Mutations in the carboxy-terminal part of IS*30* transposase affect the formation and dissolution of $(IS30)_2$ dimer

Ferenc Olasz[a,b,*], Tibor Farkas[b], Rolf Stalder[a], Werner Arber[a]

[a] *Abteilung Mikrobiologie, Biozentrum der Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland*
[b] *Institute for Molecular Genetics, Agricultural Biotechnology Center, Szent-Györgyi Albert utca 4, H-2101 Gödöllő, Hungary*

Received 6 June 1997; revised version received 14 July 1997

Abstract The transposase of IS*30* catalyses different transpositional rearrangements via the dimer $(IS30)_2$ intermediate structure. Mutation analysis provides evidence that the C-terminal part of IS*30* transposase is required for the formation and dissolution of $(IS30)_2$ dimer. C-terminal mutants are also defective in transpositional fusion; however, this deficiency can be 'suppressed' by addition of the final product of site-specific dimerisation, the core $(IS30)_2$ intermediate structure. The transposase part studied shows significant homologies in three highly conserved regions to proteins of IS*30*-related mobile elements.

Key words: IS*30*; Domain structure of transposase; Intermediate in transposition; Site-specific deletion formation; Insertion sequence; *Escherichia coli*

1. Introduction

The mobile element IS*30* (Fig. 1A) is resident in the genome of *Escherichia coli* K-12 and C strains [1]. The 1221-bp-long insertion sequence contains one large open reading frame (ORF-A, 63–1212 bp), which is transcribed from the P30A promoter [2] and codes for the 44.3-kDa transposase [3]. The 17-kDa N-terminal part of the protein, the product of a truncated gene (*ORF-A-N*), was found to bind specifically to the terminal inverted repeats (IR) of the element [4]. IS*30* has a pronounced, dual target specificity; not only well-characterised 'natural' hot spots but its own IRs are recognised as targets (Olasz et al., manuscript submitted).

The transposition of IS*30* is a conservative process that proceeds through an $(IS30)_2$ intermediate [5]. To avoid any misunderstanding, a new terminology was used for transpositional reactions mediated by IS*30* (Fig. 1B, [5]). The introduction of the term 'transpositional fusion' instead of co-integration seemed to be evident, since IS*30* is not duplicated in this reaction. Accordingly, to distinguish the two site-specific deletion processes that result in the formation and segregation of the $(IS30)_2$ dimer, these reactions were designated by the new terms 'site-specific dimerisation' and 'dimer dissolution', respectively.

The intermediate structure, which plays a central role in the transposition of IS*30*, is formed by site-specific dimerisation (Fig. 1B) so that the left and right IRs of two elements join to each other [6]. This structure is frequently produced from composite transposons and, in general, from DNA segments that carry two IS*30* elements [5]. The dimer structure is highly unstable and generates a burst of transpositional rearrangements. $(IS)_2$-like transpositional intermediates have also been observed with IS*21* [7] and IS*2* [8].

Similar to the homologous proteins of other insertion sequences, IS*30* transposase is likely to be involved in several reactions: in recognition of the terminal inverted repeats of the element, and in cutting and joining of the termini to the transpositional target that also requires specific recognition and cleavage of the target sequences (for review see [9]; [10]). A detailed study of transposase domains enables one to obtain insight into the structural and functional organisation of the protein and into the transposition process itself.

We present here experimental evidence that another putative domain, the C-terminal region of IS*30* transposase, is responsible for the formation and dissolution of $(IS30)_2$ intermediates.

2. Materials and methods

2.1. Media, enzymes, chemicals and DNA techniques

LB medium was prepared as described by Miller [11]. Restriction endonucleases, T4 DNA ligase and DNA polymerase I Klenow fragment were obtained from Boehringer Mannheim (Germany) and Minotech (Greece). Antibiotics (Sigma, Germany) were used at final concentrations: kanamycin (Km) 20 μg/ml; ampicillin (Ap) 200 μg/ml; chloramphenicol (Cm) 20 μg/ml. [^{35}S]dATP was obtained from Amersham (UK). Sequencing was performed with the chain termination method of Sanger et al. [12] using the T7 sequencing kit of Pharmacia (Sweden). Commonly used DNA techniques were carried out according to Sambrook et al. [13].

2.2. Plasmid constructs

In the transpositional fusion experiments two groups of plasmids were distinguished: the donor plasmids were pACYC177 [14] derivatives and were able to produce IS*30* transposase, while the target plasmids contained preferred target sequences (hot spots) cloned into a pEMBL vector [15]. The structure and relevant properties of the plasmids are shown in Figs. 2 and 3.

2.2.1. Target plasmids. pAW782 and pAW1016 carried hot spots derived from bacteriophages lambda (LHS) and P1 (P1HS), respectively [5]. pAW758 and pAW1105 contained the right IR of IS*30* joined to the lambda hot spot region (LHS-RIR) and the core $(IS30)_2$ structure, respectively (Olasz et al., manuscript submitted).

2.2.2. Donor plasmids. The structures of plasmids pAW472, pAW1019, pAW1020, pAW1026, pAW1027, pAW1152 and pAW1180 are described in Section 3 and presented in Fig. 2. The construction of pTFA89 and pTFA91 (Fig. 2): a new *Eco*RI site was introduced in the right IR of the element by inserting an additional T base into IS*30* position 1216. The in vitro mutagenesis (Amersham, RPN 1523) was carried out on a template containing an intact IS*30* element (pAW1038 [5]) using the following mutant oligonucleotide: GACAGATT**GAATTC**TACAGCCTGC. The *Hin*dIII–*Eco*RI fragment of the mutant IS*30* (positions 459–1216) was inserted into the *Hin*dIII–*Eco*RI site of pBluescript II KS$^+$ (Stratagene), and the Cm^R marker gene from pAW302 [16] was cloned adjacent to the truncated RIR. In the final step, the *Nco*I fragments of pAW1026 and pAW1019, containing the right end–Cm^R gene junction, were replaced with the *Nco*I fragment of the newly constructed

*Corresponding author. Fax: (36) 28-430-416.
E-mail: olasz@abc.hu

PII S0014-5793(97)00947-2

plasmid. Since the latter segment carried the right IS*30* end until the newly introduced *Eco*RI site (position 1216) the progeny control plasmids pTFA89 and pTFA91 were able to produce intact transposase protein, but their 'inner' right IRs were truncated and inactive in transposition (data not shown).

