Repression of IS200 transposase synthesis by RNA secondary structures

Carmen R. Beuzón, Silvia Marqués and Josep Casadesús*

Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Apartado 1095, E-41080 Sevilla, Spain

Received May 28, 1999; Revised and Accepted August 2, 1999

ABSTRACT

The IS200 transposase, a 16 kDa polypeptide encoded by the single open reading frame (ORF) of the insertion element, has been identified using an expression system based on T7 RNA polymerase. In wild-type IS200, two sets of internal inverted repeats that generate RNA secondary structures provide two independent mechanisms for repression of transposase synthesis. The inverted repeat located near the left end of IS200 is a transcriptional terminator that terminates read-through transcripts before they reach the IS200 ORF. The terminator is functional in both directions and may terminate ~80% of transcripts. Another control operates at the translational level: transposase synthesis is inhibited by occlusion of the ribosome-binding site (RBS) of the IS200 ORF. The RBS (5’-AGGGG-3’) is occluded by formation of a mRNA stem–loop structure whose 3’ end is located only 3 nt upstream of the start codon. This mechanism reduces transposase synthesis ~10-fold. Primer extension experiments with AMV reverse transcriptase have provided evidence that this stem–loop RNA structure is actually formed. Tight repression of transposase synthesis, achieved through synergistic mechanisms of negative control, may explain the unusually low transposition frequency of IS200.

INTRODUCTION

IS200 is an insertion sequence of 707–710 bp, originally found in the genome of Salmonella typhimurium LT2 (1). IS200 elements are abundant in the genus Salmonella, both on the chromosome and on plasmids (1–6). Insertion elements with various degrees of relatedness to IS200 have also been described in Shigella (3), Escherichia coli (4), Yersinia (7,8), Vibrio (9) and Clostridium (10). Furthermore, a recent survey of IS structure and organisation has suggested that IS200 may be related to the IS605 element of Helicobacter pylori (11).

A long-known and enigmatic trait of IS200 behaviour is its poor contribution to spontaneous mutagenesis (12,13), which can be correlated with the extremely low transposition frequency of the element (14,15). In fact, only a few IS200 insertion mutations have been characterised in S.typhimurium: the hisD984 mutation whose analysis led to the discovery of IS200 (1), an insertion in the gpt gene (13), and an insertion in the pef operon of the pSLT virulence plasmid (6). Hunts for IS200-induced mutants in S.typhimurium, sometimes involving positive selection strategies, have confirmed that transposition is rare (12–15). Furthermore, surveys carried out in field isolates have indicated that IS200 transposition is likewise infrequent in natural populations of Salmonella (5).

Previous studies have shown that IS200 elements from Salmonella have a highly conserved structure (6,16). Relevant traits shared by all IS200 copies are the absence of terminal repeats, direct or inverse (6), and the presence of an open reading frame (ORF) encoding a putative peptide of 152 amino acids (6,16,17). This ORF is driven by a weak promoter located between nucleotides 1–41 (15). Structural analysis has also shown that the ‘left’ end of IS200 contains two sets of inverted repeats, which are highly conserved among IS200 copies. Upon transcription, the repeats can be expected to form a hairpin and a stem–loop (6).

This study shows that the internal inverted repeats of IS200 behave as regulatory signals which repress synthesis of the single IS200-encoded protein, the putative transposase. The hairpin located near the IS200 left end acts as a transcriptional terminator which terminates impinging transcripts, while the stem–loop element located near the beginning of the IS200 ORF occludes the ribosome-binding site of the IS200 transposase gene. The existence of signals for both transcriptional termination and translational repression, combined with the weakness of the IS200 transposase promoter (15), seem to bring an explanation for the extremely low transposition frequency of IS200.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Escherichia coli DH5α (18) was used for cloning purposes. LT2 was used as the wild-type strain of S.typhimurium. Salmonella abortusovis SAO44 (19) was provided by S. Rubino, Istituto di Microbiologia, Università degli Studi di Sassari, Sassari, Italy. Unless indicated otherwise, bacteria were grown in LB at 37°C. When necessary, antibiotics were used at the following concentrations: ampicillin, 100 μg/ml; kanamycin, 50 μg/ml; chloramphenicol, 20 μg/ml.

