes, or more properly the 30S ribosomal subunit, must be capable of initiating in the interior of a polycistronic mRNA, and thus there must be a mechanism to distinguish start codons from AUG codons internal to a reading frame. A few exceptions serve to reinforce the rule. Some genes, such as *lacZ*, have a second Met codon within the first few codons and both the primary AUG and, in the case of *lacZ*, the third codon AUG, are used for translational starts *in vivo* (Munson *et al.*, 1984). In these cases, however, a single Shine−Dalgarno sequence serves both starts, and there appears to be no biological function dependent on the production of the shorter gene product. Presumably, the presence of the second Met in these instances simply serves to rescue initiations that would otherwise be lost because of a failure to form the ternary complex over the first AUG. Two translation starts, leading to gene products with different functions, have been shown for the transposase reading frame of IS50 (Rothstein *et al.*, 1980) but in that case the starts are served from different promoters and are thus utilized on entirely different transcripts. Other two-start genes have been described, including the ϕX174 *A* gene (Dressler *et al.*, 1978), T7 gene *4* (Dunn and Studier, 1983), and the *cheA* gene of *Escherichia coli* (Smith and Parkinson, 1980). In all these cases, the two translational starts are far apart, leading to polypeptide products of greatly different size, and clearly involve unrelated ribosome binding pathways.

Previously, a model derived from genetic data has been presented in which dual translational initiation sites in the *S* gene are utilized, generating two *S* polypeptides *in vivo*, differing only by two residues at the amino terminus (Raab *et al.*, 1988). The polypeptides were proposed to have opposing roles in lysis. Here we have presented direct evidence for the utilization of Met_1 and Met_3 as translational starts *in vitro* and *in vivo*. Toeprint analysis shows that both Met_1 and Met_3 can support ternary complex formation *in vitro* (Figure 3). In addition, the *lacZ* fusion studies demonstrated that both start codons are utilized *in vivo* (Table I). Induction of the wild-type and start codon mutant alleles cloned under p_L control results in lysis profiles entirely consistent with the notion that only starts at Met_3, and thus production of S105, can result in lysis (Figure 2). Scheduling of lysis *in vivo* is shown to be well correlated with the ratio of the toeprint signals for Met_1 and Met_3. In general, when the toeprint signal corresponding to Met_3 ternary complex formation is in excess over the signal corresponding to Met_1 events (i.e. when the Met_1:Met_3 ratio is <1), lysis occurs, and the greater the excess, the more rapid the onset of lysis (Table II). In contrast, when the Met_1 signal was in excess over the Met_3 signal, lysis does not occur (Table II). Thus, the scheduling of lysis is seen to be dependent on the partition of ribosome initiations between the two start sites.

What controls the utilization of the two AUG codons? The determinants of the strength of a translational start site are not fully understood, but one component of the fitness of a start is the complementarity of the Shine−Dalgarno sequence to the CCUCC sequence of nt 1535−1539 in the 16S rRNA (Gold *et al.*, 1981). Two adjacent sequences, GGGGG (−11:−7) and UAAG (−6:−3) can be identified that are complementary to the 3′ terminal region of the 16S rRNA and, given the spacing considerations (Gold, 1988), are the only reasonable Shine−Dalgarno sequences for Met_1 and Met_3 respectively. The UAAG sequence is not complementary to the CCUCC rRNA sequence but instead to the extreme 3′ end of the rRNA (Shine and Dalgarno, 1974). This interaction is likely to be weak and possibly contributes strain to the ternary complex. Nevertheless, altering the UAAG sequence to reduce its complementarity or to create complementarity with the CUCC (nt 1536−1539) respectively decreases and increases the toeprint signal for Met_3 (Figure 6B). Analogously, increasing the spacing between the GGGGG and the Met_1 codon (Figure 6A) reduces initiation at Met_1 but not Met_3. Thus the two

