Mol Gen Genet (1992) 235:325–332

MGG

DNA sequences required for translational frameshifting in production of the transposase encoded by IS*1*

Yasuhiko Sekine and Eiichi Ohtsubo

Institute of Applied Microbiology, University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113, Japan

Received April 21, 1992 / Accepted June 27, 1992

Summary. The transposase encoded by insertion sequence IS*1* is produced from two out-of-phase reading frames (*insA* and B′-*insB*) by translational frameshifting, which occurs within a run of six adenines in the −1 direction. To determine the sequence essential for frameshifting, substitution mutations were introduced within the region containing the run of adenines and were examined for their effects on frameshifting. Substitutions at each of three (2nd, 3rd and 4th) adenine residues in the run, which are recognized by $tRNA^{Lys}$ reading *insA*, caused serious defects in frameshifting, showing that the three adenine residues are essential for frameshifting. The effects of substitution mutations introduced in the region flanking the run of adenines and in the secondary structures located downstream were, however, small, indicating that such a region and structures are not essential for frameshifting. Deletion of a region containing the termination codon of *insA* caused a decrease in β-galactosidase activity specified by the *lacZ* fusion plasmid in frame with B′-*insB*. Exchange of the wild-type termination codon of *insA* for a different one or introduction of an additional termination codon in the region upstream of the native termination codon caused an increase in β-galactosidase activity, indicating that the termination codon in *insA* affects the efficiency of frameshifting.

Key words: Adenine run – Cointegration – Secondary structure of mRNA – Termination codon – $tRNA^{Lys}$

Introduction

Accurate translation is ensured in general by regular progression of the ribosome in a triplet manner. Since disturbance of the regular triplet rule results in abortive translation, it is believed that the translational apparatus is actively constrained to maintain the correct reading frame during polypeptide chain elongation. Some functional proteins, however, can be synthesized only if specific translational frameshifting occurs (for a recent review, see Atkins et al. 1990; Sekine and Ohtsubo 1991). Insertion sequence IS*1* (Ohtsubo and Ohtsubo 1978) uses such programmed frameshifting to produce IS*1* transposase from two out-of-phase reading frames, *insA* and B′-*insB*, where B′ is an open reading frame extending from the ATG initiation codon of the *insB* frame and is in the −1 frame with respect to *insA* (Sekine and Ohtsubo 1989). The frameshifting event in the −1 direction occurs at an AAA codon for Lys in *insA* within a run of six adenines present in the overlapping region between *insA* and B′, and produces the InsA-B′-InsB fusion protein, IS*1* transposase, with the segment Leu-Lys-Lys-Leu encoded by the region containing the run of adenines (Sekine et al. 1992).

In this paper, we report determination of the nucleotide sequence required for the frameshifting in IS*1*. We introduced substitution mutations in the run of adenines and the regions neighboring it, and then examined the effects of the changed context on frameshifting. Based on the results obtained, we discuss mechanisms underlying the event of translational frameshifting in IS*1*. We discuss the fact that, although other genetic systems use sequences similar to that in IS*1* as frameshift signals, our results show differences between IS*1* and others in the mechanism governing the frameshifting event.

Materials and methods

Bacterial strains and plasmids. Bacterial strains used were *Escherichia coli* K12 derivatives, MV1184 (Vieira and Messing 1987), BW313 (Kunkel et al. 1987), MC1000 (Casadaban and Cohen 1980), JE5519 (Ohtsubo et al. 1981), and YS202 (Sekine et al. 1992).

Plasmid pSEK17, a pUC18 derivative, carries one copy of IS*1* (Sekine et al. 1992). Plasmid pSEK117, a pUC119 derivative, which carries one copy of IS*1* was constructed from pSEK17 (Sekine et al. 1992). Plasmid

Correspondence to: E. Ohtsubo

pHS1 is a temperature-sensitive replication mutant of the tetracycline-resistance plasmid pSC101 (Hashimoto-Gotoh and Sekiguchi 1977). Plasmid pR-pMLB (a gift from D. Bastia) is a pBR322 derivative, from which IS*1*-*lacZ* fusion plasmids were constructed as described below. Plasmid pSEK6000, a pR-pMLB derivative, is an IS*1*-*lacZ* fusion plasmid having a DNA fragment of wild-type IS*1* corresponding to IS*1* coordinates 292–353 (Sekine et al. 1992).

Media. Culture media used were L broth, L-rich broth, ϕ-medium (Yoshioka et al. 1987) and 2×YT broth (Messing 1983). ϕ-medium was used for transformation of plasmid DNA, and 2×YT broth was used for mutagenesis in constructing mutant plasmids. L-agar plates contained 1.5% (w/v) agar (Eiken Chemical) in L broth. Antibiotics were added in L-agar plates, if necessary, at concentrations of 150 μg/ml ampicillin (Wako Junyaku) and 5 or 10 μg/ml, tetracycline (Sigma). Peptone dilution buffer (0.1% peptone (Kyokuto Seiyaku) in 0.3% NaCl) was used for dilution of cell cultures.

Enzymes. Restriction endonucleases (*Bam*HI, *Bgl*II, *Mlu*I and *Pst*I), bacterial alkaline phosphatase, T4 DNA polymerase, T4 polynucleotide kinase, and T4 DNA ligase were obtained from Takaro Shuzo. Restriction endonuclease *Bst*EII was obtained from New England Biolabs. RNase A was purchased from Sigma. These enzymes were used in the buffers recommended by their suppliers.

DNA preparation. Strain MV1184 or MC1000 harboring a plasmid was grown in L-rich broth. A crude lysis method (Machida et al. 1982) was used to isolate and examine small amounts of plasmid DNA from many cell cultures. The alkaline lysis method (Maniatis et al. 1982) was used to prepare plasmid DNA for cloning and nucleotide sequencing.