The structures of pAW1137 [5], pAW1158, pFOL3, pFOL4 and pFOL6 are described in Section 3 and shown in Fig. 4.

2.3. Transpositional fusion experiments

The method to detect rare transposition events with the aid of mobilizable target plasmids and the helper phage R408 was described earlier [5]. In these experiments the donor plasmid was introduced into the *recA*$^-$ JM109 [17] host already harbouring a target plasmid. After incubation bacteria were infected with the helper phage R408, target plasmids and their rearrangement derivatives were packed into phage particles that were used to infect new JM109 host. Applying appropriate selection (donor plasmids: Km^RCm^R, target plasmids: Ap^R) fusion products were detected as $Cm^RKm^RAp^R$ colonies.

2.4. Detection of site-specific dimerisation and dimer dissolution

Plasmids pAW1026, pAW1027, pAW1137, pFOL3, pFOL4 and pTFA89, respectively, were introduced into the *recA*$^-$ strain JM109, and transformants were selected for the resistance markers carried by the vector (Km^R) and the transposable element (Cm^R). In each experiment at least five transformant colonies were grown in LB under selection for Km^R. From these cultures overnight subcultures were subsequently inoculated several times with the same selection. After each passage plasmid DNA was purified from an aliquot and was used for the transformation of JM109. The number of Km^RCm^S and Km^RCm^R transformants was determined by replica plating. The frequency of Cm^S segregants was determined as the ratio of Km^RCm^S colonies versus total number of Km^R colonies. In each experiment plasmid DNA was purified from at least 10 Km^RCm^S colonies and was analysed with restriction endonucleases. The ratio of confirmed segregants versus total tested Km^RCm^S colonies was used for the correction of the frequency data (Fig. 4). In the stability assay of pAW1137 and its derivatives, the expression of the transposase gene from the *tac* promoter was induced with 0.1 mM IPTG.

3. Results

3.1. Functional analysis of mutations affecting the IS30 transposase protein

To determine the role of the C-terminal part of the IS*30* transposase protein, mutations were introduced in this region and their effect was investigated in different transpositional reactions. In the first set of experiments the chloramphenicol resistance marker gene (*Cm*^{R}) was cloned into the donor plasmids, into the *Bgl*II site of IS*30*. In the resulting plasmid pAW1019 the two inverted repeats of the element are intact, but the last 31 amino acids of the transposase are substituted by 25 unrelated amino acids (Fig. 2A). Additional mutagenesis was carried out in pAW1019 to produce the isogenic derivatives pAW1020 and pAW1180 (Fig. 2A). The control plasmid pTFA91, which was also analogous structurally, contained the intact *IS30 transposase* gene, but the last 5 bp of the element were deleted and the 'inner' right IR was therefore inactive in transposition (see Section 2.2.).

In the transpositional fusion experiment (Fig. 2A) the element producing the wild-type IS*30* transposase (pTFA91) was active with all the target plasmids. In contrast, no fusion product was detectable in the reaction of the donor plasmid carrying the *Bgl*II* mutation in ORF-A (pAW1019) and the target plasmids containing hot-spot sequences from phages lambda (LHS) or P1 (P1HS), and the right IS*30* IR joined to LHS (LHS–RIR). However, this donor plasmid was active with the joined IS*30* ends that correspond to the core region of the $(IS30)_2$ intermediate structure. In this case, the frequency of rearrangement was about 4-fold lower than with the control plasmid pTFA91. Structural analysis revealed that the right and left IR of pAW1019 interacted with the core $(IS30)_2$ structure with about equal probability (LIR: 9/16 cases, RIR: 7/16 cases). All 16 independent isolates were transpositional fusion products (Fig. 3A). No simple transposition or any other type of rearrangement was observed in this experiment. According to our data, introduction of a mutation at the *Bgl*II site of IS*30* (1115 bp position) has abolished the ability of the element to undergo transposition; activity was only observed with the target consisting of the core $(IS30)_2$ structure.

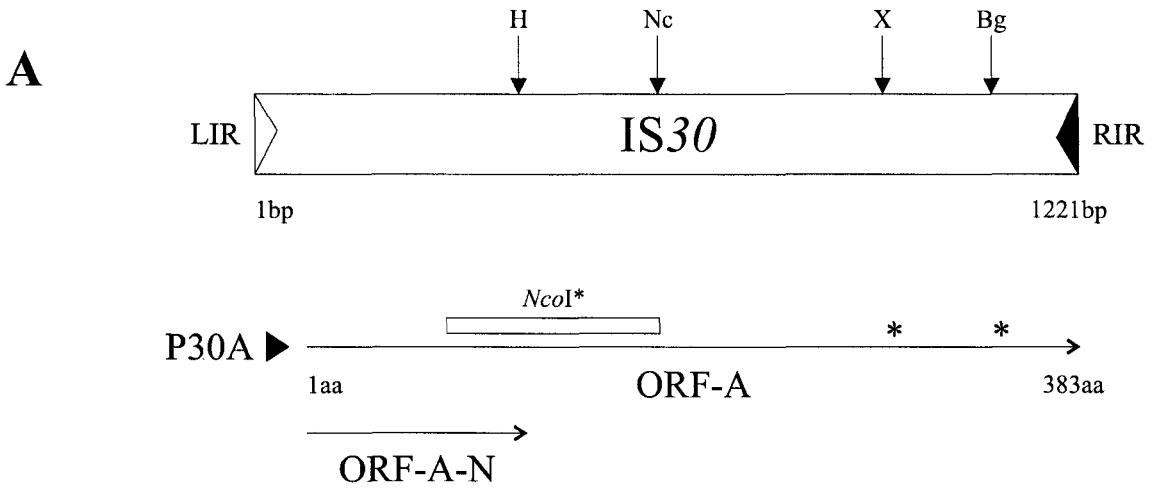
The *Xba*I* mutant, in which the filling-in of the unique *Xba*I site (965 bp position) affected the last 81 amino acids in the C-terminal part of IS*30* transposase, was functionally similar to the *Bgl*II* mutant, but the frequency of rearrangements was 3–4-fold lower in the presence of joined IRs (pAW1020, Fig. 2A). The structure of the rearranged products was analogous to those described for pAW1019.

