Resolvase-like recombination performed by the TP901-1 integrase

Anne Breüner, Lone Brøndsted† and Karin Hammer

Author for correspondence: Karin Hammer. Tel: +45 45 25 24 96. Fax: +45 45 88 26 60.
e-mail: karin.hammer@biocentrum.dtu.dk.

Department of Microbiology, Technical University of Denmark, DK-2800 Lyngby, Denmark

The site-specific recombination system of temperate lactococcal bacteriophage TP901-1 is unusual in several respects. First, the integrase belongs to the family of extended resolvases rather than to the λ integrase family and second, in the presence of this integrase, a 56 bp *attP* fragment is sufficient for efficient recombination with the chromosomal *attB* site in the host *Lactococcus lactis* subsp. *cremoris* MG1363. In the present work, this *attB* site was analysed and a 43 bp *attB* region was found to be the smallest fragment able to participate fully in recombination. *In vitro* studies showed that the TP901-1 integrase binds this 43 bp *attB* fragment, the 56 bp *attP* and a larger *attP* fragment with equal affinity. Mutational analysis of the 5 bp common core region (TCAAT) showed that the TC dinucleotide is essential for recombination, but not for binding of the integrase, whereas none of the last three bases are important for recombination. When a number of *attL* sites, obtained by recombination between an *attB* site containing a mutation in this TC dinucleotide and a wild-type *attP* site, were sequenced, a mix of sites with the wild-type or the mutated sequence was obtained. These results are consistent with the hypothesis that the TC dinucleotide constitutes the TP901-1 overlap region. A 2 bp overlap region has been observed in recombination reactions catalysed by all other members of the resolvase/invertase family tested so far. By selecting for *attB* sites with a decreased ability to participate in recombination, two bases located outside the core region of *attB* were shown to be involved in the *in vitro* binding of the TP901-1 integrase.

Keywords: lactococcal bacteriophage, extended resolvase, site-specific recombination, mechanism of recombination, chromosomal attachment site

INTRODUCTION

In most temperate bacteriophages, the phage genome is integrated in the host chromosome to ensure stable inheritance during lysogenic growth. Integration is obtained by site-specific recombination between the attachment sites *attB* and *attP*, located on the bacterial and the phage genomes respectively. In the resulting lysogenic strain, the prophage is integrated on the host chromosome between *attL* and *attR*, which are the recombination products of *attB* and *attP*.

TP901-1 is a temperate bacteriophage, induced by UV light from the lysogenic strain *Lactococcus lactis* subsp. *cremoris* 901. The attachment sites of TP901-1 before and after recombination are shown in Fig. 1(b). A 13 bp identical region with a mismatch at position 6 was identified by comparison of the attachment sites. Since the mismatch from *attB* was consistently found in *attR* and not *attL*, the core region, the region of homology between the attachment sites, was defined as being the 5 bp TCAAT region (Christiansen *et al.*, 1994).

The TP901-1 integrase, the only phage-encoded protein required for site-specific integration, belongs to the new protein family of extended resolvases (Christiansen *et al.*, 1996). The extended resolvases contain a region with

†**Present address:** Department of Dairy and Food Science, Rolighedsvej 30, The Royal Veterinary and Agricultural University, DK-1958 Frederiksberg, Denmark.

Abbreviations: Ap, ampicillin; Cm, chloramphenicol; Ery, erythromycin; Kn, kanamycin.

The GenBank accession number for the sequence reported in this paper is Y15043.

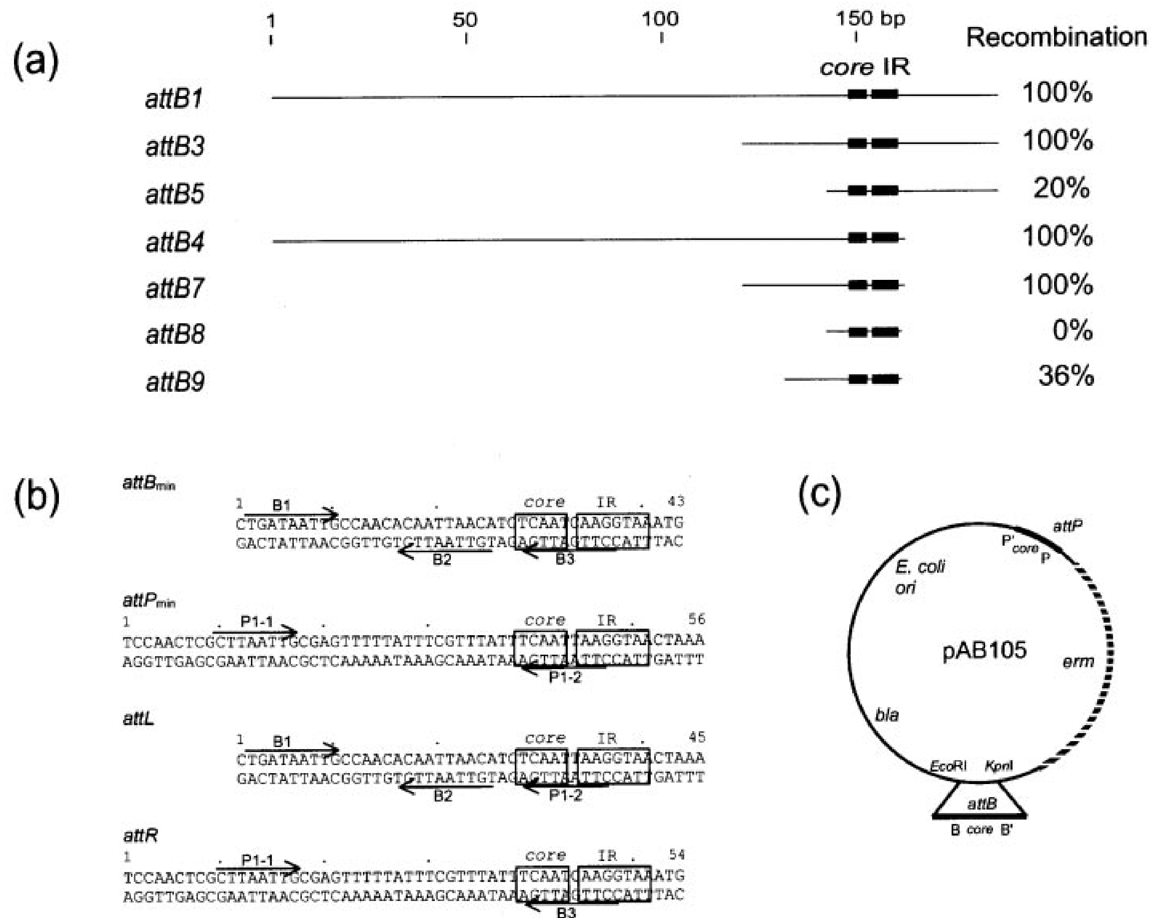

Fig. 1. (a) *attB* fragments tested in the deletion analysis of the *attB* region of the temperate bacteriophage TP901-1. The ability of each fragment to participate in site-specific recombination with *attP* is shown in percentages, 100 % indicating that recombination takes place in all cells. (b) Nucleotide sequence of the minimal *attB* and *attP* sites, and of the corresponding *attL* and *attR* sites. The core and identical region (IR) are boxed. The repeats identified in *attB* (this work) and in *attP* (Brøndsted & Hammer, 1999) are shown as arrows. The sites are written as: *attB*, B-*core*-B′; *attP*, P-*core*-P′; *attL*, B-*core*-P′; *attR*, P-*core*-B′. (c) Map of the plasmid pAB105, which is used in the *attB* analyses. *attB* can be inserted in the *Eco*RI and *Kpn*I sites as shown. *attB* and *attP* are shown as black lines, with the orientation of the sites indicated, *erm* as a broken line. Recombination between *attB* and *attP* leads to *attL* being present on the plasmid and loss of *erm*.

homology to the catalytic domain of the members of the resolvase family of resolvases and invertases, including the serine residue forming a covalently bound intermediate, and it is therefore likely that these proteins perform recombination in a similar way (resolvase-type recombination). During this type of recombination, double-stranded cuts are introduced into the DNA backbone, leading to 2 bp single-stranded overhangs. The strands are then rotated relative to one another, the overhangs are paired with their new partner, and the strands are religated. The overlap region of the resolvase family, the region between the cleavage sites on opposite strands, which consists of heteroduplex DNA after recombination, is thus 2 bp (Stark *et al.*, 1989).

