The EMBO Journal vol.4 no.10 pp.2687–2693, 1985

Promotion of RNA transcription on the insertion element IS30 of *E. coli* K12

Brian Dalrymple[1] and Werner Arber

Department of Microbiology, Biozentrum of the University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

[1]Present address: CSIRO, Division of Molecular Biology, PO Box 184, Sydney 2113, Australia

Communicated by W.Arber

Two promoters of RNA transcription have been identified on IS30 by an *in vivo* assay, in which various DNA fragments with IS30 sequences were inserted in front of the promoterless *galK* gene of plasmid pFD51. Both promoters have a similar activity of ~10% of the activity of the *lacUV5* promoter. Promoter P30A precedes the long open reading frame (ORFA), and its proposed −35 region lies within the left-hand terminal inverted repeat of IS30. However, the apparent activity of promoter P30A is significantly reduced when measured in the 3′ region of ORFA. Thus, either the activity of promoter P30A is controlled by an IS30-encoded product from the same element, or some termination of transcription from P30A occurs within the coding region of ORFA. Promoter P30C precedes a short open reading frame (ORFC) in-frame with ORFA, but in the opposite strand. Reading frame ORFC is closely followed by a terminator of RNA transcription, T30C. None of the other potential open reading frames predicted from the DNA sequence, with one possible exception, are preceded by a promoter of RNA transcription active in the assay. No significant transcription was detected out of the left-hand end of the complete element. However, a small amount, probably due to read-through from promoter P30A, was detected out of the right-hand end of a complete copy of IS30. In addition the right-hand end of IS30 has been shown to have the potential to create promoters by insertion .

Key words: galactokinase/insertion element IS30/promoters/transcription

Introduction

IS30 is a resident IS element of *Escherichia coli* K12. It is unrelated at the level of the DNA sequence to IS1, IS2, IS3, IS4 and IS5, the other well-characterized resident IS elements of *E. coli* (see Dalrymple *et al.*, 1984; Timmerman and Tu, 1985). The 1221 bp long sequence of IS30 predicts one long open reading frame (ORFA), which could encode a protein of 383 amino acids and which is preceded by possible transcription and translation control signals. The −35 region of this potential promoter sequence is contained within the left-hand terminal inverted repeat, while the −10 region is contained in an internal symmetrical sequence (Dalrymple *et al.*, 1984).

Using a variety of techniques, a number of promoters of RNA transcription carried by IS elements have been characterized. Promoters *ins*P_L and *ins*P_R of IS1 (Chan and Lebowitz, 1982; Machida *et al.*, 1984), two promoters $P_{IS(II)}$ and $P'_{IS(II)}$ in front of the long open reading frame of IS2 (Hinton and Musso, 1983), and a promoter preceding the long open reading frame of IS50 (Johnson and Reznikoff, 1981) have been localized by *in vitro* techniques. In addition, three promoters have been identified on IS10, one of which precedes the long open reading frame (Simons *et al.*, 1983). Using operon fusions to the *galK* gene, three promoters have also been identified on IS5 (Rak *et al.*, 1982; Rak and von Reutern, 1984). We have used the same system, in which a series of fragments of IS30 were cloned into pFD51 (Rak and von Reutern, 1984), a derivative of pKO1, which carries a promoterless *galK* gene preceded by a cloning array and by a long leader sequence to ensure uniform translation efficiency (McKenney *et al.*, 1981; Rosenberg *et al.*, 1983). Three stop codons, one in each possible open reading frame, ensure that galactokinase, rather than a fusion protein, is synthesized when a promoter of RNA transcription, together with part of a gene, is cloned in front of the *galK* gene.

Two mechanisms for controlling the activity of the promoter which precedes the long open reading frame of an IS element have been demonstrated. In the first system an anti-RNA is synthesized from pOUT of IS10 and inhibits transcription/translation of the mRNA from pIN (Simons and Kleckner, 1983; Simons *et al.*, 1983). In the second system methylation of an A residue in the −10 region of the promoter pIN of IS10, by the *dam* gene product, reduces the activity of the promoter ~10-fold (Kleckner *et al.*, 1984). It has also been proposed that transcription from the promoter preceding the long open reading frame of IS903, in Tn903, is affected by methylation, and additionally IS4, IS5 and IS50 have one or more sites for Dam methylation in the region of the promoters preceding their long open reading frames (Kleckner *et al.*, 1984).

Several IS elements activate genes adjacent to their site of insertion (see Iida *et al.*, 1983). IS3 and IS10, for example, carry an internal promoter which can transcribe adjacent genes (Charlier *et al.*, 1982; Simons *et al.*, 1983), while IS2 (Glansdorff *et al.*, 1980) and IS140 (probably identical to IS26) (Bräu *et al.*, 1984) can create promoters upon insertion into a new site. In these latter cases a −35 region carried on the IS element and a −10 region near the site of insertion of the IS element comprise a new promoter (Hinton and Musso, 1982; Jaurin and Normark, 1983).

We present here the results of experiments designed to locate and characterize functionally the promoters of RNA transcription present on IS30. This should help to identify the genes or RNA molecules which may be involved in transposition or in the control of expression of genes located adjacent to transpositional insertion sites. We also address the question of potential regulation mechanisms for the expression of the presumed transposase gene.