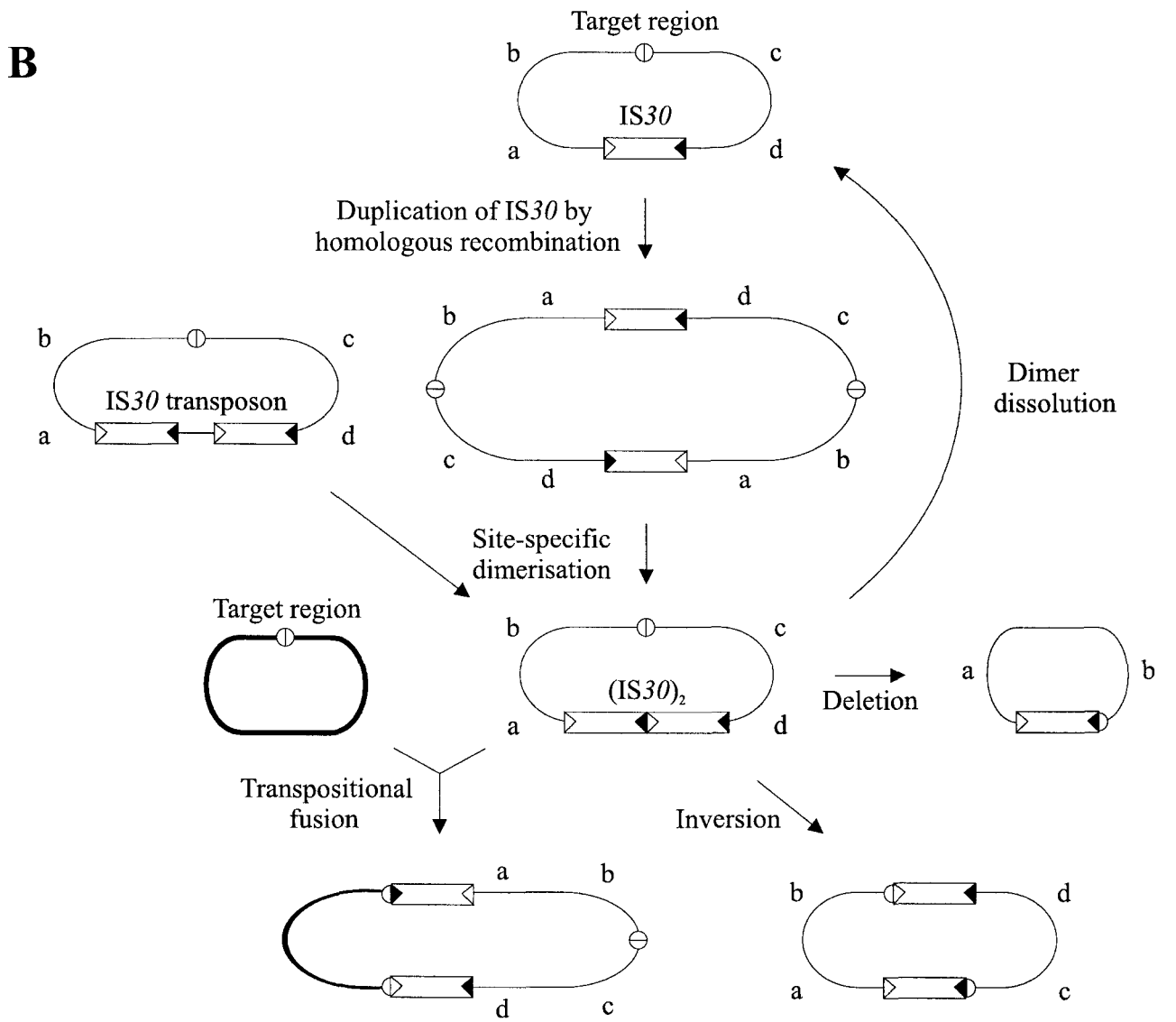
The *Nco*I* mutation contains a deletion that creates a TAA stop codon in ORF-A at position 354 bp. The construct carrying this mutation (pAW1180) did not reveal any activity with the four target sequences tested, including the joined ends of IS*30*; therefore the *Nco*I* mutant was considered as a negative control.

Since no complementation was observed with the transposase expressed from the four genomic IS*30* copies of JM109 (see donor pAW1180), the DNA rearrangements were obviously catalysed by the mutant IS*30* transposases encoded on the donor plasmids. The stability of the wild-type and mutant transposase proteins was also examined (data not shown), and no detectable differences were found; therefore it is unlikely that the reduced half-lives of the truncated proteins would be responsible for the changes in the transpositional activities of the mutants.

