The EMBO Journal vol.8 no.9 pp.2761 – 2773, 1989

Characterization of the staphylococcal β-lactamase transposon Tn*552*

S.-J.Rowland and K.G.H.Dyke

Microbiology Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

Communicated by P.Nurse

The staphylococcal β-lactamase transposon Tn*552* is a member of a novel group of transposable elements. The organization of genes in Tn*552* resembles that of members of the Tn*21* sub-group of Tn*3* family transposons, which transpose replicatively by cointegrate formation and resolution. Thus, a possible resolution site ('*resL*') and a resolvase gene (*tnpR* or '*binL*') have been identified. However, consistent with the fact that Tn*552* generates 6 bp (rather than 5 bp) flanking direct repeats of target DNA, neither the putative transposase protein, nor the terminal inverted repeats of Tn*552* are homologous to those of Tn*3* elements. Tn*552*, like phage Mu and retroelements, is defined by the terminal dinucleotides 5′ TG . . CA 3′. A naturally occurring staphylococcal plasmid, pI9789, contains a Tn*552*-derived resolution system ('*resR*–*binR*') that acts as a 'hotspot' for Tn*552* transposition; insertion creates a segment of DNA flanked by inversely repeated resolution sites, one (*resR*) on pI9789 and the other (*resL*) on Tn*552*. The putative Tn*552* resolvase, the most closely related of known resolvases to the homologous DNA invertases, initially was identified as a DNA invertase ('Bin') as a result of its ability to mediate efficient inversion of this segment *in vivo*.

Key words: Staphylococcus aureus/β-lactamase transposons/site-specific recombination/invertible DNA/Tn*552* resolvase

Introduction

The translocation of β-lactamase-mediated resistance to β-lactamase antibiotics in the Gram-positive bacterium *Staphylococcus aureus*, from the chromosome of strain PS80 (NCTC 9789) to a specific site on the resident, naturally occurring β-lactamase negative plasmid pI9789, was first described by Asheshov (1969). Translocation appeared to be replicative since the β-lactamase genes were not lost from the chromosome. The β-lactamase region on a related plasmid, pI524, comprising the β-lactamase structural gene (*blaZ*), genes for the control of β-lactamase production (*blaI*, *blaR1*; Dyke, 1979) and a reversibly invertible segment of DNA (*inv*; Murphy and Novick, 1979), was judged by restriction mapping to be identical to the inserted element, which has been designated 'Tn*552*' (Figure 1).

Electron microscopic analysis of pI524 revealed that *inv* was flanked by ~650-bp inverted repeat sequences (IR_L and IR_R). These were presumed to contain the *cis*-acting sites required for reversible, *rec*-independent (site-specific) inversion and one was believed already to be present at the site of insertion of Tn*552* on pI9789 (Shalita *et al.*, 1980). Thus it has been suggested that the site-specific translocation of Tn*552* to pI9789 occurred by recombination between IR_R and IR_L, rather than by actual transposition (Murphy, 1988). However, three other β-lactamase elements in *S.aureus*, Tn*4002*, Tn*3852* (possibly identical to Tn*4002*) and a 'Tn*4201*-like' element, which are highly homologous to Tn*552* but are not associated with invertible segments, each contains *inv*, frozen in the (−) orientation because of the absence of IR_R, plus IR_L (Lyon and Skurray, 1987; Gillespie *et al.*, 1988). There is substantial evidence, both genetic and physical, that these elements are transposable (Gillespie *et al.*, 1988): e.g. short inverted repeats define both Tn*4002* and the 'Tn*4201*-like' element. Tn*4002* occurs at different sites and in different orientations in members of the pSK1 family of staphylococcal plasmids, and a Tn*4002*-like element has been identified on the *S.aureus* chromosome free of flanking plasmid DNA sequences; furthermore, Tn*4002* (like Tn*552*) has been observed to translocate from chromosome to plasmid during storage of strains. Conclusive proof of transposition (*rec*-independent translocation) has so far been published only for Tn*4201* (Weber and Goering, 1988).

The plasmid pI9789::Tn*552* is referred to in this paper as pI9789blaIblaZ, or 'pS1 parent'. Johnston and Dyke (1971) treated *S.aureus* PS80d.[pS1 parent] with the chemical mutagen ethyl methanesulphonate (EMS) and isolated a plasmid with a mutation that caused β-lactamase production to be temperature sensitive. This plasmid was designated 'pS1' (pI9789blaIblaZseg-1). The *S.aureus* strain 'PS80d.' provided a convenient background in which to perform the mutagenesis because the chromosomal *blaIblaZ* genes are inactive or have been deleted. The nature of the *seg-1* mutation remains unknown. A β-lactamase negative deletion derivative of pS1, 'pΔD', subsequently was produced by UV-irradiation of transducing phage propagated on PS80d.[pS1] (Johnston and Dyke, 1971).

The region immediately downstream of *blaZ* in pS1 previously was found to encode 'Bin', a member of a closely related 'superfamily' of site-specific recombinases that includes the DNA invertases Hin (*Salmonella typhimurium*), Gin (phage Mu), Cin (phages P1 and P7) and Pin (*Escherichia coli*) and the resolvases of Tn*3*-class transposons (Rowland and Dyke, 1988). The DNA invertases mediate 'genetic switches' (Plasterk and van de Putte, 1984; Glasgow *et al.*, 1989): inversion of a DNA segment flanked by inversely repeated recombination sites switches expression between two alternative sets of genes. However, no such function has been established for 'Bin'. Replicative transposition of Tn*3*-class elements proceeds via the formation and resolution of a cointegrate intermediate, which consists of the donor and recipient replicons fused by directly repeated copies of the transposon. The two replicons, both now containing a copy of the transposon, are regenerated

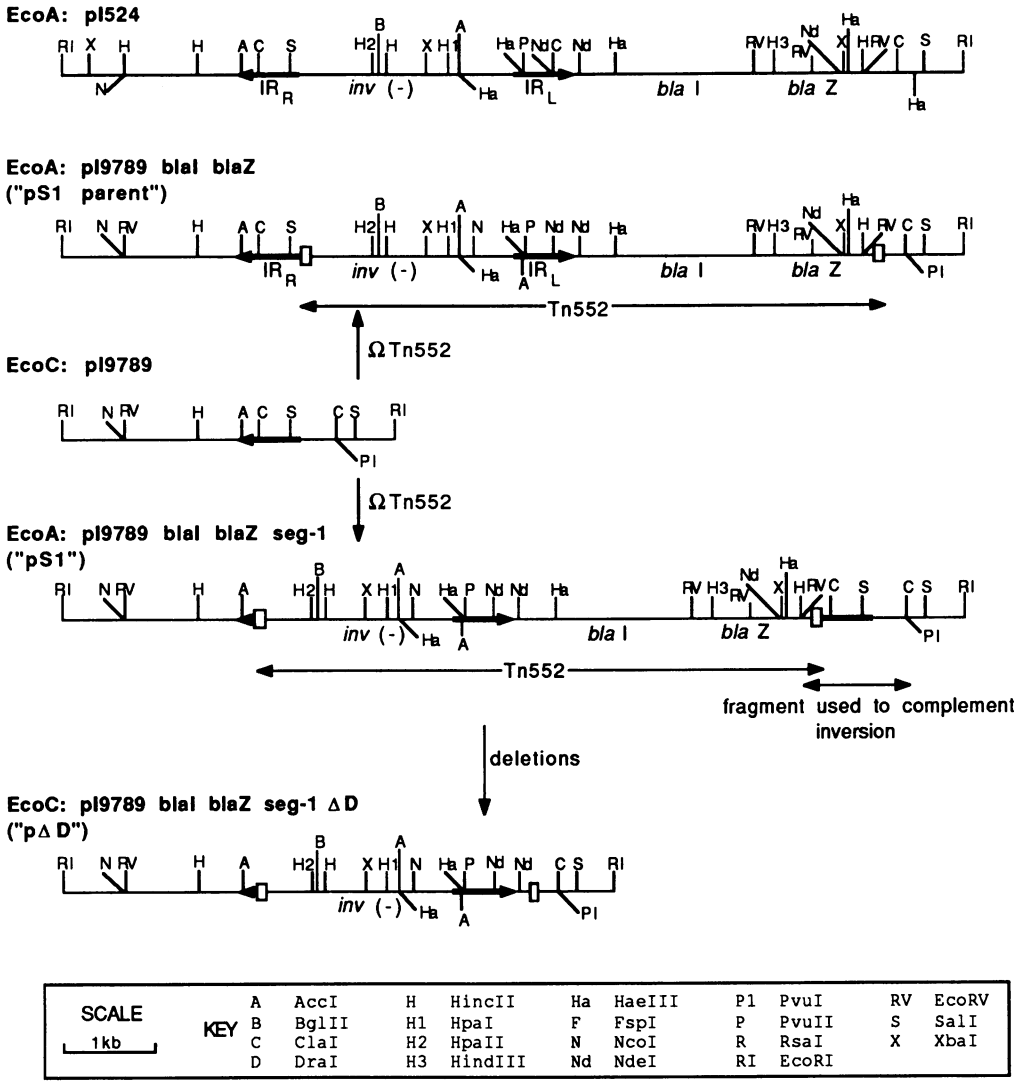

Fig. 1. Restriction endonuclease maps of the invertible (*inv*) and β-lactamase (*blaIblaZ*) regions of the staphylococcal plasmids pI524, pI9789, pI9789blaIblaZ ('pS1 parent'), pI9789blaIblaZseg-1 ('pS1') and pI9789blaIblaZseg-1ΔD ('pΔD'). *EcoA* and *EcoC* are the largest and third largest respectively of the four fragments produced upon digestion of each plasmid with *Eco*RI. These plasmids also carry genes that confer resistance to various heavy metals and inorganic ions (Shalita *et al.*, 1980). The relationships between pI9789, pS1 parent, pS1 and pΔD, with respect to Tn*552* sequences, are shown. The open boxes represent the terminal inverted repeats of Tn*552* that have been identified by sequencing (sequence data are not available for pI524). The recombination sites of *inv* (which is shown in the (−) orientation in each case) are represented by arrowheads. IR_L and IR_R, the ~750 bp inverted repeats that flank *inv* in pI524 and pS1 parent, are represented by bold lines (IR_L and IR_R were defined for *inv* in the (+) orientation). The sequences that comprise IR_L and IR_R in pI524 and pS1 parent are also highlighted at their locations in the other three plasmids. The 1.1 kb pS1 *Eco*RV−*Pvu*I restriction fragment used to complement inversion in *E.coli* is indicated.

by resolvase-mediated recombination between the transposon-borne resolution (*res*) sites (Grindley and Reed, 1985; Sherratt, 1989). This paper describes firstly, the evidence that 'Bin' mediates the inversion of *inv*, and secondly, how the elucidation of the relationships between pI9789, pS1 parent, pS1 and pΔD has provided conclusive proof that Tn*552* is a novel transposable element whose transposition may also involve cointegrate resolution; we suggest that 'Bin' is actually a homologue of the Tn*552* resolvase.