Shine–Dalgarno sequences clearly serve separate start codons. Purely on comparative grounds, the GGGGG sequence should provide a much 'better' Shine–Dalgarno domain than the UAAG. Nevertheless, it is the Met_3 start served by UAAG that is apparently stronger, as judged by toeprinting and *lacZ* fusion analysis. The *sdi* stem–loop structure, which we have shown to exist *in vitro* (Figure 4), apparently serves to reduce Met_1 initiations by occluding the GGGGG Shine–Dalgarno sequence. There are a number of precedents for local secondary structure having a negative effect on translation initiation, and in general the stronger the secondary structure, the more pronounced the reduction of initiation (Hall *et al.*, 1982; Halling *et al.*, 1982; McPheeters *et al.*, 1986; Schmidt *et al.*, 1987). Thus *sdi* mutations, which were isolated as lysis-defective alleles showing significant dominant character, can now be understood in terms of the stability of the secondary structure. Each of the *sdi* mutations should destabilize the stem–loop and thus make the Shine–Dalgarno sequence for Met_1 more accessible. In fact, the toeprints indisputably show an increase in Met_1 ternary complex formation at the expense of Met_3 (Figure 5).

The data presented here suggest a plausible sequence of events for the ribosome-binding pathway that leads to the partition of initiation events. Since changes that destabilize either *sdi* or the 31–47 downstream stem–loop all result in a significant reduction in ribosome binding (Table II), we suggest that these two structures, and the intervening 37 nt of single-stranded RNA, define a domain that serves for initial binding of the 30S ribosome. Moreover, we have shown that altering the stability of either stem–loop causes complementary changes in the Met_1:Met_3 ratio. Increasing the stability of *sdi* and decreasing the stability of the downstream structure decreases the Met_1: Met_3 ratio. Strengthening the downstream structure and destabilizing *sdi* causes an increase in Met_1:Met_3. This suggests that breathing of the *sdi* or the downstream stem loop is required for initiations at Met_1 or Met_3 respectively, presumably by allowing lateral diffusion of the 30S particle in the 5′ or 3′ direction. It is also possible that the close proximity of the bound 30S ribosome may facilitate denaturation of either stem–loop structure. In any case, most ribosome-binding events result in ternary complex formation over the Met_3 start codon, in spite of the poor topology and complementarity of the UAAG Shine–Dalgarno sequence that serves Met_3 (Figure 8). Note that substitution of the canonical GAGG for UAAG results in increased initiations at Met_3, at the expense of Met_1. Thus many, if not most, initiations at Met_1 should come from 30S particles that have failed to initiate at Met_3 in spite of adequate downstream lateral diffusion. In contrast, breathing of the *sdi* structure would result in a high efficiency of ternary complex formation over Met_1 by exposing the strongly complementary GGGGG sequence at nt −11 to −7 (Figure 8). Thus this rationale suggests that the weakness of the pairing of the UAAG sequence with the 16S rRNA is important for the partitioning mechanism. Support of this idea is the fact that the UAAG sequence is conserved in the analogous control region for the P22 lysis gene 13 (Renell and Poteete, 1985; Raab *et al.*, 1988; K.Nam and U.Bläsi, unpublished results).

Why has such an unprecedented mechanism for control of the *S* gene evolved? In a separate report, we will show that both S105 and S107 can be detected *in vivo* as stable, membrane-embedded proteins. Furthermore, S107 can act in *trans* to inhibit lysis by S105, presumably by co-oligomerizing and poisoning pore-formation. It is also clear from our previous work (Altman *et al.*, 1985) and from the *lacZ* fusion data presented here (Table I) that the *S* gene is expressed at a very low level, in spite of being the first cistron on the very abundant λ late transcript. Therefore the important features of *S* expression are the low total level of translation and the production of a particular proportion of S105 and S107 molecules. Evolving a domain consisting of the two stem–loop structures separated by one 30S particle-width of single-stranded mRNA generates the potential for low total translation and simultaneously allows, as we have shown, the partition of initiation events between the first and third codons. Efforts are currently underway to characterize analogous control regions in the closely related lysis gene *13* of P22 and the completely non-homologous lysis gene *S* of phage 21.

Materials and methods

Strains and plasmids

A list of *E.coli* strains and plasmids used is given in Table III. Cultures were grown in Luria broth (Miller, 1972) supplemented with ampicillin (100 μg/ml) if a plasmid was present. Growth and lysis profiles were monitored by measuring the A_{550} in a Gilford Stasar spectrophotometer fitted with a liquid sampling device.

