Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 5505–5510, May 1998
Biochemistry

In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family

HELENA M. THORPE AND MARGARET C. M. SMITH*

Department of Genetics, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH, United Kingdom

Edited by Allan Campbell, Stanford University, Stanford, CA, and approved March 5, 1998 (received for review November 24, 1997)

ABSTRACT **The genome of the broad host range *Streptomyces* temperate phage, ϕC31, is known to integrate into the host chromosome via an enzyme that is a member of the resolvase/invertase family of site-specific recombinases. The recombination properties of this novel integrase on the phage and *Streptomyces ambofaciens* attachment sites, *attP* and *attB*, respectively, were investigated in the heterologous host, *Escherichia coli*, and in an *in vitro* assay by using purified integrase. The products of *attP*/*B* recombination, i.e., *attL* and *attR*, were identical to those obtained after integration of the prophage in *S. ambofaciens*. In the *in vitro* assay only buffer, purified integrase, and DNAs encoding *attP* and *attB* were required. Recombination occurred irrespective of whether the substrates were supercoiled or linear. A mutant integrase containing an S12F mutation was completely defective in recombination both in *E. coli* and *in vitro*. No recombination was observed between *attB*/*attB*, *attP*/*attP*, *attL*/*R*, or any combination of *attB* or *attP* with *attL* or *attR*, suggesting that excision of the prophage (*attL*/*R* recombination) requires an additional phage- or *Streptomyces*-encoded factor. Recombination could occur intramolecularly to cause deletion between appropriately orientated *attP* and *attB* sites. The results show that directionality in ϕC31 integrase is strictly controlled by nonidentical recombination sites with no requirement to form the topologically defined structures that are more typical of the resolvases/invertases.**

In site-specific recombination a recombinase interacts with a specific site in the DNA, brings the sites together in a synapse, and then catalyzes strand exchange so that the DNA is cleaved and religated to opposite partners (1, 2). The reaction can result in integration, inversion, or resolution/excision depending on the position and orientation of the recombination sites, their interactions with recombinase, and the presence or absence of accessory factors or sites. Site-specific recombinases in bacteria fall into one of two very distinct families, the λ integrase-like enzymes and the resolvase/invertases, on the basis of amino acid sequence similarities and their different mechanisms of catalysis (1–3). Recombination by members of the λ integrase family (e.g., λ integrase, P1 Cre-*loxP*) is well understood and involves the formation and resolution of a Holliday junction intermediate during which the DNA is transiently attached to the enzyme through a phosphotyrosine linkage (4–6). The resolvase/invertase family of enzymes (e.g., Tn*3* or γδ resolvases, Mu Gin invertase) act via a concerted, four-strand staggered break and rejoining mechanism during which a phosphoserine linkage is formed between the enzyme and the DNA (2, 7). The crystal structure of γδ resolvase bound to a cleavage site reveals a unique arrangement of the catalytic and DNA-binding domains in that they bind to different faces of the helix (8). Although two models have been proposed (9–11), the structure of the synapse and the changes in the conformation of resolvase that bring about strand exchange are still a mystery (12).

Recently a new subgroup of the resolvase/invertases has been identified that have the resolvase/invertase catalytic domain in their N-terminal regions but are much larger in molecular mass (between 50.7 kDa and 82 kDa; refs. 13–20) than, and extremely diverged from, the typical enzymes of the family (≈20 kDa; refs. 2 and 21). This subgroup includes enzymes involved in excision of DNA during spore (13) or heterocyst (14) development, excision of transposons (15, 16), or integration and excision of bacteriophage genomes (17–20). The ability of some of this group of enzymes to catalyze integration, i.e., the integrases from the lactococcal phage TP901–1, and actinophages R4 and ϕC31, is a significant departure from the usual activities of the resolvase/invertases (2).

The recombination sites acted on by this new subgroup also differ fundamentally from the more typical resolvase/invertases. In the latter, the two DNA sites involved in recombination are identical for each system. In Tn*3*, the *res* sites are 114 bp, and in Mu, the *gix* sites are 30 bp (2, 22). The sequences are not symmetrical and therefore sites can be in direct or inverse orientation. The reactions catalyzed by Tn*3* resolvase and Gin, however, are topologically defined, with precise arrangements of the components in the synaptic complex and in the strand-exchange reaction itself (2, 9, 10, 22, 23). Thus, in the resolvase system only *res* sites that are in direct repeat recombine, and in the Gin system only inverted sites recombine; recombination is totally blocked if a synapse with the "wrong topology" is formed because of sites being in the incorrect orientation. In the λ integration system (5, 24), the recombination sites have different sequences, i.e., the 25-bp *attB*, chromosomal attachment site, and the 240-bp *attP*, phage-encoded attachment site, and recombination results in hybrid sites *attL* and *attR*. The sites have a 15-bp of "core" sequence identity. Accessory factors are required for integration (integration host factor, or IHF) and excision (IHF, Fis, and Xis), which also can be described as *attP*/*B* and *attL*/*R* recombination, respectively. Crossover sites for the enzymes belonging to the high molecular weight subgroup of the resolvase/invertases lie within very short regions of identity between substrate sites [3 bp for ϕC31 integrase (ref. 20; Fig. 1); 5 bp for CisA (13, 25); 5 bp for XisF (14); 12 bp for R4 Sre (19); 2 bp for Tn*4451* TnpX (15, 16)]. These probably represent core-like sequences from longer, nonidentical, recombination sites. If this group of enzymes uses the same catalytic mechanisms as the typical resolvase/invertases, then, in theory, only 2 bp of sequence identity between the two sites is required. Recent work by Crellin and Rood (16) showed that mutation of the 2-bp GA sequence at the center of the recombination

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/955505-6$2.00/0
PNAS is available online at http://www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviation: IHF, integration host factor.
*To whom reprint requests should be addressed. e-mail: maggie.smith@nottingham.ac.uk.

sites for TnpX confirmed the need for both sites to have the same sequence where the proposed staggered break occurs.