Seven members of the family of extended resolvases have been investigated functionally. These proteins are involved in the following site-specific recombination events: excision of transposons (TnpX, Crellin & Rood, 1997; TndX, Wang *et al.*, 2000), excision of fragments from the chromosome (CisA, Sato *et al.*, 1990; Popham & Stragier, 1992; XisF, Carrasco *et al.*, 1994) and phage integration (ϕC31 Int, Kuhstoss & Rao, 1991; Thorpe & Smith, 1998; TP01-1 Int, Christiansen *et al.*, 1996; R4 Sre, Matsuura *et al.*, 1996). For three of the recombinases (TnpX, ϕC31 Int and R4 Sre) the conserved serine residue in the N-terminal part of the protein has been shown to be vital for the catalysis, and for TnpX excision a 2 bp overlap region has been demonstrated. These findings support the hypothesis that the extended resolvases perform resolvase-type recombination.

In contrast to the extended resolvase integrases, most phage integrases identified to date are related to the λ integrase and are thus expected to perform integrase-type recombination, which involves a covalently bound tyrosine–DNA intermediate. Four integrase monomers bind to the substrate DNA sites, two of which are

responsible for the introduction of specific single-stranded breaks in the DNA backbone of both substrate sites. The DNA strands are ligated to the new partners, creating a Holliday junction located in the approximate centre of the overlap region. Isomerization of the Holliday intermediate activates the second pair of integrase monomers, allowing cleavage and exchange of the second pair of DNA strands (Azaro & Landy, 1997; Gopaul & Van Duyne, 1999; Guo *et al*., 1999).

To ensure proper directionality of the recombination processes, the recognition sites *attB* and *attP* for phage integration are always of different sequence, in contrast to the substrate DNA sites for resolution or inversion processes. However, in the family of extended resolvases the ϕC31 *attB* and *attP* sites are of similar size (minimum sizes 34 and 39 bp, respectively; Groth *et al*., 2000), and the ϕC31 integrase has the same affinity for the two sites (Thorpe *et al*., 2000). In contrast, the minimal *attB* of *Escherichia coli* phage λ is 23 bp in length, and thus much shorter than the 235 bp minimal *attP* site (Mizuuchi & Mizuuchi, 1985). The affinity of the λ integrase for the *attB* site is much lower than for the *attP* site (Richet *et al*., 1988).

The TP901-1 minimal *attP* was found to be 56 bp long (Brøndsted & Hammer, 1999), and is thus substantially smaller than the minimal *attP* regions reported for bacteriophages performing integrase-type recombination. To support the hypothesis that the TP901-1 integrase performs resolvase-type recombination further, we decided to investigate the boundaries of the TP901-1 *attB* site and the affinity of the integrase for the *attP* and *attB* sites. Furthermore, single bases involved in recombination were identified by mutational analysis, and the importance of some of these bases for the *in vitro* binding of the TP901-1 integrase was examined. Our data strongly support the hypothesis that the TP901-1 integrase performs resolvase-type recombination.

METHODS

Bacterial strains and culture conditions. The bacterial strains used are listed in Table 1. *Lactococcus lactis* subsp. *cremoris* MG1363 was grown without shaking at 30 °C in M17 broth (Terzaghi & Sandine, 1975) supplemented with 0·5% (w/v) glucose (GM17). When required, erythromycin (Ery) was added to a final concentration of 2 mg l^{-1} and chloramphenicol (Cm) to 5 mg l^{-1}. *Escherichia coli* cells were propagated at 37 °C with shaking in Luria–Bertani (LB) broth (Miller, 1972), and Ery was added to a final concentration of 150 mg l^{-1}, Cm to 25 mg l^{-1}, kanamycin (Kn) to 50 mg l^{-1} and ampicillin (Ap) to 100 mg l^{-1}. IPTG was added to a final concentration of 1 mM. To prepare plates, all media were solidified by adding 1·5% Bacto agar.

DNA preparation. Plasmid DNA was isolated from both lactococcal and *E. coli* cells by the alkaline lysis technique (Sambrook *et al*., 1989). To promote lysis, lactococcal cells were treated with lysozyme at a final concentration of 20 g l^{-1} at 37 °C with shaking before the addition of NaOH. If required, the DNA was further purified on Qiagen columns. Chromosomal DNA from lactococcal cells was prepared as described for *E. coli* (Sambrook *et al*., 1989), except that the cells were frozen for 30 min at −80 °C after harvesting, thawed to room temperature and treated with lysozyme at a final concentration of 20 g l^{-1} for 30 min at 37 °C to promote lysis. Phage DNA was prepared according to Sambrook *et al*. (1989).

Recombinant DNA techniques. DNA manipulations were performed by standard techniques (Sambrook *et al*., 1989). Enzymes and corresponding buffers were supplied by Pharmacia Biotech or New England Biolabs. The Ampli*Taq* DNA polymerase was used for PCR amplification of DNA fragments (Perkin Elmer Cetus). DNA sequences were determined as described by Sanger *et al*. (1977), modified according to the Thermo Sequenase Radiolabelled terminator cycle sequencing kit (Amersham Life Science).

Oligonucleotides used in this study. Oligonucleotides were supplied by T-A-G-Copenhagen, Symbion, Copenhagen, Denmark or by Pharmacia Biotech. Degenerate oligonucleotides were supplied by P. Hobolth, Lyngby, Denmark.

The sequences of the primers used were (written 5′–3′): T7, TAATACGACTCACTATAGGG; #1201, AACAGCTATGACCATG; #1211, GTAAAACGACGGCCAGT; perm, GTTACACGTTACTAAAGGG; perm1, GCAAGTCACGAACAC; pINT-1, CCCCGGATCCAGAAATGAGGTACAAAAAC; pINT-2, CCCCTCGAGTCGACGCAATTAAGCGAGTTGGAATTT; p1attB, CCCCGAATTCGATCCAACTCATAAAGTT; p3attB, CCCCGAATTCTGATAATTGCCAACAC; p4attB, GGGGGGGTACCATTTACCTTGATTGAGATG; p5attB, CCCCGAATTCATCTCAATCAAGGTAAATG; p6attB, *AATT*CATCTCAATCAAGGTAAAT*GTAC*; p7attB, ATTTACCTTGATTGAGATG; p8attB, *AATT*ACACAATTAACATCTCAATCAAGGTAAAT*GTAC*; p9attB, ATTTACCTTGATTGAGATGTTAATTGTGT; pG10C, *AATT*CTGATAATTCCCAACACAATTAACATCTCAATCAAGGTAAATG*GTAC*; pG10C-rev, CATTTACCTTGATTGAGATGTTAATTGTGTTGGGAATTATCAG; pC33T, *AATT*CTGATAATTGCCAACACAATTAACATCTCAATTAAGGTAAATG*GTAC*; pC33T-rev, CATTTACCTTAATTGAGATGTTAATTGTGTTGGCAATTATCAG; pT28M, *AATT*CTGATAATTGCCAACACAATTAACATCMCAATCAAGGTAAATG*GTAC*; pC29D, *AATT*CTGATAATTGCCAACACAATTAACATCTDAATCAAGGTAAATG*GTAC*; pA30B, *AATT*CTGATAATTGCCAACACAATTAACATCTCBATCAAGGTAAATG*GTAC*; pA31B, *AATT*CTGATAATTGCCAACACAATTAACATCTCABTCAAGGTAAATG*GTAC*; pT32V, *AATT*CTGATAATTGCCAACACAATTAACATCTCAAVCAAGGTAAATG*GTAC*; pattB-Low and pattB-High, *AATT*CTGATAATTGCCAACACAATTAACATCTCAATCAAGGTAAATG*GTAC*; pattB-Low-rev and pattB-High-rev, CATTTACCTTGATTGAGATGTTAATTGTGTTGGCAATTATCAG. Bases in italics constitute the overhangs after annealing, underlined bases are changed relative to the wild-type sequence: M, A or C; D, G, A or T; B, G, T or C; V, G, A or C; pattB-Low, pattB-High, and pattB-High-rev are degenerate. Except at the bases constituting the single-stranded overhangs the level of degeneracy is reported later in this section.

Transformation of *E. coli* and *L. lactis*. Plasmid DNA was introduced into *E. coli* cells by making the cells competent with $CaCl_2$ and transforming as described by Sambrook *et al*. (1989). To introduce DNA into lactococcal cells, they were made electrocompetent and electroporated (Holo & Nes, 1989).