*To whom correspondence should be addressed. Tel: +34 95 455 7105; Fax: +34 95 455 7104; Email: genbac@cica.es
Present address:
Carmen R. Beuzón, Department of Infectious Diseases, Imperial College of Science, Technology and Medicine, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK
Plasmids

Plasmid pIZ44 is a pUC9 derivative bearing a His6II fragment of the S.typhimurium chromosome which includes the insertion hisD984::IS200 (3). pIZ250 is a derivative of pBluescript II SK(+) carrying a 1 kb StuI–SstI fragment of pIZ44 cloned in the Smal site (15). The cloned fragment includes the whole IS200 element and flanking his regions. pIZ949 carries a Sall–SacI fragment of pIZ50 cloned in the multiple cloning site (MCS) of pT7-6 (20). The construction places the single IS200 ORF, including the stem–loop-generating structure, under the control of the T7 promoter. To obtain a mutated version of IS200 unable to form the mRNA stem–loop, a 256 bp HindIII–EcoRI fragment which included the stem-forming repeats and 154 bp of the IS200 ORF was cloned onto pBluescript II SK(+) digested with the same enzymes; this was the origin of GATCTGCA-AAAGTTCA-3'. The resulting plasmid was used in site-directed mutagenesis using the single-strand procedure of Bio-Rad (Richmond, CA) to generate a 70 bp deletion in the region which included the stem–loop structure. The primer used was 5'-CGACGAATC-ACGGGATCCATCATGCAG-3', complementary to nucleotides 235–256 of IS200 (6; EMBL accession number X56834), was end-labelled with [γ-32P]ATP, and annealed to 100 µg of total RNA prepared from S.typhimurium LT2 bearing plasmid pIZ1500. The oligonucleotide used is complementary to an internal region of the single IS200 ORF (6). For annealing, 102 c.p.m. of the oligonucleotide were used. cDNA was synthesised using AMV reverse transcriptase, as described elsewhere (28). The products of reverse transcription were analysed in urea–polyacrylamide sequencing gels. For autoradiography, gels were exposed to a Kodak Biomax MR film.

DNA manipulations

Plasmids were extracted by the alkaline lysis method (24). Enzymes were purchased from Boehringer Mannheim. DNA transformation and electroporation were performed using standard methods (25).

β-galactosidase assays

Cells were grown in LB broth until mid-exponential phase. Assays were performed by the method of Miller (26).

RNA manipulations

RNA was extracted with a modification of the guanidium isothiocyanate–phenol method (27). The oligonucleotide 5'CTGC-CTACTGCCCTACGCTTCT-3', complementary to nucleotides 235–256 of IS200 (6; EMBL accession number X56834), was end-labelled with [γ-32P]ATP, and annealed to 100 µg of total RNA prepared from S.typhimurium LT2 bearing plasmid pIZ1500. The oligonucleotide used is complementary to an internal region of the single IS200 ORF (6). For annealing, 102 c.p.m. of the oligonucleotide were used. cDNA was synthesised using AMV reverse transcriptase, as described elsewhere (28). The products of reverse transcription were analysed in urea–polyacrylamide sequencing gels. For autoradiography, gels were exposed to a Kodak Biomax MR film.

In vivo radioactive labelling

For protein expression experiments, strains carrying plasmid pGPI-2 and either of the plasmids pT7-6 and pIZ895 were grown to an OD600 = 0.4 in LB with IPTG (1 mM) at 30°C. The cultures were washed twice in E minimal medium with the corresponding antibiotics and incubated at 42°C for 90 min. Rifampicin (200 µg/ml) was then added to the cultures and incubation at 42°C was continued for 30 min. Ten µCi of 35S-labelled methionine and cysteine (labelling grade L-, Amersham) were added to each culture, and the preparations were incubated during 5 min at 37°C. The cells were then collected by centrifugation and lysed with 4% SDS.