Here we describe the properties of the ϕC31 integrase expressed in *Escherichia coli* and acting on plasmid-encoded *att* sites and in an *in vitro* reaction by using purified integrase. The results show that directionality in ϕC31 integrase is strictly controlled by nonidentical recombination sites with no requirement to form the topologically defined structures that are more typical of the resolvase/invertases. Furthermore, for recombination between *attP* and *attB,* the enzyme does not require any accessory proteins or supercoiled substrates. The properties of this integrase are discussed in comparisons with other, well studied, site-specific recombination systems.

MATERIALS AND METHODS

Strains and Plasmid Constructions. *E. coli* DH5α [F′/*endA1 hsdR17* (r_K^- m_K^+) *supE44 thi-1 recA1 gyrA*(NalR) *relA1*Δ(*lacIZYA-argF*) *U169 deoR* (ϕ80*dlac*Δ(*lacZ*)*M15*); ref. 26] was used routinely as a host for plasmid constructions, transformations, and plasmid preparations (27). *E. coli* BL21(DE3) [F$^-$ *ompT (lon) hsdS*$_B$ (r_B^- m_B^-) DE3 lysogen; a λ prophage carrying the T7 RNA polymerase gene; ref. 28] was used as a host for overproduction of wild-type and mutant integrases. Plasmids used are listed in Table 1. Detailed information on their construction can be obtained from the communicating author. The crossover site in the 0.5-kbp *attB*-containing fragment in pHS21, pHS23, pHS34, and pHS44 is approximately in the middle of the fragment whereas the crossover site in the 0.5-kbp *attP*-containing fragment in pHS20, pHS22, pHS33, and pHS44 is 112 bp from the *Eco*RI site shown in Fig. 1. pHS282 was constructed by insertion into pGEM7 of an 84-bp amplification product obtained by PCR using oligonucleotides HS3 (5′-AGGTCTCGAGAAGCGGTTTTCGG) and HS4 (5′-CGCCCTAGGTGTCATGTCGGCGACCC) designed by using the published *attP* site (20). The overexpression plasmid, pHS62, containing the *int* gene inserted downstream of the T7 promoter was constructed as follows: A 2-kbp *Bsr*I-*Tth*111I fragment containing *int* was inserted into pT7–7 (31) to form pHS61. Primers were designed by using the *int* sequence (20) for PCR amplification of the N-terminal region to place the start codon downstream of the ribosome-binding site for maximal expression in pT7–7. The primers used were HS1, 5′-AGGTCATATGGACACGTACGCGGGTGC-3′; and HS2, 5′-CCGGCCCCCGTCGCGCTCGA-3′. Amplification yielded a 148-bp fragment that was cut with *Nde*I and *Eco*RI, and a 55-bp fragment was used to replace the *Nde*I-*Eco*RI fragment from pHS61 to give pHS62. The S12F mutation in the *int* gene (in pHS63) was generated by using a mutagenic primer, HS11 (5′-AGGCATATGGACACGTACGCGGGTGCTTACGACCGTCAGTTCCGCGAG) in place of HS1, and the resulting amplification product was used to replace the *Nde*I-*Eco*RI fragment in pHS61 to give pHS63. The PCR-derived fragments were sequenced to ensure that the correct sequences were present.

***In Vitro* Recombination.** Integrase was purified from *E. coli* BL21(DE3) containing either pHS62 or pHS63 by ammonium sulfate precipitation, ion exchange chromatography, and heparin agarose affinity chromatography (H.M.T., S. E. Wilson, and M.C.M.S., unpublished data). Approximately 1 μg each of substrate DNAs was mixed with 1 μg of approximately 90% pure integrase in a volume of 100 μl. Final reaction conditions were 20 mM Tris, pH 7.5/100 mM NaCl/1% glycerol/0.1 mM EDTA, incubated at 30°C. The reactions were stopped by phenol extraction and ethanol precipitation, and the pellets were resuspended in 50 μl of 1× restriction buffer (27). After restriction, 10 μl of the reaction products was separated by electrophoresis on 0.8% 1× TBE (90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 8.3) agarose gels. DNA for use in the recombination reactions was purified by alkaline/lysis and polyethylene glycol precipitation (27). To prepare linearized substrates, plasmids were cleaved with *Sca*I and the linear DNA was purified after electrophoresis in a 0.8% 1× TBE agarose gel.