Results

To identify promoters of RNA transcription carried by IS30, a series of fragments of the element were cloned into pFD51 (Rak and von Reutern, 1984). According to the experimental design, most of the fragments contained only IS30 DNA or sequenced linker DNA from the vectors used in the constructions. This

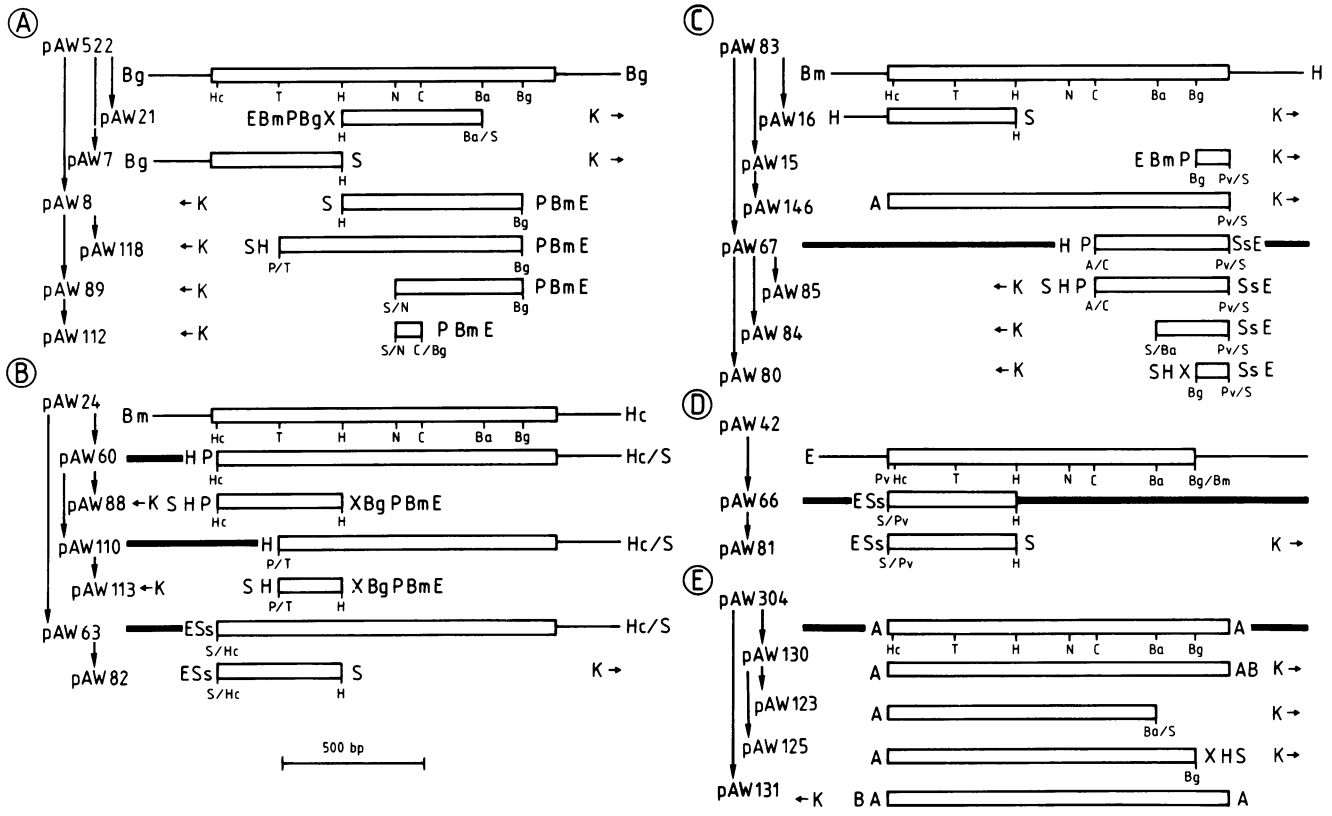

Fig. 1. The method of construction of plasmids with IS30 sequences, derived respectively from (**A**) pAW522, (**B**) pAW24, (**C**) pAW83, (**D**) pAW42 and (**E**) pAW304. The vertical arrows show the lineages of the plasmids. The open boxes represent IS30 DNA, the thick lines represent pUC7 or pUC12 DNA and the thin lines represent P1, NR1-Basel or pBR322 DNA. Only relevant parts of the plasmids are shown and only IS30 DNA is drawn to scale. K represents the galactokinase gene together with its direction of transcription. All the sites for the restriction enzymes shown are drawn on the maps of the parent plasmids, but not necessarily on the maps of their derivatives. A, represents half of the symmetrical linker from pUC7 (*Pst*I, *Sal*I, *Bam*HI, *Eco*RI), and B represents the linker from pFD51 (*Eco*RI, *Bam*HI, *Pst*I, *Bgl*II, *Xba*I, *Hind*III, *Sma*I). The following abbreviations are used for restriction enzymes cleavage sites, Ba-*Bal*I, Bg-*Bgl*II, Bm-*Bam*HI, C-*Cla*I, E-*Eco*RI, Hc-*Hinc*II, H-*Hind*III, N-*Nco*I, P-*Pst*I, Pv-*Pvu*II, S-*Sma*I, Sa-*Sal*I, Ss-*Sst*I, T-*Tth*111I and X-*Xba*I. The symbol / represents the fusion of two sites after blunt ending.

should prevent ambiguities due to the presence of uncharacterized DNA.

Detection and localization of promoter P30A

Plasmid pAW81 (Figures 1D and 2) carries the first 462 bp of the left-hand end of IS30 and contains, at coordinate 63, the beginning of the long open reading frame (ORFA) predicted from the DNA sequence, together with sequences with homology to the promoter consensus sequence (Hawley and McClure, 1983). pAW81 directs the synthesis of 30 ± 2 units of galactokinase (Figure 2). The IS30 sequence is preceded by a short segment of linker DNA with no known promoter activity, thus the observed promoter of transcription must lie between the left-hand end of IS30 and the *Hind*III site (coordinate 462). Plasmid pAW82 contains an almost identical fragment to pAW81 (Figures 1B, 2 and 3A), except that the first 17 bp of IS30 from the junction with the linker to the *Hinc*II site have been removed. The DNA sequence which now occupies the site of the first 17 bp of IS30 has little sequence homology with the original IS30 sequence (Figure 3A). Plasmid pAW82 directs the synthesis of only 5.5 ± 1.6 units of galactokinase, which is significantly less than pAW81 and only slightly above the background. Thus removal of the small segment of DNA between the left-hand end of IS30 and the *Hinc*II site severely reduces the observed promoter activity. The best fit to the *E. coli* promoter consensus sequence (Hawley and McClure, 1983) in this region of the IS30 sequence is shown in Figure 3A. This promoter is called P30A and the *Hinc*II site lies between its proposed -35 and -10 regions.

