The EMBO Journal vol.6 no.9 pp.2849–2853, 1987

Construction of hybrid Tn*501*/Tn*21* transposases *in vivo*: identification of a region of transposase conferring specificity of recognition of the 38-bp terminal inverted repeats

L.R.Evans and N.L.Brown[1]

Department of Biochemistry and Unit of Molecular Genetics, University of Bristol, University Walk, Bristol BS8 1TD, UK

[1]Present address: Department of Genetics, University of Melbourne, Parkville, Victoria 3052, Australia

Communicated by D.Sherratt

In order to study the transposase enzymes of Class II prokaryotic transposable elements, we have constructed genes encoding hybrid transposase proteins. This was done by recombination *in vivo* between the *tnpA* genes of transposons Tn*501* and Tn*21*. These hybrid genes can complement *in trans* a transposition-defective mutant of Tn*501*. The structures of the products of this complementation indicate whether the specificity of the hybrid transposase in recognising the 38 bp terminal inverted repeats is that of Tn*501* or that of Tn*21*. The determinant of this specificity is in the N-terminal region of the transposase protein, between amino acids 28 and 216. The predicted amino acid sequences so far determined of transposases from the Class II family reveal an area of homology in this region.

Key words: class II prokaryotic transposons/DNA-protein recognition/hybrid proteins/illegitimate recombination/protein engineering

Introduction

There is a large class of prokaryotic transposons related to the transposon Tn*3* (Kleckner, 1981; Heffron, 1983; Grindley and Reed, 1985). These transposons have 38-bp inverted terminal repeats, which show some homology to those of Tn*3*, and are flanked by 5-bp direct repeats generated from the recipient DNA during the transposition process. The major mechanism of transposition of this class occurs in two steps: the formation of a co-integrate between donor and recipient replicons, catalysed by the *tnpA* gene product (transposase); and resolution of the cointegrate to generate two replicons each carrying a copy of the transposon, catalyzed by the *tnpR* gene product (resolvase). There are at least two subgroups of the Tn*3* family (Heffron, 1983; Grindley and Reed, 1985); one represented by Tn*3* itself and Tn*1000* (gamma-delta), in which the *tnpR* and *tnpA* genes are transcribed divergently, and another, represented by Tn*501*, Tn*21* and Tn*1721*, in which these genes are transcribed in the same direction and in the order *tnpR–tnpA*. The two-step transposition process occurs in both subgroups.

The reaction catalysed by the transposase from bacteriophage Mu has been subject to detailed *in vitro* study (Cragie *et al.*, 1985; Craigie and Mizuuchi, 1985). However, co-integrate formation by Tn*3*-like transposons is a different process, and has not proved very amenable to *in vitro* analysis. Only binding studies of Tn*3* transposase to Tn*3* inverted repeats have been reported so far (Wishart *et al.*, 1985). The reaction catalysed by transposase is likely to be complex; in a minimal model of this reaction (Shapiro, 1979), the enzyme must recognise the 38-bp inverted repeats, catalyse a specific cleavage or nicking reaction at each inverted repeat, catalyse a less-specific 5-bp staggered cleavage of the recipient DNA and facilitate religation of these termini, resulting in replicon fusion. Host replication functions can then convert the initial product into a fully double-stranded co-integrate, containing two copies of the transposon in direct repeat.

Studies of co-integrate formation *in vivo* have shown that the *tnpA* gene of one transposon may complement *in trans* the transposition of TnpA$^-$ mutants of a related transposon, provided the mutants have intact inverted repeats (Gill *et al.*, 1978; Grinsted *et al.*, 1982). These complementation events occur only between certain transposons (Grinsted *et al.*, 1982), and individual transposases show different specificity of recognition of terminal inverted repeats. Previous studies (Grinsted and Brown, 1984) have shown that the defective transposon Tn*819* (a Tn*501* derivative; Figure 1a) can form cointegrates when complemented *in trans* by Tn*501* or by the related transposons Tn*1721* or Tn*21*, but not when complemented by Tn*3*. Furthermore, complemen-