Table 1. Plasmids used

Plasmid	Comments	Ref.
pGem-7fZ(−)	Vector, resistant to ampicillin	Promega
pZMR100	Defective λ vector, resistant to kanamycin	30
pSP72	Vector, resistant to ampicillin	Promega
pT7-7	Expression vector, resistant to ampicillin	31
pKC796	*Cla*I-*Kpn*I fragment (3.9 kbp) from ϕC31 encoding *attP-int*, resistant to apramycin	29
pKC1034	*Kpn*I fragment (4 kbp) from *S. ambofaciens* encoding *attB*	S. Kuhstoss, Eli Lilly
pMS211	*Bam*HI fragment (1.5 kbp) from ϕC31 repressor region in pUC19	37
pϕC183	*Sau*3A-*Eco*RI fragment (0.5 kbp) encoding *attP* in an *E. coli-Streptomyces* shuttle plasmid	M. Brawner, SmithKline Beecham
pHS15	pZMR100 containing 2-kbp *Bsr*I-*Tth*III fragment encoding ϕC31 *int* expressed from the *tac* promoter	This work
pHS21	*Kpn*I-*Sal*I *attB* fragment (0.5 kbp) from pKC1034 in pGEM7	This work
pHS23	*attB* fragment (0.5 kbp) in pSP72	This work
pHS20	*Bgl*II-*Eco*RI *attP* fragment (0.5 kbp) from pϕC183 in pGEM7	This work
pHS22	*attP* fragment (0.5 kbp) in pSP72	This work
pHS282	*attP* fragment (84 bp) in pGEM7	This work
pHS33	*attP* fragment from pHS22 in pHS15	This work
pHS34	*attB* fragment from pHS23 in pHS15	This work
pHS44	*attP* and *attB* in pGEM7 in direct (POP′–BOB′) orientation	This work
pHS50	*Eco*RI fragment (0.4 kbp) encoding *attL* in pGEM7	This work
pHS52	*Aat*II-*Sal*I fragment (0.4 kbp) encoding *attR* in pGEM7	This work
pHS55	*attL* and *attR* in pGEM7 in orientation BOP′–POB′	This work
pHS61	pT7-7 containing *int*	
pHS62	pT7-7 containing *int* for maximal expression	This work
pHS63	pHS63 encoding mutant S12F integrase	This work

RESULTS

The ϕC31 Integrase Catalyzes the Integration Reaction in *E. coli*. Integrase expressed from the *tac* promoter in pHS33, a derivative of the autonomously replicating defective λ vector, pZMR100 (30), catalyzed recombination between an *attP* site inserted downstream of the *int* gene in pHS33 and the *attB* site on a compatible plasmid (pHS21) in *E. coli* DH5α (Fig. 2). Fragments (approximately 0.4 kbp) containing the products of the *attP/B* recombination, i.e., *attL* and *attR*, were inserted into pGEM7 vectors to form pHS50 and pHS52, respectively, and sequenced to confirm that the recombination in *E. coli* yielded the same products as is normally obtained in *Streptomyces ambofaciens* (20). The *int* gene carrying a frameshift mutation at the unique *Nco*I site failed to yield recombinants, indicating the dependence of the recombination on the expression of ϕC31 integrase (data not shown). The ability of integrase to catalyze *attP/B* recombination in *E. coli* demonstrated that no

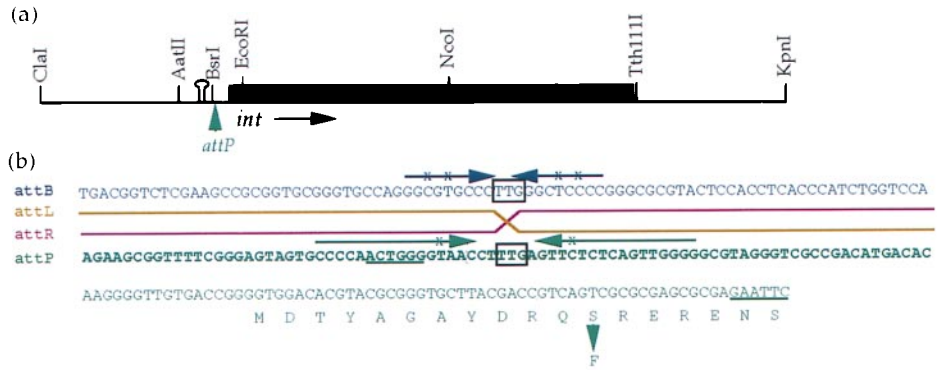

FIG. 1. Organization of the region encoding *int* in ϕC31 and its comparison with the *attB* sequence from *S. ambofaciens* (adapted from ref. 20). (*a*) Organization of *attP* and *int* from ϕC31 and restriction sites relevant to this study. The black box represents the *int* ORF reading left to right, and the green, vertical arrowhead represents the position of *attP*. The stem-loop icon shows the position of a putative rho-independent transcription terminator. (*b*) The sequences of the *attP* and the N-terminal region of *int* (green) and *attB* (blue). The crossover occurs within the boxed nucleotides (20). The extent of the 84-bp *attP* site used in recombination in *E. coli* (Fig. 2) is shown in bold. Underlined sequences in *attP*/*int* show the positions of the *Bsr*I and *Eco*RI restriction sites. Horizontal arrows indicate the positions of inverted repeats in *attP* and *attB*. In the *attP*/*int* sequence, the start of translation, the N-terminal 18 residues and the position of the S12F mutation are indicated. The *attL* and *attR* sequences are shown in yellow and pink, respectively.