Site-directed mutagenesis

Site-directed mutagenesis was carried out as described by Kunkel *et al.* (1987). The basic single-stranded templates used were a M13mp18 construct and a Bluescript M13 (Table III) derivative which carried the *Eco*RI–*Hin*dIII fragment of plasmid pRG1 (Raab *et al.*, 1986). This fragment contains the region of the λ genome from the *Eco*RI site at 44 972 to the *Cla*I site at 46 438, a segment that spans the three lysis cistrons *S*, *R* and *Rz*. The synthetic oligonucleotides listed below were synthesized by the Oligonucleotide Service Facility, Texas Agricultural Experiment Station, and were used as mutagenic primers after gel purification. The mismatches are underlined. The mutations generated with the corresponding oligonucleotide are given in parentheses.

A: 3′-CATTCTGGACTTCTACGGTCTT-5′ (M_1 → L)
B: 3′-CATTCTGTACTTCGACGGTCTT-5′ (M_3 → L)
C: 3′-CATTCTGGACTTCGACGGTCTT-5′ ($M_{1,3}$ → L)
D: 3′-GGATTAATTTATCTCGTTTCCCCCAATAAGGGGAATT-CTGTACTTCTACGGTC-5′ (inversion of *sdi* structure)
E: 3′-AACCCCCATTTTGTACTTCTA-5′ (G_{-3} → A)
F: 3′-GAATAACCCCCCTCCTGTACTTCTACGGTC-5′ (TAAG → GAGG at nt −6 to −3)
G: 3′-GGACAACCGCCGGTAAGAGCGCCGCTTCCTTGTTCC-5′ (C_{33} → G; A_{48} → G)
H: 3′-GACAACCGGTGGTAAGAG-5′ (G_{34} → A; creation of *Bal*I site)
I: 3′-ATCTCGTTTGGGGGAATA-5′ (T_{21} → C)
J: 3′-GTACTGGACAACCGGCGCTAAGAGCGCCGGTTCCTT-GTTCCGTAGCCCC (C_{36} → G; A_{48} → C)

Candidate clones were sequenced as described earlier (Raab *et al.*, 1986) and mutated DNA sequences were subcloned into *in vitro* transcription vectors (Table III) for toeprinting analysis.

Construction of the SφlacZ fusions

M13mp18 clones carrying the wild-type, the M_1 → L, the M_3 → L or the $M_{1,3}$ → L alleles of *S* were used as a template for site-directed mutagenesis utilizing primer H, generating a *Bal*I site at codon 11 in all four constructs. From each of the M13mp18 clones a 419 bp *Pvu*II–*Bal*I fragment containing the first 10 codons of gene *S* under control of the lac promoter was isolated and inserted into the *Sma*I site of the *lacZ* fusion

vector pSKS107 (Table III). β-Galactosidase was measured as described by Miller (1972).

Cloning of different S alleles under p_L control

To demonstrate the phenotypical effect of the mutant *S* alleles, the *SRRz* genes were reisolated on *Eco*RI−*Hin*dIII fragments and inserted by standard methods (Maniatis *et al.*, 1982) into the p_L expression vector pLc236 (Table III). The host for these constructs was pop2135 (Table III), which carries a thermosensitive allele of the λ *cI* repressor gene inserted into the *malPQ* locus of the *E.coli* chromosome.

In vitro transcription

The SP6 and T7 *in vitro* transcription vectors carrying the different *S* alleles were found to be interchangeable for the purpose of generating mRNA for toeprinting. The λ DNA inserted downstream of the SP6 or T7 promoter was the *Eco*RI−*Hin*dIII fragment described above, carrying the *SRRz* genes. Before addition of the phage polymerase, 5 μg of each plasmid DNA was subjected to restriction with *Rsa*I, which cleaves in the *S* gene at codon 31 (λ bp 45276). *In vitro* transcription was carried out with the Riboprobe Gemini System II (Promega Biotec) according to the manufacturer's instructions. An 0.5 μg aliquot of each *Rsa*I-restricted plasmid was used as a template for marker RNA synthesis in the presence of [α-^{32}P]UTP. The RNA products, which extend from the SP6 or T7 transcription initiation site through the first 31 codons of *S* (to the *Rsa*I cleavage) were purified by PAGE, resuspended in 50 μl 0.1 mM EDTA and used directly for toeprinting.