Table 1. Bacterial strains and plasmids

Strain or plasmid	Relevant genotype	Co-ordinates of inserted fragment(s)		Reference
		TP901-1*	*attB*†	
E. coli				
DH5α	*ϕ80lacZ*ΔM15 Δ(*lacZYA–argF*) U169 *recA1 endA1 hsdR17 supE44 thi-1 gyrA96 relA1*			Laboratory stock
NF1830	*recA* F′ *lacI*^{q1}Z::Tn5 (Kn^R)			Bonekamp *et al.* (1984)
L. lactis* subsp. *cremoris				
MG1363	Plasmid-free			Gasson (1983)
Plasmids				
pGEM-3Zf(−)	*E. coli* cloning vector (Ap^R)	–	–	Promega
pGEM-7Zf(+)	*E. coli* cloning vector (Ap^R)	–	–	Promega
pUC7,erm	*E. coli* vector::*erm* (Ery^R)	–	–	de Vos, NIZO, Netherlands
pACYC184	*E. coli* cloning vector (Cm^R)	–	–	New England Biolabs
pUHE23-2	*E. coli* overexpression vector (Ap^R)	–	–	Lutz & Bujard (1997)
pCI372	*E. coli–Lactococcus lactis* shuttle vector (Cm^R)	–	–	Hayes *et al.* (1990)
pAK80	*E. coli–Lactococcus lactis* shuttle vector (Ery^R)	–	–	Israelsen *et al.* (1995)
pBF12	pMOS*blue*-T::333 bp TP901-1 *attP* (Ap^R Ery^R)	−104–228	–	Brøndsted & Hammer (1999)
pBF30	pMOS*blue*-T::$attP_{min}$ (Ap^R Ery^R)	−12–43	–	Brøndsted & Hammer (1999)
pLB54	pGEM-7Zf(+)::224 bp *attB* (Ap^R Ery^R)	–	1740–1963	This study
pLB58	pCI372::TP901-1 *attP*, *orf1-3* (Cm^R)	−98–2708	–	Breüner *et al.* (1999)
pLB81	pACYC184::TP901-1 *orf1* (Cm^R)	25–1910	–	Breüner *et al.* (1999)
pBC144	pGEM-7Zf(+)::*erm* (Ery^R)	–	–	Christiansen *et al.* (1994)
pBC193	pGEM-7Zf(+)::224 bp *attB* (Ap^R)	–	1740–1963	Christiansen (1995)
pAB30	pMOS*blue*-T::$attP_{min}$ (Ap^R Ery^R)	−12–43	–	This study
pAB105	pMOS*blue*-T::333 bp TP901-1 *attP* (Ap^R Ery^R)	−104–228	–	This study
pAB107	pGEM-7Zf(+)::*attB1* (Ap^R)	–	1740–1926	This study
pAB120	pGEM-3Zf(−)::*attB7* (Ap^R)	–	1859–1901	This study
pAB121	pAB105::*attB1* Ap^R (Ery^R)	−104–228	1740–1926	This study
pAB123	pAB105::*attB3* Ap^R (Ery^R)	−104–228	1859–1926	This study
pAB124	pAB105::*attB4* Ap^R (Ery^R)	−104–228	1740–1901	This study
pAB125	pAB105::*attB5* Ap^R (Ery^R)	−104–228	1882–1926	This study
pAB127	pAB105::*attB7* Ap^R (Ery^R)	−104–228	1859–1901	This study
pAB128	pAB105::*attB8* Ap^R (Ery^R)	−104–228	1882–1901	This study
pAB129	pAB105::*attB9* Ap^R (Ery^R)	−104–228	1872–1901	This study
pAB130	pGEM-3Zf(−)::*attB7* (Ap^R Ery^R)	–	1859–1901	This study
pAB131	pGEM-3Zf(−)::*attB7* and $attP_{min}$ (Ap^R Ery^R)	−12–43	1859–1901	This study
attB-mutants	pAB105::mutated *attB* (Ap^R Ery^R)	−104–228	1859–1901‡	This study
pAB201	pGEM-3Zf(−)::TP901-1 *orf1* (Ap^R)	25–1508	–	This study
pAB202-4	pUHE23-2::TP901-1 *orf1* (Ap^R)	25–1508	–	This study

* TP901-1 co-ordinates: positive co-ordinates correspond to the sequence reported in EMBL accession no. Y14232. Negative co-ordinates −1 to −104 correspond to bases 2727 to 2831 in the sequence reported in EMBL accession no. X85213.

† Chromosomal co-ordinates correspond to GenBank accession no. Y15043.

‡ Mutated *attB* fragments inserted: G_3A, T_9A, $G_{10}C$, $C_{17}A$, $C_{17}G$, $A_{25}T$, $T_{28}G$, $T_{28}C$, $T_{28}A$, $C_{29}G$, $C_{29}A$, $C_{29}T$, $A_{30}G$, $A_{30}T$, $A_{30}C$, $A_{31}G$, $A_{31}T$, $A_{31}C$, $T_{32}G$, $T_{32}C$, $T_{32}A$, $C_{33}T$.

Construction of pAB105 and deletion analysis of *attB*. The plasmids used and constructed in this work are shown in Table 1 and the *attB* fragments are listed in Table 2. To construct plasmid pAB105 (Fig. 1c), pBF12 was digested with *Eco*RI, the larger fragment was religated, and the 1·1 kb *erm* cassette obtained by digesting pUC7,erm with *Bam*HI was inserted in the *Bam*HI site.

Plasmids pAB121, pAB123, pAB124, pAB125 and pAB127 contain *attB* fragments 1, 3, 4, 5 and 7, respectively, inserted in the *Eco*RI–*Kpn*I sites in pAB105. The fragments were obtained by PCR with pAB107 as template and digested with *Eco*RI and *Kpn*I (primers: Table 2). pAB107 was obtained by digesting pBC193 ($attB_{224}$) with *Hin*dIII and religating, resulting in *attB1* being present. The *attB8* and *attB9* fragments in plasmids

Table 2. *attB* fragments

attB fragment	Co-ordinates*	Mutations†	Obtained by	Oligos used
attB1	1–187	None	PCR	T7 and p1attB
attB3	120–187	None	PCR	T7 and p3attB
attB4	1–162	None	PCR	p1attB and p4attB
attB5	143–187	None	PCR	T7 and p5attB
attB7 ($attB_{min}$)	120–162	None	PCR	p3attB and p4attB
attB7 ($attB_{min}$)	120–162	None	Annealing	pattB-Low and pattB-Low-rev
attB8	143–162	None	Annealing	p6attB and p7attB
attB9	133–162	None	Annealing	p8attB and p9attB
G_3A	120–162	$G_3 \rightarrow A$	Annealing	pattB-High and pattB-High-rev
T_9A	120–162	$T_9 \rightarrow A$	Annealing	pattB-High and pattB-High-rev
$G_{10}C$	120–162	$G_{10} \rightarrow C$	Annealing	pG10C and pG10C-rev
$C_{17}G$	120–162	$C_{17} \rightarrow G$	Annealing	pattB-Low and pattB-Low-rev
$C_{17}A$	120–162	$C_{17} \rightarrow A$	Annealing	pattB-Low and pattB-Low-rev
$A_{25}T$	120–162	$A_{25} \rightarrow T$	Annealing	pattB-Low and pattB-Low-rev
$T_{28}G$	120–162	$A_{28} \rightarrow G$	Annealing	pattB-Low and pattB-Low-rev
$T_{28}C$, $T_{28}A$	120–162	$T_{28} \rightarrow$ C,A	Annealing	pT28M and pattB-Low-rev
$C_{29}G$, $C_{29}A$, $C_{29}T$	120–162	$C_{29} \rightarrow$ G,A,T	Annealing	pC29D and pattB-Low-rev
$A_{30}G$, $A_{30}T$, $A_{30}C$	120–162	$A_{30} \rightarrow$ G,C,T	Annealing	pA30B and pattB-Low-rev
$A_{31}G$, $A_{31}T$, $A_{31}C$	120–162	$A_{31} \rightarrow$ G,C,T	Annealing	pA31B and pattB-Low-rev
$T_{32}G$, $T_{32}A$, $T_{32}C$	120–162	$T_{32} \rightarrow$ G,C,A	Annealing	pT32V and pattB-Low-rev
$C_{33}T$	120–162	$C_{33} \rightarrow T$	Annealing	pC33T and pC33T-rev

* Co-ordinates correspond to the numbering in Fig. 1(a).

† Numbering of bases corresponds to Fig. 2.

pAB128 and pAB129 were obtained by annealing pairs of oligonucleotides (Table 2), resulting in double-stranded *attB* fragments surrounded by *Eco*RI and *Kpn*I overhangs. The sequence of all *attB* fragments was verified before further analysis.

Introducing site-specific mutations in *attB*. When introducing base-substitutions in *attB*, again complementary oligonucleotides were annealed and the resulting fragment was inserted in the *Eco*RI–*Kpn*I sites in pAB105. For the fragments G10C and C33T both oligonucleotides were of the desired sequence. The remaining *attB* fragments, which were obtained by introducing site-specific mutations in *attB*, were mutated in the core region ($T_{28}C$, $T_{28}A$, $C_{29}G$ $C_{29}A$, $C_{29}T$, $A_{30}G$, $A_{30}T$, $A_{30}C$, $A_{31}G$, $A_{31}T$, $A_{31}C$, $T_{32}G$, $T_{32}C$, $T_{32}A$). To obtain the mutations by using a minimum of oligonucleotides, one of the pair of oligonucleotides annealed was a mix of primers with all bases except the original one at the desired position, while the other was wild-type.