other *Streptomyces*-encoded accessory proteins were required. Constructs in which the *attP* site from pHS33 and the *attB* site from pHS21 had been switched (i.e., pHS34 and pHS20) and when introduced together into *E. coli* also gave rise to recombinants (Fig. 2), confirming that the location of the *attP* site with respect to the *int* gene did not affect its activity. Although the precise extent of DNA required for recombination at the *attP* site has not been determined, a plasmid, pHS282, con-

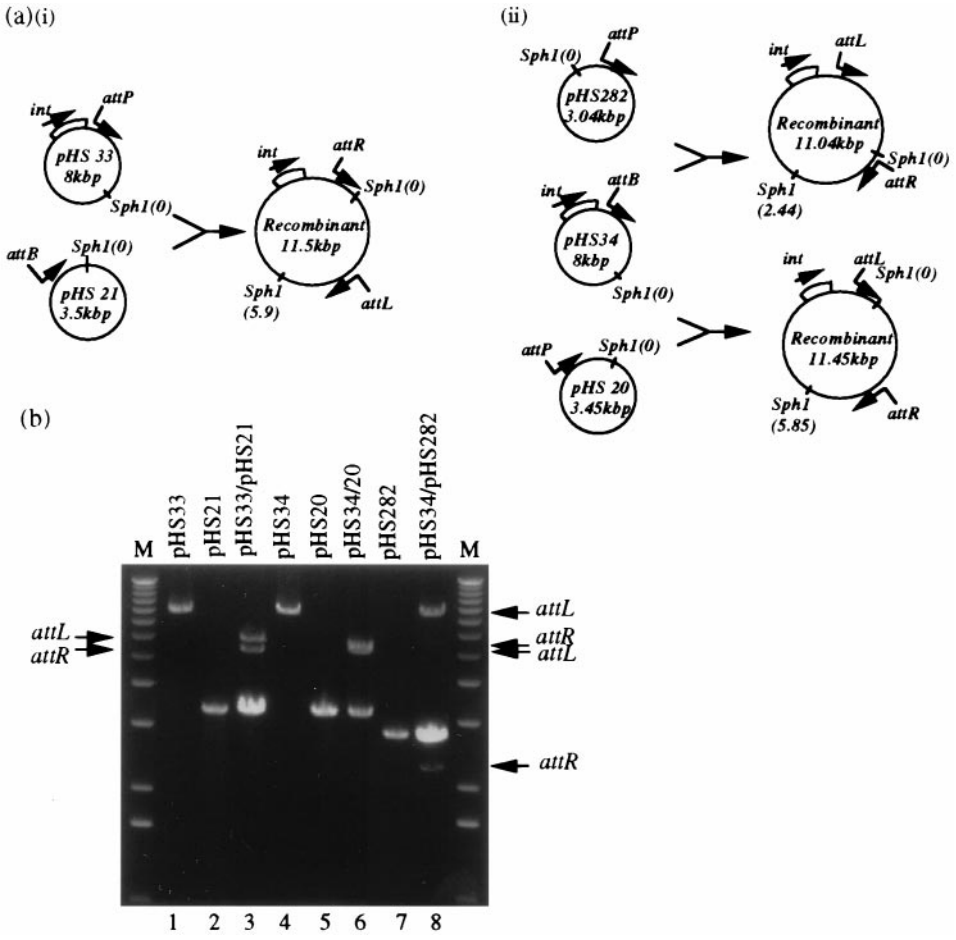

FIG. 2. Recombination by ϕC31 integrase between *attP* and *attB* sites in *E. coli* DH5α *recA*$^-$. (*a*) Restriction maps of plasmids used and expected recombinant products. *E. coli* containing pHS33 and pHS34 encode *attP* and *attB*, respectively, and express the ϕC31 *int* gene from the *tac* promoter located within the vector sequences; the vector in each case is a λ defective plasmid, pZMR100. The *attB* (*i*) and *attP* (*ii*) sites have been placed on compatible plasmids. (*b*) Restriction analysis by *Sph*I of parental and recombinant products after extraction of plasmid DNA from *E. coli*. Bands containing recombination products *attL* and *attR* are indicated by arrows. In lane 6 the recombinant fragments migrate close together but can be resolved after a longer electrophoresis. In lane 8 the *attL* recombinant product and linearized pHS34 comigrate. Molecular weight markers (M) are provided by 0.5 μg of 1-kbp ladder (Life Technologies).

taining 84 bp of DNA encoding the *attP* crossover site (located centrally within the insert) was recombinationally active in *E. coli* containing pHS34 (Fig. 2).

***In Vitro* Recombination Catalyzed by ϕC31 Integrase.** From the experiments above it seemed highly likely that the recombination reaction would occur *in vitro* without any additional proteins other than integrase. The *int* gene was overexpressed and purified by ion-exchange and heparin agarose affinity chromatography (H.M.T., S. E. Wilson, and M.C.M.S., unpublished data). A purified extract containing approximately 90% integrase in 20 mM Tris, pH 8.0/1 M NaCl/1 mM EDTA was obtained that contained little or no DNases. Standard conditions for recombination involved mixing substrate *attP* and *attB* DNA with integrase in Tris buffer, pH 7.5, and incubation at 30°C. A time course of the *in vitro* reaction suggested that as much as 50% of the initial substrate DNA recombined to form *attL* and *attR* after 10–16 hr incubation (Fig. 3*b*). Addition of 10 mM $MgCl_2$ did not stimulate or inhibit recombination as measured after 16 hr of incubation (Fig. 3*c*), and addition of 10 mM EDTA had no apparent inhibitory effect (not shown). Both supercoiled and linearized DNAs were suitable substrates for recombination between *attP* and *attB* (Fig. 3*d*).