→

Fig. 1. The structure (A) and simplified transpositional model (B) of IS*30*. The element is shown as a large box. The 26-bp-long left (LIR) and right (RIR) inverted repeats are illustrated as blank and filled triangles, respectively. The coordinates of the first and last nucleotides are given below the element, while the recognition sites of relevant restriction endonucleases are marked by small arrows: *Bgl*II (Bg); *Hin*dIII (H); *Nco*I (Nc); *Xba*I (X). The ORF of IS*30* transposase (ORF-A, horizontal arrow) is expressed from the P30A promoter (black arrowhead), the first and last amino acids (aa) of the protein are given. Asterisks and a small box above ORF-A show the location of investigated mutations: *Nco*I*, *Xba*I*, and *Bgl*II*. The truncated transposase protein expressed from ORF-A-N binds to the IRs of IS*30*. The simplified transpositional model of IS*30* mediated DNA rearrangements (B). Duplication of IS*30* means different, non-transpositional reactions that result in the presence of two IS*30* copies on the same replicon: e.g. homologous recombination between two IS*30* carrying replicon that leads to the formation of a dimer replicon (for the detailed model and transpositional steps see [5]). Similarly, the presence of a composite transposon represents the same topology. The $(IS30)_2$ intermediate is formed by site-specific dimerisation, and induces further inter- and intramolecular reactions. The reconstruction of the original structure via the dissolution of the dimer provides the possibility for further cycles. The target region for transposition is shown as a blank circle, the site of insertion is represented by a vertical line. The different regions of the vector molecule were marked (a, b, c and d).


A
H
Nc
X
Bg
LIR
IS30
RIR
1bp
1221bp
NcoI*
P30A
1aa
ORF-A
383aa
ORF-A-N


B
Target region
IS30
Duplication of IS30 by homologous recombination
IS30 transposon
Dimer dissolution
Site-specific dimerisation
Target region
(IS30)2
Deletion
Transpositional fusion
Inversion

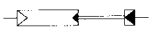
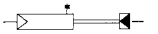
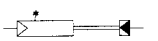
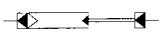
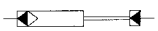
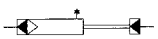
Donor plasmid	Schematic structure	Mutation in the ORF-A	Transpositional frequency with the target plasmid (10^{-9}): pAW782 lambda hot spot	pAW1016 P1 hot spot	pAW758 LHS-right IR	pAW1105 $(IS30)_2$*
A						
pTFA91		wild type, 1-383 aa	**11.0**	**5.0**	**12.0**	**195.0**
pAW1019		*Bgl*II*, 1-352 aa	**<0.7**	**<0.6**	**<1.5**	**45.0**
pAW1020		*Xba*I*, 1-302 aa	**<0.5**	**<0.9**	**<0.8**	**14.0**
pAW1180		*Nco*I*, 1-97 aa	**<0.5**	**<0.7**	**<2.0**	**<0.6**
B						
pTFA89		wild type, 1-383 aa	**589**	**343**	**414**	**2390**
pAW1026		*Bgl*II*, 1- 352 aa	**176**	**108**	**205**	**708**
pAW1027		*Bgl*II*, 1-352 aa	**96**	**78**	**126**	**n.d.**
pAW1152		*Xba*I*, 1-302 aa	**11**	**9**	**n.d.**	**17**
pAW472		*Nco*I*, 1-97 aa	**<3**	**<2**	**n.d.**	**<3**

Fig. 2. Effect of C-terminal mutations in the *IS30 transposase* gene on transpositional fusion reactions. Transpositional frequencies were measured between donor plasmids producing wild-type or mutant IS*30* transposases and plasmids containing different target regions: lambda and P1-derived hot spots; LHS–RIR or $(IS30)_2$*. The latter designates to the joined IRs, the core intermediate structure. In the wild-type donor replicons the ORF-A is intact, but the 'inner' right IR of the element was truncated (small filled triangle) and therefore inactive in transposition (see Section 2.2.). Data represent the average values of at least five independent experiments (see Section 2.3.). n.d. = not determined; aa: amino acid. Indicated amino acid positions refer to wild-type transposase sequence. Asterisks above the schematic structures show the location of additional mutations. For other symbols see Fig. 1.

3.2. Joined IS30 IRs carried in cis can suppress mutations in the C-terminal region of IS30 transposase

The phenomenon, that the *Bgl*II* mutant displayed transpositional activity when the target plasmid carried the core $(IS30)_2$ structure *in trans*, incited us to insert joined IRs into the donor plasmid (see pAW1026 and pAW1027, Fig. 2B). The new set of donor plasmids was therefore analogous to pAW1019, but they contained the core intermediate structure *in cis*.