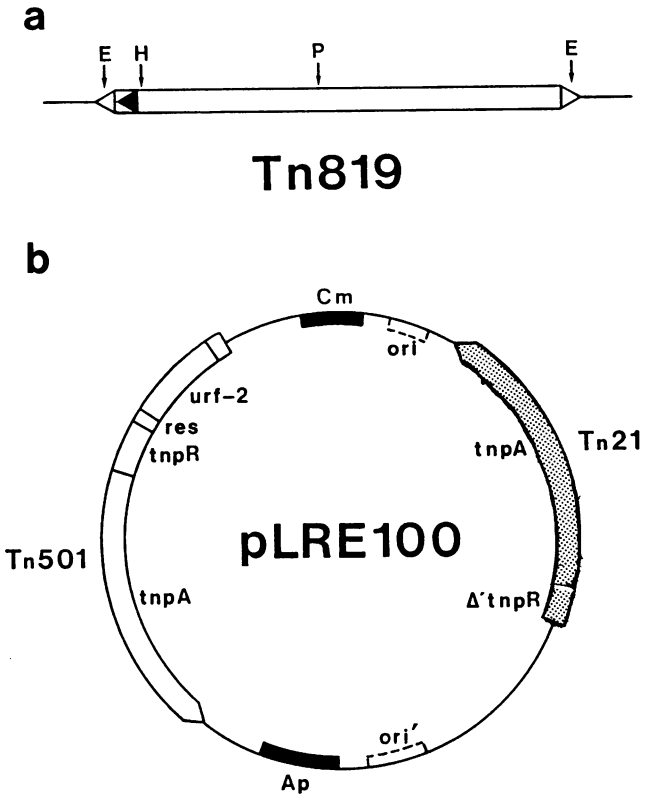

Fig. 1. (**a**) The structure of transposon Tn*819*. This is a Mer$^-$, Res$^-$, TnpR$^-$, TnpA$^-$ deletion mutant of Tn*501*, in which the central 6 kb of the transposon have been deleted and replaced by a *Pst*I linker (P) (Grinsted and Brown, 1984). The terminal 38 bp inverted repeats (open triangles) contain single *Eco*RI sites (E); the internal Tn*21* IR sequence (closed triangle) is at nucleotides 81–118, and contains no *Eco*RI site. The single *Hin*dIII site is marked (H). (**b**) The structure of plasmid pLRE100. Cm = *cat* gene; Ap = *bla* gene; ori and ori' identify the replication functions of pACYC184 and pBR325 respectively.

Table I. The specificity of *trans*-complementation by hybrid transposases

Hybrid transposase[a]	Position of cross-over[b]	Inverted repeat used in complementation[c]
C1-41	70– 80	7/7 Tn*21*
C1-31	676– 683	3/3 Tn*501*
C2-02	700– 852	7/7 Tn*501*
C8-09	700– 852	3/3 Tn*501*
C2-08	1182–1468	8/8 Tn*501*
C1-32	1182–1468	3/3 Tn*501*
C3-08	1561–2983	6/6 Tn*501*
C4-05	None (Tn*501*)	21/21 Tn*501*
A5-25	−335– −150	16/16 Tn*501*
A6-32	700–1182	6/6 Tn*21*
A7-31	700–1182	12/13 Tn*21*
A7-14	1182–1561	9/9 Tn*21*
A5-17	2388–2983	24/24 Tn*21*
A2-11	None (Tn*21*)	11/11 Tn*21*

[a]The C-series of hybrid transposases are Tn*501*-proximal, and were isolated as Cm^r,Ap^s recombinant plasmids. These hybrid transposases are on pACYC184-derived replicons. The A-series of hybrid transposases are Tn*21*-proximal, and were isolated as Ap^r,Cm^s recombinants. These are on pBR325-derived replicons.

[b]The positions of the recombination sites are given relative to the start of the gene at position 1, and are between discriminatory restriction sites at the stated positions (except for C1-41 and C1-31 which have been sequenced, and refer to the limits of the homologous sequence within which the cross-over occurred). A negative number refers to a cross-over before the gene. The gene (including the termination codon) extends to position 2967 in both Tn*501* and Tn*21*. Note that in C3-08 and A5-17, the recombination site could be just beyond the gene. The restriction enzymes corresponding to these discriminatory sites are: in Tn*501*, *Eco*RI 2983, *Hgi*AI 852, *Not*I 2388, *Pvu*II 700, 1561, *Rsa*I −335, −150; in Tn*21* *Eco*RI 1182, *Rsa*I 1468. Not every discriminatory site could be used with each hybrid gene due to ambiguity in the interpretations of the digest patterns.

[c]The numbers refer to the number of transposition products (resolved co-integrates) isolated and the fraction of these using the named inverted repeat. In the case of complementation by hybrid A7-31 one of the products was aberrant, and not the result of a simple transposition event.

tation by Tn*21* involved recognition of an internal sequence within Tn*819*, which was identical with a Tn*21* inverted repeat, and this internal inverted repeat sequence became the new terminus of the transposed element (Grinsted and Brown, 1984). When Tn*501* was used for *trans*-complementation, the Tn*501* terminal repeats present on Tn*819* were used.

Here we describe the construction of genes encoding hybrid transposase proteins by recombination *in vivo*, and the use of these genes to complement the transposition of Tn*819*. Correlation of the primary structures of the hybrid proteins with their specificity of recognition of the terminal (Tn*501*) or internal (Tn*21*-like) inverted repeats of Tn*819* has allowed us to identify a region of transposase conferring this recognition specificity.

Results

Construction of the double replicon pLRE100 and isolation of recombinant transposase genes

We have constructed a plasmid, pLRE100 (Figure 1b), containing the *tnpA* genes of Tn*501* and Tn*21* which are 74% identical in nucleotide sequence (Brown *et al.*, 1985; Ward and Grinsted, 1987). The genes are in direct repeat and are separated from one another at each end by non-homologous replication origins and resistance markers. The plasmid is stable in *E. coli* $RecA^-$ strains, but in $RecA^+$ strains genetic recombination can occur between the *tnpA* genes. We have transformed *E. coli* CSH26*recA*$^-$ with plasmid DNA prepared from the equivalent $RecA^+$ strain, and screened for segregation of the Cm^r and Ap^r markers, thus isolating plasmids in which recombination has occurred.