During strand exchange a recombinase belonging to the resolvase/invertase family transiently forms a phosphoserine linkage between the enzyme and the 5′ end of a cleaved DNA substrate at the crossover site. The active site serine residues are found very close to the N-terminal end and are surrounded by other well conserved residues (21). In the ϕC31 integrase, Ser-12 is most likely to be the active-site serine (21). A mutation was introduced in the *int* gene to change Ser-12 to a phenylalanine (S12F). The mutant protein, encoded by pHS63, could not catalyze recombination between *attP* and *attB* sites either in *E. coli* or in the *in vitro* reaction by using purified protein obtained via the same procedure as the wild-type protein (Fig. 3*c*), although binding of S12F integrase to *attP* and *attB* was unaffected (not shown).

In the Absence of Other Factors ϕC31 Integrase Only Catalyzes Recombination Between *attP* and *attB*. Could the ϕC31 integrase catalyze excision, i.e., the reaction between *attL* and *attR*? We use here the same system to describe the organization of the *attP, B, L*, and *R* sites as is used for the λ attachment sites (5, 24). The *attB* site therefore is BOB′, *attP* is POP′, where B and B′ and P and P′ describe the left and right arms at each site and O describes the homologous core. After recombination *attL* therefore is BOP′ and *attR* is POB′. The *attL* and *attR* sites were isolated from the products of *attP*/*attB* recombination in *E. coli* as approximately 400-bp fragments containing 112 and 196 bp of phage DNA, respectively. The *attL*- and *attR*-containing fragments were inserted into pGEM7 to produce pHS55 in the orientation BOP′–POB′, i.e., as they would occur in the integrated prophage. To prevent any constraints on recombination because of *attL* and *attR* located too close, pHS55 contained a 1.5-kbp fragment derived from the region encoding the repressor in ϕC31. No recombination was observed between the *attL* and *R* sites in pHS55 either in *E. coli* containing a compatible plasmid expressing integrase (pHS15; not shown) or *in vitro* (Fig. 4). Similarly, no recombination was detected in *E. coli* or *in vitro* between *attL* and *attR* when located on different plasmids, or between *attP* or *attB* and *attL* or *attR* (Fig. 4). Using plasmids of different sizes to facilitate detection of the recombinant form, no recombination was detected between two *attP* sites or between two *attB* sites (Fig. 4).

Intramolecular Recombination Between *attP* and *attB*. Although in nature the location of the *attP* and *attB* sites on the same DNA molecule is unlikely to occur, we looked for any constraint on recombination under these conditions that might be indicative of a topological influence on formation of the synapse. A substrate containing the *attP* and *attB* sites in so-called direct orientation (BOB′:POP′) was constructed (pHS44) and recombination was observed in *E. coli* (also containing pHS15) and *in vitro*. Recombination between the *attP* and *attB* sites in the same molecule is expected to give two