Results

Analysis of inversion in E.coli and identification of a 'DNA invertase'

For each of the four plasmids of which the invertible segments were investigated (pI524, pS1 parent, pS1 and pΔD) the *Eco*RI restriction fragment, containing the invertible segment (*inv*) and the adjacent *blaIblaZ* loci, was cloned in *E.coli* MC1061 and a detailed restricton endonuclease map was constructed (Figure 1). Thus it was discovered that, in the time taken to isolate plasmid DNA from transformed cells (by rapid alkaline lysis), efficient *recA*-independent inversion leading to a plasmid population consisting of an approximately equimolar mixture of molecules with *inv* in the (+) and (−) orientations had occurred only for *inv* of pS1. The invertible segments of pI524, pS1 parent and pΔD were almost completely 'frozen' in either the (+) or the (−) orientation at an equivalent stage, and subsequent sub-culturing had no significant observable effect on the fraction of plasmid molecules (estimated at <5%) with *inv* in the opposite orientation. Inversion of pS1 *inv* was found to be independent of the *E.coli* chromosomal *pin* (Plasterk and van de Putte, 1985) and *fim* (Abraham

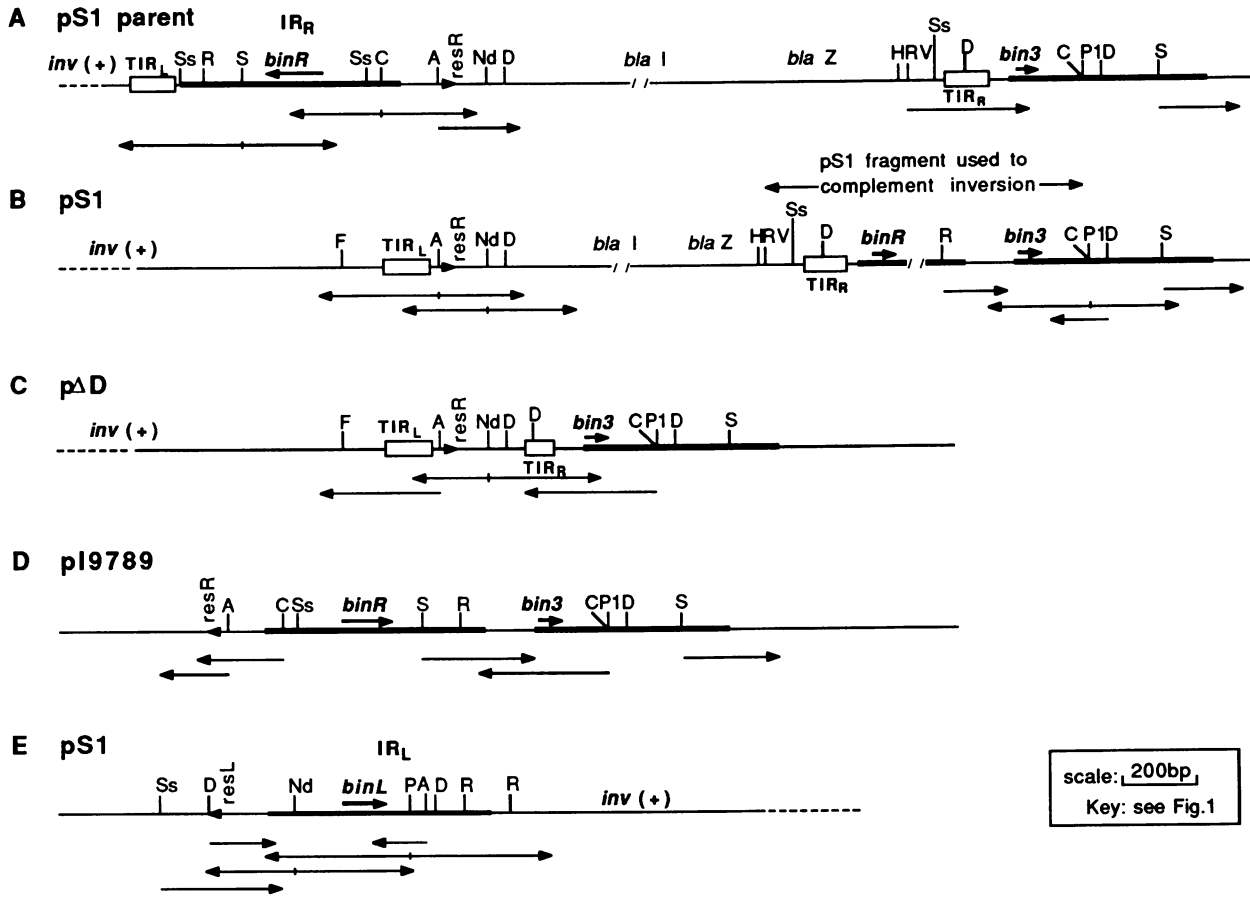

Fig. 2. Sequencing strategies for the *bin* genes (*binR*, *binL* and *bin3*) and the recombination sites of *inv* (*resL* and *resR*) in pS1 parent, pS1, pΔD and pI9789. Each arrow indicates the direction and extent of sequence obtained. The sequencing of *binR* in pS1 has been described previously (Rowland and Dyke, 1988). The inverted repeats of Tn*552*, TIR_L and TIR_R, are represented by open boxes (as in Figure 1), and the genes (bold lines) and recombination sites (arrowheads) are labelled.

et al., 1985) inversion functions (which are inactive in strains WA3782 and HB101 respectively).

Simple complementation experiments therefore were devised to locate the gene for the suspected 'DNA invertase' (detailed in Materials and methods). The strategy was to use the 'frozen' *inv* of pΔD to assay the ability of various pS1 restriction fragments to promote inversion in *trans*. Thus, aliquots of transformation-competent *E.coli* cells that harboured the cloned pΔD *inv* were supertransformed separately with sub-fragments of *Eco*A:pS1 cloned in a compatible vector. Plasmid DNA (consisting of a mixture of the two constructs) was prepared from overnight cultures derived from single transformants and complementation was defined as the appearance of restriction fragments characteristic of the opposite orientation of the 'frozen' *inv*. When the pS1 fragment tested itself contained *inv* (*Eco*A:pS1), both pΔD *inv* and pS1 *inv* exhibited inversion, but equilibration of the latter occurred much more slowly than previously had been observed (four to six rounds of sub-culturing were required), consistent with titration of a *trans*-active DNA invertase by the extra copies of pΔD *inv* present; in this particular experiment the copy number of the pΔD clone was at least five times that of the pS1 clone (as judged by restriction mapping). A 1.1 kb pS1 *Eco*RV−*Pvu*I restriction fragment located downstream of *blaZ* (Figure 1) was identified as the source of the putative invertase, which also promoted the inversion of pS1 parent *inv* in *trans*; *Eco*A:pS1 parent and *Eco*A:pI524 were unable to support detectable inversion. Complementation was independent of the initial orientation of the 'frozen' *inv*. The region was sequenced and found to contain '*bin*', a gene encoding a member of the resolvase−DNA invertase superfamily of site-specific recombinases (Rowland and Dyke, 1988). Potential transcriptional promoter and ribosome binding site (RBS) sequences were identified upstream of '*bin*' in pS1 (Rowland and Dyke, 1988). Further evidence for the role of this gene was the inability of a pS1 *Eco*RV−*Rsa*I fragment, in which the extreme 3′ end of *bin* [encoding helix 2 of the postulated DNA-binding helix-turn-helix structural motif (Pabo and Sauer, 1984; Anderson *et al.*, 1987)] was deleted, to complement inversion.

The invertible DNA segments of pS1 parent and pS1

Comparative restriction mapping of the invertible and β-lactamase regions of pS1 parent, pS1 and pΔD suggested that *bin*, the 'DNA invertase' gene located downstream of *blaZ* in pS1, comprises most of IR_R, one of the inverted repeats of *inv* in pS1 parent (Figure 1). This rearrangement was confirmed by DNA sequencing, and henceforth '*bin*' is referred to as '*binR*' (Figure 2A and B). Predictably, therefore, IR_L of *inv* encodes a second potential DNA invertase, 'BinL', in both pS1 parent and pS1 (Figures 2E and 3). Compared with *binR*, *binL* has 48 silent and 15 coding base changes, none of the latter involving amino acid

A

```
                 _________res siteI___________
IRL (-) AGATGAGAGTGTACAGAAAATAACCATTTTCTGTACACTCTCTTTTTATGCCGAAATA-AGTGAGTAATACATATATTTTCGTGTTCAAAAA
IRR (-) TTAAAACT                                 A     G  T T T   T  CT    -      --    AA       T-
                                                                   RBS     AccI
ACCCAACAAAAAATCTATGTTCACAAAAGACCTTGAAATTAAGTTATGGTACACTGTATATGTATATAGGAGGTTTGTATT TTG AAA ATA GGT
        C              T   TATAG  AG            A    A  A A  AC            C
                                                                              NdeI
TAT GCA AGA GTT TCA ACT GGA TTA CAA AAT CTG AAT TTA CAG GAG GAC CGA TTA AAC GCA TAT GGT TGT GAA AAG
                          T   G         T                 A   T           T CA                    A
                                                                ClaI

ATA TTT AGC GAC CAT ATT AGT GGT TCA AAA AGT AAA AGA CCT GGA TTA GAT AAA GCA ATT GAA TTT GCG AGA TCA
              T   C   A       A G               C T       T          G    C                   G

GGA GAT ACG ATT GTT GTT TGG AGA CTC GAC AGA CTA GGA CGT AAT ATG GAG GAT TTA ATC ACA TTA GTT AAT GAG
                                  A   T                   C      CA                               A

CTT AAC GAA CGA GGT GTA AGT TTC CAT AGT TTA GAA GAA AAT ATA ACG ATG GAT AAA TCA AGT TCT ACA GGA CAA
      T A T       C   G       T                               A
                                                                                    PvuII HaeIII
TTA TTA TTT CAT TTA TTC GCA GCG TTT GCA GAA TTT GAA CGA AAT TTA ATT TTA GAA CGT TCT TCA GCT GGC CGA
  G                           A           G           T                                   A   T
                                          AccI
ATT GCA GCA CGT GCT AGA GGA CGC TAT GGA GGT AGA CCC GAA AAA TTA AAT CAA AAA GAT TTA AAC TTA CTT AAG
          T                                 C                       A   C         G  CT       C   A
                                          SalI                                  RsaI
ACA CTT TAT GAT AAT GGC ACG CCA ATT AAG ACA ATT GCA GAG CAG TGG CAA GTT TCA CGT ACA ACG ATT TAT CGT
                      A   C           A   T               A     A                     A   C

TAC CTC AAT AAA TTG GAG GAA AAG GAA GAT GAA AAA CAA GGA GAA GTA TCT AAC TAA TTTTTCTGAAGCCAAACGTAAAG
  T           G   A A T A C C A     A   A   G T A   TA  AAATATTACTAAGTTTTAAGAAAGCTAAAAACTAGACAGGAT

AAGCTACTCAAAAGTACAATATTATAAAACCTTTTATCTTGGGCAAGCAATCTTTATC...
ACACCCTAATCATACAAATGAAGTTGTTGAATTTATAACAACTGTTACTTTACTTG...
```