Restriction analysis showed that in the majority (~75%) of these plasmids, recombination had occurred between the *tnpA* genes in the C-terminal half of the coding sequence. The C-terminal region shows the greatest homology between Tn*21* and Tn*501* (Ward and Grinsted, 1987), and also between Tn*3* and Tn*501* (Brown *et al.*, 1985). However, some 20% of the recombinants were generated by cross-overs within the first half of the *tnpA* gene. The remaining fraction (~5%) of the cross-overs was between the residual *tnpR* gene sequences, upstream of the *tnpA* gene. These mapping data were initially crude, allowing the recombinants to be sorted by approximate position of the recombination site. Once complementation data had been obtained, finer mapping of important recombinants was done. Data on the mapping of some of the sites of crossover is shown in Table I and Figure 2.

Complementation of transposition of Tn819

Several different recombinants were selected to represent different classes of crossover, and assayed for their ability to complement transposition of Tn*819* by the 'mate-out' assay (see Materials and methods). These experiments (Table I) showed that the hybrid *tnpA* genes could complement co-integrate formation between a donor plasmid carrying Tn*819* and the conjugative plasmid R388 (Datta and Hedges, 1972). The frequency of complementation varied between constructs and between experiments, and ranged from 10^{-4} to 10^{-8}. The reasons for this variation are not known, but no specific efforts to express the hybrid transposases at constant and consistent levels were made. It is possible that some of the hybrid proteins were intrinsically unstable, or only poorly functional.

The specificity of use of the terminal (Tn*501*) or internal (Tn*21*) inverted repeats of Tn*819* by the complementing transposase was examined by *Eco*RI restriction analysis of the products of transposition. Both the co-integrate between the donor plasmid and R388 and the resolved R388 recipient were examined. In Tn*819*, the terminal inverted repeat (of Tn*501*) contains an *Eco*RI site (Brown *et al.*, 1980), whereas the internal, Tn*21*-like inverted repeat does not (Grinsted and Brown, 1984). Thus, if the Tn*501* inverted repeat was used, the co-integrate molecule would have six *Eco*RI sites (one from the donor plasmid backbone, one from R388 and four from two copies of the transposon), and the resolved product would have three sites. However, if the Tn*21* inverted repeat is used, the co-integrate would have five *Eco*RI sites (one donor backbone, one R388, two from the 'parental' copy of Tn*819* and one from the new Tn*819* derivative with a Tn*21* terminus), and the resolved product would have two sites. In practice, due to the variable sizes of the fragments with the target site on R388, it is more reliable to look for the presence or absence of the 2.4 kb *Eco*RI fragment from Tn*819* in the resolved product.

The results of one such series of experiments are shown in Figure 2. These data show that the transposases C1-31, C8-09, C1-32, C2-07, A5-25 and C4-05 have recognised the Tn*501* inverted repeat, as the Tn*819* *Eco*RI fragment is present (arrowed) in the resolved products of complementation by these hybrids. The analysis of other, separately isolated products of complementation by these hybrids shows that the 2.4-kb internal *Eco*RI fragment of Tn*819* is reproducibly present (data not shown).

The 2.4-kb fragment cannot be seen in the resolved products of complementation by C1-41, A2-11, A6-32, A7-14 and A5-17,