Toeprinting analysis

Toeprinting was performed as described (Hartz *et al.*, 1988) except that gel-purified RNA was used. A 12.5 pmol aliquot of a synthetic oligonucleotide primer (5′-GGATTGCCCCGATGCCT-3′), complementary to nucleotides +57 to +73 of gene *S* was labeled with 20 nmol of [γ-^{32}P]ATP; 1.5 pmol of the labeled primer was annealed to 0.5 pmol purified RNA in a 20 μl reaction. The primer extension inhibition reactions were carried out in standard buffer (SB: 10 mM Tris−acetate pH 7.4, 60 mM ammonium chloride, 10 mM magnesium acetate, 6 mM mercaptoethanol) by mixing 1/10 of the volume of the corresponding annealing mixes, 1 μl 3.75 mM dATP, dCTP, dGTP, dTTP, 10 pmol of 30S ribosomes and 50 pmol of tRNA^{fMet}, both in SB buffer. This mixture was pre-incubated at 37°C for 10 min. Then 200 U MMLV reverse transcriptase (BRL) were added to each sample, and the incubation continued for 15 min. The reactions were stopped by addition of 1 vol of gel loading dye and heat inactivation for 3 min at 95°C. Gel loading dye contains, per ml, 940 μl deionized formamide, 40 μl 10 × TBE gel buffer, and 20 μl of a dye stock containing 20 mg/ml each of xylene cyanol and bromphenol blue. TBE is as described in Maniatis *et al.* (1982). Sequencing reactions were performed in parallel on aliquots of the gel-purified RNA using AMV reverse transcriptase (Life Sciences) and dideoxynucleotide triphosphates at a final concentration of 200 μM, without pre-incubation or the addition of ribosomes or tRNA. Toeprint and sequencing samples were analyzed in parallel on 8% polyacrylamide−urea gel (Maniatis *et al.*, 1982). The gels were developed by autoradiography. The toeprint bands corresponding to the Met_1 and Met_3 starts were quantitated by densitometry. At the top of each lane, a band corresponding to read-through of reverse transcriptase to the end of the mRNA was also scanned (not shown). This read-through signal was used as a measure for unbound 30S ribosomes (see Table II).

Ribonuclease T1 digestion

The same concentration of wild-type RNA annealing mix used for toeprinting was subjected to T1 digestion. The T1 reactions contained 2 μl of annealing mix, 1 μl of a 3.75 mM deoxynucleotide triphosphate mix, 1 μl T1 RNAse (5, 1 or 0.5 U) and 5 μl SB buffer as given above. The reactions were kept on ice for 10 min. Then 0.3 U AMV reverse transcriptase was added and the reactions were incubated at 48°C for 15 min. Reactions were terminated by heating to 95°C and addition of 20 μl loading dye.

Acknowledgements

We thank R.Traut for providing purified ribosomes. Drafting was provided by Lisa Lohman and photographic services by TAMU Biomedical Communications. U.B. was supported by a Forschungsstipendium from the Deutsche Forschungsgemeinschaft (B1 262/1-2). This work was supported by Public Health Service grants GM27099 to R.Y. and GM28685 to L.G. R.Y. was also supported by funding from the Texas Agricultural Experiment Station.

References

Adhya,S., Sen,A. and Mitra,S. (1971) In Hershey,A.R. (ed.), *The Bacteriophage Lambda*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 743−746.

Altman,E., Young,K., Garrett,J.M., Altman,R. and Young,R. (1985) *J. Virol.*, **53**, 1008−1011.

Bienkowska-Szewczyk,K.B., Lipinska,B. and Taylor,A. (1981) *Mol. Gen. Genet.*, **139**, 111−114.

Daniels,D.L., Schroeder,J.L., Szybalski,W., Sanger,F., Coulson,G., Hong,G.F., Hill,D.F., Peterson,G.B. and Blattner,F.R. (1983) In Hendrix,W., Roberts,J.W., Stahl,F.W. and Weisberg,R.A. (eds), *Lambda II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 519−676.

Dressler,D., Hourcade,D., Koths,K. and Sims,J. (1978) In Denhardt,D.T., Dressler,D. and Ray,D.S. (eds), *The Single-stranded DNA Phages*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 187−215.

Dunn,J.J. and Studier,F.W. (1983) *J. Mol. Biol.*, **166**, 477−535.