Plasmids pAW81, pAW7 and pAW16 (Figure 1A and C), which have totally distinct sequences just 5′ to the proposed IS30 transposase promoter (Caspers *et al.*, 1984), have very similar promoter activities (Figure 2). This suggests that, at least in the absence of external promoters, sequences directly upstream of the promoter do not greatly affect its activity.

Plasmids pAW82 and pAW21 (Figure 1A) direct the synthesis of only very small amounts of galactokinase (Figure 2), thus no additional active promoters are present between the start of ORFA and the *Hind*III site and between the *Hind*III site and the *Bal*I site (coordinate 957), which is still within ORFA.

Detection and localization of partial promoter P30BΔ

Plasmid pAW130 (Figure 1E), which carries a complete copy of IS30 inserted in front of the *galK* gene, directs the synthesis of a small but possibly significant amount of galactokinase (Figure 2). This would suggest that transcription from P30A may, at a low frequency, read across the end of the element, or that a weak promoter is present in the right-hand end of the element. That the transcription is unlikely to be due to the short restriction site linker sequence is shown by pAW131 (Figure 1E) which has the same linker between the end of IS30 and the pFD51 sequence and directs the synthesis of insignificant amounts of galactokinase (Figure 2). However, plasmid pAW15 (Figure 1C), which carries the last 104 bp of the right-hand end of IS30 and only one additional base pair before the pFD51 sequence (Figure 3B)

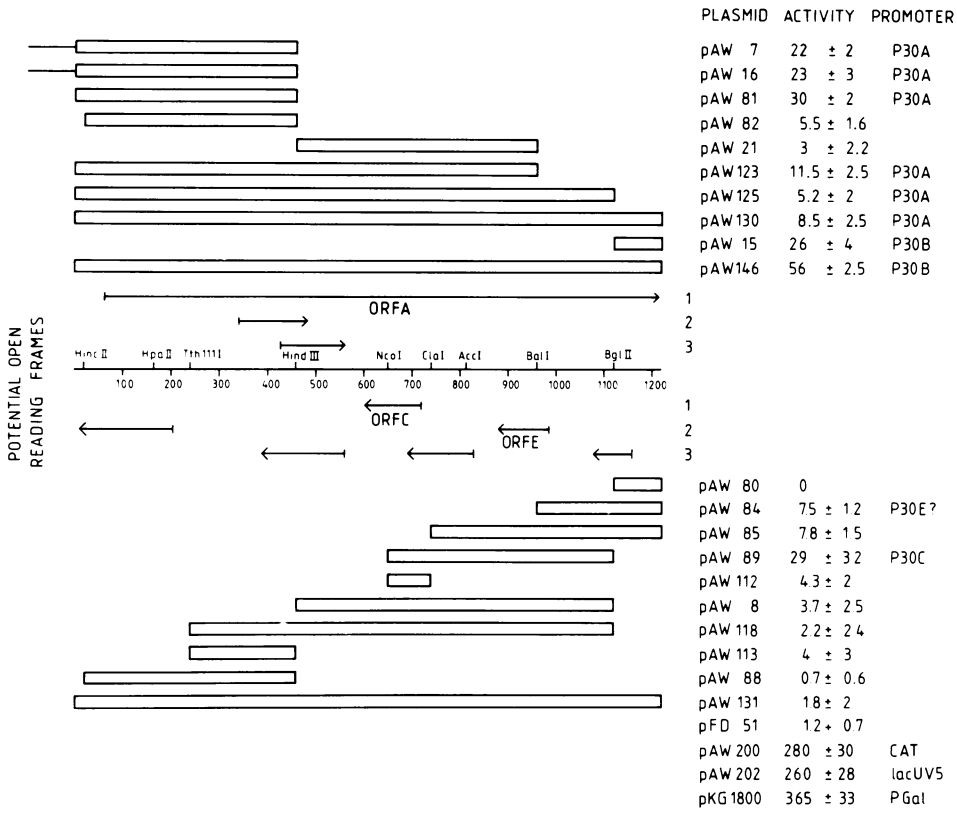

Fig. 2. The segments of IS30 and adjacent DNA cloned into pFD51 to measure promoter activity. In the center of the figure is the map of IS30 with the restriction sites utilized in the constructions and the potential open reading frames (ORF) of >25 amino acids predicted from the sequence (Dalrymple *et al.*, 1984). Frames above the line read 5′ to 3′ left to right, those below the line read 5′ to 3′ right to left; frames of the same number above and below the line are in the same relative frame. In the top half of the figure initiation of transcription from left to right is measured. In the bottom half of the figure initiation of transcription is measured from right to left. For details of sequences separating the end of the IS30 sequences from the *galK* gene see Figure 1. Galactokinase units are expressed as nmol of galactose phosphorylated/min/ml of cells at OD_{650} = 1.0. All values are the mean of four or more independent measurements and are given with their standard deviations. The amount of galactokinase directed by pAW80 was never above the background.

directs the synthesis of ~26 units of galactokinase (Figure 2). When the rest of IS30 is cloned into pAW15 to create plasmid pAW146 (Figure 1C), which then contains an intact copy of IS30, even more galactokinase synthesis is directed than by pAW15 (Figure 2). The only major difference between pAW146 and pAW130 is that the sequence between the end of IS30 and the pFD51 sequence is different (Figure 3B). Thus either the sequence in pAW130 between the end of IS30 and the pFD51 DNA has a strong terminating activity or the promoter observed on pAW15 and its derivative pAW146 was created by the juxtaposition of a suitable −35 sequence on IS30 and a suitable −10 sequence on the adjacent DNA from plasmid pFD51. We have evidence that the linker sequence between the ends of IS30 and the *Bam*HI sites in pAW130 (Figure 1E) does not have any terminator activity (to be published). Thus it is most likely that the promoter carried on pAW15 was created by the cloning process. Such a possible promoter of RNA transcription is indicated in Figure 3B. This fusion promoter is called P30B, and its −35 segment residing on IS30 is called P30BΔ.