Introducing random single-base-pair mutations in *attB*. One procedure for obtaining random mutations in *attB* involved selecting for non-functional sites. One of the oligonucleotides annealed, pattB-Low, was degenerate except in the regions constituting the single-stranded overhangs, while the second primer was not (pattB-Low-rev). To select for non-functional sites, the pAB105-$attB_{mut}$ ligation mix was introduced directly into *E. coli* strain DH5α/pLB81 and plated directly on Ery plates. The nucleotide sequence was determined of all the non-functional *attB* sites to identify fragments with single-base-pair substitutions. Four such *attB* fragments were obtained: $C_{17}A$, $C_{17}G$, $A_{25}T$ and $T_{28}G$. The remaining non-functional fragments contained multiple substitutions and/or deletions (data not shown).

The oligonucleotide pattB-Low is 2·4% degenerate. Since $P(X)$, the probability of an oligonucleotide of length N containing X mutations at a level of degeneracy C, equals $(N!/[X!(N-X!)]!) \times C^{X}(1-C)^{N-X}$, $P(0)$, the probability of one pattB-Low oligonucleotide being of the wild-type sequence is 35·0%, $P(1)$ is 37·2%, and thus the probability of pattB-Low containing more than one mutation is 27·8%. Since pattB-Low was annealed to pattB-Low-rev, which is not degenerate, the probability of obtaining mutations is approximately half this.

In the second strategy for obtaining random single-base-pair mutations in *attB*, no selection for non-functional sites was involved, and therefore the oligonucleotides annealed had to be more degenerate than the pattB-Low-primers to avoid obtaining a high number of non-mutated sites. Both pattB-High and pattB-High-rev are 7·2% degenerate; thus $P(0)$ is 4·0%, $P(1)$ is 13·4%, $P(2)$ is 21·8%, and the probability of pattB-High or pattB-High-rev containing more than two mutations is 60·8%. The sequence of a number of *attB* fragments inserted in pAB105 was determined, and two were found to contain single mutations (G_3A and T_9A); the remaining fragments analysed contained either no or multiple substitutions (data not shown).

Construction of remaining plasmids. Plasmid pAB131 contains $attB_{min}$ and $attP_{min}$ separated by the *erm* gene, and was obtained by ligating the 150 bp *Acc*I–*Hin*dIII fragment of pAB30 to the 4 kb *Acc*I–*Hin*dIII fragment of pAB130. Plasmid pAB30 is identical to pBF30, except that *erm* is inserted in the opposite orientation by digesting pBF30 with *Bam*HI and reannealing. The pAB130 plasmid was obtained by inserting *erm* on a 1·1 kb cassette from pUC7,erm in the *Bam*HI site in

pAB120. Plasmid pAB120 contains $attB_{min}$, inserted in the *Eco*RI–*Kpn*I sites in pGEM-3Zf(−). The $attB_{min}$ fragment was obtained by annealing pattB-Low and pattB-Low-rev. pattB-Low is degenerate, but a wild-type fragment was obtained after sequencing *attB* in two plasmids.

The integrase overexpression plasmid pAB202-4 was constructed by inserting the 1·4 kb *Bam*HI–*Sal*I fragment of pAB201, containing TP901-1 *orf1*, in *Bam*HI–*Sal*I in pUHE23-2. To construct pAB201, the TP901-1 *orf1* gene and SD sequence was PCR amplified with primers pINT-1 and pINT-2 and TP901-1 DNA as template. The 1·4 kb PCR fragment was digested with *Bam*HI and *Sal*I, inserted in *Bam*HI–*Sal*I in pGEM3-Zf(−) and sequenced.

Investigation of the frequency of site-specific recombination. Plasmid DNA of the pAB105 derivatives was introduced into DH5α/pLB81, in which the TP901-1 integrase is present and site-specific recombination between *attP* and *attB* can take place. The transformants were selected on LB plates containing Cm and Ap and, after growth overnight at 37 °C, a number of colonies were streaked both on LB with Cm and Ap and on LB with Cm, Ap and Ery. After another overnight growth the number of cells which had retained resistance to Ery could be determined. If zero to five single colonies were present on the plate with Ery, the colony was considered erythromycin sensitive (Ery^S); otherwise the colony was considered to be resistant (Ery^R). We estimate that approximately 10^4 cells are transferred in the streaking from the transformation plate without Ery to the plate with Ery, and thus this level of discrimination means that colonies which contain one, or fewer, *erm*-containing cell per 2000 cells will be considered Ery^S.

For the deletion analysis, 100 colonies were tested for each *attB* fragment (1, 3–5 and 7–9), and for the remaining fragments 50 colonies were tested.

Overexpression of the TP901-1 integrase in *E. coli*. Overproduction of the TP901-1 integrase was induced in a culture of NF1830/pAB202-4 cells by addition of IPTG. The cells were resuspended in elution buffer (20 mM Tris-HCl, pH 7·5; 1 mM EDTA, pH 7·5; 100 mM NaCl; 5 mM $MgCl_2$; 1 mM DTT; 5 %, v/v, glycerol) and sonicated until lysis on ice. The extract was desalted on a HiTrap Desalting Column (Pharmacia Biotech) and eluted with 5 × 400 µl elution buffer. To monitor overexpression of the integrase, the extracts were applied on a SDS-PAGE gel (Laemmli, 1970). A protein of approximately 55 kDa, the predicted size of the TP901-1 integrase, was visible in the extract prepared from cells grown with IPTG, but absent in the extract prepared from cells grown without IPTG (results not shown). The integrase was estimated to constitute approximately 10 % of the total protein content, which was determined to be 9·4 g l^{-1} by Lowry analysis, using bovine serum albumin as reference.

Gel mobility-shift assays. Fragments were labelled by PCR with [α-^{32}P]dATP. The following templates and primers were used: $attP_{min}$, template pBF30, primers perm1 and T7; $attB_{min}$, $T_{28}G$ and $C_{17}A$, template pAB105 with the relevant fragment, primers perm and #1211. Binding was carried out in binding buffer (elution buffer with 40 mg bovine serum albumin l^{-1}) with competitor DNA added to a final concentration of 0·5 mg l^{-1} (the *erm* cassette amplified by PCR on pUC7,erm plasmid DNA with primers #1201 and #1211). Protein extracts were added and the reaction was incubated for 10–15 min at room temperature. Labelled fragment was added to a final concentration of 0·06 nM and allowed to bind for another 25–30 min at room temperature (final volume: 10 µl). The reactions were run on a 5 % polyacrylamide gel at ~7 V cm^{-1}. After vacuum-drying onto 3 MM Whatman paper, the gel was acquired and analysed using a Packard Instant Imager.

RESULTS

Strategy for analysis of *attB*

To analyse the TP901-1 *attB* site the *E. coli* vector pAB105 was constructed. This plasmid carries a functional *attP* site of 333 bp positioned next to the *erm* gene, and a polylinker region for insertion of a DNA fragment containing the *attB* sequences of interest on the other side of *erm*, relative to *attP*, resulting in the gene order *attP-erm-attB* (Fig. 1c). Thus, recombination between the *attP* and *attB* sites will result in excision of the *erm* gene and in erythromycin sensitivity of the strain carrying the plasmid. Since pAB105 also contains the *bla* gene, this plasmid can be selected for by ampicillin (Ap) independently of recombination. After insertion of a fragment containing the *attB* sequences of interest into pAB105, the plasmid is introduced into the *E. coli* strain DH5α already containing plasmid pLB81, from which the TP901-1 integrase is expressed. The ability of the inserted *attB* fragment to participate in recombination is then determined by testing the sensitivity of the Ap^R transformants towards Ery. If all transformants are Ery^S, recombination between *attP* and *attB* has taken place in all cells (100 % recombination), while no recombination results in all cells being Ery^R.

Identification of the minimal *attB*

This system was used to investigate the boundaries of *attB* by testing the ability of *attB* sites of different lengths to recombine with *attP* (Fig. 1a). Both the 187 bp *attB1* fragment and the 68 bp *attB3* fragment are fully able to recombine with *attP*. By deletion of additional 23 bp (*attB5*) only 20 % recombination is observed. Thus, these 23 bp stimulate, but are not required for, recombination. When 25 bp are deleted from the right end of *attB1*, *attB3* and *attB5*, giving rise to *attB4*, *attB7* and *attB8*, respectively, *attB4* and *attB7* are still able to recombine 100 %, whereas no recombination is observed for *attB8*. Thus, the 43 bp *attB7* fragment is the smallest *attB* fragment identified giving rise to full recombination, and is therefore termed the minimal *attB* or $attB_{min}$. With *attB9*, in which 13 bp are deleted from the left of *attB7*, 36 % recombination is obtained. Thus, the 13 bp located between the left ends of *attB9* and *attB7* are required for full recombination.