Protein electrophoresis

Cellular extracts (5–10 µl) were loaded on polyacrylamide–SDS gels (29) and run at 15–20 mA in the stacking gel and 30–35 mA in the separating gel. Gels were dehydrated with 1% glycerol and 10% acetic acid, and dried in a Slab Gel Dryer, model SE-1160 (Hoeffer, San Francisco, CA) during 1 h at 80°C. Radioactively labelled material was detected by exposure to a Kodak Biomax X-ray film.

RESULTS

Transcription termination upstream of the IS200 ORF

A highly conserved structure found in all IS200 elements is an inverted repeat near the ‘left’ end (Fig. 1). If transcribed, the inverted repeat can be expected to give rise to the typical hairpin structure of a Rho-independent transcriptional terminator, with a predicted ΔG° = −14.1 kcal/mol (6). In fact, it is well known...
that insertions of IS200 are polar at least in one orientation, and this polarity depends on a DNA stretch which includes the hairpin-forming structure (30).

To examine whether the leftmost inverted repeat of IS200 had the properties of a transcriptional terminator, a 194 bp DNA fragment including the hairpin-forming structure was probed in the terminator-test vector pIC551 (21). This expression system relies on the PR promoter of bacteriophage λ, located upstream of a reporter lacZ gene preceded by an MCS. A linker containing a Rho-independent transcription terminator is present upstream of the PR promoter to stop transcription from external (plasmid) promoters (21). This design permits unequivocal detection of transcriptional terminators: if a DNA stretch containing a transcriptional terminator is cloned in the MCS, the presence of the terminator between the promoter and the lacZ gene will prevent RNA polymerase access into lacZ; hence a decrease in β-galactosidase activity will be observed (21). Plasmids pIZ882 and pIZ883 bear the 23 bp hairpin-forming sequence cloned in the Smal site between PR and lacZ (each plasmid in one orientation). Figure 2 shows the β-galactosidase activities of these constructs compared with a control that lacks the hairpin-forming fragment. While the latter yielded high levels of β-galactosidase activity, the DNA fragment containing a terminator prevented RNA polymerase from accessing the lacZ gene, resulting in a 10-fold decrease in β-galactosidase activity compared with the control (21).

Repression of IS200 mRNA translation

Downstream of the hairpin structure discussed above, two additional, larger inverted repeats are found (Fig. 1). When transcribed, these repeats can give rise to a stem–loop structure with a predicted ΔG = −24.8 kcal/mol (6). The rightmost repeat of this structure finishes 3 bp upstream of the start codon of the IS200 ORF. If formed, the stem of the structure can be expected to occlude the predicted ribosome-binding site of the ORF (6).

To examine whether this stem–loop structure affected translation of the IS200 ORF, translational lac fusions were constructed on plasmid pIZ888. This vector carries a lacZ gene (without its own transcriptional and translational initiation signals) under the control of the lac UV5 promoter, and an MCS located between the promoter and the lacZ gene. Use of the lac UV5 promoter instead of the IS200 native promoter is justified because the latter yields β-galactosidase levels which are too low to permit easy detection of differences in expression (15). For cloning, a 256 bp EcoRI–HindIII fragment carrying the stem–loop structure of IS200 and the initial 154 bp of the IS200 ORF was introduced in the MCS of pIZ888 to generate pIZ891. Using site-directed mutagenesis, the left arm of the stem and half of the loop sequence were deleted, and the resulting DNA fragment was cloned in the same site of pIZ888 to obtain pIZ892. Plasmids pIZ888, pIZ891 and pIZ892 were examined for their ability to direct synthesis of β-galactosidase in the presence of IPTG. The host strain was an LT2 derivative bearing a lacP allele on plasmid pZZ227 (23). Relevant results, summarised in Figure 3, were as follows: (i) plasmid pIZ888 (the vector without cloned DNA) gave nearly undetectable levels of β-galactosidase activity. (ii) pIZ891, in which the lacZ gene is preceded by the stem–loop-generating repeats of IS200, yielded low levels of β-galactosidase. (iii) pIZ892, which contains the ribosome binding site and the start codon of the IS200 ORF but cannot form the stem–loop structure, yielded a β-galactosidase activity 10-fold higher than pIZ891.
These data suggest that the stem–loop structure is actually formed in vivo, and that it represses translation of any protein-coding sequence starting at the ATG located immediately downstream of the stem.