Screen for promoters of transcription on the opposite strand of IS30

Unlike many IS elements such as IS2, IS4, IS5, IS102, IS903 and ISH1 (see Iida *et al.*, 1983; Rak and von Reutern, 1984), IS30 does not have an open reading frame encoding a protein of >100 amino acids in-frame with the putative transposase gene, but on the opposite strand (Dalrymple *et al.*, 1984). However, a number of shorter potential open reading frames on this strand of IS30 are predicted by the sequence, one of which is in-frame with ORFA. In the lower half of Figure 2 are shown the constructions which test whether any of these potential open reading frames is preceded by a promoter of RNA transcription (see plasmids pAW80, pAW84, pAW85, pAW89, pAW112, pAW8, pAW118, pAW113 and pAW88). No promoters detectable in this system were found preceding any of the potential open reading frames, except for ORFC and possibly ORFE (see below). In addition, no promoter directing transcription out of the left-hand end of IS30 was identified (see pAW131).

Detection and localization of P30C and T30C

Plasmid pAW89 (Figure 1A), which carries the *Bgl*II (coordinate 1116) to *Nco*I (coordinate 652) fragment of IS30 cloned into pFD51, directs the synthesis of 29 ± 3.2 units of galactokinase, while pAW85 (Figure 1C), which carries the complete right-hand end of IS30 to the *Cla*I site (coordinate 737), directs the synthesis of less than a third of the amount of galactokinase (Figure 2). Therefore a promoter defined as P30C lies between the *Cla*I and the *Nco*I sites, or across the *Cla*I site (Figure 3C). Removal of the DNA preceding the *Cla*I site in pAW89, to generate pAW112 (Figure 1A), also removes the majority of the promoter activity, thus positioning P30C across the *Cla*I site. This construction also rules out the possibility of P30C being generated in the ligation of the filled-in *Nco*I site to the *Sma*I site of plasmid pFD51.

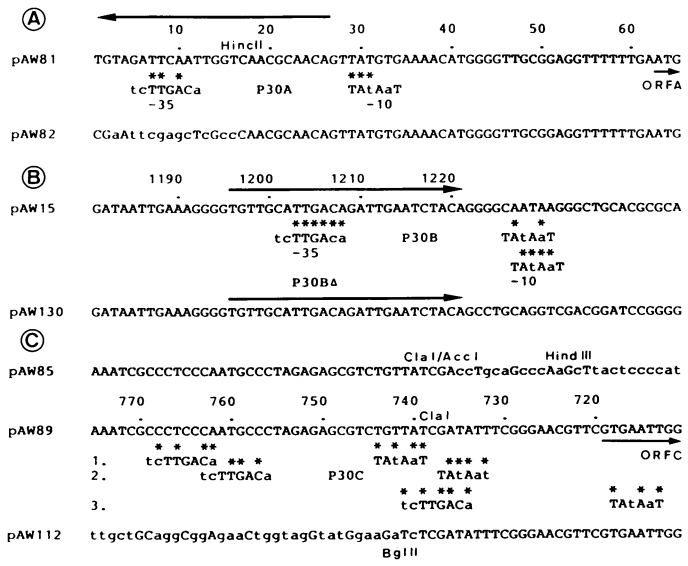

Fig. 3. The sequences of the constructions used to identify and characterize the IS30 promoters P30A, P30B and P30C. The large arrows indicate the terminal inverted repeats and the small arrows the potential start codons of open reading frames ORFA and ORFC together with the direction of translation. Matches with the *E. coli* promoter consensus sequence are identified by *. In the consensus sequence capital letters represent highly conserved bases and small letters represent less conserved bases (Hawley and McClure, 1983). Coordinate numbers refer to the IS30 sequence (Dalrymple *et al.*, 1984). **(A)** The sequence of the left-hand end of IS30 in plasmids pAW81 and pAW82 and the sequence of the linker from pUC12 (Vieira and Messing, 1982; Messing and Vieira, 1982) inserted at the *Hinc*II site of IS30 in pAW82. The IS30 sequence and homologous bases from the substituted sequence are in capital letters, while non-homologous bases are in small letters. **(B)** The sequence of the right-hand end of IS30 in plasmids pAW15 and pAW130, the sequence of the adjacent pFD51 sequence in pAW15 (McKenney *et al.*, 1981) and the sequence of the adjacent linker sequence from pUC7 and pAW130 (Vieira and Messing, 1982). The possible promoter sequence, P30B, which in pAW15 overlaps the junction between IS30 and pFD51 sequence is shown with two potential −10 regions. P30BΔ is the part of this sequence carried on IS30. **(C)** The sequences of plasmids pAW85, pAW89 and pAW112 which define the position of promoter P30C. The IS30 sequence and homologous bases from the substituted sequence are in capital letters, while non-homologous bases are in small letters. The non-IS30 sequence in pAW85 is from pUC12 (from the *ClaI*/*AccI* to *Hind*III sites) and from pFD51 after the *Hind*III site. The non-IS30 sequence in pAW112 is from pFD51 (Seed, 1983; Rak and von Reutern, 1984). Three possible matches to the *E. coli* promoter consensus sequence are shown. Only one sequence, No.2, is probably inactive in both pAW85 and pAW112 and is thus the most likely candidate for P30C.