Nucleotide sequencing. Nucleotide sequences were determined by the dideoxynucleotide method (Sanger et al. 1977; Messing 1983) using a 7-DEAZA sequencing kit (Takara Shuzo). We used M4 primer (Takara Shuzo) for sequencing derivatives of pSEK6000. Synthetic oligodeoxyribonucleotide D_2, described in Sekine and Ohtsubo (1989), was also used as primer for sequencing derivatives of pSEK117. The DNA chains were labeled with α-[^{32}P]dCTP (15 TBq/mmol, Amersham) and separated in 6 or 8% polyacrylamide gels containing 8 M urea.

Plasmid construction. pSEK17 is the parental plasmid for all plasmids carrying mutant IS*1* described in this paper. Each of the pSEK17 derivatives carrying mutant IS*1* with a substitution(s) or a 1 bp insertion was constructed as follows. Using pSEK117 as template and oligodeoxyribonucleotides synthesized using a DNA synthesizer 380B (Applied Biosystems) as primers, the IS*1* sequence in pSEK117 was mutagenized by site-directed mutagenesis according to Kunkel et al. (1987). The sequences mutated were confirmed by DNA sequencing. Two pSEK17 derivatives carrying IS*1*-48 or IS*1*-49 were constructed by replacing the *Pst*I-*Mlu*I fragment of IS*1* in pSEK17 with the *Pst*I-*Mlu*I fragment of the corresponding mutagenized derivative from pSEK117. The other pSEK17 derivatives were constructed by replacing the *Pst*I-*Bst*EII fragment of IS*1* in pSEK17 with the *Pst*I-*Bst*EII fragment of the corresponding mutagenized derivative from pSEK117.

pSEK6000 derivatives, shown in Fig. 2 and Fig. 4, each with a substitution mutation in the IS*1* fragment were constructed as follows. Plasmid pSEK4000 was first obtained by introducing two *Bgl*II recognition sites into positions 286–291 and 354–359 of the IS*1* sequence in pSEK117 by site-directed mutagenesis, as described above. Using pSEK4000 as template, a substitution was introduced within the run of adenines or in a neighboring position by site-directed mutagenesis as above, and the *Bgl*II fragment, which contained the substitution mutation in the resulting plasmid, was inserted into the *Bam*HI site of vector plasmid pR-pMLB.

Purification of β-galactosidase (LacZ) fusion proteins and amino acid sequencing. Strain YS202 harboring each of the IS*1*-*lacZ* fusion plasmids, pSEK6151, pSEK6087 and pSEK6061, was grown in 4 l of L-rich broth containing 0.2% (w/v) glucose at 30° C until the OD_{600} reached 0.5–0.6, and then the culture was incubated with aeration at 40° C for 60 min to induce the LacZ fusion protein. Purification and amino acid sequencing of the protein was carried out as described (Sekine et al. 1992).

Cointegration assay. Each of the ampicillin-resistance plasmid pSEK17 derivatives carrying mutant IS*1* was introduced by transformation into strain JE5519 which already harbored the tetracycline-resistance plasmid pHS1. Cointegration between a pSEK plasmid and pHS1 was assayed according to the method described by Sekine and Ohtsubo (1989).

LacZ assay. Each of the pSEK6000 derivatives, the ampicillin-resistance plasmids carrying the ATG_{cro}-IS*1*-*lacZ* construct, was introduced by transformation into strain MC1000. The LacZ activity in the resulting strain was determined as described (Sekine et al. 1992). Each value presented in this paper is the mean of those obtained from at least three independent experiments; standard errors in all cases were less than 15%.

Results

IS1 mutants with single substitution mutations and their cointegration ability

1. Saturation mutagenesis of the run of adenines. The run of six adenines present at the region corresponding to IS*1* coordinates 307–312 has been shown to contain the frameshift site (Fig. 1; Sekine and Ohtsubo 1989; Sekine et al. 1992). The amino acid sequence at residues 84–87 of IS*1* transposase has been shown to be Leu-Lys-Lys-Leu (LKKL), which is encoded by the DNA seg-

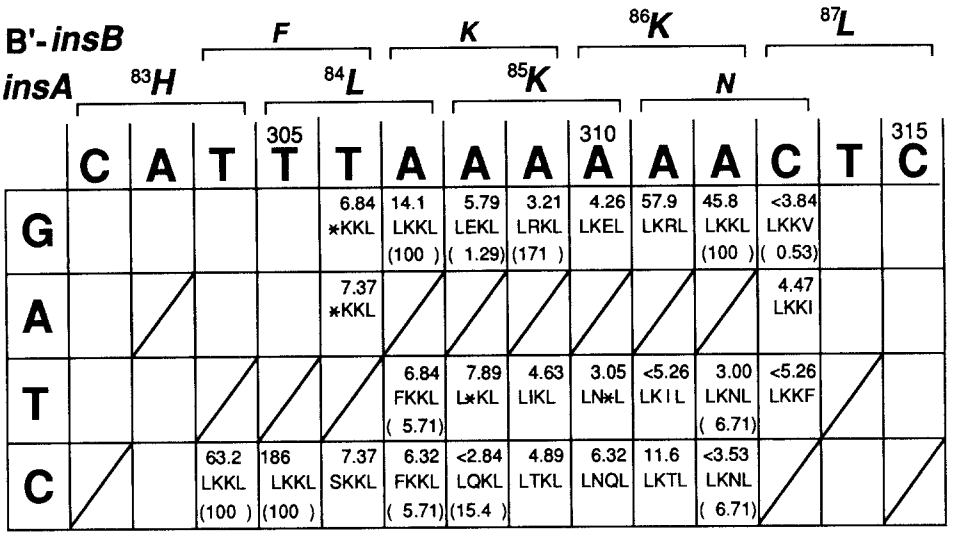
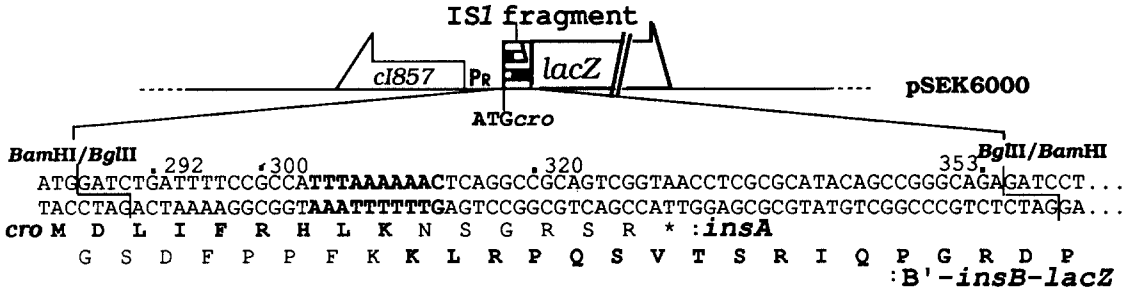