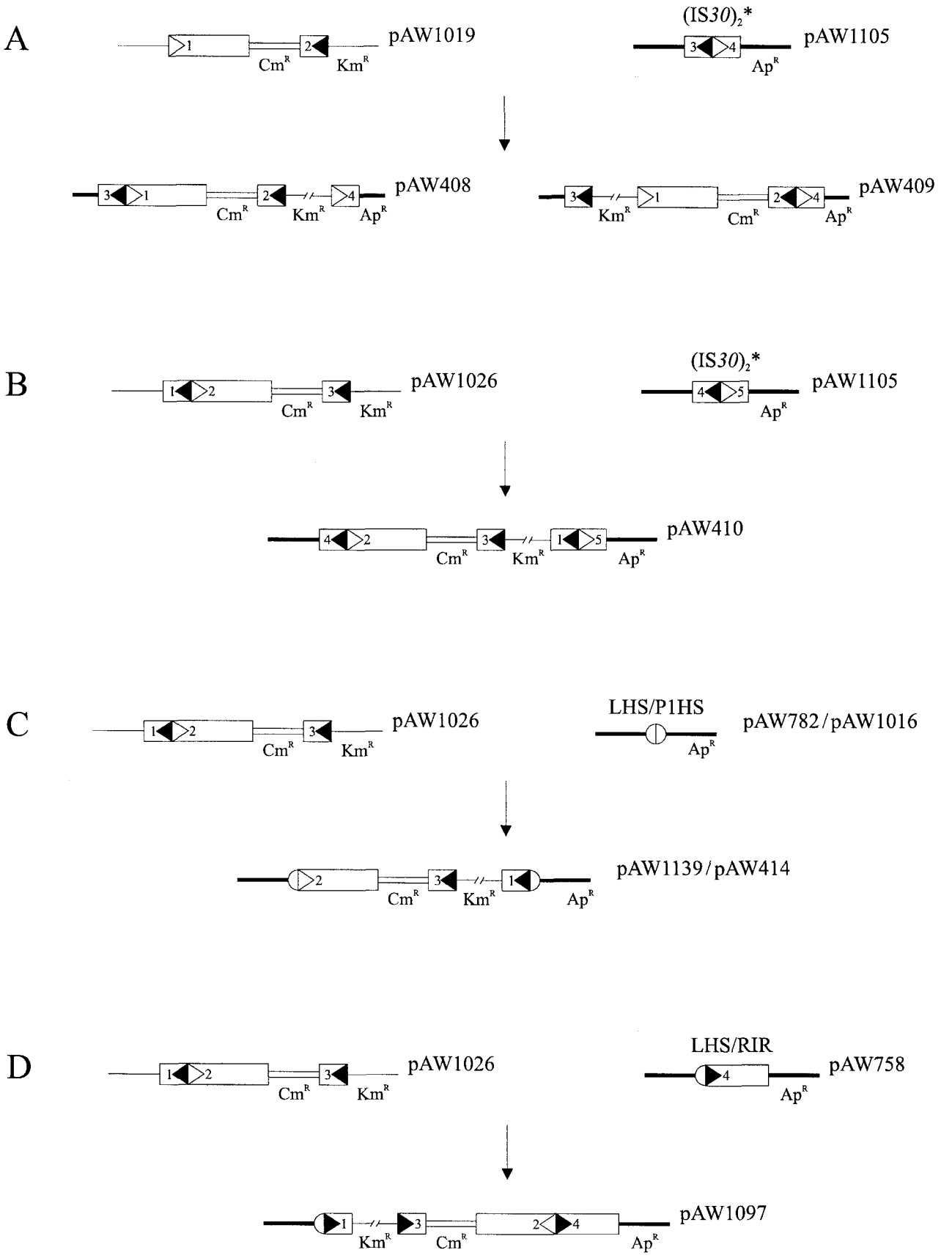
Donor plasmids containing only the *Bgl*II* mutation (pAW1026 and pAW1027) mediated transpositional fusions not only with the joined IRs of IS*30* (pAW1105) but also with the previously inactive targets. The *Bgl*II* mutant resulted in about a 2–3-fold lower transpositional activity than the wild type when they were carried in a comparable structure (pAW1026 and pTFA89, Fig. 2B. Note that pTFA89, similarly to pTFA91, contains the intact ORF-A, but the 'inner' RIR of the element is truncated.). The two *Bgl*II* mutants (pAW1026 and pAW1027) differing in the location of the $(IS30)_2$ structure showed identical pattern in their transpositional frequencies with the offered target sequences; however, pAW1026 was almost twice as active. Note that this pattern is significantly different from what we observed between pAW1019 and the same targets. These findings indicate that the presence of the joined IRs can considerably suppress the effect of mutations in the C-terminal region of IS*30* transposase regardless of its location.

Several rearrangement products of pAW1026 were analysed and their structure showed that they were formed by the interactions of the joined IS*30* ends carried by the donor plasmid and the offered target (Fig. 3B-D). The sequence analysis of both junctions in four independent pAW782–pAW1026 fusion products revealed that 2 bp of the target sequence had been duplicated while the 2 bp separating the joined IRs of IS*30* on plasmid pAW1026 had been lost upon the reaction. It demonstrates that this transpositional fusion is a conservative process as was described by Olasz et al. [5].

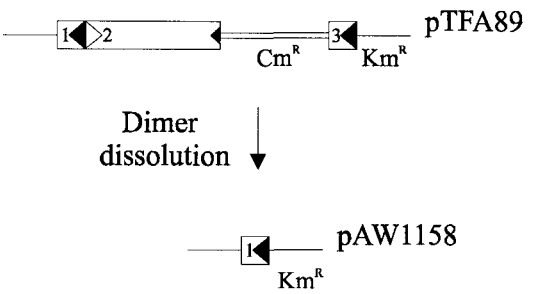
Introducing additional mutations into the ORF-A of the *Bgl*II* mutant pAW1026 we found that only the *Xba*I* mutant mediated transpositional reactions with the targets,

→

Fig. 3. Schematic illustration of experimentally identified transpositional reactions mediated by donor plamids presented in Fig. 2A,B. Characteristic interactions between the transposase producing plasmids (pAW1019 and pAW1026) and target replicons containing $(IS30)_2$*, i.e. the joined IRs, lambda hot spot (LHS), P1 hot spot (P1HS) and LHS–RIR sequence are shown in A, B, C and D, respectively. LHS and P1HS are represented by a blank circle, LHS-RIR as a half circle joining to the right IR. The inverted repeats of the element are numbered. Cm^R: chloramphenicol resistance marker (chloramphenicol acetyltransferase); Km^R: kanamycin resistance marker (aminoglycoside-3′-phosphotransferase); Ap^R: ampicillin resistance marker (β-lactamase). Thin line: pACYC177 derivatives; thick line: pEMBL18 or pEMBL19 derivatives; double line: Cm^R gene. For other symbols and abbreviations see Figs. 1 and 2.

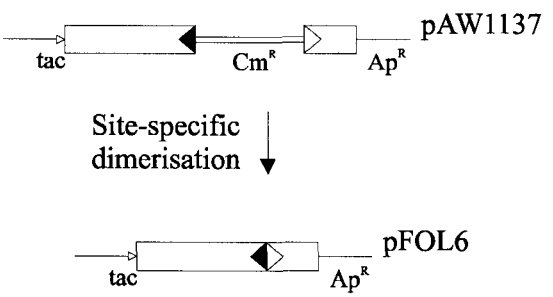

A
pAW1019
Cm^R
Km^R
(IS30)2*
pAW1105
Ap^R
pAW408
pAW409
B
pAW1026
pAW1105
pAW410
C
pAW1026
LHS/P1HS
pAW782/pAW1016
pAW1139/pAW414
D
pAW1026
LHS/RIR
pAW758
pAW1097

A



Plasmid	Mutation in the ORF-A	Frequency of Cm^S segregants after n passages (%)			
		n=1	2	3	4
pTFA89	wild type, 1-383 aa	**2.5**	**6.7**	**20.8**	**65.7**
pAW1026	*Bgl*II*, 1-352 aa	**<0.4**	**<0.2**	**<0.2**	**<0.3**
pAW1027	*Bgl*II*, 1-352 aa	**<0.2**	**<0.1**	**<0.1**	**<0.2**