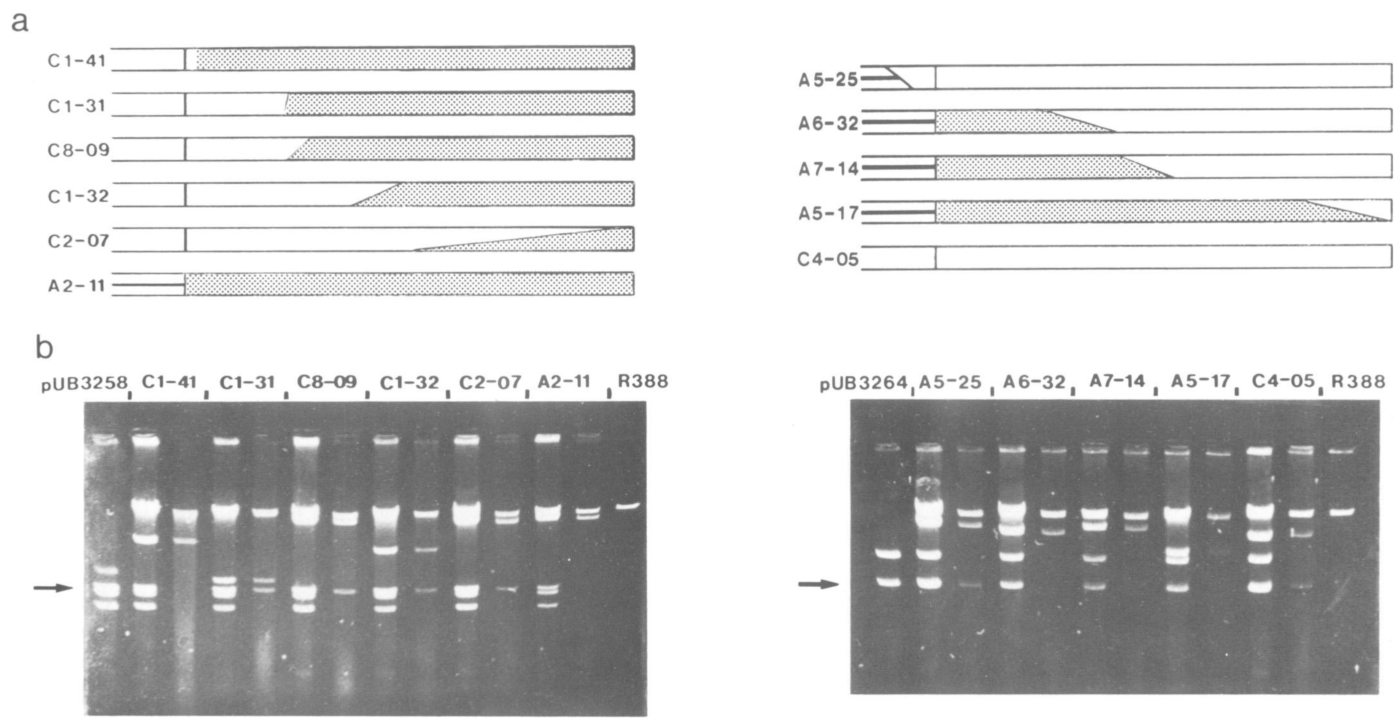

Fig. 2. (**a**) Structures of some of the hybrid *tnpA* genes used to complement transposition of Tn*819*. The sites of cross-over between sequences from Tn*501* (unshaded) and Tn*21* (shaded or with horizontal line) were mapped between sites at the ends of the diagonal lines. The positions of these sites are given in Table I. The vertical lines mark the extent of the *tnpA* gene. (**b**) Gel electrophoretic analysis of the products of transposition. The samples are marked according to the identity of the complementing transposase; the left of each pair of digests is of the co-integrate, and the right is of the resolved R388 derivative. On each gel the samples are flanked by markers of the donor plasmid (pUB3258 or pUB3264) and R388. The arrow marks the position of the 2.4 kb *Eco*RI fragment of Tn*819*. Note that there is a very small *Eco*RI fragment in pUB3264 and its co-integrates that is not seen on these gels. Due to the poor yield of the resolved product complemented by A5-17 in this particular experiment, the smaller fragment (~5 kb) is very faint.

```
                     .        .   C1-41 .       .       .       .       .       .       .  97
p501a   MPRRLILSATERGTLLALPESQDD-LIRYYTFNDSDLSLIRQRRGDANRLGFAVQLCLLRYPGYALGTDSELPEPVILWVAKQVQTDPA-SWTKYGERD
p21a    MPRRSILSAAERESLLALPDSKDD-LIRHYTFNDTDLSIIRQRRGPANRLGFAVQLCYLRFPGVILGVDELPFPPLLKPVADQPKV-GVESWNEYGQRE
p3a     MPVDFLTTEQVESYGRFTGEPDELQLARYFHLDEADKEFIGKSRGDHNRLGIALQIGCVRFLGTFLTDMNHIPSGVRHFTARQLGIRDITVLAEYGQRE
p917a  MAMKRILTTSQREQLLSVDHLSEEDFKA-YFSFSDYDLEVINQHRGKVNKLGFAIQLCLARYPGCSLSNWPIKSTRLTSYVSRQLHLDAID-LNSYDHR-
p2501a  MPRRQILSSEEQERLLVIPDDEIILTRMCF-LNEQDIALINKHRRPANRLGFAVLLCYLRGPGFIPDKSNVPHSSVISRIASRLKL-QPDLWPEYASRE
             : :                          :   : *   *   *   *:** *: :  :*  *          :            * *

          .       .       .       .       .       .       .       .       .       .  196
p501a   VTRREHAQELRTYLQLAPFGLSD-FRALVRELTELAQQTDKGLLLAGQALESLRQKRRILPALSVIDRACSEAIARANRRVYRALVEPLTDSHRAKLDEL
p21a    QTRREHLSELQTVFGFRPFTMSH-YRQAVQMLTELAMQTDKGIVLASALIGHLRRQSVILPALNAVERASAEAITRANRRIYDALAEPLADAHRRRLDDL
p3a     NTRREHAALIRQHYQYREFAWPWTFRLTRLLYTRSWISNERPGLLFDLATGWLMQHRIILPGATTLTRLISEVREKATLRLWNKLALIPSAEQRSQLEML
p917a   NTRANHFNEILEVFNYHRFGSANTQKQLIEYLIELALENDDSIYLMKKTIDFLTRKRIIFPSIATLEDIISRCRDKAEKNLFSILLCSLTDIQIEKLESL
p2501a  QTRWEHLTELYRYLKLSPFMPIM-QKDCIRHLHPYAMRTDKGFMLAEEMLNWLHSNNVIFPSVDVIERTLAEAVTLADRAYFTHLPHNWKSSINQHSTVC
         ** *   :       *      :           :  *     *   * * : :         *    *

          .       .       .   C1-31       .       .       .       .       .       .  296
p501a   LKLKAGSSITWLTWLRQAPLKPNSRHMLEHIERLKTFQLVDLPEVLGRHIHQNRLLKLAREGGQMTPKDLGKFEPQRRYATLAAVVLESTATVIDELVDL
p21a    LKRRDNGKTTWLAWLRQSPAKPNSRHMLEHIERLKAWQALDLPTGIERLVHQNRLLKIAREGGQMTPADLAKFEPQRRYATLVALATEGMATVTDEIIDL
p3a     LGPTDCSRLSLLESLKKGPVTISGPAFNEAIERWKTLNDFGLHAENLSTLPAVRLKNLARYAGMTSVFNIARMSPQKRMAVLVAFVLAWETLALDDALDV
p917a   FQIYEETKITKLAWLKDIPGKANPESFMSICKKVEVIASMGLGTINVSHINRNRFLQLARLGENYDAYDFSRFELEKRYSLLIAFLVNHHQYLIDQLIEI
p2501a  SNQRVNKL
              : *  *:  *          :      *      :  *  :**         :   :*  *:*    : *  :::
```