As mentioned, pAB105 contains an *attP* site of 333 bp. To ensure that the size of the *attP* site does not affect the results of the deletion analysis, we constructed a plasmid (pAB131) containing the 43 bp $attB_{min}$ and the 56 bp $attP_{min}$ separated by the *erm* gene. When recombination was tested with this plasmid, all cells tested (100) were Ery^S, confirming that the minimal *attB* and *attP* fragments do contain the information required for recombination.

Table 3. Transformation frequencies of *L. lactis* subsp. *cremoris* MG1363 strains

Plasmid	MG1363/pLB58 (*attP* and integrase) [c.f.u. μg^{-1}] (relative to pBC144)*	MG1363/pCI372 (control) [c.f.u. μg^{-1}] (relative to pBC144)*	MG1363/pLB58 relative to MG1363/pCI372
pAB130 ($attB_{min}$)	$3{\cdot}0 \times 10^6$ (2×10^4)	$3{\cdot}8 \times 10^2$ (7)	$7{\cdot}9 \times 10^3$
pLB54 ($attB_{224}$)	$6{\cdot}0 \times 10^6$ (4×10^4)	$3{\cdot}5 \times 10^3$ (70)	$1{\cdot}7 \times 10^3$
pBC144 (no *attB*)	$1{\cdot}4 \times 10^2$ (1)	$5{\cdot}2 \times 10^1$ (1)	2·7

* The frequencies are corrected for the difference in competence between the strains, observed when a freely replicating, *erm*-carrying plasmid (pAK80) was introduced in the two strains. The number in parentheses shows the relative transformation efficiency, compared to pBC144 in the same strain.

	Fragment	Sequence (core, IR)	% Recomb.
	$attB_{wt}$	CTGATAATTGCCAACACAATTAACATC**TCAATCAAGGTAA**ATG	100
(a)	T28G	CTGATAATTGCCAACACAATTAACATCGCAATCAAGGTAAATG	0
	T28C	CTGATAATTGCCAACACAATTAACATCCCAATCAAGGTAAATG	0
	T28A	CTGATAATTGCCAACACAATTAACATCACAATCAAGGTAAATG	10
	C29G	CTGATAATTGCCAACACAATTAACATCTGAATCAAGGTAAATG	22
	C29A	CTGATAATTGCCAACACAATTAACATCTAAATCAAGGTAAATG	76
	C29T	CTGATAATTGCCAACACAATTAACATCTTAATCAAGGTAAATG	76
	A30B	CTGATAATTGCCAACACAATTAACATCTCBATCAAGGTAAATG	100
	A31B	CTGATAATTGCCAACACAATTAACATCTCABTCAAGGTAAATG	100
	T32V	CTGATAATTGCCAACACAATTAACATCTCAAVCAAGGTAAATG	100
(b)	C17A	CTGATAATTGCCAACAAAATTAACATCTCAATCAAGGTAAATG	21
	C17G	CTGATAATTGCCAACAGAATTAACATCTCAATCAAGGTAAATG	50
	A25T	CTGATAATTGCCAACACAATTAACTTCTCAATCAAGGTAAATG	88
	G3A	CTAATAATTGCCAACACAATTAACATCTCAATCAAGGTAAATG	100
	T9A	CTGATAATAGCCAACACAATTAACATCTCAATCAAGGTAAATG	100
	G10C	CTGATAATTCCCAACACAATTAACATCTCAATCAAGGTAAATG	100
	C33T	CTGATAATTGCCAACACAATTAACATCTCAATTAAGGTAAATG	100

Fig. 2. *attB* fragments containing single-base-pair mutations and the ability of these fragments to participate in recombination in *E. coli*. The 43 bp $attB_{min}$ is shown at the top, bases numbered 1 to 43, core and identical region in bold. The changed bases are highlighted. The numbers to the right indicate the percentage of the isolates tested in which complete recombination had taken place. (a) *attB* fragments containing single-base-pair mutations in the core region. B: any base but A, V: any base but T. (b) *attB* fragments containing single-base-pair mutations outside the core region.

Functional analysis of the minimal *attB* in *L. lactis*

To investigate recombination between TP901-1 *attP* and *attB* in *L. lactis*, *E. coli* plasmids pAB130 (*erm*, $attB_{min}$) and pLB54 (*erm*, an *attB* fragment of 224 bp), were introduced into *L. lactis* subsp. *cremoris* MG1363 which already contained the plasmid pLB58 carrying a lactococcal origin of replication, a functional TP901-1 *attP* site and *orf1* encoding the TP901-1 integrase. Since the plasmids pAB130 and pLB54 contain no lactococcal origin of replication, Ery^R transformants will only be obtained if they are integrated in pLB58 or the host chromosome. When pAB130 or pLB54 were introduced into MG1363/pLB58, the transformation frequencies were four orders of magnitude higher than when the control plasmid pBC144, containing no *attB* sequences, was introduced (Table 3). Thus, the *attB* sequences of pAB130 and pLB54 are responsible for the increase in transformation efficiency. Recombination was verified by PCR analysis on plasmid DNA prepared from Ery^R MG1363/pLB58 transformants (results not shown). In contrast, the presence of *attB* in pAB130 and pLB54 did not increase the transformation frequency compared to pBC144 in strain MG1363/pCI372, which contains neither *attP* nor the integrase. Thus, the increased frequency of transformation was only observed when both functional *attP* and *attB* sites are present in the cell, as well as the phage integrase.

The efficiency of transformation of MG1363/pLB58 was very similar using the *attB* plasmids pLB54 or pAB130 (Table 3). Thus, the minimal *attB* fragment in pAB130 participates in integrative recombination with the same efficiency as the 224 bp *attB* fragment in pLB54. This confirms that the *attB* fragment in pAB130 contains the information required for efficient recombination by the TP901-1 integrase.

Mutational analysis of the *attB* core

Since sequencing of a large number of recombination products has revealed that the sequences of the *attL* and *attR* sites are always as shown in Fig. 1(b), recombination must take place in the 5 bp core (Christiansen *et al.*, 1994, and results not shown). To define the precise size of the overlap region, we performed an analysis of

Table 4. Product after recombination between wild-type *attP* and mutated *attB*

attB fragment*	No. tested	*attP* (wt)	*attB* (mutated)	% Recombination	Result of recombination†		
					wt	Mutated	Mix
$T_{28}A$	12	TCAAT	ACAAT	10	1	11	0
$C_{29}G$	4	TCAAT	TGAAT	22	2	1	1
$C_{29}A$	4	TCAAT	TAAAT	76	1	2	1
$C_{29}T$	4	TCAAT	TTAAT	76	2	2	0
$A_{30}G$	18	TCAAT	TCGAT	100	18	0	0
$A_{31}G$	17	TCAAT	TCAGT	100	17	0	0

* Sequences of *attB* fragments: Fig. 2(a).

† The number of plasmids with either wild-type (wt), mutated or 'mix' sequence of *attL*. 'Mix' indicates that the sequence showed a band for both of the two bases.

the importance of the bases of the *attB* core for recombination. Each of the five bases was changed by site-specific mutagenesis, and the effect on *in vivo* recombination was investigated in *E. coli* strain DH5α carrying pLB81 expressing the integrase, as described for the deletion analysis. The results are presented in Fig. 2(a).

When the first T in the core region is changed to G or C ($T_{28}G$ and $T_{28}C$), recombination is abolished completely, but for $T_{28}A$ 10 % recombination can take place. Changing C_{29} also reduces recombination, although not as profoundly as changing the first base. The remaining three bases of the core, AAT, can be changed to any of the other three bases without any effect on recombination. Thus, only the two first bases of the core region were observed to be important for recombination, indicating that this TC dinucleotide could constitute the TP901-1 overlap region.

Identification of the TP901-1 overlap region

To obtain further evidence that the TC dinucleotide of the TP901-1 attachment sites does constitute the overlap region, we used the plasmids containing *attB* sites $T_{28}A$, $C_{29}G$, $C_{29}A$ and $C_{29}T$, since the core mutation in these sites does not prevent recombination completely (Fig. 2a). The plasmids were introduced in the integrase-producing *E. coli* strain DH5α/pLB81, also used in the previous experiments, to allow recombination between the mutated *attB* and wild-type *attP*. If the mutation is located within the overlap region, recombination will lead to the pairing of two non-complementary bases. DNA replication or mismatch repair will correct this mismatch, enabling both bases to be present after recombination. The sequences of the *attL* site retained on the plasmids after recombination were determined from a number of Ery^S isolates, and are summarized in Table 4.