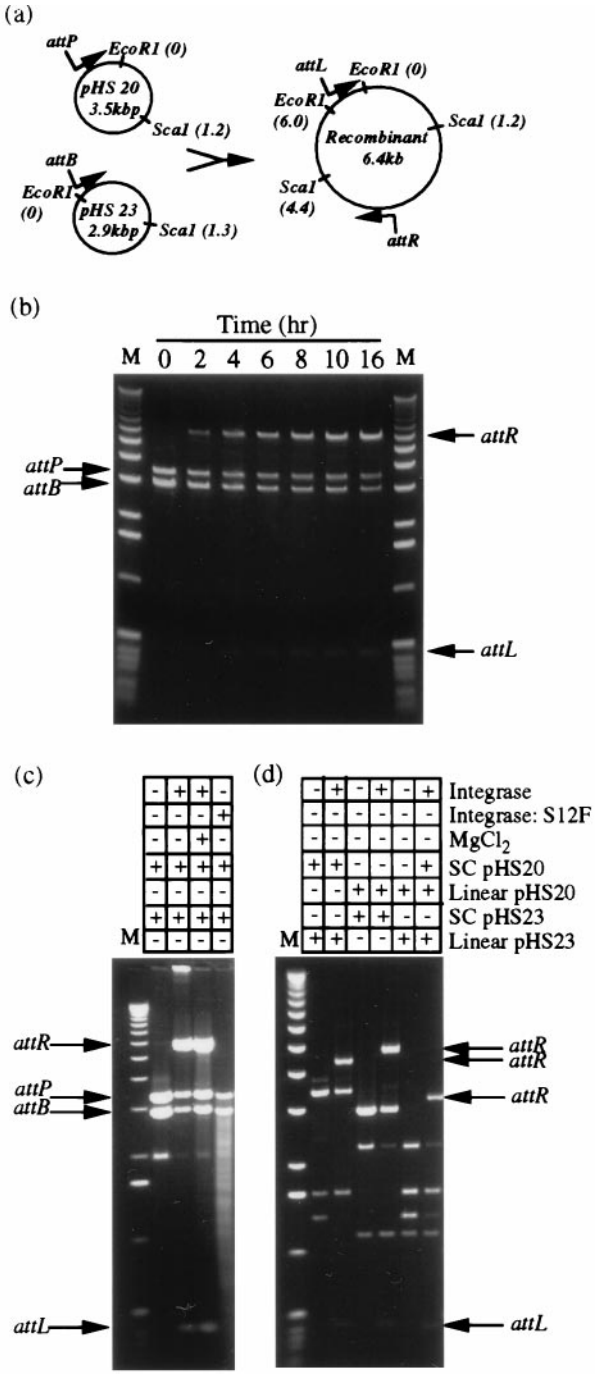

FIG. 3. Recombination by ϕC31 integrase between *attP* and *attB* sites *in vitro*. (*a*) Restriction maps of plasmids used and expected recombinant products. Detection of the recombination products in *b–d* was by restriction with *Eco*RI followed by agarose gel electrophoresis. Parental and recombinant products are indicated. (*b*) A time course of recombination *in vitro*. (*c*) No effect on recombination by $MgCl_2$ and abolition of activity by the S12F mutation. Recombination reactions were incubated at 30°C for 16 hr before analysis. The smear of degraded DNA in the presence of the S12F mutant integrase is probably because of contaminating nucleases. (*d*) Recombination between linear substrates. Linearized pHS20 and pHS23 were prepared by cleaving with *Sca*I, and the fragments were purified before use as substrates in *in vitro* recombination. Recombination reactions were incubated for 16 hr before analysis. Molecular weight markers (M) are as in Fig. 2.

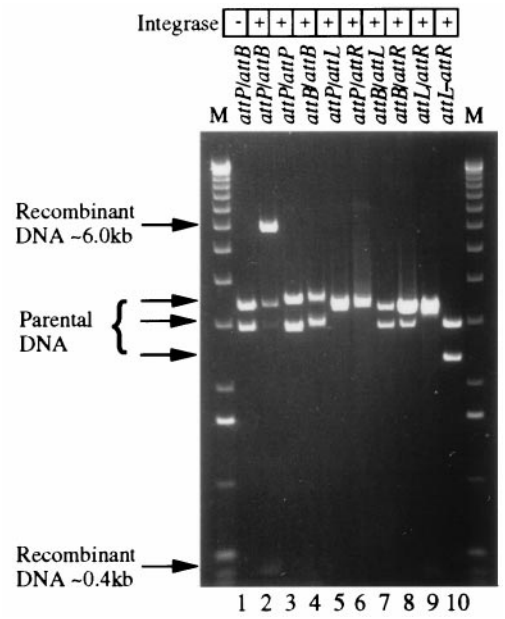

FIG. 4. ϕC31 integrase only catalyzes *attP*/*B* recombination. Plasmids encoding *attP* (pHS20 or pHS22), *attB* (pHS21 or pHS23), *attL* (pHS50), or *attR* (pHS52) or *attL* and *attR* together (pHS55) were used as substrates for recombination with ϕC31 integrase *in vitro*. Recombinant products were obtained only in lane 2 that contained *attP* and *attB* as substrates, and the reaction is the same as that shown in Fig. 3. If recombination had occurred, the predicted recombinant products between *attP*/*attP* (lane 3) and *attB*/*attB* (lane 4) cut with *Nsi*I and *Pst*I and between *attB*/*attL* (lane 7) and *attB*/*attR* (lane 8) cut with *Bam*HI would have been 6 kbp and 0.4 kbp, respectively. For *attP*/*attL* (lane 5) and for *attP*/*attR* (lane 6) recombination, the predicted products would have been 6.4 kbp and 0.4 kbp when cut with *Xho*I and *Apa*I, respectively. For *attL*/*attR* (lane 9) on different plasmids, the predicted product would have been 3.6 kbp when cut with *Sph*I and *Sma*I, and when *attL*–*attR* (lane 10) were on the same plasmid the predicted recombinant product would have been 0.4 kbp detected by restriction with *Kpn*I. Note that pHS52 cut with *Sph*I and *Sma*I (lane 9) yielded a parental band of a size (3.4 kbp) similar to the predicted recombinant product in the pHS50/pHS52 reaction, but still no recombinant could be detected in gels that had undergone electrophoresis for a longer period. Note also that the 0.4-kbp *Sph*I-*Sma*I fragment in pHS50 is a parentally derived fragment. Molecular weight markers (M) are as in Fig. 2.