Fig. 1. IS*1* mutants with 1 bp substitutions and their cointegration ability. A critical portion of the nucleotide sequence of wild-type IS*1* at positions 302–315 is shown horizontally, and bases substituted are shown vertically. The amino acid sequences encoded by the *insA* and B'-*insB* reading frames are shown at the top, where the *boldface* letters with numbers are the amino acids of IS*1* transposase produced by −1 frameshifting (Sekine et al. 1992). The first line in each box shows the relative frequency of cointegration mediated by each IS*1* mutant, taking the value for wild-type IS*1* (1.9×10^{-8} per division cycle; Sekine and Ohtsubo 1989) as 100. The second line in each box shows the amino acid sequence at residues 84–87 of the transposase produced from each IS*1* mutant by successful frameshifting at the same site as that of wild-type IS*1*. *Asterisks* indicate positions with no amino acid residue due to generation of termination codons by substitution mutations. The number in *parenthesis* in the third line in each box shows the relative cointegration frequency of each IS*1* mutant with a 1 bp insertion, designed so as to produce a mutant transposase with the polypeptide segment shown in the second line without frameshifting; the cointegration frequency of the IS*1* mutant with a single adenine insertion within the run of adenines to produce wild-type transposase (1.4×10^{-6} per division cycle; Sekine and Ohtsubo 1989) was taken as 100

ment 5'-TTAAAAAACTC-3' including the frameshift site, and the last codon recognized in *insA* during translation of IS*1* transposase has been estimated to be codon AAA for Lys at the second, third and fourth position in the run of adenines (Fig. 1; Sekine et al. 1992). To confirm the importance of the run of adenines and to determine the critical nucleotides within it, we constructed IS*1* mutants each with a 1 bp substitution within the run of adenines and examined their ability to mediate cointegration. Figure 1 shows all the mutants constructed and summarizes the frequencies of cointegration mediated by these mutants. Note here that each mutation may alter the mRNA context responsible for translational frameshifting, causing variations in the efficiency of frameshifting. (This effect is referred to as the context effect.) Also, it may alter the amino acid specified by the mutated triplet, producing a mutant transposase with an altered degree of activity upon successful frameshifting. (This effect is referred to as the amino acid effect.) The cointegration frequency may be lowered by the context effect and/or by the amino acid effect. The degree of the amino acid effect can be estimated by examining the cointegration frequency mediated by an IS*1* mutant with a 1 bp insertion, which produces, without frameshifting, an altered transposase that is expected to correspond to the one produced by frameshifting from a relevant IS*1* mutant with a substitution (Fig. 1; Sekine et al. 1992). With these mutants, we can approximately estimate the context effect by subtracting the extent of the amino acid effect from the cointegration frequency measured.

In the mutants with a substitution in the run of adenines, the IS*1* mutant with a G substitution for nucleotide A at position 307 (designated as 307G) retained the ability to mediate cointegration, although at a reduced level (Fig. 1). This substitution changes codon TTA at positions 305–307 (designated as ^{305}TTA) in *insA* for Leu at the 84th residue (designated as ^{84}L) to the synonymous codon TTG, and thus this mutation is assumed to give wild-type transposase with the LKKL segment upon successful frameshifting. The result above suggests that the substitution at 307 shows some context effect, but that this position is not critically important for frameshifting. The mutants with each of the other substitutions, 307T and 307C, did not efficiently mediate cointegration (Fig. 1). The poor cointegration ability of these IS*1* mutants is likely to be due to alteration of the amino acid from ^{84}L (Leu) to F (Phe) by substitution mutations to give a mutant transposase with the FKKL segment, since the mutant with a 1 bp insertion which produces such a mutant transposase without frameshifting in fact showed less cointegration ability than that of the 1 bp insertion mutant producing wild-type transposase with the LKKL segment (Fig. 1). Note here that a mutant IS*1* with two substitutions, 305C and 307T, which changes the codon ^{305}TTA for ^{84}L (Leu) to the synonymous codon CTT, retained cointegration ability at 62% of the level of wild-type IS*1* (Sekine et al. 1992). This strongly supports the suggestion above that the nucleotide at 307 is not essential for frameshifting.