**Formation of the RNA stem–loop**

Results presented in the former section suggested that the predicted stem–loop structure that precedes the start codon of the IS200 ORF was actually formed in vivo: removal of the structure by site-directed mutagenesis caused a 10-fold increase in β-galactosidase synthesis. Direct evidence for the formation of such a structure was provided by RNA analysis. It is known that RNA regions with highly stable secondary structure, if used as templates for reverse transcription, can cause premature termination of RNA synthesis by AMV reverse transcriptase (33,34). In these cases, the 5' end of the shorter, prematurely terminated cDNA molecules maps at the bottom of the structure (34). To investigate whether the IS200 RNA stem–loop terminated AMV–RT synthesis, total RNA was isolated from *S.typhimurium* cells bearing plasmid pIZ1500 and reverse primer extension of 100 µg of total RNA was carried out, using an end-labelled oligonucleotide complementary to the IS200 ORF. Figure 4 shows that >90% of the extended products had their 3' end downstream of the predicted structure, and that the longest of them finished exactly 2 nt downstream of the 3' end of the stem. Smaller products with 5' ends at 6, 12 and 20 nt downstream of the start codon were also detected. No other band of similar intensity was observed with normal exposure times. No band was observed after primer extension of the same amount of total RNA isolated from a strain lacking plasmid pIZ1500 (not shown). In an attempt to destabilise the double-stranded RNA structures formed, the extension was repeated at three different temperatures, 44, 50 and 60°C (the highest temperature compatible with AMV–RT activity), but no differences in the band pattern were observed (data not shown). This was not unexpected, since Zucker’s algorithm (35) predicts a Tm of 80°C for the RNA stem, far above the highest temperature used in our assays. Similar results were obtained when the reverse primer extension assays were carried out on RNA templates from *S.typhimurium* LT2 and *S.abortusovis* SAO44, although the intensity of the bands was lower (data not shown). We thus conclude that the predicted mRNA stem–loop is actually formed, and that it represses translation of the IS200 ORF. Confirmation of this model will require further experiments of site-directed mutagenesis, to introduce mutations that perturb formation of the stem–loop structure but do not remove large amounts of genetic material.

**Cloning and expression of the IS200 ORF**

To ascertain whether the IS200 ORF was indeed capable of directing protein synthesis, we cloned an IS200 fragment containing the ORF in the MCS of the expression vector pT7-6 (20). The ORF had been previously mutagenised to delete one of the stem-forming repeats located upstream of the ORF. In
the resulting plasmid, pIZ895, the ORF was placed under the control of the strong T7 promoter present in pT7-6 (20). When supplied with T7 RNA polymerase (provided in trans by plasmid pGP1-2), pIZ895 was able to direct rifampicin-resistant synthesis of a major, 16 kDa polypeptide (Fig. 5). The autodigestion study shows the only radiolabelled protein detected upon a pulse–chase experiment (not shown). The size of the protein detected (16 kDa) agrees with the size of the IS200 after incubation in the presence of rifampicin. The size of this protein increases IS200 TnpA (transposase), on the following grounds: (i) overproduction of the ORF, which predicts a peptide of 152 amino acids. Therefore, the size of the protein product was inferred from a protein size marker run in parallel (not shown). Protein expression from plasmid pT7-6 without insert was used as a negative control.

DISCUSSION

Most, if not all, prokaryotic insertion elements undergo a tight control of transposase synthesis, which can be viewed as part of a self-restraint strategy and may provide evolutionary advantages (36). Control of transposase synthesis is achieved through a variety of mechanisms: negative control of the transposase promoter by trans-acting factors encoded by the insertion element (37,38) or by the host (39); repression of the transposase promoter by DNA methylation (40); inhibition of transposase mRNA translation by antisense RNA (41); programmed translational frameshifting (42); and assembly of a transposase promoter upon programmed DNA rearrangements (43). Mechanisms that avoid uncontrolled transposase synthesis resulting from impinging transcription are also known (31,32), and will be discussed below.