B



Plasmid	Mutation in the ORF-A	Frequency of Cm^S segregants after n passages (%)				
		no induction			IPTG induction	
		n=1	2	3	4	5
pAW1137	wild type, 1-383 aa	**1.9**	**2.6**	**3.2**	**9.1**	**50.5**
pFOL3	*Bgl*II*, 1-352 aa	**<0.09**	**<0.14**	**0.09**	**0.11**	**0.26**
pFOL4	*Xba*I*, 1-302 aa	**<0.11**	**<0.09**	**<0.09**	**<0.10**	**<0.08**

Fig. 4. The intact C-terminal part of IS*30* is responsible for dimer dissolution (A) and site-specific dimerisation (B). The frequency of segregation was measured in both experiments as described in Section 2. One passage represents an overnight culture. pTFA89 and its C-terminal mutant derivatives (A) are carrying the joined terminal inverted repeats of IS*30* and an additional IS*30* end as a structural requirement for dimer dissolution. The most preferred segregation pathway of pTFA89, the loss of the element carrying the Cm^R marker, is presented. pAW1137 and its mutant derivatives (B) contain the left and right IRs of IS*30* separated by the *Cm^R* gene. ORF-A was placed under the control of the inducible *tac* promoter (small blank arrowhead). The characteristic reaction of the plasmid, the formation of $(IS30)_2$-like intermediate is shown. After the third passage IPTG was added in a final concentration of 0.1 mM in order to induce the expression of the *transposase* gene. For other symbols and abbreviations see Figs. 1 and 2.
←

although the frequency of rearrangements was 10–40-fold lower in this case than with pAW1026 (Fig. 2B). This observation suggests that in the presence of the core intermediate structure the amino acid positions 303–353 are important, but not absolutely required for the transposition of the element. The *Nco*I* mutant proved to be inactive in this experiments as well.

From the data discussed so far we draw the following conclusion: if the mutation affects the last 31–81 C-terminal amino acids of the protein, transpositional fusion is only observed in the presence of the two joined IRs of IS*30*. Therefore, according to our results, the core $(IS30)_2$ structure can suppress the effect of mutations in the C-terminal region of IS*30* transposase.

3.3. *Dimer dissolution is defective in the C-terminal mutants of IS30 transposase*

One characteristic of the wild-type intermediate is the high degree of instability, since it rapidly segregates by means of dimer dissolution to result in a product that carries only one IS*30* element (Fig. 1B, [5]). Plasmids pTFA89, pAW1026 and pAW1027 (see Fig. 2B) have the structural requirements to form such site-specific deletions (see Fig. 4A).

The ability of wild-type and mutant transposase to mediate dimer dissolution was investigated (Fig. 4A). In the control plasmid, which contained the intact transposase gene (*pTFA89*), loss of the element was observed at high frequencies, up to 65%. The most abundant segregant that represented more than 90% of the products is shown in Fig. 4A (pAW1158), while minor segregation products were identified as deletions, inversions or bacteria without detectable plasmids.

On the contrary, the *Bgl*II* mutant plasmids pAW1026 and pAW1027 (see Fig. 2B), which were also expected to yield the same or analogous segregation products, were stable and no segregants were found even after four passages. In this case, the frequency of dimer dissolution was at least 400-fold lower than with the control plasmid.

The integrity of the C-terminal of IS*30* transposase is therefore essential for the dissolution of the intermediate structure.

3.4. *Site-specific dimerisation is blocked in the C-terminal mutants*

$(IS30)_2$ intermediate is formed by the site-specific deletion of the DNA segment located between two IS*30* elements carried on the same replicon (Fig. 1B, [5]). pAW1137, which contains the intact *transposase* gene under the control of the inducible *tac* promoter, was constructed to study this process (Fig. 4B): the IRs present in the plasmid join to each other to form an $(IS30)_2$-like intermediate (see pFOL6 in Fig. 4B).