B

```
BinL   MKIGYARVSTGLQNLNLQEDRLNQYGCEKIFSDHISGSKSKRPGLDKAIEFARSGDTIVVWRLDRLGRN
BinR             **  *   **  **    *   *** *A   * *  R*    *         **    *

BinL   MEDLITLVNELNERGVSFHSLEENITMDKSSSTGQLLFHLFAAFAEFERNLILERSSAGRIAARARGRYG
BinR    A       * *N ** *       *         *      *  *  *        **   *

BinL   GRPEKLNQKDLNLLKTLYDNGTPIKTIAEQWQVSRTTIYRYLNKLEEKEDEKQGEVSN
BinR    *     KQ *T **     **  **   * K    **  *  **NNQ NKDK
                              helix1      helix2
```

Fig. 3. (**A**) Alignment of the nucleotide sequences of IR_L and IR_R, the inverted repeats of *inv* in pS1 parent, which encode *binL* and *binR*, respectively; *inv* is in the minus (−) orientation. Only nucleotides which differ from those in IR_L are shown for IR_R, and gaps (shown as −) have been introduced to maximise homology. Strand exchange occurs at the centre of *resL* and *resR* sub-sites I (possible sub-sites II and III are identified in Figure 6A). The probable ribosome binding sites (RBS) and the codons of the putative DNA-binding α-helices are underlined. The sequences of the genes and recombination sites *per se* are identical in pS1 parent and pS1. The sequence in bold type downstream of *binR* is the left-hand end of Tn*552* (Figure 6A). (**B**) Alignment of the predicted amino acid sequences of BinL and BinR (using the single-letter amino acid code); only residues which differ from those in BinL are shown for BinR. The sequence of BinR has been reported previously (Rowland and Dyke, 1988). An asterisk represents a silent codon change in *binR* (cf. *binL*). Helix 1 and helix 2 of the DNA-binding motif are indicated.

```
                          left arm                  right arm

Tn2501           T G A C C G - A A A - C C G G T C G G - T T T - C G T A C A

Tn21             C G T C A G G T T G A G G C A T A C C C T A A C C T G A T G   } Tn21
Tn501            C G T C A G A T T G A G G C A T A C C C T A A C T A G A T G   } family
Tn1721           T G T C A A T C T A G G C T A T A C C C T A A C T T G A T G   }

Tn3              C G T T C G - A A A T A T T A T A A A T T A T - C A G A C A   } Tn3
Tn1000           C G T C C G - A A A T A T T A T A A A T T A T - C G C A C A   } family
IS101            T G T C T G - A T A T A T C G A A A C A T A A - T G T A C A   }

pIP404           A G T A T C - A A A A A T C C A C A T T T T T - G A T A C A   } Gram
Tn917            G G T A C C - A T A A A C G A C C G T T T A T - G G T A C T   } +ve
Tn552 resL(R)    T G T A C A - G A A A A T A A C C A T T T T C - T G T A C A   }

hixL               T T A T C - A A A A A C C T T G G T T T T C - A A G A A
hixR               T T A T C - A A A A A C C T T C C A A A A G - G A A A A
cixL               C A A T C - C T A A A C C T T G G T T T A A - G A G A A
cixR               T T A T C - C A A T A C C T T G G T T T A A - G A G A A
pixL               C T C T C - G A A A A C C T T G G T T T G G - G A G A A
pixR               T T T T C - A T A A A C G T T G G T T T G G - G A G A A
gixL(R)            T T A T C - C A A A A C C T C G G T T T A C - A G G A A
```

Fig. 4. Alignment of the recombination site sequences of resolvases (*res* sub-sites I) and DNA invertases (*hix, cix, pix, gix*). The central dinucleotide (across which strand exchange occurs) is in bold type, and inverted repeat sequence is underlined.

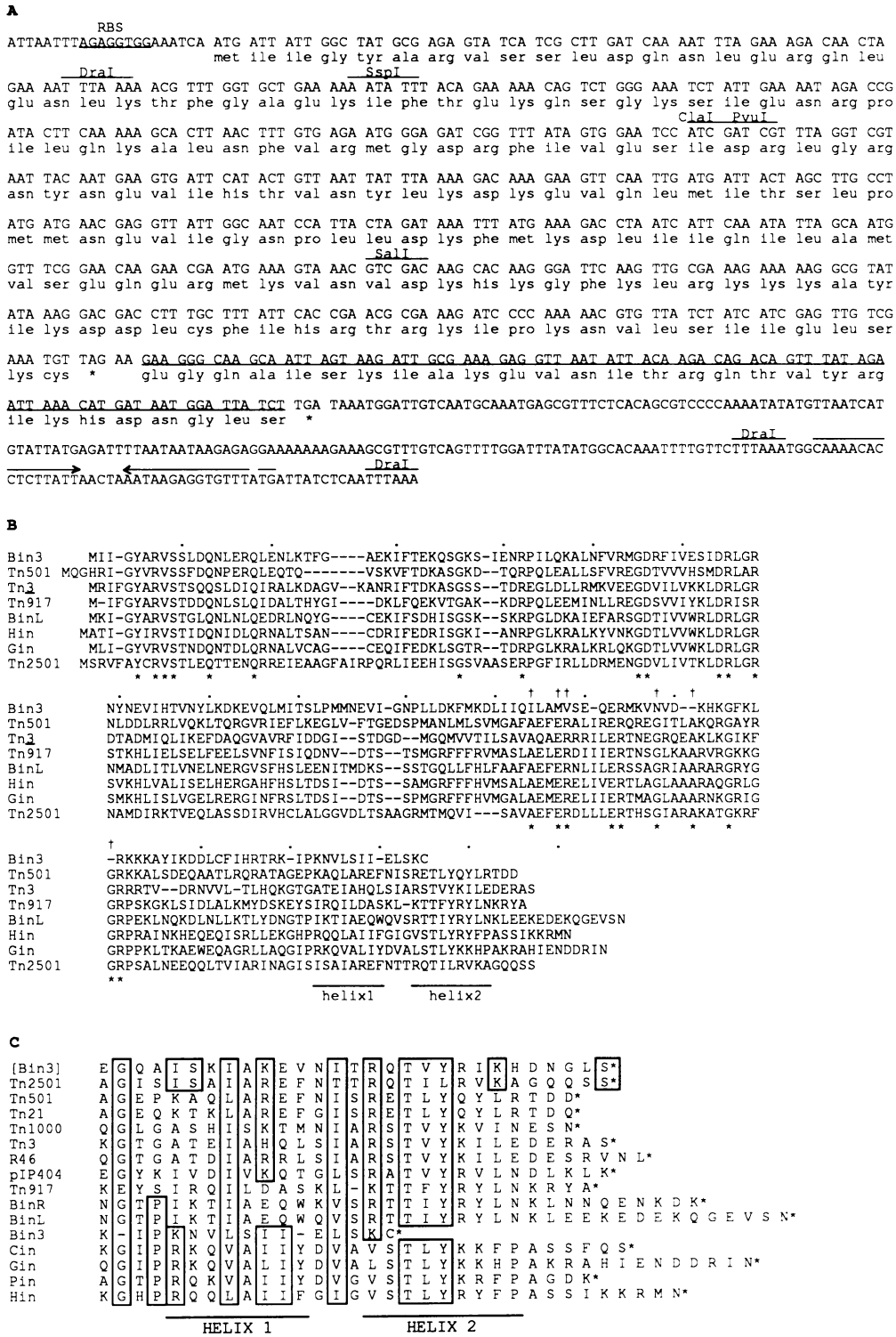

Fig. 5. (**A**) The DNA sequence of *bin3* and the predicted amino acid sequence of the protein (three-letter code). The 29 codons of the out-of-frame ORF downstream of the stop codon of *bin3* are underlined. Each stop codon is marked with an asterisk. The probable ribosome binding site of *bin3* and the downstream inverted repeat are indicated. (**B**) Alignment of the predicted amino acid sequence of Bin3 with the sequences of representative DNA invertases and resolvases [Hin, Gin, BinL (TnpR Tn*552*) (see Figure 3B) and TnpR of transposons Tn*917*, Tn*2501*, Tn*501* and Tn*3*]; gaps have been introduced to maximize homology. The third line of each sequence corresponds to the C-terminal DNA-binding domain. The 23 residues that are totally conserved in all members of the superfamily (excluding Bin3) are indicated by * below the Tn*2501* sequence; six of these (indicated by † above the Bin3 sequence) are not conserved in Bin3 in this alignment. The Tn*917* C-terminal sequence was obtained by 'deleting' an A from the published DNA sequence (Sherratt, 1989). (**C**) Alignment of the putative DNA-binding helix-turn-helix motifs of resolvases and DNA invertases with the C-terminus of Bin3 and the motif encoded downstream of *bin3* (designated [Bin3]). An asterisk indicates a C-terminal amino acid residue. Highly conserved residues and some notable homologies are boxed.

residues expected to be critical for function (Hatfull *et al.*, 1988). Seven of the amino acid substitutions are the consequence of nucleotide sequence divergence at the extreme 3′ end of the gene, immediately following the codons of the putative helix-turn-helix motif. The size of each IR, ~750 bp from sequencing, agrees quite well with