Freier,S.M., Kierzek,R., Jaeger,J.A., Sugimoto,N., Caruthers,M.H., Neilson,T. and Turner,D.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9373−9377.

Garrett,J.M. and Young,R. (1982) *J. Virol.*, **44**, 886−892.

Gold,L. (1988) *Annu. Rev. Biochem.*, **57**, 199−233.

Gold,L. and Stormo,G. (1987) In Neidhardt,F.C., Ingraham,J.L., Brooks Low,K., Magasanik,B., Schaechter,M. and Umbarger,H.E. (eds), *Escherichia coli and Salmonella typhimurium*. American Society for Microbiology, Washington, DC, pp. 1302−1307.

Gold,L., Pribnow,D., Schneider,T., Shinedling,S., Swebilius,B. and Stormo,G. (1981) *Annu. Rev. Microbiol.*, **35**, 365−403.

Grantham,R., Gautier,C. and Gouy,M. (1980) *Nucleic Acids Res.*, **8**, 1893−1912.

Hall,M.N., Gabay,J., Debarboville,M. and Schwartz,M. (1982) *Nature*, **295**, 616−618.

Halling,S.M., Simons,R.W., Way,J.C., Walsh,R.B. and Kleckner,N. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2608−2612.

Hartz,D., McPheeters,D.S., Traut,R. and Gold,L. (1988) *Methods Enzymol.*, **164**, 419−425.

Herskowitz,I. and Hagen,D. (1980) *Annu. Rev. Genet.*, **14**, 399−445.

Itakura,K., Hirose,R., Crea,A., Riggs,A., Heynecker,H.L., Bolivar,F. and Boyer,H.W. (1979) *Science*, **198**, 1056−1063.

Königsberg,W. and Godson,G.N. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 687−691.

Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) *Methods Enzymol.*, **154**, 367−382.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

McPheeters,D.S., Christensen,A., Young,E.T., Stormo,G. and Gold,L. (1986) *Nucleic Acids Res.*, **14**, 5813−5826.

McPheeters,D.S., Stormo,G.D. and Gold,L. (1988) *J. Mol. Biol.*, **201**, 517−535.

Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309−21.

Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 13−23.

Munson,L.M., Stormo,G.D., Niece,R.L. and Reznikoff,W.S. (1984) *J. Mol. Biol.*, **177**, 663−683.

Raab,R., Neal,G., Garrett,J., Grimaila,R., Fusselman,R. and Young,R. (1986) *J. Bacteriol.*, **167**, 1035−1042.

Raab,R., Neal,G., Sohaskey,C., Smith,J. and Young,R. (1988) *J. Mol. Biol.*, **199**, 95−105.

Reader,R.W. and Siminovitch,L. (1971) *Virology*, **43**, 623−627.

Remaut,E., Stanssens,P. and Fiers,W. (1981) *Gene*, **15**, 81−93.

Renell,D. and Poteete,A.R. (1985) *Virology*, **143**, 280−289.

Rothstein,S.J., Jorgenson,R.A., Postle,K. and Reznikoff,W.S. (1980) *Cell*, **19**, 795−805.

Schmidt,B.F., Berkhout,B., Overbeek,G.P., van Strien,A. and van Duin,J. (1987) *J. Mol. Biol.*, **195**, 505−516.

Shapira,S.K., Chou,J., Richaud,F.V. and Casadaban,M.J. (1983) *Gene*, **25**, 71−82.

Shine,J. and Dalgarno,L. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 1342−1346.

Smith,R.A. and Parkinson,J.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5370−5374.

Stanssens,P., Remaut,E. and Fiers,W. (1985) *Gene*, **36**, 211−223.

Wilson,D.B. (1982) *J. Bacteriol.*, **151**, 1403–1410.
Winter,R.B., Morrissey,L., Gauss,P., Gold,L., Hsu,T. and Karam,J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 7822–7826.
Young,R., Way,J., Way,S., Yin,J. and Syvanen,M. (1979) *J. Mol. Biol.*, **132**, 307–322.
Young,K.D. and Young,R. (1982) *J. Virol.*, **44**, 993–1002.
Zagotta,M.T. (1989) *Ph.D. dissertation, Cornell University.*

Received on May 31, 1989; revised on July 7, 1989