The mutants with substitution 308G, 308T or 308C did not mediate cointegration (Fig. 1). The loss of coin-



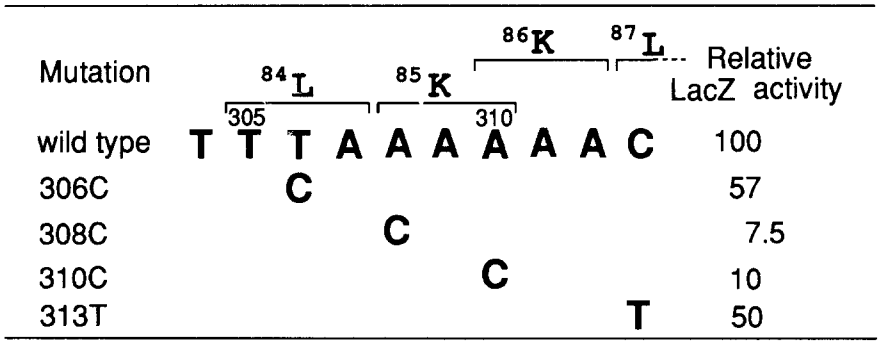


Fig. 2. The nucleotide sequence of a critical region in plasmid pSEK6000 and the activity of β-galactosidase (LacZ) produced in cells harboring pSEK6000 or each of its derivatives. pSEK6000 carries an IS*1* fragment containing a part of the *insA* and B′-*insB* reading frames, each of which is in-frame with ATG_{cro} and *lacZ*, respectively. In the amino acids encoded by the two reading frames, those which are required to give the InsA-B′-InsB-LacZ fusion protein are indicated by *boldface* letters. Production of the fusion protein is controlled at promoter p_R by a thermosensitive repressor, the product of *cI857*. A nucleotide sequence containing the run of six adenines is shown by *boldface* letters. In each pSEK6000 derivative, only the nucleotide different from that of pSEK6000 is shown. The relative LacZ activity of the LacZ fusion protein produced upon heat induction in cells harboring pSEK6000 (28.7 units, taken as 100) or each of its derivatives is shown.

tegration ability of the mutants with 308G or 308T might be due to the amino acid effect, since 308G and 308T cause serious changes of the codon ^{308}AAA for ^{85}K (Lys) to codon GAA for E (Glu) and to the termination codon TAA, respectively. However, the loss of the ability of the mutant with 308C is likely to be mainly due to the context effect, since the relevant 1 bp insertion mutant, which encodes transposase with the polypeptide segment LQKL, still retained the ability to mediate cointegration, but at a reduced frequency (Fig. 1). This suggests that the nucleotide at position 308 is essential for frameshifting. This suggestion will be confirmed by another experiment to be described below.

Substitution 309G severely inhibited cointegration (Fig. 1). This must be due to the context effect, but not to the amino acid effect, because mutant transposase with ^{85}Arg, which would be produced from the mutant with 309G upon successful frameshifting, showed an even higher activity than wild type (Fig. 1). The mutants with each of the other substitutions also did not mediate cointegration (Fig. 1). This loss of cointegration ability is quite likely to be due to the context effect, considering the result for the mutant with 309G.

The substitution for nucleotide A at position 310 of any one of the other nucleotides inhibited cointegration (Fig. 1). The substitution causes a serious change of the codon ^{308}AAA in *insA* or/and ^{310}AAA in B′-*insB*; so the context effects of these substitutions cannot be estimated exactly.

The IS*1* mutants with substitution 311G or 312G retained cointegration ability (Fig. 1). Substitution 311G changes the codon ^{310}AAA for ^{86}K (Lys) to AGA for R (Arg), which is a conservative change, whereas 312G does not cause alteration of the transposase. The results above, therefore, suggest that the nucleotides at 311 and 312 are not critically important for frameshifting. The other mutants with a substitution of the nucleotide at 311 or 312 failed to mediate cointegration (Fig. 1). These may be mainly due to the amino acid effect producing inactive transposases; the transposase with ^{86}N (Asn) in place of ^{86}K (Lys) in fact has poor cointegration ability (Fig. 1).

2. Mutagenesis of the regions flanking the run of adenines. In the IS*1* mutants with a substitution in the region preceding the run of adenines, the mutant with 304C or 305C, each of which causes no alteration of transposase, could mediate cointegration (Fig. 1). This indicates that the nucleotides at positions 304 and 305 are not important for frameshifting.

Either 306G, 306A or 306C inhibited cointegration (Fig. 1). Substitutions 306G and 306A change codon ^{305}TTA for ^{84}L (Leu) to termination codons, whereas 306C changes codon ^{305}TTA for ^{84}L (Leu) to codon TCA for S (Ser), resulting in a serious amino acid change. We assume therefore that the loss of the cointegration ability of these mutants is due to the amino acid effect, but not to the context effect.

Changes of the nucleotide C at 313 following the run of adenines to any other nucleotides inhibited cointegration (Fig. 1). We consider that the loss of the cointegration ability of these IS*1* mutants is mainly due to the amino acid effect, since a mutant transposase with ^{87}V (Val) in place of L (Leu), which would be produced from a mutant with substitution 313G upon successful frameshifting, showed a comparably poor cointegration ability (Fig. 1). This and above assumptions will be confirmed by another experiment described below.