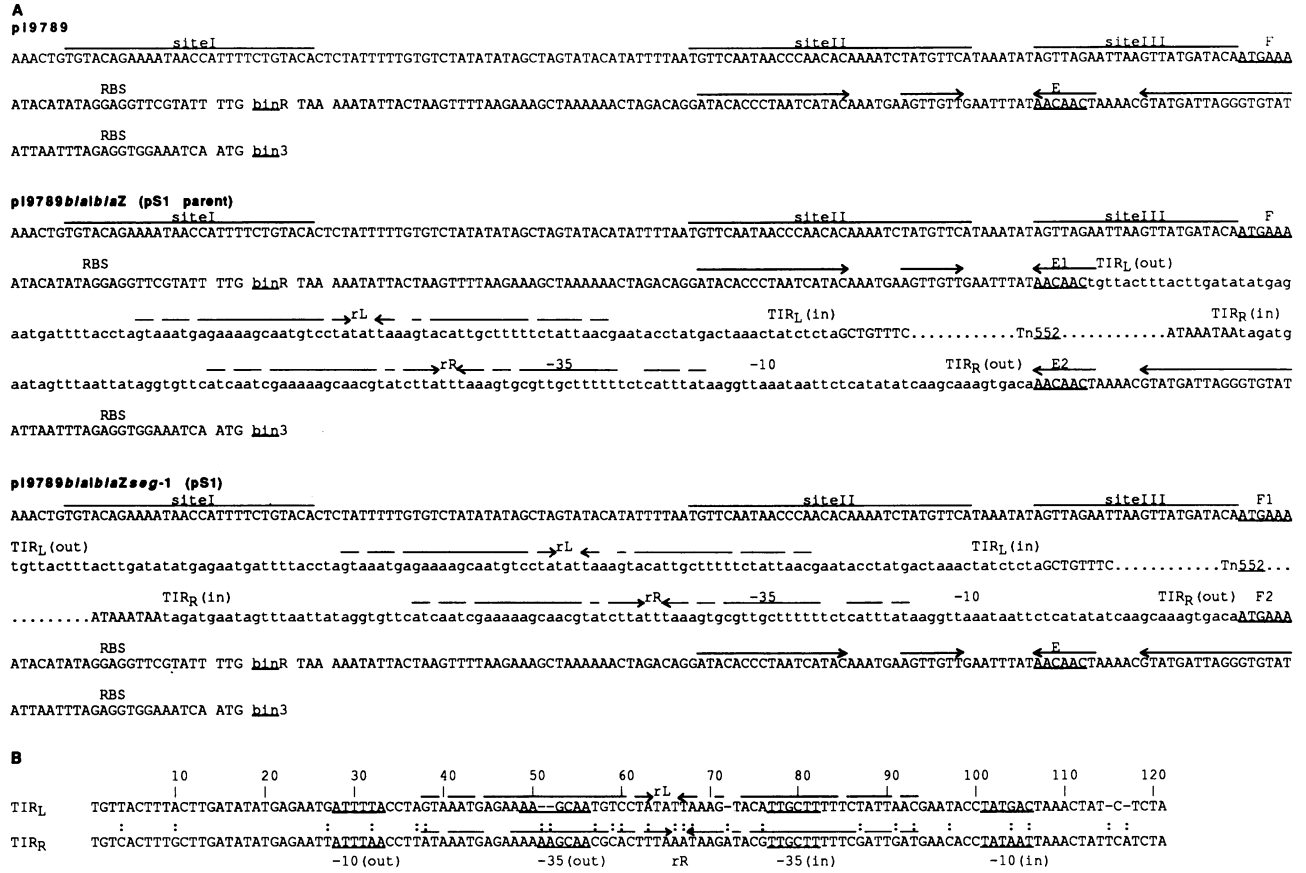

Fig. 6. (**A**) The DNA sequence relationships between pI9789, pS1 parent and pS1; the reference point is *resR* (refer to Figure 7). The target sites for the insertion of Tn*552* in pI9789 (E and F, downstream and upstream of *binR* respectively) and the flanking direct repeats generated (E1 + E2 and F1 + F2) are underlined. Site E is within a large inverted repeat located between *binR* and *bin3*. The internal sequence of Tn*552* (unpublished data) is not shown and, apart from the initiation and termination codons, the sequences of *binR* and *bin3* have been omitted for clarity. The probable ribosome binding sites of these genes are indicated. The terminal inverted repeats of Tn*552* (TIR_L and TIR_R) are in lower case, and their outer and inner ends are defined; each contains a large, imperfect inverted repeat sequence (rL and rR). The putative −35 and −10 promoter sequences of *binR* in pS1 are within TIR_R. (**B**) Alignment of the nucleotide sequences of the terminal inverted repeats of Tn*552*; colons denote nucleotide differences. The sequence shown for TIR_R (to which the numbering refers) is complementary to that in Figure 6A. Putative transcriptional promoters, reading in and out of each repeat, are identified by homology with that labelled in Figure 6A, for which there is functional evidence (see text).

the 650 bp deduced by heteroduplex mapping (Murphy and Novick, 1979). Thus, each of the flanking inverted repeats of *inv* in pS1 parent (and presumably also in pI524) contains a recombinase gene.

The probable recombinational crossover sites of *inv* were localized to ~60 bp sequences upstream of *binL* and *binR* in IR_L and IR_R respectively, by sequencing from an *Acc*I site (mapped close to one end of *inv*; Figure 1) out into the flanking, non-invertible DNA for both the (+) and (−) orientations of *inv*. Identical, perfect inverted repeat sequences at the extreme outer boundaries of *inv* could be aligned with the known crossover sites of the various resolvases (*res* sub-sites I) and the DNA invertases Hin (*hix*), Gin (*gix*), Cin (*cix*) and Pin (*pix*) (Figure 4); the recombination sites of *inv* have been designated '*resL*' and '*resR*' of Tn*552*. The invertible segments of pS1 parent, pS1 and pΔD are defined by exactly the same recombination sites.

In pS1 parent, pS1 and pΔD, *binL* and *binR* are preceded by excellent potential RBS sequences (Figure 3A). However, the only gene with obvious adjacent −35 and −10 transcriptional promoter sequences that conform reasonably well to the *E.coli* consensus is *binR*, in its non-invertible location downstream of *blaZ*, in pS1 (see below). The complementation analysis described above suggests that *binR* indeed is expressed from this putative promoter in *E.coli* (which probably, therefore, also is active in *S.aureus*). In summary, the potential for *binR* to be expressed efficiently provides a plausible explanation for the unique ability of pS1 *inv* to invert efficiently in *E.coli*. The existence of staphylococcal promoter signals for *binR(L)* that are poorly recognized in *E.coli* (and which have not yet been identified) would explain both the general ability of *inv* to invert in *S.aureus* and the much lower rates of inversion of pS1 parent (pI524) *inv* and pΔD *inv* in *E.coli* (*binR* has been deleted in pΔD); it is possible that the inversion observed in these cases represents efficient recombination confined to a short interval following transformation [as suggested for Cin-mediated recombination events by Iida (1986)]. Readthrough transcription of *blaZ* in pS1 might contribute to the expression of *binR* in *E.coli*, because deletion of the upstream region and 5′ end of the former gene significantly reduces the rate of inversion of *inv* in *Eco*A:pS1 (S.-J.Rowland, unpublished data). The promoters of other DNA invertase and resolvase genes have been proposed, or demonstrated, actually to overlap the respective recombination sites [a 'pseudo-recombination site' in the case of *cin* (Hiestand-

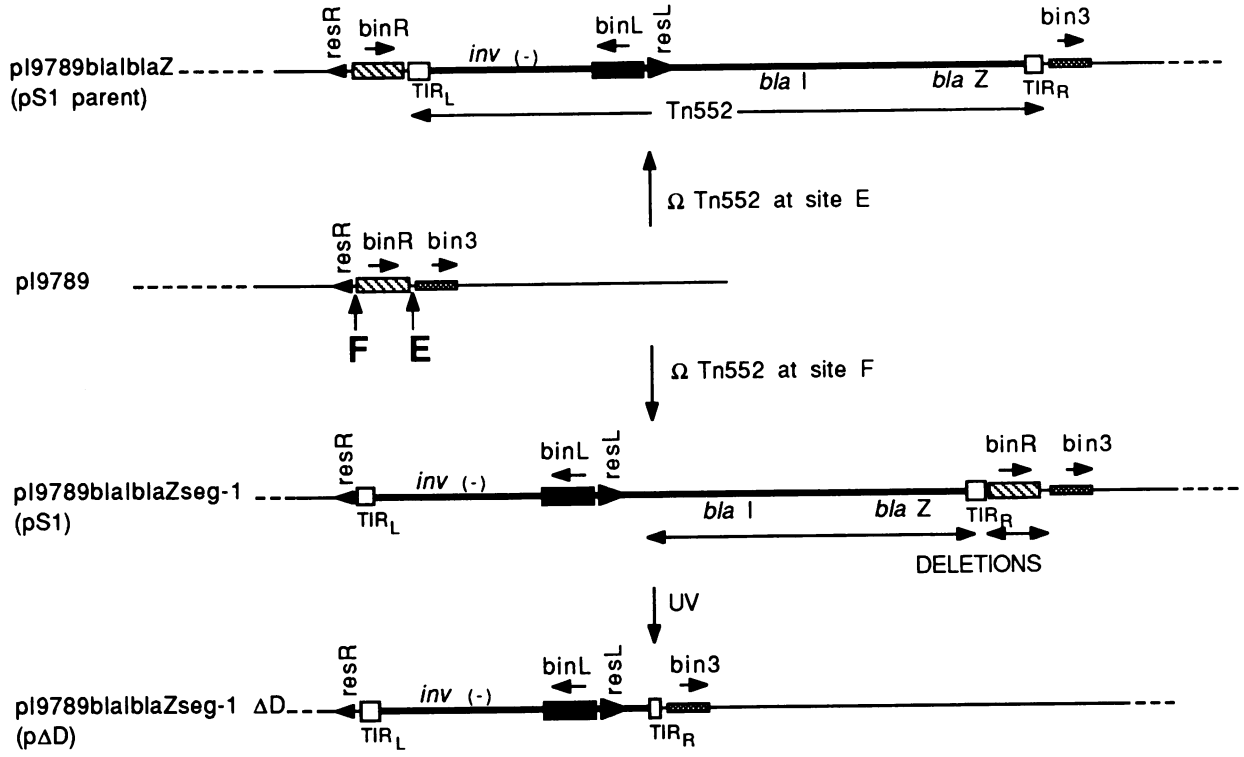

Fig. 7. Schematic representation of the relationships between the plasmids pI9789, pS1 parent, pS1 and pΔD. Tn*552* has inserted (Ω) in the same orientation, with respect to *resR–binR*, at two different sites in pI9789: site E (to generate pS1 parent) and site F (to generate pS1). The terminal inverted repeats of Tn*552* (TIR_L and TIR_R), which exists when *inv* is in the (−) orientation, are represented by open boxes. The recombinase genes associated with Tn*552* (*binR*, *binL* and *bin3*) and the recombination sites of *inv* (*resR* and *resL*) are indicated. pS1 and pΔD were isolated following treatment of pS1 parent and pS1 respectively, with EMS and UV irradiation (see text for details).

Nauer and Iida, 1983)], providing the means by which *tnpR* expression can be autoregulated; this also may be true for the expression of *binL* and *binR* in *S.aureus*.

'Bin3', another member of the resolvase–invertase family

Downstream of *blaZ* in pS1 parent, and of *binR* in pS1 and pI9789, is a gene ('*bin3*') that encodes a divergent member of the resolvase-invertase family (Figure 5). Resolvases and DNA invertases share considerable amino acid sequence homology. However, out of 23 totally conserved residues, only 17 also are conserved in Bin3 (Figure 5B). Analysis of γδ resolvase (Hatfull *et al.*, 1988) and Hin (Bruist *et al.*, 1987) has shown that a hinge region, containing a conserved glycine that is missing in Bin3, connects two major structural domains: an N-terminal catalytic and dimerization domain and a C-terminal DNA-binding domain; the postulated helix-turn-helix motif is at the extreme C-terminus. Bin3, however, is truncated: helix 2 of the motif has been deleted. It is notable that helix 1 resembles that of a DNA invertase (Hin, Gin, Cin, Pin). Furthermore, a complete helix-turn-helix motif is encoded immediately after the stop codon of, and in the −1 reading frame with respect to, *bin3*, and this motif is more resolvase-like (Figure 5C).