Fig. 3. Parsimonious alignment of the predicted N-terminal sequences of transposases from members of the Class II family of transposons. The sequences are from Tn*501* (Brown *et al.*, 1985), Tn*3* (Heffron *et al.*, 1979), Tn*2501* (Michelis *et al.*, 1987), Tn*917* (Shaw and Clewell, 1985) and Tn*21* (Ward and Grinsted, 1987). Each transposase is indicated by an acronym containing the transposon number. Hyphens are padding to assist the alignment of identical residues, which are marked with *; highly conserved residues are marked with : The numbering refers to the Tn*501* amino acids. The sites of cross-over from Tn*501* sequence to that of Tn*21* in the C1-31 and C1-41 hybrid transposases are marked with overlines. (Tn*917* has been shown to contain a *res* site and *tnpR* gene homologous to those of Tn*501*, and the published DNA sequence shows two open reading frames with homology to the Tn*501* *tnpA* gene; Shaw and Clewell, 1985. The predicted product of the first of these is aligned here.)

```
IR-21-L      GGGGGCACCTCAGAAAACGGAAAATAAAGCACGCTAAG
IR-21-R      GGGGTCGTCTCAGAAAACGGAAAATAAAGCACGCTAAG
IR-501-L     GGGGGAACCGCAGAATTCGGAAAAAATCGTACGCTAAG
IR-501-R     GGGGGGCTCGCAGAATTCGGAAAAAATCGTACGCTAAG
                ***  *     **        * ** *
```

Fig. 4. Alignment of the inverted repeat sequences of Tn*501* and Tn*21*, showing the differences between them. The Tn*21* inverted repeat which is found at position 81–118 of Tn*501* has a sequence identical to the right inverted repeat of Tn*21* (21R). Thus, discrimination between Tn*501* and Tn*21* inverted repeats during complementation of transposition of Tn*819* is discrimination between the sequences marked IR-501-L and IR-21-R, and presumably involves one or more of the positions marked with *.

nor is it seen when other resolved products of complementation by these hybrids are examined. More detailed restriction analysis of these resolved products (data not shown) indicates that they contain the Tn*819* *Hin*dIII site (Figure 1a). In all cases tested the transposon terminus maps close to (and presumably at) the outside end of the Tn*21* inverted repeat, indicating that these hybrid transposases recognise the Tn*21* inverted repeat.

Therefore, the specificity of recognition of the inverted repeat is determined by the first component of the hybrid *tnpA* gene (except for hybrid C1-41, see below). Thus, the N-terminal region of transposase determines this specificity. Within the limits of our assays (Table I), we have detected no mixed or reduced specificity in the hybrid transposases. Only in one case, hybrid A7-31, have we found a product that was not the product of a normal transposition event.

The two Tn*501*-proximal hybrid *tnpA* genes C1-31 and C1-41 (Figure 2, Table I) were of particular interest, as C1-31 has Tn*21* specificity and C1-41 has Tn*501* specificity (Figure 2). Thus, their recombination sites must lie to either side of the determinant for discriminatory recognition of the inverted repeats. These recombination sites were the closest to the start of gene of those detected; the recombination sites on C1-41 and C1-31 were sequenced and found respectively to be between nucleotides 70 and 80 and between nucleotides 676 and 683 of the *tnpA* genes (Figure 2, Table I). The limits on the accuracy of mapping the cross-overs are determined by the extent of identical sequences in the two parental *tnpA* genes. Thus, the discriminatory specificity determinant must be encoded between nucleotides 80 and 676. As the differences in nucleotide sequence between positions 648 and 676 are third position changes, with no effect on the encoded amino acid sequence (see Figure 3), the specificity determinant must be between amino acids 28 and 216 of the transposase. Sequences outside this region may also be involved in recognition of the inverted repeats.