Three possible good fits to the *E. coli* promoter consensus sequence are present in the vicinity of the *Cla*I site preceding the reading frame ORFC (Figure 3C). Only one of these sequences, No. 2, would definitely be inactive as a promoter in both plasmids pAW85 and pAW112. However, promoter sequence No. 2 may not be sufficiently far in front of the GTG start codon of ORFC to allow efficient translation of this 38 amino acids long open reading frame from the message (Dalrymple *et al.*, 1984).

Plasmid pAW8 (Figure 1A) only directs the synthesis of a small amount of galactokinase suggesting that the open reading frame ORFC is followed by a terminator of transcription, which we call T30C. This role could be taken by a small, fairly stable, potential stem and loop structure located in this region (Figure 4). Thus the synthesis of a short RNA molecule of <270 nucleotides is probably initiated at P30C. At present there is no evidence that this RNA is translated.

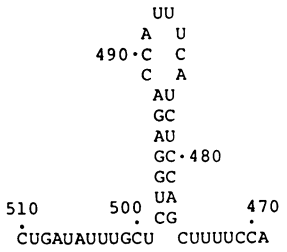

Fig. 4. The possible RNA secondary structure of the inverted repeat which follows the open reading frame ORFC and which is a candidate for T30C. This structure has a predicted stability of −8.4 Kcal (Tinoco *et al.*, 1973). Coordination numbers refer to the IS30 sequence (Dalrymple *et al.*, 1984).

Is ORFE expressed?

The insertion of any fragment of IS30 into pFD51, with the exception of the fragments in pAW80 and pAW88, leads to the synthesis of more galactokinase than pFD51. We have assumed that less than four units of galactokinase is probably not significant. The activity of pAW85 could be explained by a partially active promoter P30C, however, the activity directed by pAW84 is significantly higher than that directed by pAW80, which was never above the background of the assay (Figure 2). Taking pAW80 as the background for pAW84 would suggest that a weak promoter of RNA transcription may be present in the region of IS30 between the *Bgl*II and *Bal*I sites, or across the *Bgl*II site. A number of sequences with weak homology to the *E. coli* promoter consensus sequence (Hawley and McClure, 1983) are present in this area. However, further experiments are required to determine with certainty whether a real IS30-borne promoter is present in this region of IS30.

The relative strengths of the IS30 promoters

The strength of the IS30 promoters was compared with the strength of a number of other, well-studied promoters cloned into the same system (Figure 2). P30A and P30C, which have almost identical strengths, are ~10% the strength of the *lacUV5*, CAT and P_{Gal} promoters. These are all intermediate strength promoters, hence the IS30 promoters can be considered as weak promoters.

Does the activity of P30A reflect the number of full-length RNA molecules synthesized?

The activity of promoter P30A was measured at the *Bal*I (pAW123) and *Bgl*II (pAW125) sites of IS30 and compared with the activities at the *Hind*III site (pAW81) and at the end of IS30 (pAW130) (Figure 2). Plasmid pAW123 directs about a third of the pAW81-directed galactokinase and pAW125 directs about a sixth. This suggests that expression of the *galK* gene in pAW130 is due to read-through from promoter P30A rather than a promoter in the end of the element. ORFC and promoter P30C are present, and presumably expressed, on both pAW123 and pAW125, thus it is possible that the activity of P30A is controlled by the potential protein product of ORFC or the RNA. However, it is also possible that termination of transcription in this region of IS30 may be responsible for the observed reduction in transcription. To test whether the potential products could be supplied from an adjacent copy of IS30, the IS30 cartridge from pAW304 was cloned into plasmids pAW81 and pAW82. This generated plasmids pAW96 and pAW97, respectively, which contain IS30 in the same relative orientation as the already present segment of IS30. Plasmid pAW97 allows us to correct the activity of pAW96 for the low, but significant, expression

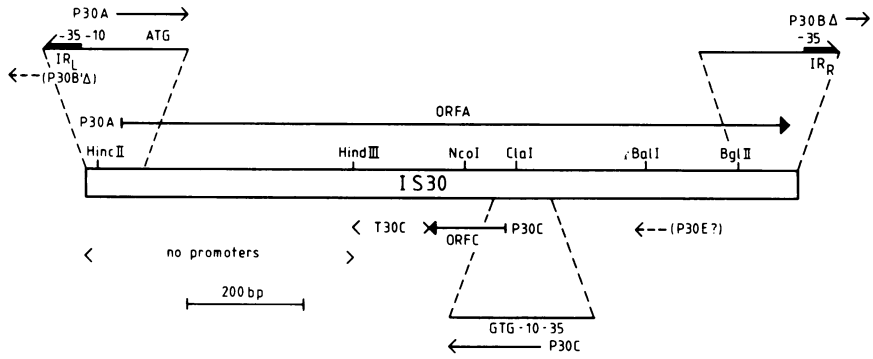

Fig. 5. The organization of IS30, showing positions of identified promoters of RNA transcription, terminator T30C and relevant open reading frames (ORF). Symbols in parentheses refer to possible additional promoters. IR_L and IR_R are the terminal inverted repeats of IS30.

out of the right-hand end of IS30. After this correction 26 ± 8 units of galactokinase activity directed by pAW96 are attributed to P30A. This is essentially the same activity as obtained with the parental plasmid, pAW81. This suggests that IS30 does not encode a product which in stoichiometric amounts can significantly affect the activity of the promoter P30A of another element, even though it lies directly adjacent on the same replicon as the measuring system.

Discussion

From the data presented above one can draw the following picture of the functional organization of IS30, which shows, in addition to a number of differences, striking similarities to the organization of the other IS elements (Figure 5).