There is a large (16 bp) inverted repeat downstream of *bin3* which might (i) participate in transcription termination (although it is somewhat distant from the end of the gene) or (ii) constitute part of the adjacent *asa-asi-ant* operon (encodes resistance to arsenate, arsenite and antimony) which is flanked by ~200 bp inverted repeats and may once have been transposable (Shalita *et al.*, 1980; Lyon and Skurray, 1987).

Tn552 defined

The DNA sequence data that describe the relationships between pI9789, pS1 parent, pS1 and pΔD (Figures 6A and 7) is considered to prove that Tn*552* is a transposable element. Thus, pS1 parent and pS1 appear to be the products of independent transposition events to sites downstream and upstream of *binR* in pI9789 respectively: Tn*552* is flanked by 6 bp direct repeats of target DNA in both locations. This 6.5 kb element, which has imperfect terminal inverted repeats of ~120 bp (designated 'TIR_L' and 'TIR_R'; Figure 6B), comprises *inv* in the (−) orientation, IR_L and the β-lactamase genes (*blaIblaZ*). Thus, one of the recombination sites of *inv* (*resL*) is within Tn*552* and the other (*resR*) is plasmid borne. In pI9789, *binR* and *bin3* are separated by a large inverted repeat, identical to that in pS1, which contains the original target site for transposition from the chromosome of *S.aureus* PS80 ('site E' in Figures 6A and 7). Transposition probably was replicative since the *blaIblaZ* loci were not lost from the chromosome (Asheshov, 1969). pS1 was isolated following treatment of *S.aureus* PS80d.[pS1 parent] with EMS (Johnston and Dyke, 1971). If these two plasmids are indeed directly related, as their names suggest, there must have been two genetic rearrangements: (i) the precise excision (and loss) of Tn*552* from site E by recombination across the flanking 6 bp direct repeats (E1 and E2), thus regenerating pI9789; and (ii) the transposition of Tn*552* to a different target site upstream of *binR* (site F; Figures

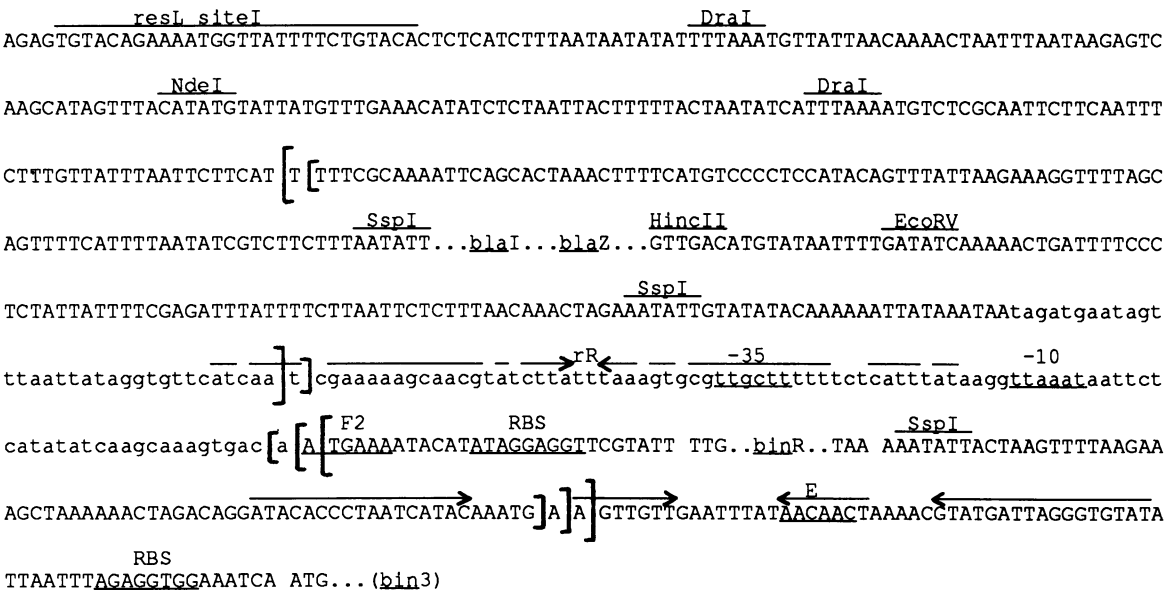

Fig. 8. The deletions which occurred within, and adjacent to, Tn*552* during the generation of pΔD from pS1 (Figure 7). Deleted sequences are enclosed in brackets and cannot be defined precisely due to repetition of nucleotides at the boundaries. The right terminal repeat of Tn*552* (TIR_R) is in lower case; other features of interest are defined in the legend to Figure 6A.

6A and 7). Functional β-lactamase genes are absent from the chromosome of PS80d. (the host for the construction of pS1), therefore the source of this latter element must have been another molecule of the pS1 parent plasmid. Subsequently the new plasmid, pS1, would have been segregated. Excision and transposition would have occurred in the order given above if Tn*552* has the property of transposition immunity. Precise excision (perhaps induced here by the EMS treatment) similarly has been suggested to account for a particular deletion of Tn*551*, a Gram-positive member of the Tn*3* family (Khan and Novick, 1980). It does not require transposition functions and is favoured by the presence of large inverted repeats immediately inside the direct repeats (Egner and Berg, 1981). Nevertheless, precise excision is a rare event and the possibility must be considered that pS1 arose directly from pI9789 (rather than *via* 'pS1 parent').

The pS1-derived plasmid, pΔD, may provide an example of a typical transposon induced 'adjacent deletion'. Thus, the smaller of the two identified deletions suffered by pS1 upon UV-irradiation probably extended from the exact outer boundary of TIR_R to a site within the inverted repeat downstream of *binR* (Figures 2C and 8). The deletion of *binR* explains why the invertible region of pΔD is 'frozen' in *E.coli*. The larger deletion, for which there is no obvious rationale based on DNA sequence, specifically removed the β-lactamase genes of Tn*552* (*blaIblaZ*), plus nearly one third of TIR_R. This deleted element might still be functional, although now cryptic.

Each of the flanking inverted repeats of Tn*552* contains a large, imperfect inverted repeat (designated 'rL' and 'rR'. Figure 6). The outer arm of rR contains the −35 component (TTGCTT) of the putative transcriptional promoter of *binR* in pS1, for which there is functional evidence (the complementation of inversion in *E.coli*). Thus, the insertion of Tn*552* may have activated *binR* in pS1 (and therefore *bin3* in pS1 parent), at least with respect to expression in *E.coli*. Additional potential promoter sequences can be identified, but only by homology (Figure 6B).

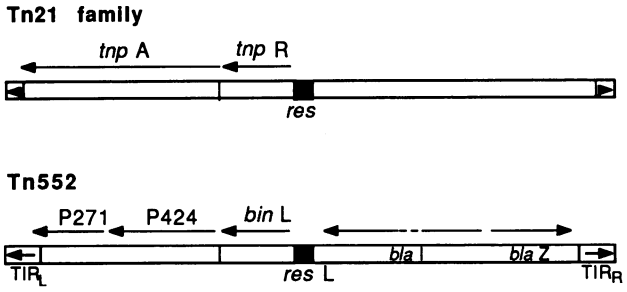

Fig. 9. Schematic comparison of the structures of members of the Tn*21* sub-group of Tn*3* family elements and the staphylococcal β-lactamase transposon, Tn*552* (not to scale). The *tnpA* and *tnpR* genes of Tn*3* elements encode the transposase and resolvase respectively; the rest of the transposon may carry additional genetic markers (e.g. ampicillin resistance in Tn*3*). The solid arrows indicate the directions and extents of ORFs so far determined for Tn*552*; the *blaI* region might encode one or two proteins (S.-J.Rowland, unpublished data).

Discussion

A new class of transposable elements

Tn*552* is a member of a novel group of transposable elements. The organization of genes in Tn*552* resembles that of members of the Tn*21* sub-group of Tn*3*-family transposons, which differs from the Tn*3* sub-group in the orientation of *res*−*tnpR* with respect to *tnp*A (Figure 9). According to this hypothesis, *binL* is actually the resolvase gene (*tnpR*) of Tn*552*, and one of the recombination sites of *inv* (*resL*) is its cointegrate resolution site. It is therefore logical to suggest that Tn*552* transposes replicatively via a cointegrate intermediate that is resolved by BinL ('Tn*552* resolvase'). The naturally occurring plasmid, pI9789, contains an isolated 'resolution system' (the origin of which is discussed below) that provides the other recombination site of *inv* (*resR*). The divergence of IR_L (*resL*−*binL*) and IR_R (*resR*−*binR*) apparently occurred under pressure to maintain intact both *binL* and *binR*. The possibility therefore