IS200 is a transposable element with extreme self-restraint, as indicated by the scarcity of IS200-induced mutations (12,13), the low frequency of IS200-induced rearrangements (14,15) and the constancy of IS200 fingerprints in natural populations (5). This behaviour correlates with the low expression of the single IS200 ORF. One cause of this low expression is the existence of a weak promoter (15), a feature common to most transposase genes. In addition, this study shows the existence of two additional mechanisms which contribute to keep IS200 transposase synthesis low: the termination of read-through transcripts and the repression of transposase synthesis by an mRNA secondary structure. The DNA sequence elements that embody these structural motifs are as follows: (i) the inverted repeats located near the ‘left’ end of IS200 (nucleotides 9–32 in the hisD984::IS200 copy; see ref. 6) encode a transcriptional terminator active in both orientations. Its structure is reminiscent of that of a Rho-independent terminator (44), but its independence from Rho has not been tested. This terminator is not located at the end of any gene but upstream of the IS200 transposase gene, and terminates >80% of impinging transcripts. Interestingly, this region overlaps with the IS200 tnpA promoter, but the significance of this fact (if any) has not been analysed. (ii) The larger inverted repeats located downstream of the IS200 tnpA promoter (nucleotides 66–85 and 120–139 in the hisD984::IS200 copy; see ref. 6) give rise to a stem–loop structure which represses translation of the tnpA gene. The stability of the stem–loop is such that the structure remains double-stranded at 60°C in vitro.

If transcriptional termination prevents >80% of read-through transcripts from reaching the transposase gene and translational repression decreases 10-fold the efficiency of translational initiation, any impinging transcript which escapes the terminator will encounter the translational inhibitor, and the combination of two separate hurdles will secure repression. Of course, transcripts initiated at the IS200 tnpA promoter will encounter only one hurdle, the mRNA stem–loop, but the combination of low transcription levels and efficient translational repression can be expected to allow transposase synthesis only rarely. Interestingly, a mechanism to inhibit transposase synthesis from impinging transcripts has been previously described in both IS10 and IS50 (31,32). The mechanism is based upon formation of an RNA stem–loop structure which sequesters translation initiation signals. Shorter transcripts initiated at the transposase promoter lack one of the repeats required to make the stem, which is thus not formed (31,32). Therefore these
insertion elements make use of translational inhibition to counter the effects of inward transcription, while IS200 has two separate mechanisms for transcription termination and repression of translation. This design appears efficient both for avoiding uncontrolled transposase synthesis and for keeping transposase levels low even when transcripts are initiated at the IS200 promoter. In fact, unequivocal detection of the 16 kDa polypeptide predicted by sequence analysis required both the use of a heterologous promoter and the removal of the stem–loop mRNA structure by site-directed mutagenesis.

A question unanswered in this study is whether the translational repression exerted by the mRNA stem–loop can be lifted under mutational events (6). Mechanisms for disruption or removal of RNA secondary structures have been discussed in the literature. These bursts are not a consequence of IS200 transposition by disrupting the stem–loop. The possibility that IS200 is regulated receives support from the old and enigmatic observation that IS200 undergoes ‘bursts’ of transposition in stab cultures kept at room temperature for months or years (1,15). These bursts are not a consequence of IS200 ‘activation’ by environmental signals can trigger IS200 replication when inward transcription, while IS200 is a transposable element with intrinsically low activity, could induce IS200 transposition. An alternative view, that IS200 is a transposable element with intrinsically low activity, is supported by the observation that IS200 transposition is infrequent not only in the laboratory but also in natural populations of Salmonella (5). According to Doolittle et al. (36), a strategy of this kind will minimise the risk of causing mutations detrimental to the host, thereby favouring the maintenance of the element in natural populations. IS200 may fit in this scenario because sequence comparisons indicate that the element is extremely old (4,6,16). The possibility that the evolutionary endurance of IS200 has been favoured by the possession of tight mechanisms of transposase control is an attractive speculation.

ACKNOWLEDGEMENTS

We thank John Roth, Ken Haack, Fernando Govantes and Eduardo Santero for helpful discussions. This work was supported by grant PB93/649 from the DGICYT, Spain.

REFERENCES