Discussion

The data presented above show that one region of the transposase protein can provide discriminatory recognition of the Tn*501* and Tn*21* inverted repeat sequences. The limits of this region are defined by the sites of recombination in the hybrid genes C1-41 and C1-31; thus, the determinants of this recognition specificity lie between amino acids 28 and 216 of the Tn*501* and Tn*21* transposase proteins. The predicted primary structures of these two proteins show 53% identity of amino acid sequence in this region, and it is not possible to assign specific groups of residues as potentially important in specificity. We do not know whether this is the only region of the protein molecule involved in recognition of the inverted repeats. Indeed, with such a large (38-bp) recognition site on the DNA, different domains of the protein may well contribute to the recognition process. The N-terminal region may provide only discriminatory recognition, and not participate in the major interactions between the transposase and the inverted repeat.

Tn*21* transposase will act on Tn*21*, Tn*1721* and Tn*501* inverted repeats, but Tn*501* transposase will only act on Tn*501* and Tn*1721* inverted repeats (Grinsted *et al.*, 1982). Tn*21* shows a broader specificity of recognition than does Tn*501* (J.Grinsted, unpublished data). Replacement of the N-terminal sequences of Tn*501* transposase by those of Tn*21* alters the specificity of the transposase, possibly by reducing the constraints on recognition specificity imposed by this region of the Tn*501* transposase. However, the difference in Tn*501* and Tn*21* specificity is not simply a 'tight' and 'loose' specificity, as in all our experiments (Table I; Grinsted and Brown, 1984) a preference was shown for one or the other type of inverted repeat.

The simplest, but not the only, explanation of the discriminatory recognition of the Tn*501* and Tn*21* inverted repeats by the hybrid transposases is that the N-terminal region of transposase directly interacts with the base pairs in the inverted repeat. Comparison of the predicted sequences of the N-terminal region from several transposases from the Tn*3* family (Figure 3) reveals a limited amino acid sequence homology. There is a small region, residues 33–65, in which ~50% of the residues are identical or closely homologous between the five transposase sequences shown. We can speculate that this small region of homology may define sequences involved in recognition of the 38-bp inverted repeats, which themselves show some sequence conservation (Figure 4).

The known DNA binding domains of other proteins (mainly repressors) contain helix–turn–helix motifs (Pabo and Sauer, 1984), which confer DNA recognition specificity. This conserved region in the five transposases shows some limited homology with these helix–turn–helix sequences, centered around the glycine at position 50. However, the homology is not strong, and in Tn*21* transposase the sequence that would correspond to helix II contains a proline residue. We have no direct evidence that the N-terminal region is involved in DNA-binding. In any event, the binding of DNA by transposase may not involve a single helix–turn–helix structure, as this could not interact with all 38 bp of the inverted repeat.

Alternative explanations of the role of the N-terminal region of transposase in discriminatory recognition of the inverted repeats can be made. For example, altering the N-terminal region may slightly change the tertiary structure of transposase, so that the DNA-binding regions elsewhere on the enzyme preferentially recognise one or other inverted repeat. The identification of the precise residues involved in DNA recognition and binding by transposase awaits further experiments.

We have not examined in detail the catlaytic properties of the hybrid transposases. There are differences in the catalytic behaviour of the Tn*501* and Tn*21* transposases *in vivo*, in that Tn*21* will catalyse replicon fusion by 'single-ended' transposition (Motsch *et al.*, 1985) whereas Tn*501* will not do so with significant frequency. Preliminary results (not shown) suggest that the hybrid transposases do show differences in their abilities to catalyse single-ended transposition, and should provide a useful means of studying this phenomenon.

The method used to generate the hybrid Tn*501*/Tn*21* transposases could be applied to other proteins for which the genes have been cloned, and which show regions of significant homology. Recombination *in vitro* has the disadvantage that there must be an appropriate distribution of restriction sites, or suitable

restriction sites must be introduced. This *in vivo* approach may be most useful in identifying functional regions of a protein as a prerequisite for directed mutagenesis.

Materials and methods

Restriction endonucleases and Klenow fragment of DNA polymerase I were obtained from Boehringer Corporation (London) Ltd or from Anglian Biotechnology Ltd. DNA ligase was prepared by the method of Murray *et al.* (1979). Rapid small-scale preparations of plasmid DNA were performed by the method of Birnboim and Doly (1979). Plasmid DNA purified by CsCl-ethidium bromide isopycnic centrifugation was used for the construction of pLRE100. DNA manipulations and transformation of *E. coli* with plasmid DNA were performed as described in Maniatis *et al.* (1982). Conjugation assays were performed as described by Grinsted and Brown (1984).