The long open reading frame ORFA of the putative transposase gene is preceded by a promoter, P30A, which directs the synthesis of ~25 units of galactokinase in the assay system used. The proposed −35 region of the promoter is part of the left-hand, 26-bp long, terminal inverted repeat and is thus conserved between the ends. The proposed −10 region is in the unique sequence, and no similar sequence is present in the right-hand end of IS30. Accordingly no transcription into IS30 was observed from the right-hand end (see plasmid pAW80, Figure 2). Thus, in common with IS1 (Chan and Lebowitz, 1982; Machida *et al.*, 1984), IS2 (Hinton and Musso, 1983), IS5 (Rak *et al.*, 1982), IS10 (Simons *et al.*, 1983), IS26 (B.Mollet, personal communication) and IS50 (Johnson and Reznikoff, 1981), IS30 has a detectable promoter of RNA transcription in front of its long open reading frame. In several cases, on IS1, on IS2, proposed for IS3 (Timmerman and Tu, 1985), on IS26 and now on IS30, the promoters are partially contained within one of the terminal inverted repeats of the element. The termini of transposable elements are specifically recognized in transpositional recombination processes. Thus the binding site of the RNA polymerase and one of the presumed binding sites of the transposase must have considerable overlap. This might suggest that, in these elements at least, the transposase may be involved in the control of its own expression. According to this proposal, binding of the transposase to the ends of the element could potentially suppress the promoter of the transposase gene, in a manner analogous to the control of the resolvase gene expression in Tn3 and related transposons (see Heffron, 1983). However, so far there is no experimental evidence that the product of the long open reading frame of an IS element does control its own expression, nor do experiments presented in this paper allow any firm conclusions to be drawn one way or the other.

The mechanisms described for the control of expression of the long open reading frame of IS10 (Simons and Kleckner, 1983; Simons *et al.*, 1983; Kleckner *et al.*, 1984) are probably not utilized by IS30. No promoter equivalent to pOUT of IS10 has been identified on IS30 (see plasmid pAW88, Figure 2), although it is possible that transcripts from such a promoter could be terminated before the *Hinc*II site. Additionally, IS30 does not have any sites for *dam* methylation in the vicinity of either promoter P30A or P30C; thus, control of the activity of the promoter by methylation is very unlikely. However, the apparent activity of P30A is significantly reduced when measured at the *Bal*I or *Bgl*II sites. The data presented in the last section of the Results suggests that either premature termination of a proportion of the mRNA molecules within ORFA, or inhibition of the activity of P30A by a very poorly diffusable product of IS30 is involved. We have evidence that the reduction in observed activity is due to a terminator located within ORFA (to be published). Thus the activity of P30A measured at the *Hind*III site does not accurately reflect the number of full-length mRNA molecules synthesized.

IS30 has been shown to be inserted into a number of different sites in the *E. coli* chromosome (Arber *et al.*, 1978; Caspers *et al.*, 1984). Promoter P30A is so close to the end of IS30 that in different insertion sites different sequences will be directly upstream of the −35 region. It is possible that sequences in the region −40 to −80 of the promoter could affect the activity of promoter P30A, similar to what has been demonstrated for tRNA promoters (Travers *et al.*, 1983; Bossi and Smith, 1984) and suggested for other promoters (Siebenlist *et al.*, 1980; Bujard *et al.*, 1982). Thus, despite the similarity of the activities of P30A in plasmids pAW7, pAW16 and pAW81 (Figure 2), which have totally distinct sequences external to IS30, it cannot be ruled out that under certain circumstances sequences upstream of IS30 could affect the activity of P30A, thereby altering the transposition frequency of the element at different sites of insertion.

Some transcription is observed out of the right-hand end of IS30, which is probably due to read-through from the promoter P30A. On the other hand, no significant RNA transcription was detected out of the left-hand end. Thus by this mechanism IS30 can initiate transcription into adjacent genes in only one orientation and only at a low rate. In addition, under certain conditions, IS30 is able to generate a novel promoter, P30B, by insertion. This occurs by the suitable juxtaposition of a −35 region (P30BΔ) in the right-hand end of IS30 with a −10 region in the region of insertion, as is demonstrated with pAW15. The proposed −35 region of this promoter is also conserved between the ends of IS30, with one change (Caspers *et al.*, 1984), so it is possible that the left-hand end may also have the potential to generate a promoter (P3OB'Δ in Figure 5). However, so far no natural IS30

insertions have been isolated in which a promoter would have been identified in this orientation as due to the insertion process. This mechanism for activating adjacent genes is very similar to that proposed for IS2 (Hinton and Musso, 1982; Jaurin and Normark, 1983) and IS140 (Bräu *et al.*, 1984).

Another promoter, P30C, is found near the middle of IS30. It precedes an open reading frame (ORFC), which is in-frame with, but in the opposite strand to, the long open reading frame ORFA. ORFC has the potential to encode a short protein of 38 amino acids. This is considerably shorter than the equivalent reading frames of other IS elements (see Iida *et al.*, 1983; Rak and von Reutern, 1984). The most likely sequence for the P30C promoter (Figure 3) would direct the synthesis of a mRNA with only few nucleotides in front of the start codon (coordinate 721), thus the potential protein may be only poorly expressed. It is also possible that P30C initiates the synthesis of an RNA transcript of direct functional relevance for transposition. A potential sequence for the terminator of transcription, T30C, has been located ~250 bp downstream of promoter P30C. For IS5 a reading frame equivalent to ORFC has been shown to be preceded by a promoter and its protein product has been demonstrated (Rak *et al.*, 1982). However, at present, the function of these analogous potential genes is not clear.

There are no obvious sequence homologies between the DNA sequences surrounding promoters P30A and P30C suggesting that they would not be regulated by a common mechanism. Additionally there is no extensive complementarity between the RNA synthesized from P30C and the 5′ end of the mRNA synthesized from P30A, suggesting that the RNA product of P30C is not involved in inhibiting the initiation of translation of the presumed transposase. Alternatively, hybridization of the two RNA molecules in the region of complete complementarity could perhaps reduce translation efficiency. This function has been proposed for the anti-RNA synthesized from IS1, although control of translation has not been shown (Machida *et al.*, 1984).