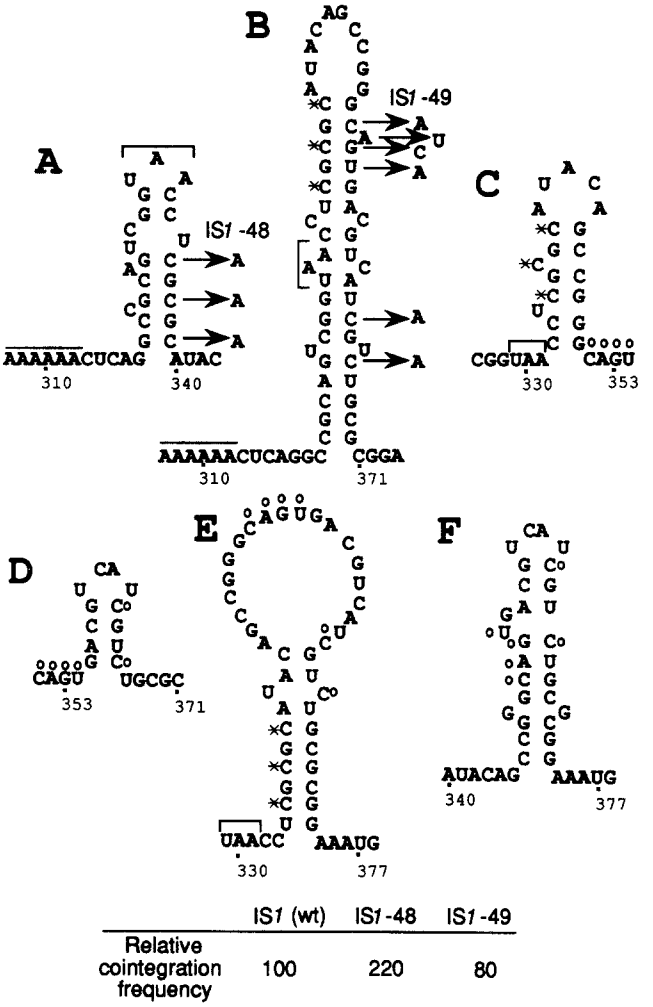
Analysis of the effects of base substitutions on frameshifting by means of lacZ *fusion*

Based on examination of the cointegration ability of the IS*1* mutants in the previous sections, the context effects of substitutions at positions 306, 308, 310 and 313 were ambiguous because their amino acid effects could not be estimated exactly. To determine the context effects of the substitutions at these positions in an alternative way, we constructed plasmid pSEK6000, which contains the 62 bp DNA fragment of IS*1* (corresponding to IS*1* coordinates 292–353), that includes the run of six adenines, flanked by the ATG codon of the *cro* gene of phage λ and the *lacZ* gene, such that *insA* is fused in-frame with ATG_{cro} and B′-*insB* in fused in-frame with *lacZ* (Fig. 2). This DNA fragment is considered to include a region required for efficient frameshifting (Sekine et al. 1992). We also constructed mutant derivatives each with a substitution for the nucleotide at the position of interest. The expression of the reading frame connected to ATG_{cro} is under the control of a thermosensitive repressor, the product of *cI857*, which is also carried by the plasmid. We then measured the β-galactosidase (LacZ) activity in the lysate of cells harboring each plasmid to determine the efficiency of -1 frameshifting required to give the InsA-B′-InsB-LacZ fusion protein upon heat induction.

Substitutions 308C and 310C inhibited the production of LacZ activity, but substitutions 306C and 313T were not completely inhibitory (Fig. 2). These results indicate that the nucleotides at 308 and 310 are both important for frameshifting, but the nucleotides at 306 and 313 are not.

The role of the region located downstream of the run of adenines in frameshifting

Secondary structures downstream of the frameshift site have been reported to be essential for -1 frameshifting in other genetic systems (Jacks et al. 1987, 1988; Brierley et al. 1989; Flower and McHenry 1990; Tsuchihashi and Kornberg 1990; Dinman et al. 1991; Vögele et al. 1991). IS*1* also has possible secondary structures in the region downstream of the run of adenines, as shown in Fig. 3. To determine whether these structures are essential for frameshifting or not, we constructed two IS*1* mutants, IS*1*-48 and IS*1*-49, each with several substitutions which are all silent for amino acids in IS*1* transposase, but



	IS*1* (wt)	IS*1*-48	IS*1*-49
Relative cointegration frequency	100	220	80

Fig. 3. Possible secondary structures of mRNA in the region downstream of the run of six adenines. Numbers in the structures are IS*1* coordinates. The run of six adenines is indicated by a *line* above the nucleotide sequence. The termination codon of *insA* is shown with a *bracket*. Structures (C), (D) and (F) are taken from Hübner et al. (1987). The free energy (ΔG) at 25° C of each structure is calculated according to Tinoco et al. (1973), as follows: (A), -6.4 kcal; (B), -14.4 kcal; (C), -13.2 kcal; (D), -3.6 kcal; (E), -13.8 kcal; (F), -15.2 kcal. IS*1*-48 has a 3 bp substitution which destabilizes the structures (A), (B), (C) and (E), whereas IS*1*-49 has a 6 bp substitution which destabilizes the structures (B), (D) and (F). The base substitutions in IS*1*-48 and IS*1*-49 are shown in structures (A) and (B), respectively; only the positions of substitutions are indicated in the other structures by * and o. The ΔGs of all the resulting structures are greater than 0 kcal. All these substitution mutations cause no changes in the amino acids of the transposase. The frequencies of cointegration mediated by each IS*1* mutant relative to that of wild-type (wt) IS*1* (1.9×10^{-8} per division cycle) are shown

destroy some of the possible secondary structures shown in Fig. 3. IS*1*-48 has substitutions which destroy structures A, B, C and E (Fig. 3), while IS*1*-49 has mutations that destroy structures B, D and F (Fig. 3). Neither of the mutants lost the ability to mediate cointegration (Fig. 3). This shows that these secondary structures are not required for frameshifting in IS*1*.