For all four mutated *attB* sites tested, *attL* sites of both wild-type or mutated sequence were identified after recombination, as would be expected if the two bases are part of the overlap. For the three *attB* fragments with C_{29} mutated, an approximately 50–50 distribution of the wild-type and changed base was obtained (Table 4), as would be expected if the mismatch created during recombination is corrected by a random mechanism. In contrast, for the plasmid containing T28A, the *attL* sites of plasmids from 12 Ery^S isolates had to be sequenced before a wild-type *attL* site was obtained.Thus, there seem to a strong bias towards the base originating from *attB* being part of the *attL* site after recombination (Table 4).

Even though the bases to the right of this TC do not seem to be important for recombination (Fig. 2a), they could still be part of the overlap region. We therefore repeated the above experiment using the plasmids with the $A_{30}G$ and $A_{31}G$ *attB* sites. We sequenced the *attL* sites of 18 and 17 plasmids obtained after recombination, respectively, and found the wild-type sequence TCAAT in all cases (Table 4). The number of sequences determined rules out the possibility that the wild-type sequence obtained could be the result of mismatch repair. Thus, recombination takes place to the left of these bases, and the data are consistent with the hypothesis that the TC dinucleotide constitutes the overlap region.

Analysis of bases outside the core region

To identify those bases in the *attB* region outside the core that are important for recombination, the *attB* site was mutagenized randomly, and non-functional sites were selected as described in Methods. Approximately 450 *attB* fragments were tested in this way, of which 21 displayed a reduced ability to participate in recombination. Sequence analysis showed that four of these contained single-base-pair substitutions (Fig. 2b). In two of the fragments C_{17} was mutated, and it was observed that an A at this position affected the ability of the fragment to participate in recombination more profoundly than a G. In the third fragment A_{25} was changed to a T, which had only a slight effect on recombination. The fourth fragment obtained by this procedure, the $T_{28}G$ *attB* fragment, contained a substitution in the core

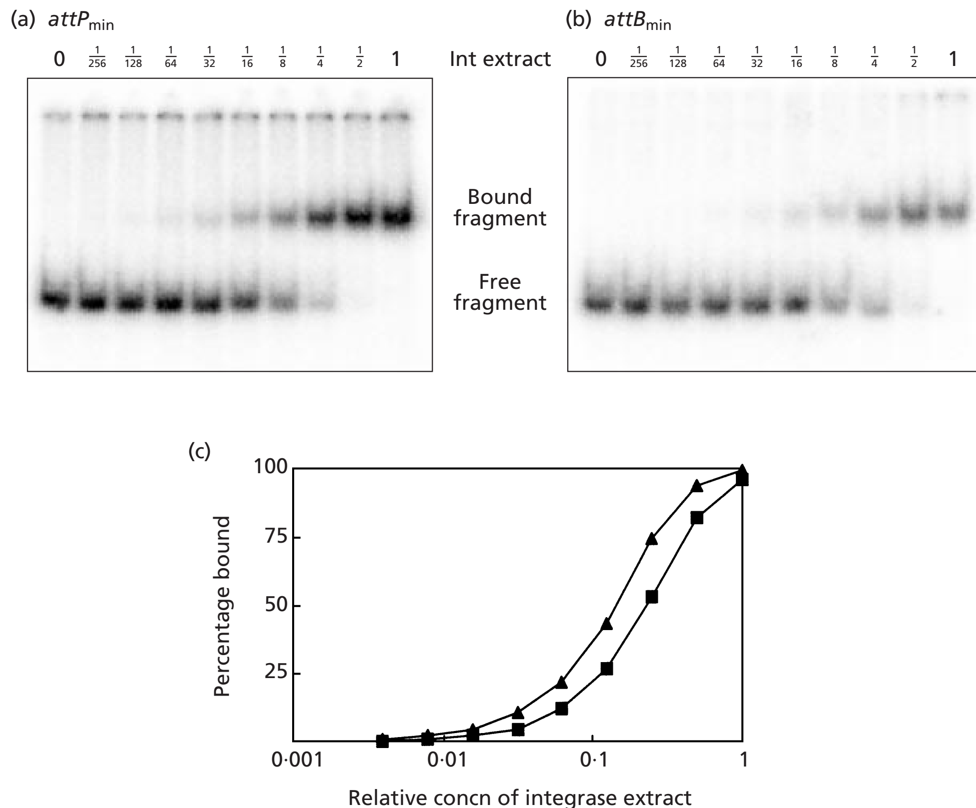

Fig. 3. Binding of the TP901-1 integrase to the minimal *attP* and *attB* fragments. Gel retardation experiments with labelled $attP_{min}$ (a) and $attB_{min}$ (b) fragments and the TP901-1 integrase extract. Relative amounts of integrase extract added are indicated. Undiluted extract corresponds to a final concentration of total protein in the assay of 1·9 g l^{-1}. The positions of the bound and free fragments on the gels are indicated. (c) Percentage of the fragments bound (relative to total amount of labelled fragment) as a function of the amount of protein added. ▲, $attP_{min}$; ■, $attB_{min}$.

region and was discussed in the previous section (Fig. 2a).

To identify single bases that are not important for recombination, *attB* was mutagenized without selecting for a non-functional site. Four *attB* fragments with single-base-pair substitutions were isolated, which were mutated in either G_3, T_9, G_{10} or C_{33} (Fig. 2b).

Binding of the integrase to the minimal *attP* and *attB* sites

To investigate the binding of the TP901-1 integrase to the attachment sites, we decided to perform gel mobility-shift assays, using a crude extract containing the TP901-1 integrase (see Methods). In Fig. 3, binding of the integrase to the 56 bp $attP_{min}$ site and the 43 bp $attB_{min}$ site is shown. No retardation of the fragments was observed with the control extract containing no integrase, nor did the integrase bind to a labelled fragment of similar size containing no *attP* or *attB* sequences (data not shown).

The integrase binds to $attB_{min}$ and $attP_{min}$ with nearly the same affinity (Fig. 3). The affinity of the integrase for the two fragments was further compared in a competition assay, in which both fragments were present in the same binding reaction. The two fragments were discriminated in size by elongation of one with approximately 200 bp of the flanking *erm* cassette. This experiment verified that the affinity of the integrase for the two fragments is identical (results not shown).

An additional binding assay showed that the integrase binds a 333 bp *attP* fragment with the same affinity as $attP_{min}$ and $attB_{min}$ (data not shown). This confirms that the minimal attachment sites contain all the information required for binding of the integrase.

Effect of *attB* mutations on integrase binding

A number of mutated *attB* fragments with reduced ability to participate in recombination have been isolated. The reduction could be due either to an altered binding of the integrase to the fragment, or to an impairment later in the recombination process. To discriminate between these possibilities, the *in vitro* binding of the integrase to selected, mutated *attB* fragments was investigated.

The $T_{28}G$ *attB* fragment is not able to participate in recombination at all (Fig. 2a). However, the integrase