The stability of pAW1137 and its mutant derivatives was monitored upon induction of *transposase* gene expression. Experimental data demonstrated that pAW1137, the control plasmid encoding the wild-type transposase, was very unstable; the frequency of segregation was 2–3% without induction (Fig. 4B, passages 1–3) and raised to 50% upon induction (Fig. 4B, passages 4–5). On the contrary, plasmids pFOL3 and pFOL4, which resulted from pAW1137 by filling-in the

```
IS1086   238 QTLTYDQGRE MARHAELSAA TGVRVYFCDP HSPWQRGTCE NTNGLLRQ.Y LPKGTDLSVY SQEELDAIAD SLNGRPRKTL NWHSPLQVLA QVLANPTDRL PVQ 339
IS1394   236 KSMTYDQGRE MARHAEITQK TGVAIYFCDP HSPWQRGSNE NINGLIRQ.Y LPKGTDLSVY SQEQLDAIAY ELNIRPRKRF NWKCPIEVMT EVVALQHDAP ASIQ 338
ISAS2    245 HTITFDNGGE FAEHKAIEEA LGAETYFAHP YSSWERGLNE NSNGLLRQ.F IPKGTDLREV TDEDVRRAEQ WLNLRPRKCL GFRQPVKVFE EYRQAA 339
IS4351   226 HTITADNGKE FAKHEEIAQK LEIKFYFCKP YHSWERGANE NTNGLIRQ.Y IPKGKDFSEV TNKQIKWIEN KLNNRPRKRL GYLTPNEKFK QIINQNSVAF AS 326
IS1070   211 KSITVDRGRE FARYNEIEQK LGIPVYFAHP YSPEERGSNE ILNRYVRR.F IPKERKIETV SAKELDQINH WINARPMKTL NWQSPRKFFQ QFTVFG 305
IS1161   252 KSITADNGNE FSSLSKI.E. .GLDVYFAHA YSSYERGTNE NFNGLLRE.F IPKGCSLKEL NQNLLEDYTK AINERPRRIH GYQSAKKLFE LTQTA 342
IS1239   231 NSITADNGTE FSRLSDVFP. .ASDIYYAHP YASWERGTNE NHNRLIRR.C LPKG..TKTT TPKEVAAIEH WINHYPKRLF NYKSPVEMLK LANFNLKFAK R 326
IS6770   218 KSITFDCGKE FSNWKAISNQ HDIDIYFADP GTPSQRPLNE NSNGILRRNG LPKSMDFREV NQTFISSVSN QRNHIPRKSL NYRTPIEIFL SYVQEAFYSN LI 313
Spv1     231 KGIITDRRKE FSKWREMEIF AETQVYFCDA GSPQQKPLIE YMNSELR.HW FPKGTDFNKV SQKQIDWVVN VINDKLRPCL NWISSKEMFL QNI 322
IS1062   218 KSMTFDCGKE FSNWKSISNI NDIDIYFADP GTPSQRGLNE NSNGLLRKDG LPKQMDFNEV DESFIQSIAS KRNNIPRKSL NYKTPIEVFL SHICKEELSN LI 319
                 * * * * *              **  *      ** ** * *** *     **                    *  *** * *   *
                   *   *                *             *   *         **                    *
consensus        T D G E F              YF  P      RG NE N NGL R     PK                  N  PRK L N   P
                  region I                          region II                              region III
IS30     288 KSLTWDRGME LARHLEFTVS TGVKVYFCDP QSPWQRGTNE NTNGLIRQ.Y FPKKTCLAQY TQHELDLVAA QLNNRPRKTL KFKTPKEIIE RGVALTD 383
BglII*   288 KSLTWDRGME LARHLEFTVS TGVKVYFCDP QSPWQRGTNE NTNGLIRQ.Y FPKKTCLAQY TQHELD 352
XbaI*    288 KSLTWDRGME LARHL 302
```

Fig. 5. Homology between the C-terminal regions of IS*30*-related transposases. Asterisks mark the highly conserved amino acids: *, similarity (at least 8 of 11 amino acids were identical in the given position); **, identity in all 11 proteins. Regions I, II and III (shaded) are referring to conserved stretches of amino acid sequences. The consensus sequence of each region is presented. The C-terminal part of the *Xba*I* and *Bgl*II* IS*30* transposase mutants is shown below the wild-type protein. The references for the presented amino acid sequences are the following: IS*1086* [18]; IS*1394* [19]; ISA*2* [20]; IS*4351* [21]; IS*1070* [22]; IS*1161* [23]; IS*1239* [24]; IS*6770* [25]; Spv1: ORF3 protein of the single-stranded virus [26]; IS*1062* [27]; IS*30* [3]. The alignment was made by the GCG program package [28].

unique *Bgl*II and *Xba*I sites, respectively, remained stable during prolonged growth. Only a few Km^RCm^S derivatives were found in their segregation tests (pFOL3: 11/4570; pFOL4: 9/5308 colonies), and they were mostly point or deletion mutants in the Cm^R gene. However, four pFOL6-type plasmids were also found among the segregants of pFOL3, which contains the *Bgl*II* mutation, but they all derived from one of the five parallel experiments. Possibly, the *Bgl*II* mutant either retained a small residual activity or it was complemented by the host-resident copies of the element, although it should be mentioned here that the *Bgl*II* mutant did not show any residual activity in dimer dissolution (Fig. 4A).

The very high stability of the plasmids carrying mutations in the C-terminal part of the IS*30* transposase indicates that an intact C-terminus is required for site-specific dimerisation that leads to the formation of $(IS30)_2$ intermediates.

4. Discussion

The mutant analysis described in this paper shows that the C-terminal part of the IS*30* transposase is involved in various transpositional rearrangements. The introduction of mutations in this region abolished the ability of the element to mediate site-specific dimerisation, dimer dissolution and transpositional fusion. However, the deficiency in transpositional fusion can be surpassed by the presence of the core $(IS30)_2$ structure. This special kind of suppression, where the addition of a structure leads to regaining a function, was observed both *in trans* (Fig. 2A) and *in cis* (Fig. 2B). A construct containing only one IS*30* IR (LHS–RIR, see pAW758, Fig. 2A) was unable to substitute the core $(IS30)_2$ intermediate and therefore to suppress the effect of mutations in the C-terminal region.

The fact that the effect of mutations in the C-terminal of the IS*30* transposase is surpassed by the presence of the core $(IS30)_2$ structure suggests that this part of the protein is responsible for the joining of the IRs, i.e. for site-specific dimerisation. According to this hypothesis the formation of $(IS30)_2$ dimer should be affected by the C-terminal mutations — as was verified experimentally (Fig. 4B). The phenomenon that C-terminal mutants are defective both in site-specific dimerisation and dimer dissolution can refer to the common nature of these processes. Indeed, these reactions represent site-specific deletions between the IRs of the element [5]. It is also worthwile to mention in this context that dimer dissolution occurs roughly 1000-fold greater efficiency than transpositional fusion.

The specificity of the C-terminal mutations is strongly suggested by the fact that the various types of reactions were not affected to the same extent. This conclusion is based on comparison of the transpositional frequencies mediated by the *Bgl*II* mutant and wild-type control in different reactions. In the experiment presented in Fig. 2A, the activity of the *Bgl*II* mutant was reduced compared to the wild type at least 8–15-fold with all the targets, except the core $(IS30)_2$ structure, for which the decrease was only 4-fold. Note that the difference can be more significant, since no transpositional fusion product was detected at all with the targets LHS, P1HS and LHS–RIR, although we reached the detection limit of the system (transpositional frequency is lower than 10^{-9}). The suppression effect of the core intermediate structure was detected both *in trans* and *in cis*. When the core $(IS30)_2$ structure was situated in the donor replicon (Fig. 2B), the difference between the activities of the *Bgl*II* mutant and wild-type constructs was uniformly 2–3-fold with all the targets. This level corresponded to what was observed when the donor replicon carried no $(IS30)_2$ structure, and joined IRs were present only on the target molecule, i.e. *in trans* (pAW1105, Fig. 2A). These findings confirm the specific function of the C-terminal part of IS*30* transposase, namely formation of the $(IS30)_2$ dimer. Indeed, in the experiments focusing on dimer dissolution (Fig. 4A) and site-specific dimerisation (Fig. 4B), the *Bgl*II* mutant was about 200-fold less active. These observations rule out the possibility that the C-terminal mutants of IS*30* transposase reduces the transpositional ability in general. If this hypothesis were true, all types of reactions should be affected to about the same extent, in contrast with the data presented above.