Sekine et al. (1992) have found that the region 315–332 contains a sequence stimulating the frameshifting

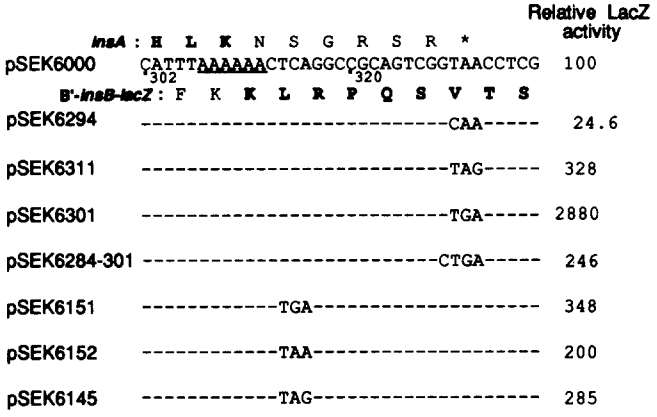


Fig. 4. Effect of the termination codon in *insA* on the LacZ activity specified by pSEK6000 derivatives. The nucleotide sequence of the critical region, 302–336 of IS*1*, is shown together with the amino acids encoded by the two reading frames, *insA* and B'-*insB-lacZ*. In the sequences of pSEK6000 derivatives, only the altered codons are shown. The relative activity of LacZ is calculated taking the LacZ activity (16.9 units) specified by the wild type, pSEK6000, as 100

event. We suspected that the termination codon ^{329}TAA of *insA*, which is located within this region at 17 bp downstream of the run of adenines, is important for efficient frameshifting. To test this, we constructed several derivatives of the IS*1*-*lacZ* fusion plasmid pSEK6000 each with a base substitution in the termination codon, and examined the LacZ activity specified by each plasmid. Plasmid pSEK6294 has codon CAA instead of the ochre codon ^{329}TAA (Fig. 4). This plasmid gave a reduced activity, 24.6% of that of the wild type, pSEK6000 (Fig. 4). This 4.1-fold decrease is almost equal to the degree of reduction in the efficiency of frameshifting (4.5-fold decrease) observed in the plasmid deleted for the termination codon (Sekine et al. 1992). This suggests that the termination codon of *insA* is responsible for efficient frameshifting in IS*1*.

Plasmid pSEK6311, with a substitution changing ^{329}TAA to the amber codon TAG, gave an increased activity (328%, Fig. 4). Plasmid pSEK6301 with a substitution changing ^{329}TAA to the opal codon TGA gave a much increased activity (2880%, Fig. 4). This increase in considered to be mainly due to the phenomenon called translational coupling (see Discussion) in that translation was unusually initiated from a possible initiation codon, ^{328}GTG, created in-frame with *lacZ* by 330G. In fact, plasmid pSEK6284-301, with a sequence altered from ^{328}GTGA to CTGA, which abolished the putative initiation codon GTG, decreased LacZ activity (Fig. 4). Note here, however, that pSEK6284-301 still gave more LacZ activity than the wild type, pSEK6000. These results further indicate that the termination codon of *insA* is important for frameshifting.

Our previous genetic analysis (Sekine and Ohtsubo 1989) has revealed that an IS*1* mutant with 315G, which generates the opal codon ^{314}TGA in *insA* immediately downstream of the run of adenines and upstream of the native termination codon ^{329}TAA of *insA*, mediated cointegration at a frequency 3 times higher than that observed with wild-type IS*1*. This result led us to assume that this termination codon introduced into *insA* at a position upstream of the native termination codon affects the efficiency of frameshifting. To test this assumption directly, we constructed plasmid pSEK6151, a derivative of pSEK6000, with the opal codon ^{314}TGA in *insA* at the same position as that in the IS*1* mutant described above. This plasmid gave a LacZ activity 3.5 times higher than the wild-type, pSEK6000 (Fig. 4), which parallels the increase in cointegration frequency in the IS*1* mutant with the same mutation. We purified the protein with LacZ activity from the cells harboring pSEK6151 and determined its N-terminal sequence. The 17 cycles of Edman degradation revealed a sequence identical to that of the protein specified by the plasmid with the DNA fragment of wild-type IS*1* (data not shown). This indicates that the LacZ activity specified by pSEK6151 was dependent on frameshifting and that the site of frameshifting was not changed.

Other plasmids (pSEK6152 and pSEK6145 in Fig. 4) with a different termination codon ^{314}TAA or ^{314}TAG also gave increased activities (Fig. 4). These results support the suggestion that the termination codon in *insA* downstream of the run of adenines has an important role in efficient frameshifting.

Discussion

The data presented in this paper indicate that nucleotides at positions 308, 309 and 310, corresponding to the second, third and fourth residue in a run of six adenines, respectively, are crucial for contexts responsible for frameshifting in IS*1*. This is consistent with our estimate that the last codon recognized in *insA* during translation of the IS*1* transposase gene is ^{308}AAA, at which the −1 frameshifting event occurs (Sekine et al. 1992). Mutations in the other adenines in the run also reduced the efficiency of frameshifting to a certain extent. This suggests that those adenines are not critically important, but are necessary for frameshifting at the most efficient level.

In many genetic systems, secondary structures downstream of the frameshift site are supposed to cause translating ribosomes to pause at the frameshift site, thereby increasing the probability that the tRNA on the ribosomes will change reading frames (Jacks et al. 1987, 1988; Brierley et al. 1989; Flower and McHenry 1990; Tsuchihashi and Kornberg 1990; Dinman et al. 1991; Vögele et al. 1991). We have shown here that several secondary structures downstream of the run of adenines in IS*1* are, however, not required for frameshifting. This is consistent with the results obtained from the deletion analysis described by Sekine et al. (1992). We have further shown here that the termination codon TAA of *insA*, located 17 bp downstream of the run of adenines, is important for efficient frameshifting. In IS*1*, the termination codon in *insA* might play a role in causing ribosomes to pause, resulting in enhancement of the efficiency of frameshifting. Enhancement of −1 frameshifting by a termination codon immediately downstream