A third, but very weak promoter, may be present on IS30 preceding ORFE. Thus IS30 may also have a third reading frame capable of expressing a protein, analogous to open reading frame *ins*L of IS5 (Rak and von Reutern, 1984). None of the other potential open reading frames of IS30 is preceded by a promoter with *in vivo* activities significantly above the background.

In conclusion, IS30 encodes at least two RNA molecules which could be expressed in cells in which IS30 is resident. IS30 also has the potential to activate expression of adjacent genes by an element born sequence, which can generate an active promoter by fusion to adjacent sequences. Thus IS30, despite being unrelated at the level of its DNA sequence to the other IS elements, shares characteristic functional features in common with these elements.

Materials and methods

Strains and plasmids

The following *E. coli* K12 strains were used: N100, *pro, his, galK, recA* (McKenney *et al.*, 1981) and JM83, *ara, Δlac-pro, strA, thi*, lysogenic for φ80d*lacZ* M15 (Messing, 1979), Plasmid pFD51 (Rak and von Reutern, 1984) carrying a promoterless galactokinase gene *galK* was obtained from B.Rak. Plasmids pUC7, pUC8 and pUC12 (Vieira and Messing, 1982) were obtained from J.Messing. Plasmids pAW302 and pAW304 were constructed in our laboratory by R.Stalder (to be published). pKG1800 was obtained from D.Schümperli (McKenney *et al.*, 1981).

Despite the three different origins of the examples of IS30 used in this work, no sequence differences have so far been detected between them (Caspers *et al.*, 1984; Dalrymple *et al.*, 1984).

Construction of plasmids

From pAW522. (Figure 1A; see Caspers *et al.*, 1984). pAW21 was constructed by cloning the internal *Hind*III-*Bal*I fragment of IS30 from pAW522 into *Hind*III-*Sma*I-cleaved pFD51. pAW7 and pAW8 were constructed by cloning *Bgl*II-*Hind*III fragments of pAW522 into *Bgl*II-*Hind*II-cleaved pFD51; pAW89 was derived from pAW8 by *Nco*I and *Sma*I cleavage, blunt ending and religation; and pAW112 was derived from pAW89 by *Cla*I and *Bgl*II cleavage, blunt ending and religation. N100 is *dam*$^+$ and the *Cla*I site in the *galK* gene is not cleaved in DNA prepared from this strain due to methylation of an A residue (Debouck *et al.*, 1985). pAW118 was constructed by cloning the small *Hind*III fragment from pAW110 (see below) into *Hind*III-cleaved pAW8.

From pAW24. (Figure 1B). pAW24 was constructed from pAW209 (Caspers *et al.*, 1984) by partial *Bgl*II digestion and religation, pAW60 was constructed by cloning the ~1600-bp *Hinc*II fragment from pAW24 into *Hinc*II-*Sma*I-cleaved pUC12. pAW88 was constructed from pAW60 by *Hind*III digestion and ligation of the short *Hind*II fragment into *Hind*III-cleaved pFD51. pAW110 was constructed from pAW60 by *Tth*111I and *Pst*I digestion, blunt ending and religation. pAW113 was constructed from pAW110 by *Hind*III digestion and ligation of the short *Hind*III fragment into *Hind*III-cleaved pFD51. pAW63 was constructed by cloning the ~160-bp *Hinc*II fragment from pAW24, in the reverse orientation to pAW60, into *Sma*I-cleaved pUC12. pAW82 was constructed from pAW63 by insertion of the small *Eco*RI-*Hind*III fragment into *Eco*RI-*Hind*III-cleaved pFD51.

From pAW83. (Figure 1C; see Caspers *et al.*, 1984). pAW16 was constructed from pAW83 by cloning the *Hind*III fragment containing the left-hand end of IS30 into the *Hind*III site of pFD51. pAW15 was constructed by cloning the *Bgl*II-*Pvu*II fragment containing the right-hand end of IS30 into *Bgl*II-*Sma*I-cleaved pFD51. pAW146 was constructed by cloning the *Eco*RI-*Bgl*II fragment containing the majority of IS30 from pAW304 (see below) into *Eco*RI-*Bgl*II-cleaved pAW15. pAW67 was constructed by cloning the *Pvu*II-*Cla*I fragment of pAW83 into *Sma*I-*Acc*I-cleaved pUC12. pAW80, pAW84 and pAW85 were constructed, respectively, by cloning the *Eco*RI-*Bgl*II, the *Eco*RI-*Bal*I and the *Eco*RI-*Hind*III fragments of pAW67 into *Eco*RI-*Sma*I and *Eco*RI-*Hind*III-cleaved pFD51.

From pAW42. (Figure 1D). pAW42 was constructed by cloning the *Eco*RI-*Bgl*II fragment containing the majority of IS30 from phage P1::IS30 D7 (Caspers *et al.*, 1984) into *Eco*RI-*Bam*HI-cleaved pBR322. pAW66 was constructed by inserting the *Pvu*I-*Hind*III fragment from pAW42 into *Sma*I-*Hind*III-cleaved pUC12 (the *Pvu*II site is created by the insertion of IS30 into phage P1 DNA); and pAW81 was constructed by cloning the *Eco*RI-*Hind*III fragment from pAW66 into *Eco*RI-*Hind*III-cleaved pFD51.

From pAW34. (Figure 1E). pAW130 and pAW131 were generated by cloning the *Eco*RI IS30 cartridge from pAW304 into *Eco*RI-cleaved pFD51 in both possible orientations. pAW123 and pAW125 were constructed from pAW130 by *Bal*I and *Sma*I and by *Bgl*II digestion, respectively, followed by religation. pAW96 and pAW97 were constructed by cloning the *Eco*RI IS30 cartridge from pAW304 into *Eco*RI-cleaved pAW81 (Figure 1D) and pAW82 (Figure 1B), respectively. In both plasmids IS30 is in the same orientation relative to the already present IS30 sequences.