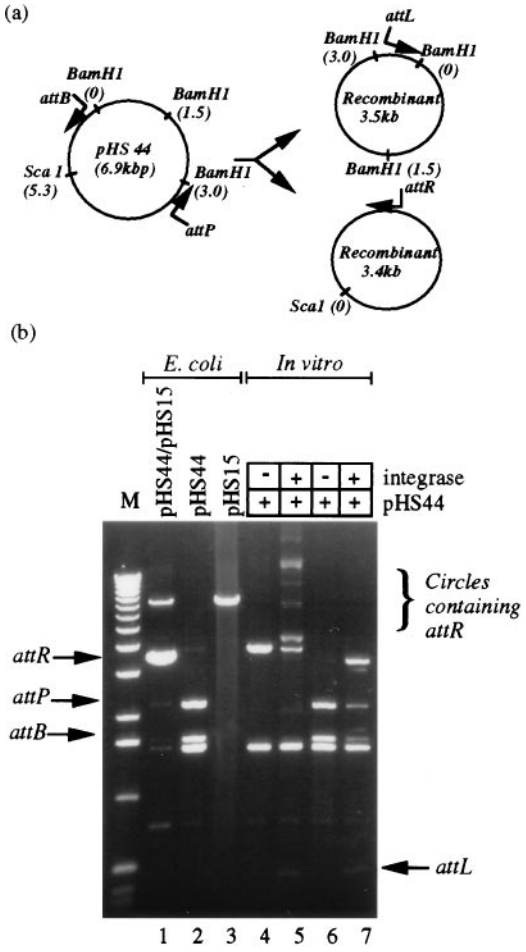

FIG. 5. Intramolecular recombination between *attP* and *attB*. (*a*) Restriction maps of pHS44 and the expected recombinant products. (*b*) Restriction analysis of parental and recombinant products after extraction of plasmid DNA from *E. coli* (lanes 1–3) or after recombination assays *in vitro* (lanes 4–7). *Bam*HI and *Sca*I were used to analyze the plasmids from *E. coli* (lanes 1–3) and after *in vitro* recombination (lanes 7 and 8). In *E. coli* the recombinant product containing *attL* was barely visible (even though *attR* is abundant) because it does not contain an origin of replication and is lost. To detect the *attR*-containing circles obtained after recombination *in vitro, Bam*HI alone was used (lanes 4 and 5), and this cut both the parental DNA and the *attL*-containing product. Molecular weight markers (M) are as in Fig. 2.

circular molecules, only one of which contains a replication origin (*attR*). In *E. coli* the *attL*-containing recombination product is lost (Fig. 5*b*, lanes 1–3). *In vitro* circles were detected after restricting the products of recombination, by using pHS44 as a substrate, with an enzyme (*Bam*HI) that does not digest the *attR*-containing product (Fig. 5*b*, lane 5). The circles then were cleaved with *Sca*I to give the expected linear fragments containing *attR* (Fig. 5*b*, lane 7). The migration properties of the circles suggests that they are relaxed monomers, dimers, trimers, and higher multimers of the *attR*-containing DNA. We presume that the multimers form by recombination between *attP* and *attB* on different plasmids before deletion events occurring by intramolecular recombination. Resolution of two directly repeated *res* sites by resolvase results in singly linked (catenated) supercoiled circular products (9, 10). Curiously, little if any of the *attR* circles obtained by recombination between *attP* and *attB* in pHS44 (Fig. 5*b*) appeared to be supercoiled, although this may be an artifact because of nuclease contamination of the integrase preparation.

DISCUSSION

The ϕC31 integrase is a member of a new subfamily of resolvase/invertases that have, on the basis of amino acid sequence comparisons, the resolvase/invertase catalytic domain at their N termini but that are much larger than the typical members of the family (13–20). Consistent with this classification is the observation that mutation of the putative nucleophile, Ser-12, of ϕC31 integrase completely abolished recombination both *in vivo* and *in vitro* (Fig. 3*c*) whereas previously mutation of Tyr-174 and Tyr-181 (chosen because tyrosine is the nucleophile in enzymes belonging to the λ integrase family of recombinases) had no effect on recombination (20). Similar data were obtained after mutation of Ser-15 of TnpX, which prevented excision of Tn*4451*, but mutation of tyrosines thought to form part of a λ integrase-like catalytic domain had no effect on excision (16). ϕC31 integrase displays unusual properties for a recombinase of the resolvase/invertase family. First, it does not display a strong preference for the position or orientation of recombination sites because it caused integration and deletion depending on the positions of the *attP* and *attB* sites (Figs. 2–5). Second, the reaction does not depend on supercoiling because linear substrates recombined *in vitro* (Fig. 3). These properties are reminiscent of

certain mutants of Gin invertase that, although selected to be Fis-independent, simultaneously gained the ability to integrate and delete DNA between appropriately placed *gix* sites (32, 33). The Fis-independent Gin mutants demonstrated that strand exchange is not dependent on the topologically defined complex of wild-type Gin.

Whereas the resolvase/invertases have evolved to control the direction of recombination via topologically defined nucleoprotein structures, ϕC31 integration is strictly controlled by nonidentical recombination sites (Fig. 4; refs. 20 and 34). This aspect of ϕC31 site-specific recombination is reminiscent of λ recombination but with significant differences that probably reflect the formation of very different synaptic structures. In λ, the minimal *attP* sequence is 240 bp in length and an accessory factor, IHF, is required for recombination (5, 24). At the λ *attP* site, which must be supercoiled, integrase and IHF bind to sites on the arms of *attP* (arm-type binding sites) at some distance from the core sequence (the site of crossing over). λ integrase, which has two DNA-binding motifs, then can contact the low-affinity, core-type binding sites immediately flanking the core sequence (5, 24). The synapse is formed when this *attP*–integrase–IHF complex or "intasome" combines with the much shorter (25-bp) *attB* site (5, 24). In the ϕC31 system we envisage that (like $\gamma\delta$ and Tn*3* resolvase, Gin, etc.) integrase binds the imperfect inverted repeats that flank the crossover sites in *attP* and *attB* (refs. 20 and 34; Fig. 1). This model accounts for the smaller *attP* site (which can be equal to or less than 84 bp) and the lack of accessory factors required for *attP*/*B* recombination. Asymmetries in integrase binding to *att* sites would determine the directionality because only integrase bound to the *attP*/*B* combination of sites could form a productive synapse. We currently are studying these interactions and other features of the *att* sites that might contribute to the control of directionality in this system. We expect that the excision reaction catalyzed by ϕC31 integrase (i.e., *attL* and *attR* recombination) will require accessory proteins, probably phage-encoded, thus resembling excision catalyzed by XisF in *Anabeana* (35) or plasmid pSM19035 monomerization by the β protein (36). The structure and activities of ϕC31 and other members of this new group of recombinases show that the resolvase/invertase family is much more diverse than originally thought (2).