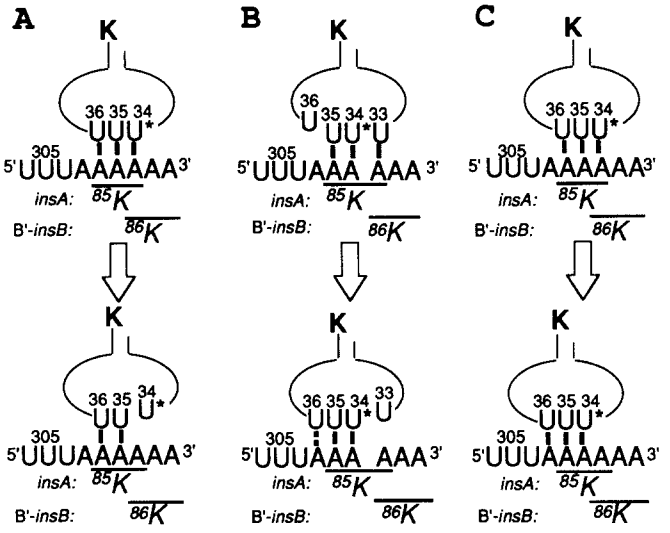


Fig. 5A–C. Possible models for translational −1 frameshifting in IS*1*. Base pairing between the anticodon of $tRNA^{Lys}$ and the mRNA encoding the IS*1* transposase is shown at the top of each panel. Nucleotide U* at position 34 in $tRNA^{Lys}$ denotes 5-(methylaminomethyl)-2-thiouridine ($^{mnm5s2}U$) (Chakraburtty et al. 1975). **A** Transient disengagement of base pairing between U*-34 of $tRNA^{Lys}$ and A-310, the third base of codon ^{308}AAA for ^{85}K (Lys). **B** Unusual base pairings of codon ^{308}AAA in mRNA with U-35, U*-34 and U-33 of $tRNA^{Lys}$, followed by disengagement of U-33. Base pairing between A-307 of mRNA and U-36 in the normal anticodon of $tRNA^{Lys}$ may be formed. **C** Slippage of $tRNA^{Lys}$ in the −1 direction. Note that each event permits A-310 to be available for base pairing with the anticodon of $tRNA^{Lys}$ reading codon ^{310}AAA

of the frameshift site has also been reported in an artificial context (Weiss et al. 1987).

The results in this paper were obtained by two methods, cointegration mediated by IS*1* and the LacZ assay of the products from IS*1*-*lacZ* fusion genes. In the LacZ assay, we observed that LacZ activity measured in one case derives from an unusual translation initiation event at a codon in-frame with *lacZ*, as well as from translation of *lacZ* which is dependent on −1 frameshifting. Introduction of the termination codon TGA into *insA* instead of TAA resulted in an unusually high activity of LacZ due to translation initiated from codon GTG in B′-*insB*-*lacZ*, which overlaps the introduced TGA codon. Since a reasonable ribosome binding site is absent in the upstream region of the initiation codon created, initiation of translation from the new initiation codon in B′-*insB*-*lacZ* might have occurred in a manner that depends on translation of the upstream reading frame (*insA*). Such a phenomenon is known as translational coupling, and was first described for the tryptophan operon in *E. coli* (Oppenheim and Yanofsky 1980) and subsequently in many other genetic systems. Obviously in our analysis described above, we have not considered this case of unusual initiation of translation in our interpretation of the requirements for frameshifting in IS*1*.

Based on our results described above, it is reasonable to presume that a tRNA for Lys ($tRNA^{Lys}$) recognizing codon ^{308}AAA in *insA* plays a key role in frameshifting. How does $tRNA^{Lys}$ mediate −1 frameshifting? It is expected that the anticodon of $tRNA^{Lys}$ interacts with the codon ^{308}AAA for ^{85}K (Lys) in *insA*, as shown in Fig. 5A (top), and that the nucleotide A at position 310 (designated as A-310) is somehow recognized by the next $tRNA^{Lys}$ reading ^{310}AAA for ^{86}K (Lys) in B′-*insB* to initiate the −1 frameshifting event. It is possible that after $tRNA^{Lys}$ recognizes codon ^{308}AAA, the base pairing between U-34 in the anticodon and A-310 of the codon is disengaged transiently to permit A-310 to became available for the first base of codon ^{310}AAA in B′-*insB* for ^{86}K (Lys) (Fig. 5A). Note here, however, that U-34 in the anticodon of $tRNA^{Lys}$ is modified to 5-(methylaminomethyl)-2-thiouridine ($^{mnm5s2}U$; see U* in Fig. 5) in *E. coli* (Chakraburtty et al. 1975), and that such modifications in U stabilize base pairing with A (Scheit and Faerber 1975; Yokoyama et al. 1985). Therefore, disengagement of this stable base pairing is assumed not to be a frequent event in induction of a −1 shift of the reading frame.

It is also likely that the nucleotide at a position preceding the normal anticodon in a tRNA may participate in recognition of the codon, as has been proposed by Weiss (1984), such that in IS*1*, U-33 preceding the anticodon U*UU of $tRNA^{Lys}$ is engaged in unusual pairing with codon ^{308}AAA in mRNA together with U*-34 and U-35, and then disengaged to free A-310 for base pairing with the anticodon of the next tRNA (Fig. 5B). In this model, it is possible that at the same time as the disengagement of U-33, another pairing between U-36 in the normal anticodon and A-307 of the mRNA may occur to stabilize the $tRNA^{Lys}$ on the mRNA (see Fig. 5B). However, since alteration of A-307 permitted frameshifting to occur, as described in Results, the base pairing between U-36 and A-307 of mRNA may not be essential.

Slippage of $tRNA^{Lys}$ reading the codon ^{308}AAA in *insA* back one nucleotide into B′-*insB* is also possible (Fig. 5C). The four nucleotides at positions 307–310 in the mRNA are all adenines, and therefore the anticodon of $tRNA^{Lys}$ can base pair with the mRNA after slippage. If the slippage model is correct, a substitution at A-307 would inhibit frameshifting due to disruption of the site to which the $tRNA^{Lys}$ slips. However, as described in Results, the mutants with a substitution at position 307 still retained their ability to mediate cointegration. We assume, therefore, that the slippage model shown in Fig. 5C is questionable.