Comparing the amino acid sequences of the transposases of IS*30* and related IS elements (listed in Fig. 5), the C-terminal region reveals to be highly conserved, as has already been described by Dong et al. [18], Vaughan and de Vos [22], Kapur et al. [24] and Yeo and Poh [19]. This part of the transposases show a 51% similarity and 40% identity to IS*30* transposase in average. The corresponding values for the remaining part are 37% similarity and 28% identity. The conserved amino acid residues in the C-terminal part of the proteins are concentrated in three major regions that may be considered as domains. While region I correspond to the N3 motif (D-x-G/A-x-x-Y/F, [29]), which was found in transposases of various origins; the other two regions show no significant homology to known motifs. The observation that the frequencies of rearrangements were about one order of magnitude higher in the case of the *Bgl*II* mutant compared to the *Xba*I* mutant (Fig. 2B) might be related to the fact that the *Bgl*II* mutation affects only region III, while in the *Xba*I* mutant regions II and III were both destroyed. However, at this stage of the investigations other explanations cannot be ruled out.

Further genetical and biochemical analysis of the C-terminus can clarify the molecular mechanism of intermediate formation, which is required for the efficient transposition of IS*30* and other $(IS)_2$-forming insertion sequences.

Acknowledgements: The authors thank Angelo Guidolin, Philipp Hübner and Heinrich Sandmeier for fruitful discussions. We are grateful to Rosemarie Hiestand-Neuer, Ágota Bakos-Nagy and Ilona Keresztúri-Könczöl for their highly skilled technical assistance. This work was supported by Grants T019365, T6054, F016426 and F017090 of the Hungarian Research Foundation OTKA and by Grant 3.624-0.87 from the Swiss National Science Foundation.

References

[1] Caspers, P., Dalrymple, B., Iida, S. and Arber, W. (1984) Mol. Gen. Genet. 196, 68–73.
[2] Dalrymple, B. and Arber, W. (1985) EMBO J. 4, 2687–2693.
[3] Dalrymple, B., Caspers, P. and Arber, W. (1984) EMBO J. 3, 2145–2149.
[4] Stalder, R., Caspers, P., Olasz, F. and Arber, W. (1990) J. Biol. Chem. 265, 3757–3762.
[5] Olasz, F., Stalder, R. and Arber, W. (1993) Mol. Gen. Genet. 239, 177–187.
[6] Dalrymple, B. (1987) Mol. Gen. Genet. 207, 413–420.
[7] Reimmann, C. and Haas, D. (1987) Genetics 115, 619–625.
[8] Szeverényi, I., Bodoky, T. and Olasz, F. (1996) Mol. Gen. Genet. 251, 281–289.
[9] Galas, D.J. and Chandler, M. (1989) in: Berg, D.E. and Howe,

M.M. (Eds.), Mobile DNA, American Society for Microbiology, Washington, DC, pp. 108–162.
[10] Syvanen, M. (1988) in: Kucherlapati, R. and Smith, G.R. (Eds.), Genetic Recombination, American Society for Microbiology, Washington, DC, pp. 331–356.
[11] Miller, J.H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[12] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[13] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[14] Chang, C.J. and Cohen, S.N. (1978) J. Bacteriol. 134, 1141–1156.
[15] Dente, L., Cesareni, G. and Cortese, R. (1983) Nucl. Acids Res. 11, 1645–1655.
[16] Stalder, R. and Arber, W. (1989) Gene 76, 187–193.
[17] Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103–119.
[18] Dong, Q., Sadouk, A., van der Lelie, D., Taghavi, S., Ferhat, A., Nuyten, J.M., Borremans, B., Mergeay, M. and Toussaint, A. (1992) J. Bacteriol. 174, 8133–8138.
[19] Yeo, C.C. and Poh, C.L. (1996) Gene 175, 109–113.
[20] Gustafson, C.E., Chu, S. and Trust, T.J. (1994) J. Mol. Biol. 237, 452–463.
[21] Rasmussen, J.L., Odelson, D.A. and Macrina, F.L. (1987) J. Bacteriol. 169, 3573–3580.
[22] Vaughan, E.E. and de Vos, W. (1995) Gene 155, 95–100.
[23] Giffard, P.M., Rathsam, C., Kwan, E., Kwan, D.W., Bunny, K.L., Koo, S.P. and Jacques, N.A. (1993) J. Gen. Microbiol. 139, 913–920.
[24] Kapur, V., Reda, K.B., Li, L.L., Ho, L.J., Rich, R.R. and Musser, J.M. (1994) Gene 150, 135–140.
[25] Thorisdottir, A.S., Carias, L.L., Marshall, S.H., Green, M., Zervos, M.J., Giorgio, C., Mermel, L.A., Boyce, J.M., Medeiros, A.A., Fraimow, H. and Rice, L.B. (1994) J. Infect. Dis. 170, 1539–1548.
[26] Renaudin, J., Aullo, P., Vignault, J.C. and Bove, J.M. (1990) Nucl. Acids Res. 18, 1293–1298.
[27] Hirt, H., Wirth, R. and Muscholl, A. (1996) Mol. Gen. Genet. 252, 640–647.
[28] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucl. Acids Res. 12, 387–395.
[29] Rezsöhazy, R., Hallet, S., Delcour, J. and Mahillon, J. (1993) Mol. Microbiol. 9, 1283–1295.