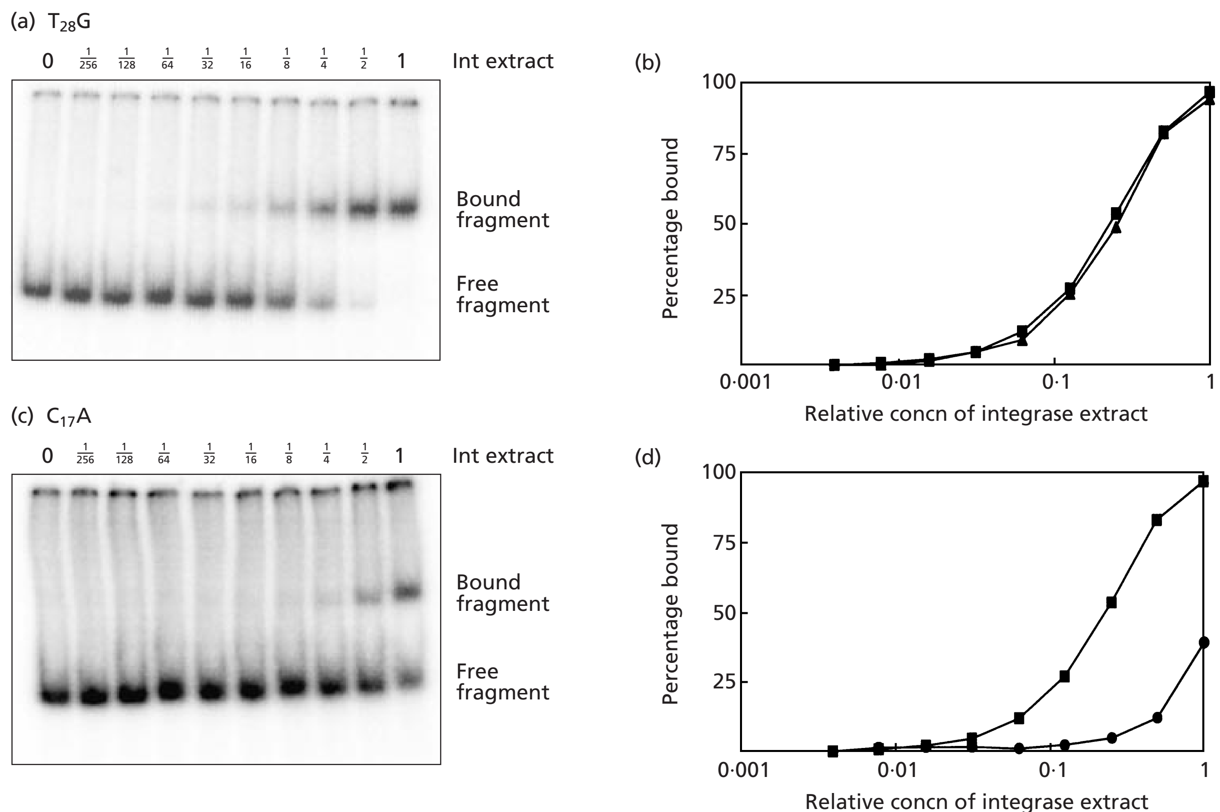

Fig. 4. Binding of the TP901-1 integrase to the $T_{28}G$ and $C_{17}A$ fragments, compared to binding to the *attB*$_{min}$ fragment. (a) Gel retardation experiments with the $T_{28}G$ fragment and the TP901-1 integrase extract. Relative amounts of integrase extract added, and the positions of the bound and free fragments on the gels, are indicated. (b) Percentage of the $T_{28}G$ fragment bound (relative to total amount of labelled fragment) as a function of the amount of protein added (▲). The similar results for the wild-type *attB*$_{min}$ fragment are included for comparison (■). (c) Gel retardation experiments with the $C_{17}A$ fragment and a crude extract containing the TP901-1 integrase. (d) Percentage of the $C_{17}A$ fragment bound (●) compared to the wild-type *attB*$_{min}$ (■).

binds this fragment with the same affinity as wild-type *attB* (Fig. 4a, b). Thus, the inability of the $T_{28}G$ *attB* fragment to participate in recombination is not due to a decreased affinity of the integrase for the fragment. In contrast, binding of the integrase to $C_{17}A$, which can participate in recombination at a level of 21 % relative to wild-type *attB* (Fig. 2b) is severely reduced (Fig. 4c, d), and binding to the $A_{25}T$ fragment, which leads to a less pronounced reduction in recombination, is slightly reduced (data not shown). Thus, the substitutions in the $C_{17}A$ and $A_{25}T$ fragments result in reduced affinities of the TP901-1 integrase for the fragments, and the relative reduction in affinity correlates well with the relative decrease in the ability of the fragments to participate in recombination.

Repeats in the minimal *attB* and *attP* sites

The nucleotide sequences of the 56 bp minimal *attP* and the 43 bp minimal *attB* were inspected for repeated sequences, since these could be integrase recognition sites. Three copies of a 9 bp sequence (TGTTAATTG), with a number of mismatches, were observed in the minimal *attB* (Fig. 1b). These repeats are almost identical to the two P1 repeats previously reported in the minimal *attP* (Brøndsted & Hammer, 1999). If these repeats are integrase recognition sites, it would be expected that mutations that decrease the affinity of the integrase for the fragment are located in these repeats. And indeed, the two bases that have been found to be required for proper binding of the integrase (C_{17} and A_{25}) are located at both ends of the middle repeat. However, mutations have been introduced in the two other repeats without an effect on recombination (B1: G_3A, T_9A, $G_{10}C$; B3: $A_{30}B$, $A_{31}B$, $T_{32}V$, $C_{33}T$).

Location of *attB* on the lactococcal chromosome

Many temperate bacteriophages integrate their genomes in intergenic sequences or in the 3′ end of tRNA genes (for review see Campbell, 1992); however, it has earlier been reported that this is not the case for TP901-1 (Christiansen *et al.*, 1994). We determined the nucleotide sequence of the region of the lactococcal chromosome surrounding the TP901-1 *attB* site (GenBank accession number: Y15043), and found that *attB* is located within

what appears to be an operon, the first three genes of which encode the proteins Orf311, Orf283 and Orf125. These proteins show very high homology to the ComYA, ComYB and ComYC proteins of *Streptococcus gordonii*, and to the ComGA, ComGB and ComGC proteins of *Bacillus subtilis*, which are involved in the development of natural competence in both strains (Lunsford & Roble, 1997; Chung & Dubnau, 1998). Integration of TP901-1 in *attB* results in *orf125* being disrupted, but since *L. lactis* is not known to be able to develop natural competence, there exists no obvious way to investigate the function of this putative protein or the physiological effects of disruption of the ORF.

DISCUSSION

Identification of the TP901-1 minimal *attB*

Recombination by the TP901-1 integrase can take place in *E. coli* as well as in *Lactococcus*. This capacity to function in hosts other than the original one has previously been described for another member of the family of extended resolvases, the integrase of *Streptomyces* temperate bacteriophage ϕC31 (*E. coli*: Thorpe & Smith, 1998; *Homo sapiens*: Groth *et al*., 2000). It has furthermore been shown that only the purified ϕC31 integrase is required for recombination *in vitro* (Thorpe *et al*., 2000).

The smallest fully functional *attB* site identified is 43 bp, and is termed the minimal *attB* or $attB_{min}$ (Fig. 1a). This fragment includes 27 bp upstream and 11 bp downstream of the 5 bp core. When 23 bp are deleted from the upstream part of this fragment, no recombination is possible. Curiously, the presence of an additional 25 bp downstream of *attB8* restores recombination to a level of 20%, even though these bases are not part of the minimal *attB*. Upon inspection of the sequence it is observed that these 25 bp contain two T-stretches of 5 and 7 bp, respectively, separated by a GC dinucleotide. One possible explanation for the stimulatory effect of this region on recombination could be the formation of some special conformation that might enhance binding of a protein required for the process – possibly the integrase.

The TP901-1 overlap region

A substitution of either of the first two bases (TC) of the 5 bp *attB* core reduces the frequency of *in vivo* recombination. Furthermore, when a number of *attL* products, obtained after recombination between such a mutated *attB* fragment and a wild-type *attP*, were analysed, a mix of plasmids with and without the substitution was obtained. In contrast, when the *attB* site contains a mutation in A_{30} or A_{31}, located just downstream of the TC dinucleotide, only wild-type recombination products were obtained. The number of recombinants analysed was high enough that we feel safe to conclude that recombination does take place just downstream of the TC dinucleotide. With the T_{28}A mutant only one wild-type recombination was obtained of the 12 plasmids sequenced (Table 4). However, the inhibitory effect on recombination of any kind of substitution in this position strongly suggests that T_{28} belongs to the overlap region. Furthermore it was shown that the T_{28}G substitution did not affect binding of the integrase to *attB*; hence it has to affect the catalytic activity. Taken together, our results strongly suggest that the TC dinucleotide constitutes the overlap region, which again confirms the hypothesis that the TP901-1 integrase performs resolvase-type recombination.

TC/GA overlap regions have been described both for the Gin invertase of bacteriophage Mu (Mertens *et al*., 1988) and for the extended resolvase TnpX, which is involved in site-specific excision of the conjugative transposon Tn*4451* (Crellin & Rood, 1997; Bannam *et al*., 1995). A mutation analysis revealed that changing the T of the Tn*4451* overlap region has a more profound effect on recombination than changing the C, in agreement with the observations in the present study.

Binding of the TP901-1 integrase to the *attP* and *attB* sites

The TP901-1 integrase binds the minimal *attB* and *attP* sites with equal affinity, resulting in a single retarded band in gel mobility-shift assays. Thus, a likely mechanism for bringing together the DNA species in recombination is that integrase proteins bind to both *attP* and *attB* and the DNA–protein complexes are then paired. This mechanism for DNA binding resembles that of the resolvases and invertases, rather than that of the λ integrase (Richet *et al*., 1988). Likewise the integrase of *Streptomyces* phage ϕC31, which is another extended resolvase, has also been shown to bind *attP* and *attB* with similar affinity (Thorpe *et al*., 2000).