Promoter plasmids. pAW200 was constructed from pAW302, which carries a chloramphenicol acetyl transferase gene cartridge (to be published), by insertion of the short *Eco*RI fragment containing the promoter and the N-terminal end of the gene into the *Eco*RI site of pFD51. pAW202 was generated by cloning the short *Pvu*II fragment, which contains the *lacUV5* promoter, from pUC8 into *Sma*I-cleaved pFD51.

Enzyme assays

Galactokinase activity was measured as described by McKenney *et al.* (1981). D-[1-^{14}C]galactose was purchased from Amersham International.

Acknowledgements

We wish to thank Rolf Stalder for the gift of plasmids pAW302 and pAW304, Bodo Rak for the gift of pFD51, Patrick Caspers for the gift of pAW522, pAW83 and phage P1::IS30 D7 DNA and Daniel Schümperli for the gift of pKG1800. We are very grateful to Thomas Bickle and Shigeru Iida for advice and fruitful discussions. This project was supported by grant 3.392.83 from the Swiss National Science Foundation.

References

Arber,W., Iida,S., Jütte,H., Caspers,P., Meyer,J. and Hänni,C. (1978) *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 1197-1208.

Bossi,L. and Smith,D.M. (1984) *Cell*, **39**, 643-652.

Bräu,B., Pilz,U. and Piepersberg,W. (1984) *Mol. Gen. Genet.*, **193**, 179-187.

Bujard,H, Niemann,A., Breunig,K., Roisch,U., Dressel,A., von Grubain,A., Grentz,R., Stüber,D. and Weiher,H. (1982) in Rodriguez,R.L. and Chamberlin,M.J. (eds.), *Promoters, Structure and Function*, Praeger, NY, pp. 121-140.
Caspers,P., Dalrymple,B., Iida,S. and Arber,W. (1984) *Mol. Gen. Genet.*, **196**, 68-73.
Chan,P.T. and Lebowitz,J. (1982) *Nucleic Acids Res.*, **10**, 7295-7311.
Charlier,D., Pietle,J. and Glansdorff,N. (1982) *Nucleic Acids Res.*, **10**, 5935-5948.
Dalrymple,B., Caspers,P. and Arber,W. (1984) *EMBO J.*, **3**, 2145-2149.
Debouck,C., Riccio,A., Schümperli,D., McKenney,K., Jeffers,J., Hughes,C., Rosenberg,M., Heuterspreute,M., Brunel,F. and Davison,J. (1985) *Nucleic Acids Res.*, **13**, 1841-1853.
Glansdorff,N., Charlier,D. and Zafarullah,M. (1980) *Cold Spring Harbor Symp. Quant. Biol.*, **45**, 153-156.
Hawley,D.K. and McClure,W.R. (1983) *Nucleic Acids Res.*, **11**, 2237-2255.
Heffron,F. (1983) in Shapiro,J.A. (ed.), *Mobile Genetic Elements*, Academic Press Inc., pp. 223-260.
Hinton,D.M. and Musso,R.E. (1982) *Nucleic Acids Res.*, **10**, 5015-5031.
Hinton,D.M. and Musso,R.E. (1983) *J. Mol. Biol.*, **169**, 53-81.
Iida,S., Meyer,J. and Arber,W. (1983) in Shapiro,J.A. (ed.), *Mobile Genetic Elements*, Academic Press Inc., pp. 159-221.
Jaurin,B. and Normark,S. (1983) *Cell*, **32**, 809-816.
Johnson,R.C. and Reznikoff,W.S. (1981) *Nucleic Acids Res.*, **9**, 1873-1883.
Kleckner,N., Morisato,D., Roberts,D. and Bender,J. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 235-244.
Machida,C., Machida,Y. and Ohtsubo,E. (1984) *J. Mol. Biol.*, **177**, 247-267.
McKenney,K., Shimatake,H., Court,D., Schmeissner,U., Brady,C. and Rosenberg,M. (1981) in Chirikjian,J.C. and Papas,T.S. (eds.), *Gene Amplification and Analysis. Vol. 2. Analysis of Nucleic Acids by Enzymatic Methods*, Elsevier, NY, pp. 383-415.
Messing,J. (1979) *Recombinant DNA Technical Bulletin*, NIH Publication No. 79-99, **2**, 43-48.
Messing,J. and Vieira,J. (1982) *Gene*, **19**, 269-276.
Rak,B., Lusky,M. and Hable,M. (1982) *Nature*, **297**, 124-128.
Rak,B. and von Reutern,M. (1984) *EMBO J.*, **3**, 807-811.
Rosenberg,M., Chepelinsky,A.B. and McKenney,K. (1983) *Science (Wash.)*, **222**, 734-739.
Seed,B. (1983) *Nucleic Acids Res.*, **11**, 2427-2447.
Siebenlist,U., Simpson,R.B. and Gilbert,W. (1980) *Cell*, **20**, 269-281.
Simons,R.W. and Kleckner,N. (1983) *Cell*, **34**, 683-691.
Simons,R.W, Hoopes,B.C., McClure,W.R. and Kleckner,N. (1983) *Cell*, **34**, 673-682.
Timmerman,K.P. and Tu,C.-P.D. (1985) *Nucleic Acids Res.*, **13**, 2127-2139.
Tinoco,I., Borer,P.N., Dengler,B., Levine,M.D., Uhlenbeck,O.C., Crothers, D.M. and Gralla,J. (1973) *Nature, New Biol.*, **246**, 40-41.
Travers,A.A., Lamond,A.I., Mace,H.A.F. and Berman,M.L. (1983) *Cell*, **35**, 265-273.
Vieira,J. and Messing,J. (1982) *Gene*, **19**, 259-268.

Received on 28 May 1984; revised on 8 July 1985