Bacterial strains and plasmids

The bacterial strains used in this work are: UB281 (*pro met nalA*; Bennett and Richmond, 1976); UB5201 (UB281 *rec*A56; de la Cruz and Grinsted, 1982); UB1636 (*his lys trp strA;* = JC3272, Achtman *et al.*, 1971); UB1637 (UB1636 *recA56*; de la Cruz and Grinsted, 1982); CSH26 (Miller, 1972); and CSH26Δ*recA* (Lund *et al.*, 1986). The following plasmids were used: R388 (Tp^r, Su^r, Tra^+; Datta and Hedges, 1972); pUB3258 (pBR322::Tn*819*, Ap^r; Grinsted and Brown, 1984); pUB3264 (pACYC184::Tn*819*, Cm^r; Grinsted and Brown, 1984). The plasmid pLRE100 and its derivative recombinants were constructed from the plasmids pUB2413 (de la Cruz and Grinsted, 1982) and pUB2579 (J.Grinsted, unpublished data) as described below.

Construction of the double-replicon pLRE100

Plasmid pUB2579 was a kind gift of Dr J.Grinsted (unpublished data); it had been constructed by inserting a 6.0 kb *Eco*RI fragment of Tn*501* DNA containing all the transposition functions (Brown *et al.*, 1985) into the *Eco*RI site of pBR325. Plasmid pLRE75 was constructed by deleting a 0.65 kb *Bam*HI–*Hin*dIII fragment containing the *res* site and part of the *tnpR* gene of Tn*21* from plasmid pUB2413 (pACYC184 tet::Tn*21*, with a *Bam*HI deletion; de la Cruz and Grinsted, 1982); this was done by controlled *Hin*dIII digestion of *Bam*HI-linearised DNA, filling-in of the single-stranded termini and ligation of the product. Both plasmids pLRE75 and pUB2579 were linearised with *Sal*GI, and ligated. The double replicon in which the *tnpA* genes were in direct orientation was identified by restriction analysis.

Duplicated sequences, which would be preferred sites of homologous recombination, were deleted in two stages. First, an *Ava*I fragment containing the tetracycline resistance determinant from pACYC184 and pBR325 (across the *Sal*GI junction), the distal part of the chloramphenicol resistance (*cat*) gene from pBR325 and 1051 nucleotides of the Tn*501* sequences (Brown *et al.*, 1985), was removed by partial digestion and religation. Second, an *Eco*RI–*Pst*I fragment, which contains the proximal regions of the *bla* and *cat* genes of pBR325, was replaced by the corresponding *Eco*RI–*Pst*I fragment of pBR322, which contains no *cat* sequences. The resulting plasmid, pLRE100, is shown in Figure 1b, and contains 4934 bp of Tn*501* DNA and 3300 bp of Tn*21* DNA.

Isolation and mapping of hybrid tnpA genes

Plasmid pLRE100 was stable in $RecA^-$ strains, but in $RecA^+$ strains gave rise to Cm^r,Ap^s and Ap^r,Cm^s segregants. *E. coli* CSH26 was transformed with pLRE100, and plasmid DNA was made. This DNA was used to transform *E. coli* CSH26Δ*recA*. Transformants were selected on either Cm or Ap, then screened for the absence of the other marker. The A-series of recombinant plasmids were Ap^r,Cm^s and the C-series was Cm^r,Ap^s. The structures of the plasmids were determined by restriction analysis.

The sites of cross-over between sequences from Tn*501* and Tn*21* were mapped by the presence or absence of restriction fragments from the *tnpA* genes of the two transposons. The discriminatory restriction sites used are given in Table I.

The site of crossover which generated the C1-41 hybrid transposase was determined by chain-termination sequencing of double-stranded *Eco*RI-cut plasmid DNA (Korneluk *et al.*, 1985), using as primer a synthetic oligonucleotide corresponding to positions 5335–5355 of the coding strand of Tn*501* DNA (Brown *et al.*, 1985); i.e. positions −21 to −1 relative to the gene, using the numbering of Table I. The site of crossover in the C1-31 hybrid was determined by chain-termination sequencing of a *Nsp*HI–*Eco*RI junction fragment (positions 665–1182 of the gene) cloned in *Sph*I/*Eco*RI-cut bacteriophage M13mp18.

Complementation of transposition of Tn819

Transposition of Tn*819* was detected using the 'mate-out' assay. The *E. coli* strains UB281 and UB1636 (or their *recA* derivatives UB5201 and UB1637) were used alternatively as donor and recipient strains, with appropriate antibiotic counter-selection against the donor. Co-integrate molecules were isolated by mating R388 from a donor strain (containing R388, a Tn*819*-donor plasmid, and the pLRE100 derivative encoding the hybrid transposase) into a $RecA^-$ strain and selecting for the markers on R388 (Tp^r) and on the Tn*819*-donor plasmid (Cm^r or Ap^r), as described elsewhere (Grinsted and Brown, 1984). Each co-integrate was resolved by mating into a $RecA^+$ strain, in which homologous recombination between the two copies of the transposon could occur, and transconjugants were screened to identify resolved R388 derivatives containing a copy of the transposable element.