The TP901-1 integrase binds to the T_{28}G fragment, which contains a mutation in the overlap region that prevents recombination completely, as well as to the wild-type *attB* (Fig. 4a, b). This observation confirms the suggestion that this TC dinucleotide constitutes the TP901-1 overlap region. Both of the mutations outside the core which affect recombination were found also to lead to a decrease in the affinity of the integrase for the fragment [C_{17}A: Fig. 4(a, b); A_{25}T: results not shown), suggesting that the decrease in recombination is the result of impaired binding of the integrase. Both C_{17} and A_{25} are situated in the central part of *attB*, at both ends of the B2 repeat (Fig. 1b). The suggestion that the repeats identified in *attB* are involved in integrase binding is further strengthened by the observations that the deletion of the B1 repeat in *attB* fragment 9 reduces recombination to 36% (Fig. 1a), and that related sequences are found in *attP* (Fig. 1b).

Other proteins binding to the attachment sites

The small size of both of the TP901-1 attachment sites suggests that the number of proteins binding to the attachment sites is limited. This observation correlates well with the suggestion that the integrase of TP901-1

performs resolvase-type recombination, since neither the resolvases nor the invertases require the binding of additional proteins to the *res* or *inv* sites for recombination. However, if host-encoded proteins are required to bind the attachment sites during integration, they must be present both in *E. coli* and *L. lactis*, since recombination can take place in both hosts. Another candidate for a protein binding to the attachment sites is the TP901-1 excisionase, encoded by TP901-1 *orf7*, which is required for efficient excisive recombination (Breüner *et al*., 1999). In the TP901-1 prophage, binding of Orf7 to either *attL* or *attR* could stimulate excision by inducing a change in the conformation of the DNA, facilitating excisive recombination.

ACKNOWLEDGEMENTS

We sincerely appreciate the expert technical assistance of Lotte Bredahl and Lise Sørensen. This work was supported by the FØTEK program through The Centre of Advanced Food Studies and by grants from the EC STARLAB program (BIO4-CT96-0402) and the Carlsberg Foundation.

REFERENCES

Azaro, M. A. & Landy, A. (1997). The isomeric preference of Holliday junctions influences resolution bias by lambda integrase. *EMBO J* **16**, 3744–3755.

Bannam, T. L., Crellin, P. K. & Rood, J. I. (1995). Molecular genetics of the chloramphenicol resistance transposon Tn*4451* from *Clostridium perfringens*: the TnpX site-specific recombinase excises a circular transposon molecule. *Mol Microbiol* **16**, 535–551.

Bonekamp, F., Clemmensen, K., Karlström, O. & Jensen, K. F. (1984). Mechanism of UTP-modulated attenuation at the *pyrE* gene of *Escherichia coli*: an example of operon polarity control through the coupling of translation to transcription. *EMBO J* **3**, 2857–2861.

Breüner, A., Brøndsted, L. & Hammer, K. (1999). Novel organization of genes involved in prophage excision identified in the temperate lactococcal bacteriophage TP901-1. *J Bacteriol* **181**, 7291–7297.

Brøndsted, L. & Hammer, K. (1999). Use of the integration elements encoded by the temperate lactococcal bacteriophage TP901-1 to obtain chromosomal single-copy transcriptional fusions in *Lactococcus lactis*. *Appl Environ Microbiol* **65**, 752–758.

Campbell, A. M. (1992). Chromosomal insertion sites for phages and plasmids. *J Bacteriol* **174**, 7495–7499.

Carrasco, C. D., Ramaswamy, K. S., Ramasubramanian, T. S. & Golden, J. W. (1994). *Anabaena xisF* gene encodes a developmentally regulated site-specific recombinase. *Genes Dev* **8**, 74–83.

Christiansen, B. (1995). *Site-specific integration of the lactococcal temperate phage TP901-1*. PhD thesis, Technical University of Denmark.

Christiansen, B., Johnsen, M. G., Stenby, E., Vogensen, F. K. & Hammer, K. (1994). Characterization of the lactococcal temperate phage TP901-1 and its site-specific integration. *J Bacteriol* **176**, 1069–1076.

Christiansen, B., Brøndsted, L., Vogensen, F. K. & Hammer, K. (1996). A resolvase-like protein is required for the site-specific integration of the temperate lactococcal bacteriophage TP901-1. *J Bacteriol* **178**, 5164–5173.

Chung, Y. S. & Dubnau, D. (1998). All seven *comG* open reading frames are required for DNA binding during transformation of *Bacillus subtilis*. *J Bacteriol* **180**, 41–45.

Crellin, P. K. & Rood, J. I. (1997). The resolvase/invertase domain of the site-specific recombinase TnpX is functional and recognizes a target sequence that resembles the junction of the circular form of the *Clostridium perfringens* transposon Tn*4451*. *J Bacteriol* **179**, 5148–5156.

Gasson, M. J. (1983). Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. *J Bacteriol* **154**, 1–9.

Gopaul, D. N. & Van Duyne, G. D. (1999). Structure and mechanism in site-specific recombination. *Curr Opinion Struct Biol* **9**, 14–20.

Groth, A. C., Olivares, E. C., Thyagarajan, B. & Calos, M. P. (2000). A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci U S A* **97**, 5995–6000.

Guo, F., Gopaul, N. & Van Duyne, G. D. (1999). Asymmetric DNA bending in the Cre-*loxP* site-specific recombination synapse. *Proc Natl Acad Sci U S A* **96**, 7143–7148.

Hayes, F., Daly, C. & Fitzgerald, G. F. (1990). Identification of the minimal replicon of *Lactococcus lactis* subsp. *lactis* UC317 plasmid pCI305. *Appl Environ Microbiol* **56**, 202–209.

Holo, H. & Nes, I. (1989). High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media. *Appl Environ Microbiol* **55**, 3119–3123.

Israelsen, H., Madsen, S. M., Vrang, A., Hansen, E. B. & Johansen, E. (1995). Cloning and partial characterization of regulated promoters from *Lactococcus lactis* Tn*917*-*lacZ* integrants with the new promoter probe vector, pAK80. *Appl Environ Microbiol* **61**, 2540–2547.

Kuhstoss, S. & Rao, R. N. (1991). Analysis of the integration function of the streptomycete bacteriophage ϕC31. *J Mol Biol* **222**, 897–908.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Lunsford, R. D. & Roble, A. G. (1997). *comYA*, a gene similar to *comGA* of *Bacillus subtilis*, is essential for competence-factor-dependent DNA transformation in *Streptococcus gordonii*. *J Bacteriol* **179**, 3122–3126.

Lutz, R. & Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements. *Nucleic Acids Res* **25**, 1203–1210.

Matsuura, M., Noguchi, T., Yamaguchi, D., Aida, T., Asayama, M., Takahashi, H. & Shirai, M. (1996). The *sre* gene (ORF469) encodes a site-specific recombinase responsible for integration of the R4 phage genome. *J Bacteriol* **178**, 3374–3376.

Mertens, G., Klippel, A., Fuss, H., Blöcker, H., Frank, R. & Kahmann, R. (1988). Site-specific recombination in bacteriophage Mu: characterization of binding sites for the DNA invertase Gin. *EMBO J* **7**, 1219–1227.

Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Mizuuchi, M. & Mizuuchi, K. (1985). The extent of DNA sequence required for a functional bacterial attachment site of phage λ. *Nucleic Acids Res* **13**, 1193–1208.

Popham, D. L. & Stragier, P. (1992). Binding of the *Bacillus subtilis spoIVCA* product to the recombination sites of the element interrupting the σ^K-encoding gene. *Proc Natl Acad Sci U S A* **89**, 5991–5995.

Richet, E., Abcarian, P. & Nash, H. A. (1988). Synapsis of attachment sites during lambda integrative recombination involves capture of a naked DNA by a protein-DNA complex. *Cell* **52**, 9–17.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463–5467.

Sato, T., Samori, Y. & Kobayashi, Y. (1990). The *cisA* cistron of *Bacillus subtilis* sporulation gene *spoIVC* encodes a protein homologous to a site-specific recombinase. *J Bacteriol* **172**, 1092–1098.

Stark, W. M., Boocock, M. R. & Sherrat, D. J. (1989). Site-specific recombination by Tn*3* resolvase. *Trends Genet* **5**, 304–309.

Terzaghi, B. E. & Sandine, W. E. (1975). Improved medium for lactic streptococci and their bacteriophages. *Appl Microbiol* **29**, 807–813.

Thorpe, H. M. & Smith, M. C. M. (1998). *In vitro* site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci U S A* **95**, 5505–5510.

Thorpe, H. M., Wilson, S. E. & Smith, M. C. M. (2000). Control of directionality in the site-specific recombination system of the *Streptomyces* phage ϕC31. *Mol Microbiol* **38**, 232–241.

Wang, H., Roberts, A. P., Lyras, D., Rood, J. I., Wilks, M. & Mullany, P. (2000). Characterization of the ends and target sites of the novel conjugative transposon Tn*5397* from *Clostridium difficile*: excision and circularization is mediated by the large resolvase, TndX. *J Bacteriol* **182**, 3775–3783.

Received 21 December 2000; revised 19 March 2001; accepted 11 April 2001.