We are grateful to S. Kuhstoss and M. Brawner for gifts of plasmids. This work was funded by Medical Research Council Grant G9502658.

1. Sadowski, P. (1986) *J. Bacteriol.* **165,** 341–347.
2. Stark, W. M., Boocock, M. R. & Sherratt, D. J. (1992) *Trends Genet.* **8,** 432–439.
3. Argos, P., Landy, A., Abremski, K., Egan, J. B., Haggard-Ljunquist, E., Hoess, R. H., Kahn, M. L., Kaliones, B., Narayana, S. V. L., Pierson, L. S., III, *et al*. (1986) *EMBO J.* **5,** 433–440.
4. Nunes-Duby, S., Azaro, M. A. & Landy, A. (1995) *Curr. Biol.* **5,** 139–148.
5. Landy, A. (1989) *Annu. Rev. Biochem.* **58,** 913–949.
6. Guo, F., Gopaul, D. N. & Van Duyne, G. D. (1997) *Nature (London)* **389,** 40–46.
7. Hatfull, G. F. & Grindley, N. D. F. (1988) in *Genetic Recombination*, eds. Kucherlapati, R. & Smith, G. R. (Am. Soc. Microbiol., Washington, DC), pp. 357–396.
8. Yang, W. & Steitz, T. A. (1995) *Cell* **82,** 193–207.
9. Stark, W. M., Boocock, M. R. & Sherratt, D. J. (1989) *Trends Genet.* **5,** 304–309.
10. Stark, W. M., Sherratt, D. J. & Boocock, M. R. (1989) *Cell* **58,** 779–790.
11. Rice, P. A. & Steitz, T. A. (1994) *EMBO J.* **13,** 1514–1524.
12. Oram, M., Szczelkun, M. D. & Halford, S. E. (1995) *Curr. Biol.* **5,** 1106–1109.
13. Sato, T., Samori, Y. & Kobayashi, Y. (1990) *J. Bacteriol.* **172,** 1092–1098.
14. Carrasco, C. D., Ramaswamy, K. S., Ramasubramanian, T. S. & Golden, J. W. (1994) *Genes Dev.* **8,** 74–83.
15. Bannam, T. L., Crellin, K. & Rood, J. I. (1995) *Mol. Microbiol.* **16,** 535–551.
16. Crellin, P. K. & Rood, J. I. (1997) *J. Bacteriol.* **179,** 5148–5156.
17. Mahairas, G. G., Sabo, P. J., Hickey, M. J., Singh, D. C. & Stover, C. K. (1996) *J. Bacteriol.* **178,** 1274–1282.
18. Christiansen, B., Bronsted, L., Vogensen, F. K. & Hammer, K. (1996) *J. Bacteriol.* **178,** 5164–5173.
19. Matsuura, M., Noguchi, T., Yamaguchi, D., Aida, T., Asayama, M., Takahashi, H. & Shirai, M. (1996) *J. Bacteriol.* **178,** 3374–3376.
20. Kuhstoss, S. & Rao, R. N. (1991) *J. Mol. Biol.* **222,** 897–908.
21. Leschziner, A. E., Boocock, M. R. & Grindley, N. D. F. (1995) *Mol. Microbiol.* **15,** 865–870.
22. Van der Putte, P. & Goosen, N. (1992) *Trends Genet.* **8,** 457–462.
23. Kanaar, R., van der Putte, P. & Cozzarelli, N. R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 752–756.
24. Nash, H. A. (1990) *Trends Biochem. Sci.* **15,** 222–227.
25. Stragier, P., Kunkel, B., Kroos, L. & Losick, R. (1989) *Science* **243,** 507–512.
26. Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., Decruz, E., Noyerwiedner, M., Smith, S. S., Michael, M. Z. & Graham, M. W. (1989) *Nucleic Acids Res.* **17,** 3469–3478.
27. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
28. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60–89.
29. Kuhstoss, S., Richardson, M. A. & Rao, R. N. (1991) *Gene* **97,** 143–146.
30. Rogers, M., Ekaterinaki, N., Nimmo, E. & Sherratt, D. J. (1986) *Mol. Gen. Genet.* **205,** 550–556.
31. Tabor, S. & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1074–1078.
32. Klippel, A., Cloppenborg, K. & Kahmann, R. (1988) *EMBO J.* **7,** 3983–3989.
33. Klippel, A., Kanaar, R., Kahmann, R. & Cozzarelli, R. (1993) *EMBO J.* **12,** 1047–1057.
34. Rausch, H. & Lehmann, M. (1991) *Nucleic Acids Res.* **19,** 5187–5189.
35. Ramaswamy, K. S., Carrasco, C. D., Fatma, T. & Golden, J. W. (1997) *Mol. Microbiol.* **23,** 1241–1249.
36. Rojo, F. & Alonso, J. C. (1994) *J. Mol. Biol.* **238,** 159–165.
37. Smith, M. C. M. & Owen, C. (1991) *Mol. Microbiol.* **5,** 2833–2844.