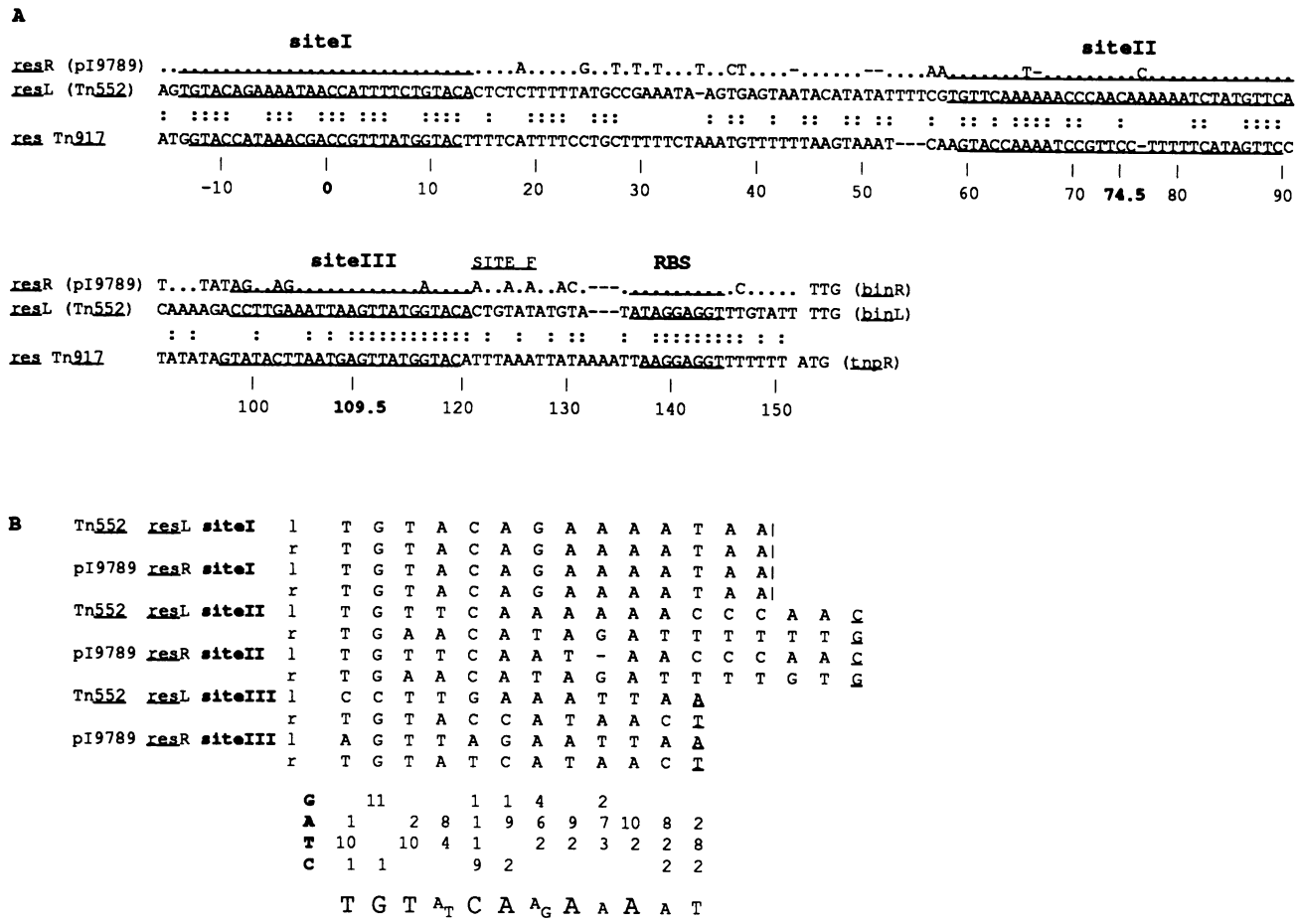

Fig. 10. (**A**) Alignment of the putative *res* regions of Tn*552* (*resL*), pI9789 (*resR*) and Tn*917*. Only nucleotides which differ from those in *resL* are shown for *resR*; nucleotide identities are represented by dots. Colons indicate identity between *resL* (Tn*552*) and *res* Tn*917*. Gaps have been introduced to maximise homology. Phosphodiester bond coordinates refer to Tn*552 resL* and are numbered from the centre of sub-site I; the coordinate at the centre of each site is in bold type. The initiaton codon of each resolvase gene and its probable ribosome binding site (RBS) are indicated. Tn*552* transposed to site F to generate pI9789blaIblaZseg-1 (pS1) (see Figures 6A and 7). (**B**) Alignment of the proposed half-site sequences comprising sub-sites I, II and III of *resL* (Tn*552*) and *resR* (pI9789), and a derived consensus. A centre of dyad symmetry is represented by a vertical line or an underlined nucleotide.

exists that, by analogy with the *E.coli* plasmid R46 (Dodd and Bennett, 1987), *resR*−*binR* also has a function: to contribute to the stable maintenance of pI9789.

Tn*552* encodes two putative proteins in the location expected for a transposase gene, 'P271' and 'P424' (271 and 424 amino acids respectively: S.-J.Rowland and K.G.H.Dyke, unpublished data; Figure 9). Although neither shows significant homology with known Tn*3*-family transposases (TnpA), it is notable that the P271 ORF extends into one of the terminal inverted repeats of Tn*552*; this is a common feature of transposase genes. Consistent with the absence of a *tnpA* homologue, the 116/121 bp terminal inverted repeats of Tn*552*, to which a transposase would bind (e.g. Wiater and Grindley, 1988), show no homology with the well-conserved, ~40 bp inverted repeats of Tn*3* elements, nor has an EMBL database search revealed homologies with the termini of other known transposons. Interestingly, each of the terminal inverted repeats of Tn*552* contains a large, imperfect inverted repeat, the arms of which might contain binding sites for a transposase. Nevertheless, the terminal dinucleotides of Tn*552*, 5′ TG . . CA 3′, are common among both prokaryotic and eukaryotic elements (e.g. phage Mu and retroelements). Furthermore, Tn*552* transposition generates 6 bp flanking direct repeats of target DNA (a specificity which is also shown by some retroviruses), rather than the 5 bp characteristic of Tn*3* elements (and Mu).

Resolution and inversion

Resolvases and DNA invertases execute the same basic strand exchange reaction; their specificities for directly and inversely repeated recombination sites respectively are believed to derive from topological constraints that depend on interactions between the proteins bound at accessory sites and at the crossover sites, in a supercoiled substrate (Boocock *et al.*, 1986; Johnson *et al.*, 1987; Benjamin and Cozzarelli, 1988; Kanaar *et al.*, 1988; Salvo and Grindley, 1988). A transposon *res* site comprises three separate resolvase binding sites spanning 100−150 bp upstream of the resolvase gene (*tnpR*). Strand exchange occurs at the centre of the most *tnpR*-distal, sub-site I (homologous to the crossover sites of the inversion systems; Figure 4) and is preceded by the formation of transient covalent bonds between a conserved serine (residues 9 and 10 in Gin and γδ resolvase, respectively) and the recessed 5′ phosphoryl ends produced by 2 bp staggered cleavage (Hatfull *et al.*, 1988; Klippel

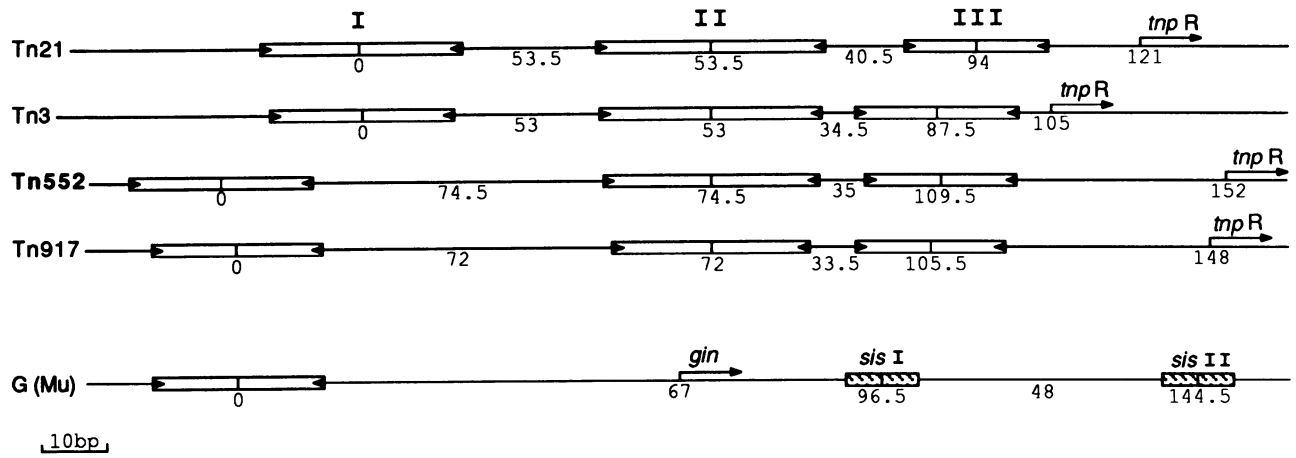

Fig. 11. The structures of the recombination sites of representative resolvases (Gram-negative: Tn*3* and Tn*21*, and putative Gram-positive: Tn*552* (*resL*) and Tn*917*) and the DNA invertase Gin. A typical *res* region contains three sub-sites (I, II and III; represented by boxes), each comprising an inverted repeat sequence and variable spacer (see Figure 4). The phosphodiester bond coordinates are defined from the centre of sub-site I and the distances between the centres of adjacent sub-sites are indicated. The recombination site of the invertible G segment that is adjacent to the DNA invertase gene (*gin*) is shown and is equivalent to *res* sub-site I. The beginning of each recombinase gene is indicated by a short arrow. A recombinational enhancer sequence lies within each DNA invertase gene and comprises two FIS binding sites, *sis* I and *sis* II [*gin* may actually contain a third site (C.Koch and R.Kahmann, unpublished, in Klippel *et al.*, 1988a)].

et al., 1988b; Mertens *et al.*, 1988). The accessory sub-sites II and III, essential for efficient resolution, appear to be analogous to the FIS-binding domains (*sis*I and *sis*II) of the recombinational enhancer sequences found within each of the DNA invertase genes *hin, gin, cin* and *pin*, but absent from *binL*, *binR* and *bin3* [a possible motif for *sis* is CAa/g-a--tGA-C (Johnson and Simon, 1987)]. Probable sub-sites II and III have been identified for *resL* (Tn*552*) and *resR* (pI9789); the homology between these *res* regions and the sequence upstream of another putative Gram-positive *tnpR*, that of Tn*917*, suggests that Tn*917 res* has a similar structure (Figure 10). Tn*917 res* is more homologous with *resL* than with *resR*; e.g. the right arms of the proposed sub-sites III of Tn*552* and Tn*917* are identical. It is particularly striking that site F, to which Tn*552* transposed to generate pS1, is precisely adjacent to sub-site III (Figures 6A and 10A). The *res* region of pIP404, which is very unusual [and also Gram-positive (Garnier *et al.*, 1987)], cannot be included in this alignment (apart from homology between the sub-sites *per se*). The structures of these sites are consistent with the model of Boocock *et al.* (1986, 1987) for the path of *res* DNA in a productive synaptic complex, with respect to the relative sizes and spacings of sites (Figure 11).

It has not yet been demonstrated that BinL ('Tn*552* resolvase') can recombine directly repeated *res* sites (i.e. perform cointegrate resolution). However, BinR (and very probably, therefore, BinL) can recombine inversely repeated *res* sites efficiently *in vivo*; indeed BinR was discovered as a result of its ability to complement inversion in *trans*. Dodd and Bennett (1986) and Altenbuchner and Schmitt (1983) also observed relatively efficient inversion, but only when *tnpR* (from Tn*802* or R46, and Tn*1721* respectively) was present in *cis*; in *trans*, inversion has been found to be very inefficient and dependent on the provision of *tnpR* at high copy number (Reed, 1981; Altenbuchner and Schmitt, 1983). These two factors (*tnpR* in *cis* and at high copy number) are not always sufficient for inversion (Chiang and Clowes, 1982). At present, explanations for resolvase-mediated inversion, which is known to require at least one sub-site in each *res* for Tn*3* resolvase (A.Bednarz and M.Boocock, personal communication), must be speculative. Firstly, the very strong preference for resolution over inversion or fusion (intermolecular recombination) shown by the resolvases of γδ, Tn*3* and Tn*21 in vitro* can be reduced by altering the reaction conditions (Boocock *et al.*, 1987), suggesting that perturbation of resolvase–resolvase interactions can influence topological specificity. A second possibility is a substrate topology that favours or allows the formation of a productive synaptic complex between inverted *res* sites [e.g. a linear substrate *in vitro* (Boocock *et al.*, 1986)]. Finally, inversion could be the indirect result of replicon fusion and deletion: for inversely repeated *res* sites designated A and B, on a circular substrate, recombination between *resA* on one molecule and *resB* on another is followed by recombination between either of the two pairs of directly repeated *res* sites on the product. Intermolecular events normally are proscribed for both resolvases and DNA invertases; nevertheless BinL(R) might be responsible for the unusually high frequency of intermolecular site-specific recombination events that occur between the closely related plasmids pI524 and pI258 in *S.aureus* (Murphy and Novick, 1980); pI258, like pI9789, probably contains a single *res* site and *bin* (L or R) gene.