The C-series of hybrid transposases are on pACYC184-derived replicons, isolated as Cm^r,Ap^s recombinant plasmids. Therefore, in the complementation assay pUB3258 (pBR322:Tn*819*, Ap^r; Brown and Grinsted, 1984) is used as a compatible donor carrying Tn*819*. The A-series of hybrid transposases are on pBR322-derived replicons (Ap^r,Cm^s), and the compatible donor plasmid in this case is pUB3264 (pACYC184::Tn*819*, Cm^r; Grinsted and Brown, 1984).

The frequency of co-integrate formation was measured as the ratio of Tp^r,Ap^r recombinants to the total number of Tp^r recombinants when pUB3258 was used as the Tn*819* donor, and as Tp^r,Cm^r to total Tp^r when pUC3264 was used.

Analysis of the products of transposition

Both the co-integrate and resolved plasmids were digested wtih *Eco*RI, and the products were analysed by gel electrophoresis in 0.8% agarose in 90-mM Tris–borate, 2.5 mM EDTA, pH 8.3. Control digests of the donor plasmid (pUB3258 or pUB3264) and R388 were also performed. The retention of a fragment of the same size as Tn*819* (2.4 kb) from donor to resolved product, and present as a doublet in the co-integrate, occurs only if the Tn*501* inverted repeat is used. (This fragment is arrowed in Figure 2). If the diagnostic fragment was not seen, further restriction digests of the resolved product were performed. The presence of transposon sequences (from at least position 136 of the transposon) was indicated by a *Hin*dIII–*Eco*RI fragment of ~2.2 kb. For each hybrid transposase several independent transposition products were isolated and characterised.

Acknowledgements

We are particularly grateful to John Grinsted for the unstinting provision of advice, strains and unpublished data, and to Pete Lund and Steve Halford for discussion. This work was supported by the Medical Research Council and by the Royal Society, of which N.L.B. is an EPA Cephalosporin Fund Senior Research Fellow.

References

Achtman,M., Willetts,N. and Clark,A.J. (1971) *J. Bacteriol.*, **106**, 529–538.
Bennett,P.M. and Richmond,M.H. (1976) *J. Bacteriol.*, **126**, 1–6.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513–1523.
Brown,N.L., Choi,C.-L., Grinsted,J., Richmond,M.H. and Whitehead,P.R. (1980) *Nucleic Acids Res.*, **8**, 1933–1945.
Brown,N.L., Winnie,J.N., Fritzinger,D. and Pridmore,R.D. (1985) *Nucleic Acids Res.*, **13**, 5657–5669.
Craigie,R. and Mizuuchi,K. (1985) *Cell*, **41**, 867–876.
Craigie,R., Arndt-Jovin,D. and Mizuuchi,K. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7570–7574.
Datta,N. and Hedges,R. (1972) *J. Gen. Microbiol.*, **72**, 349–355.
de la Cruz,F. and Grinsted,J. (1982) *J. Bacteriol.*, **151**, 222–228.
Gill,R., Heffron,F., Dougan,G. and Falkow,S. (1978) *J. Bacteriol.*, **136**, 742–756.
Grindley,N.D.F. and Reed,R.R. (1985) *Annu. Rev. Biochem.*, **54**, 863–896.
Grinsted,J., de la Cruz,F., Altenbuchner,J. and Schmitt,R. (1982) *Plasmid*, **8**, 276–286.
Grinsted,J. and Brown,N.L. (1984) *Mol. Gen. Genet.*, **197**, 497–502.
Heffron,F. (1983) In Shapiro,J. (ed.), *Mobile Genetic Elements*. Academic Press, New York, pp. 223–260.
Heffron,F., McCarthy,B.J., Ohtsubo,H. and Ohtsubo,E. (1979) *Cell*, **18**, 1153–1163.
Kleckner,N. (1981) *Annu. Rev. Genet.*, **15**, 341–404.
Korneluk,R.G., Quan,F. and Gravel,R.A. (1985) *Gene*, **40**, 317–323.
Lund,P.A., Ford,S.J. and Brown,N.L. (1986) *J. Gen. Microbiol.*, **132**, 465–480.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Michiels,T., Cornelis,G., Ellis,K. and Grinsted,J. (1987) *J. Bacteriol.*, **169**, 624–631.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
Motsch,S., Schmitt,R., Avila,P., de la Cruz,F., Ward,E. and Grinsted,J. (1985) *Nucleic Acids Res.*, **13**, 3335–3342.
Murray,N.E., Bruce,S.A. and Murray,K. (1979) *J. Mol. Biol.*, **132**, 493–505.
Pabo,C.O. and Sauer,R.T. (1984) *Annu. Rev. Biochem.*, **53**, 293–321.
Shapiro,J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1933–1937.
Shaw,J. and Clewell,D.B. (1985) *J. Bacteriol.*, **164**, 782–796.
Ward,E. and Grinsted,J. (1987) *Nucleic Acids Res.*, **15**, 1799–1806.
Wishart,W.L., Broach,J.R. and Ohtsubo,E. (1985) *Nature*, **314**, 556–558.

Received on January 20, 1987; revised on June 9, 1987