The simultaneous slippage model has been proposed as the mechanism for frameshifting on retroviral mRNA (Jacks et al. 1988). In this model, two adjacent tRNAs in the 0-frame, resident in the ribosomal P site and A site, respectively, slip back by one nucleotide, where the A site codon is the site of frameshifting. Wild-type IS*1* has nucleotide sequence T TTA AAA at positions 304–310 which seems to fit the sequence requirement of this model. However, as described in Results, the mutant IS*1* with 304C or 305C substitution, where the tRNA in the P site cannot make stable base pairs after a −1 slippage, still mediated cointegration. This suggests that the simultaneous slippage model cannot explain the −1 frameshifting in IS*1*.

Acknowledgements. We would like to thank H. Nagasawa for helping with amino acid sequencing and S. Yokoyama for valuable information. We thank D. Bastia for a generous gift of plasmid pR-pMLB. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan. Y.S. was the recipient of a JSPS Fellowship for Japanese Junior Scientists.

References

Atkins JF, Weiss RB, Gesteland RF (1990) Ribosome gymnastics – Degree of difficulty 9.5, style 10.0. Cell 62:413–423

Brierley I, Digard P, Inglis SC (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell 57:537–547

Casadaban MJ, Cohen SN (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J Mol Biol 138:179–207

Chakraburtty K, Steinschneider A, Case RV, Mehler AH (1975) Primary structure of $tRNA^{Lys}$ of *E. coli* B. Nucleic Acids Res 2:2069–2075

Dinman JD, Icho T, Wickner RB (1991) A −1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a gag-pol fusion protein. Proc Natl Acad Sci USA 88:174–178

Flower AM, McHenry CS (1990) The γ subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. Proc Natl Acad Sci USA 87:3713–3717

Hashimoto-Gotoh T, Sekiguchi M (1977) Mutations to temperature sensitivity in R plasmid pSC101. J Bacteriol 131:405–412

Hübner P, Iida S, Arber W (1987) A transcriptional terminator sequence in the prokaryotic transposable element IS*1*. Mol Gen Genet 206:485–490

Jacks T, Townsley K, Varmus HE, Majors J (1987) Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary virus *gag*-related polyprotein. Proc Natl Acad Sci USA 82:2829–2833

Jacks T, Madhani HD, Masiarz FR, Varmus HE (1988) Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. Cell 55:447–458

Kunkel TA, Roberts JD, Zakour RA (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol 154:367–382

Machida Y, Machida C, Ohtsubo E (1982) A novel type of transposon generated by insertion element IS*102* present in a pSC101 derivative. Cell 30:29–36

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular Cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Messing J (1983) New M13 vectors for cloning. Methods Enzymol 101:20–78

Ohtsubo E, Zenilman M, Ohtsubo H, McCormick M, Machida C, Machida Y (1981) Mechanism of insertion and cointegration mediated by IS*1* and Tn*3*. Cold Spring Harbor Symp Quant Biol 45:283–295

Ohtsubo H, Ohtsubo E (1978) Nucleotide sequence of an insertion element, IS*1*. Proc Natl Acad Sci USA 73:2316–2320

Oppenheim D, Yanofsky C (1980) Translational coupling during expression of the tryptophan operon of *Escherichia coli*. Genetics 95:785–795

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Scheit KH, Faerber P (1975) The effects of thioketo substitution upon uracil-adenine interactions in polyribonucleotides. Eur J Biochem 50:549–555

Sekine Y, Ohtsubo E (1989) Frameshifting is required for production of the transposase encoded by insertion sequence *1*. Proc Natl Acad Sci USA 86:4609–4613

Sekine Y, Ohtsubo E (1991) Translational frameshifting in IS elements and other genetic systems. In: Kimura M, Takahata N (ed) New Aspects of The Genetics of Molecular Evolution. Japan Scientific Societies Press, Tokyo/Springer-Verlag, Berlin, pp 243–261

Sekine Y, Nagasawa H, Ohtsubo E (1992) Identification of the site of translational frameshifting required for production of the transposase encoded by insertion sequence IS*1*. Mol Gen Genet 235:317–324

Tinoco I Jr, Borer PN, Dengler B, Levine MD, Uhlenbeck OC, Crothers DM, Gralla J (1973) Improved estimation of secondary structure in ribonucleic acids. Nature New Biol 246:40–41

Tsuchihashi Z, Kornberg A (1990) Translational frameshifting generates the γ subunit of DNA polymerase III holoenzyme. Proc Natl Acad Sci USA 87:2516–2520

Vieira J, Messing J (1987) Production of single stranded plasmid DNA. Methods Enzymol 153:3–11

Vögele K, Schwartz E, Welz C, Schiltz E, Rak B (1991) High-level ribosomal frameshifting directs the synthesis of IS*150* gene products. Nucleic Acids Res 19:4377–4385

Weiss RB (1984) Molecular model of ribosome frameshifting. Proc Natl Acad Sci USA 81:5797–5801

Weiss RB, Dunn DM, Atkins JF, Gesteland RF (1987) Slippery runs, shifty stops, backward steps, and forward hops: −2, −1, +1, +2, +5, and +6 ribosomal frameshifting. Cold Spring Harbor Symp Quant Biol 52:687–693

Yokoyama S, Watanabe T, Murao K, Ishikura H, Yamaizumi Z, Nishimura S, Miyazawa T (1985) Molecular mechanism of codon recognition by tRNA species with modified uridine in the first position of the anticodon. Proc Natl Acad Sci USA 82:4905–4909

Yoshioka Y, Ohtsubo H, Ohtsubo E (1987) Repressor gene *finO* in plasmids R100 and F: Constitutive transfer of plasmid F is caused by insertion of IS*3* into F *finO*. J Bacteriol 169:619–623

Communicated by M. Sekiguchi