The available data therefore imply that the putative Tn*552* resolvase, BinL (and BinR) can recombine both directly and inversely repeated intramolecular, and intermolecular, *res* sites *in vivo*. Significantly, a single amino acid substitution in either Gin (Klippel *et al.*, 1988a) or Cin (Haffter and Bickle, 1988) can cause FIS- and enhancer-independence, accompanied by the complete relaxation of topological specificity. Klippel *et al.* (1988a) have suggested that the protein–protein contacts between Gin-bound recombination sites are altered so that the enhancer–FIS complex is superfluous for efficient recombination; in the absence of 'interference' by FIS, mutant Gin is no longer restricted to inversely repeated sites. Analogously, a novel hybrid *cer* site, which differs from wild type by only 2 bp, supports unconstrained recombination that requires <50 bp (rather than 250 bp) sites and only one of the three chromosomal *xer* gene products used in specific recombination (Summers,

1989). It is likely that a detailed analysis of the recombinational proficiency of BinL(R) will be an important contribution to the study of 'resolvases' and 'DNA invertases'.

Staphyloccal β-lactamase transposons and plasmid evolution

With the exception of Tn*4201*, the *inv* region of which is somewhat divergent (Weber and Goering, 1988), the restriction maps of the *inv* and *bla* regions of Tn*552*, Tn*4002*, Tn*3852* and the 'Tn*4201*-like' element are almost indistinguishable, indicating that these β-lactamase transposons are very closely related (Lyon and Skurray, 1987). The accuracy of the map of the Tn*552 inv* region (Figure 1), which differs in minor details from those already published, has been confirmed by sequencing (S-J.Rowland, unpublished data). The β-lactamase elements associated with pS1 parent and pI524 are not identical and are believed to have diverged from that present on a plasmid equivalent to pS1 parent (pI9789::Tn*552*), the proposed evolutionary precursor of pI524 (Shalita *et al.*, 1980). The genetic structures and properties of Tn*4002*, Tn*3852* and the 'Tn*4201*-like' element are probably identical to those of Tn*552*.

The unique locus occupied by the *bla* genes on the *S.aureus* chromosome (Pattee, 1987) coincides with the site to which Tn*4201* has been shown to transpose from pCRG1600 (Weber and Goering, 1988). Furthermore, hybridization analysis indicates that Tn*4002* occupies exactly the same position on the chromosomes of six different strains (Gillespie *et al.*, 1988). The nature of this site is unknown. Tn*552* transposed to pI9789 both site- and orientation-specifically. Thus, at least two independent transpositions generated invertible segments apparently identical to that of pI524 (Murphy and Novick, 1979) and probably represent insertions within the large inverted repeat which separates *binR* and *bin3* (e.g. site E, Figure 6A). The origin and possible function of this repeat are intriguing. The second identified target site (F) is displaced to a location just upstream of *binR*, immediately adjacent to *resR*; insertion here generated the novel invertible region of pS1. It is interesting that each target site lies upstream of a resolvase-related gene. None of the other β-lactamase elements are associated with an invertible segment, consistent, for example, with the 'regional' specificity of Tn*4002* transposition to pSK1 (Gillespie *et al.*, 1988), a plasmid that lacks *resR*−*binR*. The resolution system of Tn*552* on pI9789 (*resR*−*binR*) is therefore a 'hotspot' for Tn*552* transposition.

Transposons mediate a variety of genetic rearrangements including duplications, deletions, inversions and translocations of neighbouring sequences. The structures of naturally occurring members of the staphylococcal α- and γ-family plasmids indicate that a variety of large-scale rearrangements involving the *inv* and *bla* regions have occurred during evolution (Shalita *et al.*, 1980). The γ-type *bla* region was derived from Tn*552* by unknown events that led to the deletion of *inv* (leaving one of the inverted repeats, *resR*−*binR* or *resL*−*binL*) and the inversion of *blaIblaZ*. Furthermore, most of the present members of the α-family were derived from a pI524-like plasmid by rearrangements (deletions, inversions, substitutions and deletion−inversions) focused at the ends of Tn*552*.

Further evidence that the site of insertion of Tn*552* in pI9789 has been a 'hotspot' for recombination during plasmid evolution is provided by *bin3*, which encodes a divergent member of the resolvase−invertase family and appears to have recombined at some stage with a homologous gene. Thus, a resolvase-like helix-turn-helix motif is encoded immediately downstream of, but not in frame with, *bin3*, which itself encodes a single more DNA invertase-like, helix 1 (Figure 5). This downstream 'silent' motif (there is no evidence from the DNA sequence that it might be expressed at the protein level) is most homologous with that of Tn*2501* resolvase, which, like Bin3, apparently diverged early during evolution (Figure 12; Stern and Kamp, 1989). The putative Bin3 protein lacks helix 2 and consequently should be unable to bind specifically to any *res* site; nevertheless, the integrity of the gene suggests that it does have a function that perhaps involves the potential N-terminal strand exchange activity. The site-specific recombinase encoded by the presumed precursor of *bin3* might have participated in recombination events during the evolution of α- and γ-family staphylococcal plasmids, for which Shalita *et al.* (1980) have presented a hypothetical pathway. Furthermore, the evidence that Tn*552* resolvase can recombine replicons (Murphy and Novick, 1980) suggests that it too might have played a significant role. Thus recombination between pI524- and pI258-like plasmids could explain the structure of pI836, an 'α plasmid with a γ *bla* determinant'. The early stages of this evolutionary pathway can now be updated: an ancestral plasmid specifying cadmium resistance acquired various other heavy metal and inorganic ion resistance genes and a site-specific recombinase gene (the precursor of *bin3*), perhaps introduced on a Tn*3*-like transposon. The Tn*552*-derived resolution system (*resR*−*binR*) carried by the present pI9789 could have arisen by the deletion of all other Tn*552* sequences subsequent to insertion upstream of *bin3*; homology with Tn*552* extends no further than the boundaries of *resR*−*binR* [contrast the *E.coli* plasmid R46, in which *res*−*tnpR* is associated with vestigial *tnpA* sequence (Dodd and Bennett, 1987)]. Transposition of Tn*552* to a site between *binR* and *bin3* then created a pI9789::Tn*552*-like plasmid, from which the α- and γ-families both could have evolved (Shalita *et al.*, 1980).

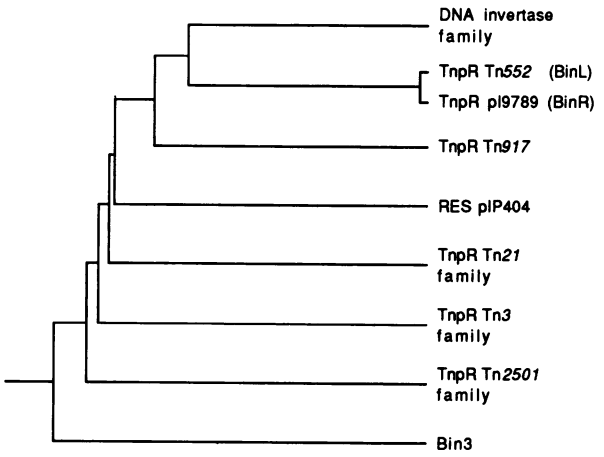

Fig. 12. Phylogenetic tree describing the evolution of the DNA invertase and resolvase superfamily. This was constructed by B.Stern (personal communication) using the computer program ALIGN and the Mutation Data Matrix of Dayhoff to analyse protein sequence similarity (as described by Stern and Kamp, 1989).

Tn552: a 'missing link'

The putative Tn*552* resolvase is a representative of the branch of resolvases from which the DNA invertases appear to have evolved (Figure 12; B.Stern, personal communication); this might be relevant to the question of topological specificity. In addition it has been suggested that an ancestral site-specific recombinase gene resided on a Tn*3*-like transposable element, from which evolved members of the Tn*3*-family and a Mu-like phage; the Hin, Gin, Cin and Pin invertible systems could have evolved from the latter (Kamp *et al.*, 1984; Kamp, 1987; Stern and Kamp, 1989). It is therefore extremely significant that Tn*552*, from which the staphylococcal *inv* has evolved, encodes a protein ('P271'; Figure 9) partially homologous with the Mu B transposition protein (S.-J.Rowland and K.G.H.Dyke, unpublished data). Furthermore, consistent with the facts that Tn*552* has the terminal dinucleotides 5′ TG . . CA 3′ and generates 6 bp flanking direct repeats of target DNA (both are characteristics of retroviruses), the suspected Tn*552* transposase ('P424'; Figure 9) shows excellent homology with part of the integrase (endonuclease) function of retroelements (S.-J.Rowland and K.G.H.Dyke, unpublished data).

Materials and methods

Bacterial strains, plasmids and growth conditions

S.aureus strains RN1753[pI524], PS80[pI9789], PS80d.[pI9789blaIblaZ], PS80d.[pI9789blaIblaZseg-1] and PS80d.[pI9789blaIblaZseg-1ΔD] (Johnston and Dyke, 1971) were the sources of plasmid DNA and were grown in CY medium [the latter two strains at 30°C since they harbour temperature-sensitive replication (*tsr*) plasmids]. 'PS80d.' is a β-lactamase negative derivative of PS80N (NCTC9789) (Johnston and Dyke, 1971).

E.coli strain MC1061 (Casadaban and Cohen, 1980) was the host for constructions that used the plasmid vectors pBR322, pRW33 [a pBR322-derived vector in which chloramphenicol has been substituted for ampicillin resistance (Mezes *et al.*, 1983)], pACYC184 (Chang and Cohen, 1978) and an ampicillin-sensitive derivative of pBR322 constructed by destroying the *Pst*I restriction site (this work). These vectors were used for the preliminary cloning and characterization (restriction mapping and complementation analysis) of the invertible and β-lactamase regions. Other *E.coli* strains used were WA3782 [a *pin*$^-$ strain (Iida, 1984) supplied by S.Iida] and HB101 [deleted for the chromosomal *fim* genes (Boyer and Roulland-Dussoix, 1976)]. *E.coli* JM107 (Yanisch-Perron *et al.*, 1985) was the host for restriction fragments sub-cloned for DNA sequencing into the 'phagemids' pTZ18R and pTZ19R. *E.coli* strains were grown in standard LB or 2 × TY media. Selective antibiotics were used at the following concentrations: ampicillin (50 μg/ml; 150 μg/ml for pTZ clones), chloramphenicol (10 μg/ml) and tetracycline (12.5 μg/ml).

DNA manipulations

Plasmid DNA was purified from *S.aureus* by the method of Novick and Bouanchaud (1971). Large-scale and rapid isolations by plasmid DNA from *E.coli* (by alkaline lysis preceded, when appropriate, by plasmid amplification) and all subsequent recombinant DNA techniques (cleavage of DNA with restriction endonucleases, agarose gel electrophoresis, filling in protruding 5′ ends with Klenow polymerase and ligation) were performed using standard procedures (Maniatis *et al.*, 1982). Enzymes were obtained from various suppliers (BRL, New England Biolabs, Amersham, Boehringer Mannheim) and used as recommended. *E.coli* JM107 was transformed using a modification of the method of Fort and Piggot (1984), and *E.coli* MC1061, subsequent to the preparation of frozen competent cells [cells at OD_{600} = 0.2 were harvested, washed in 0.5 vol ice-cold 50 mM $CaCl_2$, 15% glycerol (v/v), 10 mM Pipes pH 6.6, then resuspended in 0.05 vol of the same, divided into 150-μl aliquots and stored at −70°C].

DNA sequence analysis was performed on single-stranded plasmid DNA templates isolated from phage particles produced following superinfection of cells with the helper phage M13K07 (Vieira and Messing, 1987), using [α-^{35}S]dATP (Amersham) and Sequenase^{TM} as recommended by the supplier (United States Biochemical Corporation; Cambridge Bioscience).

Complementation analysis

The 6.2 kb pΔD *Eco*RI fragment containing *inv* (*Eco*C:pΔD) was cloned into the *Eco*RI site of pACYC184, the host *E.coli* MC1061. Aliquots of frozen competent cells were prepared for supertransformation by individual sub-fragments of *Eco*A:pS1 cloned into suitable sites in the tetracycline resistance gene of the compatible vector, pBR322. Selection was therefore for ampicillin (over tetracycline) resistance. Controls (e.g. the vector plasmid) were included. Complementation was assayed by restriction analysis of plasmid DNA isolated after the growth of single transformants in liquid medium (without selection). The 'titration effect' described in the text was observed when a similar procedure was followed for MC1061[pRW33/*Eco*C:pΔD] and pACYC184/*Eco*A:pS1. Aliquots of competent *E.coli* MC1061 harbouring the smallest fragment found to promote inversion in *trans* (the 1.1 kb pS1 *Eco*RV−*Pvu*I fragment cloned into the *Ssp*I−*Pvu*I sites of pBR322) were then supertransformed separately with pACYC184-based clones containing 'frozen' *inv* regions (both (+) and (−) orientations) from pΔD (the 5.0 kb *Eco*RV−*Pvu*I fragment), pS1 (the 6.2 kb *Eco*RV fragment) and pS1 parent (the 10 kb *Eco*RI fragment), selecting for chloramphenicol or ampicillin resistance, as appropriate. Complementation was assayed as described above.

Acknowledgements

We are very grateful to B.Stern for constructing the phylogenetic tree, to M.Boocock for valuable discussion and to D.Sherratt, E.Murphy and A.Glasgow for communicating manuscripts prior to publication. This work was supported by a Science and Engineering Research Council studentship for training in research methods to S.-J.R.

References

Abraham,J.M., Freitag,C.S., Clements,J.R. and Eisenstein,B.I. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5724−5727.
Altenbuchner,J. and Schmitt,R. (1983) *Mol. Gen. Genet.*, **190**, 300−308.
Anderson,J.E., Ptashne,M. and Harrison,S.C. (1987) *Nature*, **326**, 846−852.
Asheshov,E.H. (1969) *J. Gen. Microbiol.*, **59**, 289−301.
Benjamin,H.W. and Cozzarelli,N.R. (1988) *EMBO J.*, **7**, 1897−1905.
Boocock,M.R., Brown,J.L. and Sherratt,D.J. (1986) *Biochem. Soc. Trans.*, **14**, 214−216.
Boocock,M.R., Brown,J.L. and Sherratt,D.J. (1987) In Kelley,T.J. and McMacken,R. (eds), *DNA Replication and Recombination*. Alan R.Liss, Inc., New York, pp. 703−718.
Boyer,H.W. and Roulland-Dussoix,D. (1976) *J. Mol. Biol.*, **41**, 459−472.
Bruist,M.F., Horvath,S.J., Hood,L.E. and Steitz,T.A. (1987) *Science*, **235**, 777−780.
Casadaban,M.J. and Cohen,S.N. (1980) *J. Mol. Biol.*, **138**, 179−207.
Chang,A.C.Y. and Cohen,S.N. (1978) *J. Bacteriol.*, **134**, 1141−1156.
Chiang,S.-J. and Cowes,R.C. (1982) *Mol. Gen. Genet.*, **185**, 169−175.
Dodd,H.M. and Bennett,P.M. (1986) *J. Gen. Microbiol.*, **132**, 1009−1020.
Dodd,H.M. and Bennett,P.M. (1987) *J. Gen. Microbiol.*, **133**, 2031−2039.
Dyke,K.G.H. (1979) In Hamilton-Miller,J.M.T. and Smith,J.T. (eds), *Beta-lactamases*. Academic Press, London, pp. 291−311.
Egner,C. and Berg,D.E. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 459−463.
Fort,P. and Piggot,P.J. (1984) *J. Gen. Microbiol*, **130**, 2147−2153.
Garnier,T., Saurin,W. and Cole,S.T. (1987) *Mol. Microbiol.*, **1**, 371−376.
Gillespie,M.T., Lyon,B.R. and Skurray,R.A. (1988) *J. Gen. Microbiol.*, **134**, 2857−2866.
Glasgow,A.C., Hughes,K.T. and Simon,M.T. (1989) In Berg,D. and Howe,M. (eds), *Mobile DNA*. American Society of Microbiology, Washington, DC, pp. 637−659.
Grindley,N.D.F. and Reed,R.R. (1985) *Annu. Rev. Biochem.*, **54**, 863−896.
Haffter,P. and Bickle,T.A. (1988) *EMBO J.*, **7**, 3991−3996.
Hatfull,G.F., Salvo,J.J., Falvey,E.E., Rimphanitchayakit,V. and Grindley,N.D.F. (1988) In Kingsman,A.J., Chater,K.F. and Kingsman,S.M. (eds), *Transposition*. Society for General Microbiology Symposium Vol. 43, Cambridge University Press, Cambridge, pp. 149−181.
Hiestand-Nauer,R. and Iida,S. (1983) *EMBO J.*, **2**, 1733−1740.
Iida,S. (1984) *Virology*, **134**, 421−434.
Iida,S. (1986) *FEMS Microbiol. Lett.*, **17**, 183−187.
Johnson,R.C. and Simon,M.I. (1987) *Trends Genet.*, **3**, 262−267.
Johnson,R.C., Glasgow,A.C. and Simon,M.I. (1987) *Nature*, **329**, 462−465.
Johnston,L.H. and Dyke,K.G.H. (1971) *J. Bacteriol.*, **107**, 68−73.

Kamp,D. (1987) In Symonds,N., Toussaint,A., van de Putte,P. and Howe,M. (eds), *Phage Mu*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 259–269.
Kamp,D., Kardas,E., Ritthaler,W., Sandulache,R., Schmucker,R. and Stern,B. (1984) *Cold Spring Harbour Symp. Quant. Biol.*, **43**, 301–311.
Kanaar,R., van de Putte,P. and Cozzarelli,N.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 752–756.
Khan,S.A. and Novick,R.P. (1980) *Plasmid*, **4**, 148–154.
Klippel,A., Cloppenborg,K. and Kahmann,R. (1988a) *EMBO J.*, **7**, 3983–3989.
Klippel,A., Mertens,G., Patschinsky,T. and Kahmann,R. (1988b) *EMBO J.*, **7**, 1229–1237.
Lyon,B.R. and Skurray,R.A. (1987) *Microbiol. Rev.*, **51**, 88–134.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Mertens,G., Klippel,A., Fuss,H., Blocker,H., Frank,R. and Kahmann,R. (1988) *EMBO J.*, **7**, 1219–1227.
Mezes,P.S.F., Wang,W., Yeh,E.C.H. and Lampen,J.O. (1983) *J. Biol. Chem.*, **258**, 11211–11218.
Murphy,E. (1988) In Kingsman,A.J., Chater,K.F. and Kingsman,S.M. (eds), *Transposition*. Society for General Microbiology Symposium Vol. 43, Cambridge University Press, Cambridge, pp. 59–89.
Murphy,E. and Novick,R.P. (1979) *Mol. Gen. Genet.*, **175**, 19–30.
Murphy,E. and Novick,R.P. (1980) *J. Bacteriol.*, **141**, 316–326.
Novick,R.P. and Bouanchaud,D. (1971) *Ann. New York Acad. Sci.*, **182**, 279–294.
Pabo,C.O. and Sauer,R.T. (1984) *Annu. Rev. Biochem.*, **53**, 293–321.
Pattee,P.A. (1987) In S.J.O'Brien (ed.) *Genetic Maps 1987: a Compilation of Linkage and Restriction Maps of Genetically Studied Organisms*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Vol. 4.
Plasterk,R.H.A. and van de Putte,P. (1984) *Biochim. Biophys. Acta*, **782**, 111–119.
Plasterk,R.H.A. and van de Putte,P. (1985) *EMBO J.*, **4**, 237–242.
Reed,R.R. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3428–3432.
Rowland,S.-J. and Dyke,K.G.H. (1988) *FEMS Microbiol. Lett.*, **50**, 253–258.
Salvo,J.J. and Grindley,N.D.F. (1988) *EMBO J.*, **7**, 3609–3616.
Shalita,Z., Murphy,E. and Novick,R.P. (1980) *Plasmid*, **3**, 291–311.
Sherratt,D. (1989) In Berg,D. and Howe,M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 163–184.
Stern,B. and Kamp,D. (1989) *Protein Seq. Data Anal.*, in press.
Summers,D.K. (1989) *EMBO J.*, **8**, 309–315.
Vieira,J. and Messing,J. (1987) *Methods Enzymol.*, **153**, 3–11.
Weber,D.A. and Goering,R.V. (1988) *Antimicrob. Agents Chemother.*, **32**, 1164–1169.
Wiater,L.A. and Grindley,N.D.F. (1988) *EMBO J.*, **7**, 1907–1911.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.

Received on March 6, 1989; revised on May 22, 1989

Note added

We have now demonstrated that BinR can resolve an artificial co-integrate *in vivo*.